# The Atlas of Reality

# The Atlas of Reality

## A Comprehensive Guide to Metaphysics

**Robert C. Koons and Timothy H. Pickavance**

WILEY Blackwell

This edition first published 2017


*Registered office*
John Wiley & Sons, Ltd, The Atrium, Southern Gate, Chichester, West Sussex, PO19 8SQ, UK

*Editorial Offices*
350 Main Street, Malden, MA 02148-5020, USA
9600 Garsington Road, Oxford, OX4 2DQ, UK
The Atrium, Southern Gate, Chichester, West Sussex, PO19 8SQ, UK

For details of our global editorial offices, for customer services, and for information about how to apply for permission to reuse the copyright material in this book please see our website at www.wiley.com/wiley-blackwell.



*Library of Congress Cataloging-in-Publication Data applied for.*

Hardback ISBN: 9781119116127

A catalogue record for this book is available from the British Library.

Cover image: *Atlas of Reality*, graphite and ink drawing (2016), © Tim High, www.timhigh.com

Set in 10/12.5pt MinionPro by Aptara Inc., New Delhi, India

1 2017

To our children, Emily, Betsy, and Ben Koons, and Lyle and Gretchen Pickavance.

# Contents

# Acknowledgements

We would like to thank our commissioning editor at Wiley-Blackwell, Deirdre Ilkson, our project editors, Roshna Mohan and Manish Luthra, and four anonymous reviewers. We would also like to thank Al Martinich, who got us started on this project by inviting us to contribute to the Fundamentals of Philosophy series. Our special thanks to Benjamin Koons for his help in proofreading the manuscript. We are grateful to Peter Forrest, L. A. Paul, Daniel Korman, Josh Parson, Graham Priest, and Trenton Merricks for comments on an early draft, Alexander Pruss and Josh Rasmussen for their help on causation and composition, and Robert Garcia for his insights on tropes and constituent ontology.

We have both benefited greatly from the assistance of our students, including the students of eight sections of PHL 323K and 327 at the University of Texas from 2010 through 2015. Graduate students at both the University of Texas and Biola University have contributed much to this project, including Bryan Pickel, Nicholas Mantegani, Brian Cutter, Nora Berenstain, Richard Lawton Davis, Nicole Garcia, Trevor Nyman, Kate Ritchie, Travis Hobbs, Briggs Wright, Janice Chik, Matthew M. O'Brien, and Max Goss.

RCK adds: I would like to acknowledge the support during the 2014-2015 academic year of the James Madison Program in American Ideals and Institutions at Princeton University (for a Visiting Fellowship) and the University of Texas at Austin (for a faculty research grant).

THP adds: I would like to thank Scott Rae and Clint Arnold, as well as the administration at Biola University, for a Sabbatical during the Fall of 2015.

# Part I

# Foundations

# 1

# Introduction

Metaphysics, or first philosophy, is that branch of philosophy concerned with the nature of reality in its most fundamental aspects: existence, the part/whole relation, space, time, causality, possibility and necessity, similarity and dissimilarity. It includes ontology, the study of what exists, as well as the investigation of the most general features of reality. Metaphysicians seek to understand the real structure and the unity of the world and to catalog the ways in which its parts relate to each other.

In this chapter, we begin with a short history of metaphysics (1.1), followed by a discussion of some reasons why metaphysics matters (1.2). We conclude with some guidance about how best to use this book (1.3).

## 1.1 A Brief History of Metaphysics

Metaphysics is the oldest branch of philosophy. The early, pre-Socratic philosophers of Ionia (off the coast of Turkey) and southern Italy proposed theories about the universal nature of things and about change and the explanations of change. Many of the early philosophers, including Empedocles (c. 490–430 BC) and Democritus (c. 460–370 BC), approached these questions from a materialistic point of view, assuming that wisdom comes primarily from understanding what things are made of. In the fourth century BC, the great Greek philosopher Plato (428/427 or 424/423–348/347 BC) developed a theory of "forms" as a deep explanation for what makes things of a kind similar to each other, as an alternative to the earlier materialism. Plato's student, Aristotle (384–322 BC), built upon the work of all of his predecessors in creating the first comprehensive and systematic metaphysical theory in a work that acquired (for the first time) the title *Metaphysics*. Aristotle describes his subject as "primary" or "first" philosophy and as the study of being as such. Aristotle examined the nature of change and of powers to change, and he built

*The Atlas of Reality: A Comprehensive Guide to Metaphysics*, First Edition.
Robert C. Koons and Timothy H. Pickavance.

a theory of categories to use in classifying all of the constituents of reality. Like Plato, Aristotle rejected simple materialism and emphasized the qualitative and holistic features of the world, especially of living organisms.

Both Plato and Aristotle founded schools of philosophy, and their students and their students' students extended their philosophical work over many generations. During the Hellenistic period (between the conquests of Alexander and the rise of Rome), three additional major schools of philosophy appeared—the Stoics, the Epicureans, and the Skeptics. Both the Stoics and Epicureans revived a more materialistic approach to understanding life and human action. During the Hellenistic and Roman periods, such metaphysical investigations continued, but gradually attention turned to ethics, politics, and the theory of knowledge (epistemology). The problem of defending the very possibility of knowledge against the challenge of the Skeptics became a major preoccupation, and Plato's Academy began to defend (at least in public) a moderate form of skepticism.

In Late Antiquity and the Middle Ages, philosophical work in the Mediterranean basin and in Europe fell predominantly into the hands of Christians, Jews, and Muslims, and during this period philosophers returned to metaphysics as their central focus. Aristotle's influence grew, as more of his work was translated and commented upon in both Arabic and Latin. The resulting philosophical movement, known as 'scholasticism', achieved the status of being the consensus view for many hundreds of years.

This consensus began to dissolve in the seventeenth and eighteenth centuries, as a result of the success of modern science, which returned in important respects to the materialism of Democritus and the Epicureans. At the same time, the French philosopher René Descartes (1596–1650) re-introduced a fixation on the problem of refuting the skeptics. Descartes recommended answering the skeptics by turning inward, building the foundations of science and philosophy firmly upon the indubitable contents of one's own mind and experience. This inward or subjective turn profoundly affected the course of metaphysics for hundreds of years, leading to the dominance of various forms of *idealism*, according to which all of reality is fundamentally mental or experiential in character.

In the early twentieth century, a number of philosophers began turning away from idealism and from any attempt to build an indubitable foundation for knowledge that would be immune to the challenge of the skeptic. The British philosopher G.E. Moore (1873–1958) argued that our ordinary knowledge of the world is rationally more secure than any skeptical challenge. Ludwig Wittgenstein (1889–1951), an Austrian who made his career in England, pointed out that doubt stands in no less need of justification than does belief. Wittgenstein concluded that skeptical doubts lacked adequate justification. Many philosophers, in Britain, the United States, and continental Europe, argued that science requires no foundation other than that provided by ordinary observations, which embody knowledge about our physical environment. Thus, philosophy began to turn outward again, in a way that supported the revival of more traditional approaches to metaphysics—materialistic, Platonic, and Aristotelian or scholastic.

For a brief period at the end of the nineteenth century and beginning of the twentieth century, metaphysics fell out of favor among philosophers. Some (such as Friedrich Nietzsche (1844–1900), Karl Marx (1818–1883), William James (1842–1910), and Søren Kierkegaard (1813–1855)), because of theories in psychology and cultural studies, raised doubts about the ability of the metaphysician to escape the prejudices and interests of one's class and time or the idiosyncratic influence of one's personal constitution. Others (such as those in the Vienna Circle and Ludwig Wittgenstein) embraced an extreme empiricism, arguing that all meaningful assertions must be directly verifiable or falsifiable by the senses, a standard which relegated metaphysical theory to the category of the nonsensical. Yet another group limited the task of the philosopher to analyzing the underlying grammar and logic of ordinary language.

But the middle of the twentieth century witnessed one of the most remarkable rebirths in Western philosophy: a dramatic renaissance of interest in pure metaphysical theory. The impetus for this revival came in part from circles that had once been hostile to the metaphysical enterprise and in part from philosophers working within older traditions that had survived despite that hostility. Some philosophers of physics found themselves inquiring into the structure of space, time, and causation in ways that revived ancient debates. Some who had studied the logical structure of ordinary language found that they could not avoid questions about ontology—questions about which sorts of things really exist. Others returned to the Aristotelian and scholastic traditions that had survived. Significant circles of metaphysical research began in the 1940s in Australia, at Oxford and Cambridge, and at Harvard. Logical research that had dominated philosophy in the early twentieth century matured naturally into metaphysical investigations into the nature of possibility and necessity and of time. By the early twenty-first century, metaphysics had reclaimed its place at the very center of philosophy.

## 1.2 Why Do Metaphysics?

The practice of metaphysics is controversial within philosophy itself. This controversy stems from two primary sources: skepticism and pragmatism. Anti-metaphysical skeptics question whether it is possible to reach knowledge or even reasonable opinion about metaphysical questions. Our response to the skeptic is simply that the proof is in the pudding. The best rebuttal of those who claim that metaphysics is impossible is simply to do it.

The pragmatic challenge to metaphysics is perhaps even more widespread. Even if metaphysics is possible, the pragmatist asks, why is it important? There are many more urgent philosophical questions, questions about ethics and politics (the good and the right), and questions of epistemology (what do we know, and how do we know it?).

Our response to the pragmatist is twofold. First, we would argue, with Aristotle, that philosophy begins with a sense of wonder and curiosity about the world, a wonder and a curiosity that inevitably led to puzzling over the metaphysical questions: what sorts of things really exist, and how do these things relate to one another? Much of what we do in science and scholarship generally is motivated by pure curiosity about our selves and our world. Not everything can be evaluated in terms of cash value.

Second, metaphysical questions are relevant to other questions, both in value theory and in epistemology and philosophy of science, as we will argue in this chapter. Even when philosophy is primarily engaged in ethical or epistemological reflection, the issues of metaphysics cannot be avoided.

### 1.2.1 Fatalism and alternative possibilities

The making of decisions is a characteristic feature of human life. Much of our time is consumed in considering and deliberating about what to do, and our emotions are much engaged with questions of the correctness of our past, present, and future choices. The practice of making choices seems to presuppose that the future could take any one of many alternative courses, and that which course it takes is to some extent up to each of us. This presupposition is metaphysical in nature. Suppose that we inhabit the one and only one possible world—that is, suppose that nothing has happened in the past or could happen in the future other than the one way in which things must, of necessity, unfold. On such a fatalistic picture, human decision-making would seem to be pointless and devoid of significance.

If the future is indeed open, this fact would raise further metaphysical questions. How and why is the future open in a way that the past is not? Why does it make sense to deliberate about what to do in the future, but not to deliberate about what to have done in the past? What does the direction of time consist in, and how do we know which direction is which? What does it mean to say that something is possible, impossible or necessary? Are there merely possible things that do not actually exist, but might have existed? These are questions in the area of metaphysics known as *modality*. The underpinnings of modality, dispositions and powers, are discussed in Part II, on dispositions (Chapters 4–6). We turn to the questions of modality proper in Part V, Chapters 14–16.

### 1.2.2 Causation: rights, responsibilities, and knowledge

The practices of deliberation and decision-making also seem to presuppose that we have some kind of influence over future events: that we can, in some cases, cause things to happen or prevent their occurrence. If there were no such causation, then, even if the future were open to many alternatives, much of our deliberating would be pointless, since none of our actions would have consequences. Perhaps things just happen, for no reason whatsoever. Is the impression we have that some things cause other things a mere illusion?

If none of our actions or decisions had consequences, this would have radical implications for our understanding of our moral responsibility. We generally take care to ensure that our actions do not impinge, without adequate justification, on the rights and welfare of others. Whenever we do harm others, we think of ourselves as being under burdens of guilt, remorse, and the obligation to make amends and to compensate our victims. Conversely, we believe that choosing to benefit others creates some reciprocal responsibilities of gratitude and thanksgiving. Without cause and effect, all of these practices would be rendered unintelligible. The nitty-gritty details of causation matter morally. For example, are we morally responsible for the consequences of our omissions? Can omissions have consequences?

Since Edmund Gettier's famous paper (Gettier 1963), most philosophers have accepted that there is a difference between knowledge and justified or reasonably held true belief—knowledge involves some real and non-accidental connection between the act of knowledge and the thing known. In most cases, this connection seems to involve causation:

our sense experience, for example, must be influenced by the features of the object being sensed if it is to constitute perceptual knowledge.

We deal with causation in Part VIII, Chapters 26–28.

### 1.2.3 The foundations of science: laws, space, and time

Some prominent scientists have expressed skepticism about the value and the very possibility of metaphysics, in light of the robust success of the natural sciences. Given the coherent and well-supported account of the world provided by modern physics and cosmology, what is the point of indulging in the archaic practice of metaphysical speculation and argumentation, a practice which provides little evidence of secure forward progress?

However, the very success of modern science itself presses forward certain metaphysical questions. What are the laws of nature that play so prominent a role in modern physics? Are those laws of nature necessary or contingent? Can they themselves be explained by more fundamental facts or are they the rock bottom? David Armstrong (1983, 1993) and Fred Dretske (1977) have argued that an adequate understanding of the laws of nature must consider them to involve objective relations between *universals* (natural properties like mass and charge). Others (Ramsey 1928/1978, Lewis 1973a), following in the tradition of Scottish philosopher David Hume (1711–1776), argue to the contrary that laws consist merely in certain kinds of regularities among particular facts. We consider various theories about the laws of nature in Chapters 4 and 5.

Moreover, the description of reality current in physics is often incomplete or indeterminate with respect to certain unavoidable questions. Are space and time infinitely divisible or are they made up of very small, indivisible units? Are there absolutely fundamental units of matter (electrons, quarks, or whatever) or is every kind of material thing decomposable into still smaller units? Does time itself have an absolute beginning or end? Could the universe be infinite in extent? Are the distinctions between past, present, and future of absolute significance or do they have meaning only in relation to a particular location in space and time or a particular relative velocity? Scientific theories do not typically entail answers to these questions, and yet it seems that, if the world is as science describes it, such questions must have answers. It is the vocation of the metaphysician to tackle such questions as these. We consider these foundational questions about space and time in Part VI, Chapters 17–21.

### 1.2.4 Mind and body

Modern physics has been apparently quite successful in telling us about the fundamental building blocks of matter: electrons, photons, quarks, and all of the associated fields and forces (nuclear, electromagnetic, and gravitational). Human beings and the most characteristic features of our experience and action—conscious experience, feelings and emotions, our sense of free will and agency, the normative standards of reasonableness and morality—play no role in the physicist's "complete" description of the world. This raises a host of questions about the relationship between what the American philosopher Wilfrid Sellars (1912–1989) called the "manifest image of the world" and its "scientific image". (Sellars 1962). Do the personal states of conscious experience, thought, decision, and intention pull any weight in explaining the actual course of events or are they all merely *epiphenomena*, a colorful decoration of a reality that is fully determined (insofar as it is determined at all) by the micro-physical facts? Do the special properties of consciousness somehow "emerge" from the underlying physical facts, introducing some genuine novelty, some "addition of being"? The relation between mind and body arises in two contexts in this book: in Chapter 13 (on idealism and the nature of perception), and in Chapters 22 and 25, on composite things (like persons and sentient organisms).

Many of these questions fall within the scope of the philosophy of mind, but answers to these questions often depend upon prior answers to purely metaphysical questions. If certain events can be accurately described in human terms (psychological and social—normative and rational), what must reality be like in its most fundamental aspects to make these higher level descriptions true? What does it take for complex entities to exist and to have real properties or features? What are properties and features, and what is it for something to *have* them? Is this merely a linguistic matter, a matter of how we describe things or how they appear to us or is the having of properties part of the fundamental constitution of things? The nature of properties is the subject of Chapters 7 and 8.

### 1.2.5 Personal identity and persistence

Are there large, complex, and enduring things or is everything microscopic and fleeting in existence? This classical question of metaphysics has great import in our everyday lives since we ourselves, we human beings, are, if we exist at all, large, complex, and enduring things. If the world consists only of subatomic particles or instantaneous events, then we are all mere fictions or illusions. This of course raises the question: if we are illusions, who is deluded? Are we human beings dreams dreamt by protons and electrons? The ancient African thinker Augustine of Hippo (354–430) and Descartes argued that it is impossible for each of us to be deceived in thinking that he or she exists, since, in order to be deceived, we must first exist.

Further, the unreality of human persons would have profound implications for our lives, since much of what we value in life consists in the qualities of our relationships to other persons. We care about the endurance and maturing of our friendships over time, but this concern would have no real object if human persons themselves never endure beyond a single instant, if what we popularly call a 'person' is merely a chain of ephemeral entities.

The foundations for an account of persons and personal identity are laid, first of all, in Chapter 8 on particulars. Since persons and other organisms are apparently complex things, with many material parts, Chapters 22 and 23 (on Composition) are essential to understanding how such composite things can exist. The question of personal identity through time is a special case of persistence, the subject of Chapter 25.

The set of fundamental truths provides us with a complete description of the world in terms of the most basic, irreducible facts. Derived truths can be derived from the set of fundamental truths by means of logic and ontological definitions, definitions specifying in terms of the arrangements of fundamental entities what the existence of derived entities consists in. For example, suppose that tables are not fundamental entities. Then all truths about tables would be derivable from the fundamental truths—say, truths about the arrangements of bits of wood—given a suitable definition of what it is for some bits of wood to constitute a table. A fundamental thing is something mentioned or referred to in some fundamental truths; derived things show up only in connection to derived truths. We examine the nature of grounding and fundamentality in Chapter 3.

## 1.3 How to Use the Book

THE TABLE OF METAPHYSICAL THESES AND ANTITHESES We intend in this book to explore, as completely as possible, the "logical space" of metaphysics: to say at least something about every possible theory on the important questions in metaphysics. We try to present the best arguments for and against each position, as fairly and impartially as we can. Undoubtedly, the careful reader will be able to detect that our own sympathies lie in a broadly Aristotelian region, but we hope that those with more Humean or Platonic sympathies will find little or no grounds for complaint.

We've used a unique method of labeling our theses and antitheses in order to represent perspicuously this ongoing ambition. We have listed all of the theses and antitheses of the book (along with the necessary definitions of technical terms) in Appendix A. The theses are numbered first by chapter: so, thesis 3.1T is the first thesis to be considered in Chapter 3. The negation of 3.1T is labeled '3.1A', with the letter 'A' indicating that it is the antithesis of the thesis 3.1T. Thesis 3.2T is the second major thesis considered in Chapter 3, and 3.2A is its antithesis.

Metaphysical theories that adopt a particular position on one of the theses or antitheses can often be usefully subdivided, depending upon the stance that they take on some subsidiary issue. So, for example, the antithesis 13.3A, perceptual realism, is divided into two sub-theories: indirect realism (13.3A.1T) and its antithesis, direct realism (13.3A.1A). Direct realism is further subdivided into perceptual dualism (13.3A.1A.1T) and unitary direct realism (13.3A.1A.1A):

13.3A Perceptual Realism.
  13.3A.1T Indirect Realism.
  13.3A.1A Direct Realism.
    13.3A.1A.1T Perceptual Dualism.
    13.3A.1A.1A Unitary Direct Realism.

In other cases, the metaphysical theories that share a common commitment cannot be simply divided in a binary way, based on their position on some one subsidiary issue. Instead, there may be three or more different ways of making a given position more determinate. In those cases, we follow the name of a thesis with a period and a numeral (1, 2, 3, or 4), without adding any additional T's or A's. For example, we break down Reductive Nominalism (8.1T) into four sub-theories:

8.1T.1 Predicate Nominalism
8.1T.2 Concept Nominalism
8.1T.3 Class Nominalism
8.1T.4 Resemblance Nominalism

THE TABLE OF METAPHYSICAL PRINCIPLES A second ambition that we have for the book is to keep track with great care the metaphysical first principles that we appeal to in developing arguments for or against a particular position. We have listed all of the metaphysical principles that appear in more than one section of the book in Appendix B. The principles are divided into six major categories: principles of methodology (PMeth 1 through 4), principles of knowledge or epistemology (PEpist 1 through 5), principles of truth (PTruth 1 and 2), principles of metaphysics (PMeta 1 through 6), principles of natural philosophy (PNatPhil 1 and 2), and axioms of mereology, the formal theory of parts and wholes (MA 1 through 6), for a grand total of 25 principles. Some of the principles take more than one form or are associated with a number of corollaries. The first principle of methodology, Ockham's Razor, has six corollaries (and one addendum), while the second principle, scientific realism, takes two distinct forms, objectivity (PMeth 2.1) and reliability (PMeth 2.2).

There are also a number of other first principles that occur only once in the text: these are always given a name (distinguished by boldface type). Chapter 29, the final chapter, includes a table in which the principles appealed to by each of four philosophical "packages" are listed. The four packages, Aristotelian, Ludovician (for David Lewis), Fortibrachian (for David Armstrong), and Flatlander (for Quine, Chisholm, Plantinga, and van Inwagen), represent bundles of philosophical theses and antitheses that cohere together naturally in terms of their rationales and methodological commitments, as the table helps to reveal.

THE ORGANIZATION OF THE BOOK This book is divided into eight parts, with a total of 29 chapters. The first three chapters, including this one, are introductory in character. Chapters 2 and 3 introduce the two notions of *truthmaking* and of *grounding*, ideas that lie at the heart of a significant number of metaphysical projects. They can be skipped by those who are willing to plunge into the project of positive metaphysics, armed only with an intuitive grasp of such notions as something's making a proposition true, or of one truth holding in virtue of or wholly grounded by another. The chapters provide useful details about the methodological foundations of much of the rest of the text.

Part II comprises three chapters, each developing an account of dispositions: conditionals (Chapter 4), laws of nature (Chapter 5), and intrinsic powers (Chapter 6). This part is really foundational for the rest of the book and cannot be omitted.

After completing Chapter 6, the reader can take a number of different paths. Part III, on universals and particulars, is largely independent of the rest of the book, as is Part IV, on the scope of existence and the question of idealism. Parts V and VI, on modality (possibility and necessity) and space and time, are highly interdependent and should ideally be read as a unity. The final two parts, VII (on the unity of things) and VIII (on causation), depend on much of what has gone before them and should be read at the end, as should the conclusion, Chapter 29.

We have also written a much shorter introduction to metaphysics, *Metaphysics: The Fundamentals* (Wiley-Blackwell 2015), which could serve very well as an introduction and orientation to this volume. In almost every case where there is overlap between the two volumes, we go into more detail and consider more theoretical alternatives in this volume. Our discussions in Part IV and of the structure of space and the nature of causation are almost entirely without precedent in the earlier work. *Metaphysics: The*

*Fundamentals* contains a final chapter ("The Concluding Unmetaphysical Postscript") in which we defend metaphysics against various skeptical challenges. We assume, in this volume, that our reader has already exorcised such skeptical demons.

To return to the organization of this volume, in Part III, we turn to four chapters on the ancient problems of universals and particulars. Chapter 7 includes our treatment of the arguments for and against the existence of *universals*—things that are common to things that are similar to one another. Chapter 8 examines the alternative view of Nominalism, according to which everything real is particular and unshareable, including the form of Nominalism that posits individualized properties or *tropes*. In Chapter 9, we consider the internal constitution of ordinary things and the ways in which they can be distinguished from one another. We take up the special case of relational and quantitative properties in Chapter 10.

Part IV includes three chapters in which we consider theories about the nature of reality as a whole. We look first (in Chapter 11) to the question of how many things exist: none, one or more than one? Then, we consider (in Chapter 12) the place in reality of the non-existent, the merely possible, and the impossible. Finally, we examine (in Chapter 13) the case for supposing that all of reality is fundamentally mental or *ideal* by looking carefully at the structure of human sense perception.

In Part V, we take up the questions of *modality*: necessity, possibility, contingency, and actuality. Chapter 14 contains our treatment of David Lewis's theory of possible worlds as concrete, material universes. In Chapter 15, we examine the opposing view, according to which possible worlds are abstract representations, properties or states of affairs. Chapter 16 concerns the problems of *de re* possibility, the realm of possibility that concerns the potentialities of particular things rather than of the whole world, and of our knowledge of modality.

We deal in Part VI with the nature of space and time, with two chapters devoted to space and three to time. In Chapter 17, we consider whether space is a thing in its own right or whether it consists merely in the holding of spatial relations between bodies. We look in Chapter 18 at the structure of space: whether it consists fundamentally in dimensionless points or in extended regions. In Chapter 19, we ask similar questions about the structure of time: does it consist primarily of durationless instants or in extended periods? Chapters 20 and 21 concern two competing theories about the flow or passage of time, the A and B Theories. According to the A Theory, the differences between past, present, and future are absolute and fundamental, while for the B Theory the differences consist entirely in differences in one's perspective from within time.

In Part VII, we take up the question of the unity of things, both at one time and through time. Chapters 22 and 23 concern the unity of composite things, things made up at a single time of many distinct parts. When do things make up a single whole (Chapter 22), and what is it for them to do so (Chapter 23)? The next two chapters concern the unity of things that persist through change and time. We take up the nature of change in Chapter 24 and the nature of those things that persist through change in Chapter 25.

The final part, Part VIII, concerns the metaphysical problems of causation. In Chapter 26, we consider the question of whether causation exists at all. Assuming there is such a thing as causation, we must then consider what things does causation relate: truths or concrete events? Is it a relation between existing things or merely a logical relation between truths (Chapter 27)? Finally, we examine the relations between causation and time in Chapter 28. How do earlier events influence later ones: by a direct connection across time, or by being part of a single, temporally extended process?

The book concludes with Chapter 29, in which we describe the four competing philosophical packages that have emerged in the course of the rest of the work.

Importantly, the reader should be aware that the divisions in the book are not meant to demarcate disconnected sub-fields of metaphysics, but are rather helpful divisions that make the metaphysical task a bit more manageable. One's views in one area can impact one's views in another; we do our best to make those connections clear when they are especially important. Further, and maybe more importantly, there are rarely if ever deductive arguments with unassailable premises for or against a metaphysical position. What one is faced with, rather, is a stock of evidence that one must weigh in order to form a considered opinion. Therefore, one must not only think about the evidence for and against a particular system but also make comparative judgments about which system does best on the evidence taken together. This is very difficult to do well, especially in light of the explosion of activity in metaphysics in recent years, and the interconnectedness of the various regions in that sprawl.

There are two important consequences of this picture of metaphysics. First, the reader who has yet to form views in one or another area would do well to abstain from forming an opinion in that area until she has fully digested the connections to other areas and the strengths and weaknesses of views there. And second, despite the length of this book, we have been unable to carry every dispute to its furthest boundaries. We truncate the journey, sometimes by a good deal, in almost every direction. For those readers committed to a view that is underexplored or shortchanged, we ask your forgiveness.

# 2

# Truthmakers

## 2.1 Introduction

Metaphysics is at least an attempt to build a theory that makes sense of many of the more abstract and general features of reality. Metaphysicians try to understand which things are fundamental, those things' natures, and the fundamental ways those things relate to one another. But where does one start such an endeavor?

There is no generally accepted answer to that question, and so we start with a mundane yet fruitful observation: we say and write lots of *truth-evaluable* things. That is, we say and write things that have a truth-value, usually if not always either True or False. We can and regularly *do* say all sorts of true things. We also can and regularly do say all sorts of false things. Putting these two together, we can note that 'Grass is green', 'Some trees are over one hundred feet tall', 'Two plus two is five', 'Kicking puppies for no reason at all is morally good', and 'Either the Triune God of Christianity exists or he doesn't' are all either true or false.

We can take this mundane observation, however, and put it to work. Philosophers have long argued about how to understand the nature of truth; that is, they have argued about how it is that a sentence comes out true or false. Here is how Aristotle put it in his *Metaphysics*:

> [Thus] we define what the true and the false are. To say of what is that it is not, or of what is not that it is, is false, while to say of what is that it is, and of what is not that it is not, is true.
> Metaphysics 1011b25–8 (Aristotle 1984: 1597)

Aristotle might naturally be taken to be suggesting that a sentence is true whenever the world is the way that the sentence says it is. The sentence 'Grass is green' is true, according to this suggestion, if and only if grass really is green. Any view of truth that incorporates this suggestion as a central feature is a *Correspondence Theory of Truth*. A

*The Atlas of Reality: A Comprehensive Guide to Metaphysics*, First Edition.
Robert C. Koons and Timothy H. Pickavance.

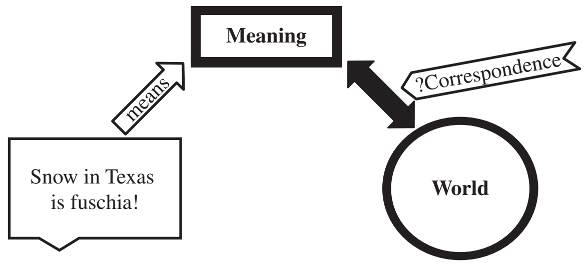


**Figure 2.1** The Correspondence Relation

Correspondence Theory is one that claims a sentence is true if and only there is a match, a correspondence, between what a sentence says and the world. Different Correspondence Theories are distinguished by their respective takes on what "matching" or "correspondence" consists in.

There are two sides to this simple correspondence story. We can picture the view as in Figure 2.1.

The first side of the correspondence story involves the sentence and what it says about the world. The second side is the world itself. We'll say a bit about the first side and then turn to a more sustained discussion of the second. In both cases, the task is to say something more careful and deep about the natures of the items on either side of such a correspondence relation.

### 2.1.1 Propositions

The first way that a discussion of truth gets one going in metaphysics is via its connection to *propositions*. Propositions, according to standard philosophical lore, are the meanings of (many, most, or all) assertoric (indicative) sentences and are the objects of belief and knowledge (and maybe other mental attitudes like desire). The idea is that a sentence token, say a token utterance of 'Texas is super-duper' expresses a proposition, that Texas is super-duper.[1] This proposition represents the world as being a certain way. (More on representation below.)

If the world is as the proposition represents the world as being, if Texas really is super-duper, then the proposition is true and, thereby, so is the sentence token that expressed it. The truth or falsity of a sentence token, on this view, is derivative of the truth or falsity of the proposition it expresses. Similarly, a token belief is a relation to a proposition that functions as the content of that belief, and again, that proposition represents the world as being a certain way. One might, for example, believe that Texas is super-duper or believe that Southern California is overcrowded. These beliefs are relations to the propositions that Texas is super-duper and that Southern California is overcrowded, respectively. As with sentences, beliefs are derivatively true or false in virtue of the truth or falsity of the propositions that are their contents.

Propositions are standardly taken to be both *non-linguistic* and *representational*. Indeed, these two features are crucial to propositions playing their characteristic roles in language and thought.

To motivate the idea that propositions must be *non-linguistic*, consider the possibility of not introducing propositions at all. You might think that sentences can do anything that propositions can do, like being the contents of language and thought and being the fundamental bearers of truth-value. If so, one could simply cut out the propositional middle man. However, there are two reasons why philosophers have thought it necessary to posit propositions in addition to sentences, and crucially, both reasons plausibly demand a non-linguistic object to serve as the content of language and thought. First, sentences in different languages can mean the same thing, and second, native speakers of different languages can share belief contents. Two people who speak different languages can express the same idea. For example, 'Grass is green' means in English just what 'L'herbe est verte' means in French. These two sentences express the same proposition. If this is the case in general, then sentences cannot themselves function as meanings because sentences are language-bound but meanings appear not to be. Similarly, two people who natively speak different languages can have thoughts with the same content. Consider a native English speaker who has a thought she would express with the sentence 'Grass is green' and a native French speaker who has a thought he would express with the sentence 'L'herbe est verte'. These two thoughts have the same content, the same meaning. For the same reason, therefore, sentences cannot function as the meanings of these thoughts because sentences are language-bound but contents appear not to be. In light of the inability of sentences to serve as meanings for sentences and contents for thoughts, philosophers often postulate propositions.

We can turn these two observations into reasons to think that propositions must be non-linguistic. Suppose that propositions, like sentences, were language-bound objects. If so, the English sentence 'Grass is green' would express an English-proposition, and the French sentence 'L'herbe est verte' would express a French-proposition. But that can't be right, for the same reason that sentences cannot function as meanings: these two sentences express the same proposition! Thus, for propositions to function as meanings, they must be non-linguistic.

In addition to being non-linguistic, propositions must *represent* the world. This is necessary in order to explain, for example, how it is that we have true and false beliefs and how it is that we say true and false things. Propositions have representational properties in virtue of which they can be accurate or inaccurate. These representational features are something like the representational features of a realist painting, drawing, or sculpture or like the representational features of a photograph. Insofar as an artist is attempting to mirror the features of the world in his or her art, the resulting piece can be evaluated for its accuracy. In a similar way, beliefs and sentences represent the world in virtue of their relationships to propositions. Beliefs are true or false because they have propositions as their content, propositions that represent the world accurately or not; the sentences we use are true or false because they express propositions that represent the world accurately or not. See Figures 2.2 and 2.3.

The representational features of propositions are plausibly different from the representational features of realist artwork and photographs, however, in at least two ways: (i) propositions represent what they do independently of the intentions of conscious agents, and (ii) propositions do not represent by being similar to that which is represented. One important metaphysical question that arises at this stage, then, is just how it is that propositions represent what they do. Another metaphysical question arises as well:

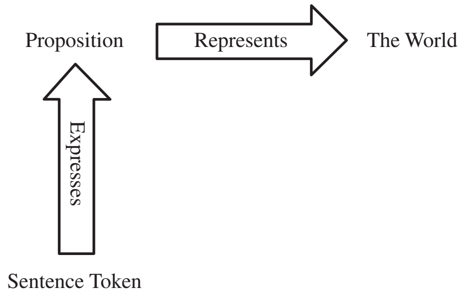


**Figure 2.2** Propositions as Contents of Sentence Tokens

what, exactly, *are* propositions? So far, we have identified two features of propositions that constrain what sort of thing a proposition could be. They must be non-linguistic and representational. Further, we have given propositions two roles to play, namely, being the meanings of (at least many) assertoric sentences and the contents of belief (and other mental attitudes).[2] But this leaves open the question of what sort of non-linguistic, representational thing propositions are. These two questions are importantly interrelated. By answering the second one may find that one has arrived at, or anyway constrained the possibilities regarding an answer to the first.

Philosophers have taken a number of views about the true nature of propositions. Some claim they are fundamental abstract objects. Among those who take this view, some maintain that propositions have no internal structure and have primitive representational features, while others maintain that propositions have some sort of quasi-linguistic structure and represent what they do because of this structure. Others claim that propositions are certain sorts of fact in the world, like facts about which objects have which features or which objects are related to which other objects or facts about objects and features of objects in combination with facts about some language or other. Still others claim that propositions are just arbitrary classes of "possible worlds." (We discuss possible worlds in Chapters 14–16.) There are pluses and minuses to each of these views of propositions, and unfortunately we do not have the space to discuss them in detail here. Happily, which view of propositions one takes will not have a substantial impact on how our discussion proceeds from this point forward; we simply want to flag these issues and suggest that the interested reader take up these interesting questions for him- or herself.

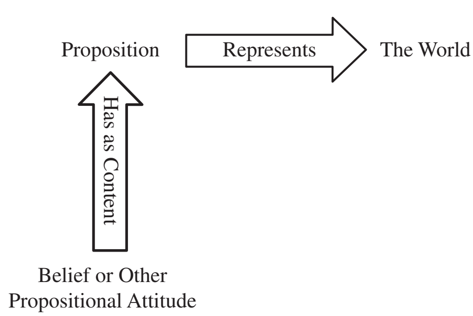


**Figure 2.3** Propositions as Contents of Attitude States

To wrap things up, we can articulate a simple and basic theory of propositions. Propositions are:

1 Abstract (immaterial, not located in space or time).
2 Capable of being true or false.
3 Capable of being believed, doubted, considered or assumed.
4 Capable of being expressed by simple indicative statements.
5 Things that stand in various logical or quasi-syntactic relations to other propositions: e.g., each proposition has a unique negation; two propositions have a conjunction, etc.
6 Things that contain one or more *concepts*, either individual or general, where concepts are also abstract things that stand in some sort of intentional relation (a relation of representation or *aboutness*) to some particular thing, general property or class of things.

This theory leaves much about the exact nature of propositions up for grabs, but articulates certain crucial features of propositions that any adequate theory must accommodate.

### 2.1.2 Truthmakers

What can we say about the world side of the correspondence story? The central question concerns what propositions must correspond to in order to be true. It seems fairly clear that we shouldn't say that a proposition must correspond to the *whole* of the world in order to be true. Consider the proposition that grass is green. It shouldn't take the whole world for that proposition to come out true!

There was a tradition in metaphysics during the nineteenth century, however, that would claim that in an important sense every truth was necessarily connected to every truth. These British Idealists and Hegelians believed the world was a single, interconnected, inseparable whole. That is, these philosophers were *holists*. Holists maintain that the interconnection of the world is so pervasive that if one truly understands one part of reality, one thereby understands the whole of reality. Because of this interconnectedness, any sentence must correspond to the whole world in order to be true.

The early part of the twentieth century saw a strong reaction against holism, led prominently by Bertrand Russell and Ludwig Wittgenstein. These philosophers argued that holism failed to conform to common sense and to our best scientific practices and that there must be a set of fundamental, logically independent facts, the "logical atoms." Two facts *a* and *b* are logically independent if and only if all four of the following are possible: (i) *a* and *b* both exist, (ii) neither *a* nor *b* exists, (iii) *a* exists but *b* doesn't, and (iv) *b* exists but *a* doesn't. Logical atoms, according to logical atomists, were the proper parts of the world to which propositions must correspond in order to be true. Supposing, for example, that grass's being green is a logical atom, the proposition that grass is green is true because it corresponds to the logical atom of grass's being green. Further, and unlike holists, logical atomists would say that the fact that grass is green is utterly independent and distinct from the fact that the sky is blue, and thus that the proposition that the sky is blue need not correspond to the fact that grass is green in order to be true.

Logical atomists like Russell and Wittgenstein, as well as other philosophers like G.E. Moore, further developed the Correspondence Theory of Truth by asserting that facts are what *make* propositions true. In so doing, they built on an idea first expressed by Aristotle in his *Categories*:

> [I]f there is a man, the statement whereby we say that there is a man is true, and reciprocally—since if the statement whereby we say that there is a man is true, there is a man. And whereas the true statement is in no way the cause of the actual thing's existence, the actual thing does seem in some way the cause of the statement's being true: it is because the actual thing exists or does not that the statement is called true or false." (Aristotle, *Categories* 14b15–22; 1984: 22)[3]

Because the fact that grass is green is what makes the proposition that grass is green true, the fact that grass is green is a *truthmaker* for the proposition that grass is green.

**Definition (Def) D2.1 Classical Truthmaker.** A thing $x$ is a (classical) *truthmaker* for proposition $p$ if and only if (i) necessarily, if $x$ exists, then $p$ is true and (ii) necessarily, if $x$ exists and $p$ is true, then $p$ is true at least in part in virtue of $x$'s existence.

The fact that grass is green seems to satisfy these constraints for being a truthmaker of the proposition that grass is green. It's necessarily the case that if the fact that grass is green exists, then the proposition that grass is green is true; thus (i) is satisfied. And it's necessarily true that the proposition that grass is green is true, at least in part, in virtue of the existence of the fact that grass is green; thus (ii) is satisfied. Similarly, the proposition that the University of Texas is the largest university in Texas is true whenever the fact that UT is the largest university in Texas exists, and it is true in virtue of the existence of that fact. (What, exactly, are facts? Good question. We take it up in Chapter 9.)

This truthmaking story is plausible as a way of working out the Correspondence Theory of Truth. Though we aren't interested at this point in the tug-of-war between holism and atomism that initiated the philosophical discussion of truthmakers, we are interested in truthmakers themselves. The remainder of this chapter is devoted to whether we ought to go in for truthmaker theory, how we ought to formulate truthmaker theory if we do go in for it, and what truthmakers might be.

Any view that believes there are things satisfying Def 2.1, that is, any view that believes in (classical) truthmakers, is a 'Classical Truthmaker Theory'.

**2.1T Classical Truthmaker Theory.** There are classical truthmakers for all truths or for some very large sub-classes of truths.

**2.1A No Classical Truthmakers.** There are few, if any, truths with classical truthmakers.

There are a number of varieties of Classical Truthmaker Theory, and we consider them below (Sections 2.4 and 2.5). Before we do that, we consider why we should believe in Classical Truthmaker Theory in the first place (Section 2.2), as well as a fundamental challenge to the very foundation of truthmaker theory: the *deflationary* theory of truth (Section 2.3).

If at the end of Section 2.3, you find Classical Truthmaker Theory convincing, or if you find the whole idea of truthmakers mysterious and unhelpful, then you need not consider the variations in Sections 2.4 and 2.5. If, in contrast, you find the basic motivation for truthmakers to be compelling but are unsatisfied with the Classical picture, then Sections 2.4 and 2.5 offer some useful alternatives. There is *Atomic Truthmaker Theory*, in which the truthmaker principle is applied only to the simplest truths, those that lack negation or any other logical connective. We also consider two more radical departures from Classical Truthmaker Theory: *Spectral Truthmaker Theory* and *Truth Supervenes on Being*. Spectral Truthmaker Theory is based on the observation that reality does not seem to be uniformly *binary* in character—namely, that truthmakers simply exist or do not—as Classical Truthmaker Theory supposes. Truth Supervenes on Being tries to capture some of the leading motivations for truthmaker theory while dispensing with truthmakers altogether.

## 2.2 Five Arguments for Classical Truthmaker Theory

Why should we believe that there are truthmakers? Here are five arguments for truthmakers.

1. The Correspondence Theory of Truth leads naturally to truthmakers. This is something we've already seen, but the point can be turned into an argument for truthmaker theory. First point, one that's already been made: the Correspondence Theory of Truth is initially highly plausible. It seems to make sense of our usual ways of thinking about most, if not all, truths. It also explains why it is that we investigate the world in order to ascertain whether a claim is true or not. We seem to take our knowledge of the meaning of a sentence whose truth is in question and compare it to what we find in the world; if there is a match, we think the sentence is true, and if not, then it's false. This is precisely what we would expect if the Correspondence Theory were correct. So suppose the Correspondence Theory of Truth is in fact correct.[4]

Second point: The chunk of the world to which a proposition must correspond in order to be true satisfies clause (i) of the definition of a truthmaker. The proposition that grass is green simply *cannot* fail to be true if grass is in fact green![5]

Third point: There seems to be an asymmetric dependence of the truth of a proposition on the chunk of the world to which that proposition must correspond. In other words, it is *because* grass is green that the proposition that grass is green is true. It is difficult to argue for this, but it is even more difficult to deny! Compare: it is because the proposition that grass is green is true that grass is green in fact. That seems wrong. But then, the proposition that grass is green is true in virtue of the existence of the fact that grass is green. Thus, the chunk of the world to which a proposition must correspond in order to be true satisfies clause (ii) of the definition of a truthmaker.

Let's approach this question from another tack. The Correspondence Theory proposes that truth consists in a certain kind of relation, call it 'truthful correspondence', between a proposition and some real thing or things. What kind of relation is truthful correspondence?

One question we can ask is this: is truthful correspondence an internal relation? G.E. Moore articulated a distinction between internal and non-internal relations according to which an *internal* relation is a relation *R* which is such that, whether it holds of two entities, *x* and *y*, depends only on the intrinsic characters of *x* and *y*, considered individually.

**Def D2.2 Internal Relation.** *R* is an *internal relation* if and only if, necessarily, for every *x* and *y*, whether *R* holds between *x* and *y* depends only on the intrinsic properties of *x* and of *y*.

Relations are non-internal if and only if they are not internal. Of course, to understand this definition, one must understand what it takes for a property to be intrinsic. An object's intrinsic character is determined by its intrinsic properties, where a property is intrinsic to an object if that object has the property because of the way it is or its parts are, rather than in relation to something that is not itself or its parts.[6]

**Def D2.3 Intrinsicality.** *x* is *intrinsically F* if and only if nothing that is not *x* or a part of *x* is part of the ground of *x*'s being *F*.

The class of internal relations includes the relation of *being the same size as*, that of *being twice as tall as*, *being smarter than*, *contains more parts than*, etc. Consider the relation of *being smaller than*. To tell if one thing is smaller than another, all we need to know are the sizes of each of the two things, and size is an intrinsic property of a thing. Relations that are not internal include those of *being at least a meter apart from*, *being to the north of*, *is moving away from*, etc. If a relation is not internal, we have to know something beyond the intrinsic characters of the two things to know if the relation holds.

Truthful correspondence seems to be an internal relation. If we know everything there is to know about the intrinsic character of a proposition (including its content), and if we know everything there is to know about the intrinsic character of the part of the world that the proposition is about (its *worldly correlate*), we should able to tell whether the proposition is true or false. Moreover, the intrinsic character of the proposition is eternal and essential to it: propositions don't vary intrinsically from one time or one situation to another. So, the truth of the proposition must depend on the character of its worldly correlate. This is exactly the asymmetric dependency of the truth on the world that Aristotle mentions in the *Categories*, and it is why we find it natural to speak of 'truthmakers'.

Therefore, the Correspondence Theory is a theory about what truth consists in. For a proposition to be true is for that proposition to be in the truthful correspondence relation to some part of the world, its worldly correlate. So, we should be able to say something like this:

(1) For *p* to be true is for *p* to stand in the truthful correspondence relation to the worldly correlate of *p*.

Since the truthful correspondence relation is an internal relation, for some proposition *p* to stand in this relation to something is just for that thing to have the right kind of intrinsic character. For each proposition *p*, we should be able to say the following, where *X* stands for this intrinsic character or property:

(2) For *p* to be true is for the worldly correlate of *p* to be *X*.

If truth is a real property in the world, then truthful correspondence must also be a real relation. By 'real relation', we mean a relation that corresponds to some real, intrinsic similarity between different cases. A real relation has some unifying theme that ties together all of its instances. An unreal or merely nominal relation occurs in a highly variegated, gerrymandered set of instances. So, *being spatially next to* seems to be a real relation. Any two cases of one thing's being in contact with another are very similar to each other. On the other hand, being next to something in an arbitrary list, like a laundry or grocery list, is merely nominal.

For truthful correspondence to be such a real relation, the character *X* should be the same character for every proposition *p*. For it to be true that the Moon is cubical is for something to be *X*, and for blood to be red is for something to be *X*—the same *X* in both cases. Classical Truthmaker Theory provides just such an account. The worldly correlate of a proposition is some possible fact or state of affairs, and the *X* factor is existence or actuality. Thus, the worldly correlate of the proposition that blood is red is the possible fact of blood's being red, and for that proposition to be true is for that fact to exist. Similarly, the correlate of the proposition that the Moon is cubical is the possible fact of the Moon's being cubical, and for that proposition to be true is for that fact to exist. Since that fact does not exist (the Moon isn't cubical), the proposition is not true.

Is there any other *X* that could work as well as existence? Propositions can be about anything, and they can predicate any sort of property. This is why *X* cannot be a feature that is specific or limited to a particular topic or set of questions. For example, *X* couldn't be redness. Redness might work as an account of the truth of the proposition that blood is red. We could suppose that the correlate of the proposition is simply blood, and that for that proposition to be true is just for blood to be red. But then what could we do with the proposition that grass is green? There isn't anything *z* in the world such that for it to be true that grass is green is for *z* to be red. Nothing's being red can make it true that grass is green. For similar reasons, the *X* factor can't be any specific shape, size, or material composition. A feature like existence can work because it is generic and universal. Every proposition can be correlated with some possible fact, and it makes sense to suppose that what it is for that proposition to be true is for the possible fact to exist.

Are there any other alternatives to existence as the *X* factor? Are there other features that are equally generic and universal in scope? We might consider the relation of *exemplification* or *instantiation* (which we will discuss in more detail in Chapter 7). This is the relation that holds between a thing and a property when that thing has the property. So, we might suppose that the proposition that the Moon is round has two things as its correlate, the Moon and the property of roundness, and we might then hypothesize that what it is for the proposition that the Moon is round to be true is for one of the correlates (the Moon) to instantiate the other (the property of roundness). However, this account will only work for simple or *atomic* propositions. It doesn't tell us what to do with logically complex propositions, like negations or disjunctions. In addition, it requires us to assume that all atomic propositions have a subject-predicate or individual-property structure, and this might be disputed. Classical Truthmaker Theory doesn't have to make any such assumptions about the internal structure of propositions.

If the Correspondence Theory of Truth is correct, and if it holds for all propositions, and if the correspondence relation is a real relation that unifies all the cases of truth, then Classical Truthmaker Theory seems inescapable. However, many defenders of truthmaker theory argue that the theory has nothing to do with truth! (For example, see Horwich 1998 and Lewis 2001.) As the following four points will make clear, there are arguments for something like truthmaker theory that do not depend on the Correspondence Theory of Truth. In addition, there is a sixth argument for a specific type of Classical Truthmaker Theory, namely, *Atomic Truthmaker Theory*. Atomic Truthmaker Theory can provide a grounding for the distinction between fundamental and non-fundamental truths.

2. Without truthmakers we would be unable to discern whether theories differ with respect to what they say about the world. Theories describe at least part of the world. We can understand the claim that theories describe the world by noting that a theory is true if and only if the world is a certain way, that is, if and only if certain things exist, have certain intrinsic features, and stand in certain relations to other things. Clearly, different theories can describe the world differently. But equally, different theories can describe the world similarly by using different sets of fundamental terminology. Theories can be formulated, not just in two different languages, but by means of two different sets of propositions, with two different sets of fundamental concepts. For example, geometry can be formulated in terms of fundamental points, but equivalently by taking spheres or other solids as fundamental. This is a sort of cognitive or notional difference between the theories, as the two theories describe the world in equivalent ways. Though this may seem insubstantial, we can describe cases that don't seem quite so trivial.

Truthmakers offer a way to understand the difference between pairs of theories that describe the world similarly and pairs of theories that describe the world differently. Two theories are *ontologically* distinct when they describe the world differently, and two theories are merely *ideologically* distinct when they describe the world similarly using different terminology. Using truthmakers, we can give a clear analysis of these notions. Two theories are ontologically distinct if and only if there is a set of possible truthmakers that would make one theory true and the other false. Two theories are merely ideologically distinct if and only if any set of possible truthmakers that would make one theory true also would make the other true.

Suppose one theory uses the concept *bachelor* as one of its fundamental notions, and the other contains only the complex concept *unmarried male adult*. The first theory might entail that there are bachelors, while the second theory could not have any such implication (although it might imply that there are unmarried male adults). If every truthmaker for the proposition that there are unmarried male adults were also a truthmaker for the proposition that there are bachelors, then the two theories would differ only ideologically, not ontologically.

To take another example, suppose that one theory includes the implication that there are crowds, while a second refers only to people and their spatial relationships. Is this difference ontological or merely ideological? The correct answer to the question would depend on whether a truthmaker for the proposition that there is a crowd in the plaza must include some single entity corresponding to the crowd or could the truthmaker merely consist of a large number of individual people in close proximity in the plaza.

Is there any way to tell whether two theories differ ontologically without appealing to truthmakers? We might think that we could do this by simply asking whether it is possible for one of the theories to be false while the other is true. If we can construct a possible scenario in which one theory would be true while the other would be false, then we would know that the two theories were ontologically distinct, while if we could not construct such a scenario, we might reasonably conclude that the difference is merely ideological or verbal. This is the *modal method* of testing for ontological difference. It seems that we can employ the modal method without making any reference to truthmakers.

Suppose that theory *A* says that Austin is a part of Texas and that Texas is a part of the United States. Theory *B* says only that Austin is a part of the United States; theory *B* says nothing about the state in which Austin is located. We can easily show that there is an ontological difference between the two. Theory *A* has more information about the world than theory *B* because we can imagine a possible situation in which everyone would agree that theory *B* is true while theory *A* is false, like a situation in which Austin is made part of Oklahoma.

However, the modal method will not always give the right answer. For example, it won't work if we are considering two ontologically distinct theories about mathematical entities like numbers. Numbers do not and cannot change, so it makes no sense to talk about different possible scenarios involving them. The numbers 3 and 2 exist in every scenario, and 3 is greater than 2 in every scenario. Nonetheless, philosophers have proposed what seem to be clearly ontologically different theories about the numbers. For example, one might propose that numbers are really sets: the number zero is the empty set, and the number 1 is the set containing just the empty set, and so on. Another philosopher might disagree, arguing that numbers are not sets but are a quite disjoint class of mathematical objects. This is clearly an ontological difference, since the two theories disagree about what the numbers are. But there are no possible, alternative scenarios involving the numbers that we could use to show that the two theories carry different information about the world. If the number zero is identical to the empty set, then it would be so in every possible situation, and if it isn't identical to that set, it would be non-identical to it in every situation. So, each theory is (by the lights of its proponents) true in every possible situation. Thus, the modal method fails.

In contrast, truthmakers can be used to show that the two theories of the numbers are ontologically different. What is the truthmaker for the proposition that zero exists? The first theory entails that the truthmaker is the empty set, while the proponent of the second theory will deny this.

Another case in which the modal method will fail involves theories about *supervenient* facts. G.E. Moore (1903a) proposed that moral goodness is a special, non-natural quality. To say that a person or an action has this quality is to assert the existence of a moral fact, and such moral facts are distinct from all *natural* or non-moral facts (like facts about psychology or physiology). Nonetheless, Moore believed that moral qualities like goodness supervene on non-moral qualities, in such a way that it is impossible for two situations to differ morally without some non-moral difference. Now consider two possible theories, *M* and *N*. *M* is a theory about all of the moral and non-moral facts in the world. *N* contains exactly the same non-moral facts as *M*, but it contains absolutely nothing about moral facts. It seems clear that there is an ontological difference between the two theories. According to *M*, the ontology of the world is richer than it is according to *N*. However, given Moore's claim that moral facts supervene on non-moral facts, there is no possible situation in which *N* is true but *M* is not, and obviously there is no possible situation in which *M* is true and *N* is not (since *M* contains *N*). Thus, the modal method fails.

In contrast, Truthmaker Theory can be used to demonstrate the ontological difference between *M* and *N*. For *M* to be true, there must be some moral truthmakers, containing moral qualities, while no such moral truthmakers are needed to make *N* true.

It is difficult to mark the difference between ontologically distinct theories and theories that differ only ideologically without appealing to truthmakers. This gives us some reason to think that there are truthmakers, since it is clear that we *can* tell whether theories are ontologically distinct or are merely ideologically distinct.

3. Truthmaker Theory is needed to rule out metaphysical "cheaters." Some philosophers have argued that without truthmakers we have no way to combat theories which help themselves to putative truths without wanting to say anything metaphysically serious about the bits of the world to which those truths correspond.

(See Armstrong (1968: 85), and Sider (2001). For discussion, see Merricks (2007).) Metaphysical "cheaters" are people who appeal to such theories, theories that appear to capture some truth or class of truths without a serious metaphysical undergirding.

Here is an example of putative metaphysical cheating from the metaphysics of time (see Chapters 19–21 for more): Presentism (20.2T.4) is (roughly) the view that there are no objects that do not exist at the present moment. Eternalism (20.2A.1T), on the other hand, is (again, roughly) the view that there is nothing ontologically special about the present moment; all moments in time are equally real, and objects that exist only in the past or future really do exist. Presentists and Eternalists agree, for example, that dinosaurs don't exist *now*, but they disagree about whether dinosaurs exist. Eternalists think they do. Presentists think they don't.

A notorious problem for Presentists is making sense of the truth of ordinary, uncontroversially true sentences like 'Dinosaurs existed in the past'. This sentence seems to require the existence of some things—namely, the dinosaurs—that exist only in the past. But if there are these things, the dinosaurs, then dinosaurs exist (even if not in the present)! Some Presentists, in the face of this sort of problem, make something like this speech:

> The best way to understand the sentence in question is to say that there are irreducible tense operators in English. To say, "Dinosaurs existed in the past" is just to say, "PAST(Dinosaurs exist)," where PAST($p$) is true if and only if it was the case that $p$. But the thing about the PAST operator is that one is not ontologically committed to the existence of things that fall under its scope. It's like a belief operator in this way. If one allows that a person can believe that the Fountain of Youth exists, it doesn't thereby follow that one must believe there is this thing, the Fountain of Youth! Likewise, PAST (Dinosaurs exist) doesn't commit one to the existence of dinosaurs.

Suppose, however, that we're committed to the Correspondence Theory of Truth. Presentists have told us nothing about how it is that the sentence, 'Dinosaurs existed in the past', corresponds to reality. And it's hard to see how they could, given that there is no dinosaur and no past to correspond to if Presentism is true. Some argue, on this basis, that unless Presentists supply a truthmaker for this claim, a really existent thing that grounds the truth of the claim, then they are simply cheating. Truthmaker Theory gives us a way to catch metaphysical cheaters by demanding that theorists specify the truthmakers for claims they are committed to the truth of.

4. Truthmaker Theory can provide an explanation of possible truth. Ludwig Wittgenstein, in his *Tractatus Logico-Philosophicus* (Wittgenstein 1921/1974), employed truthmakers (in the form of *states of affairs*) as a way of explaining what it is for propositions to be possibly true. Suppose that each of the truthmakers is utterly separate and independent from the others. If so, it would seem natural to think that the existence or non-existence of each truthmaker has nothing to do with the existence or non-existence of the others. We could assign the values 'Exists' or 'Does not exist' to each possible truthmaker without reference to the assignment of these same values to the others. Consequently, every permutation of such values would represent a real possibility. This Tractarian account of possibility is sometimes called 'Combinatorialism' (15.3T), since each combination of logical atoms constitutes a real possibility (see Section 15.2.1.1).

5. Truthmakers as causes and effects. Some truths causally explain other truths. For example, that there were sectional differences in the United States over slavery is a cause of the truth that the Civil War occurred in the 1860s. Causation seems to be non-circular: facts don't causally explain themselves. One account of causation takes it to be a relation of a certain kind between truthmakers. When the truthmaker of the proposition that $p$ causes the truthmaker of the proposition that $q$, then the truth of $p$ is a causal explanation of the truth that $q$. We'll look at this account of causation in Chapter 27, on the relata of causation.

## 2.3 The Challenge of Deflationism

One of the fundamental motivations for truthmaker theory of any form is, as we have seen, the Correspondence Theory of Truth. But it is not clear that we really need the Correspondence Theory of Truth. Many philosophers defend a *minimal* or *deflationary* conception of truth, in which the predicate "is true" is explicated entirely in terms of Tarski's truth-schema (see **Principle of Truth 2**), first developed by the Polish logician, Alfred Tarski (1901–1983):

(3) 'Snow is white' is true if and only if snow is white.
(4) 'Grass is green' is true if and only if grass is green.

According to such deflationary accounts (see Ramsey 1927, Grover 1992, Horwich 1998), to say that some sentence $S$ is true is simply to assert $S$ itself indirectly, or by proxy. For example, we might say that everything Einstein said about relativity was true. To do so is not to ascribe some sort of property to Einstein's statements or to claim that these statements *correspond* somehow to reality or the facts. It is simply to endorse those statements, to commit oneself as if one had asserted those very things.

Importantly, deflationists are anti-realists only about one thing: the existence of a real property of truth. Deflationists can be as realist as one might like about other things, like the past, unobservable physical entities, consciousness, moral facts, and so on. Equally, deflationists can be anti-realist about any of those things. A commitment to deflationism involves only a commitment to the unreality of a metaphysically special property of truth and of metaphysically fundamental facts involving the truth of propositions. Typically, deflationists embrace classical logic, including the Law of Excluded Middle (for all *p*, either *p* or not-*p*). Thus, deflationists can accept that reality outruns our understanding or knowledge of it. They simply deny that we should flesh out our "robust sense of reality" in terms of a special property of truth.

The specter of deflationism is clearly a serious worry for advocates of Truthmaker Theory, given that the Correspondence Theory of Truth is such a crucial motivation for believing in truthmakers. Is there reason to conclude that deflationism is false?

The form of deflationism that is relevant to truthmaker theory is a metaphysical thesis: the thesis that there is no real property of truth (or of falsity). For deflationists, true propositions have nothing in common with one another.[7] They are as dissimilar to one another as the various parts of reality are. The truth that snow is white is as dissimilar to the truth that black holes are massive as the whiteness of snow is dissimilar to the massiveness of black holes. Deflationists deny that true propositions bear *any* resemblance to one another just by being true. Each truth is true in its own unique and incomparable way. The only thing that true propositions have in common is that they all bear the name "true" by virtue of the relevant pair of instances of the Tarski schema. Consequently, deflationists must deny that there is any property of truth (or relation of truthmaking) in any metaphysically robust sense. To use David Lewis's (1983) term, deflationists deny that truth is a *natural* property. The thinness of deflationist truth disqualifies truth from playing certain kinds of explanatory roles. Truth as such cannot figure (for deflationists) in any causal or metaphysical explanation, nor can it appear in a causal law or other law of nature.

There have been five major objections to deflationism in this metaphysical sense.

1. The methodology of science requires truth as a natural property. In science and other forms of inquiry, we are always trying to find methods, practices, and authorities that are reliable sources of information. For example, we have come to accept that many of the methods of established science are reliable in the sense that they have a high probability of generating true results (or, at the very least, a probability of doing so that is greater than mere chance). We take our past experience with such methods as providing good grounds for expecting future applications of the method to be similarly reliable in generating truths. We take our experience with science to provide us with "projectible" generalizations, to use the term introduced by Nelson Goodman (1954).

A projectible generalization must, as Goodman pointed out, make use of projectible predicates. We can make inferences about the future using predicates like "green" but not bizarrely gerrymandered predicates like Goodman's "grue" (a thing is "grue" if it has or will have been first observed before 2100 AD and is green, or will not be observed until 2100 AD and is blue). Lewis (1983) has argued that projectible predicates are those that signify natural properties. Since deflationists deny that truth is natural, they must deny that we can make projectible generalizations about truth, including generalizations about the relative reliability of various sources and methods.

2. The pursuit of truth is *constitutive* of belief and inquiry. Many philosophers have recognized that the pursuit of the truth is a *norm* for certain kinds of cognition and uses of language (Williams 1973, Dummett 1980). We are *supposed to* aim at truth when we make assertions, form beliefs or decide upon methods of inquiry. Deflationism is consistent with recognizing the normative character of truth, since it isn't obvious that a norm must make use only of natural properties. However, it seems that the pursuit of truth is a special kind of norm: a constitutive or essential norm. It is of the very essence of assertion and belief that they aim at truth. Acts of belief that have nothing to do with pursuit of truth seem to be a metaphysical impossibility. It seems that the property of truth is partly constitutive of our powers to believe and to assert things. In other words, the property of pursuing the truth is a part or constituent of the power of believing, and the property of truth is a constituent of the pursuit of the truth.

Deflationists may reply that social practices can be constituted by norms that make use of non-natural properties, like truth. There is no limit to how complex socially constructed norms can be. Consider, for example, fantastically complicated games like cricket or baseball. Why couldn't belief and assertion also be constituted by such complex norms, norms that apply each instance of the Tarski schema to our practices?

Truthmaker Theorists can reply that belief and assertion are themselves fundamental and natural functions, not wholly created by arbitrary conventions. Conventions can vary widely from one social context to another, but we would expect to find belief and assertion in a wide range of actual and counterfactual situations, even situations involving non-human species. It seems reasonable to suppose that such natural functions are constituted by natural norms, norms involving simple, real properties, not grue-like constructions.

3. The property of truth is needed to account for *truth-value gaps*. Some apparent assertions fail to be either true or false. Such assertions suffer from *truth-value gaps*. Some examples of truth-value gaps involve failures of presuppositions of various kinds:

(5) The present king of Texas is bald.
(6) It is taboo to step on the emperor's shadow.
(7) It is noon on the surface of the sun.

Since there is no present king of Texas, it is neither true nor false to say of him that he is bald. If the supposed property of *being taboo* doesn't really exist, then it is neither true nor false to say that it characterizes some action. Statements about the hour of the day presuppose that the statement is being made within some time zone on the Earth's surface.

Some philosophers have argued that statements involving vague predicates or referring to vague entities are neither truth nor false:

(8) Robert Duvall is bald.
(9) Austin is a large city.

Deflationists have trouble accounting for truth-value gaps. It is natural for deflationists to explain the use of the word 'false' in a way parallel to their account of 'true':

(10) 'Snow is white' is false if and only if snow is not white.
(11) 'Snow is green' is false if and only if snow is not green.

If deflationists accept the laws of classical logic, including the Law of Excluded Middle, they will be forced to endorse every instance of the following argument:

1 Either *p* or not-*p* (the Law of Excluded Middle).
2 *S* is true if and only if *p*. (Tarski's schema, where '*S*' is the name for '*p*'.)
3 *S* is false if and only if not-*p*. (the falsity counterpart to Tarski's schema.)
4 Either *S* is true or *S* is false.

In order to avoid this argument, deflationists have to restrict the application of Tarski's schema to those sentences that succeed in *expressing a proposition*. Then they could attribute truth-value gaps to (5) through (9) by denying that those sentences succeed in expressing propositions. However, deflationists would then owe some account of what a proposition is. It would seem to be part of the very essence or nature of a proposition that it be true or false, and such a metaphysical account of propositions would again require truth to be a natural property.

4. A property of truth is required for the contextualist solution to the Paradox of the Liar. Finally, we turn to the ancient Paradox of the Liar (first discussed by the Greek philosopher Epimenides). Consider (12):

(12) Statement (12) is not true.

If we assume that (12) is true, then we quickly find ourselves in the contradiction that (12) is both true and not true. However, it is equally difficult to affirm that (12) is not true (for whatever reason), since the non-truth of (12) seems to be exactly what (12) is affirming. If (12) is not true, then things are as (12) states them to be, so (12) is true after all. Tarski proved this result formally:

1 'Statement (12) is not true' is true if and only if statement (12) is not true. (T-schema)
2 'Statement (12) is not true' = statement (12). (Stipulated identity of 'statement (12)')
3 Statement (12) is true if and only if statement (12) is not true. (1, 2, the substitution of identicals)

Introducing truth-value gaps won't resolve this problem. Suppose we say that (12) fails to express a proposition. If (12) doesn't express a proposition, then it isn't true. But this is just what (12) says—so it seems to express a proposition, in fact a true one, after all.

There have been many attempts to solve this paradox, but there is no consensus about which is correct. Most approaches to the Liar are consistent with deflationism, but there is one popular approach that is not: the context-sensitive solution of Burge (1979), Barwise and Etchemendy (1987), Koons (1992), Simmons (1993), and Glanzberg (2001, 2004).

On the context-sensitive approach, when we assert that some assertion is true, we are claiming that it corresponds to some *part* of the world. Which part of the world we can refer to shifts from one context to another. Thus, we must interpret statement (12) as tacitly asserting something of this form:

($12_S$) Statement ($12_S$) does not correspond to any part of *S*.
(Where *S* is some contextually indicated part of the world.)

We can now recognize that statement ($12_S$) is true because it corresponds to some part *S'* of the world outside of *S*. *S'* is a part of the world that includes (as *S* does not) the fact that no part of *S* corresponds to ($12_S$). This solution requires the existence of truthful correspondence as a real relation between statements (or propositions) and parts of the world and is therefore inconsistent with deflationism.

5. The theory of truth is not a conservative extension of truth-free theories. Stewart Shapiro (1998) points out that reasoning by means of truth "non-conservatively extends" theories that do not involve truth. That is, we are able to reach novel conclusions by deploying obvious facts about truth. For example, the mathematician Kurt Gödel proved that no consistent mathematical system (like Peano arithmetic) can prove its own consistency. However, we can use truth to do so:

1 All of the axioms of Peano arithmetic are true.
2 Every set of truths is mutually consistent (in the sense that no contradiction can be proved formally from them).
3 Therefore, Peano arithmetic is consistent.

Deflationists can account for the truth of premise 1: asserting 1 is simply to assert once again the axioms of arithmetic themselves. However, premise 2 is a problem for deflationists. It is a mathematical assertion about a class of sets of sentences. The axioms of Peano arithmetic are complicated enough that it is by no means obvious that one couldn't derive a contradiction from them, assuming that one sets aside the facts expressed by premises 1 and 2. Hence, the property of truth provides us with real insight into mathematics in a way that deflationists cannot explain.

Deflationists could respond to this argument (as have Field 2001 and Azzouni 2008) by claiming that the device of using the word 'true' enables us to express more than we could express without it, even though there is no real property of truth. Consider especially premise 2 of the argument above. Deflationists should claim that we can assert premise 2, even though the content of premise 2 goes beyond our present mathematical knowledge, because we know in advance that however knowledgeable we become about mathematics in the future, we will never assert two contradictory claims about the numbers. We are in effect committing ourselves (and all future mathematicians) to a certain general policy: never to assert both of a pair of contradictory statements.

The problem for deflationists is explaining how it is *reasonable* for us to commit ourselves in advance to such an open-ended policy. Shouldn't we have to consider each contradictory pair on a case-by-case basis, if deflationism is true? This raises an even more fundamental problem for deflationists of explaining how we know the fundamental laws of logic, like the law of non-contradiction. The logical pioneer Gottlob Frege classified these fundamental laws as "laws of truth," and with apparently good reason. It is because we grasp something about the property of truth (and falsity) that we can say with confidence that no proposition whatsoever could be both truth and false. In light of that confidence, we can sensibly affirm the law of non-contradiction as a general law without ever having to consider the specific cases to which we apply it. Deflationists lack any similar story about what grounds our knowledge of the absolute generality of classical logic. They can always posit that such knowledge is simply constitutive of rationality, but such a stipulation counts against the simplicity of their theory.

We can appeal here to a principle known as 'Ockham's Razor', after the English scholastic philosopher William of Ockham. Ockham's Razor directs us to prefer the simplest theory consistent with the known facts. One way in which a theory can be simpler is in positing fewer basic postulates of reason. This is the first corollary of Ockham's Razor:

**Principle of Methodology (PMeth) 1 Ockham's Razor.** Other things being equal, adopt the simplest theory.

**PMeth 1.1 First Corollary of Ockham's Razor: Minimizing Rational Postulates.** Other things being equal, prefer the theory that posits the fewest primitive, underivable postulates of reason.

## 2.4 Truthmaker Maximalism

We have seen some reason to believe that there are truthmakers, so let's suppose that there are truthmakers. (We worry about the reasons, and consider an alternative to classical truthmakers, in Section 2.4.1.) The simplest, most natural understanding of Truthmaker Theory is the view that every truth has a truthmaker. This is 'Truthmaker Maximalism' (sometimes just 'Maximalism'):

**2.1T.1 Truthmaker Maximalism.** Every truth has a classical truthmaker.

Truthmaker Maximalism possesses some impressive theoretical virtues. In particular, it has great simplicity and explanatory power. Given Maximalism, no question arises about how it is that any truth is true, and Maximalism accounts for each truth in a uniform, straightforward manner. We will need to keep these virtues in mind as we consider alternatives to Maximalism.

### 2.4.1 Fundamentality and logically complex propositions

The first worry for Maximalism stems from the intuition that there must be a set of fundamental truths upon which all other truths depend. Take, for example, the sentence 'The cat is on the mat, and THP is sitting'. According to Maximalism, we must believe there is a truthmaker for this sentence *over and above* the truthmakers for 'The cat is on the mat' and 'THP is sitting'. This is counter-intuitive. Shouldn't the existence of the truthmakers for 'The cat is on the mat' and 'THP is sitting' be enough? The conjunctive truth just seems less fundamental than the two simpler, subject-predicate truths, a seeming that is emphasized when we realize that we don't seem to need another truthmaker for it. Why go in for the extra truthmaker, when the two we had are already sufficient?

However, there is a relatively simple change to Truthmaker Maximalism that will take care of this problem: require that, for every true proposition, there is either one thing that makes it true or there are *some* things that *jointly* make it true. In the case of a conjunction like 'The cat is on the mat, and THP is sitting', we can suppose that there are two facts that jointly make the conjunction true without having to suppose that there is a single, fundamental conjunctive fact.

What about negations? Suppose that Fido is not a cat or is not gray. What would the truthmakers for (13) or (14) have to be like?

(13) Fido is not a cat.
(14) Fido is not gray.

Given Maximalism, even of the modified variety, one would apparently need metaphysically fundamental *negative* truthmakers. Raphael Demos (1917) suggested that these could be made true by positive facts about Fido, facts that are incompatible with Fido's being a cat or being gray. Fido's being a dog or Fido's being white seem to fit the bill nicely. However, there are two problems with this suggestion. First, as Bertrand Russell pointed out (Russell 1918–1919, 213–215), it is not clear that we can make sense of the incompatibility relation without making use of purely negative truthmakers. Second, there are some negative predications that seem to be pure privations, in the sense that they don't require the thing to have any positive property at all. Consider, for example, (15) and (16):

(15) John is not thinking of anything right now.
(16) Mary does not remember Paris.

John does not have to be doing anything in order not to be thinking. Similarly, there doesn't have to be any relevant, positive state of affairs involving Mary's mind for it to be true that she simply doesn't remember Paris. Thus, negative propositions seem to require special, negative truthmakers. This is a serious enough problem to deserve its own section, which we will move to next.

### 2.4.2 The problem of negative existentials

Negative existentials, such as 'There are no unicorns' or 'There are no golden mountains', are an especially serious version of the negativity problem for Truthmaker Maximalism. (This problem involves as well universal statements, like 'Obama is the only president of the United States'. This sentence is equivalent to, 'There does not exist any president of the United States other than Obama'.) A truthmaker for such a sentence must be something that, by its very nature, excluded the possibility of adding a unicorn (or a golden mountain) to the world. In addition, it would have to be something that, as a matter of metaphysical necessity, had to exist whenever there are no unicorns (or golden mountains). Positing such truthmakers involves populating our theory with a large number of brute metaphysical necessities connecting separate things. The existence of the truthmaker for 'unicorns do not exist' somehow excludes the existence of any unicorn, and the absence of truthmakers for the existence of unicorns somehow entails the existence of the truthmaker of the negative existential claim. But Ockham's Razor demands that, other things being equal, we should minimize the class of brute necessities that we posit; this is the second corollary of Ockham's Razor:

**PMeth 1.2 Second Corollary of Ockham's Razor.** Other things being equal, adopt the theory with the fewest brute, inexplicable impossibilities and necessities.

Truthmaker Maximalism requires many such brute, inexplicable impossibilities and necessities in order to account for the truth of negative existentials. These necessary connections between positive and negative truthmakers are the sort of things that a simple and elegant metaphysical theory must minimize.

Further, Maximalism entails that at least one contingent thing exists. For each possible but non-actual contingent being, there must exist a truthmaker for the claim that *that* being does not exist. Here is an example. Let 'Winnie' name a particular unicorn that might have existed. The sentence 'Winnie doesn't exist' is true, since Winnie doesn't exist. According to Maximalism, there exists a truthmaker that makes this sentence true. Call this truthmaker, 'Un-Winnie'. Un-Winnie is a contingent thing. If Winnie had existed, the sentence 'Winnie does not exist' would have been false. But that sentence can't be false if Un-Winnie exists, since Un-Winnie necessitates the truth of that sentence, since Un-Winnie is that sentence's truthmaker. Thus it is possible that Un-Winnie fails to exist, and this is because Winnie might have existed. So Un-Winnie is a contingent being. It follows, then, that for every contingent being that fails to exist, there exists a truthmaker for the claim that that very contingent being fails to exist. These truthmakers are themselves contingent. Thus, at least one contingent being must exist. If Winnie doesn't exist, then Un-Winnie does, and vice versa. Even God couldn't create a world devoid of contingent beings, on the assumption that Maximalism is true.

On Maximalism, one cannot create a new possibility simply by thinning down an old one. If one deletes an entity, one must simultaneously add a truthmaker that makes it true that the deleted entity doesn't exist. This seems implausible. It should be possible to thin out the population of entities without being forced to introduce new ones in the process. Putting these together, Truthmaker Maximalism entails that there is some number *N* such that the number of contingent beings that exists is necessarily *N*. Implausible, indeed.

Another worry is that Maximalism requires a truthmaker for every universal generalization. Consider (17):

(17) Every living organism is terrestrial.

Suppose that (17) is true and that every living organism lives on the Earth. What would the truthmaker for (17) have to be like? At the very least, the truthmaker for (17) would have to include a part that is a truthmaker for proposition ($17_x$), for every existing thing $x$:

($17_x$) Either $x$ is not a living organism, or $x$ is terrestrial (or both).

However, even this massive truthmaker (call it 'Max') is not sufficient to be a truthmaker for (17). The existence of Max is consistent with the existence of a new entity, one that does not in fact exist at all (call it ET), and that is both alive and non-terrestrial. Max determines that all the things that actually exist are either terrestrial or not alive, but it does not exclude the existence of additional entities, entities that could but do not in fact exist. Since Max does not exclude such "new" entities, it is compatible with a world that is just like this one, except that some new entities exist that are living extra-terrestrials.

Thus, we have to add to Max what David Armstrong calls "the totality fact," a truthmaker that guarantees that nothing exists except the things that actually exist. Suppose the set *T* contains every actually existing thing. Then the totality fact is the fact that nothing exists except what is in *T*. The totality fact "says," in effect, "That's all, folks. Nothing but the members of *T*." However, the totality fact is quite a strange entity. As we've seen with negative facts, to believe in the totality fact we would have to believe in brute necessities. The existence of totality would have to be metaphysically incompatible with the existence of anything outside of *T*.

In addition, (as Armstrong admits) it is hard to see how we could ever come to know the totality fact in any detail. Only an omniscient being (like God) could possibly know the set *T*, the set containing exactly the things that exist. However, if we do not know the totality fact, how could we know any universal generalization, even one as simple and everyday as (18)?

(18) All ravens are black.

The truthmaker for (18) will also include the totality fact, which is needed to exclude the existence of any "new" entities that are non-black ravens. If we can't know the totality fact, we can't know the truthmaker of (18), which seems a problematic result. We could put the problem in the form of a dilemma: either there is just One Big Totality Fact for all of reality, or there are many, merely local totality facts, like the fact that *S* contains every raven, or that $S'$ contains every mammal in this building. If there were only *One Big Totality Fact*, then every bit of knowledge that we have of negative existential facts would have to involve some familiarity with this Big Fact, which seems wildly implausible. On the other horn of the dilemma, if we suppose that there are many localized totality facts, then knowledge of negativities is unproblematic, but we must posit a huge number of brute necessary connections between the different totality facts. For example, if there is the fact that *S* contains every raven, and another fact that *T* contains every existing thing, then it is necessary that *S* be a subset of *T*. Moreover, if *E* is an atomic fact, for example, the fact that some particular raven, Edgar, is a raven, then *E* must (of necessity) be a part of *S* (the actual totality fact for ravens). Other things being equal, we should try to minimize the class of necessary connections (**PMeth 1.2**). Finally, as Merricks has pointed out, truths like (18) do not seem to be *about* the positive character of the whole universe. (18) is not about how many quasars there are, for example. Yet the totality fact would include every detail about every part of the universe.

There is, however, at least one way around these difficulties. We could suppose that all facts are totality facts, and that each totality fact is associated with some fundamental property (the sort of thing that we will identify as a "universal" in Chapter 7). Thus, every fact will have the following form:

The instances of *U* are (exactly): the *x*'s.

Or, equivalently:

The instances of *U* are exactly the members of class *C*.

So, if being a raven and being black are two fundamental properties, there will be totality facts for ravens and for black things. However, there need be no totality fact for the whole universe, or for complicated, non-fundamental properties (like being a black raven). In addition, there need be no atomic facts at all. If Edgar is a particular raven, then the totality fact for ravens will be a truthmaker for the proposition that Edgar is a raven (and similarly for all other atomic predications of ravenhood), since Edgar will be a member of the class of ravens. Now, we won't have any necessary connections between distinct truthmakers. Each totality fact will be logically independent of all the other totality facts. Totality facts will be truthmakers both for positive predications (like 'Edgar is a raven') and for negative ones (like 'Fido is not a raven'). More precisely, the truthmaker for the proposition that Fido is not a raven will consist of the totality fact for ravenhood along with Fido itself. A truthmaker for the proposition that some things are not ravens will be the totality fact for ravenhood, together with one or more particular things that aren't ravens.

There is now only one necessary connection left: each universal must have only one totality fact. There cannot be two different sets each of which includes all of a universal's instances. However, this isn't a brute necessity, since we could suppose that it is something to do with the nature of universals that explains why no universal can have two different totality facts.[8]

We have arrived to a version of Truthmaker Maximalism: Totality Fact Maximalism.

**Def D2.4 Totality Fact.** A *totality fact* is a connection between a fundamental property (or universal) *U* and a class of entities *C* that is a strict truthmaker for the proposition that the class *C* contains all of the instances of *U*.

**2.1T.3 Totality Fact Maximalism.** Every true proposition has a truthmaker, which includes one or more totality facts, possibly together with one or more ordinary existing things. Each universal is associated with at most one totality fact.

What do we do with the problem of negative existentials? The true proposition that there is no golden mountain isn't a problem: it will be made true (jointly) by the totality fact for mountains and the totality fact for golden things (assuming that these are both fundamental properties). Similarly, the true proposition that all ravens are black will be made true jointly by the totality facts for ravens and for black things. What about a simple negative existential, like the proposition that there are no unicorns? If we assume that being a unicorn is a fundamental property, this could be a problem. The totality fact for unicorns will "say": the instance of unicornity are: ______. What can we use to fill in this blank? There are no unicorns to put there!

There are at least six options here, none of which is entirely satisfactory:

1 The totality fact for an uninstantiated universal connects the universal to the empty set. This would mean that all totality facts are fundamentally about sets, which are abstract, mathematical objects. This doesn't seem right: the proposition that all ravens

are black shouldn't require positing a set, so neither should the proposition that no unicorns exist.

2 We could assume that each universal is self-instantiating. So, the universal UNICORN is a unicorn. But this seems to give the wrong answer to the question, How many unicorns are there? (There would be one, rather than zero.)

3 There are special negative facts, in addition to totality facts, one negative fact for each uninstantiated universal. These special facts would have to have necessary connections with other totality facts, in the sense that when a universal has an associated negative fact, it cannot have a totality fact (and vice versa).

4 There are no uninstantiated universals. So, UNICORN doesn't really exist (as a fundamental, natural property), and so there is no true proposition of the form 'there are no unicorns'. Many philosophers, following a popular interpretation of Aristotle's theory of properties, have embraced this view. However, it produces some inconveniences, as we shall see when we consider non-existent things in Chapters 12 and 15.

5 We could hypothesize that there is an *uninstantiated-universal totality fact*. This involves treating the property of *being a first-order universal with no particular instances* as having its own, higher-order universal U*.[9] This uninstantiated-universal universal will have its own totality fact of the form: These are all the uninstantiated first-order[10] universals: $U_1$, $U_2$, … If T* is this totality fact, then it will be a truthmaker for the non-existence of unicorns, since UNICORN will be among the universals contained in T*. This seems to be a viable solution, although it requires us to complicate our story by adding at least one higher-order universal (a universal with other universals, and not particulars, as its instances) to our theory. In addition, there would have to be necessary connections between the totality facts for second- and first-order universals.

6 We could make an exception Maximalism for true propositions asserting that a universal is uninstantiated. These true propositions have no truthmaker at all. Fortunately, these propositions are relatively rare. All ordinary true propositions will still have truthmakers. Still, making an exception undermines the claim that truth is absolutely natural and unified.

So far, we have been focusing on a special case of negative existentials: those that simply deny that anything in the world has a certain natural property (or universal). Let's look briefly at a more general case, that of universal generalization. To say that everything is Φ, where Φ is some complex property, is equivalent to saying that nothing is not-Φ. However, we can't dodge this problem, as we did in option 4 above, by simply supposing that there is no not-Φ universal, since the proposition that everything is Φ will still exist, even though there is no not-Φ universal. For example, suppose it is true that everything is material or spiritual. This means that nothing is neither material nor spiritual. Since the universals MATERIAL and SPIRITUAL exist, this true proposition exists and requires a truthmaker.

Similarly, option 3 (special negative facts for uninstantiated universals) won't work in this case, unless we are willing to add special negative facts for every uninstantiated property, no matter how complex. This would result both in a large number of fundamental facts and a very large number of necessary connections between distinct facts, the very things we are trying to avoid.

Consider a relatively simple universal generalization, such as the truth that all ravens are black. This truth is made true by the combination of the totality fact for ravens and the totality fact for black things (assuming that RAVEN and BLACK correspond to simple universals), since the totality of black things includes the totality of ravens. In fact, for any true generalization of the form 'all ravens are Φ', the corresponding truthmaker will consist of the totality fact for RAVEN, plus a truthmaker for each proposition of the form '*x* is Φ', where *x* is a raven. More generally, any true universal generalization whose antecedent clause is a conjunction or disjunction of universals will have a truthmaker of the same kind. We can call these the *positively bounded generalizations*.

Thus, the only difficult cases are those without antecedents (such as 'everything is concrete or abstract') and those whose antecedents that are not positively bounded, such as 'all non-ravens are beautiful'. Here again, we will have to resort either to the uninstantiated-universal totality fact (option 5), or to the supposition that there are no uninstantiated universals (option 4), or else make an exception to the truthmaker principle for such cases (option 6).

If we take a variant of option 5, we will need to add still one more totality fact: the Universal Totality Fact, one that lists all of the universals. We can now assemble a global totality fact G, consisting of the totality of all universals, the totality of all uninstantiated universals, and the totality facts for each of the instantiated universals. The union of all of the particulars showing up in the totality facts for instantiated universals will be the complete set of all actual particulars.

Alternatively, if we suppose that there cannot be any uninstantiated universals (just as there cannot be any particulars that do not instantiate anything), then we could instead make do with a first-order universal totality: a totality fact that includes all of the first-order universals. This universal totality, together with all of the totality facts for the first-order universals themselves, would define for us a universal domain of particular objects. Every particular will belong to at least one totality fact, and the universal totality fact ensures that every universal has been surveyed.

If we were to take option 6 instead, we could say that unbounded universal generalizations are true, not by having a truthmaker, but by virtue of the non-existence of a fact that would make them false. However, it seems that most of the universal truths in which we are interested in science and ordinary life are positively bounded generalizations. Thus, we could still maintain that all "ordinary" truths have truthmakers.

## 2.5 Alternatives to Truthmaker Maximalism

Given the difficulties with Truthmaker Maximalism, philosophers have proposed three alternative truthmaker theories that do not require classical truthmakers for every truth:

1 Atomic Truthmaker Theory. Only logically atomic or simple propositions have (classical) truthmakers.
2 Spectral Truthmaker Theory. Logically atomic propositions have *spectral* truthmakers: entities that make the atomic truth true, not by simply existing, but by existing and having an intrinsic character of a certain kind. (Josh Parsons 1999).
3 Truth Supervenes on Being. The truths of the world are fixed by fixing which things exist, and what natural properties and relations those things have (Lewis 2001).

### 2.5.1 Atomic Truthmaker Theory

The two worries we discussed in the last section (concerning logically complex truths and negative existentials) motivate the idea that we ought to restrict our truthmaker theory in a way that upsets Maximalism. Maximalism demands that every truth, even ones that seem to be far from fundamental, have a unique classical truthmaker. Both of the worries above suggest that this is wrong. One might, in response to these worries, simply restrict Maximalism in the most minimal ways possible. For example, you might keep Maximalism except for denying that negative existentials have unique truthmakers. This would accommodate the second worry. You might keep Maximalism except for denying that non-fundamental truths have unique truthmakers. This would accommodate the first worry. In the face of other worries along similar lines, one can just keep restricting one's Maximalism in the most minimal ways possible to avoid the worries.

If we were to make all these restrictions, the result would plausibly be Atomic Truthmaker Theory.

**2.1T.4 Atomic Truthmaker Theory.** Every atomic (simple, positive) truth has a (classical) truthmaker.

If only atomic truths have truthmakers, how can we account for the truth of complex propositions? Conjunctions (propositions involving 'and') and disjunctions (propositions involving 'or') pose no real problem. If $p$ and $q$ are simple truths, then we can explain why the conjunction '$p$ and $q$' is true: it is true by virtue of each conjunct's having a truthmaker. Similarly, the disjunction '$p$ or $q$', if it is true, is true by virtue of one or the other of its disjuncts having a truthmaker.

What about negations? If $p$ is a simple proposition, and Not-$p$ is true, what account can we give of its truth? It won't be true by virtue of having a truthmaker. If Not-$p$ is true, then $p$ is false, and so $p$ does not have a truthmaker. This gives us our answer: the ground of the truth of Not-$p$ is to be found in the *absence* of a truthmaker for $p$. Not-$p$ is true, not because it has a truthmaker, but because it doesn't have a falsity-maker.

One interesting fact that follows from Atomic Truthmaker Theory is that the set consisting of the property of being a true complex proposition *weakly supervenes* on the set consisting of the property of being a true atomic proposition.[11]

**Def D2.4 Weak Supervenience.** A set of properties $A$ *weakly supervenes* on a set of properties $B$ if and only if it is impossible for any two worlds to agree on which things have which $B$-properties but to disagree about which things have which $A$-properties. That is, two situations that are indiscernible in respect of the $B$-properties must also be indiscernible in respect of the $A$-properties.

An important special case of weak supervenience is that in which a set containing a single property, say $\{F\}$, supervenes on a set containing another single property, $\{G\}$. In that case, we shall say, for simplicity's sake, that $F$ supervenes on $G$. This means that whether anything is $F$ or not-F is determined by the set of things that are $G$ and the set of things that are not-$G$.

For example, you might think that the extension of the property of *being a true proposition about clouds* is completely determined by the extension of the property of *being a true proposition about the location of water molecules in the atmosphere*. It is plausible that one couldn't get a difference in truths about clouds without a difference in truths about the location of water molecules in the atmosphere. Once you've settled the truths about the water molecules, you've settled the truths about clouds. Similarly, if one thought that one couldn't get a mental difference without a brain difference, then one thinks that mental properties weakly supervene on brain properties.

If we know what the set of truthmakers for positive atomic truths is, then we know what the set of positive atomic truths is. Once we know what the set of positive atomic truths is, we know what the set of negative truths is (and similarly for the sets of all the logically complex truths). A negated proposition Not-*p* belongs to the set of negative truths just in case its positive counterpart *p* does *not* belong to the set of positive truths. You can't have a difference in the set of negative truths without a corresponding difference in the set of positive truths, and you can't have a difference in the set of positive truths without a difference in the set of truthmakers. Thus, the property of *being a negative truth* weakly supervenes on the property of *being an existing truthmaker*. Once we know which possible truthmakers exist, we know all there is to know about which negative propositions are true.

If we move from Truthmaker Maximalism to Atomic Truthmaker Theory, do the five arguments for truthmakers still apply?

1. Catching cheaters. Trenton Merricks (2007) argues that Atomic Truthmaker Theory cannot be used as a weapon against metaphysical cheaters. Atomic Truthmaker Theorists admit that some truths lack truthmakers: all complex truths, including especially negative truths. Thus, Atomic Truthmaker Theorists are themselves guilty of cheating metaphysically. How can they complain when, for example, Presentists deny that past-tensed truths have truthmakers?

Here is a possible response to Merricks. According to the Atomic Truthmaker Theory, the extension of the property of *being a complex truth* (*being a true negation*, *conjunction*, *disjunction*, and so on) is completely determined by the extension of the property of *being an existing truthmaker* (even though only positive, atomic truths have truthmakers). Once you have the information about which possible truthmakers exist, you can determine which propositions (including negative ones) are true. In contrast, assuming Presentism (see Section 20.4), there is little or no reason for thinking that the set of past-tensed truths is determined in this way by the set of truths about what truthmakers exist now. This lack of supervenience of the properties of past- or future-tensed truth on the actual properties of what really exists seems more problematic than the mere absence of truthmakers for every truth. Thus, Presentism seems guilty of somewhat worse cheating than Atomic Truthmaker Theory.

However, this difference doesn't seem so impressive once we realize that the definition of supervenience has built into it the assumption that negative facts are unproblematic: the set of negative truths is determined by the facts about which atomic positive propositions are *and are not* true. By the same token, Presentists can point out that the set of past truths is determined by the set of facts about which propositions are *and were and will be* true. There doesn't seem to be any difference here. Just as Atomic Truthmaker

Theorists help themselves to negation, Presentists can help themselves to the past and future tenses. Ultimately, Atomic Truthmaker Theorists likely cannot appeal to the catching cheaters argument.

2. Distinguishing ontological and ideological differences. Atomic Truthmaker Theorists can claim that this argument still works, despite the restriction to atomic truths. Two theories differ ontologically if they differ in what truthmakers do and *do not* exist.

3. Correspondence Theory. Atomic Truthmaker Theorists must give up, to some extent, their commitment to the Correspondence Theory of Truth, at least in the form we initially described. It is no longer the case that every truth corresponds to some truthmaker. Instead, we could say that every positive atomic truth corresponds to reality, and that the class of all other truths weakly supervenes on the class of atomic truths. This means that truth is a disjunctive or complex property, made up of two or more quite different components. This would come close to supporting deflationism.

4. Accounting for possibility. As we will see in Chapter 15, Atomic Truthmaker Theory is quite compatible with Logical Atomism's account of possibility.

5. Causation. If we adopt Atomic Truthmaker Theory, and we suppose that it is truthmakers (and only truthmakers) that can serve as causes and effects, this will have significant effect upon our conception of causation. In particular, it will entail that all causation is fundamentally positive. This will create a difficulty, as we shall see in Chapter 27, in accounting for *negative causation*, such as prevention or causation by omission.

6. Accounting for fundamental truths. It is natural to think, as we saw above, that some truths are more fundamental than others. For example, conjunctions usually depend for their truth on the truth of their parts (or *conjuncts*). Atomic Truthmaker Theory provides a simple account of which truths are fundamental: each fundamental truth corresponds to the existence of a single possible truthmaker. In other words, if a truth is fundamental, then it is made true by just one thing in the actual world, and, moreover, it is made true by the existence of that same thing in every world in which it is true. Suppose, for example, that we have a fundamental truth of the form '*b* is *F*'. According to Atomic Truthmaker Theory, this is made true by a single, simple truthmaker: *b*'s *F*-ness or the fact that *b* is *F*. Moreover, this truth can be made true only by the existence of this same truthmaker. There is no possibility in which '*b* is *F*' is true but *b*'s *F*-ness does not exist. We can call any proposition that could be made true only by one possible truthmaker a *fundamental proposition*.

Next, we have those truths that are negations of fundamental propositions. Suppose that *b* is not in fact *F*, but that '*b* is *F*' would be a fundamental truth if it were true at all. In that case, the proposition '*b* is not *F*' is true by virtue of the non-existence of the appropriate truthmaker for the more basic proposition that *b* is *F*.

Finally, all other truths derive their truth from the fundamental truths and from the true negations of fundamental propositions. For example, a disjunction of two more fundamental propositions *p* and *q* would be true by deriving its truth from that of *p* or of *q* or both. Such disjunctive truths would not be fundamental, since there is no one possible truthmaker that must exist in order for such propositions to be true: it is enough if one of a range of possible truthmakers exists. Similarly, if the predicate '*F*' is definable as '*G* and *H*', then the proposition that '*b* is *F*' would, if true, derive its truth from the two more fundamental truths (viz., that *b* is *G* and *H*). We take up this idea of fundamentality in more detail in the next chapter.

### 2.5.2 Spectral Truthmaker Theory

We are now in a position to see a serious problem for an important motivation we had for believing in truthmakers. Recall the first argument for truthmakers, from the Correspondence Theory of Truth. That argument implicitly assumed that the things to which propositions correspond are facts.[12] This is how we guaranteed that clause (i) of Def D2.1 (the definition of a truthmaker) was satisfied by the things to which propositions correspond. How could, we asked rhetorically, the fact that grass is green exist without the proposition that grass is green's being true? What this assumes is that *facts* (in this sense) are among the metaphysically fundamental bits of the world. But we have seen no reason to think this is true!

Suppose, for example, that it's just *the grass* that's among the metaphysically fundamental things. At rock bottom, metaphysically, you've just got the grass with its features. In that case, the fundamental things of the world might not satisfy clause (i) of the definition of a truthmaker. Grass might not have been green, so the sheer existence of the grass is not sufficient to make true the proposition that grass is green. There is another requirement: the grass must also be a certain way.

**2.1A.1T Non-Classical Truthmaker Theory.** Propositions are made true by the way things are in the world, but there are no classical truthmakers.

**2.1A.1A No Truthmakers.** Truths have no truthmakers, classical or otherwise.

If we give up the idea that is is just the existence or non-existence of certain things that make propositions true and false—an assumption we might call Binarity—but we want to hold on to as much of Truthmaker Theory as possible, we end up with something we call 'Spectral Truthmaker Theory'. This name is meant to denote the fact that the fundamental parameters form a spectrum of possible values, not simply a binary choice between existence or non-existence. A theory of this kind has been proposed by Josh Parsons (1999). Parsons argues that the mere existence of a truthmaker is not sufficient to entail the truth of the corresponding proposition. Instead, we should think of a proposition's truthmaker as an entity of such a kind that its being intrinsically the way it is, is sufficient for the truth of the proposition. That is, it is both the existence and the actual intrinsic character of the fundamental thing that, taken together, ground the truth of the proposition.

**2.1A.1T.1 Spectral Truthmaker Theory.** Every fundamental atomic truth is made true by something's existing *and being a certain way intrinsically*.

It is the italicized phrase that distinguishes Spectral Truthmaker Theory from Classical Truthmaker Theory. A classical truthmaker makes a proposition true simply by existing. In contrast, a spectral truthmaker makes a proposition true both by existing and by having a certain intrinsic character. So, for example, a blade *B* of grass cannot be a classical truthmaker for the proposition *that B is green*, since *B* could exist without being green (it could be brown, for example). However, blade *B* could be a spectral truthmaker for this same proposition, since whether *B* is green or not is intrinsic to *B*.

We can illustrate the advantage of Spectral Truthmaker Theory over Classical Truthmaker Theory by considering the phenomenon of *determinates* and *determinables*. Determinable properties are properties that come in different kinds, and those different kinds are a determinable's determinates. For example, the property of *being a certain temperature* is a determinable property, with determinates like the properties of *being 10° Celcius*, of *being 30° Celcius*, and so on. Similarly, the property of *being red* is a determinable property, with determinates like the properties of *being scarlet*, of *being burgundy*, and so on. If an entity has a determinable property (e.g., temperature), then it must have exactly one corresponding determinate (a specific value of the temperature variable, like 10° C). It cannot have two determinate temperatures or none at all. On Classical Truthmaker Theory, the independence of temperature from other parameters or determinables is explained by the existence of a separate truthmaker for a thing's temperature, an accident or trope of temperature. However, Classical Truthmaker Theory predicts that each specific temperature is independent of every other. If body *B* has a truthmaker for its having a temperature of 10°, it should be possible for it to have simultaneously a truthmaker for the determinate temperature of 15°. The impossibility of having two determinate temperatures at the same time is, relative to Classical Truthmaker Theory, a brute or unexplained necessity.

In contrast, Spectral Truthmaker Theory does not over-generate interdependencies, since it doesn't posit distinct and separate truthmakers for each determinate property. Instead, it entails the existence of just one possible truthmaker for each parameter of each entity, with the specific or determinate property corresponding to some *internal state* of that truthmaker (not simply to its existence). This gives us no reason to expect it to be possible for one thing to have two determinate temperatures at once, since there is no reason to think that the internal state of one truthmaker could correspond to two different temperatures.

The main disadvantage for Spectral Truthmaker Theory, as compared with Classical Truthmaker Theory, is that it does not provide a unified definition of the correspondence relation. Consequently, it isn't clear that correspondence is a real relation on Spectral Truthmaker Theory, a relation corresponding to a real similarity between distinct cases of truth. Consider the propositions that blade *B* is green and that the Moon is round. The spectral truthmaker of the first proposition is the blade *B*, and the spectral truthmaker of the second is the Moon. What it is for the first proposition to be true is for its truthmaker to be green, and what it is for the second to be true is for its truthmaker to be round. Thus, truth for the first proposition is something entirely different from truth for the second proposition. This would seem to support a deflationist view of truth.

If we accept Spectral Truthmaker Theory, we can ask how many truthmakers we need, given a set of fundamental truths. Could there be just one Big Truthmaker that is responsible for making true all of the truths of the world? This possibility runs contrary to the underlying spirit of all truthmaker theory, which is to ground truths in relevant parts of reality. At the very least, it seems that if there are two different, equally fundamental properties being predicated in two truths, those truths ought to have distinct truthmakers:

**Principle of Truth (PTruth) 1 One Truthmaker per Fundamental Property.** If $p$ is the true predication of a fundamental property $P$ to $x_1$ through $x_n$, and $q$ is the true predication of a different fundamental property $Q$ to the same things $x_1$ through $x_n$, then $p$ and $q$ have distinct truthmakers.

(One Truthmaker per Fundamental Property will become important in later chapters.)

### 2.5.3 Truth Supervenes on Being

There is another, more radical way of replacing individual truthmaking with plural or joint truthmaking. This view drops the truthmaking relation entirely and makes use instead of the idea of weak supervenience introduced above. Truth Supervenes on Being (TSB) is such a view. TSB claims that there is a special class of ultimately real or fundamental entities $E$ and a special class of perfectly natural or fundamental properties *and relations* $N$, and that the fundamental propositions are those atomic predications involving only fundamental entities and natural relations.

**2.1A.1A.1T Truth Supervenes on Being (TSB).** The property of *being true* weakly supervenes on the property of being a truth about what things exist (and don't exist) and about exactly which natural properties they have and which natural relations they stand in.

**2.1A.1A.1A Truth does not supervene on being.**

The main difference between Spectral Truthmaker Theory and TSB is that Spectral Truthmaker Theory is committed to One Truthmaker per Fundamental Property while TSB is not. Suppose that the various shapes and colors are natural properties and that we have a ball that is spherical and red. For Spectral Truthmaker Theory, there must be two entities, which we can call the ball's 'shape' and its 'color', one of which is a spectral truthmaker for the ball's being circular and the second of which is a spectral truthmaker for the ball's being red. TSB has no such commitment. The ball's being red and its being circular must both supervene on how things are, but there is no implication that the things involved include such finely individuated things as the ball's shape and its color. TSB theorists can simply suppose that the things involved include only the ball itself, which, by being a certain way, ensures that both propositions are true.

We've seen what TSB permits. What sort of things does it forbid? TSB prohibits any fundamental truth that does not involve actually existing things. Consider, for example, Bucephalus, the horse owned by Alexander the Great. Bucephalus no longer exists, having long since died. If we suppose that the only things that exist are things that exist now, that is, assuming Presentism, then we would have to say that nothing is Alexander's horse. Consider now (19):

(19) Bucephalus was fierce.

If Bucephalus does not exist, TSB theorists must deny that (19) is a fundamental truth. They must say one of two things about (19). Either it is in fact true that Bucephalus exists (even though it is not now alive)—a denial of Presentism—or the truth of (19) must consist in facts about other existing things, such as memories, records, or other traces of Bucephalus. The truth of (19) cannot simply float free of all the facts about existing things and their natural properties and relations. Since the second option is implausible (it is hard to believe that a fact like (19) is really made true by facts about remains and records), TSB is often taken to be incompatible with Presentism (an issue we will take up again when we consider theories of time in Chapters 19–21).

Here's another illustration of what TSB forbids, taken from Gilbert Ryle's theory of behavioral dispositions (Ryle 1949). According to Ryle, behavioral dispositions correspond to conditionals such as (20):

(20) If Roy is frustrated, he will curse loudly.

According to Ryle, conditionals like (20) can be true, even though there is nothing further to be said about why they are true or what makes them true. That is, there might be nothing about Roy's mind, body, or brain that could be expressed categorically (without making use of conditional statements) that is sufficient to guarantee the truth of (20): nothing about Roy's current memories, moods, feelings, thoughts, or neurochemistry. Assuming that conditional dispositions do not count as natural properties, then Ryle's theory entails that (20) can be true even though its truth does not supervene on any natural properties of Roy or any other things in Roy's environment. This violates the constraints of TSB.

The bite of TSB depends on the content of the idea of natural properties and relations. If just any relation is natural, then TSB ceases to be a substantive doctrine. Suppose, for example, that there were some truth $p$ that intuitively doesn't supervene or depend upon being. If we were to count the property denoted by 'being an $x$ such that $p$' as natural, then any proposition $p$ whatsoever could easily satisfy the constraints of TSB: in every world in which $p$ is true, if at least one thing exists in that world, then that thing has this property of being something such that $p$ is true (or, in other words, being something that coexists with the truth of $p$). In the case of (20), suppose we thought that the property of *being such that Bucephalus once was fierce* was a natural property of presently existing things, like the Moon or the Eiffel Tower. On that assumption, TSB and Presentism would be compatible, after all. The same result would ensue if we treated as natural such properties as 'are some atoms that once composed a fierce horse known as "Bucephalus"'. If that relation among atoms is natural, then the truth of (20) supervenes on this relation's holding between some currently existing atoms. So, a serious defender of TSB should not count such bizarre properties as natural.

Let's go back one more time to the six arguments for truthmaker theory and see which of them provides some support for TSB (as we did above for Atomic Truthmaker Theory).

1. Catching metaphysical cheaters. Catching metaphysical cheaters is the principal motivation for TSB. As we've seen, some supposed "cheaters" violate TSB. However, as we saw in the discussion of Atomic Truthmaker Theory, it is unclear that either the Atomic Truthmaker Theorist or TSB theorists can consistently apply the TSB standard, given that both allow negative truths to be ungrounded by any existent thing.

2. The Correspondence Theory of Truth. TSB Theorists can claim, with some plausibility, that TSB captures something of our common sense intuition that truths "correspond" to reality. TSB theorists cash out this correspondence in terms of difference-making: in order for a true proposition to be false, or a false proposition to be true, different things would have to exist, or existing things would have to stand in different natural relations to one another. A true proposition is one that corresponds with how things in general are.

3. Distinguishing ontological and ideological differences between theories. On TSB, two theories will be merely ideologically different when there is no difference between the theories with respect to what exists and what natural properties those things have and what natural relations they stand in. Thus, TSB can be used to distinguish ontological and ideological differences between theories.

4. Accounting for possibility.

5. Accounting for the relata of causation.

6. Accounting for fundamental truths. We'll lump the final three arguments together, since TSB theorists cannot appeal to any of them. Without truthmakers, they have no basis for a combinatorial theory of possibility. Without truthmakers, they have no special account to provide for the relata of causation nor can they account for the difference between fundamental and derived truths.

Thus, TSB theorists must rely on one motivation, a sense that the TSB constraint is needed to take seriously the intuition that truths "correspond" to reality.

## 2.6 Conclusion and Preview

Recall that there were five initial arguments for truthmaker theory: (1) the appeal to the idea of correspondence, (2) distinguishing between ideological and ontological differences, (3) catching metaphysical "cheaters," (4) accounting for metaphysical possibility, and (5) providing the relata for the causal relation. Along the way, we added a problem for

**Table 2.1** Comparing Truthmaker Theories

| *Advantages* | *Truthmaker Maximalism* | *Atomic Truthmaker Theory* | *Spectral Truthmaker Theory* | *Truth Supervenes on Being* |
|---|---|---|---|---|
| 1. Correspondence Theory of Truth | Yes | Yes, but only for atomic truths | Yes, for atomic sentences, given instantiation | Yes, but only in terms of difference-making |
| 2. Distinguishing theories | Yes | Yes | Yes | Yes |
| 3. Catching "cheaters" | Yes | Arguably by supervenience | Arguably, by supervenience | Arguably, by supervenience |
| 4. Simple Combinatorial Theory of Possibility | Yes | Yes | Yes | No |
| 5. Causal relata | Yes | Only for positive causation | No | No |
| 6. Identifying fundamental truths | No | Yes | Yes | No |

Maximalism that motivated Atomic Truthmaker Theory, namely, identifying the fundamental truths. We've considered four versions of truthmaker theory: Truthmaker Maximalism, Atomic Truthmaker Theory, Spectral Truthmaker Theory, and Truth Supervenes on Being. Table 2.1 summarizes our results.

We have postponed a discussion of some issues directly relevant to various truthmaker theories until later portions of the book (e.g., the nature of the causal relation, the issue of so-called 'natural' properties and relations, the question of just what facts might be). And issues about truthmaking will continue to crop up throughout later chapters. In particular, Part II takes up one area where truthmaking looms large: conditional statements, powers, and laws.

In the next chapter (Chapter 3), we will look at a way of extending the idea of fundamentality beyond the relation between propositions into the world itself, through the idea of metaphysical *grounding* or *explanation*.

## Notes

1 Philosophers distinguish between types and tokens. Standardly, tokens are time- and place-specific instances of some type of thing. For example, in the following list, there are two tokens of a single word type: BEN, BEN. An individual cat is a token of the type, CAT. And so on.

2 Some philosophers would question whether a single group of things, the *propositions*, really does play these two roles. David Lewis (1986a) famously did so, arguing that the roles were at cross-purposes.

3 It is important to note that Aristotle's use of 'cause' here is not limited to so-called 'efficient' causation, that type of causation involved when a baseball causes a window to break or when one billiard ball causes another to move.
4 We'll consider an alternative to the Correspondence Theory—the deflationary theory of truth—in Section 2.3 below. While there are alternatives to both the Correspondence and deflationary theories, notably the 'coherence' and 'pragmatic' theories, we believe the two we consider are the only theories with much hope for success at the end of the day. Thus we ignore their alternatives.
5 This argument assumes that if something is a truthmaker for a proposition, then it is essentially so. In other words, it assumes that if something $x$ makes it true that grass is green in the actual world, then $x$ makes it true that grass is green whenever $x$ exists. As we shall see below, this is exactly the step that will be challenged by defenders of Spectral Truthmaker Theory.
6 See Section 9.3.1.1 for a further discussion of intrinsicality.
7 There is a possible intermediate position, in which there is no one property of truth, but rather a finite number of distinct kinds of truth (such as, perhaps, physical truth and historical truth). We will count such a position as a version of anti-deflationism. Thus, we will assume that the deflationist denies that there is either a single property of truth or a finite number of such properties. We'll simply ignore the view according to which there are an infinite number of fundamental properties of truth, since such a theory would involve an egregious violation of Ockham's Razor.
8 In fact, we might even consider the possibility that some universals do have more than one associated totality fact, resulting in a kind of ontological indeterminacy or even inconsistency, that could be useful as a model of vagueness (see Chapter 12.2.2). Perhaps the universals that are of most interest and use to us (especially in science) are the ones that tolerate no more than one totality fact.
9 A first-order universal is a universal instantiated by ordinary objects; higher-order universals are instantiated by other universals.
10 Why can we limit ourselves to first-order universals? What can we do about uninstantiated higher-order universals. It is plausible that there are no such uninstantiated higher-order universals to worry about. All universals have their properties essentially, and it is plausible to suppose that they all exist necessarily. If so, all higher-order universals will have their non-empty classes of instances essentially.
11 Why *weakly* supervenes? Philosophers have defined a family of supervenience relations. Jaegwon Kim (1993) was responsible for classifying them into weak and strong versions.
12 This assumption was not entirely unwarranted, since it is plausible to think that truthmakers will have a metaphysical structure that mirrors the syntactic structure of the truth it makes true.

# 3

# Grounding, Ontological Dependence, and Fundamentality

Metaphysicians seek to understand the world, and a large part of that project is building an *ontology*: a theory of what exists, and of what those existing things are like. However, metaphysicians have traditionally sought more than this. They have also sought to uncover the fundamental structure of reality. This includes understanding what depends upon what, and how.

In addition, metaphysicians want to know, not just what exists, but what exists most *fundamentally* and *really*. These last two tasks may be closely related. It may be that what most fundamentally or really exists are those things that exist independently, and upon which all other things depend. At the very least, it seems clear that what exists fundamentally cannot depend for its existence on things that do not exist fundamentally. So, a theory of metaphysical dependence would greatly constrain our theory of what is fundamental.

Finally, as we have seen, an appeal to ontological parsimony or economy (PMeth 1) plays an important, perhaps indispensable, role in evaluating metaphysical theories. However, we want to minimize our ontological commitments to fundamental entities, not to derived ones. At the very least, economy with respect to fundamental entities seems to be of much greater importance than economy with respect to less fundamental ones:

**PMeth 1.0 The Zeroth Corollary of Ockham's Razor.** Other things being equal, prefer the theory that posits the smallest number of fundamental entities.

In recent years, many metaphysicians, following the lead of Kit Fine, have used the term 'grounding' to represent a relation of metaphysical dependency: if $x$ is *grounded* in $y$, then $x$ (in a certain sense) depends upon $y$, for its existence, or truth, or nature. We could identify fundamental entities or truths with those that are not grounded in other

*The Atlas of Reality: A Comprehensive Guide to Metaphysics*, First Edition.
Robert C. Koons and Timothy H. Pickavance.

entities or truths, either by being absolutely ungrounded or by being in some special way grounded without being grounded in or by anything.

Some new terminology will aid the present discussion. If $x$ grounds $y$, call $x$ the 'fundans' (plural: 'fundantia') and $y$ the 'fundatum' (plural: 'fundata'), from the Latin verb 'fundare', to found or ground.

There are a number of different theories about what the relation of grounding is like. They are not all pair-wise incompatible. Some of these theories may be combined, others cannot. We can identify at least three conceptions of grounding:

(1) Grounding is a kind of explanation: the metaphysical explanation of some facts or truths in terms of others.
(2) Grounding is a relation of dependency between entities that is generated by the essences or real natures of those entities: $x$ depends on $y$ just in case $y$ occurs as a constituent within the definition of $x$'s essence or real definition. We will call this relation 'ontological dependence'.
(3) Grounding is a relation of constitution or construction between entities: $x$ depends on $y$ just in case $x$ is constructed from $y$, or $y$ is an essential element in the constitution of $x$.

In this chapter, we focus primarily on the first conception of grounding, grounding as metaphysical explanation. We also discuss briefly (in Section 3.4) the relation of ontological dependency and its connections with grounding as explanation. We set aside the third conception, primarily because it presupposes an answer to a question that we examine critically in future chapters, namely, do wholes always depend on their parts, or are there cases where a part can depend upon some whole to which it belongs?

If grounding is a kind of explanation, what kind of explanation is it? Jon Litland (2013) suggests that it is a matter of explaining *how* something is the case. When $p$ grounds $q$, that $p$ is the case is a *way for $q$* to be the case. As we shall see, this talk of 'what is the case' can be understood in either of two ways: (i) in a conceptual or logical way, explaining the *truth* of one proposition in terms of the truth of other propositions, by reference to the essences of the conceptual components of the propositions involved, or (ii) in a purely ontological way, explaining the *existence* of one fact or the *actuality* of some state of affairs in terms of other facts or arrangements of things in the world. We investigate both conceptual and ontological grounding in Section 3.5.

We also explore (in Section 3.2) the relationship between grounding and truthmaking. Can one be defined in terms of the other? Do they form a kind of inter-definable circle? Is there any reason to prefer one over the other? We also consider various alternatives or competitors to grounding theory in Section 3.3.

Sections 3.6 through 3.8 take up several important questions about grounding:

- Can facts about grounding themselves be grounded? (3.6)
- Do grounds entail what they ground? (3.7)
- How is grounding different from causal explanation? (3.8)

We conclude in Section 3.9 by examining the connection between grounding, fundamentality, and Ockham's Razor (PMeth 1). We also point to some recent literature on the formal properties of grounding (transitivity, asymmetry, and well-foundedness).

## 3.1 Is Grounding Real?

In this section, we consider whether there is any reason to think that there is such a thing as grounding or metaphysical explanation.

**3.1T Real Grounding.** There is a relation of *metaphysical grounding*.

There are several reasons for thinking so, which we take up presently.

### 3.1.1 Connections with natural language

The notion that some facts are grounded in others is a matter of common sense, well marked in the conventional features of natural language. In English, we express a relation of grounding by means of expressions like 'by virtue of', 'in virtue of', 'thereby', and 'makes' (in a certain sense).

(1) Four is an even number *by virtue of* its divisibility by two.
(2) The word 'wounded' is in the past tense *in virtue of* its '-ed' ending.
(3) Mary crossed the finish line first and *thereby* won the race.
(4) Jamie's valid election to the office *makes* her the president of this club.

Latin actually has a separate case, the ablative case, to capture just this sort of meaning. It is important to note that none of these cases seems to involve a cause-and-effect relationship between two events or conditions. Being divisible by two doesn't *cause* four to be even. Two's evenness just *consists in* it being divisible by two. Similarly, it would be odd to say that Mary's crossing the finish line *caused* her to win the race. The causes of her winning must be separate, typically earlier events and conditions, like her rigorous training regimen or her well-regulated diet. To cross the finish line (in the right circumstances) just *is* to win the race.

### 3.1.2 Plausible examples of grounding

There are many actual cases that seem to be cases of grounding between facts or between other kinds of entities. Jonathan Schaffer (2009) gives some examples:

1 A singleton set, like $\{2\}$, the set containing just the number 2, is grounded in the existence of its member. The set exists, at least in part, because its member exists, but the member does not seem to exist because the set exists.
2 The shape or topological form of a piece of Swiss cheese is the ground for the existence, the number, and the arrangement of its holes. There is an obvious asymmetry here. Bits of Swiss cheese can exist and have a definite shape without any Swiss-cheese holes, but Swiss-cheese holes cannot exist in the absence of Swiss cheese. However, even if we consider *holey* Swiss cheese (which by definition cannot exist

without holes), it still seems clear that the holes depend on the location of the bits of cheese and not vice versa.[1]

3 The existence and shape of a heap of sand depends on the existence and the location of each of the grains of sand that make it up. Even though it may be true that the location of some grains of sand can be causally explained by the locations of other grains, it would be very odd to think that the location of all the individual grains consists in or depends on the overall shape of the heap, instead of the other way around.
4 Consider the property of *being made of wood or aluminum*. The existence of this complex property presupposes (in some sense) the *prior* existence of the properties of *being made of wood* and of *being made of aluminum*. In addition, to have or instantiate the complex property just is to have one or the other of the simpler ones. All of the facts involving the instantiation of the complex properties can be derived from facts about the instantiation of the two simpler components.
5 We argued, in Chapter 2, that in some cases the truth of a proposition is grounded in the existence of a truthmaker. The truth of the proposition that atoms exist is explained by the existence of atoms, not vice versa.

Jaegwon Kim (1994) noted other cases involving determination without causal determination. Given the right social context and conventions, one's signing a check is a way of paying a debt that one owes another. One's signing the check does not cause the debt to be discharged, though in some sense, it *is* the discharging of the debt. And yet, there seem to be two facts here. First, there is the fact that one signed the check, and second, there is the fact that one paid the debt. One could have paid the debt in many other ways, but as things actually turned out, one paid the debt by signing the check. One's signing the check was the way in which one paid the debt. The relation between these two facts seems to be one of metaphysical grounding: the paying of the debt is grounded in the signing of the check.

### 3.1.3 Philosophical debates that turn on grounding claims

Debates about grounding are a recurring theme in the history of Western philosophy. This has been especially true in the field of moral philosophy or ethics. To take a famous example, let's consider Plato's dialogue, the *Euthyphro* (Plato 1997: 1–17). In this dialogue, Socrates and Euthyphro are seeking the definition or essence of piety. They agree that every holy (pious) act is loved by the gods, and that everything that is loved by the gods is holy. However, Socrates is not satisfied with this extensional coincidence: the fact that the same acts that can be called 'holy' are loved by the gods is not enough to make being loved by the gods the *definition* of holiness. He wouldn't be satisfied even if it turned out that the two properties are necessarily coincident (which logicians call *intensionally* equivalent). Instead, Socrates wants to know if an act is holy because it is loved by the gods, or if the gods love an act because it is holy. Plato seems to be investigating a question of grounding: is the holiness of an act grounded in the gods' love of it, or is the gods' love of it grounded in its holiness?

Much of Aristotle's metaphysical method also presupposes the existence of a grounding relation. He assumes in the *Metaphysics* (Aristotle 1984: 1552–1728) that when we have discovered the definition of a species in terms of a genus and a differentia, then the fact that something belongs to that species is grounded in its belonging to the genus and its having the differentiating feature. This is true even when the membership in the species coincides exactly and necessarily with the class of things possessing the differentia. For example, if the correct definition of the class of mammals is hairy animal, then a dog's being a mammal is grounded in its being an animal and in its being hairy, even if being hairy and being a mammal coincide in extension. It would still be the case that being mammal is partly grounded in being hairy, and not vice versa.

The concept of grounding is also needed to make sense of debates in more contemporary philosophy. For example, moral philosophers ask whether *the right* is grounded in *the good*. Are reasons grounded in other normative truths, or vice versa? Are legal facts partly grounded in moral truths? Questions of grounding arise in epistemology as well. Is the fact that a belief is justified always partly grounded in other beliefs that one has, or can it be wholly grounded in one's experience, or can it be totally ungrounded?

### 3.1.4 Fruitfulness of grounding

As Schaffer (2009) points out, the notion of grounding can be very useful for defining other metaphysical notions. For example:

*x* is fundamental = *x* is not grounded by anything
*x* is derivative = *x* is grounded by something
*x* is an integrated whole = *x* wholly grounds each of *x*'s proper parts
*x* is a mere heap = *x* is wholly grounded by its proper parts

We could also define fundamental truths or facts as those that are ungrounded, or perhaps either ungrounded or "zero-grounded," as Kit Fine (2012a: 47–48) says, meaning grounded by the null or empty set of facts).

If we take the relation of grounding between facts or truths as fundamental, we could perhaps define a relation of grounding between other entities in something like the following way:

**Def D3.1 Entitive Grounding.** Entity *x (entitively) grounds y* if and only if *y*'s existence and all of the facts intrinsic to *y* are wholly grounded in the fact of *x*'s existence and all of the facts intrinsic to *x*.

We can also use grounding (as an explanatory relation between facts) to define which properties or facts are intrinsic to an entity, as we did in Chapter 2:

**Def D2.3 Intrinsicality.** *x* is *intrinsically F* if and only if nothing that is not *x* or a part of *x* is part of the ground of *x*'s being *F*.

Grounding as a relation between facts is normally taken to be *factive*: that is, if the fact that *p* grounds the fact that *q*, then both *p* and *q* are true. On this conception, if the shape *S* of the Swiss cheese grounds its having seven holes, then the cheese really does have shape *S* and seven holes. We could also introduce a non-factive grounding relation between propositions or possible facts. Proposition *p* non-factively grounds proposition *q* if and only if it is possible that *p* (factively) grounds *q*. (Or, perhaps, it is necessarily true that if *p* and *q* are both true, then *p* factively grounds *q*. The two definitions are probably equivalent. See Section 3.7.) If we instead take non-factive grounding as primitive, we could define the factive grounding of *p* by *q* in this way: *p* and *q* (are both true), and *q* non-factively grounds *q*.

There is also a useful distinction between one fact's wholly grounding versus only partially grounding another. For example, one might think that the meaning of a word on a particular occasion is partly grounded in the intentions of the speaker, but not wholly grounded in those intentions, since the meaning also depends on current linguistic practices and conventions. Let's suppose that one fact can be wholly grounded (jointly) by a set of facts, and not just by one other fact taken alone. For example, the fact that the set $\{0,1\}$ has two members is wholly grounded in the fact that is has 0 as a member together with the fact that it has 1 as a member (and perhaps also by the fact that 0 and 1 are different). It is then easy to define partial grounding in terms of whole grounding:

**Def D3.2 Partial Grounding.** The fact that *p* is a *partial ground* for the fact that *q* if and only if there is a set $\Gamma$ of facts containing the fact that *p* such that $\Gamma$ (wholly) grounds the fact that *q*.

Fine (2012a) also defines a relation of weak grounding, a relation that is useful in defining the logic of grounding:

**Def D3.3 Weak Grounding.** That *p* is a *weak ground* for the fact that *q* if and only if either the fact that *p* grounds the fact that *q* or $p = q$.

### 3.1.5 Grounding is different from conceptual priority, reduction, and supervenience

Let's grant that we have pretty good reason to believe that there is metaphysical grounding. We still need to investigate whether this is an independent, indispensable notion, not definable in terms of other familiar relations.

**3.1T.1T Indefinable Grounding.** The grounding relation is unique and indefinable.

**3.1T.1A Definable Grounding.** The grounding relation is definable in terms of other relations.

We might think, for instance, that grounding either is or is definable in terms of conceptual analysis, reduction, or supervenience.

Consider conceptual analysis or conceptual priority first.

**3.1T.1T.1 Conceptual Analysis and Indefinable Grounding.** Grounding is not definable in terms of conceptual analysis.

What is *conceptual analysis*? What is it to *analyze* one thought or proposition *into* another? These are difficult questions to answer, because the notion of *analysis* has been used in somewhat different ways by different philosophers since at least the time of Leibniz in the seventeenth century. We nonetheless take a stab at it. First, a conceptual analysis of one proposition in terms of another preserves the semantic content or cognitive significance of the analyzed proposition (the analysandum). At the same time, the structure of the two propositions should be similar, with the exception that simple elements in the analysandum have been expanded into more complex expressions in the analysans (the product of the analysis). So, for example, one might analyze the proposition that Brown is a ophthalmologist into the slightly more complex proposition that Brown is a medical doctor specializing in the treatment of the eyes, with the underlying thought that the cognitive significance of being an opthamologist just is (for those who have that concept) exactly the same as that of being an eye doctor.

The first thing to note is that there seems to be a mismatch between the domains of grounding and conceptual analysis. Grounding, at its core, seems to relate facts or entities "in the world" (as we might say), while conceptual analysis is clearly not such a relation. Rather, conceptual analysis relates propositional contents, thought of as objects of thought and understanding. On some views, there is no real difference between facts and true propositions, but on many other views they are fundamentally different.

More importantly, there are some instances of grounding that have nothing to do with conceptual analysis. For example, suppose that it is part of the very essence of water to consist of $H_2O$ molecules (as Saul Kripke 1980 argues). Then the fact that the pitcher is full of water is wholly grounded in the fact that it is full of $H_2O$ molecules, despite the fact that the concept of *being water* cannot be analyzed as consisting of $H_2O$ molecules. Conceptual analysis is supposed to be *a priori*, something that we can verify from the philosophical armchair. We should be able to recognize the truth of the analysis simply on the basis of possessing the relevant concepts, while facts about grounding can be empirical discoveries (although Hofweber 2009 demurs).

The upshot is that grounding does not seem definable by conceptual analysis.

What about reduction? Historically, there have been two conceptions of reduction in analytic philosophy. The first was *theoretical* reduction (in the sense of Ernst Nagel 1961): the reduction of one scientific theory (like chemistry or geology) to another (like physics). The second was *ontological* reduction: reducing one object or class of objects to another, by showing that objects of the first kind are really nothing but objects of the second kind.

**3.1T.1T.2 Theoretical Reduction and Indefinable Grounding.** Grounding cannot be defined in terms of theoretical reduction.

**3.1T.1T.3 Ontological Reduction and Indefinable Grounding.** Grounding cannot be defined in terms of ontological reduction.

Grounding and theoretical reduction seem quite disparate. Grounding relates facts to other facts (or to sets of facts). Theoretical reduction relates theories to other theories, where these theories are understood to be products of human activity, in the form of sentences of a language or some other sort of representational model. Grounding has nothing to do with human theories of this kind. Still, one might think that where a theoretical reduction is possible, the reduction shows that one set of facts (the facts represented by the reduced theory) are grounded in another set of facts (the facts represented by the reducing theory).

Unfortunately for this supposition, most philosophers of science are agreed that theoretical reductions as conceived of by Ernst Nagel and other contemporaries of his are impossible. Nagel (1961: 352) hoped that we could discover a set of "bridge" principles that would enable us to deduce the sentences of the reduced theory by means of sentences in the reducing theory. But in fact the relation between two successful scientific theories (like molecular biology and quantum physics) cannot be captured by such a simple deductive process. Still, one might hope that some or all facts of molecular biology might be grounded (at least partially) in microphysical facts.

Ontological reduction is much closer to grounding. In fact, one might reasonably suppose, as Gideon Rosen (2010) does, that ontological reduction is a special case of grounding: if entity $x$ is reducible to the entities in set $S$, then intrinsic facts about $x$ are wholly grounded in those involving the members of $S$. Still, as Rosen recognizes, there might be cases in which some fact is grounded in some other facts in the absence of ontological reduction. One might think, for example, that some facts about a thing are grounded in other facts about that same thing. For instance, one might think that in each case in which one has a justified belief, the fact that the belief is justified is grounded in facts about other beliefs, or in other facts about one's memory or sense organs. This wouldn't require an ontological reduction of justification to other things.

Finally, consider the relation of supervenience, which we introduced in Chapter 2.

**3.1T.1T.4 Supervenience and Indefinable Grounding.** Grounding is not definable in terms of metaphysical supervenience.

We defined weak supervenience in Chapter 2 (D2.6). That definition gave us a relation between two sets of properties. We could modify it slightly to get a relation between two classes of possible facts:

**Def D3.4.1 Weak Supervenience (Facts).** A set of possible facts $A$ weakly supervenes on a set of possible facts $B$ iff it is impossible for two worlds to agree on which $B$ facts exist but to disagree about which $A$ facts exist.

We could talk similarly about supervenience for propositions. To bring this even closer to grounding, consider the special case in which a single proposition weakly supervenes on a set of propositions:

**Def D3.4.2 Weak Supervenience (Truth-Value of a Single Proposition).** The truth-value of a proposition $p$ *weakly supervenes* on a set of possible facts $B$ if and only if it is impossible for two worlds to agree on which $B$ facts exist but to disagree about the truth-value of $p$.

There is also a notion of strong supervenience in the literature, but the distinction between strong and weak supervenience is irrelevant for our current purposes. Supervenience expresses a kind of covariation or correlation between possible facts.

There are at least three reasons why grounding cannot be defined in terms of supervenience. First of all, grounding is an asymmetric relation, while supervenience is not (see Stoljar 2015). It is possible for two sets of facts to supervene on each other. For example, the set of facts about which numbers are less than or equal to other numbers supervenes on the set of facts about which numbers are greater than or equal to other numbers, and vice versa. It won't help to try to define grounding in terms of one-way supervenience. That is, we can't say that set *A* of facts grounds set *B* if only if *B* supervenes on *A* and not vice versa. Consider three sets *A*, *B*, and *C*, such that any one of the sets supervenes on the union of the other two, but not vice versa. This would be a case of circular partial grounding, using the proposed definition of one-way supervenience: *A* would partially ground *B*, and *B* would partially ground A.

Second, and more fundamentally, the *granularity* of the two relations are different. Grounding is more fine-grained than supervenience, able to distinguish between facts and sets of facts that supervenience must count as equivalent. Supervenience is an intensional relation. If propositions *p* and *q* are necessarily in agreement in truth-value (true or false in exactly the same worlds), supervenience must treat the corresponding possible facts as interchangeable. In contrast, grounding is hyperintensional, able to draw distinctions between metaphysically equivalent facts. So, for example, we saw that the existence of $\{2\}$ is grounded in the existence of the number 2 (and not vice versa), even though 2 and its singleton exist in exactly the same worlds.

Third, grounding implies an explanatory relation between the fundans and the fundatum, while supervenience requires no such explanatory relation between the supervening facts and their base. It is the presence of this explanatory relation that makes grounding relevant to issues of ontology. Explained facts can be treated as less fundamental, and so less subject to considerations of economy, like Ockham's Razor (**PMeth 1**).

## 3.2 Relation between Grounding and Truthmaking

If grounding is a kind of explaining how (Litland 2013), then to say that *p* grounds *q* is to say that for *p* to be the case is a way for *q* to be the case. This could be explicated in terms of truthmakers: all of the truthmakers for *p* are also truthmaker for *q*, but not vice versa. This is a much more stringent condition than the condition of *p*'s entailing or necessitating *q*. Proposition *p* could necessitate *q* without any of *p*'s truthmakers being truthmakers for *q*, for example, if truthmakers for *p* contain facts that are superfluous for *q*'s truth.

### 3.2.1 Interdefinability?

Given this connection between truthmaking and grounding, we might hope that we could simply define grounding in terms of truthmaking. But this supposition faces an immediate problem. Truthmaking was defined in terms of grounding: *x* is a truthmaker for *p* only if *p*'s truth is (wholly) grounded in the existence of *x* (see Rodriguez-Pereyra 2006: 960). To define truthmaking in terms of grounding and grounding in terms of truthmaking is viciously circular. Still, it would be interesting if the two notions were inter-definable, that is, if we could either define truthmaking in terms of grounding or grounding in terms of truthmaking. This would leave us with the difficult problem of deciding which notion is metaphysically prior to the other.

As we suggested above, it seems that one fact grounds another just in case any truthmaker for the proposition corresponding to the first fact is also a truthmaker for the proposition corresponding to the second.

**3.2T Truthmaker to Grounding Link.** The fact that $p$ (wholly) grounds the fact that $q$ if and only if: (i) $p$ & $q$, (ii) the truth of $p$ necessitates the truth of $q$, and (iii) necessarily, if $p$ & $q$, then every truthmaker for $p$ is a truthmaker for $q$.

We also want to allow for the case in which a *set* of facts (and not just an individual fact) grounds a further fact. For example, we want the facts that $p$ and that $q$ to jointly ground the fact that ($p$ & $q$).

**3.2T.1 Generalized Truthmaker to Grounding Link.** The facts corresponding to set $\Gamma$ of propositions (wholly) grounds the fact that $q$ if and only if (i) $q$ and every member of $\Gamma$ is true, (ii) the joint truth of the members of $\Gamma$ metaphysically necessitates the truth of $q$, and (iii) necessarily, if both $q$ and every member of $\Gamma$ are true, then, for every $p$ in $\Gamma$, every truthmaker for $p$ isis a part of a truthmaker for $q$.

It seems that if we were to take truthmaking as a primitive or indefinable notion, we could in this way define grounding in terms of it. Suppose, for example, that a ball's being colored is grounded in its being scarlet. If there is some truthmaker for the ball's being scarlet, it seems that this would also necessarily be a truthmaker for its being colored. Suppose, for another example, that the moral attributes of an action are wholly grounded in certain of its non-moral attributes. On that assumption, it seems plausible to suppose that the truthmakers for the predications of those non-moral attributes to the action will jointly constitute a truthmaker for the predication of the moral attribute. As a final example, suppose that the fact that ($p$ *or* $q$) is always wholly grounded in any fact corresponding to one of its disjuncts. So, if $p$ is true, then the fact that ($p$ *or* $q$) is wholly grounded in the fact that $p$. Clearly, any truthmaker for $p$ will also be a truthmaker for the disjunction, satisfying the Truthmaker to Grounding Link.

So Truthmaker to Grounding Link seems to capture at least some cases of grounding in terms of truthmakers. However, we shall see (in Section 3.4) that other cases of grounding (extra-conceptual grounding) cannot be so captured.

### 3.2.2 Can grounding theory replace truthmaker theory?

Kit Fine (2012a) has argued that we should simply replace truthmaker theory with the theory of grounding. Fine finds two problems with truthmaking. First, truthmaker theory is too restrictive in terms of what can be grounded: it only provides a metaphysical explanation of the truth of propositions. It doesn't provide an explanation for the existence of things or for any properties other than truth. Second, Classical Truthmaker

Theory (2.1T) is too restrictive in terms of what can act as a ground: all truthmaking is done only by attributing existence to some truthmaker.

As we discussed in Chapter 2, there are Non-Classical Truthmaker Theories (2.1A.1T), like Spectral Truthmaker Theory (2.1A.1T.1), that avoid Fine's second complaint. On Spectral Truthmaker Theory, truthmaking is done not by just the existence of the truthmaker but also by the truthmaker's being a certain way intrinsically. Thus, Non-Classical Truthmaker Theory and Fine's grounding theory come very close to one another.

Nonetheless, Fine's first complaint still applies. Truthmaker Theory provides explanations for the truth of certain propositions but *not* for the existence of things, nor for propositions' having properties other than truth, nor for things' (other than propositions) having any properties. Insofar as we find metaphysical explanations in such wider cases plausible, we are going to need a theory of grounding that goes beyond Truthmaker Theory. In addition, Fine points out that Truthmaker Theory is not well suited to providing certain kinds of chains of grounding chain. For example, social facts are plausibly grounded in psychological facts, psychological facts in neural facts, and neural facts in physical facts. It isn't the fact that certain psychological propositions are true that grounds the truth of social propositions. Rather, it is the psychological facts themselves that ground both the truth of the psychological propositions as well as the social facts, which in turn ground the truth of the social propositions. Truthmaker theory isn't enough to capture these relations. In order for truthmaker to form a chain, in each link the maker and the thing made would have to belong to the same category, but this never happens in truthmaker theory.

However, it is one thing to supplement Truthmaker Theory with a wider theory of metaphysical grounding; it is quite another thing to dispense with Truthmaker Theory altogether. Truthmaker Theory provides, as we saw in Chapter 2, a link between propositions and the world. Suppose, for example, that WhiteSnow is the truthmaker for the proposition that some snow is white. Is (1) some snow white because WhiteSnow exists, or (2) does WhiteSnow exist because some snow is white?

Truthmaker Theory seems committed to answer (1). So, on the one hand, it does seem that grounding is conceptually prior to truthmaking: we have used the notion of grounding in our definition of truthmakers. But, on the other hand, truthmaking seems to require some link between propositions and the world; so, if we try to simply translate truthmaker talk into grounding talk, the resulting translation will be in some tension with the thesis that grounding is always a form of explanation. Explanation is a relation between true propositions, or between the facts they correspond to, while truthmaking is a relation between propositions and entities in the world at large.

Consider, for example, a simple and relatively uncontroversial case of truthmaking: electrons are truthmakers for the proposition that electrons exist. The proposition is true *because* the electrons exist. If we try to translate this into the language of grounding (as a relation between facts), we might translate it into this claim: the fact that electrons exist explains the fact that electrons exist. But that is an obvious falsehood. Facts don't explain themselves!

Of course, we could say that the fact that electrons exist grounds the fact that *it is true that* electrons exist, and that fact grounds the fact that *it is true that it is true that* electrons exist, and so on. Perhaps that shows that grounding theory can link truth to the world in the right way after all, so long as we take grounding to be primarily a relation between facts (and not between propositions). This suggests that we take factive grounding to be more fundamental than non-factive grounding.

## 3.3 Relation between Grounding and Ontological Dependence

Fine (2012a) distinguishes between *grounding* and *ontological dependence*. Grounding is an explanatory relation between facts. Ontological dependence is a relation between entities or things: *x* is dependent on *y* iff *y* is contained in the *essence* of *x*.

### 3.3.1 Two kinds of essence

Historically, philosophers have talked about the essence of things in at least two different ways. For example, Aristotle insisted that, for each thing, there is a unique answer to the question: What sort of thing is this? This was later translated into the Latin word *quidditas*, for the what-ness of a thing. For Aristotelians, two things that belong to the same fundamental, natural kinds have the same quiddity or essence. All human beings have the common essence of humanity. All differences between individual human beings are inessential or "accidental" (*per accidens*, in Latin).

Other philosophers, like Duns Scotus or Gottfried Leibniz, have spoken in terms of the individual essences of particular things. On this conception of essence, every particular thing has its own unique essence, the what-it-is-to-be that particular thing. Scotus introduced the Latin term *haecceitas* or "thisness" for such individual essences. Socrates has one haecceity, and Plato necessarily has another, since distinct things must have distinct individual essences (see Section 9.3 for more on haecceities).

So, we might distinguish between the *quiddity* of *x* (*x*'s species, a nature or what-it-is-to-be *x* that is shared by things with the same form), and the *haecceity* of *x* (the thisness of *x*, what it is to be *x* in particular). Quiddities are shareable; haecceities are not. If so, we should distinguish between two different kinds of ontological dependence: *quidditistic* ontological dependence ('q-dependence') and *haecceitistic* ontological dependence ('h-dependence'). Socrates is q-dependent on his animality and his rationality, and on his soul and body, but not on his parents, while he might be h-dependent on his parents and on the circumstances of his conception, if we assume that these particular parents and the particular event of his conception are in some sense essential to Socrates' particular individuality or identity. In fact, many metaphysicians (following Kripke 1980) subscribe to what is called *origins essentialism*, meaning that a thing's particular origin is essential to its individual identity (i.e., part of its haecceity).

So, if we believe in origins essentialism, particular events (like conceptions) might be included in the haecceities of particular things, but not in their quiddities (although the property of having some conception-event or other might be included in the quiddity).

### 3.3.2 Circular dependency?

Fine (1994b) suggests that, while grounding is always asymmetric, two entities might be ontologically dependent on each other. Each of two identical twins might be h-dependent on the other, in the sense that the haecceity or thisness of each includes the other. If Sam and Dan are twins, it might be essential to Sam that he have Dan as a twin, and vice versa. This wouldn't require any vicious circularity, since Sam is a part of the haecceitistic essence of Dan, not of Dan himself. We don't have to suppose that a thing's essence (quiddity or haecceity) is literally a part of it (although we will consider these possibilities in Chapter 9, on particulars).

Can there be circles of q-dependency? Quite possibly. For example, it seems reasonable to suppose that the properties of being a male member of a species and being a female member of that species are interdependent: to be male is to have some sort of potential to mate with females, and vice versa. Similarly, if two species exist in an essential symbiosis with each other, then it might be that each species is part of the definition of the other.

Although there can be circles of ontological dependency, grounding is asymmetric. So, we shouldn't say that *A* ontologically depends on *B* if and only if *A*'s existence is grounded in *B*'s existence. The existence of *A* can be fundamental and so not grounded in the existence of *B*, even if the individual essence of *A* somehow includes *B*. To return to our identical twins, Sam's existence is not grounded in Dan's existence, even though Sam's identity essentially involves Dan. Similarly, q-dependency between two species does not entail that the existence of the one species is even partly grounded in the existence of the other. The two species might be causally interdependent over time, but the existence of the members of one species would not be grounded in the existence of members of the other.

Thus, there must be a real difference between ontological dependence and existential grounding, as Fine assumed.

### 3.3.3 Two kinds of fundamental entity

So, we must distinguish between two kinds of *fundamental* entities, namely, those whose existence is *ungrounded*, and those which are *not ontologically dependent* on any other entity. These are, respectively, G-fundamental and O-fundamental entities.

**Def D3.5 G-Fundamentality.** An entity *x* is *G-fundamental* if the existence of *x* is ungrounded.

**Def D3.6 O-Fundamentality.** An entity *x* is *O-fundamental* if and only if neither the quiddity (generic essence) nor the haecceity (individual essence) of *x* contains any entity other than *x* itself.

Something that isn't G-fundamental will be G-derivative, and, similarly, something that isn't O-fundamental will be O-derivative.

Can we say that if *A* ontologically depends on *B*, then *B* cannot be wholly entitively grounded in *A*? Is it impossible for an entitive fundans to be ontologically dependent on its corresponding fundata? It is natural to think, at least at first glance, that if both entitive grounding and ontological dependence exist in a particular case, the two dependency relations must run in the same direction. However, one might think that if *B* is wholly grounded in *A*, then *A*'s essence must somehow contain *B*, as a natural emanation or effluence from *A*. Otherwise, how could *A*, all by itself, be a complete explanation of *B*? If that were so, then whenever *A* grounds *B*, *A* would ontologically depend on *B*, and not vice versa. If so, then the ground or fundans would depend ontologically on the fundatum, even though the existence and character of the fundatum is wholly grounded in the existence and character of the fundans. The two dependency relations would run in opposite directions!

But, in fact, this can't be right. Our first intuition was the correct one. The two forms of dependency (grounding and ontological dependency) must run in the same direction. In many cases, it is the nature or essence of the fundatum, not that of the fundans, that explains the grounding relation. After all, *p* and *q* need "know nothing of" conjunction in order to ground (*p* & *q*), to use Fine's phrase. If this is right, we have found an important relationship between G-fundamental and O-fundamental entities:

**3.3T O-Fundamentality Entails G-Fundamentality.** Necessarily, all O-fundamental entities are also G-fundamental.

O-Fundamentality Entails G-Fundamentality itself entails that all G-derivative entities are O-derivative.

For another example of this relationship, the existence of a singleton set like {Socrates} seems to be both *grounded* in the existence of its member, Socrates, and to be *ontologically dependent* on him.

But how is this possible? How *can* the existence of the member all by itself *explain* the existence of a singleton set, if we must appeal to the essence of the fundatum in justifying the explanation? *How can* the explanation relation itself depend on the nature of the very thing to be explained? Perhaps we should say instead that singletons are only *partly* grounded in their members, with the set-hood and membership properties somehow contained implicitly in the other part of the ground. We might need to include some set-making principle in the ground of the existence of the set.

Let's consider one more time the relation between Socrates and his singleton, {Socrates}. It seems that the existence and character of {Socrates} is at least partially grounded in Socrates himself. It's also true that the individual essence or haecceity of {Socrates} somehow contains or involves Socrates himself, and so we have both kinds of dependency running in the same direction. However, Socrates seems so "know nothing" of the existence of sets. We have to combine Socrates with some kind of singleton-producing principle in order to have a complete ground of the existence and character of the set {Socrates}. That is, we seem to need something prior in existence to {Socrates} that explains the nature of that set, given the existence of Socrates.

Let's consider the relations between the counting (or natural) numbers. Suppose we take 1 to be the first number (for present purposes, it wouldn't matter if we counted 0 or 2 to be the first number). It would seem that it is of the very essence of 2 to be the successor of 1 (in this case, it is the individual essence or haecceity of 2 that seems relevant). So, 2 is ontologically dependent on 1. Similarly, it does seem that the existence of 2 is grounded in the existence of 1, and not vice versa. Of course, we might worry that the nature of 1 does not all by itself explain why 1 has a successor at all. But, perhaps we should say that 1's having the generic nature of a counting number does explain its having a successor—it may well be of the very essence of counting numbers to have successors. In any case, it seems that we have a clear case of ontological grounding here.

In contrast, consider a case of the grounding of a disjunction. Let's suppose that Austin is weird. This fact would seem to ground the disjunctive fact that either Austin is weird or $0 = 1$. In other words, the fact corresponding to (3) grounds the fact corresponding to (4):

(3) Austin is weird.
(4) Either Austin is weird or $0 = 1$.

However, there is no ontological dependency of any components of (4) on components in (3). In this case, it is the nature of the disjunctive element (the truth-function corresponding to *either…or…*) in (4) that explains the grounding relation between (3) and (4). However, this disjunctive element is not an entity "in the world": it is simply part of the *proposition* whose truth is being explained. Thus, what we are really doing is explaining the *truth* of a disjunctive proposition in terms of the *truth* of one of its disjuncts. Such grounding of truth doesn't seem to require any ontological dependence of the fundatum on the fundans. Instead, it is the logical essence of the proposition whose truth is to be explained that provides the motive force for the grounding explanation.

It seems that we can distinguish two importantly different cases of grounding: (1) those in which we explain the *truth* of one proposition in terms of the truth of another, appealing to the nature of the constituents of both propositions to explain the connection, and (2) those in which we explain facts about one kind of entity in the world in terms of one or more other entities or facts, appealing to the nature of the grounded entity in explaining its ontological dependence on some element or element in the fundans. The first is conceptual or logical grounding, while the second corresponds to what we will call *extra-conceptual* grounding. We look at this distinction in more detail in Section 3.4.

### 3.3.4 Grounding and essence

Rosen (2010: 131) has argued that all grounding facts are generalizable. That is, all grounding facts are instances of some general rule or other. We could put this point in terms of formality. Whenever one fact, say, the fact that $p$, grounds another fact, the fact that $q$, the two corresponding propositions, $p$ and $q$, instantiate general forms, $\Phi$ and $\Psi$, respectively, of such a kind that any fact that instantiates $\Phi$ grounds a fact that instantiates $\Psi$. This is Formality of Grounding:

**3.4T.1 Formality of Grounding.** Whenever $p$ grounds $q$, there exist forms $\Phi$ and $\Psi$ such that (i) for some $x$, $p = \Phi(x)$ and $q = \Psi(x)$, and (ii) for all true propositions $r$ and $s$ such that for some $y$, $r$ is $\Phi(y)$ and $s$ is $\Psi(y)$, the fact that $r$ grounds the fact that $s$.

In addition, Rosen proposes that these general rules are themselves derivable from facts about the essences of the constituents of the two facts.

**3.4T.2 Mediation of Grounding by Essence.** Every grounding fact of form (ii) (in Formality of Grounding) is itself grounded in or derivable from some *essential truths* about the constituents of $\Phi$ and $\Psi$.

Can we *define* grounding in terms of essence (Fine 2012a: 75)?

**3.6T.1 Grounding-Essence Correspondence.** The fact that $p$ grounds the fact that $q$ if and only if (i) ($p$ & $q$), and (ii) some generalization of the form *For all x, if p(x), then q(x)* follows logically from the essences of $p$ and $q$, together with the essences of things that the facts $p$ and $q$ contain.

As Fine notes, there are some problems with this suggestion. First, it won't enable us to distinguish between (i) cases where one fact is grounded by a plurality of grounds and (ii) cases where one fact is grounded by a single, conjunctive ground. For example, we might want to say that every conjunctive fact of the form ($p$ & $q$) is grounded (jointly) by the fact that $p$ and the fact that $q$. We don't want to say that the fact that ($p$ & $q$) is grounded in the fact that ($p$ & $q$), even though it is certainly true that *If* ($p$ & $q$)*, then* ($p$ & $q$) does follow from the essence of the proposition ($p$ & $q$). Still, one might get around this by adding the condition that the fundatum and fundans can't be exactly the same fact. Then, we could expand the condition to include plural grounds:

**3.5T.2 Plural Grounding-Essence Correspondence.** The facts that $p_1$, that $p_2$,..., and that $p_n$ jointly ground the fact that $q$ if and only if (i) $p_1$ & $p_2$ &... & $p_n$ & $q$; (ii) some generalization of the form *For all x, if* $p_1(x)$ & $p_2(x)$ &.... & $p_n(x)$, *then q(x)* follows logically from the essences of $p_1$, $p_2$,..., $p_n$, and $q$, together with the essences of things that the facts $p$ and $q$ contain; and (iii) if $n = 1$, i.e., there is only one grounding fact, then $p_1 \neq q$.

Second, Fine argues that Grounding-Essence Correspondence wrongly predicts that ($p$ & $p$) grounds $p$. In fact, the problem is more general than this. Grounding-Essence Correspondence predicts that any conjunction of the form ($p$ & $q$) will ground the fact that $p$, and that seems wrong. It is rather the case that the conjunctive fact is partly grounded in the facts corresponding to each of its conjuncts. Grounding-Essence Correspondence ignores facts about the relative priority or fundamentality of the two facts.

However, none of this shows that an essential link might not be a necessary condition of grounding, even if it cannot be a sufficient condition. Thus, we can replace the "if and only if" of Grounding-Essence Correspondence with a simple "only if," which brings us back to Formality of Grounding and Mediation of Grounding by Essence. It seems that we cannot define grounding in terms of essences, but all grounding relations seem to depend on facts about the essences involved.

## 3.4 Conceptual vs. Extra-Conceptual Grounding

As we saw at the end of Section 3.3.3, we can distinguish between two ways in which grounding can require a link to essences. First, there are those in which it is the truth of one proposition that grounds the truth of a second, due to the essences of the conceptual and logical elements in the two propositions. Second, there are those in which one fact grounds another fact, due to the essences of the (worldly) objects and properties that are constituents of the two facts. Of course, this distinction depends on having some fairly precise notion of both propositions and concepts:

**Def D3.7 Proposition.** A *proposition* is an object or vehicle of thought that represents things as being a certain way and is capable of being true or false.

**Def D3.8 Concept.** A *concept* is a constituent of a proposition or a capacity for thinking propositionally that plays no other role in the world other than as an object or vehicle for thought.

Philosophers have conceived of propositions in different ways. For example, Bertrand Russell supposed that propositions contain as part the very same properties and objects (like *being snow*y or Mt. Blanc) that they are about. Gottlob Frege, in contrast, believed that propositions (*Gedanke*, usually translated 'thoughts') contained only modes of presentations (*Sinne*, or 'senses') of such worldly properties and objects. So, on the Fregean view, there are far more concepts in our sense than there are on the Russellian view. Nonetheless, even on the Russellian view there are elemens of propositions (traditionally known as *syncategorematic*) which are merely conceptual, like the logical connectives corresponding to 'and' (conjunction), 'or' (disjunction), and 'not' (negation), as well as the quantifiers corresponding to 'all' or 'some'.

We could imagine a hyper-Russellian picture according to which every element of any proposition is something that also exists in the world beyond the sphere of thought, even disjunction or negation. David Armstrong, for example, believed that the world contains conjunctive facts, corresponding to the 'and' operation on propositions. Even so, Armstrong was not a hyper-Russellian, because he thought there were no disjunctive or general facts. For true hyper-Russellians, there might be little or no room for merely conceptual grounding. For the rest, however, the distinction is a significant one.

Some cases of grounding seem clearly conceptual, like the grounding of the truth of conjunctions, disjunctions, and general propositions. Other cases seem to be cases of extra-conceptual dependence, like the relations between entities and singletons.

**3.6T Conceptual-Grounding to Essence Link.** The truth of the proposition *p* grounds the truth of the proposition *q* only if *p* and *q* are both true, and *if p, then q* follows logically from the essences of the conceptual elements of the propositions *p* and *q*.

The distinction between conceptual and extra-conceptual grounding turns on a very subtle difference. Compare the following two claims, where '[*Fa*]' abbreviates the proposition *a is F* and '[*Ga*]' abbreviates *a is G*:

(5) [*Fa*]'s truth is grounded in [*Ga*]'s truth.
(6) [*Fa*]'s truth is grounded in *a*'s being *F*, and *a*'s being *F* is grounded in *a*'s being *G*.

In both cases, the truth of [*Ga*] is in some sense prior to, more fundamental than the truth of [*Fa*]. In the first case, the dependency is propositional or conceptual, in the latter case, extra-conceptual. To distinguish between the two, we have to look carefully at what licenses or justifies the explanatory connection between *a is F* and *a is G*: is it licensed by the essence of the *property* designated by '___is *F*' or the *concept* the predicate expresses? Is the essence involved in something in the mind-independent world, or is it merely in the mind? We will examine this question in much more detail in Chapters 7 and 8, where we will see that this distinction between properties and concepts is very important for most metaphysicians but is completely absent for others (in particular, for those who adopt the position of Concept Nominalism).

Is conceptual grounding the same thing as conceptual analysis? If the truth of proposition $p$ is grounded in the truth of $q$, does this mean that $q$ is the correct *analysis* of the meaning of $p$? We don't think so, since we've said nothing to suggest that conceptual grounding is always discoverable *a priori*, simply be reflecting on the meanings of our concepts "from the inside," so to speak. It could turn out that the facts about what grounds the application of a concept might be discoverable a posteriori, that is, by some scientific investigation of the human mind and its environment, while conceptual analysis is supposed to be something that can be carried out entirely in the philosophical armchair. Where conceptual analysis is possible, we would have excellent grounds for a claim about conceptual grounding, but there might be cases of conceptual grounding without conceptual analysis. Conceptual grounding could be fleshed out in terms of truthmaking rather than conceptual analysis, as in the Grounding to Truthmaker Link discussed above (Section 3.2.1).

### 3.4.1 The ontological import of grounding

What is the ontological import of grounding? Does grounding show that the fundatum doesn't "really" exist? Let's call the view that the fundatum doesn't really exist the 'eliminativist' option.

The eliminativist option certainly breaks down in the case of extra-conceptual grounding, where the grounding depends in part on the essence of the fundatum. If the fundatum doesn't really exist, it can't have an essence. This is true even if the fundans contains a complete explanation of that nature. If $x$ is grounded in $y$ in part because of $x$'s essence, then it seems that $x$'s derivative status does not count against its really existing.

Thus, once again it is crucial that we distinguish merely conceptual from extra-conceptual grounding:

(i) That $p$ is grounded in the fact that $q$ because of the essence of some worldly component of the facts corresponding to $p$ and $q$.
(ii) That $p$ is true is grounded in the fact that $q$ is true because of the essence of some conceptual constituents of the propositions $p$ and $q$.

Case (ii), conceptual grounding, would seem to have very strong deflationary consequences. We can still believe in the conceptual constituents of $p$, but we would no longer have to believe in any of the apparent worldly commitments of $p$ at all (e.g., things that terms in $p$ seem to refer to, properties that predicates of $p$ seem to express). However, case (i), extra-conceptual grounding, would seem to have much weaker deflationary implications. For example, if we believe in the conceptual grounding of the truth of ethical propositions in the truth of non-ethical or *natural* propositions, then we could simply deny the *real existence* of ethical properties and facts. In contrast, if we believe that ethical facts are non-conceptually grounding in the natural facts, in part because of the very essence of ethical properties, then we are fully committed to the real existence of ethical facts and ethical properties. Nonetheless, there would still be some gain, since we will have reduced the number of fundamental or ungrounded properties.

### 3.4.2 Support for the One Truthmaker per Fundamental Property Principle

In Chapter 2, we argued that defenders of truthmaker theory should accept the principle that two truths that involve different fundamental properties must have different truthmakers (PTruth 1). We are now in a position to define fundamental property:

**Def D3.9 Fundamental Property.** *F*-ness is *fundamental* if and only if, necessarily and for all *x*, if *x* is *F*, then that *x is F* is ungrounded (or zero-grounded).

Attributions of fundamental properties are always ungrounded. If we combine this definition with a grounding-to-truthmaking link, then we can provide a criterion for individuating fundamental properties, that is, for saying when two fundamental properties are identical or distinct.

If two propositions have of necessity the same truthmaker, then either the truth of one is grounded in the other or the truth of both is grounded in the truth of some third proposition. So, if *x is F* and *x is G* are both necessarily made true (when either is true) by the same truthmaker, then at least one of the two propositions must have its truth grounded in the truth of some other proposition. Consequently, it is impossible, given these assumptions, for *F*-ness and *G*-ness to be distinct *fundamental* properties.

**3.7T Distinct Fundamental Properties.** *F*-ness and *G*-ness are distinct fundamental properties if and only if it is necessarily the case that, for any *x*, a truthmaker for the proposition *x is F* is distinct from any truthmaker for the proposition *x is G*.

From a slightly more general version of this thesis, we can derive One Truthmaker Per Fundamental Property (PTruth 1):

**PTruth 1 One Truthmaker per Fundamental Property.** If *p* is the true predication of a fundamental property *P* to $x_1$ through $x_n$, and *q* is the true predication of a different fundamental property *Q* to the same things $x_1$ through $x_n$, then *p* and *q* have distinct truthmakers.

## 3.5 Alternatives to Grounding?

We've shown that grounding is a useful notion, distinct from reduction, supervenience, truthmaking, and conceptual analysis. In this section, we will consider three competitors to grounding theory. These competitors attempt to do the same work without treating *grounds* as an undefined primitive. The first follows the method of philosophical paraphrase defended by Willard van Orman Quine. The second relies on Kit Fine's notion of *reality*. The third seeks to take *fundamentality* as primitive, defining *grounding* in terms of *fundamentality*, instead of the other way around.

### 3.5.1 Quine's method of paraphrase

As Quine (1953/1980) pointed out, even if we have no notion of truthmaking or grounding or fundamentality, we can still ask the basic ontological question, *What is there?* As Quine also noted, there is the obvious and simple answer, *Everything*. However, we can ask what exactly this "everything" includes. Does it include minds, organisms, artifacts, or composite entities of any kind? Events or processes? Numbers, sets, or propositions? Quineans deny that there is any ontological order to be placed on what exists. If a supposed "something" is really distinct from but dependent on one or more other objects, then "it" doesn't exist at all.

A complication enters when we consider Quineans who wish to deny the existence of things whose existence seems to be taken for granted in much of our everyday speech. Peter van Inwagen (1990a), for example, denies that there are any artifacts, including tables, forks, and automobiles. But even van Inwagen, in his ordinary life, will speak in a way that seems to imply that such things exist, like 'I own an automobile' or 'Please pass a fork'. Grounding theorists can make sense of this by supposing that the truth of the propositions expressed by these ordinary sentences are grounded in facts that don't include anything corresponding to the apparent reference to artifacts—that is, they are grounded in or made true by facts that don't include any actual automobiles or forks. Quineans, in contrast, must deny that these statements are, strictly speaking, true at all. They must suppose that, when we seem to refer to such things, we are speaking in a "loose and popular way," with an insouciant disregard for the actual truth of the matter. Within the ontological seminar room, in contrast, we must be careful to speak only the strict and literal truth.

One difficulty with this paraphrase theory is that it seems to ascribe very sophisticated communicative intentions to ordinary people who have never considered sophisticated philosophical or scientific arguments purporting to show that many ordinary objects don't really exist. It is probably better for Quineans to adopt a stance of agnosticism toward the psychological question of what ordinary people really believe or really intend to convey by sentences of natural language. Perhaps ordinary people really do believe the propositions that some ontologists reject. It may be, as Theodore Sider (1999) argues, that people are relatively indifferent to the question of the real, strict truth of what they and others say. People are content if what is said is *quasi-true*, the sort of thing that would be true if the general ontological framework embedded in common sense really corresponded to the way the world is. In everyday contexts, although we speak with the intention of being understood strictly and literally, we are in fact content if we achieve the expression of things that are quasi-true, close enough to the truth for practical purposes. If so, we might be tolerant of van Inwagen's speaking, in ordinary contexts, as if artifacts existed, even though he believes that they don't, so long as van Inwagen believes that what he says would be true, if (counterfactually) artifacts did exist.

In many cases, when we are considering what truths and entities are fundamental, it won't matter whether we understand that issue in terms of grounding theory or in terms of a Quinean theory. As we shall see, grounding theorists believe that we should be economical with respect to the fundamental entities of our ontology, and Quineans believe that we should be economical with respect to the entities we accept as existing in the context of doing serious ontology. In both cases, metaphysicians will inquire about what theoretical and explanatory work can be done by the entities in question.

### 3.5.2 Fine's *really* operator

Fine (2001) argues that philosophers should recognize the usefulness of a *really* operator, a linguistic device that would enable them to say that certain ordinary objects (like forks or automobiles) exist, even though they do not *really* exist. The adverb 'really' signals that we intend to speak of entities that make up what Fine calls "the intrinsic structure of reality." Other things *exist*, in a sense, but only in a qualified and attenuated sense.

Fine did not offer the *really* operator as an alternative or competitor to the notion of grounding. He thought that we needed both notions as primitive terms of metaphysical theory. Nonetheless, we can still ask: would *really* make grounding redundant? Can we define grounding in terms of the *really* operator? Here is a possible definition:

**Grounding-Really Link (Fine).** The fact that $p$ grounds the fact that $q$ if and only if the truth of $p$ entails the truth of $q$, *Really*($p$), and not *Really*($q$).

The idea is that only real facts can ground anything, and only unreal facts can be grounded. A real fact grounds an unreal fact if and only if it metaphysically entails that unreal fact. Fine notes a number of difficulties with this proposal. First, it does not respect the idea that grounds should be as minimal as possible. For example, suppose that the fact that $p$ is grounded in the fact that ($q$ & $r$), and that $p$ is unreal, while $q$ and $r$ are real, with $r$ a real truth that is completely irrelevant to the truth of $p$. The Grounding-Really Link wrongly entails that $p$ is grounded in the conjunction ($q$ & $r$). This problem could be solved by suitable repairs to the Grounding-Really Link. We could require, for example, that $p$ is one of the logically weakest propositions satisfying the Link. This would rule out conjunctions with irrelevant conjuncts.

There is, however, a more fundamental problem with the Grounding-Really Link: it entails that no grounded fact can be a part of reality. Fine argues that we shouldn't assume this:

> We may grant that some things are more explanatorily basic than others. But why should that make them more real?… We cannot read off what is nonreal from what is nonbasic. Indeed, it is possible to imagine metaphysical scenarios in which the nonbasic, or grounded, is plausibly taken to be real. Suppose [for illustration] compound events enter into both basic and grounded causal relations (suppose their being caused is always grounded but their causing some effect is sometimes basic). Then we should count the grounded causal facts as real. *Source:* (Fine 2001: 25, 27)

Fine suggest that there is a general *presumption* in favor of the grounded not being real. However, this presumption need not hold in every case, as the Grounding-Really Link assumes.

### 3.5.3 Primitive fundamentality, naturalness, and primitive structure

Jessica Wilson (2014) argues that we can replace grounding with a primitive notion of fundamentality. Once we have fundamentality in place, we can define '*x* grounds *y*' in terms of *x*'s being fundamental, *y*'s being non-fundamental, and *x* and *y*'s standing together in one of a small number of possible relations, like part/whole, member/set, determinate/determinable, realizer/realized. Call these latter relations the 'small-g grounding relations'. Wilson proposes the following definitional link:

**Fundamentality-Grounding Link (Wilson).** The fact that *p* grounds the fact that *q* if and only if *p* is fundamental, *q* is non-fundamental, and *p* and q (or the entities involved in *p* and *q*) stand in some small-g grounding relation.

All of Wilson's small-g grounding relations seem to be relations between worldly entities and not explanatory relations among facts or true propositions, except perhaps for the realization relation (one fact might *realize* another). It is not obvious that every explanatory relation involving essences can be fit into one of the categories that Wilson provides.

Wilson's proposal also raises the question of why just these relations are cases of *grounding*. What is it that they have in common? The most plausible answer to this question is that, at least in certain cases, these small-g grounding relations are instances of grounding (simpliciter).

In addition, it is doubtful that *x* grounds *y* can really be defined as (i) *x* is fundamental, (ii) *y* is not fundamental, and (iii) *x* and *y* stand in one of a list of relations. For example, we can imagine cases in which *x* is fundamental, *y* is non-fundamental, *x* is a part of *y*, and yet *y* is not even partly grounded in *x*. Imagine, for example, a theory in which non-fundamental entities are grounded in the fundamental wholes which contain them, and not in their fundamental parts. In a view like van Inwagen's (1990a), the only fundamental things are simple particles and living organisms. Biological facts concerning an organic, functional part of an organism, like an eye, are grounded in the structure of the whole organism of which it is a part, and not in facts about the subatomic parts of which the eye is composed. Wilson's linkage is unable to make the relevant distinction between those facts that are grounded in facts about a non-fundamental thing's parts, and those that are grounded in facts about the fundamental whole that includes it. In addition, there is the worry that we might discover new ways for facts to be grounded in other facts, ways that do not correspond to any of the relations on Wilson's official list of small-g grounding relations.

Finally, as Litland has suggested,[2] we can have grounding without a difference in fundamentality. Suppose the world is homogeneous and infinitely divisible. It could still be the case that wholes are wholly grounded in their parts, even though there is no fundamental level (or all levels are equally fundamental).

A number of other philosophers have developed accounts of a kind of primitive or indefinable fundamentality. For example, David K. Lewis introduced the idea of *perfectly natural* properties (Lewis 1986a: 60–69). A perfectly natural property is one that corresponds with a maximal degree of intrinsic similarity. In addition, a property has to be perfectly natural in order to figure in our most fundamental laws of nature. A natural property is to be contrasted with an *unnatural* property, one that has been constructed in a highly artificial, gerrymandered way. For example, *being spherical* or *having negative charge* are fairly natural properties, while *being two-feet from an avocado* or *being green or oblong or well-liked* are quite unnatural. Lewis's notion could be extended to define notions of fundamental entity and fundamental fact. A fundamental entity might be defined as something whose essence or defining characteristic is a natural property, and a fundamental fact could be one that consists in attributing a perfectly natural property to one or more fundamental things.

Lewis's theory could be seen as a version of grounding theory, rather than as a competitor to it, especially since Lewis suggested that the naturalness is degreed, that some properties are more natural than others. We could combine Lewis's notion of degree of naturalness with Wilson's list of small-g grounding relations, and define grounding as follows:

**Grounding-Naturalness Link (Lewis).** The fact that $p$ grounds the fact that $q$ if and only if $p$ is more natural than $q$, and $p$ and $q$ (or the entities involved in $p$ and $q$) stand in some small-g grounding relation.

This will have some of the defects of Wilson's Grounding-Fundamentality Link, but it will be able to handle cases in which there is no fundamental level (no perfectly natural properties), and it can make sense of grounding relations between non-fundamental facts and entities.

Sider (2011) recently defended another notion of primitive fundamentality, which he labels 'structural'. For Sider, the term 'structural' can be applied to entities of many different categories. There are structural properties, entities, and propositions. Even logical elements, like 'and', 'some', 'not', or 'all' can count as structural. In fact, Sider allows 'structure' to be applied as a linguistic operator to any meaningful component of an assertion.

In place of grounding, Sider recommends that we proceed by means of metaphysical semantics (Sider 2011: 118–125). Sider's notion of metaphysical semantics corresponds very closely to what we have called conceptual grounding. Sider in effect denies that there is any such thing as ontological or non-conceptual grounding. Therefore, Sider's approach to metaphysics is largely Quinean in spirit. We should acknowledge as really existing only those entities and kinds that are part of the fundamental structure of the world. What is most innovative about Sider's approach is his inclusion of logical elements in the world's fundamental structure.

## 3.6 Can Grounding Relations be Grounded?

Are grounding facts themselves grounded? There seem to be four options:

1 Grounding facts are fundamental: simply and absolutely ungrounded.
2 Grounding facts are *zero-grounded*, grounded but not grounded by anything.
3 The *straightforward account* (using Litland's label): whenever the fact that $p$ grounds the fact that $q$, the fact that $p$ also grounds the fact that the fact that $p$ grounds the fact that $q$.

4 All grounding facts are grounded in facts about essences (both quidditive and haecceitistic), and at least some of these essential facts are absolutely ungrounded.

There is an obvious problem with the first option. It entails that derivative (non-fundamental) entities are constituents of certain fundamental facts, namely, those facts concerning the grounding of those derivative entities in fundamental ones. It seems natural to assume that fundamental facts should be *pure*, to use Sider's (2011) term. That is, fundamental facts should involve only fundamental entities and fundamental properties.

Fine (2012b) and Litland (forthcoming) defend option 2, zero-grounding, as a solution to this problem. They rely on an analogy between theorems of pure logic and zero-grounded facts. We don't have to accept each theorem of logic as absolutely ungrounded. Instead, we accept that each theorem is true on the basis of a valid proof. However, the proof of a theorem of pure logic (unlike theorems of a theory like arithmetic or geometry) doesn't depend on beginning with any premises. A theorem of logic can be proved from the empty set of premises. In the same way, a zero-grounded fact has a metaphysical explanation, but the explanation does not involve appealing to any other, more fundamental facts. The fact that one fact is grounded in another can be demonstrated simply by showing how the first can be metaphysically explained in terms of the other.

There are a couple of problems with option 3, the straightforward account. First, it does not seem plausible that in *all* cases in which the fact that $p$ grounds the fact that $q$, it is also the case that the fact that $p$ grounds the grounding fact itself. Suppose that the existence of a shadow is grounded in a certain pattern of illumination on the sidewalk. It doesn't seem right that a mere pattern of light and dark on the sidewalk would have the metaphysical resources to explain how it is that it is able to explain the existence of a shadow. We need to know something about the very essence of light and darkness under these conditions in order to be in a position to know what can be metaphysically explained by the pattern, and those essential facts are not contained by the fact of the actual pattern of illumination and darkness.

Second, as Litland (forthcoming) points out, we need to have an account, not only of the status of positive grounding facts, but also of the status of negative grounding facts. That is, we must account for the falsity of positive grounding claims, such as the fact that $p$ does not ground $q$. We certainly can't claim that whenever $p$ does not ground something, it is $p$ that grounds the absence of a grounding relation. The non-grounding fact $p$ could be almost any fact, such as the fact that $2 + 2 = 4$. That fact doesn't ground the fact that snow is white, but the fact that $2 + 2 = 4$, all by itself, doesn't explain the fact that $2 + 2 = 4$ doesn't explain the fact that snow is white.

Option 4 seems the most viable alternative to the Fine-Litland zero-grounding account. On this view, if $p$ grounds $q$, then the fact that $p$ grounds $q$ is grounded in certain facts about the essences of the components of the facts $p$ and $q$. In some special cases, the straightforward account is true. Suppose that $p$ grounds $q$, and suppose that $r$ contains all the facts about the essences of the constituents of $p$ and $q$ that are needed to ground the grounding relation between $p$ and $q$. In this sort of case, it is plausible to suppose that $r$ also grounds the fact that $r$ grounds the grounding relation between $p$ and $q$, since $r$ contains all of the information about essences needed to explain the relation between $p$ and $q$. Consequently, it would seem also to contain all the information needed to explain the relation between itself and the grounding claim. Similarly, if $r$ is the ground of the absence of a grounding relation between *p* and *q*, it is plausible to suppose that *r* also grounds the fact that *r* grounds the absence of this grounding relation, since *r* would contain all the necessary information about the relevant essences (see Rosen 2010, Dasgupta 2014).

Paul Audi (2012) objects to the claim that the straightforward account could be right about *any* grounding relation, since no fact can ground its own ability to ground further facts. But why think this? Perhaps the thought is that we would need to add some information about the essence of the grounding relation itself. However, we might reasonably doubt whether there is such an essence. Or, we might suppose that every essence contains implicitly the relevant information about the essence of the grounding relation, since it is the very essence of an essence (we might say) to ground certain grounding relations.

## 3.7 Connections between Grounding and Entailment

Does grounding entail metaphysical entailment? That is, if *p* grounds *q*, does *p* also metaphysically entail *q*? Does it follow that necessarily if *p* then *q*? Let's call this the thesis of Grounding-Entailment Entailment, or GEE.

**3.8T Grounding-Entailment Entailment (GEE).** Necessarily, if the fact that *p* grounds the fact that *q*, then *p* metaphysically entails *q* (i.e., necessarily if *p* is true, then *q* is true).

Kelly Trogdon (2013) argues in favor of GEE. If *p* grounds *q*, then there should be no unbridgeable gap between *p* and *q*, since otherwise *p* offers no metaphysical explanation of *q*. An explanatory connection between the two requires that the grounding relation itself be grounded in essential properties (in Fine's sense) of *p*, *q*, or the entities they involve. But whatever connection is grounded in such essential truths must itself be a necessary truth. The last step seems to be a non sequitur, or perhaps a begging of the question. Essential truths are necessary, but how do we know that all truths grounded by essential truths are themselves necessary, without assuming the Grounding-Entailment Entailment thesis?

In addition, there are some possible counter-examples to GEE. Jonathan Dancy (2004) argues that *q* can be grounded in *p*, even if *p* does not entail *q*, since there could be other conditions that *enable* the grounding to take place. For example, my obligation to do *A* might be wholly grounded in the fact that I've promised to do *A*, even though there are other conditions (such as my being able to do *A*) that must be present in order for this grounding to take place. Dancy's objection is especially apt, since it might be that the necessity for these conditions is itself grounded in the essences of *p* and *q*. There might also be grounding *disablers*. Grounding might sometimes be a defeasible relation, one that holds in normal conditions but that can be defeated by abnormalities.

The possibility of grounding enablers and disablers could also be relevant to questions about truthmakers for universal truths. Perhaps the fact that everything is *F* is grounded in the conjunction of facts of the form *x is F*, for each actually existent *x*, with the relevant total totality fact (the fact that the existent things are *all* the existent things) as an enabler of this grounding relation, and not part of the ground.

If GEE is false grounding can still count as a kind of explanation. It is not obvious that an explanans must always *entail* its corresponding explanandum. It is enough if the explanans is sufficient in the actual circumstances (given the presence and absence of enablers and disablers) for the explanandum.

## 3.8 How is Grounding Different from Causal Explanation?

Jonathan Schaffer (2016) has argued that, structurally and formally, causation and grounding are very similar. Both support non-accidental generalizations, both delimit a specific kind of necessity (causal or natural in the one case, metaphysical in the other), and both can back explanations. However, Schaffer also notes some differences. First, a cause is *distinct* or *separate* from its effects in a strong sense. In fact, we can use grounding to define the relevant sort of separateness:

**Def D3.10 Separate Existences.** *x* and *y* are *separate existences* if and only if *x* is not identical *y*, *x* does not ground *y*, *y* does not ground *x*, and nothing grounds both *x* and *y*.

When one fact grounds another, the two facts are definitely not separate existences in this sense. This brings out one very important difference between the two.

Second, assuming GEE, grounding entails metaphysical supervenience, while causation does not. Finally, Schaffer (2014) argues that there can be indeterministic causation, but not indeterministic grounding. In fact, causation can be probabilistic in nature, as when some cause makes a certain effect more likely than it would otherwise be. Nothing like that seems to happen in the case of grounding.

## 3.9 Conclusion: Grounding and Ontological Economy

When applying Ockham's Razor, do we look at all entities posited by the theory or just the fundamental ones (i.e., those that are either ungrounded or zero-grounded)? Here was our initial statement of Ockham's Razor:

**PMeth 1 Ockham's Razor.** Other things being equal, adopt the simplest theory.

What does it mean for a theory to be simpler? One widely accepted criterion is that the simpler theory posits fewer entities (quantative economy) and fewer types of entities (qualitative economy). In seeking such simplicity, do we minimize only fundamental things and types of things, or all things and types?

It might seem that a theory's having a large number of both derived (non-fundamental) entities and derived types or properties is a virtue rather than a vice. Schaffer argues that the correct version of Ockham's Razor demands that we maximize the ontological "bang for the buck," that is, that we achieve an optimal balance of minimizing fundamental entities while maximizing derivative entities (especially useful ones) (see Schaffer 2014: 9).

There's some reason for thinking, however, that we should minimize the number and variety of derived entities, as well. The so-called special sciences, like chemistry, geology, and metereology, deal almost exclusively with derived entities, like mountains and hurricanes. Nonetheless, such social scientists still make use of Ockham's Razor, minimizing the postulation of theoretical entities, even though those entities are not fundamental ones. Schaffer argues that special scientists apply the Razor only when the new entities threaten to require new fundamental forces, as in cases of telekinesis (Schaffer 2014: 18).

However, ontological economy (the reduction of the number of things and types) isn't the only domain of application of Ockham's Razor. We don't just try to minimize the number of theoretical entities; we also prefer relatively simple *explanations* of derived facts, even when the more complicated explanations add no new fundamental entities. This demand for simpler explanations could (in some cases) bring with it a demand for fewer derived entities. Moreover, if grounding always involves an appeal to the natures or essences of the fundata, then it seems that a rich set of fundata requires a comparably rich set of fundamental facts about derived-entity essences. This should also count as a cost of the theory.

What does seem clear is that the minimizing of fundamental entities should always take *priority* over the minimizing of derived entities.

What if there is no fundamental level? What if there are no ungrounded or zero-grounded facts? On this supposition, all facts, and therefore all entities, would be derived. Schaffer argues that we should deny that such a scenario is really possible (Schaffer 2014: 18). Nothing could exist unless there is a fundamental explanation of its being in metaphysical terms. If, however, we do take seriously the possibility of a world without a metaphysically fundamental level, we could still use the relation of grounding to define a kind of relative fundamentality (as Schaffer 2014 demonstrates). Then we could simply give priority to the minimizing of more fundamental facts over less fundamental ones.

## Notes

1 In this example and in the following one, the grounded feature is a holistic one, while the grounding facts pertain to the locations and relations of the microscopic parts. As we shall see in Chapters 18 and 22, it is not in fact so obvious that grounding should always run in this "bottom-up" direction.

2 In a seminar on grounding at the University of Texas at Austin, fall of 2015.

# Part II

# Dispositions

# 4

# Conditionals

In Part II (consisting of Chapters 4–6), we are concerned with the dispositions and powers of things. From the time of the ancient Greeks, especially from the time of Aristotle onward, philosophers have recognized that the properties or features of things seem to come in two kinds: the way things are in fact (actuality), and the way things could be but aren't (potentiality). Some things are actually hot or actually red, while some things aren't hot but are potentially hot or aren't red but are potentially red.

However, this distinction between actuality and potentiality does not really correspond to a dichotomy between actual properties and potential properties, since even potentialities are properties that things have actually or in fact. I may only be only potentially a speaker of Basque, but I have the potentiality in fact. The potentiality itself is actual, even if the property that is had potentially (the speaking of Basque) is only potential. There is, though, a dichotomy between two kinds of properties: those properties that attribute some kind of *mere potentiality* to things, and those that don't. At least, at first glance, a property like *being red* describes how things are in fact, without saying anything about how things might be potentially. Other properties, like *being teachable* or *being fragile*, do involve the attribution of potentialities to things. Metaphysicians use the term 'categorical' for those verbs, adjectives, and predicates that designate properties that are only about the actual world and 'dispositional' for those terms that are also about what could potentially be the case. We will say that categorical properties are designated by categorical terms, while dispositional properties are designated by dispositional terms.

Dispositional properties typically involve potential facts by way of some condition. For example, a fragile object is one that is potentially broken, in the sense that it *would* break *if* struck with sufficient force. Consequently, one popular approach to the metaphysics of dispositional properties (adopted, for example, by Gilbert Ryle in his influential book, *The Concept of Mind* (1949)) takes them to involve ascribing a conditional property, a property corresponding to a conditional statement (one involving 'if' or a synonym of

*The Atlas of Reality: A Comprehensive Guide to Metaphysics*, First Edition.
Robert C. Koons and Timothy H. Pickavance.

'if'). Ryle's conditional account of dispositions raises a further question: what is it in the world that makes a conditional statement true? Answers to this question fall into four categories: (i) conditional statements are fundamental truths, of which no further account can be given (Hypotheticalism), (ii) conditional statements are made true by the world's laws of nature (Nomism), (iii) conditional statements are made true by the actual distribution of categorical properties (Neo-Humeism), and (iv) conditional statements are made true by the powers of things (Powerism).

We will begin by looking at some recent work on the semantics and logic of conditionals (Section 4.1), followed by a consideration of Hypotheticalism (Section 4.2 through Section 4.4), Nomism (Section 5.1), Neo-Humeism (Secion 5.2), and Powerism (Chapter 6).

## 4.1 Counterfactual Conditionals: Semantics, Logic, and Metaphysics

A counterfactual conditional is a conditional expressed (in English) in the subjunctive, as opposed to the indicative mood. The difference in moods corresponds to a difference in truth-conditions. Consider the following pair of propositions, used illustratively by David K. Lewis in his seminal work, *Counterfactuals* (Lewis 1973b:. 3):

(1) If Oswald did not kill Kennedy, then someone else did.
(2) If Oswald had not killed Kennedy, then someone else would have.

(1) is in the indicative mood, while (2) is subjunctive. Let's suppose that we all believed that the Warren Commission was right, and Oswald acted alone. In that case, we would certainly count (1) as true (since our confidence that Kennedy was killed by someone is independent of our belief that Oswald was in fact the killer), and we would count (2) as false (since, believing Oswald to have been acting alone, we suppose that had he not carried his plans, Kennedy would not have been assassinated at all). The obvious falsity of (2) also tells us that we cannot interpret the conditional in (2) as the *material conditional* of standard propositional logic.

In standard logic, the statement 'if *p*, then *q*' is true just in case *p* is false or *q* is true. This rule corresponds to reading the 'if/then' statement as a material conditional, in the sense that the truth-value of the conditional depends only on the facts about the actual truth or falsity of its parts (its logical 'matter'). Since the antecedent of both (1) and (2) is false (Oswald did in fact kill Kennedy), the corresponding material conditional must be true. Since (2) is false, a counterfactual or subjunctive conditional cannot be a material conditional.[1]

Counterfactual conditionals must also be distinguished from what C.I. Lewis called the "strict conditional" (Lewis 1932). A *strict conditional* 'if *p*, then *q*' is true if and only if it is impossible for *p* to be true and *q* to be false. (3) would be an example of a strict conditional:

(3) If this box were red and square, then it would be red.

If the box is red and square, it is impossible for it not to be red. However, there are examples of true counterfactuals that would be false if they were read as strict conditionals. Consider (4):

(4) If McCain had won the vote in Michigan and Florida, he would have been elected President in 2008.

(4) seems true, given the closeness of the electoral vote in 2008. However, it would not be true, if read as a strict conditional. It is surely possible for McCain to win those two states without winning the Presidency. This could have happened were he to have won Michigan and Florida but lost states which he did in fact win. Hence, the counterfactual conditional is not a strict conditional.

Quine classified counterfactual conditionals as "creatures of darkness," precisely because they did not correspond to any of the formulas of classical extensional logic (Quine 1960). The truth-value of a counterfactual conditional is not a function of the truth-value of its parts. The parts of (2) are both false, but there are many false counterfactuals that also have false antecedents ('if'-clauses) and consequents ('then'-clauses):

(5) If Oswald had not killed Kennedy, the earth would have stopped revolving on its axis.

Work by Robert Stalnaker (Stalnaker 1968) and David Lewis (Lewis 1973b), however, rescued counterfactuals from Quine's ban. The rehabilitation of counterfactual conditionals led to a revival of the conditional theory of dispositions, which had been proposed by Gilbert Ryle (1949) but had been largely abandoned as a result of Quinean anxieties about the meaning of the relevant conditional. The conditional theory of dispositions suggests that having a disposition like fragility consists in making true a conditional. Thus, the truth of (6) was supposed to be grounded in (7):

(6) The vase is fragile.
(7) If struck, the vase would break. (The word 'fragile' is derived from a Latin word meaning 'breakable'.)

Ryle applied the conditional theory very widely, since he believed that nearly all mental states, such as beliefs, desires, and even experiences, were merely behavioral dispositions of one kind or another. Whether there is hope for this sort of analysis, then, is a very important question for a number of philosophical areas.

First, we must ask, is there any reason to believe that (7) is the fundamental truth and (6) derived, rather than the other way around? Can there in fact be fundamental conditional truths? There seem to be two options here.

**4.1T Hypotheticalism.** There are fundamentally conditional or hypothetical truths.

**4.1A Anti-Hypotheticalism.** All conditional truths are grounded in non-conditional facts.

The sort of grounding that is relevant here seems to be what we called conceptual grounding in Section 3.4: can reflecting on the essence of the logical *if-then* construction reveal that conditional truths are grounded in non-hypothetical facts and entities? Or must there be some entities or primitive facts in the world that correspond directly to that *if-then* structure?

But before we examine directly the question whether Hypotheticalism or Anti-Hypotheticalism (also called 'categoricalism') is correct, it will be worth making a foray into how we evaluate counterfactual conditionals.

So: how do we evaluate counterfactual conditionals? By 'evaluate' we don't mean: how do we know or judge them to be true? What we mean is: what are the conditions under which such counterfactual conditionals are in fact true? In the mid-twentieth century, Nelson Goodman (1954) and J.L. Mackie (1973) both proposed argument models for the truth or falsity of counterfactual conditionals. A conditional of the form 'if *p*, then *q*' is true if *q* can be validly deduced from:

(a) *p* itself,
(b) the laws of nature, and
(c) other true propositions that are "co-tenable" (in Goodman's terms) with *p*.

If *q* cannot be deduced from this set of propositions, then the conditional is false. Both (b) and (c) are important: it is very rare to find a conditional in which the consequent can be deduced from the antecedent alone. Statement (3), noted above, would be one of these rare examples.

Laws of nature are relevant. Such laws are said to support corresponding counterfactuals. Consider (8):

(8) If this sample of water were cooled to 0° C, then it would freeze.

In (8), the consequent ('it would freeze') is a consequence of the antecedent plus certain laws of chemistry, together with facts about the chemical composition of water. This seems sufficient to make (8) true. In contrast, mere accidental regularities do not support counterfactuals in this way. Suppose it is true that every coin in my pocket is made of bronze. (9) would certainly not follow:

(9) If this dime were in my pocket, it would be made of bronze.

Finally, we must also consider so-called 'co-tenable' propositions. These are background facts that we believe would not have been disturbed had the antecedent been true. Recall (4):

(4) If McCain had won the vote in Michigan and Florida, he would have been elected President in 2008.

Again, (4) seems true, given the closeness of the electoral vote in 2008. However, we can deduce the consequent from the antecedent only with the additional supposition that, in the imagined counterfactual scenario, the constitutional rules for electing the President would not have been different, the other forty-eight states voted as they actually had, and so on. As Goodman noted, it is quite difficult to give a non-circular account of which true propositions are co-tenable with an actually false antecedent. Intuitively, we want to include just those propositions that still *would have been* true had the antecedent been true. If we do that, we have slipped into the subjunctive mood ('would have been') in defining the co-tenable proposition. Clearly, this is not going to give us a non-circular account of the truth-conditions of the subjunctive conditionals.

If we want to give an informative account of what it is for these subjunctive conditionals to be true, we must find something more to say about these co-tenable truths. Even if we think (as Hypotheticalists do) that the truth of subjunctive conditionals is irreducible to other truths, it would be nice to have some more information about the relationship between the truth-conditions of one conditional and those of another or between subjunctive conditionals and non-conditional statements about modality (possibility and necessity). For example, (11) seems to follow from (10), and (13) follows from the conjunction of (10) and (12):

(10) If Michigan had voted for McCain, then McCain would have won and Obama would have lost.
(11) If Michigan had voted for McCain, then McCain would have won.
(12) It was possible that Michigan vote for McCain.
(13) It was possible that McCain win.

Ideally, we would like some sort of theory, even if a relatively abstract one, that could explain why these propositions stand in these logical relations to one another.

In 1968, Robert Stalnaker introduced a new approach, which built on the notion of a possible world. Gottfried Leibniz, the seventeenth-century German rationalist, made use of the idea of possible worlds in trying to clarify issues in the metaphysics of possibility, necessity, and actuality, which philosophers refer to as 'modalities'. A possible world is a total or maximal possibility: a way that the whole of reality could have been. We will take up modality in more detail in Chapters 14–16, but for the purposes at hand we will need to say something about *possible worlds semantics*. (A *semantical* theory is a theory about linguistic meaning, including the truth of sentences.) Possible worlds semantics grew out of the work of C.I. Lewis (1932), Rudolf Carnap (1947), and Saul Kripke (1963) as an attempt to exploit the Leibnizian idea of a possible world in giving an explicit, regimented representation of our modal talk, our talk about possibility, necessity, and actuality.

Possible worlds semantics involves specifying a mathematical model with three characteristic features. First, the model must contain a domain of entities that play the role of possible worlds. Second, every *atomic* sentence—every sentence that ascribes a property to an object or asserts that a relation holds between or among some objects—is assigned a truth-value at each of these worlds. In other words, every sentence like 'Elsie is a dachshund', 'The cat is on the mat', and 'The boys formed a circle' are variously true and false according to these worlds. Finally, one of the worlds is taken to represent the actual world. The atomic sentences true according to our world are just the atomic sentences that really are true. Given a model of this sort, we can specify rules that will tell us which modal claims are true, given the facts about which sentences are true according to which worlds. The rules go like this. A statement of the form 'possibly *p*' is true in the actual world just in case $p$ is true according to at least one world. A statement of the form 'necessarily $p$' is true just in case $p$ is true according to every world.[2]

Stalnaker applied these models to elucidating the semantics of counterfactual conditionals. He assumed that we can employ a world-selection function, a function which selects a world on the counterfactual supposition that $p$. When we apply this function to a world (representing the actual world) and a set of worlds (representing some proposition $p$, the antecedent of the conditional), the world-selection function gives us a unique world—the world that would have been actual had $p$ been the case. We then check to see if the consequent of the conditional is true in the selected world: if it is, the conditional is true; if not, it is false. Here are the truth-conditions for the counterfactual conditional in Stalnaker's semantics (we'll use Lewis's symbol '$\Box\!\!\rightarrow$' to represent the subjunctive conditional connective):

**Stalnaker.** '$(p \Box\!\!\rightarrow q)$' is true if and only if $q$ is true according to $^*(p)$.[3]

The right-hand side of Stalnaker's definition includes the world-selection function *. This *-function takes a proposition as an input and yields a world as its output. Intuitively, $^*(p)$ is the world that would have been actual had proposition $p$ been true. The world-selection function * always picks a world in which the actual laws of nature hold, but it need not maintain the truth of merely accidentally true generalizations. In this way, laws support counterfactuals in a way that accidental regularities (like the coins in my pocket's being bronze) do not. The world-selection function * also changes background facts as little as possible. We make the smallest change necessary to make $p$ true. In this way, the semantics preserves all of the co-tenable truths.

Let's run through an example to clarify Stalnaker's machinery. Recall (4) once again:

> (4) If McCain had won the vote in Michigan and Florida, he would have been elected President in 2008.

In this case, $p$ is 'McCain won the vote in Michigan and Florida' and $q$ is 'McCain is elected President in 2008'. According to **Stalnaker**, for (4) to be true, the proposition that McCain is elected President in 2008 must be true according to *(McCain won the vote in Michigan and Florida), which is, intuitively, just the world most like ours but according to which McCain won the Michigan and Florida votes. Given that the *-function changes background facts as little as possible in order to make its input proposition true at the output world, *(McCain won the vote in Michigan and Florida) will be a world according to which McCain won all the states he actually did win (in our world). But given that the proposition that McCain won the vote in Michigan and Florida must be true according to *(McCain won the vote in Michigan and Florida), that world will be such that McCain got all of the electoral votes he actually got (in our world), plus the electoral votes of Michigan and Florida. As it happens, this means that according to *(McCain won the vote in Michigan and Florida), McCain won the 2008 Presidential election. Which is just to say that our $q$ is true at our $^*(p)$. (4), according to **Stalnaker**, is true. This matches our intuition about (4).

David Lewis (1973b) proposed some modifications and extensions of Stalnaker's semantics. Lewis argued that we could envisage a system of concentric spheres around the actual world, representing the various degrees of similarity these worlds bear to ours. The worlds in the closest sphere to the actual world are more similar to the actual world than any other worlds, and they are each as similar to the actual world as any other world within that sphere. The worlds in the next closest sphere (and that are not in the closer sphere) are less similar to the actual world than every world in the closer sphere, but they are each as similar to the actual world as any other world within that sphere (and not within the closer sphere). And so on, out to very distant worlds that are very much unlike the actual world.

A counterfactual conditional ($p \mathrel{\Box\!\!\rightarrow} q$) is true for Lewis just in case there is a $p$-permitting sphere of worlds such that every $p$-verifying world within that sphere is also a $q$-verifying world. In other words, to evaluate the truth of ($p \mathrel{\Box\!\!\rightarrow} q$), Lewis says we must first go out to the closest spheres in which there is at least one world according to which $p$ is true. Then one checks whether every one of the worlds in that sphere according to which $p$ is true is also a world according to which $q$ is true. If every $p$-world in that sphere is also a $q$-world, then the counterfactual ($p \mathrel{\Box\!\!\rightarrow} q$) is true. If in all of the closest spheres, there is a $p$-world that is not a $q$-world, then ($p \mathrel{\Box\!\!\rightarrow} q$) is false.[4] We can summarize Lewis's view like this:

**Lewis.** '($p \mathrel{\Box\!\!\rightarrow} q$)' is true if and only if the closest sphere containing a $p$-world such that every $p$-world in that sphere is also a $q$-world.

There is one significant difference between Stalnaker's logic and Lewis's. Stalnaker assumes in effect that for every world $w$ and proposition $p$ there is a unique $p$-world $w'$ that is closest, most similar, to the actual world. Lewis instead assumes that there is a non-empty set of closest worlds. This semantic difference results in a logical difference: Stalnaker accepts, and Lewis rejects, the Law of Conditional Excluded Middle, ($p \mathrel{\Box\!\!\rightarrow} q$) $\vee$ ($p \mathrel{\Box\!\!\rightarrow} \sim q$). (14) is an instance of this law:

(14) Either: (a) if this coin had been flipped a minute ago, it would have landed heads, or (b) if this coin had been flipped a minute ago, it would not have landed heads.

(14) must be true on Stalnaker's semantics, since we can find the *one world* closest to the actual world in which the coin is flipped. In that world, the coin must either have landed heads or not. On Lewis's approach, there is typically a *set* of closest worlds, any one of which *might* equally have been actual had the coin been flipped. The coin might have landed heads in some of those worlds but not in others. For Lewis, it is (15) rather than (14) that is logically valid:

(15) Either: (a) if this coin had been flipped a minute ago, it *would* have landed heads, or (b) if this coin had been flipped a minute ago, it *might not* have landed heads.

Both agree, of course, that if the coin had been flipped, it would have landed either hand or tails. That disjunction (heads or tails) is true in all of the closest worlds. The difference is that Stalnaker is committed to the view that, even when the process is indeterministic, so the laws of nature and the intrinsic natures of the coin and its environment do not determine whether coin would lead heads or tails, it is always true that there is some one result (either heads or tails) which is the result that would happen if the coin were flipped.

One difficulty with Stalnaker-Lewis semantics: it doesn't handle conditionals with impossible antecedents well. Suppose the number 3 exists necessarily (a plausible assumption!). Consider the following:

(16) If there were no number 3, the number 2 would still have the same successor it does have.
(17) If there were no number 3, the number 2 would have no successor or a different successor from the one it does have.

If there is no possible world containing no number 3, then both (16) and (17) count as true (vacuously true) according to Lewis-Stalnaker semantics.[5]

This is a problematic result: we seem to be able to evaluate some *per impossibile* conditionals as true (like (17)) and others as false (like (16)).

## 4.2 Hypotheticalism

So much for the logic and semantics of the counterfactual conditional (see Lewis 1973b for the details). Let's get back to the metaphysical question: are counterfactual truths fundamental, or are they grounded or made true by other truths? There are, to repeat, the two obvious options: Hypotheticalism (some conditional truths are fundamental) and Anti-Hypotheticalism (no conditional truth is fundamental).

**4.1T Hypotheticalism.** There are fundamentally conditional or hypothetical truths.

**4.1A Anti-Hypotheticalism.** All conditional truths are grounded in non-conditional facts.

What would the world be like if Hypotheticalism were true? One possibility would be for there to be simple conditional truthmakers: basic, irreducible facts to the effect that if $p$ (subjunctive mood), then $q$ (subjunctive mood). Alternatively, there could be basic facts of comparative closeness between worlds, facts of something like this form: world $w'$ is one of the closest $p$-worlds to $w''$. These facts of closeness would be very much like simple conditional facts, since they tell us which worlds might have been actual (from the viewpoint of $w'$) had $p$ been the case. These facts about closeness between worlds could exist, if possible worlds were themselves fundamental entities, as David Lewis believed them to be (a position we will discuss more thoroughly in Chapter 14). Or, to use Stalnaker's semantics, there might be fundamental truths about which one world would have been actual, had some member of a set of worlds been actual.

Hypotheticalism has had its defenders, both historically (Luis de Molina, 1535–1600) and among contemporary metaphysicians (Alvin Plantinga). We don't know of any knockdown objections to it. It does stand in some tension with both truthmaker theory and Truth Supervenes on Being. However, it is possible to believe in Hypotheticalism while embracing either TSB or a stronger truthmaker theory. One way of doing this would be to treat possible worlds as real, concrete entities existing in "logical space," a view called Concretism (which we will discuss in detail in Chapter 14). Alternatively, one could posit simple hypothetical truthmakers in the actual world whose existence is sufficient to make some counterfactual conditional true. One way of doing this would follow the lead of Luis Molina, introducing special, primitive properties, which Molina called 'habitudes'. If it's true, to use Plantinga's example, that Curley the mayor would have accepted a $300,000 bribe, had he been offered one, we could suppose that this conditional is made true by virtue of Curley's having the appropriate habitude property: the property of *being the sort of person who would accept a $300,000 bribe if offered one*.[6]

An important point: although these habitudes are properties of individuals, they are not intrinsic properties of them (see our discussion of intrinsic properties in Section 3.1.4). If a habitude were part of the intrinsic nature of a thing, then the unique result associated with the habitude (like accepting the bribe) would be pre-determined by Curry's nature. Habitudes, however, are supposed to be the sort of thing that fills the gap between what is pre-determined to happen and what would happen. If we suppose the truth of conditionals to be grounded in the intrinsic natures of things, we've moved to a Strong Powerist (4.4A.3) position (see Chapter 6), not Hypotheticalism. Consequently, for Hypotheticalists a habitude must be an external relation, rather than an internal relation (Def D2.2), between an individual (like Curley) and some non-actual state of affairs (like being offered the bribe).

What about non-actual individuals? Suppose Pope John Paul II had had children. Might any of them have accepted a $300,000 bribe, if offered one? Molina and Plantinga suppose that this question must have an answer (despite the difficulty of our finding it out). Molina supposed that even merely possible individuals can have habitudes, grounding the truth of conditional propositions.

Alvin Plantinga (1974), however, argues that only existing things can have properties. Instead, he makes use of the corresponding *haecceities* or *thisnesses*. According to Plantinga, each actually existing thing, like Socrates, has its own unique thisness. Socrates is unique in having the property of *Socrateity*, the property of *being Socrates* (that very man). Nothing else could possibly have had that property. Since Socrates might not have existed, this property of *Socrateity* might have had no actual instances. It might have been like the property of *being a unicorn*. Since there could have been things that don't actually exist, Plantinga infers that there are many thisnesses that lack their corresponding instance in the actual world. For example, there might be an un-actualized thisness, which, had it been actualized, would have been an child of Pope John Paul II, and which would have been accompanied or *co-instantiated* by the habitude of *being the sort of person who would have accepted a $300,000 bribe if offered one*. (We will discuss haecceities again in more detail in both Chapters 9 and 16.)

Plantinga's neo-Molinist account seems to be quite a lot to swallow—we have to accept both un-actualized thisnesses and the peculiar, simple, and extrinsic habitude properties. Perhaps that is the best metaphysical account of counterfactual conditionals, but it would seem reasonable to look hard for an alternative with fewer somewhat mysterious and *sui generis* entities and properties. Another way to read Plantinga is as embracing Hypotheticalism but denying that (fundamental) hypothetical truths have a ground at all. On this view, you don't embrace habitudes and ground counterfactual facts in them; instead, you just keep asserting counterfactuals ("what would have had that haecceity would have done A, had such and such occurred")!

## 4.3 Anti-Hypotheticalism and Laws of Nature

Let's turn, then, to Anti-Hypotheticalism, the view that the truth of conditionals is conceptually or logically grounded in non-conditional facts. As we've seen, the evaluation of conditionals seems to turn on two sorts of facts about the actual world: the laws of nature, and the true propositions that are co-tenable with the conditional's antecedent. If we knew these two sets of truths, we could define which antecedent-verifying worlds are 'closest' in the relevant sense to the actual world.[7]

Here's a simple proposal for identifying the co-tenable truths, offered by Frank Jackson (1987). Let's suppose that the antecedent of the conditional refers either to the non-occurrence of some specific, actual event or to the occurrence of some specific, counterfactual event. In either case, we can identify a specific time and place that is uniquely relevant to the actual falsity of the antecedent. The contents of this antecedent-falsifying location have some place in the actual network of causes and effects. Jackson proposes that any truth describing an event or state occurring entirely before the antecedent-falsifying location is co-tenable with the antecedent, as is any truth describing an event or state simultaneous with or shortly after the location, so long as this event or state is not causally *posterior* (an effect or effect of an effect, etc.) to anything in the antecedent-falsifying location. Roughly speaking, when we evaluate the counterfactual conditional ($p \mathrel{\Box\!\!\rightarrow} q$), we can treat as co-tenable with $p$ any truth pertaining to a time before the time at which $p$ was falsified, and any truth that is not causally "downstream" from $\sim p$. Here are a couple of examples:

(18) If Curley had been offered a $300,000 bribe last Monday morning, he would have accepted it.
(19) If the circuit breaker had failed at the time of the power surge this afternoon, a fire would have resulted.

In conditionals like these, some counterfactual event or condition is supposed to occur at some precise time. In evaluating the conditionals, we treat as fixed the actual conditions of the world up to the time involved in the antecedent. We then add the counterfactual condition to that set of actual conditions, and we suppose that the world would subsequently evolve according to the actual laws of nature. Thus, we take Curley's character in the actual world at 10 a.m. last Monday (not his character at any earlier or later time); we add the event of the bribe, and then we apply what we know about the laws of human psychology to predict how he would have responded. Similarly, we keep the condition of the circuit and the circuit breaker at the time of the power surge as fixed. We also assume that the power surge occurs when and how it actually occurred, and then we add the supposition that the circuit breaker fails. We use what we know about the laws of physics to predict the subsequent behavior of the system. There may be other conditionals that cannot be evaluated in this way (consider Lewis's example: 'if kangaroos had no tails, they would topple over'), but this class seems to be of special and central importance.

We will take up questions about the direction of time and of causality in later chapters. For present purposes, let's pretend that we can always determine the co-tenable truths relevant to a counterfactual conditional and focus exclusively on the role of laws of nature in supporting such conditionals.

One possibility, then, is that the truth of counterfactual conditionals is wholly grounded in truths about the laws of nature.

What are the laws of nature? To be a law of nature is not simply to be a true universal generalization. In fact, being a true universal generalization is neither necessary nor sufficient for being a law of nature. It is not sufficient because true universal generalizations like 'all the coins in my pocket are bronze' are pretty obviously not laws of nature. But even more interestingly, laws need not hold without exception. Some laws are "oaken" rather than "iron," to use David Armstrong's distinction. Laws of biology and economics, for example, seem to admit of exceptions. For example, deficit spending in a recession stimulates employment, but this effect can be neutralized by a number of factors, including anxiety about the future on the part of investors and business managers. Birds fly, but ostriches, penguins, and dodos do not. Thus, being a statement corresponding to a true universal generalization is not even a necessary condition for being a law of nature.

Are the laws of nature metaphysically fundamental or are they grounded in still more basic truths? If the laws of nature are derived truths, what do they derive from? If laws are not fundamental, there are at least three possibilities: (i) the laws of nature are grounded in the truths of counterfactual conditionals, (ii) they are grounded in truths about the powers and dispositions of particular things, or (iii) they are grounded in other truths about particular things, truths having nothing to do with powers and dispositions.

**4.2T Nomic Fundamentalism.** Some truths about the laws of nature are fundamental.

**4.2A Nomic Reductionism.** No truths about the laws of nature are fundamental.

What would the world be like if the laws of nature were fundamental? Such a view has been defended by Fred Dretske (1977), David M. Armstrong (1983), and Michael Tooley (1977). The Dretske/Armstrong/Tooley, or DAT, account of laws holds that a law consists in a special, irreducible relation of nomological necessity connecting one or more properties or universals. For example, if it were a law of nature that water freezes at 0°C, then this law would consist in a special nomic link between the property of being water at 0°C and that of becoming frozen. We will look at the DAT account of laws in Section 5.1.

Laws could, however, be reduced to counterfactuals. For instance, we could define a law of nature as a truth that would remain true under any antecedent that does not falsify any law of nature or any law of logic or mathematics. Of course, put this way, the definition would be circular. The circularity can be eliminated by defining the laws of nature as the smallest set of truths $T$ (strictly larger than the set of logical and mathematical truths) such that every member of $T$ would remain true under any antecedent not falsifying any member of $T$ (see Lange 2004). However, such a reduction would be of interest only to Hypotheticalists, and not to Anti-Hypotheticalists, since the latter assumes that the truth of conditionals is grounded in that of the laws, and not vice versa. We will look take another look at a strong version of Hypotheticalism in Section 4.4.

Could the laws of nature be grounded in the *powers* and *dispositions* of particulars? The law that ice freezes at 0° C could be grounded in the passive power of water to freeze when its temperature is lowered to 0° C. Any causal law of nature could be expressed in terms of all of the instances of one kind of thing $A$ having the active power to affect any instances of a second kind of thing $B$ when the two are in a certain relation $R$, producing the result that the second thing takes on a new characteristic $C$. The law could be taken as simply expressing the fact that all $A$'s have the active power of producing effect $C$ in any $R$-related $B$ or that all $B$'s have the passive power of being made $C$ by any $R$-related $A$.

Powers and dispositions seem to come in five varieties: active powers, passive powers, immanent powers, resultants, and tendencies.

**Def D4.1 Active Power.** A property $P$ is an *active power* if and only if, necessarily, whenever a thing has $P$ there is the possibility of its producing a specific kind of effect $E$ on some other thing under specifiable conditions by virtue of having $P$.

**Def D4.2 Passive Power.** A property $P$ is a *passive power* if and only if, necessarily, whenever a thing has $P$ there is the possibility of its being affected in some specific way $E$ by some other thing under specifiable conditions by virtue of having $P$.

**Def D4.3 Immanent Power.** A property $P$ is an *immanent power* if and only if, necessarily, whenever a thing has $P$ there is possibility of its producing some intrinsic change in itself under specifiable conditions by virtue of having $P$.

**Def D4.4 Resultant.** A property $P$ is a *resultant* if and only if, necessarily, whenever a thing $x$ has $P$ at $t$, there is some earlier time $t'$ and some passive or immanent power $M$ such that $x$ has $P$ at $t$ by virtue of its having exercised $M$ at $t'$.

**Def D4.5 Tendency.** A property $P$ is a *tendency* if and only if, necessarily, whenever a thing has $P$ there is a certain likelihood or propensity for it to exercise one of its active or immanent powers under specifiable circumstances by virtue of having $P$.

We will use the terms 'power' and 'disposition' interchangeably as general terms for any property of any of these five kinds.

We can now ask: are truths about the powers of things fundamental or derived?

**4.3T Power Fundamentalism.** Some truths about the powers of particular things are fundamental.

**4.3A Power Reductionism.** No truths about powers are fundamental.

If powers are not fundamental, there are two plausible accounts of how they can be derived from other truths. They are derivable either from truths about the laws of nature or from truths about counterfactual conditionals.

There are thus four attractive positions: Strong Hypotheticalism (conditionals are fundamental, but not powers or laws), Strong Nomism (laws are fundamental, but not powers or conditionals), Strong Powerism (powers are fundamental, but not laws or conditionals) and Neo-Humeism (none of the three are fundamental).

**4.4T Neo-Humeism.** None of the truths of counterfactual conditionals and none of the truths about laws of nature or about the powers of particulars are fundamental.

**4.4A Anti-Humeism.** Some of the truths of counterfactual conditionals and none of the truths about laws of nature or about the powers of particulars are fundamental.

**4.4A.1 Strong Hypotheticalism.** Some of the truths of counterfactual conditionals are fundamental, but no truths about particular powers or the laws of nature are fundamental.

**4.4A.2 Strong Nomism.** Some of the truths about laws of nature are fundamental, but no truths about particular powers nor any of the truths of counterfactual conditionals are fundamental.

**4.4A.3 Strong Powerism.** Some of the truths about the powers of particular things are fundamental, but no truths about the laws of nature, nor any of the truths of counterfactual conditionals, are fundamental.

Here is a diagram illustrating the difference between Neo-Humeism and Strong Powerism:

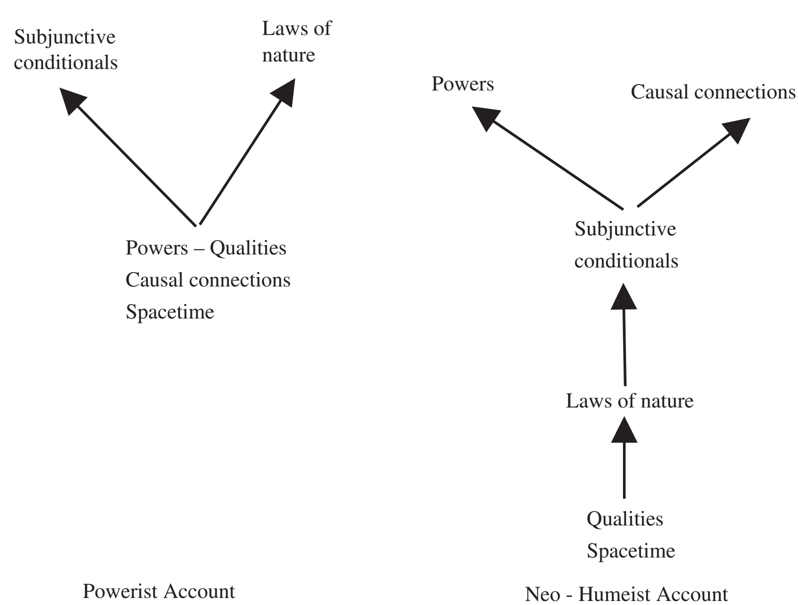


**Figure 4.1** Comparing Powerism and Neo-Humeism

Strong Hypotheticalism and Strong Nomism differ from Powerism simply by shifting either subjunctive conditionals or laws of nature from the upper, dependent position to the base. The Neo-Humeist view is unique in having a simpler base (only qualities and spacetime) and by involving two steps of reduction: reducing laws to the "Humean mosaic" of qualities in spacetime, and then reducing powers, causal connections, and subjunctive conditionals to the laws of nature (together with the mosaic).

If we were to reject all four positions, there would be only two options. First, we could suppose (along with Quine) that there are (strictly speaking) no truths of any kind in this area, no true conditionals, laws of nature or attributions of powers. Second, we could suppose that two or more of the classes of propositions have their own, distinctive truthmakers. The first option is a radical departure from science and common sense, and the second involves apparently redundant or superfluous truthmakers.

## 4.4 Strong Hypotheticalism: Counterfactual Accounts of Powers and Dispositions

Strong Hypotheticalists hold that the truths of counterfactual conditionals are fundamental, and that truths about the laws of nature and of powers and dispositions are reducible to them. We are not going to question the reducibility of laws to conditionals, but we do want to take a close look at the claim that powers can be reduced to conditional truths:

Let's consider fragility, a paradigm case of a dispositional property or power. Fragility is an example of what we have called a 'passive power', the power to be affected in a certain way by certain things. Strong Hypotheticalists are committed to the thesis that all facts about powers are wholly grounded in facts about conditionals:

**Reduction of Powers to Conditionals.** Necessarily, for any *x*, if x has the *B*/*A* power (the power to bring about *B* when in state *A*), then *x*'s having the *B*/*A* power is always wholly grounded in the truth of the subjunctive conditional: if *x* were in *A*, *x* would *B*.

Plausibly, the attribution of fragility to a thing *x* is grounded in the fact that *x* would break if struck. The very meaning of 'fragility' is tied to the connection between breaking and being struck. Fragility is the breaking/being-struck power or disposition. Reduction of Powers to Conditionals asserts that a thing *x*'s having the *B*/*A* power consists in the appropriate conditional proposition's being true—namely, the proposition that *x* would *B* if *A*. If this is so, then the truth of the corresponding conditional should be both necessary and sufficient for the truth of the predication of fragility to *x*.

It is plausible to suppose that if the truth of one proposition *p* is necessarily wholly grounded in the truth of another proposition *q*, then the two propositions are metaphysically equivalent (i.e., each is a necessary and sufficient condition for the other):

**Principle of Metaphysics (PMeta) 1 The Necessary and Sufficient Condition Test for Necessary Grounding.** If the truth of a proposition *p* is necessarily wholly grounded in the truth of proposition *q*, then the truth of *q* is a metaphysically necessary and sufficient condition for the truth of *p*.

This principle of metaphysics would be true if thesis 3.9T, the Grounding-Entailiment Entailment thesis, is true:

**3.9T Grounding-Entailment Entailment.** Necessarily, if the fact that $p$ grounds the fact that $q$, then $p$ metaphysically entails $q$ (i.e., necessarily if $p$ is true, then $q$ is true).

In Chapter 3, we found some reasons both for and against 3.9T. If we end up rejecting 3.9T, then we should replace the Necessary and Sufficient Condition Test for Necessary Grounding with the weaker Necessary Condition test:

**Principle of Metaphysics (PMeta) 1.1 The Necessary Condition Test for Necessary Grounding.** If the truth of a proposition $p$ is necessarily wholly grounded in the truth of proposition $q$, then the truth of $q$ is a metaphysically necessary condition for the truth of $p$.

C.B. Martin (1994) has produced some examples of so-called 'finkish dispositions' that demonstrate that the truth of the 'if *A* then *B*' counterfactual conditional is neither necessary nor sufficient for the truth of the attribution of the *B*/*A* disposition (see also Lewis 1997, Lowe 2010).

Here is an example of a finkish disposition which shows that the truth of the 'if *A* then *B*' counterfactual is not necessary for the truth of the attribution of the B/A disposition. Suppose there was a genius mad scientist who was obsessed with a particular vase. This vase is especially fragile because it's made of extremely thin blown glass. But our genius mad scientist's obsession with this vase caused him to invest his energies in producing a device of the following sort: if ever the vase is about to be struck, this device instantaneously alters the chemical structure of the vase, rendering it unbreakable. The vase's fragility is "finkish" in that it cannot manifest itself in the vase's breaking given the presence of the device. The device is, in that sense, a finking mechanism. In this case, we have something that is fragile, even though it is *false* that it would break if struck. The truth of the conditional is not necessary for the truth of the attribution of the *B*/*A* disposition.

Conversely, suppose we start with a vase that is not fragile. But this time there is a finking mechanism which instantaneously alters its microstructure, rendering it breakable, whenever it is about to be struck. In this case, we have something that is *not fragile*, even though it is true that it would break if struck. This shows that the truth of the 'if *A* then *B*' counterfactual is not sufficient for the truth of the attribution of the *B*/*A* disposition.

The moral of the story seems to be that a thing has the *B*/*A* disposition just in case it is the sort of thing that would *normally B* if *A*, absent any exogenous interference. We must build into the relevant counterfactual conditionals the further condition that the object in question not be altered in its intrinsic character. This means that the truth of an attribution of a B/A disposition cannot consist merely in the truth of the corresponding conditional 'if *A* then *B*'. Instead, one must have in addition an account of the intrinsic character of the object. Either the disposition itself is part of that intrinsic character, in which case we cannot reduce having the disposition to satisfying the conditional without circularity, or the attribution of the disposition involves other facts about the properties of the object, beyond the bare fact about how it would respond to the relevant stimulus. In either case, the having of a disposition does not consist merely in satisfying the conditional that we ordinarily associate with the disposition.

In 1997, David Lewis proposed a slightly more sophisticated fix for the conditional analysis of dispositions (Lewis 1997):

**Lewis's Analysis of Dispositions.** $x$ has the $B/A$ disposition if and only if $x$ has some intrinsic property $C$ such that $x$ has $C$, and, if $x$ were to be made to have $A$ and continued to be $C$ for some finite period of time, the fact that $x$ has both $A$ and $C$ would be sufficient to cause $x$ to have $B$.

If we try to employ Lewis's "fix," we have to ask ourselves: *why* must we hold the intrinsic character of the object fixed, when hypothetically testing for the appropriate reaction? Indeed, why must we suppose that the object has to have any intrinsic property, other than the disposition itself? The Lewisian Strong Hypotheticalist can offer us no explanation of this fact. Therefore, it seems that conditionals are grounded in dispositions, not the other way around. Moreover, Lewis's solution makes use of the idea of *causation*: the intrinsic property $C$ must (together with $A$) have the *power* to produce $B$. Lewis's solution can help the Strong Hypotheticalist only if the notion of causation can be reduced to conditionals, an issue that we will discuss in Section 27.1. By way of preview, it seems unlikely that such a conditional analysis of causation can succeed.

Therefore, Power Fundamentalism (4.3T) is vindicated, and we have yet another reason (in addition to worries about how conditionals could be fundamental) for rejecting Strong Hypotheticalism.[8]

In addition to finkishness, there are two additional problems with the hypotheticalist account of dispositions: dispositions can be *masked* and *mimicked*. For example, suppose that a fragile vase can be filled with a certain kind of foam in such a way that the foam absorbs all shocks delivered by striking the vase. The vase remains fragile, and yet it doesn't break when struck because the foam masks or acts as an antidote to the manifestation of the disposition (Johnston 1992, Bird 1998). Alternatively, suppose that an illusionist has set up a sonic beam that would cause a sturdy rock to break whenever it is tapped. The rock isn't in fact fragile and never becomes really fragile, but the illusionist causes the rock to mimic the disposition (Smith 1977, Prior, Pargetter, and Jackson 1982, Lewis 1997, Armstrong 1997). Cases of masking and mimicking are in fact quite common (Fara 2005).

## Notes

1 We will treat 'counterfactual conditional' and 'subjunctive conditional' as though they were synonyms. It is permissible to use a subjunctive conditional even when one is not sure that the antecedent is false (especially when it concerns the future), but it is always odd to use it when one believes the antecedent to be true.

2 There are two niceties omitted from this characterization of possible worlds semantics. First, we need the rules that take us from the truth-values of the atomic sentences to the truth-values of all sentences. Second, we need an accessibility relation on the domain of worlds. Neither of these is necessary for understanding what we're up to here, and thus the complications they would introduce are needless.

3 We can generalize this by making * a function from world-proposition pairs to worlds and changing the definition of counterfactual truth to:
Stalnaker′. ‘($p$ □→ $q$)’ is true at $w$ if and only if $q$ is true according to $*(w,p)$
4 Again, there are niceties here that we ignore. For ease of exposition, we assume “strong centering,” that the actual world is the only world in the innermost sphere.
5 The result is similar to the way that modern logic (following Gottlob Frege) treats “vacuous” quantification. If there are no unicorns, then modern logic counts both ‘all unicorns are ugly’ and ‘all unicorns are beautiful’ as true.
6 Some conditionals involve more than one individual, such as ‘if the forbidden fruit had been a kumquat, Adam and Eve wouldn’t have eaten it’. A Molinist would have to suppose that in such cases, the habitude is a special kind of relation holding between Adam and Eve. In general habitudes could involve any number of individuals, perhaps even infinitely many.
7 One important thing to bear in mind: the sense of ‘closeness’ or ‘similarity’ that is used in Lewis’s semantics for the counterfactual conditional is a special, technical sense. The words ‘close to’ or ‘similar to’ are terms of art. We don’t look for those worlds that are most similar, all things considered, to the actual world, as Lewis’s response to Jonathan Bennett makes clear (Bennett 1984, 2003). Consider the proposition: if the President were to press the nuclear button, all of civilization would end. This proposition is probably true, even though worlds in which the consequent are true are very dissimilar (we hope) from the actual world.
8 Bonevac, Dever, and Sosa have proposed altering the semantics and logic of the conditional in such a way as to avoid the refutation of the conditional account by finkish dispositions (Bonevac, Dever, and Sosa 2006). Their new “normality” conditional does indeed seem to correspond in the right way to the predication of powers and dispositions, but it seems clear that, metaphysically speaking, their work only confirms the fact that powers and dispositions are fundamental.

# 5

# Laws of Nature

## 5.1 Strong Nomism: The Dretske-Armstrong-Tooley (DAT) Theory of Laws

As we mentioned above, Fred Dretske (1977), David M. Armstrong (1983), and Michael Tooley (1977, 1987) have all proposed that the truths about the laws of nature are metaphysically fundamental, consisting in a primitive, unanalyzable relation of "necessitation" holding between two or more properties or universals. This is Strong Nomism (4.4A.2). According to Strong Nomism, the laws of nature determine which counterfactual conditionals are true, and they also determine which powers and tendencies particular things have. For example, if it is a law of nature that whenever an *F* encounters a *G* in relation *R*, the *G* becomes *H*, then we can say that the law of nature confers a corresponding active power on all *F*'s and a corresponding passive power on all *G*'s. Similarly, if it is a law of nature that all *F*'s become *H*'s within a certain span of time, then the law confers on all *F*'s the immanent power of making themselves *H*'s within that span of time.

We also noticed that laws of nature can be probabilistic and "oaken" or exception-permitting. Such probabilistic and oaken laws can be thought of as conferring *tendencies* of the appropriate kind to the relevant objects. Thus, it seems that laws can indeed ground the truths of both conditionals and the attributions of power.

Let's take a closer look at the connection between laws and powers. In particular, we need to be clear about the issue of *priority*: are the laws grounded in truths about powers or vice versa? If the laws are supposed to be fundamental, with the powers of particular things grounded in them, how exactly is this supposed to work? What is the logical form of the laws of nature, from which we are supposed to be able to derive the attributions of powers to particular things? One option has the laws of nature directly state that certain kinds of things have certain powers and tendencies. On this view, we can certainly derive the attribution of powers and tendencies to such things from those laws. But it would

*The Atlas of Reality: A Comprehensive Guide to Metaphysics*, First Edition.
Robert C. Koons and Timothy H. Pickavance.

seem that the laws are then made true by the particular attributions, not vice versa. For example, if it were a law of nature that all fire has the power to burn wood, then we could infer that *this* fire has the power to burn that wood. But if we ask what makes it true that *all* fire has this power, it would seem that this universal generalization ('all fire…') is made true by its instances or by the absence of counterexamples (fire without the power to burn wood). Either way, the attributions and non-attributions of powers to particular things would ground the truth of the law, not the other way around.

As a second option, suppose the laws of nature merely say that all instance of fire's being brought into proximity to wood are followed in time by the burning of the wood. This is logically independent of the attribution of any power to the fire, so the law of nature could be independent of and prior to the particular attributions. However, now we are no longer able to deduce that fire has the power to burn wood from the law. The law merely says that the placing of fire and wood in proximity to one another is always *followed* by the wood's being burned. It says nothing about any instance of fire's having the *power* to burn the wood, which is what is needed.

Dretske, Armstrong, and Tooley adopt a third option concerning the logical form of laws of nature. The laws of nature neither ascribe powers to particular things, nor do they merely describe the unvarying sequences of events. Instead, a law of nature attributes a special kind of relation, nomic necessitation, to an ensemble of properties or universals. This view, the Dretske/Armstrong/Tooley theory of laws, is the most prominent and plausible Strong Nomist position in the literature. Through the remainder of this chapter, we will treat Nomism as committed to the DAT theory.

Return to the fire and wood case. The supposed law of nature would state that the nomic-necessitation relation holds between the following pair of complex properties: (1) being some wood brought into proximity to some fire, and (2) being some wood that is burned. As van Fraassen (1987, 1989) has pointed out, however, this leaves it somewhat mysterious how we are supposed to derive any information about particular cases of fire and wood from the fact that this supposed relation holds between the two properties. We want there to be some connection between the attribution of powers to things and the actual sequence of events.

Strong Nomists must simply posit a brute necessity between the holding of a law and the corresponding universal generalization. Armstrong was reluctant to admit this, since he was attempting to build a metaphysical system with no brute necessities whatsoever, but this was a quixotic quest. As we've seen, it is a corollary of Ockham's Razor (**PMeth 1.2**) that we should prefer a theory with fewer *ad hoc* postulations of necessities. Where necessities must be posited, we should prefer theories that derive the necessity from the essential structure of the things involved. Nomism offers no such explanation of the necessary connection between laws and generalizations, so this must be counted as a cost of the theory.

The main compensating advantage is that Nomism provides a metaphysical explanation of the difference between lawful and accidental generalization. A lawful generalization is one that corresponds to a relation of nomic necessitation connecting the relevant properties, while accidental generalizations have no such counterpart. However, this advantage is not unique to Nomism, as we shall see. A Strong Powerist account can make a similar distinction between generalizations that are supported by powers and those that are not.[1]

A key question for Nomism is this: is the relation of nomic necessitation an *internal* relation among properties involved? Here again is the definition of an internal relation:

**Def D2.2 Internal Relation.** *R* is an *internal relation* if and only if necessarily, for every *x* and *y*, whether *R* holds between *x* and *y* depends only on the intrinsic properties of *x* and of *y*.

A non-internal relation between two things depends on more than the intrinsic qualities of the two things. Here is an example of each kind of relation. If one book is twice as long as another, then they are related by an internal relation, the relation of *being twice as long as*. As soon as we know the length of each book (where length is intrinsic to each book), we know whether or not they stand in that relation. In contrast, if one book is higher than another in the Amazon.com sales ranking, they are thereby *not* internally related. You could know all there is to know about the intrinsic features of the two books without knowing which ranks higher than the other on Amazon. We will discuss internal relations in more detail in Chapter 18, on the structure of space.

It seems clear that the Nomist must count the nomic relation as an *non-internal* relation. If it were internal, then, assuming that properties have their intrinsic features essentially, the properties involved would have active and passive powers in their very natures. Thus, powers would be among the fundamental features of the world. This would fit with a Strong Powerist picture and not with the Strong Nomist one.

If the nomic relation is not an internal relation, is it a contingent or necessary one? If it is necessary, then its holding with necessity in each case would seem to be brute necessity, with no possible explanation. Similarly, however, if it is a contingent relation, its holding contingently would also seem to be something for which there could be no further explanation. It would be a *brute contingency*.

As we do in the rest of science, we should prefer the simple metaphysical theories to ones that are overly complicated. This is Ockham's Razor (**PMeth 1**). An aspect of Ockham's Razor is that we should minimize the brute external relations posited by a theory, whether necessary or contingent. Such a *brute* relation is one that cannot be explained in any way, whether by the structures or essences or internal character of the relata or by means of some cause.

**PMeth 1.3 Third Corollary to Ockham's Razor.** Other things being equal, adopt the theory that posits the fewest inexplicable and uncaused non-internal relations between things.

On Nomism, the relation of nomic necessitation between properties is inexplicable and necessarily uncaused. A causal connection presupposes the existence of a corresponding causal law. Hence, it is impossible to give a causal explanation of all of the causal laws. For example, the Nomist cannot suppose that God has caused the laws of nature, since for God to have the power to do so, there would have to be a law of nature giving him that power, and God could not have caused that law to hold without vicious circularity. Since the relation of nomic necessitation is external, instances of it cannot be explained by appeal to the structure or internal character of the properties involved.

Why do Nomists reject the idea that laws are made true by an internal relation between properties? Their motivation seems to have involved a desire to avoid the supposed obscurities and mysteries of an Aristotelian or scholastic essentialism. Since the seventeenth century, prominent philosophers and scientists (including Descartes, Galileo, Boyle, Locke, and Hume) have rejected the idea of essential powers as "occult" and vacuous. The French comic dramatist Molière famously poked fun at scholastic essentialism when he has a scholastic alchemist (in his play "The Imaginary Invalid") "explain" the power of narcotics to cause sleep by hypothesizing that narcotics have a "dormitive virtue" or "virtus dormitiva" (which is simply Latin for 'the power to cause sleep'). However, it is far from clear that Nomism involves less, rather than greater, mystery.

To sum up, from the perspective of simplicity and economy, Nomism suffers from four defects, as compared with either Strong Powerism or Neo-Humeism:

1 Nomism must posit an additional, irreducible relation of nomic necessitation between properties.
2 Nomism must posit that whenever the nomic-necessitation relation holds between two properties, its holding between them is a brute, inexplicable, and uncaused fact.
3 As a consequence of (1), Nomism must treat properties as real entities, capable of entering into relations.
4 Nomism must posit a brute necessity between each law of nature and the corresponding generalization, whether universal (e.g., all *F*'s are followed by *G*'s) or statistical (e.g., in *x*% of the cases, *F*'s are followed by *G*'s).

There are three additional objections to be lodged against Nomism:

5 There are a variety of logical forms that laws of nature can take. Few, if any, have the simple form of a universal generalization involving just two universals, like $(x)(Fx \rightarrow Gx)$.

Laws might involve disjunctions or conjunctions in various places, like (1) or (2):

(1) $(x)((Fx \& Gx) \rightarrow Hx)$ Everything that is both *F* and *G* must be *H*.
(2) $(x)(Fx \rightarrow (Hx \text{ or } Jx))$ Everything that is *F* must be *H* or *J*.

As Tooley (1977) recognized, Nomists need a distinct necessitation relation for each logical form. In the simple case, we have one relation, $N_1$, holding between *F* and *G* and entailing $(x)(Fx \rightarrow Gx)$. In the case of laws (1) and (2), we would need two additional necessitation relations, $N_2$ and $N_3$, with $N_2$ holding between *F*, *G* and *H* and entailing (1), and $N_3$ holding between *F*, *H*, and *J* and entailing (2). If there are a large number of laws with a wide variety of logical forms, Nomism could be saddled with positing a large number of distinct necessitation relations.

6 Nomism must explain why we tend to accept simple scientific theories about the laws of nature. Whenever we find the data fitting to a simple pattern (such as a smooth, mathematically simple curve), we come to be quite confident that the data are to be explained by a correspondingly simple law of nature. We believe that gravitational

force varies inversely with the square (power 2) of the distance, not with the distance to the power 2.000003, for example. This preference for simple laws is hard to explain on Nomism. What reason would we have for thinking that the nomic-necessitation relation is more likely to hold between two simple properties than between two complex ones? As we shall see, the Neo-Humeist has a neat explanation of this fact.[2]

The Nomist could always posit that it is simply a basic, inexplicable principle of rationality that we should always prefer the theory with the simplest law, but this would be a still further cost of the theory, in light of the first corollary of Ockham's Razor:

**PMeth 1.1 First Corollary of Ockham's Razor.** Other things being equal, prefer the theory that posits the fewest primitive, underivable postulates of reason.

7 There is still another potential drawback to Nomism, one concerning functional laws that relate two or more quantities. An example would be Newton's force law, $F = ma$. When a force of quantity $x$ Newtons is applied to a mass of $q$ kilograms, the mass accelerates at a rate of $x/q$ meters per second per second. The values of $F$ or $m$ can be any real number from 0 to infinity. Thus, there are an infinite numbers of property-triples of force, mass, and acceleration, one for each pair of force and mass quantities. On Nomism, the force law $F = ma$ is in reality an infinite collection of laws, each with a fundamental nomic-necessitation connection among a different triple of force, mass, and acceleration quantities. This involves a huge inflation of the Nomist ontology.

In addition, many Nomists (like David Armstrong) want to deny the existence of uninstantiated universals. However, it seems likely that there are many specific quantities of mass, force, and acceleration that are never instantiated in the actual world (a mass greater than the total mass-energy of the entire universe, or a mass smaller than that of any mass-bearing fundamental property). These possible quantities would seem to be things that would necessarily obey the relevant functional laws, but the Nomist cannot explain this, since the "missing" universals cannot bear the nomological necessity relation because they simply do not exist in fact.

Strong Powerists can dodge objection 7 if they treat generic determinables (like mass and acceleration) as real universals, in addition to the various determinate quantities of mass (e.g., 1 kg, 10 kg) and of acceleration (1 m/sec$^2$, 5 m/sec$^2$, etc.). This might seem to involve some truthmaking redundancy: why should each massive object have both some specific mass and the general determinable of massiveness? However, this can be worked out in a plausible and elegant way (see Chapter 10).

Finally, we could describe the debate between Strong Nomists and Strong Powerists in this way: Nomists believe in the relation of nomic necessitation, which consists in the possession of certain powers on the part of universals. When universal $U_1$ is tied by nomic necessitation to $U_2$, the relation of nomic necessitation confers on universal $U_1$ the active power of making its instances be instances of $U_2$, while conferring the complementary passive power to $U_2$. In contrast, Powerists believe that the fundamental powers are possessed by ordinary particulars (like particles or people), not universals. The Powerist account seems a simpler and more natural way to think about the matter.

## 5.2 Neo-Humeism: Reduction of Conditionals, Laws, and Powers

The Scottish philosopher David Hume famously argued that we have no good reason to believe that there are any "necessary connections" in the world. The only kind of necessity or impossibility that Hume was willing to accept was that generated by the connections among our ideas or concepts. In the twentieth century, this Humean perspective has been revived by such philosophers as Frank Ramsey (Ramsey 1978/1928) and David K. Lewis (Lewis 1980b, 1986a, 1994). For the Neo-Humeist, all the truths or supposed truths about powers, counterfactual conditionals, and laws of nature are grounded in and reducible to truths about the actual distribution of ordinary, qualitative properties in space and time.[3]

The Neo-Humeist program proceeds in the following way. First, give a reductionist account of the laws of nature, and then use the laws to ground the truths of attributions of power and of counterfactual conditionals. We'll focus in this section on the Neo-Humeist or Ramsey/Lewis Theory of the laws of nature.

The key problem is to account for the difference between lawful generalizations and mere accidental generalizations. The Ramsey/Lewis Theory proposes that the difference between the two depends on the way in which a generalization does or does not fit into our best scientific theory of the world. By 'best' scientific theory, Ramsey and Lewis do not mean the one that is fundamentally true, but rather that theory that combines the most 'virtues', where the standards of theoretical virtue are fixed by the conventions and customs of our actual scientific practice. In particular, we seem to value two things in our theories: (1) *good fit* between the theory's predictions and observed experimental results (the theory predicts *all and only* the observed results), and (2) *overall simplicity* of the theory, in terms of its basic vocabulary, fundamental postulates, and mathematical form. According to Lewis (1986a), we should be willing to accept some discrepancy between the theory and the data if the theory is much simpler than its more accurate competitors, and we should be willing to accept a relatively complex theory if its predictions are much better than any simpler one. The "best" theory is the one that achieves the best trade-off between these two values.

This Neo-Humeist account has two principal advantages. First, it is ontologically very simple. It posits no fundamental truths involving powers, laws, or counterfactual conditionals. Second, it has a simple explanation for our preference for simple scientific theories. A simple account of the laws is more likely to be true, since to be a law is nothing more than to be a generalization that belongs to the simplest account of nature. A simpler theory is more likely to be "true" because simplicity is one of the two factors that make theories true.

There are three major objections to the Neo-Humeist account of laws. First, it makes the laws of nature dependent on us, on our practices and preferences, in a way incompatible with scientific realism. Second, it makes the powers of things extrinsic to those things and faces counterexamples involving hypothetical "small worlds." Third, it has difficulty explaining the rationality of induction, that is, of our confidence that unobserved cases (such as those in the future) will be relevantly similar to observed ones.

### 5.2.1 Argument from scientific realism

On the Neo-Humeist account, what makes something a law of nature is the fact that it would fit into a theory that would best satisfy our preferences, as fixed by our actual scientific practices. This seems to make scientific reality relative to the contingencies of our conventional perspective. Scientific truth would no longer be something objective and mind-independent. This would contradict what seems obvious about the progress of science, namely, that in science we discover truths about the world as it is, independent of ourselves. The reality of the law of gravity, for instance, does not depend in any way on us or on our preferences.

There is a Neo-Humeist response to this objection developed by David Lewis (1973b, Section 3.3). One can "rigidify" the reference to our practices and preferences, so it is our practices and preferences in *this*, the actual world, that fixes the meaning of 'law'. This "rigidification" makes the laws of nature independent of us, in the sense that the laws would have been the same, even if our practices and preferences had been different, since it is our practices in *this* world that determine what the laws are in all worlds, regardless of what our practices and preferences might have been in those other worlds.

A similar rigidification could take place with respect to time. Suppose that our standards and preferences have changed over the centuries and will continue to change. The Lewisians suppose that what we mean by 'law of nature' changes as the standards change. The phrase 'law of nature' meant something different in the year 1500 than it does now, and different from what it would mean in 2500. The laws of nature don't change. What we mean now by 'law of nature' incorporates in a "rigid" way today's standards, and so doesn't vary as the standards change. What is a law of nature now (by our standards) will still be a law (by those same standards) in 2500, whatever people then may mean by 'law of nature'. This is a subtle point, involving the distinction between *using* the phrase 'law of nature' and merely *mentioning* that phrase.

This rigidification comes at a steep cost, however, since it would follow from this account that we may not be at all reliable in identifying the laws of nature. Our conventions and preferences could have been different, and had they been different, we would have systematically misidentified the laws of nature. That we here and now in fact get the laws right (applying the "right" standards) is just a lucky accident, resulting from our occupying this one, special world, in which our actual practices fix the meaning of 'law'.

In order to know what the laws are, we must be reliable at detecting the true laws. This reliability is a matter of our "tracking" the correct laws with our beliefs across a span of possible variation. Let's suppose that we do in fact believe in the true laws of nature (or something close to them) in the actual world. There are three kinds of variation to consider:

1 Would we still believe in the true laws if the laws were the same, and we had the same preferences for theories, but we observed slightly different parts of the actual world?
2 Would we still believe in the true laws, if we had the same theory preferences, and we observed the same parts of the world, but the laws of nature were slightly different?
3 Would we still believe in the true laws, if the laws were the same and we observed the same parts of the world, but our theory preferences were slightly different?

It's variation 3 that potentially creates difficulties for the Neo-Humeist. It seems that slight variations in our theory preferences in nearby worlds could lead us away from the true laws of nature, since what the laws are is fixed (rigidly) by our preferences in the this world.

We may be very lucky—it may be that the very same laws would come out as "best" under a variety of theory preferences near to our actual ones, but there seems to be little reason to think that we are so lucky. We can take the actual variations in our theory preferences in different eras of history to be a good sample of the sort of variations of type 3 that we need to consider. Thomas Kuhn, in his classic work on scientific revolutions (Kuhn 1970), provided evidence that our standards and preferences for "good" theories has changed dramatically over the last 300 years, so much so that our current theories would not have counted as good theories in the relatively recent past. For example, in the time of Copernicus, it was considered critical that a good theory use only circular orbits in explaining the motions of the planets. In the nineteenth century, only deterministic theories were considered good enough to be considered, and Albert Einstein resisted quantum mechanics for most of his life for that reason.

If these worries are justified, then the Neo-Humeist position entails that one of the following two methodological principles involving scientific truth must be overridden:

**PMeth 2.1 Scientific Realism: Objectivity.** Other things being equal, adopt the theory that implies that our best scientific theories are objectively true: true independently of our scientific preferences and practices.

**PMeth 2.2 Scientific Realism: Reliability.** Other things being equal, adopt the theory that implies that we are reasonably reliable in finding scientific truth.

If the laws of nature are not determined by the rigidification of our actual standards, then we must violate objectivity, since what the laws of nature are would then vary from world to world with our varying preferences. This would make all of science a branch of sociology. Alternatively, if we rigidify our actual standards, we salvage objectivity but we put reliability into jeopardy. However, it might turn out that nature is "kind" (as Lewis put it), in that the very same scientific lawbooks come out as best under a wide range of variation surrounding our actual standards of theory choice. If so, then the Neo-Humeists could claim that their theory of laws is consistent with objectivity and reliability. It's hard to tell who's right here. We would have to be able to anticipate which system of laws does turn out to be best by our standards, and then we would have to do a great deal of historical and sociological research to determine how much variation in the standards of theory choice were really possible.

In addition, isn't reliability a problem for every account of laws? Can the Nomist or Powerist do a better job of securing our reliability at detecting the true laws? At the very least we can say that given rigidification, the Neo-Humeist can do *no better than* advocates of other theories at explaining the reliability of using our contingent standards of simplicity as a guide to the true laws. The Neo-Humeists lose a potential advantage. In addition, the Strong Powerist might have a better shot at securing our reliability, in two ways. First, the Powerist might attribute to the human mind an inherent power to recognize the true laws of nature, given enough actual experience, by a kind of rational intuition (what Aristotle called 'noûs'). Second, Powerists could argue that Neo-Humeism exaggerates the importance of global theory choice. Instead, we discover the laws of nature by means of local interactions under carefully controlled conditions (see the Powerist response to the problem of scientific knowledge in Section 6.1.2).

### 5.2.2 The extrinsicality objection and small worlds

According to Neo-Humeism, whether something has a particular power (active, passive, or immanent) depends on the laws of nature, and whether something is a law of nature depends on the overall pattern of particular facts across the entire history of the cosmos. Thus, having a power is an extrinsic feature of a thing, dependent on the pattern of events in remote parts of the cosmos in remote times, both past and future. In contrast, it seems obvious that having a power is an intrinsic feature of a thing.

**PMeta 2 The Intrinsicality of Powers.** Having a power is an intrinsic property.

This fact is confirmed by thought-experiments involving hypothetical "small worlds" (Tooley 1987). Suppose that the world consisted only of a single electron, an electron intrinsically identical to all actual electrons. It seems that such a solitary electron would still be negatively charged, and that, being negatively charged, it would still have the power of repelling other negatively charged things and attracting positively charged things. However, in such a small world, the "best" scientific theory (in the Ramsey/Lewis sense) would include no laws of nature involving charge at all, since adding such a law would make the theory more complicated without enabling any new or more accurate predictions to be made. On the Neo-Humeist account, then, it must be impossible for anything inhabiting such a small world to have any power to move other objects. We seem to be able to coherently imagine small-world possibilities in which things have powers, however, and the methodological principle of Imagination as a Guide to Possibility provides support for the conclusion that such a world really is possible, contrary to Neo-Humeism.

**Principle of Epistemology (PEpist) 1 Imagination as Guide to Possibility.** If a scenario is imaginable in great detail without evident absurdity, then we have good reason to think that it represents a metaphysical possibility.

Here is another similar example, also from Tooley. Suppose that we discovered that the world consists of ten kinds of fundamental particle, and suppose that we have observed 54 of the 55 possible kind-to-kind interactions. However, suppose that particles of type #1 have never interacted with particles of type #10, and never will (perhaps type #1 disappeared a few seconds after the Big Bang, and type #10 didn't appear until millions of years later). Since we know the laws that govern the other 54 possible interactions, it seems reasonable for us to believe that there must be some similar but unknown law governing the type #1-to-type #10 interaction. However, the Neo-Humeist theory entails that there can be no such law, since the simplest theory of the actual interactions in the world would not include any general statement about what would happen in such a case. Adding any conjectured law would only make the total theory more cumbersome without improving its fit with the actual mosaic of facts.

### 5.2.3 Objection based on induction

The final objection to Neo-Humeism concerns the rationality of induction. Once we have encountered a large body of varied data that conforms to a simple scientific theory, we seem reasonable in believing that the theory accurately describes much, if not all, of the world. In particular, we take it for granted that the theory will continue to fit the data in cases that have not yet been observed. We assume, for example, that the future will be in this respect very similar to the past. To use a familiar example, we rationally believe that the earth will continue to rotate on its axis, bringing about many future sunrises and sunsets.

On the Neo-Humeist account, the fact that the particular cases conform to the laws of nature is not explained by those laws (at least, not in the sense that the first is metaphysically grounded in the second). The ontological explanation goes the other way: the laws are laws in part because they conform to the pattern of particular events, not vice versa. For the Neo-Humeist, there are possible worlds corresponding to every imaginable pattern of events. In most of those worlds, the "best" theory is horrendously complicated. We are just lucky to inhabit a possible world where the "best" theory is relatively simple.

But how do we know that we do inhabit such a "simple" world? All we know for sure is that the observed part of our world conforms to a simple system of laws. Let's call the observed part of our world *O*. There are many possible worlds that agree with our world with respect to all of the events in *O*. It is fairly easy to show that, of the worlds that agree about *O*, the overwhelming majority (by astronomically large margins) fail to conform to any simple system of laws in the complement of *O*, Comp(*O*), which encompasses all the events other than those in *O*. Thus, it would seem that we have very good reason to believe that it is unlikely that our world is one of the few simple ones. Neo-Humeism, therefore, gives us strong reasons to doubt the reliability of induction. It gives what epistemologists call an 'undercutting defeater' of our inductive inferences.

Here is a simple, "toy" example of what we have in mind. Suppose that the world consists of 1000 marbles, each of which is either black or white. We have observed 30 of them, and all of the observed marbles have been black. By induction, we conclude that nearly all of the world's marbles are black. However, there is only one possible world in which all 1000 marbles are black. In contrast, there is an astronomically high number of worlds in which the first 30 marbles are black and the rest are half black and half white.

If we don't suppose that we have been interacting with a mechanism that provides us with a random sample of the marbles, the 30 black marbles give us virtually no reason for thinking that most of the rest are black. But Neo-Humeists have no right to talk about "mechanisms" here, since such talk implies the existence of powers and propensities as fundamental realities, exactly the sort of thing Neo-Humeists deny.

Here's another example. Suppose that we have observed a simple pattern, repeated trillions of times over: an event of type *X* followed by an event of type *O*, followed in turn by another event of type *X*. There is one possible world where this same pattern is repeated throughout the history of the world, including all of the so-far-unobserved times. There is an astronomically large number of alternative worlds, where the pattern is violated some, many or all the time beyond the window of observation. The existence of an astronomical number of "anti-inductive" possibilities seems to give us a good reason not to jump to the usual inductive conclusion. This worry is based on the following version of the Principle of Indifference:

**Principle of Indifference.** If the number of possibilities consistent with available data that contradicts a certain guess is known to be astronomically larger than the number of possibilities consistent with the data that fit it, then it is prima facie unreasonable to be certain that the guess is true.

The best Neo-Humeist response is simply to insist that reasoning inductively is part of what we mean by 'being rational'. Despite the large number of complicated worlds agreeing with observed data, it is always most "reasonable" to be confident that we in fact occupy a relatively simple world. It's just a fundamental axiom of reason to do so, in need of no justification and in danger of no refutation. Once again, this response comes at some cost. The Neo-Humeist must treat induction as a brute requirement of reason, without further justification. This can be done, but it must be counted as a cost of adopting the theory, in light of the first corollary of Ockham's Razor (**PMeth 1.1**), which requires that we minimize rational postulates.

There is an alternative solution that is somewhat less attractive. Neo-Humeists could suppose that what it is rational to believe depends on what is true. In that case, it might be reasonable to believe that our world is governed by simple worlds, so long as it actually is. The skeptic would have to prove that the world is not simple in order to prove that we are unreasonable in thinking that it is. The drawback to this proposal is that it violates a plausible constraint on rationality, namely, that what is reasonable must be to some extent independent of what is true. It must be possible to reasonably believe something false and to unreasonably believe something true.

Do Strong Nomism or Strong Powerism have an advantage here or is induction equally a problem for all? Armstrong and Tooley cautiously argued that Nomism does have some advantage. For Nomists, the problem is not to assign probabilities to possible mosaics of qualities (as it is for Neo-Humeists). Instead, Nomists can simply assign probabilities to each possible law, where the truthmaker for a law is a nomic-necessitation relation among certain universals. If we are considering whether or not there is a law that entails that all *F*'s are *G*'s (e.g., all ravens are black), it seems that we could give each possibility a *prior probability* of 50%. If we then observe many black ravens and no non-black ones, the probability that it is a law that all ravens are black should go up, approaching 100% in the limit.

There are two problems with this rosy assessment of the rationality of Nomist induction. First, each possible law will be in logical conflict with a large number of competing laws. For example, that all ravens are black is in conflict with the law that all ravens smaller than 5 kg are black and the rest are white. In fact, there are an infinite number of possible laws in conflict with the law that ravens are black, and so it seems that each possible law must begin with an infinitesimally small prior probability. These tiny priors would constitute a probability trap from which the laws cannot escape, no matter how much data we collect.

Second, this argument for induction works only if we assume that we already know in advance which universals exist and which do not. However, Armstrong and Tooley (at least) believe that we discover the universals that exist empirically, by discovering that they play a role in actual laws of nature. If the world is infinitely large, then there would be infinitely many possible universals to consider and so once again an infinite number of competing laws over which to distribute our finite prior probabilities. (It is a law of the calculus of probabilities that each set of incompatible hypotheses must have a total probability that is no greater than 1.)

Powerist theories of induction will face exactly the same problems, with one important difference. For Powerists, the laws of nature are metaphysically necessary. Hence, there really are no "chaotic" worlds in which the actual laws of nature are violated. These chaotic scenarios are conceivable and can't be ruled out a priori (prior to empirical investigation), but Powerists don't have to concede that they are really possible, and they can claim that experimentation suffices to rule them out (as we will argue in the next chapter). Hence, Powerists are under no pressure to concede that we are unreliable or merely lucky in our inductive reasoning. That is, the Principle of Indifference cannot be used against Powerists because they do not concede that we know that there are a large number of chaotic possibilities. In fact, they can claim that we know that there are no chaotic possibilities at all. However, Powerists do share with the other theories the problem of the existence of an infinite number of possible *hypotheses,* all of which are compatible with the evidence so far collected.

So, is the situation a symmetric one? Might no account of powers and laws have a solution to the problem of justifying induction? Not really. There is a difference between the challenge faced by both Neo-Humeists and Nomists and that faced by Powerists. Powerists have no simple solution to offer that explains why inductive inference is reasonable. However, they don't face a positive argument that seems to show that, if their theory is true, induction must be unreasonable. In contrast, Neo-Humeists and Nomists *do* face such an argument, based on the Principle of Indifference, since their account of possibility entails the real existence of large numbers of chaotic worlds that are really possible and fully compatible with observed data. They have the burden of explaining how induction could be reliable under those circumstances.

## Notes

1 Tooley and Armstrong also argue that Nomism provides a solution to the *problem of induction*, the problem of justifying the inference that unobserved cases are like observed ones. We'll examine this further in the section on neo-Humeism.

2 Could the Nomist suppose that there is a higher-order law, constraining the lower-order laws to connect only *simple* properties? Perhaps, but this also comes at the price of additional complexity in the theory.

3 Neo-Humeists differ from Hume himself in three ways. First, they are concerned with metaphysical issues of fundamentality and truthmaking, not with the psychological question of what our concepts of power and law may be. Second, Neo-Humeists do not deny, as Hume occasionally did, that it is true that there are powers and real lawful connections between things. Finally, Neo-Humeists offer a new account of the distinction between laws and accidental generalizations, as we discuss in the following paragraphs.

# 6

# Powers and Properties

We turn, finally, to Strong Powerism. Strong Powerists believe that attributions of power are fundamental. Laws of nature are merely expressions of the powers possessed by various kinds of things, and counterfactual conditionals are grounded in the powers and tendencies of the entities involved in the counterfactual supposition together with their counterfactual surroundings.

As we have defined them, powers and dispositions are properties of things. It seems natural to assume that there are other properties besides powers. If so, we can ask about the relationship between those properties that are powers and those that are not.

There are two versions of Strong Powerism. One takes the truthmakers for causal laws to be universals (a "Realist" version). The second takes the truthmakers for causal laws to be the particulars that fall under the laws (a "Nominalist" version). We will explore the Realism-Nominalism controversy in more detail in the following chapter (Chapter 7). In the present chapter, we will ignore this distinction, instead focusing on how Powerism compares to other views about the laws.

## 6.1 Advantages of Strong Powerism

### 6.1.1 Causal connections and causal direction

Powerists believe in real causal connections between things, causal connections that are not reducible to the Neo-Humeist's mosaic of qualities in spacetime. Instead, Powerists can rely upon the existence of causal processes, temporally extended things that unite cause and effect into a single, undivided whole. When one thing exercises an active causal power, introducing a process of change in a patient, there exists a single process

*The Atlas of Reality: A Comprehensive Guide to Metaphysics*, First Edition.
Robert C. Koons and Timothy H. Pickavance.

that begins with the agent's active power at the time of the action and that includes the subsequent process of change in the patient (for more details on this issue, see Chapter 28). Where there is symmetric overdetermination (whether deterministic or probabilistic), the question of which potential cause is a real cause is simply the question of which potential agent is actually connected, by a real process, with the effect. This may be impossible for us to determine empirically, but there will always be a fact of the matter in the things themselves.

Similarly, Powerists can appeal to the intrinsic nature of processes to fix the direction of causation. The exercise of an active power is always found at the beginning of an appropriate process in the patient, never at the end. That is, agents with appropriate active power are always joined to a process of an appropriate kind in the patient at the beginning of that process. Which terminus of the process counts as the beginning and which the end is also fixed by the nature of the active and passive powers involved. So, for example, since fire has the power to heat water, exposure to fire will typically be found at the beginning of a process of the water's becoming hotter, that is, at the terminus of the process with the lowest water temperature. We learn whether a power is one of heating or cooling by interacting with its bearer in well-designed experiments.

### 6.1.2 Strong powerism and scientific knowledge

We have considered objections to both Neo-Humeism and Strong Nomism based on the fact that they could not explain either the rationality or the reliability of our inductive and scientific methods. Do Powerist views fare any better in this regard?

First, as we gestured toward in the last chapter and as will become clear below, the causal laws of nature turn out to be metaphysically necessary on standard Powerist accounts. In addition, these are not "brute" necessities because the causal laws are manifestations of the intrinsic natures of powers. This opens up the possibility of a purely rational, a priori component to our knowledge of the laws of nature, a component that might reliably and rationally guide us to the simplest of the empirically adequate theories (see Ellis 1999 and 2001).

Second, since the causal laws are necessary, there are no worlds in which the same properties occur but obey different laws. Hence, we have no reason to think that there are any worlds that agree with our world in the distribution of properties in the observed zone but deviate from it in the unobserved zone. Thus, we lack any argument (any 'undercutting defeater') against the rationality of induction, in contrast to Neo-Humeism and Nomism.

However, the critic could charge that there a large number of conceivable and epistemically possible *scenarios* in which things are chaotic in the unobserved zone, even if there aren't any really possible worlds like that. Doesn't the existence of such counter-inductive scenarios provide an equally good defeater to the rationality of induction?

The simple answer is, No. Strong Powerists believe that we gain knowledge about the natures and powers of things through scientific investigation. This knowledge enables us to rule out those otherwise-conceivable scenarios in which things act contrary to their actual natures. For example, before we understood the nature of water (that it is composed of $H_2O$ molecules, each with a certain quantum-mechanical structure and associated powers), we might have thought it was possible for ice to exist at high temperatures. Now, we see that that conceivable scenario is just not possible.

As Nancy Cartwright (1983, 1994), Judea Pearl (2009), and Alexander Bird (2010), among others, have noted, Powerism better accounts for our actual scientific practice than does Neo-Humeism. We don't simply consider a passively received set of observations and reason inductively. Instead, we actively isolate and manipulate things in order to better understand their causal powers and propensities. For example, in order to understand the causal powers and propensities of electrons and magnetic fields, we create carefully constructed experiments in which we expose electrons to magnetism while isolating them from other influences and interferences. We force the electrons to start with a variety of speeds and directions, and we use electrons produced by a variety of sources. We change the intensity and orientation of the magnetic field, noting how these changes affect the electrons' accelerations. These active interventions give us knowledge of natures that enable us to exclude certain conceivable scenarios from the realm of real possibility.

This is not to say that Powerism has no epistemological vulnerabilities relative to Neo-Humeism. The Powerist has to admit that our ability to observe or detect real powers through experimental interaction is fallible. There is always the possibility that we have misidentified the samples we are probing (e.g., we're trying to study water but accidentally obtain a sample of hydrogen peroxide instead) or that we have failed to identify and neutralize all interfering factors in the environment. In addition, if we are trying to measure something's propensity to cause or to undergo certain effects, we may unluckily observe a frequency that deviates from the true objective propensities. Our empirical investigations of powers can never attain certainty, and this opens the door to the skeptic, who worries that we can never rule out the possibility that we are wrong. However, the Powerist can plausibly respond that it would be unreasonable to be paralyzed by mere possibilities of error, without specific reason to suspect that we are in error in this particular case. In addition, Neo-Humeists face similar skeptical challenges, since our observations of the categorical qualities of things are also fallible.

Can the Powerist explain our preference for simple laws of nature? Is this theory vulnerable to the same objection we ran against Nomism—that it makes preference for simplicity a brute fact? The simplicity of the laws of nature would be a consequence of which properties are actually instantiated in our world. In other words, the simplicity of the laws is a function of the complexity of the essences of instantiated properties. The Powerist could perhaps appeal to God's preferences for simplicity as an explanation for the simplicity of the essences of instantiated properties. The Nomist has no such move.

## 6.2 The Individuation of Properties

It also seems natural to suppose that the connections between non-powers and powers are sometimes internal relations. That is, we might suppose that in some cases there is a property that is not a power, $F$, and a power, $P$, such that it is a matter fixed by the intrinsic characters of $F$ and $P$ that everything that has $F$ also has $P$. In such a case, we could say that $P$ is part of the 'causal profile' of $F$, or, to put it in abbreviated form, that $F$ has $P$ as one of its powers. For example, suppose that the property of *being fire* is such that, by virtue of its intrinsic character, anything that has the property of *being fire* also has the power of *heating proximate things*. If so, we could say that the property of *being fire* itself has the power of *heating proximate things*.

**Def D6.1 Conferring a Power.** A property $F$ *confers* power $P$ if and only if there is an internal relation between $F$ and $P$ of such a kind that anything that has the property $F$ also has the power $P$.

In particular, we can ask whether a property is 'individuated' by the powers it confers. Is it possible for two distinct properties $P_1$ and $P_2$ to confer exactly the same powers? If not, then the powers conferred by properties individuate them.

**Def D6.2 Property Individuation.** A property $P$ is *individuated* by the features in set $A$ if and only if $A$ is a minimal set with the following property: necessarily, any property having all of the features in the set is identical to $P$ itself. In other words, if and only if $A$ is a set having that property, and no proper subset of $A$ has that property.

**6.1T Causal Individuation of Properties (Weak Thesis).** Some fundamental properties are individuated by the set of powers they confer.

**6.1A Sicceity Theory.** No fundamental properties are individuated by the set of powers they confer.

If a property $P_1$ is not individuated by its powers, then it would be possible for there to be a second property $P_2$, such that $P_1$ and $P_2$ are distinct properties that confer exactly the same powers. If $P_1$ and $P_2$ are distinct but confer the same powers, it seems that there must be something that makes them distinct properties. At the very least, $P_1$ has the property of *being identical to* $P_1$, a property not shared by $P_2$. The property of *being identical to* $P_1$ is an example of what have been called 'thisnesses' or 'haecceities' (see Chapters 7 and 9), in this case the thisness of a property. We don't want to assume that properties are particular things, so we will use a different term than 'haecceity' in this case. Some contemporary metaphysicians, including John Hawthorne (2001) and Jonathan Schaffer (2005), have used the term 'quiddity,' but this choice is inappropriate, since the word 'quiddity' (or 'quidditas' in Latin) has a long history of use in medieval and early modern philosophy with an entirely different meaning. For this reason, we prefer 'thusness' or 'sicceity' ('sic' is Latin for 'thus' or 'so,' pronounced 'sick-say-ity'). If $P_1$ and $P_2$ are distinct properties, then the being-so that corresponds to $P_1$ is different from the being-so corresponding to $P_2$. In other words, they have different thusnesses or sicceities.

**Def D6.3 Sicceity.** A property has a *sicceity* if and only if it is not individuated by its causal powers.

What sort of thing is a sicceity (in this sense)? Consider the following thought-experiment. Suppose that there is a parallel universe exactly like the actual world, but with a different color-sensation, red*, in place of the red sensations of the actual world. That is, imagine a parallel reality in which everything is exactly like the actual world, except that the human sensations that realize the red *quale* (that are *phenomenologically* red) in this world are not red at all but instead are red*, a different color-sensation (one that no one ever experiences in the actual world). Let's suppose that in this parallel world, red* sensations have all the active and passive powers that red sensations have in our world. Thus, looking red* and looking red have all of the same environmental, behavioral, intellectual, and emotional connections. The network of powers and of causal laws in the two worlds would be exactly the same, once we have substituted the name 'red*' for 'red' and vice versa. It seems plausible that the two worlds are really different, since there are different experiences in the two worlds. If so, there must be more to the property of *phenomenal redness* than its abstract causal profile. There must be a certain thusness to red sensations that is lost when red*-ness is substituted for phenomenal redness.

Here's another, similar example. Imagine a world exactly like ours except that in this alternative reality there are two kinds of electrons, instead of just one: alpha-electrons and beta-electrons. Suppose that the two kinds of electrons have exactly the same powers as electrons have in our world. Both types repel negatively charged things, attract positively charged things, have the same quantity of mass, charge and spin, interact with each other and with other particles in exactly the way electrons do, and so on. If that is a real possibility, then the properties of *being an electron* and of *being an alpha-electron and a beta-electron* are not individuated by their causal profiles. Each would have to have some sicceity, over and above its causal powers.

If there are no such sicceities, then it would be impossible for there to be two, indiscernible kinds of electron-like particles. The resulting world would just be the actual world itself, with some electrons arbitrarily labeled as 'alphas' and others as 'betas.' Without sicceities, each property would be individuated (distinguished from all other properties) only by the relative powers it confers on things. One would be committed to Strong Causal Individuation of Properties:

**6.1T.1T Strong Causal Individuation of Properties.** All fundamental properties are individuated by the set of powers they confer.

If even one property has a sicceity, one must deny this view.

Why does this issue matter? Who cares if particles are individuated by their powers or not? There is at least one way in which our position on this issue affects another important philosophical question. If properties are individuated by their causal powers, then there are no *purely qualitative* aspects to things. Every kind of thing would be definable functionally. This would make a certain materialist strategy in the philosophy of mind, namely 'functionalism,' much more attractive. Conversely, if properties have qualitative sicceities, this would be true of the secondary qualities (like colors and smells) that we experience consciously. Dualists have argued that these secondary qualities cannot be explained in purely physical terms. If these properties have sicceities, then dualism is more plausible.

The second major issue to consider is whether a property has its powers contingently or as a matter of necessity. We'll assume that if a property has its powers as a matter of necessity, this is not a brute necessity. Instead, we should suppose that the very essence of the property includes the power. Conversely, if a property has a power contingently, this is an "accident" of the property, in the sense of something super-added to its essence.

**6.2T Essentiality of Powers (Weak Thesis).** Some fundamental property has its causal profile of necessity.

**6.2A Strong Accidentality of Powers.** No fundamental property has any of its causal profile of necessity: that is, no property confers or fails to confer any power necessarily.

**6.2T.1T Strong Essentiality of Powers.** All fundamental properties have all of their causal profile of necessity: that is, each power is either necessarily conferred or necessarily not conferred by each property.

Is there a connection between Causal Individuation of Properties and Essentiality of Powers? We think so—if a property is individuated by certain features, it seems reasonable to assume that it has those features essentially. This may not be the case for concrete, material objects. For example, it is not immediately implausible to think that ordinary material objects are individuated by their spatial locations if you think it is impossible for two of them to occupy the same place at the same time. An ordinary material object's spatial location is, then, sufficient to individuate or distinguish it from all other things. But spatial locations aren't essential to those objects, since they can move around. However, it is hard to see how an abstract object like a property could be individuated by anything other than its essential features. The only plausibly fundamental contingent features of properties are what we might call *properties of extensions*, properties had by properties in virtue of the fact that they are instantiated by a certain group of things. For example, the property of *being a dog* has the property of *being exemplified by exactly the things in the following set*: {Elsie, Fido, Crackers, Spike,...}. These sorts of properties are insufficient to individuate properties because two distinct properties might be exemplified by exactly the same things without being identical. For example, the property of *being a thing with a heart* and the property of *being a thing with a kidney* have the same properties of extension because they are exemplified by exactly the same things. Since there are no contingent properties that individuate properties, then the following principle seems to be true of properties:

**Individuation Entails Essentiality.** If a property $P$ is individuated by the features in set $A$, then $P$ has all of the features in $A$ as a matter of necessity.

The converse implication is not so plausible. One might think that properties confer powers essentially even though they are individuated, not by those powers, but by their sicceities. In fact, we will argue that such a view is a metaphysically attractive position.[1]

### 6.2.1 The causal theory of properties: causal structuralism

Sydney Shoemaker is the most prominent proponent of Causal Structuralism or the Causal Theory of Properties (1980, 1998). Causal Structuralism combines Strong Causal Individuation of Properties with Strong Essentiality of Powers. The combination of the two means that all properties are pure powers. There are no sicceities, since it is impossible for two distinct powers to have the same causal profile.

Since all properties are pure powers, this means that all powers are powers to affect other powers. We can never leave the "circle of powers." In any given world, we can abstract a directed network or graph of powers. How ought we think of this network? If $P_1$ is an active power, for example, then $P_1$ is the power to affect things with some passive power $P_2$ in such a way that those things come to have some new (active, passive or immanent) power $P_3$. From these internal facts about the powers, we could abstract the law of nature L1:

(L1) It is a law of nature that: if a $P_1$-thing comes into proximity to a $P_2$-thing, the latter becomes a $P_3$- thing.

Or, in a more abbreviated form:

(L1) $P_1(x)$ & $P_2(y)$ & $C(x,y) \Rightarrow P_3(y)$

Each power, whether active, passive or immanent, would correspond to one or more such law formulas. If we put all the laws together, we constructed the world's "law-book." From the law-book, we can abstract the causal profile of each power, by simply conjoining the laws into one long statement, and replacing the name of the power with a variable. For example, the following is part of the causal profile of $P_1$, given law L1:

(P1 Profile 1) $V_1(x)$ & $P_2(y)$ & $C(x,y) \Rightarrow P_3(y)$

All we have done is replace the name of $P_1$ with a variable, '$V_1$', in law L1. Suppose that we simultaneously replace all of the names of properties with variables in the conjoined statement of the law-book. The resulting formula will be the simultaneous causal profile of all of the world's properties. Here is an example of part of such an abstracted law-book:

(L1*) $V_1(x)$ & $V_2(y)$ & $V_3(x,y) \Rightarrow V_4(y)$

Since a property can, for Shoemaker, be identified with its causal profile, to be property $P_1$ is nothing more than to be something that can play the role of $V_1$ in the fully abstract version of the world's law-book. Ditto for all of the other properties. This type of abstract network of causal relations is all there is to the being of any of the world's properties.

AN ARGUMENT FOR ESSENTIALITY OF POWERS: THE BRANCHING-WORLDS ARGUMENT

Sydney Shoemaker has offered an argument for Essentiality of Powers from the nature of possibility. It seems reasonable to suppose that the causal profile of a property is something the property has permanently. It would be very odd if the property of *being fire* were to lose its power of *heating* or if the property of *being water* were to lose its passive power of *freezing when cooled to 0 °C*. If that is right, then the permanence of power provides us with some reason for thinking that powers are essential to their properties.

Shoemaker attempts to substantiate this argument by assuming that every possible world has 'branched off' from the actual world at some time in the past. In other words, every possible world shares an initial segment of its history with the actual world. This is what we will call an 'Aristotelian' conception of possibility, an idea that we will explore further in Chapter 15. Let's call Shoemaker's assumption the 'Branch Principle'.

**Branch Principle.** For every possible world $w$, there is a time $t$ such that $w$ and the actual world are exactly alike up until time $t$.

If we accept the Branch Principle (Branch, for short), and we accept that causal profiles are permanent features of properties, then it immediately follows that every property has its causal profile as a matter of necessity. As we shall see, there is a great deal to be said on behalf of Branch. One obvious drawback is that if determinism were true, then Branch would entail that there is only one possible world. This would make every truth necessarily true. So, the plausibility of Branch requires the world to be necessarily indeterministic.

ARGUMENTS FOR CAUSAL STRUCTURALISM Shoemaker offers two arguments in favor of Causal Structuralism: one appeals to semantics and a second to epistemology. A third argument, based on simplicity, is also relevant.

1. The semantic argument. Shoemaker argues that we could not refer to properties unless they were individuated by their causal powers. Suppose, to the contrary, that there were two distinct properties $P_1$ and $P_2$ that had exactly the same causal profiles. That is, each property causes and is caused by exactly the same properties. In such a case, we could never distinguish one from the other, since we can only distinguish properties by their effects, including their effects on our senses and our minds. If we cannot tell one from the other, how could we refer to one rather than the other?

It isn't clear how this is supposed to establish that it is impossible for two properties to share their causal profiles. It's not even clear why this should give us reason to suppose that in fact there aren't two properties with the same profile. All that it establishes, if it establishes anything at all, is that we could not refer to one rather than to the other in such a case. So what? Why should such a linguistic limitation on our part have any metaphysical implications?

In any case, it seems (as Hawthorne 2001 argues) that we could in fact refer to one rather than the other, simply by encountering one rather than the other on a single occasion and making subsequent reference back to that specific occasion. Suppose that $P_1$ and $P_2$ share the same profile, and that at noon on 10 October, I encounter an instance of $P_1$ but not of $P_2$. I could stipulate that by '$P_1$' I shall mean the property an instance of which I encountered at noon on 10 October. This stipulation will enable me to refer to $P_1$ rather than to $P_2$, even though it is true that I can never be sure, on any future occasion, whether I am then encountering a case of $P_1$ or $P_2$. In effect, Hawthorne was applying to this case Saul Kripke's arguments against the descriptive theory of names in *Naming and Necessity* (1980).

2. The epistemological argument. Shoemaker's epistemological argument is also directed toward supporting Strong Individuation of Properties by Powers.

1 Suppose that properties are not individuated by their causal profiles.
2 Then we would be unable to discriminate one property $P_1$ from its causal "twin" $P_2$.

3 This entails that we could never tell whether two things have the same property (both $P_1$ or both $P_2$) or two different properties (one $P_1$ and the other $P_2$).
4 Hence, we could never know that any two things shared the same properties.
5 But, obviously we can know in many cases that two things have the same property.
6 Therefore, Strong Causal Individuation of Properties is true.

As Hawthorne (2001) points out, the weakness in this argument occurs at step 4. Suppose, for example, that property $P_1$ occurs with great regularity in our world, while $P_2$ is extremely rare. In that case, we might well know that we are encountering another instance of $P_1$, even though we cannot tell the difference between $P_1$ and $P_2$. The trouble is that Shoemaker's argument depends upon the following principle:

**Discrimination Principle.** If we cannot discriminate between cases in which *A* is true and cases in which *B* is true, then we can never know that *A* is true.

However, Discrimination leads very quickly to radical skepticism. One cannot tell cases in which one is a brain in a vat from cases in which one is not. Yet it seems reasonable to think that one can know that one is not a brain in a vat, or at least that one can know things that imply that one is not a brain in a vat, such as the proposition that one has arms and legs. It is plausible to think that one can know that *p*, even though one cannot tell whether it is the case that *p* or the case that *q*, so long as the alternative that *q* is sufficiently weird, remote, outlandish, or improbable. If so, then Shoemaker's argument establishes at best that we rarely encounter properties with identical causal profiles, not that such a situation is metaphysically impossible.[2]

3. Appeal to simplicity. Finally, the defender of Causal Structuralism can point out that it is an attractively simple theory. It posits only a single kind of property: pure powers. There are no sicceities, and no pure non-powers or "dual-aspect" properties, properties with both a causal profile and a sicceity. If we don't need to posit additional kinds of properties, why should we?

ARGUMENTS AGAINST CAUSAL STRUCTURALISM If there are good arguments *against* Causal Structuralism, then simplicity alone can't carry the day. We will now consider three such arguments: the problem of circularity or vacuity, the problem of arbitrariness, and the problem of symmetrical systems of causal laws.

1. The problem of circularity or vacuity. Bertrand Russell objected to an early version of Causal Structuralism on the grounds that it was vacuous and viciously circular:

> There are many possible ways of turning some things hitherto regarded as "real" into mere laws concerning the other things. Obviously there must be a limit to this process, or else all things in the world will merely be each other's washing. (Russell 1927: 325)

Causal Structuralists imagine a world in which every property is a pure power, a power to bring about other pure powers or to be brought about by them, with the whole network of powers never "bottoming out" in anything other than pure powers. The whole picture seems metaphysically anemic or colorless—a world of furious activity but lacking any bottom line of qualities or forms. In particular, we might object that Causal Structuralists cannot give an adequate account of the qualitative aspects of our conscious experience, which seem to have a thusness or sicceity that consists in more than their mere powers or dispositions to produce further states.

2. The problem of arbitrariness. There is one relation that Causal Structuralists treat as having a sicceity, independent of its place in the causal/nomological network: the causal relation itself. As Hawthorne points out, it would make little if any sense to move all the way to "Hyperstructuralism" in which even causation itself is individuated by its causal profile. What sort of causal profile does causation have? If we replace even the name for the causal relation in the law-book with yet another variable, the law-book would have no content at all. But if the causal relation has a sicceity, why not concede that other properties might have sicceities as well? Causal Structuralism seems inherently unstable or unprincipled.

3. The problem of symmetrical law-books. The final objection Hawthorne (2001) raises concerns the possibility of symmetrical law-books. Suppose that the world's law-book was made up of exactly the following immanent causal laws:

$A$ = {the power, when combined with property $B$, of producing a change from $D$ to $C$; the power, when not combined with $B$, of producing a change from $C$ to $D$}
$B$ = {the power, when combined with $A$, of producing a change from $D$ to $C$; the power, when not combined with property $A$, of producing a change from $C$ to $D$}
$C$ = {the power of producing a change from $B$ to $A$; the power of producing a change from $A$ to $B$}
$D$ = {the power of producing a change from $B$ to $A$; the power of producing a change from $A$ to $B$}

The picture is this: when a body has both properties $A$ and $B$, it has the power to change another body from $D$ to $C$, but not vice versa. When a body has either property $A$ or property $B$, but not both, it has the power to change another body from $C$ to $D$, but not vice versa. The properties $C$ and $D$ have powers that treat $A$ and $B$ equivalently. In such a world, there must be a difference between the two properties (or supposed power-bundles) $A$ and $B$, since there is a real causal difference between having just one of the properties and having both of them simultaneously. However, if we try to identify the two properties with their power-bundles, we face a kind of vicious circularity. The only difference between the two bundles depends on there being a difference between $A$ and $B$. If we substitute $A$ for $B$ and $B$ for $A$, we transform the $A$-bundle into the $B$-bundle, and vice versa. Figure 6.1 illustrates the symmetry of the situation. The horizontal arrows between $A$ and $B$, and between $C$ and $D$, represent possible changes, and the diagonal and vertical arrows represent the causal powers of the five relevant combinations of properties: $C$, $D$, $(A\&B)$, $(A\&{\sim}B)$, and $({\sim}A\&B)$.

The roles of $A$ and $B$ in this law-book are perfectly symmetrical. Each can be described as one of two properties that, by itself, results in the transition from $D$ to $C$, and, in combination with the other, results in the opposite transition from $C$ to $D$. The two properties

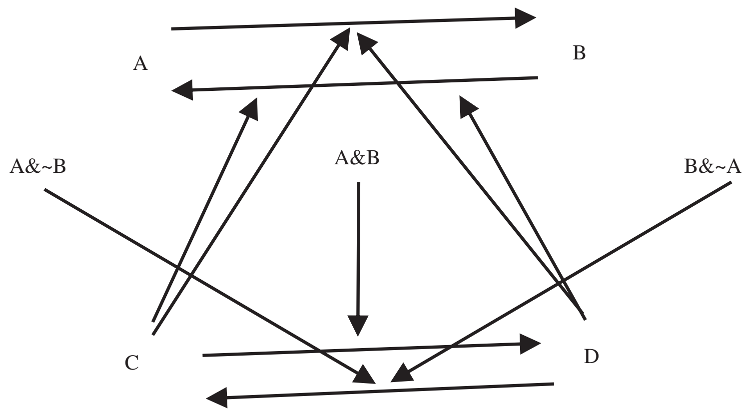


**Figure 6.1** A Symmetrical Law-Book

*A* and *B* are analogous to the two indistinguishable steel spheres in Max Black's famous thought-experiment (Black 1952) (see Section 9.3.2.1 for details). Black used his example to argue that two objects can be distinct although indiscernible. This is turn could be used in an argument for individual haecceities. In the same way, the symmetrical law-book in Figure 6.1 can be used to argue for sicceities, since the causal profiles of *A* and *B* are not sufficient to individuate one from the other.

One complication: we have to distinguish between two types of causal profile: property-relative profile and systemic profile. A *property-relative profile* is a set of powers (active and passive) that are defined in terms of the other properties with which the given property interacts. A property-relative profile assumes that we can identify the other properties, distinguishing them from each other and from the property in question. A *systemic profile* is a set of powers that are defined abstractly, in terms of positions or nodes within a network of nomic connections. The difference between the two kinds of profiles emerges in cases of symmetrical law-books, as in Figure 6.1. The property-relative profile of property *A* is that of causing the *D*-to-*C* transition (when by itself), and causing the *C*-to-*D* transition when jointly instantiated with *B*. Property *B* has a different property-relative profile: it has the profile of causing the *D*-to-*C* transition (when by itself) and causing the *C*-to-*D* transition when jointly instantiated with *A* (not with *B*).

In contrast, the two properties have exactly the same *systemic profile* in this case. Both have the profile of being the first of four properties $P_1$, $P_2$, $P_3$, and $P_4$, of causing the transition from $P_4$ to $P_3$ when instantiated by itself, and of causing the opposite transition from $P_4$ to $P_3$ when instantiated with another of the four, $P_2$. In the case of the profile of property *A*, the role of $P_2$ is played by *B*, $P_3$ by *D*, and $P_4$ by *D*, while in the case of property *B*, the role of $P_2$ is played by *A*. Nonetheless, the systemic profiles are themselves exactly the same.

Thus, we can distinguish two forms of Causal Structuralism: a strong form in which properties are individuated by their systemic causal profiles, and a modest form in which properties are individuated by their property-relative causal profiles. The modest form is really a version of Dual-Aspect Theory (discussed in the next section), since we need something like sicceities to distinguish between two property-relative causal profiles (see

Hawthorne 2001: 376). If *A* and *B* do not have different sicceities, what distinguishes the property-relative profile of *A* in this case from that of *B* if not the sicceities of *A* and *B*? If we cannot account for the real distinctness of the two property-relative profiles, we cannot account for the real distinctness of *A* and *B*.

Therefore, it seems that there must be something in addition to the powers themselves (a difference in qualitative characters or sicceities) that accounts (in a non-circular way) for the difference between *A* and *B*. Only by accounting for the difference between *A* and *B* in this way can we account for the difference (in such a world) between the power of instantiating both properties and the power of instantiating just one of them.

### 6.2.2 Dual-aspect vs. mixed two-category theories

C.B. Martin has articulated and defended what the Dual-Aspect Theory of properties (Martin 1994, Armstrong, Martin, and Place 1996). On this view, every property has two aspects, its causal profile or powers, and its sicceity. Since properties have sicceities, they cannot always be individuated by their powers. It is possible for two distinct properties to have the same causal profile. At the same time, on Martin's view, every property has at least one power and has that power (and all of its other fundamental powers) essentially.

To make this idea concrete, return to the two thought-experiments from the last section: swapping phenomenal red for red*, swapping one type of electron for two. In each case, Dual-Aspect Theorists will conclude that there is a real difference. Even if all the red-lookings became red*-lookings, and had every other causal connection to things that red-lookings actually have, they would still be phenomenally red and not red*, by virtue of having the sicceity of phenomenal redness and not that of red*-ness. Similarly, there is something about the property of *being an alpha-electron* that makes it intrinsically and essentially different from the property of *being a beta-electron*, even if the causal profiles of the two properties are indistinguishable. If, in the next moment, alpha- and beta-electrons were to swap places, the world would change dramatically, even if we couldn't tell the difference.

Like Dual-Aspect Theory, the Mixed Two-Category Theory posits the existence of dual-aspect properties. In addition, it proposes that there are pure non-powers, properties which have their causal profiles contingently. The most plausible examples of pure non-powers are spatial and temporal properties and relations, such as size, shape, distance, and duration. If there are other pure non-powers, we can call them 'pure qualities'.

**Def D6.4 Pure Quality.** A *pure quality* is a fundamental, non-spatiotemporal property that has its fundamental powers contingently.

Mixed Two-Category Theory requires either that spatial and temporal relations have their fundamental powers contingently or that there are pure qualities or both.

It seems we can imagine a perfectly inert world, in which material bodies occupy places at times, even though no material body has any active or passive power. Let's suppose that each body has as its only immanent power the power of inertia—the power to keep moving in a constant velocity. We can imagine that the bodies are made of a ghostly matter, so they are able to pass through each other without disturbance. In such a world, it would seem that the spatial and temporal relations would exist, even though they had no power of any kind, whether active, passive or immanent. If so, then all spatial and temporal relations must have all of their fundamental powers, if any, contingently.

Do spatial and temporal properties have any powers at all, even in the actual world? They seem merely to provide occasions for the exercise of the powers of other things, by providing instances of contact and proximity needed to trigger the active and passive powers of body or instances of temporal duration needed for the exercise of immanent powers. Space and time in and of themselves seem to be causally inert, never causing or preventing anything by themselves. If this is right, then the Mixed Two-Category Theory would seem to be correct.

## 6.3 Objections to Strong Powerism

The arguments for Powerism have already been presented, in the form of objections to its competitors. Let's turn to some objections to Powerism itself.

1. Implausibility of necessary laws of nature. Both Dual-Aspect Theory and the Mixed Two-Category version of Powerism entail that most of the laws of nature—all of those laws that describe the fundamental powers of fundamental kinds of things—are metaphysically necessary, since they are entailed by the essences of the kind-properties involved. This entails further that all of the derived laws that are logical and mathematical consequences of the fundamental laws are also metaphysically necessary. For example, it would turn out to be necessary that water freezes at 0 °C.

We could certainly imagine a world in which water freezes at some other temperature, a world, for example, in which water freezes at the same temperature and pressure as mercury. Since imagination is a guide to possibility (**PEpist 1**), this gives us good reason to believe that such a world is possible. The best versions of Powerism must treat these reasons as misleading in this case. This is at least a theoretical cost.

2. Determinism and recombination (Patchwork Principles). Jonathan Schaffer (2005) has argued against the thesis that the laws of nature are metaphysically necessary, on the grounds that necessary deterministic laws would drastically undermine our knowledge of alternative possibilities. If the laws of nature are both metaphysically necessary and deterministic, then there is only one possible world that begins as our world does, namely, the actual world. If the laws of nature are deterministic, then there is no possible world that agrees with our world up to some point of time and then diverges thereafter, since this would involve a violation of the world's laws, but those laws are supposed to hold in every possible world.

Worlds in which things go differently from the actual world would have to be different from the very beginning. Worlds that look like ours at the beginning but diverge would have to be populated partly or wholly by things with *alien properties*, properties that are never instantiated at all in the actual world, because only such alien properties could obey laws of nature other than the ones we are familiar with.

Such a picture of possibility would deprive us of the method of *recombination*: the ability to construct alternative possibilities by using mental "scissors and paste" to take apart pieces of the actual world and recombine them in new ways. Yet, this seems to be exactly how we go about forming our opinions about what is really possible. This way of forming our modal opinions can be captured by various *Patchwork Principles* (Principle of Metaphysics 5), which say (roughly) that recombined possibilities correspond to further possibilities. Without recombination and the associated Patchwork Principles, we would never be in a position to know anything about how things might have gone but didn't.[3] We examine Patchwork Principles in much more detail in Chapter 16.

The best response for Powerists is to deny that the laws of nature are, or even could be, deterministic. Every law of nature must ascribe certain tendencies and propensities to things, but it must never prescribe exactly one possible outcome for each interaction. This is another cost to the theory, since we seem to be able to conceive of worlds with deterministic laws, and some philosophers and scientists believe that the actual world is deterministic.

3. The mystery of "natural intentionality." As George Molnar has pointed out (2003), the thesis that powers are fundamental properties means that a kind of *intentionality*, similar to that displayed by human beliefs and intentions, is built into the very fabric of reality. A property is *intentional* if it essentially involves some relation to another thing or to a situation, without entailing that that other thing or situation be actual. For example, I can imagine a golden mountain or think about a round square, even though no such thing exists. Similarly, a power to attract negatively charged particles could exist, even if there were in fact no negatively charged particles to attract. Opponents of Powerism object to such properties as fundamental, insisting that intentionality must somehow be reducible to non-intentional facts.

However, this objection suffers from the fact that no one has yet succeeded in reducing intentionality to non-intentional facts, despite many attempts to do so. Since intentionality is an undeniable reality, taking natural intentionality as fundamental seems a live option.

In addition, the mysteriousness of the intentionality of powers can be dispelled to some extent if we embrace Possibilism (12.1A.1T), the thesis that some merely possible things (in particular, merely possible future events) do not actually exist. A powerful property could be an internal relation that ties its bearer to merely possible future events. Suppose, for instance, that powerful particulars are always involved in a process of some kind and that processes involve relations among things and events, both actual and potential. Let's consider a single electron that is in the process of motion through space, giving rise to a magnetic field with the power to exert force on other moving, charged particles. The electron's magnetic power is an aspect of its being in motion, a paradigmatic case of a *process*. This process of motion ties the electron to certain actual past events, events involving the past location of the particle and its actual effects on other particles in its neighborhood. The process also ties the electron to future, merely potential events, events involving both certain future locations of the electron and events involving the forceful acceleration of particles in its vicinity. Some of these potential events may never actually occur. In fact, if the electron's process of motion were immediately brought to an end by encountering some new force or by the electron's annihilation, all of the process's future events might be prevented from becoming actual. For example, it can be true that an armadillo was in the process of crossing the road when it was crushed by a passing truck.

The process was one of crossing the road because it related the armadillo to a future event of its reaching the other side of the road, but the passing truck ensures that that future event is never actualized (for a further discussion of processes and causation, see Chapter 28).

4. Pairing active and passive powers. Another possible objection to Powerism concerns the coincidence of finding matching pairs of active and passive powers. Why, if *F*'s have the active power to make *G*'s *H*, must *G*'s have the passive power to *H* when acted on by *F*'s? The simplest account would be one in which either active powers alone are fundamental (and all passive powers derived) or all passive powers are fundamental (and all active powers derived). But neither answer seems more plausible than the other, and it can seem implausible to take these coincidences as everywhere brute, unexplained necessities (PMeth 1.2).

The best answer to this problem involves carefully distinguishing between a causal power and our descriptions of it, or, to put it more precisely, between those descriptions that capture the true essence of the power, considered in isolation, and those descriptions that tacitly take into account the existence of other powers.

Here's a suggestion about how this disentangling might go. Let's suppose that water has the passive power of being capable of being heated, period. This fact about water doesn't entail that fire, in particular, should have the active powers of heating water, but only that there could be properties that confer such an active power. At the same time, fire might have the active causal power of heating anything with the passive power of being capable of being heated, or perhaps, of heating anything with this passive power together with certain other characteristics. Fire could have this active power, regardless of whether water, in particular, has the passive power of being capable of being heated. The two facts together entail that fire can heat water: that is, these two fundamental facts entail the two derived facts that fire has the active power of heating water, and water has the passive power of being heated by fire. No circularity or unexplained coincidence is required.

5. Negative causation. We've saved the most difficult problem for the last. We often speak about an event's being caused by an absence. In fact, as Jonathan Schaffer (2004) has pointed out, such negative causation is extremely common. Every time one fires a gun, for example, there is a case of negative causation: pulling the trigger removes an obstacle between the firing pin and the bullet. It is the absence of the obstacle that causes the gunpowder to be ignited. Every time a nerve signal passes through the brain, a case of negative causation is involved: it is the absence of certain chemicals in the synapse between two neurons that causes an electrical impulse to be transmitted. Many cases of death (perhaps all) involve negative causation: the absence of oxygen in the bloodstream to the brain.

We can also speak of absences as effects. The causation of an absence is a case of prevention. We can prevent an explosion by cooling a chemical compound, thereby causing the absence of an explosion. One can prevent a stampede by calming a jittery steer or prevent an accident by swerving one's car to the right.

How is this a problem for Powerists? Powerism requires that we attribute causal powers to absences. But only *things* can have powers (whether active or passive), and absences aren't things!

Here's another way to put the point: Powerists conceive of causation as a real connection between things. If absences are nothing, then they cannot be connected to other events, either as causes or as effects. We will take up this problem again in Section 27.2, including some strategies by which Strong Powerists might meet this challenge.

## 6.4 Conclusion

Each of the four positions we have examined—Hypotheticalism, Powerism, Nomism, and Neo-Humeism—has its advantages and disadvantages. Neo-Humeism has the leanest metaphysical theory, since its class of fundamental truths includes no conditionals, laws of nature or attributions of power. In addition, the Ramsey/Lewis account of natural laws provides a simple explanation for our preference for simple theories about the laws. However, it faces some serious challenges, both metaphysical and epistemological. The cheif metaphysical challenge is that Neo-Humiesm denies that causal powers are intrinsic to their bearers and consequently forces us to claim that certain conceivable small worlds are impossible. The chief epistemological problem is that Neo-Humeism has to treat induction as a primitive postulate of reason, despite the apparent abundance of chaotic possibilities.

Hypotheticalism has much to be said for it, given the possibility of explaining laws in terms of conditionals. However, the problems of finks, masks and mimics suggest that we cannot reduce powers or dispositions to conditional truths. As we shall see in Chapter 27, there are also difficulties in reducing causation to conditional facts. Finally, Hypotheticalists seem to be committed to the reality of merely possible things, or to the existence of haecceities, either of which would have to be considered an ontological cost to the theory.

Nomism has a good chance of grounding the truth of most counterfactual conditionals, so long as we can add to it an account of temporal or causal priority. It does, however, face some serious epistemological challenges, since it counts the laws of nature as contingent and so must (along with Neo-Humeism) acknowledge the existence of an abundance of chaotic possibilities that mimic any apparently orderly set of observations. In addition, unlike Neo-Humeists, Nomists have no explanation of our preference for simple laws. Metaphysically, Nomism is burdened with a large number of brute necessities, linking laws with actual patterns of particular fact. In the case of functional laws (like Newton's laws of motion), Nomists are forced to posit an infinite number of laws, one for each $n$-tuple of values, many of which are never instantiated in the actual world.

Consequently, Powerism, the position that would seem to be closest to our commonsense view of the world, is in a relatively strong position, despite the burden of positing both natural necessities between separate things and natural intentionality. Of the three versions, Causal Structuralism, Dual-Aspect Theory, and Mixed Two-Category Theory, the possibility of symmetrical law-books provides strong evidence against the former.

### Notes

1 In recent work, Nora Berenstain (Berenstain forthcoming) has argued that a structuralist should take into account more than just the causal role of properties. It is also important to

look at the broader nomological role that properties play, including non-causal laws of nature (such as laws about the curvature of spacetime in general relativity, or the Pauli exclusion principle in quantum mechanics).

2 See also Schaffer (2005), who shows how each of the available anti-skeptical strategies can be adopted by the defender of sicceities (what he calls 'quiddities').

3 In addition, the impossibility of such divergences would deprive us of the plausible account of the semantics of counterfactual conditionals offered by David Lewis (1979b).

# Part III

# Universals and Particulars

# 7

# Universals

## 7.1 Introduction

As has become clear in the foregoing, the world seems to be made up of both things and the ways those things are. There are things, and these things have a certain character. There are particular things, and there are the properties those things have and relations those things bear to one another. On the one hand are trees, clouds, people, fundamental particles and fields, and possibly regions or points of space. On the other hand are shapes, sizes, colors, charge, mass, and various mental states, both occurrent (*being in pain*) and habitual (*understanding geometry*). We can distinguish between *particulars* and *properties* (or attributes, features, ways), with particulars *instantiating* or *exemplifying* properties or we can speak of properties as *predicated of* or *characterizing* particulars.

However, there is substantial controversy about the nature of both particulars and properties. (The controversy about the nature of properties is sometimes called "the problem of universals" or "the problem of the one over many".) Some philosophers, for example, think that the categories of *particular* and *property* are fundamental, that at least some of the things in both are in no way derived from or dependent on things in another category. These philosophers are *Realists* about both particulars and properties. *Nominalists* (and *conceptualists*), on the other hand, think of particulars as fundamental and of properties as non-fundamental, with the latter being derived from the former. This derivation has been cashed in a number of ways, and we will consider some of those ways below. Bundle Realists think the dependence runs exactly the other way: they agree with Realists against Nominalists that (at least some) properties are fundamental, but disagree with both Realists and Nominalists in thinking that particulars are derived from properties. Particulars, on this view, are simply "bundles" of properties. In this chapter and the next, we will explore this landscape in more detail. We will start this exploration by thinking about why someone might go in for Realism about properties.

*The Atlas of Reality: A Comprehensive Guide to Metaphysics*, First Edition.
Robert C. Koons and Timothy H. Pickavance.

But before doing that, we want to note two bits of near-universal agreement. First, very many properties, maybe even all of them, are *shareable*: more than one thing can exemplify one and the same property. For example, many particulars exemplify the property of *being red*, many exemplify the property of *being a dog*, and so on. Though there is controversy about the exact nature of this sharing, that there *is* sharing is so obvious it almost needn't be mentioned!

Second, there is some connection between something's satisfying a linguistic predicate and its exemplifying a property (or some properties). For example, one of us has a dog, a dachshund in fact, named 'Elsie'. We can say a number of true things about Elsie: Elsie is a dog; Elsie is a dachshund; Elsie is sweet; and so on. Thus, Elsie satisfies the predicates 'is a dog', 'is a dachshund', and 'is sweet', among a great many others. More generally, something *x* satisfies a predicate *P* if and only if when one combines a name or specifying description for *x* with *P* in a grammatically appropriate way, one produces a true sentence.[1] For example, 'Elsie' is a name for Elsie, and when one combines 'Elsie' with 'is a dog' in a grammatically appropriate way, one produces the sentence, 'Elsie is a dog', which is true. Similarly, if one combines the specifying description 'the only pet of THP' with 'is a dachshund' in a grammatically appropriate way, one produces the true sentence, 'The only pet of THP is a dachshund.' A simple-minded account of the truth of these sentences is just this: the property of *being a dachshund* is correlated somehow with the predicate 'is a dachshund', sort of like how Elsie is correlated somehow with the name 'Elsie', and thus the sentence 'Elsie is a dachshund' is true if and only if Elsie exemplifies the property of *being a dachshund*. More generally, the sentence formed by combining a name for *x* with the predicate *P* in a grammatically appropriate way is true if and only if *x* exemplifies the property correlated with *P*. (We'll consider a complication that arises from this general strategy in Section 7.2.3.)

### 7.1.1 What properties must explain

We have already hinted at the fundamental driving force behind the development of a theory of properties. The primary datum a theory of properties must explain is the simple, undeniable fact that many things are similar to one another in certain respects. There are many such respects of similarity, and things can be similar in some respects but not in others. Consider three things, the first of which is red and square, the second of which is red and circular, and the third of which is blue and circular. Let us call these 'RedSquare', 'RedCircle', and 'BlueCircle', respectively. RedSquare and RedCircle are exactly similar in color, which is to say they are both red. They are, however, dissimilar in shape, since RedSquare is square and RedCircle is circular. RedCircle and BlueCircle, on the other hand, are exactly similar in shape, since they are both circular. They are dissimilar in color, since RedCircle is red and BlueCircle is blue. RedSquare and BlueCircle, on the other hand, are dissimilar both in color and shape, which is to say that they do not share a color or a shape property.

What we see here is, first, similarity or resemblance in certain respects (e.g., color, shape, etc.) and, second, sharing of properties (e.g., *redness*, *circularity*, etc.). We might also say that, in typical cases, when there is a respect of similarity or resemblance between or among things, there will be a sharing of properties, and when there is dissimilarity or a lack of similarity or resemblance there will be a difference of properties. For it is plausible to think that RedSquare and RedCircle are similar in color *because* they are share a property, namely the property of *being red* (or *redness*). Likewise, RedCircle and BlueCircle are similar in shape *because* they share the property of *being circular* (or *circularity*). RedSquare and BlueCircle, on the other hand, are wholly dissimilar because they share no properties. These facts—facts of similarity or resemblance—are the primary bits of data a theory of properties must explain. We will focus our attention on how the theories of properties we examine explain this data.

We do not, however, want to give the impression that facts of similarity are the only bits of data a theory of properties must explain. For example, a theory of properties must also account for the relations among properties, so-called *categorial relations* (*categorial* in that they have to do with categories). For example, the property of *being red* and the property of *being colored* are related in that being red is a way of being colored. The property of *being colored*, we will say, is a *determinable* of the property of *being red*, and the latter is a *determinate* of the former. Many properties stand in these determinable-determinate relations: the property of *being an animal* is a determinable of the property of *being a dog*, which is in turn a determinable of the property of *being a dachshund*, which is in turn a determinable of the property of *being a wire-haired dachshund*. (We can turn the order around and talk instead in terms of determinates.) A different type of categorial relation obtains between other properties. Take the property of *being red* and the property of *being blue*; these are related in that nothing can have both at once (that is, nothing can be red and blue all over). We might say that being red *precludes* being blue. It's not difficult to identify a number of these relations of preclusion. The property of *being a dog* precludes that of *being a cat*, the property of *being a circle* precludes that of *being a square*, and so on. An interesting feature of these categorial relations is that they seem to be *necessary*. Given just the existence of the properties themselves, they simply *must* stand in these categorial relations. It just couldn't be that something could be red and blue all over at once. It couldn't be that there was something that is both a cat and a dog. And so on. It is important to consider how it is that various theories of properties fare explaining the *necessity* of these categorial relations (whatever one's view of the *nature* of the relations happens to be). Unfortunately, we don't here have space to consider these issues in more detail, but we suggest that readers take them up on their own.

Closely related to these categorial relations are predications of other properties and relations to properties. For instance, it seems clear that the property of *being red* is more similar to that of *being orange* than it is to that of *being blue*. Similarly, the property of *having 1 kg of mass* is closer to *having 2 kg of mass* than it is to *having 10 kg of mass*. These relations of relative similarity and difference among properties suggest that there must be properties to serve as the relata of these relations.

Finally, there are two other phenomena that properties have been employed to explain: the possibility of intentionality in both language and thought, and the distribution of causal powers and dispositions. Intentionality concerns the aboutness of language and thought: how it is, when we predicate some property of some entity, that our thought succeeds in being about that property. If I predicate wisdom of Socrates, what is the connection between my act of speech or my thought and Socrates' wisdom or possible wisdom (as opposed, say, to his snub-nosedness or his musical ability)? One popular theory, with roots in both Plato and Aristotle, would explain this fact by making a property (like the property of wisdom) an actual component of a thought. We'll take this idea up in Section 7.3.

We have already discussed the nature of causal powers in Chapters 4–6. We'll look at the implications of these chapters on our theory of properties in Section 7.4.

## 7.2 Realism

The first view, or cluster of views, about the nature of properties we will consider is Realism. Realism has two characteristic features. First, Realist views maintain that (at least some) properties ground the character of ordinary objects like rocks and tables and dogs and people. Second, Realist views maintain that properties are universals. Any view with both of these characteristic features is Realist.

**7.1T Realism.** Universals exist and ground the character of ordinary objects.

What, exactly, is it to think that properties *ground* the character of ordinary objects? Essentially, it is that ordinary objects have the character they do *in virtue of* the properties they exemplify; it is because ordinary objects exemplify properties that they have the features, the character, that they do. This is what we called extra-conceptual grounding in Chapter 3.4: a grounding relation between entities or facts in the world.

And what, exactly, is it to think that properties are *universals*? Essentially, it is that properties (of the shareable variety discussed above) are just as fundamental as the things that have them, in the sense that they are not derived from or dependent on those things. This second characteristic feature of Realism could use some clarification, for it is common to define universals as properties that can be exemplified by more than one thing. We find this characterization inadequate to fully account for the differences between believing in universals and not. To (briefly!) see that this is so, suppose that one identifies properties with classes. Two versions of Nominalism, Class and Resemblance Nominalism (see below and Chapter 8), make just this identification between properties and classes.

**7.1A Nominalism.** Universals do not ground the character of ordinary objects.

If one insists that universals are just properties that could be exemplified by more than one thing, Class and Resemblance Nominalism turn out to be committed to universals. Properties are classes, exemplification is class membership, and some property-classes will have more than one object in them. Thus, some properties are capable of being exemplified by more than one object, and that's just to say that there are universals. This would be fine except that Nominalism is often taken to be the view that there are no universals! Even if one rejects that particular definition of Nominalism (as we do), everyone agrees that universals can't be classes of objects. Defining universals as multiply exemplifiable properties leaves open the possibility that universals are just classes. Therefore, that definition must be rejected or supplemented. Second, starting with the thought that universals must be multiply exemplifiable forecloses a certain picture of haecceities (see

Chapter 9), properties that only one possible object could exemplify. In particular, it precludes taking haecceities to be universals. We believe (though will not here defend) that one ought not foreclose this possibility at the outset. To repair these problems, we believe one ought to think of universals as *fundamental character-grounders*.

We thus arrive at the task of explaining this notion of *fundamentality*. The easiest way to appreciate what we mean by the talk of the fundamentality of properties is to contrast Realism with Class Nominalism. Class Nominalism, again, is the view that all properties—the things that are (sometimes) shared between and among things—are just classes of the things that have the property. The property of *being a dog*, on this view, is just the class of all dogs. The property of *being red* is just the class of all red things. The property of *being identical to Elsie* is just the class that includes only THP's dachshund, Elsie. And the property of *being a color* is just the class of all color properties (which would be a class of classes!). Class Nominalism, then, maintains that properties are less fundamental than the things that exemplify them, since sets are less fundamental than (at least partially grounded in) their members. There is a dependence of classes on their members that is not reciprocated. Realists deny that properties are less fundamental than other objects in this, or any other, way.

It is fairly clear that Realists accept the need to explain both character and similarity, and that they offer an explanation something like the one outlined above. The idea is that universals, in paradigm cases, ground the character of multiple objects, and in that way account for similarities among those objects. For example, the universal BLACK is exemplified by a number of things, including THP's phone, his car key fob, and the keys on his computer's keyboard. These things, according to Realism, are similar because they exemplify one and the same universal. The universal BLACK that is exemplified by THP's phone is identical to the universal BLACK that is exemplified by the "Y" key on THP's keyboard. The similarity between these two things is to be explained by the fact that each exemplifies *the very same property*. Further, the reason why the exemplifying of the very same property makes for similarity, according to Realism, is that shareable properties, like the universal BLACK, ground character. THP's phone is black because it exemplifies the universal BLACK; likewise, the "Y" key on THP's keyboard is black because it exemplifies the universal BLACK. According to Realism, universals ground character, and thereby ground similarity. The picture the Realist paints appears promising, at least when it comes to the central phenomena properties are meant to explain.

### 7.2.1 Troubles for Realism

Regardless what variety of Realism one opts for, one must face the potential trouble spots for Realism in general. We are in the game of constructing a theory to explain a certain data set, and the process of choosing a theory is more than simply considering whether a view can explain the data. Many views can explain any set of data! We must consider a theory's strengths and its weaknesses in order to evaluate whether it *best* explains the relevant data. So we turn now to arguments that have been raised against Realism.

*7.2.1.1 The Universal-Particular Distinction* The first challenge a Realist faces can be put in the form of a question: What, exactly, is the distinction between universals and particulars? What is it, in other words, for something to be a universal rather than a particular (and vice versa)? Realists must commit to the following principle:

**6.2T Universal-Particular Distinction.** There is a clear and coherent distinction between universals and particulars.

The challenge is to find a way to uphold this principle.

There are three popular candidates for making the universal-particular the distinction.

**7.2T.1 Aristotelian UP.** Universals can be *predicated* of other things, while particulars cannot be predicated of anything.[2]

**7.2T.2 Russellian UP.** Particulars are necessarily located in only *one place* at a time, while universals can be wholly present at *many places* at once.[3]

**7.2T.3 Wise UP.** Any particular can coexist with another particular indiscernible from it, but it is impossible for any universal to be indiscernible from anything else.[4]

There are some well-known problems with the Aristotelian UP and Russellian UP distinctions.

Take Aristotelian UP first. Advocates of Aristotelian UP tend to start with the observation that predicates in natural languages "correspond" to universals, in the sense discussed at the beginning of this chapter. In at least a certain class of cases, something satisfies a predicate if and only if it exemplifies the universal that corresponds to that predicate. Aristotelian UP is an attempt to turn this mundane observation into an account of the universal-particular distinction. The thought goes like this: predicates in language are attached to subjects in a language via concatenation, and in an analogous way universals are "attached" to particulars via (so-called "ontic") predication.

Though there are a number of ways to try to execute this strategy, Frank Ramsey (1925) powerfully argued that this notion of (ontic) *predication* cannot be made sense of without presupposing a distinction between universals and particulars. Therefore, one cannot use predication as a basis for *making* that distinction. Instead, the distinction must be the basis for understanding predication. And even if one could understand predication independently of the universal-particular distinction, Ramsey argued that one would not be able to define the *direction* of predication. The only reason one has for thinking that it is universals, rather than particulars, that are predicated is an arbitrary linguistic distinction between subjects and (natural language) predicates. But natural languages needn't have had this sort of structure: we can develop languages such that the bits of language that correspond to particulars are the predicates, whereas the bits of language that correspond to universals are the subjects. In other words, Ramsey argued that the linguistic evidence is simply insufficient to support a robust universal-particular distinction. But that linguistic evidence is all the advocate of Aristotelian UP has to draw on.

Consider now Russellian UP. It is not obvious that particulars cannot be multi-located. For example, if something is transported into its own past, it could occur twice over at the destination time. In addition, fission might enable one particular to exist in two places at once. But if it is possible that particulars be multiply located, then Russellian UP is false, since being possibly multiply-located would not be sufficient for being a universal. Further, if there are universals that only one thing can have, then those universals cannot be located in more than one place at once (unless particulars can, in which case there still is not the needed asymmetry). In this case, being possibly multiply-located would not be necessary for being a universal.

We'll consider Wise UP in more detail. If Wise UP is correct, then particulars and universals must satisfy the following definitions, respectively:

**Def D7.1 Particular.** Something is a *particular* if and only if it is possible for there to be two distinct but indiscernible duplicates of it.

**Def D7.2 Universal.** Something is a *universal* if and only if it is not possible for there to be two distinct but indiscernible duplicates of it.

One note before we begin our examination of these definitions in earnest: given that everything either is or is not such that it is possible that there exist two distinct but indiscernible duplicates of it, these definitions demand that everything that exists is either a universal or a particular. That is, there is nothing that is neither universal nor particular. It is an important question whether this is a reasonable assumption, but it is not a question we will explore here.

Before we are able to ask whether these definitions are *true*, we must first understand their content, what they *say*. There is one issue, in particular, that will demand sustained attention, namely, the notion of *indiscernibility*. The idea is that two things are *indiscernible* if and only if they are qualitatively the same. As we have noted, things can be alike in various ways. This happens when things have the same properties. Things can also fail to be alike, by having different properties.

We can thus begin a more careful examination with a simple-minded definition:

**Def D7.3 Indiscernibility.** Things *a* and *b* are *indiscernible* if and only if *a* and *b* have the same properties.

Given Def D7.3, we can understand Wise UP in another way. Gottfried Leibniz argued for two principles of identity:

**Indiscernibility of Identicals.** If thing *a* is identical to thing *b*, then *a* exemplifies *F* if and only *b* exemplifies *F*.

**7.3T Identity of Indiscernibles.** If (thing *a* exemplifies *F* if and only if thing *b* exemplifies *F*), then *a* is identical to *b*.

The Indiscernibility of Identicals is almost universally accepted and has come to be known as 'Leibniz's Law'. The Identity of Indiscernibles, on the other hand, has been much more controversial. Indeed, Wise UP amounts to the claim that some things, the universals, obey the Identity of Indiscernibles, while other things, the particulars, violate the Identity of Indiscernibles.

Def D7.3 turns out to be inadequate, but the reason for its inadequacy is instructive. What we are trying to capture with the notion of *indiscernibility* is what we might call "purely qualitative sameness". We've talked above about objective similarities and are thinking about properties as something that might explain those similarities. But we also noticed a connection between properties and predicates. Connecting these to our discussion of indiscernibility, though, there seem to be predicates that won't correspond to a property that makes for objective similarity (if they correspond to a property at all—see the discussion of Russell's Paradox in Section 7.2.1.2).

To see how this is so, consider the following:

(1) Lyle is standing next to a tree.
(2) Elsie is standing next to a tree.

Suppose both (1) and (2) are true. If so, then both Lyle and Elsie satisfy the predicate 'is standing next to a tree'. If there is a property of *standing next to a tree*, then Lyle and Elsie both exemplify it. But it seems quite strange to think that Lyle and Elsie objectively resemble or are objectively similar to one another just in virtue of standing next to a tree! Maybe they resemble one another in being mammals or in being alive or in being creatures or maybe even in that they are both standing. But their both standing next to a tree doesn't seem to be a sense in which they resemble each other.

This apparent lack of resemblance can be emphasized if we consider a counterfactual case. Suppose Lyle walked away from the tree and stood instead next to a tractor. It is difficult to see how this would change the resemblance relations between Lyle and Elsie. If Lyle resembled Elsie while standing next to a tree, it doesn't seem that he would resemble her less just by moving next to a tractor.

Returning to our discussion of indiscernibility, there are properties like that of *standing next to a tree* that seem to make it too easy for objects to be discernible, according to Def 7.3. Consider the property of *being located at place p*, for some definite *p*. (Let an appropriate substitution for *p* be a set of GPS coordinates.) This property, like that of *standing next to a tree*, doesn't seem to be the sort of property that ought to matter for whether things are indiscernible or not. But most things occupy different locations. If these location properties are relevant to indiscernibility, then we can use them to argue that things are discernible. But that seems wrong! A thing's location doesn't seem to be among its deep, qualitative characteristics.[5]

There are other properties that seem to make things even worse. Take the property of *being distinct from Elsie* (which anything that satisfies the predicate 'is distinct from Elsie' has). Everything but Elsie has this property. But it doesn't seem that things ought to resemble just in virtue of sharing this property. It doesn't seem like a qualitative characteristic of a thing. And if properties like this, "properties of distinctness" we might call them, are relevant to discernibility, then it seems all too easy to show that two things are discernible. Suppose you have two things, Lyle and Elsie, for example. Given that Lyle and Elsie are *two*—that is, given they are distinct—they will exemplify different properties of distinctness. Lyle has the property of *being distinct from Elsie*, which Elsie lacks. And Elsie has the property of *being distinct from Lyle*, which Lyle lacks. Thus Lyle and Elsie are discernible (or fail to be indiscernible), by the lights of Def 7.3. This seems like the wrong way to argue that two things are discernible!

What we need is to identify a restricted class of properties, a set of *sparse* properties (and relations), that are relevant to matters of discernibility. Notice that we have employed words like 'deep' and 'qualitative' in our discussion so far, and what we're looking for is a way to think and speak more clearly about these notions. David Lewis (1983, 1986a) described what he called "natural" properties, and these seem well suited to play the role we're hoping to occupy here.[6] Natural properties are properties that are implicated in an account of the fundamental features of things in the world. According to Lewis, natural properties ground either resemblance or causal powers (or both). We believe that properties implicated in an account of the fundamental features of things will sometimes make for resemblance or causal powers, but we do not want to unduly limit our attention to just those sorts of properties. It may be that there are natural properties that do other things as well.

It isn't easy to say exactly which properties are natural. One would like to have a list of the natural properties at the outset, but there is no such list. We will content ourselves to give a few examples of properties that are plausibly natural, to try to make some general remarks, and to consider each unclear case on its own merits. First, let's provide a few examples. The properties ascribed to very small physical objects are plausibly natural, like properties of spin and charge, for example. Biological kind properties like *being a dog*, *being a human*, *being a mosquito*, and so on, are also plausibly natural. So too simple mental properties like *being in pain* or *having an itch*. Properties of mass are plausibly natural, and to a lesser extent so are color and shape properties. On the relation side, relations of spatiotemporal distance are likely natural, as are simple logical and mathematical relations like *entails* and *is greater than*. Second, then, let's look at some tentative general observations, based on our earlier discussion. Logically complex properties are plausibly non-natural. Relational properties, including properties of *identity* and *distinctness*, are plausibly non-natural as well (though we'll consider a possible exception below).

With that basic understanding of natural properties on the table, we can return to our discussion of indiscernibility. Given the troubles we had with Def 7.3 and the discovery of natural properties, we can offer the following, revised definition of indiscernibility:

**Def D7.3* Indiscernibility.** Things *a* and *b* are *indiscernible* if and only if *a* and *b* have the same natural (monadic) properties.

We can similarly modify the Identity of Indiscernibles, as follows:

**7.3T.1 Restricted Identity of Indiscernibles.** If (thing *a* exemplifies *F* if and only if thing *b* exemplifies *F*), where *F* is a natural monadic property, then *a* is identical to *b*.

Finally, we can recast Wise UP as the claim that universals satisfy Restricted Identity of Indiscernibles, whereas particulars do not. That is, we read Def D7.1 and Def D7.2 as deploying this revised notion of indiscernibility.[7] (Hereafter, we will use 'Indiscernibility' to mean Def D7.3*.)

The question then becomes, *do* particulars run afoul of Restricted Identity of Indiscernibles while universals do not? We are going to leave an examination of the first part of that question, having to do with particulars, for the next chapter and our discussion of the Bundle Theory of substance. We turn presently, though, to the second half, the question whether universals satisfy Restricted Identity of Indiscernibles.

For the sake of space, we will only suggest a few ways one might defend the claim that universals are identical if indiscernible. First, one might defend Causal Structuralism (6.1T.1T plus 6.2T.1), which has an implication that every property confers on the things that have it a unique set of causal powers. If, then, the properties of *conferring such-and-such power* (for the right *such-and-such's*) are natural, then it follows that each universal exemplifies a unique set of natural properties having to do with what powers they confer. Thus, it follows that universals satisfy Restricted Identity of Indiscernibles. Second, one might defend the claim that properties of *identity* (and/or *distinctness*) for universals are natural. For example, you might claim that the property of *being the universal NEGATIVELY CHARGED* is natural. If so, and since only the universal NEGATIVELY CHARGED can exemplify this property, then each universal exemplifies a unique natural property. Thus, universals satisfy Restricted Identity of Indiscernibles.

Third, and maybe most promising, one might claim each universal has a unique set of categorial relational properties. What are categorial relational properties? Simply, they are relational properties that have to do with universals' categorial relations! Categorial relations, as we noted near the beginning of the chapter, are relations like *being a determinable of* and *being a determinate of*. Categorial relational properties, then, are properties that a universal exemplifies just in case it stands in one of these categorial relations to some specified other universal. Suppose, for example, that colors are universals, and that there is further a universal COLOR. Each color more determinate than the universal COLOR stands in the *is a determinate of* relation to COLOR, and thus each color more determinate than the universal COLOR exemplifies the categorial relational property of *being a determinate of COLOR*. Further, COLOR stands in the *is a determinable of* relation to each of these more determinate colors, so COLOR exemplifies, for example, the categorial relational property of *being a determinable of REDNESS*. REDNESS is, of course, a determinable of BURGUNDY, SCARLET, and so on, and so exemplifies the categorial relational properties of *being a determinable of BURGUNDY*, *being a determinable of SCARLET*, and so on.

Plausibly, each universal exemplifies a unique set of such categorial relational properties. No other universal has just the determinates and determinables that COLOR does, no other universal has just the determinates and determinables that SHAPE has, no other universal has just the determinates and determinables that BURGUNDY has. Clearly this list could go on, and we can think of no examples of universals that don't fit the pattern. So, since each universal stands in a unique set of categorial relations, and since which categorial relational properties a universal exemplifies is determined by these sets, each universal exemplifies a unique set of categorial relational properties.

Further, it is plausible that categorial relational properties are natural. Universals seem to be more or less similar to one another, and categorial relational properties offer a way to account for these similarities and dissimilarities. For example, colors are more similar to one another than they are to shapes, and this can be accounted for by noting that colors share more categorial relational properties (like the property of *being a determinate of color*) than they share with any shape. Since grounding resemblance is one telltale sign of naturalness, we have a reason to think that categorial relational properties are natural.

The most obvious objection to this strategy is simple. We've seen that relational properties in general are bad candidates for being natural, and so we need some reason to think that categorial relational properties are unique among relational properties for being natural. We have already seen one such reason in the preceding paragraph. And another reason is not far to seek. It is part of what it is to be a certain universal that it stands in just the categorial relations it does to just the other universals it does. A complete description of the fundamental nature of universals cannot be complete without invoking categorial relational properties. This strongly suggests that such properties are natural.

You might have the following thought. Can't one say the same thing about properties of identity and distinctness? Aren't they, too, relevant to a thing's being what it is? Though the case of properties of identity (and distinctness) is more complicated than many think (see Pickavance forthcoming), they aren't relevant *in the same way*. In particular, while it is the case that something can't be what it is without exemplifying the right property of *identity*, properties of *identity* seem to presuppose the identity in question. Take THP's dachshund Elsie. Elsie exemplifies the property of *being identical to Elsie*. But her doing so seems to presuppose that she stands in the identity relation to herself. Since one can't stand in relations to non-existent object, Elsie's exemplifying her property of identity presupposes her own existence!

Think of it this way. A full description of the world needn't appeal to Elsie's identity property. Just mentioning Elsie herself is enough to convey that she is identical with Elsie. There is no need for an appeal to the property of *being identical to Elsie*. But this is not so with categorial relational properties. No full description of the universal ORANGE, for example, can fail to mention that ORANGE is a color. Thus, a full description of that universal invokes the property of *being a determinate of COLOR*.

If the foregoing is right, then we have reason to think that universals satisfy Restricted Identity of Indiscernibles. Further, if particulars fail to do so (and we will see in the next chapter that this is likely), then Wise UP is successful as a universal-particular distinction.

*7.2.1.2 Russell's Intensional Paradox* To get at the second problem facing Realism, recall the second bit of near-universal agreement noted in Section 7.1, namely, the connection between something's satisfying a predicate and its exemplifying a property. We there noted the plausibility of the following account: the sentence formed by combining a name for $x$ with the predicate $P$ in a grammatically appropriate way is true if and only if $x$ exemplifies the property correlated with $P$. While this account seems like a natural, if not inevitable starting point, it is not plausibly universally applicable. Indeed, it *cannot* be.

To make our way to a defense of that claim, consider the true sentence 'Elsie is not a cat'. Here, we have a name for Elsie, 'Elsie', combined in a grammatically appropriate way with the predicate 'is not a cat'. If we apply the foregoing account, the result is that we are committed to there being a property of *not being a cat*. While this may be the right account of this truth, it doesn't seem so inevitable as the account of the truth of 'Elsie is a dog.' Why? Because the property of *not being a cat* is a surprising sort of property. Indeed, it seems like the property of *being a cat* ought to play a role in the truth of the sentence 'Elsie is not a cat,' which it would not were this account correct. In fact, it is initially very tempting to treat 'Elsie is not a cat' as true if and only if it is not the case that Elsie exemplifies the property of *being a cat*. This account also avoids commitment to the negative property of *not being a cat*. Other logically complex predicates admit of similar treatment. For example, disjunctive predicates like 'is *F* or *G*' seem to be satisfied by an object *x* if and only if (simplifying a bit) one gets a truth either when one combines a name for *x* with the predicate 'is *F*' or one gets a truth when one combines a name for *x* with 'is *G*'; that is, if and only if either *x* exemplifies the property correlated with 'is *F*' or *x* exemplifies the property correlated with 'is *G*'. No mention of a disjunctive property is needed! Conjunctive predicates like 'is *F* and *G*' can be treated analogously and without appeal to conjunctive properties. If we go these ways, though, then there are some predicates, among them logically complex negative, disjunctive, and conjunctive predicates, which *do not* correlate uniquely with individual properties.[8]

In fact, a version of Russell's Paradox (the so-called "intensional" version, in contrast to the "set-theoretic" version) can plausibly be used to show that it *can't* be the case that every predicate is uniquely correlated with a single property.[9] Suppose that there is a unique property for every predicate. According to many views of properties, properties do not always exemplify themselves. *Redness*, for example, is not self-exemplifying. (The property of *being red* does not exemplify itself; the property of *being red* isn't red.) Here we have a predicate, 'is not self-exemplifying', that is true of *redness*, and thus according to our supposition, there must be a property of *not being self-exemplifying*. There should be a fact of the matter, for any property, whether that property exemplifies the property of *not being self-exemplifying*. But if we ask, about the property of *not being self-exemplifying*, whether it exemplifies the property of *not being self-exemplifying*, we face a dilemma. Either this property is self-exemplifying or it is not. Suppose that the property of *not being self-exemplifying* exemplifies itself. Then it is self-exemplifying. But it exemplifies itself if and only if it has the property of *not being self-exemplifying*. If that is right, then it is *not* self-exemplifying. Contradiction. It must be that the property of *not being self-exemplifying* does not exemplify itself. Then it is not self-exemplifying. Which means, by our account of the truth of that sentence, that it exemplifies the property of *not being self-exemplifying*. Which means it *is* self-exemplifying. Contradiction again! Whether this property is self-exemplifying or not, we get a contradiction. So either there is no fact of the matter about whether this property is self-exemplifying or our account of the truth of sentences involving the predicate 'is not self-exemplifying' is wrong. Most metaphysicians are unhappy with the first option, and so go for the second. A promising way to make good on the second option is to just deny that there is any property of *not being self-exemplifying*.

If we follow these philosophers, we face a very general question. Just how many properties are there? Answers to this question fall on a spectrum from super-abundant theories to super-sparse theories. Super-abundant theories maintain that there are a great many properties, while super-sparse theories maintain that there are very few properties. Abundant theories of properties characteristically maintain that there is at least one property for every possible predicate (or concept), while spare theories characteristically claim that there are a great many predicates—logically complex predicates, for example—that are not correlated with a unique property. At any rate, Realists will need to be careful not to draw arbitrary lines, so as not to run afoul of Ockham's Razor.

*7.2.1.3 Bradley's regress* A third challenge for Realists arises because they takes the character of an object to be determined by its relationship to universals. This challenge was raised powerfully by F.H. Bradley (1897/1930). We will present one version of Bradley's objection in this section (a version somewhat loosely based on Bradley's text), but in Section 11.2.3, we will return to it as an objection to the existence of multiple entities.

First, we observe that Realists seem committed to what Michael Loux (2006) has called "Platonic Schema":

**Platonic Schema.** Something *o* is *F* if and only if *o* exemplifies the universal *F*.

There is a worry that Platonic Schema results in a regress, and this worry can take a definite shape in at least two ways.

First, consider the following:

(3) Lyle is sweet.

Here we have a true claim about THP's son, Lyle, a claim that concerns Lyle's character. (3) is a simple subject-predicate sentence of the form '*o* is *F*'. Realists, then, will offer the following account of Lyle's having this character, given their commitment to Platonic Schema:

(4) Lyle exemplifies the property of *being sweet*.

One way to read (4) is as another subject-predicate sentence of the form '*o Fs*', where 'Lyle' is '*o*' and 'exemplifies the property of *being sweet*' is '*Fs*'. But then, given Platonic Schema, Realists will be forced to say that (4) is true if and only if (5) is:

(5) Lyle exemplifies the property of *exemplifying the property of being sweet*.

But (5) is yet another subject-predicate sentence, and will generate a still further commitment given Platonic Schema, and so on. We have generated a regress.[10]

There is a second way a regress might get going. We can point out that Loux's Platonic Schema does not fully capture the Realists' commitment to properties. Realists believe not only *o* exemplifies *F*-ness whenever *o* is *F*—they also believe that *o* is *F because o* exemplifies *F*-ness. Let's call this the Strengthened Platonic Schema:

**Strengthened Platonic Schema.** Whenever something *o* is *F*, *o* is *F* because *o* exemplifies the universal *F*.

What can Realists say about these regresses? There are two ways to blunt the trouble created by them. First, one can accept that these are infinite regresses but deny that they are problematic. Second, one can find a way to stop the regress by restricting or rejecting Platonic Schema. Let's look a bit more at each of these strategies.

Suppose that one accepted these regresses but wanted to say that they are not problematic. Isn't this tack doomed from the beginning? Aren't *all* infinite regresses problematic? Maybe not. First, notice that there are infinite *sequences* at which no one balks. For example, the natural numbers (0, 1, 2, 3,...) form an infinite sequence, but no one takes this as a reason to worry about them.[11] Now, this sequence isn't really a *regress* in any important sense. Regresses occur when, in the process of explaining or explicating something, one discovers a further, similar thing that must be explained or explicated as well, and that when one goes to explain or explicate that further thing, one discovers a still further thing that must be explained or explicated, and so on. *Vicious* regresses require that one thing is *explained by* or somehow *dependent on* something else, which something else is itself explained by or dependent on another thing further down the regress. This is problematic because one can never fully ground or explain the initial thing in the regress, the thing one is trying to ground or explain. The further requirement for viciousness will become a bit clearer as we consider our two regresses.

Consider the first regress, which starts with (3) and moves to (4), (5), and so on. Given a commitment to Platonic Schema, a commitment to (3) entails a commitment to (4), (5), and so on. However, Platonic Schema does not require that Realists *explain* (3) by appealing to (4), nor must Realists say that (3) is *dependent on* (4) in any way. Realists can simply insist that (3), (4), (5), and so on are related facts, and that (3) can be clarified by invoking (4), but that this relationship shouldn't be thought of as any kind of explanation or as charting any sort of metaphysical dependence. This first regress, then, is plausibly not vicious.

What about the second regress, which corresponds to the Strengthened Platonic Schema? Here there are explanations: (3) is true in virtue of the truth of (4), and (4) is true in virtue of the truth of (5). Thus, this regress is vicious because we cannot fully ground (3) without fully grounding (4), which cannot be fully grounded without (5), and so on to infinity. This regress looks a great deal more troubling.

However, there are other ways to respond to this regress. Most of these ways reject, in one way or another, the Strengthened Platonic Schema. One seemingly easy route is to deny that there really is a fundamental, universal relation of INSTANTIATION or EXEMPLIFICATION. This view is a form of Nominalism (more specifically, Ostrich Nominalism—see the next subsection) about that relation. If it's correct, then truths like (4) are not to be treated using the Strengthened Platonic Schema. Similarities that correspond to natural properties (like *sweetness*) are grounded in the sharing of a universal, but the way in which two instantiation pairs—take the pair *a* and *F* and the pair *b* and *G*, whenever *a* exemplifies *F* and *b* exemplifies *G*—are similar is not grounded in their sharing the universal EXEMPLIFIES. One might go in more generally for a sparse theory of universals, either by denying that any relations correspond to universals (restricting universals to monadic properties) or by denying that second-order, logical relations like *exemplification* correspond to universals. If one goes on to restrict the Strengthened Platonic Schema in the same way, then one will have solved Bradley's regress.

There is still a third way of understanding Bradley's regress, a way that is in fact closer to Bradley's actual text. This approach to the regress looks at the problem of what explains the unity of the truths or facts involved. This version of the regress is not immediately relevant to the evaluation of Realism, so we will take it up again in Section 11.2.3.

*7.2.1.4 The Challenge of Ostrich Nominalism* We have seen that appealing to a sparse theory of universals is imperative for eluding some of the most challenging troubles facing Realism. The sparsest theory of universals, and one that eludes all of the troubles highlighted above, is one that posits no universals at all. An easy way to think about Nominalism is as just the denial that there are universals, and we'll think of Nominalism that way in this section. (We'll take this back a bit in our discussion of Resemblance Nominalism in Section 8.1.3.) One virtue of Nominalism is that it avoids the need to deal with the Universal-Particular Distinction and with Russell's Paradox. But further, there is an obvious challenge to the positing of any universals, grounded in Ockham's Razor.

An *Ostrich* Nominalist is one who denies the need for any general explanation of putative facts involving attribute agreement. The term was coined by opponents of the view who compared such Nominalists with the proverbial ostrich that sticks its head in the sand when faced with the similarity facts. Recall the red and square RedSquare, the red and circular RedCircle, and the blue and circular BlueCircle. Consider the following:

(7) RedSquare and RedCircle are both red/are similar with respect to color.
(8) RedCircle and BlueCircle are both circular/are similar with respect to shape.

According to most theories of properties, (7) is true because RedSquare and RedCircle exemplify one and the same (shareable) property, the property of *being red*, while (8) is true because RedCircle and BlueCircle exemplify one and the same (shareable) property, the property of *being circular*. Ostrich Nominalists, however, deny these explanations. Indeed, they deny the need for positing shared properties at all. Instead, Ostrich Nominalists simply insist that the fact that RedSquare is red is metaphysically fundamental, the fact that RedCircle is red is metaphysically fundamental, and (7) is made true by these two facts together. Similarly, the fact that RedCircle is circular is metaphysically fundamental, and the fact that BlueCircle is circular is metaphysically fundamental, and (8) is made true by these two facts together. The pattern is to explain putative cases of attribute agreement by exploiting metaphysically fundamental facts of character-having without ever appealing to properties that get instantiated. Notice that the Ostrich Nominalists cannot do their explaining of facts of attribute agreement in any *generalized* way. That is, they cannot say that claims of attribute agreement are made true because the things whose attributes agree share a property. They are precisely trying to avoid talking about properties at all! So, while there are explanations for every particular fact of attribute agreement in terms of other facts of character-having, there is no general explanation of facts of attribute agreement.

Not all Nominalists are Ostriches, and so we get a division of the Nominalist landscape:

**7.1A.1T Reductive Nominalism.** Nominalism is true, and there is a general explanation of the fact that some particulars have properties in common.

**7.1A.1A Ostrich Nominalism.** Nominalism is true, and there is no general explanation of the fact that some particulars have properties in common.

We take up Reductive Nominalism in Section 8.1. For now, we focus on Ostrich Nominalism.

Ostrich Nominalists face one obvious challenge, namely, a potential conflict with Classical Truthmaker Theory. Consider again (3):

(3) Lyle is sweet.

Ostrich Nominalists take (3) to be fundamental, in no need of further metaphysical grounding. However, what is the truthmaker for (3)? It can't be Lyle himself, since Lyle could exist and not be sweet. If we suppose that the truthmaker of (3) is something like the fact that (3), we begin moving beyond the scope of Ostrich Nominalism, and we face the problem of explaining what this fact is and how it differs from Lyle, on the assumption that there is no such thing as *sweetness*. If there are no properties, it is hard to see what could differentiate one fact from another. In light of this, some prominent philosophers (including Frege, Alonzo Church, Quine, and Donald Davidson) have argued that there could be only One Big Fact.

However, most Ostrich Nominalists reject Classical Truthmaker Theory. Spectral Truthmaker Theory and Truth Supervenes on Being seem tailor-made for Ostrich Nominalism, since on these views the truth of (3) is grounded in Lyle's existing and being as he is, which requires no further entity.

In fact, it seems that Ockham's Razor demands that we prefer Ostrich Nominalism to Realism, unless there are some further facts that the Nominalist cannot explain. This is because Realism demands more things than Ostrich Nominalism. Realists about universals and particulars—UP-Realists—posit two classes of things, particulars and universals, and a fundamental relation of *exemplification* or *instantiation* between things and universals.

**7.1T.1T UP-Realism.** There are fundamental universals and fundamental particulars, and the latter instantiate the former.

(We will consider some varieties of UP-Realism in Chapter 9.) On the other hand, Ostrich Nominalists posit only one class of things, particulars. Let's suppose that there are $k$ fundamental particulars and $n$ fundamental kinds of things. Ostrich Nominalism requires only $k$ things, falling into $n$ irreducible kinds. UP-Realism, on the other hand, requires $k+n$ things, falling into two fundamental kinds (things that instantiate something and things that are instantiated).[12] If $n$ is very large, then the domain of Ostrich Nominalism is much smaller quantitatively.

However, this argument from Ockham's Razor is too hasty, as Bryan Pickel and Nicholas Mantegani have recently argued (Mantegani 2010, Pickel 2010, Pickel and Mantegani 2012). This is because Ostrich Nominalists must posit a very large number of metaphysically fundamental *sorts* of things: spheres, cubes, blue things, red things, etc. For UP-Realists, each of these kinds of things can be reduced to universals and instantiation pairs:

(9) $x$ is a red thing if and only if $x$ instantiates the universal REDNESS.
(10) $x$ is a sphere if and only if $x$ instantiates the universal SPHERICALITY.

Thus Ostrich Nominalism is much larger qualitatively, since UP-Realism has only two fundamental kinds, whereas Ostrich Nominalism has *n* fundamental kinds. This leads us to an important question relevant to evaluating Ostrich Nominalism: is quantitative economy more important than qualitative economy or vice versa?

There are good reasons to think that qualitative economy is much more important than quantitative economy. Both are important, but the kind of simplicity that comes with reducing the number of fundamental kinds or categories of things is much more valuable. Simply adding more entities to already existing fundamental kinds is a relatively trivial addition, compared to adding entirely new kinds to our theory. This preference can be seen to be at work throughout modern science. We are always willing to add new entities (such as atoms, subatomic particles and quarks) in order to reduce the number of fundamental kinds of things (for example, reducing the 100+ elements to the three kinds of subatomic particles—protons, neutrons, and electrons). Thus, UP-Realism represents exactly the kind of scientific advance over Ostrich Nominalism that modern atomic physics holds over Daltonian chemistry.

We can codify this into a methodological principle. First, we need to articulate a fourth corollary Ockham's Razor:

**PMeth 1.4 Fourth Corollary of Ockham's Razor.** Adopt the theory that jointly minimizes the following classes: (i) facts about the existence of fundamental things, (ii) facts about the holding of fundamental (natural) properties of and relations between fundamental things, and (iii) facts about brute metaphysical necessity. If more than one theory minimizes these classes (if some trade-off between them is inevitable), adopt that theory which makes the best trade-offs.

Assuming that qualitative simplicity is more important than quantitative, we can add the following addendum.

**PMeth 1.4.1 Prioritizing Qualitative Economy (Addendum to Ockham's Razor).** When trading off quantitative economy (minimizing the number of fundamental things) and qualitative economy (minimizing the number of fundamental properties of and relations between fundamental things), always prefer the latter.

One might worry that Ostrich Nominalism makes hash of the Fourth Corollary of Ockham's Razor and Prioritizing Qualitative Economy, eschewing as it does the very existence of properties and relations. Can Ostrich Nominalists even make sense of these principles? In a word: yes. They can and should do. Ostrich Nominalists can put the principle into what is known as the 'formal' mode, speaking about 'predicates' or 'general names', in place of 'properties' or 'kinds'. Ostrich Nominalists should prefer theories that can be expressed in languages with the fewest number of basic, underived general terms or names.

But can Ostrich Nominalists make sense of names, predicates, or other linguistic types? Mustn't they refuse to accept anything but particular tokens, particular episodes of speaking, or writing? No, they need not take such an extreme position. It is true that, for Ostrich Nominalists, there is no one thing common to every occurrence of the word 'red', for example. Nonetheless, Ostrich Nominalists can make sense of the practice of counting-by-type, versus counting-by-token. For example, consider the following sentence token:

(11) The red door is red.

Ostrich Nominalists can make perfect sense of each of the following claims:

(12) Sentence (11) contains five words (counting by tokens).
(13) Sentence (11) contains four words (counting by types).

So long as the Ostrich Nominalists can make sense of counting by types, they can apply the principle of qualitative economy to the expression of scientific (and metaphysical) theories.

If that is right, then Ostrich Nominalism is in a surprising amount of trouble vis-à-vis Ockham's Razor. This is a surprising and serious problem for the Ostrich, since an important, if not fundamental driving force behind Ostrich Nominalism is the offer of a metaphysically simple theory that powerfully accommodates the data a theory of properties is meant to accommodate.

This debate over qualitative economy between Realists and Ostrich Nominalists connects with another issue, one that we discussed in Chapter 6: are universals individuated by their causal roles or by their causal roles together with their systemic categorial relations to other universals or does each universal have its own sicceity or thusness? If each universal is individuated from others by its own, unique sicceity, then these sicceities would greatly add to the qualitative complexity of Realism, neutralizing the advantage over the Ostrich Nominalism gained by positing universals and instantiation. To use this argument against Ostrich Nominalists, it seems that Realists should reject sicceities and embrace the individuation of universals by their causal and categorial properties.

## 7.3 Universals and the Problem of Intentionality

What connects the predicates of our language (including adjectives and common nouns like 'white' or 'dog') with properties in the world? What connects our corresponding mental concepts with those properties?

The Ostrich Nominalists have no general explanation of this fact. They must rely on something like the homonymic clauses of Tarski's truth theory:

(14) The predicate '_____ is white' applies to *x* if and only if *x* is white.
(15) The predicate '_____ is a dog' applies to *x* if and only if *x* is a dog.

In fact, they must claim (since they are doing metaphysics) that the applicability of the predicate '___ is white' to *x* must be grounded by *x*'s whiteness. This lack of *general* expanation of concept applicability looks like a serious defect, if we compare the Ostrich Nominalism with Realism. Realists can offer a single, general explanation:

(16) A predicate of the form '___ is *F*' applies to *x* if and only there is a concept *C* that is conventionally associated with '___ is *F*' (as its meaning), and *C* applies to *x* (and when it does, it applies because *C* applies to *x*).
(17) A concept *C* applies to *x* because there is a universal *U* such that *U* is the one and only universal that is a part of *C*, and *x* instantiates *U* (and when it does, it applies because *x* instantiates U).

Of course, this explanation will only work for those predicates that have a single meaning, and even then only for those predicates whose meaning involves a single universal. It is probable, given a sparse theory of universals, that most of our concepts involve some complicated logical structure that includes more than one universal. Nonetheless, (16) and (17) should work for an important set of basic cases of intentionality and so could form part of a general theory of the intentionality of predicates.

Ostrich Nominalists could respond by pointing out that there is one important case of predication that the Realists' account cannot handle without incurring an infinite regress: the predicate '__ instantiates ___'. We're assuming that Realists will avoid Bradley's regress by denying that there is such universal as INSTANTIATION. This means that they cannot make use of the relevant instance of schema (17):

(18) The concept *instantiation* applies to the ordered pair <*x*, *U*> if and only if there is a universal INSTANTIATION such that INSTANTIATION is the one and only universal that is part of *instantiation*, and <*x*,*U*> instantiates INSTANTIATION.

We recommend that Realists reject (18) as part of their theory of intentionality. In this one case, Realists should employ exactly the kind of Tarskian, homonymic axiom that Ostrich Nominalists use in all cases:

(19) The concept *instantiation* applies to the ordered pair <*x*, *U*> if and only if *x* instantiates *U* (and that concept applies to the pair because *x* instantiates *U*).

Thus, the Realists' account of the intentionality of simple predicative concepts will have two axioms, (19) and (20).

(20) For all simple concepts *C* other than *instantiation*, *C* applies to *x* if and only if there is a universal *U* such that *U* is the one and only universal that is part of *C*, and *x* instantiates *U*.

A theory with two axioms is simpler, and therefore better, than a theory with many axioms doing the same job. If our thought includes many simple predicative concepts, the Ostrich Nominalists' theory will be that much more complicated than the Realists'.

It would also be possible to build a relatively simple theory of predicative intentionality with the resources of Resemblance Nominalism.

(21) For all concepts *C*, *C* applies to *x* if and only if, for every mental token *c* of *C*, there is one and only one part of *c* that belongs to a resemblance class, and *x* and *c* belong to the same resemblance class.

Resemblance Nominalists will have to make one concept the exception to rule (21): namely, the concept of *resemblance* in terms of which we define *resemblance classes*. Just as the Realists had to do, Resemblance Nominalists will have to explain the application of resemblance homonymically (in terms of resemblance itself):

(22) The concept of *resemblance* applies to $x$ and $y$ if and only if $x$ resembles $y$.

Once again, a theory with two rules will be preferable to one with many independent rules.

Realists face a further problem: namely, they must explain how a concept can contain a universal without instantiating the universal. We don't want the concept of a horse to be itself a horse! I don't have a little horse running around in my mind every time I think of a horse.

Realists will need some way of distinguishing the way in which a property is part of an ordinary particular and the way in which it is part of a concept in the mind. Aristotle and many of the scholastic philosophers who followed him made this distinction by arguing that properties exist with matter in ordinary particulars and without matter in mental concepts. This seems to work for properties that are essential properties of material objects, like properties of *mass*, *velocity*, *chemical composition*, and *size*. However, what about simple properties that aren't essentially properties of a body, like *thinking* or *feeling sad*? Surely I can have the concept of *feeling sad* without actually feeling sad myself at that moment. One more potential problem: Aristotle's solution only works on the assumption that the human mind is immaterial (as Aristotle argues in *De Anima* III).

Realists have at least two further options, beyond the traditional Aristotelian solution. First, they might suppose that concepts are wholes that include, besides the universal that the concept expresses, a universalizing element, which marks the whole as a concept rather than as particular instance of the property. These universalizing or conceptualizing elements would function in a way similar to the functioning of a bare particular. They would make the concept as a whole be the truthmaker for the mind's ability to think of the contained property, rather than the truthmaker for a proposition asserting the existence of a particular that instantiates the property.

Alternatively, Realists could give up on the theory that the simple concept of *F*-ness contains the *F*-ness universal or an *F*-ness trope as a part. They could instead suppose that the intentional connection between the concept and the property it expresses involves some *other* relation. For example, one might think that a concept $C$ expresses a property $P$ just in case (i) there was some relevant exercise by some particular thing of a power conferred upon that particular by $P$ that resulted in the formation of the concept $C$ (in the minds of the members of the relevant community or other population), and (ii) the concept $C$ belongs to a system of concepts $S_1$, property P belongs to a system of properties $S_2$, and there is a formal isomorphism between $S_1$ and $S_2$ (i.e., the two systems exhibit the same formal structure). This is still the sort of general account of intentionality that would be unavailable to Ostrich Nominalists. Indeed, it would also be unavailable to Neo-Humeists, who deny the fundamental existence of powers.

Clearly, the problem of intentionality is a difficult one. Until a fully satisfactory account of intentionality in terms of universals or tropes is developed, the issue cannot count decisively against Ostrich Nominalism. However, it does point to a potentially grave challenge to that theory.

## 7.4 Properties as the Ground of Causal Powers

In Chapters 4 through 6, we examined four theories of causal powers and dispositions: Neo-Humeism, Strong Hypotheticalism, Strong Nomism, and Strong Powerism. As far as we can tell, Neo-Humeism and Strong Hypotheticalism have no implications for our general theory of properties. They are fully compatible with all of the various forms of Nominalism and Realism.

Nomism, in contrast, favors the truth of Realism. Nomists need properties as constituents of the laws of nature. This rules out Ostrich Nominalism. A law of nature is by its very nature something timeless and universal, connecting properties to one another. These properties certainly cannot be individual tropes, since then each law would be a merely particular fact about those particular tropes. One could, perhaps, take the laws of nature to be relations among resemblance classes of tropes, so long as these classes included merely possible tropes along with all of the actual ones (since laws of nature are supposed to govern merely possible, counterfactual situations). This seems like an unnatural combination of views, however. It is not surprising that Nomists overwhelmingly favor Realism.

Since Powerists take causal powers and dispositions to be grounded in the essences or natures of properties, Powerism is inconsistent with Ostrich Nominalism, which includes the denial of the real existence of properties altogether. In fact, it is hard to see how Powerism could be combined with any form of Extreme Nominalism. Mere classes of ordinary particulars would not seem to be the sort of thing capable of having a power-conferring essence. Realism and Powerism make a good fit, since universals could well have essences that confer powers and dispositions on the things that instantiate them. It seems that Powerism could be combined with Moderate Resemblance Nominalism, with the understanding that it is essential to each trope that it bestow a certain set of powers on its bearer, with exactly similar tropes conferring exactly similar powers (John Heil defends such a view).

## Notes

1 This isn't quite enough. We would need to generalize this definition to include predicates/properties with more than one place. A *specifying description* of an object is a description which only that object satisfies.
2 See Aristotle's *On Interpretation* 17a37, Hochberg (2004), Lowe (2004), Oliver (1996), and cf. MacBride (2004, 2005), Ramsey (1925).
3 See Russell (1911), and cf. Lewis (1983, 1986a), MacBride (1998).
4 This view is mentioned by Wisdom (1934). See also Ehring (2004), Williams (1986), and Pickavance (unpub.).

5 There is one account of spatial location, the Theory of Spatial Qualities (17.1T.1T), on which a thing's location is an intrinsic characteristic of it. See Chapter 17.

6 Our usage of 'natural' will diverge from the particularities Lewis attaches to the idea. But we believe the spirit is similar. Space prevents us detailing the (interesting!) literature on naturalness.

7 See Pickavance (forthcoming) for a more thorough defense of this strategy.

8 Relational predicates, like 'is two meters east of Elsie' and 'is taller than Wilt', are plausibly treated as not corresponding to a unique property. Rather than saying that the fundamental level involves a thing exemplifying the property of *being two meters east of Elsie*, one can say that the fundamental level involves a thing standing in the relation of *being two meters east of* to Elsie.

9 We ignore attempts, like that of Russell himself (cf. Russell and Whitehead 1962 or Russell 1959), to dissolve the paradox by "typing" predicates and properties.

10 The regress could also move forward by thinking of EXEMPLIFICATION as a relational universal that links Lyle and the universal SWEETNESS (see Chapter 10.1). In that version of the regress, we would move from 4 to 5*:(5*) Lyle and the universal SWEETNESS exemplify EXEMPLIFICATION.

11 Well, there are *some* people who worry about the natural numbers for this reason. But they are in a very, very small minority.

12 There is one complication to consider in applying Ockham's Razor to the evaluation of UP-Realism: how to count the primitive binary relation *x instantiates y*. Does this relation introduce a single primitive sort (a sort of ordered pair) or two primitive monadic sorts (*instantiating* and *being instantiated*)? Or does it introduce a potentially infinite number of sorts, two for every entity *A*: *instantiating A* and *being instantiated by A*? We can rule out the last of these suggestions, since counting sorts in this way would give priority to quantitative economy over qualitative economy, since every individual entity would give rise to two primitive sorts. The first suggestion seems to under-count the commitments of Classical Realism, especially since instantiation is a non-symmetric relation, with a difference between *being an instantiator* and *being instantiated*. So, we'll assume that fundamental binary relations like *instantiation* correspond to exactly two fundamental sorts, each with a variable parameter *x*: the sort of instantiating *x*, and the sort of being instantiated by *x*. These two sorts can be used to define 'universal' and 'particular': a universal is something that can be instantiated, while no particular can be instantiated.

# 8

# Reductive Nominalism and Trope Theory

## 8.1 Reductive Nominalism

As we noted, Ostrich Nominalism is only one type of property Nominalism. The other form is Reductive Nominalism:

**8.1T Reductive Nominalism.** Nominalism is true, and there is a general explanation of the fact that some particulars have properties in common. (8.1T is identical to 7.1A.1T.)

There are a number of different versions of Reductive Nominalism, versions distinguished by the way in which each accounts for facts about having and sharing properties. We will discuss three broad varieties of Reductive Nominalism: Predicate Nominalism (Section 8.1.1), Class Nominalism (Section 8.1.2), and Resemblance Nominalism (Section 8.1.3). We will spend most of the first section on Resemblance Nominalism, which is (by our lights) the best alternative to UP-Realism.

Resemblance Nominalists believe that the general explanation of the fact that particulars resemble each other appeals to the existence of a general relation of *resemblance* between particulars. This relation is a basic or fundamental one. The fact that one particular resembles another is not grounded in their instantiating any common universals. Resemblance Nominalists come in two sub-varieties, depending on whether they take the resemblance relation to hold between particular properties (called 'tropes') or particular things that have properties (which we will call 'ordinary particulars'). Trope Nominalists ground all similarity between ordinary things in a *resemblance* relation between tropes, while Extreme Nominalists take the *resemblance* relations between ordinary particulars to be fundamental.

*The Atlas of Reality: A Comprehensive Guide to Metaphysics*, First Edition.
Robert C. Koons and Timothy H. Pickavance.

Socrates' paleness and Plato's paleness would be examples of two distinct tropes, one belonging to each philosopher, even when they are pale in exactly the same way. Ordinary particulars are things like Socrates or the Moon: particular things that *have* properties, not properties themselves.

Trope Nominalists are not the only philosophers who believe in tropes. There are also Trope Realists. Trope Realists believe in two distinct types of properties: tropes (like Socrates' paleness) and universals (PALENESS itself). Extreme Realists, in contrast, believe in universals but not in tropes, just as Extreme Nominalists believe in ordinary particulars but not in tropes. In Section 8.2, we focus our attention on Trope Theory. We argue that there are two competing conceptions of what tropes are like, namely, modular tropes and modifying tropes, and so two fundamentally different sorts of Trope Theory. We consider some objections to both.

### 8.1.1 Predicate Nominalism

We sometimes talk about properties. Indeed, sometimes we quantify over properties. That is, we use phrases of quantification that seem to require properties within their range of values, as in (1) and (2):

(1) Napoleon had *all of the qualities* needed by a great general.
(2) A circle has *more properties* in common with an ellipse than it has with a triangle.

Nominalists owe some account of what such phrases as 'all of the qualities' or 'more properties' stand for in (1) and (2) if there are no universals like GENIUS or ROUNDNESS. One historically popular account is Predicate Nominalism:

**8.1T.1 Predicate Nominalism.** Reductive Nominalism is true, and whenever two particulars resemble each other, their resemblance is grounded in the fundamental fact that the two fall under some one predicate of some language.

Predicate Nominalists, then, take properties to be predicates in some language (like English). Exemplification is just predicate satisfaction. This seems to be a case of conceptual grounding (as per Section 3.4), indicating that, for the Predicate Nominalist, there is nothing in the world corresponding to resemblance among particulars.

Predicate Nominalists might offer us the following paraphrases of (1) and (2):

(1′) All of the predicates in English that are necessarily true of all great generals are also true of Napoleon.
(2′) There are more predicates in English that are true of both circles and ellipses than there are predicates true of both circles and triangles.

A close cousin of Predicate Nominalism is Concept Nominalism:

**8.1T.2 Concept Nominalism.** Reductive Nominalism is true, and whenever two particulars resemble each other, their resemblance is grounded in the fundamental fact that the two fall under some one concept.

For Concept Nominalism, properties are concepts in some mind, something like predicates in a language of thought. Exemplification is falling under a concept. This is also a case of merely conceptual grounding (Section 3.4). Concept Nominalists could paraphrase (1) and (2) as:

(1″) All of the concepts that apply with necessity to all great generals apply to Napoleon.
(2″) There are more concepts that apply to both circles and ellipses than there are concepts that apply to both circles and triangles.

Clearly, these two accounts are structurally similar, and we treat them together.

A common complaint against both Predicate and Concept Nominalism is that they confuse the intrinsic properties of things with their extrinsic relations to a language or a mind. Things wouldn't cease to *be* spherical simply because the *predicate* 'is spherical' or the concept of *sphericality* ceased to exist. Things would still have been spherical even if there had been no languages and no concepts at all.

In addition, both seem to get the order of explanation wrong. A predicate like 'is a circle' applies to various concrete shapes by virtue of the fact that they are all circular. It's not the case that the shapes are circular by virtue of the fact that the predicate 'is a circle' is true of them.

### 8.1.2 Class Nominalism

The next version of Reductive Nominalism is Class Nominalism:

**8.1T.3 Class Nominalism.** Reductive Nominalism is true, and whenever two particulars have a property in common, this fact is grounded in the fundamental fact that the two belong to some one set or class.

Thus, Class Nominalism identifies properties with classes or sets. The property of *being spherical* simply consists in the set of spherical things. This has the advantage of identifying properties with things that have an existence that is independent of language or thought. Further, exemplification is just set membership; there is nothing more to exemplifying a property than being a member of the class that is that property.

There are, however, two serious types of objection to Class Nominalism. The first type results from Class Nominalism's identification of properties with classes of things. We will call this Class Nominalism's 'extensionality problem'. This extensionality problem consists of two sub-problems: the problems of contingent predication and of co-extensive properties.

THE CONTINGENT PREDICATION PROBLEM Let's suppose that Class Nominalism is true, that properties are just classes of their instances. Here is a fundamental fact about sets: two sets are identical if and only if they have the same members. The set of dogs is different from the set of cats because the two sets have different members. And there is only one set of dogs. Any set with all of the dogs in it just is the set of dogs! The problem is that because of this connection between sets and their members, a set cannot exist if all of its members do not exist. But it seems that, in typical cases, a given property could have been exemplified by things that do not actually exemplify it. For instance, there might have been one more green apple than there is in fact, and thus one more green thing than there is in fact. So Class Nominalism seems committed to all three of (3–5):

(3) The property of *being green* is identical to a certain set.
(4) The property of *being green*, since it is a set, has its members essentially; any addition or subtraction of members is impossible.
(5) The set of green things might have had different members than it has in fact.

Unfortunately for Class Nominalism, (3–5) are inconsistent: they can't all be true together. The conflict is between the fact that sets have their members essentially and the fact that properties aren't essentially exemplified by just the things that happen to actually exemplify them. Thus, it seems that properties can't be sets of things as Class Nominalism contends. This is a serious problem, and it generalizes easily to a great many properties.

THE CO-EXTENSIVE PROPERTY PROBLEM Another extensionality problem emerges when we consider two distinct but co-extensive properties. Imagine a small world in which all and only electrons have negative charge. In such a world, the class of electrons would be identical to the class of negatively charged things, and so the two properties would also be identical. But the two properties are obviously *not* identical! For if they were identical, then it would be impossible for something to be negatively charged that wasn't also an electron. However, it *is* possible for there to be negatively charged non-electrons. So Class Nominalism gets the wrong result here as well.

One might try to solve this problem by combining Class Nominalism with Possibilism (12.1A.1T), and identifying the properties with classes of things, both actual and possible. The property of *being spherical* would be the class of all actual and possible spherical things, rather than just the class of all actually spherical things. There is a clear problem with this suggestion. Some things, like Pluto, are both possibly spherical and possibly non-spherical. If sphericality just is the class of all possibly spherical things, then Pluto has sphericality in all possible worlds, since it belongs to the class of possibly spherical things in all worlds. This would make Pluto necessarily spherical, which again is obviously wrong.

This problem can be avoided if we are willing to embrace David Lewis's theory of possible worlds as concrete universes (14.1T.1T), along with his doctrine of Worldbound Individuals (16.1A). On this view, the class of spherical things will not include our Pluto (the Pluto in our possible world), but it will include the spherical 'counterparts' (to use Lewis's term) of Pluto in other worlds. This certainly works, but it comes at the steep price of embracing Lewis's account of possible worlds.

THE PROBLEM OF THE SUPER-ABUNDANCY OF PROPERTIES The second difficulty with Class Nominalism is a consequence of the fact that it is a super-abundant theory of properties. Every class of things corresponds to a different property, and all such properties are metaphysically on a par with each other. This super-abundancy of properties leads to a series of paradoxes involving induction and causation.

One of the first such paradoxes to be proposed was the *grue* paradox of Nelson Goodman, his so-called 'new riddle of induction' (1954). Goodman defines a new color property:

(G) For all *x*, *x* is *grue* if and only if either (a) *x* was first observed before 2000 and is green, or (b) *x* was first observed after 2000 and is blue.

Every emerald observed before the year 2000 was observed to be green. Thus, all of these emeralds were also grue. Given the uniformity of experience, we had (before the year 2000) equally good reasons for accepting each of the following two hypotheses:

(6) All emeralds are green.
(7) All emeralds are grue.

However, the two hypotheses made contradictory predictions about what we would see when we uncovered the emeralds first observed after January 1, 2000. Hypothesis (6) entailed that these emeralds are green, while hypothesis (7) entailed that they are all blue. We would all agree that it is (6), rather than (7), that is the most reasonable hypothesis to embrace, apparently because '____is green' is a more natural or conceptually simpler predicate than '___is grue'.

There are similar problems involved in the assignment of meanings to words and of contents to mental acts. For example, how do we know that the word 'green' in English doesn't correspond to a weirdly gerrymandered set of the kind introduced by Goodman's definition of 'grue' in (G)? How do I know that my concept of green corresponds to one set rather than another, when we try to extend the class of instances from those that we have actually encountered in the past to those we will encounter in the future?

To solve this problem, we have to make some distinction between the two properties. The sparse theory of universals offers a simple answer: some predicates correspond to universals, and most do not. There is no universal corresponding to the predicate '____is grue', while there might be one corresponding to '___is green'. Even if there is no universal corresponding to either, we might be able to use universals to discriminate between the two classes. The class of green things probably has a simpler definition in terms of universals than the definition of the class of grue things.

Can Class Nominalists mimic a sparse theory of universals? David Lewis (1983) suggested that they can by introducing a distinction between natural and unnatural classes. The class of green things is natural, and the class of grue things is unnatural. Natural properties, properties of the kind that can figure in induction and in the interpretation of words and concepts, consist in natural classes only. Lewis suggests that it is a fundamental, irreducible truth about certain classes that they are natural.

There are two difficulties with Lewis's suggestion. First, it seems that Lewis's account has the order of explanation backward. What makes a class natural is the fact that it corresponds to the extension of some real property. It's not the case that what makes something a property is that it corresponds somehow to a class of the right kind. In addition, there is good reason to doubt that sets or classes have any natural properties or relations, except for the relation of membership:

**PMeta 3 Membership the Only Fundamental Set-Relation.** If *S* is a set, then the only fundamental relation involving *S* is the membership relation between *S*'s members and itself.

A better solution for Class Nominalists would be to define the naturalness of a class in terms of the relations that hold among its members. Roughly, a natural class is one whose members resemble each other more than they resemble non-members. Green things resemble each other in color, but grue things bear no such resemblance to one another. This is a move toward Resemblance Nominalism, a view we'll consider presently.

### 8.1.3 Resemblance Nominalism

The final version of Reductive Nominalism that we take up is Resemblance Nominalism.

**8.1T.4 Resemblance Nominalism.** Reductive Nominalism is true, and whenever two particulars have a property in common, this fact is grounded in fundamental facts of resemblance between them.

The basic idea here is that at the fundamental level, there are just particulars, particulars which stand in metaphysically fundamental relations of resemblance. This is again a case of conceptual grounding: properties and property-exemplification are not to be found in reality: all we find there is a pattern of resemblance relations among particulars. For Resemblance Nominalists, if properties exist at all, they are constructed out of these particulars and identified with classes of resembling particulars (see Section 18.1.1 for more on logical constructions). More precisely, for each property *P* (other than the property of *resemblance* itself), Resemblance Nominalists propose that what it is to have property *P* is for a thing to resemble certain particulars, namely, those particulars in a *resemblance class* for *P*, where a resemblance class for a property is, intuitively, the set of things that exemplify that property (more precision below). Thus, if property *P* and property *Q* have the same resemblance class *C*, then the two properties are identical, since to be *P* is simply to resemble exactly the members of *C*, and to be *Q* is to resemble exactly the same things.

We can bring out the difference between Resemblance Nominalists and Realists by comparing their accounts of similarity and property-exemplification. Suppose that two particulars *A* and *B* resemble each other by way of having property *P*. Realists will say that the resemblance of *A* and *B* is grounded in the fact that they both exemplify the universal for property *P*. Resemblance Nominalists, in contrast, will take the resemblance to be fundamental, and they will suppose that each particular's exemplifying *P* is grounded in its similarity to certain particulars, namely, the resemblance class for *P*, a class that includes *A*, *B*, and all other particulars that resemble *A* and *B* in this same way.

There are two broad varieties of Resemblance Nominalism. In order to distinguish them and to see Resemblance Nominalism in more detail, we must have recourse to a few definitions:

**Def D8.1 Property.** A *property* is anything (whether universal or particular) that either grounds the character of something else or is implicated in a general account of attribute agreement.

**Def D8.2 Ordinary Particular.** An *ordinary particular* is a particular that is not a property.

**Def D8.3 Trope.** A *trope* is any particular that is also a property.

The nature of tropes is controversial, but Trope Theorists agree on at least this much: for any two things, let's call them *a* and *b*, which share some property *P*, there are two (non-identical) tropes of *P* had by *a* and *b* respectively. No trope, in other words, is had by more than one thing. In this sense, tropes are particulars. But tropes also ground character and are, in that sense, properties. We will return to tropes below (in Section 8.2), so we don't say much more about them here. On the basis of these definitions, we can distinguish extreme and moderate forms of Resemblance Nominalism:

**8.1T.4.1T Extreme Resemblance Nominalism.** There are only ordinary particulars, and whenever two ordinary particulars resemble each other, their resemblance is metaphysically fundamental.

**8.1T.4.1A Moderate Resemblance Nominalism (Trope Nominalism).** There are only particulars, and whenever two ordinary particulars resemble each other, their resemblance is grounded in the fundamental fact that the two are characterized by tropes whose resemblance is metaphysically fundamental.

To see how these views differ, and to get a bit clearer about the nature of Resemblance Nominalism, consider two red things, a ripe apple and a cardinal bird. Extreme Resemblance Nominalists claim that the property of *being red* is a resemblance class of apples, cardinals, and other ordinary red things. A resemblance class is a set of things that bear strong resemblance relations to each other. The first philosopher to make use of the idea of a resemblance class was Rudolf Carnap in his *Logical Construction of the World* (1928/1967: 113). Here is Carnap's definition (he called these classes 'quality circles'):

**Def D8.4.1 Resemblance Classes (Carnap).** A class of objects *X* is a *resemblance class* if and only if (i) each member of *X* resembles every other member of *X*, and (ii) nothing outside of *X* (that is, no non-member of *X*) resembles every member of *X*.

For Extreme Resemblance Nominalists, ordinary particulars are just unstructured wholes that primitively satisfy certain predicates in natural languages like English. Properties—that is, resemblance classes of ordinary particulars—do not, according to Extreme Resemblance Nominalists, ground character. The resemblance class of red things does not ground the character of red things. Properties are, however, implicated in a general account of attribute agreement. The resemblance class of red things is implicated in the Extreme Resemblance Nominalists' claim that two things are alike in some respect if and only if they are both members of some single resemblance class.

Moderate Resemblance Nominalists, on the other hand, claim that a red apple has the character that it does because it is related somehow to some tropes, particulars that are also properties. For Moderate Resemblance Nominalists, the (shareable) property of *being red* is a resemblance class made up of tropes rather than ordinary particulars. The property of *being red* is just the class of *redness* tropes; the property of *being apple-shaped* is the class of *apple-shape* tropes. These properties do not ground character (that is the job of tropes, after all), but they are implicated in a general account of attribute agreement, just as with Extreme Resemblance Nominalism. The precise nature of Moderate Resemblance Nominalism will become clearer in Section 8.2, where we turn our attention to tropes. However, the worries we will raise here beset both varieties of Resemblance Nominalism.

It should not be difficult to see that Extreme Resemblance Nominalism is faced with Class Nominalism's extensionality problems, both the problem of contingent predication and the co-extensive property problem. The reason they inherit these problems is that Resemblance Nominalism's theory of properties ensures that two properties are identical if they have the same resemblance class. Resemblance Nominalism entails the Resemblance Theory of Properties:

**Resemblance Theory of Properties.** Properties *P* and *Q* share the same resemblance class if and only if *P* and *Q* are identical.

For Resemblance Nominalists, what it is to have a property is simply to resemble certain things (the resemblance class for that property). If what it is to have *P* is the same as what it is to have *Q*, then *P* and *Q* must be identical. This means that no property can ever change its extension or have a different extension in different possible situations, since a difference in extension or resemblance class is *ipso facto* a difference in property identity. Similarly, it must be impossible for two distinct properties to share the same resemblance class.

The introduction of tropes enables Moderate Resemblance Nominalists to mitigate these two problems without embracing *possibilia* (merely possible things). For example, even if the class of red things and the class of round things were co-extensive, the class of *redness* tropes and the class of *roundness* tropes would be different (in fact, disjoint). Thus, the co-extensive property problem seems to be solved, at least in part. The only problematic cases remaining would be cases of co-extensive properties of tropes. For example, suppose that there are shape tropes that characterize plane surfaces: *circularity* tropes, *triangularity* tropes, and so on. Suppose that all of the *triangularity* tropes are also *right-triangularity* tropes (tropes grounding the character of *being a right triangle*) because it just so happened that all the triangular shapes were also right triangles. In such a possible scenario, the property of *being triangular* and the property of *being right-triangular* would be the same because the resemblance class of *triangularity* tropes would be identical to the resemblance class of *right-triangularity* tropes.

Tropes can also help with the problem of contingent predication, since tropes ground the characters they do essentially. A *triangularity* trope, for example, cannot become a *squareness* trope. When a concrete particular changes from being triangular to being square, this change is associated with the destruction of one trope and the creation of a new one. The tropes themselves do not change. However, this example brings out the fact that Moderate Resemblance Nominalism still has a problem with contingency and change, a problem we can call the problem of contingently existing instances. If, for example, at one time or in one possible scenario there are seven *scarlet* tropes, and at another time or in another possible scenario there are eight *scarlet* tropes, then Moderate Resemblance Nominalists are forced to recognize two distinct properties of *being scarlet*: one grounded in similarity to the seven tropes, and the other grounded in similarity to the eight tropes. This is clearly the wrong answer. There is only one property of *being scarlet*, a property that can vary in its set of instances from one time to another or one possible scenario to another.

TWO ADDITIONAL PROBLEMS: COMPANIONSHIP AND IMPERFECT COMMUNITY Nelson Goodman (1951) pointed out two additional problems for Resemblance Nominalism, problems that are not shared with Class Nominalism. These are the problems of companionship and imperfect community. Given Carnap's definition of resemblance class (Def 8.4.1), and the Resemblance Nominalist theory of properties, we get too few properties (the companionship problem) and too many properties (the problem of imperfect community).

THE COMPANIONSHIP PROBLEM Consider two properties *P* and *Q* that are distinct and have different resemblance classes, but *A*, the resemblance class of *P*, is a proper subset of *B*, the resemblance class of *Q*. That is, every instance of *P* is an instance of *Q*, but not vice versa. For example, *P* might be the property of *being square* and *Q* the property of *being rectangular*. In such a case, let's say that *Q* is a companion of *P*. If Resemblance Nominalism is true, then what it is to be *P* is to resemble all of the members of *A*. However, all of the members of *B* also resemble the members of *A*, since they all resemble each other and every member of *A* is also a member of *B*. To resort to our example, all of the rectangles do resemble all of the squares, since they resemble all rectangles and the squares are rectangles. Thus, it seems that *every* member of *B* should count as having the property *P*. But then the instances of *P* should include all of *B*, not just the members of the subset *A*. In other words, we would be forced to say that all of the rectangles are square! This would mean that the resemblance class of *P* is actually *B*, not *A*, making the properties *P* and *Q* identical after all.

The obvious solution of this problem is to tweak Carnap's definition of a resemblance class. Let's recall Carnap's original definition:

**Def D8.4.1 Resemblance Classes (Carnap).** A class of objects *X* is a *resemblance class* if and only if (i) each member of *X* resembles every other member of *X*, and (ii) nothing outside of *X* (that is, no non-member of *X*) resembles every member of *X*.

The problem is that the class of square objects cannot satisfy this definition, since it fails condition (ii): some things outside the class of squares resemble all of the squares. In particular, the rectangles do. Why do we need condition (ii) at all? Why not simply drop it? If we dropped condition (ii) we would get too many resemblance classes and thus too many properties. Suppose that the shape of Times Square in New York is in fact square. Now consider the class of square things that are not identical to the shape of Time's Square. If we simply dropped condition (ii), then this class would also be a resemblance class, and *being a square that isn't identical to Time's Square* would be a property. Going down this route will result in a super-abundance of properties of the kind we wanted to avoid.

A better solution is to observe that the squares resemble each other more closely than any of the non-square rectangles resemble them. If we can replace the single, simple, binary *resemblance* relation with a family of *resemblance* relations (resembling to degree *d*, for some class of degrees) or a *comparative resemblance* relation (such as *x*'s resembling *y* more than it does *z*), we can solve the companionship problem:

**Def D8.4.2 Resemblance Classes (Comparative).** A class of objects *X* is a *resemblance class* if and only if (i) each member of *X* resembles every other member of *X* to a certain degree, and (ii) nothing outside of *X* (that is, no non-member of *X*) resembles every member of *X* to that same degree.

THE PROBLEM OF IMPERFECT COMMUNITY However, there is an additional problem for Resemblance Nominalism, the problem of imperfect community. Suppose that there are three natural properties, *being large*, *being green*, and *being cubical*. Now consider the class of things that have at least two of these three properties. This class, call it At-Least-Two, will contain particulars of the following kinds:

Large Green Cubes
Large Non-Green Cubes
Large Green Non-Cubes
Small Green Cubes

Intuitively, there is no real or natural or "sparse" property that all of these things have in common. Nonetheless, the class At-Least-Two does satisfy the comparative definition of a resemblance class: any two members of At-Least-Two do resemble each other to a certain degree (by having two or three of the properties in common), and nothing outside the class resembles every member of the class to that same degree. At least, it seems plausible that we can find cases like this that do satisfy that condition.

Tropes help with the problem of imperfect community to some degree. In the case that we just considered, there would be three different resemblance classes, the class of *largeness* tropes, the class of *greenness* tropes, and the class of *cubicality* tropes. None of these classes form an *imperfect community* in the way that At-Least-Two does.

However, David Manley has pointed out (2002) that tropes alone do not entirely solve the problem of imperfect community, as long as all tropes are tropes of fully specific, determinate properties. Consider the class that includes all of the *pink*, *baby-blue*, and *dark-purple* tropes. These tropes all have two of three generic properties: the property of *being pale* (pink and baby-blue), the property of *being bluish* (baby-blue and dark purple), and the property of *being reddish* (pink and dark purple). Thus, they all resemble each other to a certain degree, and nothing outside the class resembles all of the tropes in the class to the same degree. Thus, this class also constitutes a comparative resemblance class, even though the tropes in the class do not (intuitively speaking) share any natural property in common.

The only solution available to Moderate Resemblance Nominalists involves positing a large number of new tropes, namely, generic tropes (tropes of *mere paleness* or *mere bluishness* or *mere reddishness*). If there are tropes of generic color then we could deny that the *pink*, *baby-blue*, and *dark-purple* tropes resemble each other at all: it is only the *pale*, *bluish*, and *reddish* tropes that resemble each other. Pink and baby-blue ordinary particulars resemble each other only by having *paleness* tropes, not by having *pink* or *baby-blue* tropes. However, the introduction of generic or determinable tropes comes with a high price tag, in two ways. First, this move requires a huge increase in the number of tropes. Each case of a colored concrete particular will contain a large number, perhaps even an infinite number, of color tropes, one for each kind of color property, however specific or generic it may be. Second, the move introduces a large number of necessary connections, for which there is no further explanation or ground. You cannot have a *red* trope by itself: it must always be accompanied by tropes of mere color, mere reddishness, and so on. However, a theory should ideally minimize the number of such brute necessities connecting separate entities (**PMeth 1.2**).

There is another route available to Resemblance Nominalists, whether Extreme or Moderate. This route involves a further modification of the definition of resemblance classes. What we want is for the members of the resemblance class to resemble each other in a certain way. The members of an imperfect community do not resemble each other in the same way. *Dark-purple* tropes resemble *baby-blue* tropes in an entirely different way from that in which the *baby-blue* and *pink* tropes resemble each other.

However, introducing new relations of resemblance in certain ways threatens to collapse Resemblance Nominalism into Ostrich Nominalism. Purple things and baby-blue things resemble each other by way of being bluish, while pink and baby-blue things resemble each other by way of being pale. It seems that to resemble each other in a certain way is just to have a certain property in common. If so, there are just as many ways of resembling things as there are natural properties. Hence, the Resemblance Nominalist will have no general explanation of why things resemble things, nor any general explanation of what it is to have a property in common. To have a property in common is just to resemble each other in the way that corresponds to having that very property. This is no explanation at all, and so this strategy collapses into Ostrich Nominalism.

There is, however, another way to go. We could abandon the idea that *resemblance* is a binary relation between two things. Instead, we could think in terms of the *collective* resemblance of a large number of things. The pale things are collectively in a relation of a certain degree of *resemblance*. However, the things that are pink, baby-blue or dark purple do not collectively resemble each other at all (or only to the minimal degree that all colors resemble each other). This is the route taken by David Lewis, who combines both collective and comparative resemblance:

**Def D8.4.3 Resemblance Classes (Lewis).** A class of objects *X* is a *resemblance class* if and only if (i) there are some *z*'s such that the members of *X* collectively resemble each other and do not likewise resemble any of the *z*'s, and (ii) *X* contains every *y* such that *y* and the members of *X* collectively resemble each other and do not likewise resemble any of the *z*'s.

Instead of using a single comparative resemblance relation, we could also make do with a family of collective resemblance to a certain degree. Then we could say that a class *X* is a resemblance class just in case the members of *X* resemble each other collectively to a certain degree, and there is no *y* such that the members of *X* and *y* collectively resemble each other to the same degree.

This definition solves both the companionship problem and the problem of imperfect community, but at the cost of introducing a new, variably polyadic relation. By 'variably polyadic' we mean that the very same relation can hold between the members of a class *X* and the *z*'s, no matter how many or how few things there are in either case.[1]

Besides these three extensionality problems, there are some further worries for Resemblance Nominalism. We turn now to these.

## ADDITIONAL OBJECTIONS TO RESEMBLANCE NOMINALISM

1. The Hochberg-Armstrong objection. The first argument against Resemblance Nominalism comes from Herbert Hochberg (1999: 50–54) and David Armstrong (2004: 43–44). Consider these two propositions:

(8) *A* and *B* are similar fundamental particulars.
(9) *A* and *B* are distinct fundamental particulars.

What makes each of these statements true? For Resemblance Nominalists, the answer must be: the pair of *A* and *B*. In fact, not only do (8) and (9) have the same truthmaker in the actual world, this is the only truthmaker that either of them could have. In contrast, UP-Realists can say that (8) is made true by the fact that there is some universal *U* that both *A* and *B* instantiate and that statement (9) is made true by the pair {*A*, *B*}. In other words, (8) is made true by the two ordered instantiation pairs <*A, U*> and <*B, U*>, while (9) is made true simply by the pair {*A*, *B*}.

(8) and (9) are clearly distinct in content. (8) could be true even if *A* and *B* were identical, and (9) could be true even if *A* and *B* were dissimilar. Neither proposition can be deduced a priori from the other. However, propositions that both have the same truthmaker and could have only that same truthmaker are metaphysically equivalent, that is, true and false in exactly the same situations, since it is part of the very essence of a proposition that it be true whenever one of its truthmakers exist and false if no truthmaker for it exists.

Resemblance Nominalists can reply by insisting on a distinction between a priori or *conceptual* equivalence and *metaphysical* equivalence. Although the conceptual contents of (8) and (9) do not guarantee their equivalence a priori, if as a matter of fact both (8) and (9) are true, then they must be true in exactly the same possible situations. This answer assumes two things. First, two entities that are numerically distinct in actuality are necessarily distinct, and second, two tropes that are similar are necessarily similar. Either of these claims could be challenged, but we will accept both of them for the sake of argument here.

The Hochberg-Armstrong objection suggests that there is something unsatisfying about this response. What's relevant about the difference between (8) and (9) is not just that they are not a priori equivalent. Instead, each attributes a different *relational property* to the pair of *A* and *B*. (8) attributes the property of *distinctness*, and (9) attributes the property of *similarity*. These properties are clearly different, as they have quite different extensions, even in the actual world. Some pairs of distinct things are exactly similar to each other, and some are not. Arguably, some "pairs" of exactly similar things are identical to each other, and some are not. Given the difference between the two properties involved, it seems natural to demand two distinct truthmakers, based on One Truthmaker per Fundamental Property:

**PTruth 1 One Truthmaker per Fundamental Property.** If $p$ is the true predication of a fundamental property $P$ to $x_1$ through $x_n$, and $q$ is the true predication of a different fundamental property $Q$ to the same things $x_1$ through $x_n$, then $p$ and $q$ have distinct truthmakers.

If we accept One Truthmaker per Fundamental Property (which we showed, in Chapter 3, to follow from a plausible definition of 'fundamental property'), then we will need one truthmaker as the ground of the distinctness of $A$ and $B$ and another for their exact similarity. Resemblance Nominalists could meet this demand by supposing that the pair of $A$ and $B$ is the truthmaker for (9) but not for (8). This would require the postulation of a new entity as the truthmaker for (8). This additional entity would have to be a *resemblance* trope. If $A$ and $B$ are distinct and exactly similar, then the resemblance trope exists and has a particular intrinsic feature corresponding to $A$ and $B$'s standing in the *similarity* relation.

However, once Resemblance Nominalists posit such resemblance tropes, they immediately fall into an infinite hierarchy of such tropes. If they adopt resemblance tropes, they must now give an account of the fact that two of these resemblance tropes are both of the *resemblance* kind. Given the constraints accepted so far, this will force Resemblance Nominalists to posit yet another, higher-order resemblance trope, standing between the two, resembling *resemblance* tropes. The same argument will force the positing of third-order and fourth-order tropes, and so on ad infinitum.

This infinite series comes at a very high cost in terms of quantitative economy. Moderate Resemblance Nominalists require an infinite number of tropes whenever there are two similar but distinct particulars. Does this also incur a high cost in terms of qualitative economy? Must Moderate Resemblance Nominalists posit an infinite number of kinds of tropes? It's hard to say definitively, but the various orders of tropes (first-, second-, third-, and so on) seem to involve different kinds of trope.

More importantly, the hierarchy of resemblance tropes constitutes an infinite regress, as Bertrand Russell (1912) argued. The existence of each resemblance trope $R$ between $A$ and $B$ is supposed to be the ground of the truth that the two relata are similar, but this depends on the fact that the resemblance trope $R$ is in fact of the resemblance kind, but this latter fact is supposed to be grounded in the further fact that the trope $R$ is connected by resemblance tropes to the other members of the class of resemblance tropes. The truth of $A$ and $B$'s resemblance is never finally grounded. Its grounding is perpetually deferred.

There is a second way that Moderate Resemblance Nominalists can meet the Hochberg-Armstrong objection. They could suppose that ordinary particulars are bundles of two sorts of things, namely, tropes plus *bare particulars*. Each bundle would contain many tropes but just one bare particular. The bare particulars have just one "job" to do: that of distinguishing one bundle from another exactly similar bundle. The bare particulars would be the ultimate source of numerical distinctness. Suppose, for example, that we have two tropes of scarlet red, $S_1$ and $S_2$, and suppose that these tropes belong to two different bundles, $B_1$ and $B_2$, each containing a different bare particular, $P_1$ and $P_2$. In this situation, the distinctness of $S_1$ and $S_2$ is grounded in the distinctness of $P_1$ and

$P_2$. It is their connection to two bare particulars that makes $S_1$ and $S_2$ different. Thus, the truthmaker for the proposition that $S_1$ and $S_2$ are exactly similar is the pair $\{S_1, S_2\}$, while the truthmaker for the proposition that $S_1$ and $S_2$ are distinct is the quadruple $\{P_1, S_1, P_2, S_2\}$. We don't have an adequate ground of the distinctness of $S_1$ and $S_2$ without making reference to $P_1$ and $P_2$. We will consider this Bare Particular Theory (9.1T.1A.2A) in more detail in the next chapter.

There is a third response that Resemblance Nominalists could make to the Hochberg-Armstrong objection. They could posit *distinctness* tropes as the ground for the truths like (9), leaving the pair $\{A, B\}$ as the truthmaker for the resemblance truth (8). In other words, what makes it true that $A$ and $B$ are distinct particulars is some distinctness trope $D$ that connects them. There are two worries here. First, one might wonder how it is that a trope could connect two things if they are not already, prior to the connection, two things. The trope $D$ is what is supposed to make $A$ and $B$ two, but how can it do that? If $D$ is a binary trope by its very nature, it would seem that the twoness of its relata would have to be in existence in a way that is ontologically prior to $D$'s actually connecting them. Second, if we need $D$ to ground the distinctness of $A$ and $B$, then that must be because neither $A$ nor $B$ nor the combination of the two can do this. That means that $D$ has to be distinct from $A$, from $B$, and from the pair if it is going to be able to perform its ontological job. But this will require more distinctness tropes—one to distinguish $D$ from $A$, another to distinguish it from $B$, and perhaps a third to distinguish it from the pair. In addition, each of these distinctness tropes will require still more distinctness tropes for them to do their job. We seem to be generating an infinite regress of tropes.

A fourth response to the Hochberg-Armstrong objection is to deny that propositions like (9) require a truthmaker at all. Some philosophers (including Ludwig Wittgenstein in his *Tractatus Logico-Philosophicus*) have argued that there are no identity or distinctness propositions at all or if there are that they are only about language, not about the world. This approach dissolves the need to ground (9), but it comes at a high cost. We would have to assume that it is absolutely necessary (even a logical truth) that the world contains exactly the fundamental entities that it does contain. We would have to rule out the possibility that any fundamental thing is contingent in its existence, or that there could have been more or other fundamental things than there in fact are. Whether we consider the fundamental things to be physical particles or fields, or more complex things like people or organisms, it seems obvious that many, if not all of them, exist contingently. If so, then truths of the form 'There exist at least $n$ distinct things' should be treated as genuine assertions.

However, even if it is the case that there are truths about distinctness, we still have to ask whether these truths are among the truths that require truthmakers. If we were to accept Truthmaker Maximalism, we would have to find truthmakers for all truths, including truths of distinctness. However, we must consider restricted versions of Truthmaker Theory. Two plausible restrictions seem most relevant, namely, the restriction to positive truths, and the restriction to contingent truths.

Are truths about distinctness positive or negative truths? Consider (10) and (11):

(10) Mt. Everest is identical to Mt. Everest.

(11) Mt. Everest is distinct from K4.

Which of these is positive and which negative? It seems that 'identity' and 'distinctness' refer to contradictory properties, in the sense that the falsity of (10) corresponds to (12), and the falsity of (11) to (13):

(12) Mt. Everest is distinct from Mt. Everest.
(13) Mt. Everest is identical to K4.

If so, we could either take the logical form of (10) to be the denial of (12) or the logical form of (11) to be the denial of (13). That is, we could take (10) to be asserting the fact expressed in (14) or we could take (11) to be asserting the fact expressed in (15):

(14) Mt. Everest is not distinct from Mt. Everest.
(15) Mt. Everest is not identical to K4.

Suppose we took the identity statements to be positive. What would be the truthmaker of a truth like (10)? There would seem to be two possibilities. First, the truthmaker could be just Mt. Everest itself, or second, it could be some nexus between Mt. Everest and the property of *self-identity*. Either way, if identity statements were positive, then distinctness truths (like (11)) would be negative. (15) would be asserting the non-existence of a truthmaker for (13). However, both Mt. Everest and K4 do exist and both have the property of *self-identity*, so what possible truthmaker for (13) is in fact missing?

We seem to be forced to consider distinctness truths to be positive and identity statements to be negative. There is some truthmaker for (13), something that grounds the distinctness of Mt. Everest and K4. (12) is false because there is no ground for the distinctness of Mt. Everest from itself.

What about contingency-based restrictions of Truthmaker Theory? If statements of distinctness are necessary truths, they may not need truthmakers, even if they are positive. Are all distinctness truths necessary truths? Consider again (11):

(11) Mt. Everest is distinct from K4.

It seems unlikely that Mt. Everest could be identical to K4. If they were identical, they would be one and the same thing, and it seems plausible that any one thing must be one thing in every possible world in which it exists (see Kripke 1980 and Williamson 1996 for details). However, the fact that Mt. Everest and K4 couldn't be identical is not enough to make (11) a necessary truth, since we have to ask what would be the case if either Mt. Everest or K4 hadn't existed. It is at least arguable that in the absence of Mt. Everest or K4, (11) wouldn't have been true. If so, and if there are fundamental particulars that exist contingently (without necessity), then truths of distinctness involving such contingent particulars will not be necessary truths.

In addition, it is far from clear that no necessary truth has a truthmaker. It may be that trivially true statements or statements true by something like a stipulation, like 'All bachelors are unmarried', have no truthmaker. However, many necessary truths are far from trivial. Among the class of substantial necessary truths, some seem more fundamental than others. Truthmaker Theory can explain this fact by identifying the fundamental truths with those asserting the existence of a single, necessarily existing truthmaker.

2. Explaining the Facts of Resemblance. The second argument against Resemblance Nominalism starts with the observation that this view seems committed to large number of distinct, fundamental resemblance relations. How so? As we have seen, Resemblance Nominalists must in some way distinguish between different degrees of resemblance. This can be done in either of two ways. First, we could suppose that there are a large number of fundamental *resemblance* relations, each corresponding to a certain degree of resemblance, forming together a linear ordering of degrees. This is Gonzalo Rodriguez-Pereyra's (2002) approach. He posits an infinite number of degrees, one for each natural number. In addition, we have to suppose that it is metaphysically impossible for a group of things to resemble each other to different degrees at once. The second option is to suppose that there is a single, comparative resemblance relation: the relation of the *x*'s resembling each other more than they resemble the *y*'s. This was David Lewis's (1999a: 14) proposal.

It seems that Lewis's approach, employing a single comparative resemblance relation, is preferable in light of Ockham's Razor (**PMeth 1**). However, Lewis's proposal also brings with it a large number of necessary connections between distinct fundamental facts. For example, suppose that the members of *X* resemble each other more than they do the *z*'s, and suppose that *Y* is a proper subset of *X*. Clearly, it must be true that the members of *Y* also collectively resemble each other more than they do the *z*'s, yet the two collective resemblance-facts are equally fundamental, according to Lewis's theory. Neither can be grounded in the other.

Lewis's theory entails a second set of necessary connections, this time involving subsets of the comparison classes. Again, suppose that the members of *X* resemble each other more than they do any of the members of *Y*, and suppose that *Z* is a subset of *Y*. Again, it must be the case that the members of *X* resemble each other more than they resemble any of the members of *Z*, but this will have to be a brute necessity connecting the two fundamental facts.

In contrast, Realists have a ready explanation for both of these facts. If members of the *X* resemble each other in a way that they do not resemble the *z*'s, then the members of *X* must each instantiate some universal that is not instantiated by any of the *z*'s. But of course it then follows that all of the members of any subset of *X* must also instantiate some universal that is not instantiated by the *z*'s, and that the members of *X* instantiate a universal that is not instantiated by any member of a subset of the *z*'s.

There is a similar class of brute necessities required by Rodriguez-Pereyra's account. First, as we have seen, it has to be impossible for two things to resemble each other to two different degrees. Second, resemblance to any degree must be symmetric; that is, if *x* resembles *y* to degree *n*, then *y* must resemble *x* to degree *n*. Third, if the members of *X* resemble each other to a certain degree, and the members of *Y* resemble each other to a certain degree, and there are some things in the intersection of *X* and *Y*, then the things in that intersection must resemble each other to a greater degree than do the things that are in *X* but not *Y*, or in *Y* but not *X*. None of these things can be explained in terms of the instantiation of universals, and so they must be brute necessities. But, as we have seen, Ockham's Razor directs us to minimize such necessities (**PMeth 1.2**).

Resemblance Nominalists, in addition, must posit a large number of metaphysically fundamental similarity facts. If there are *n* distinct particulars (whether ordinary particulars or tropes), then there must be at least $2^n$ fundamental facts about the comparative degree of similarity among sets of particulars (using Lewis's definition). In contrast, if there are $n$ particulars and $m$ universals, Realists require only $n \cdot m$ fundamental facts, a much smaller number (if $n$ is at all large). In fact, Realists who embrace Totality Fact Maximalism require only $m$ fundamental facts, one for each universal. If there are only $m$ universals, then Realism entails that there can be only $m$ natural classes. However, Resemblance Nominalists have no way of putting an upper bound on the number of natural classes because any class could be a natural one, as long as its members resemble each other more than they do any non-member. Thus, Lewis's version of Resemblance Nominalism requires a large number of independent, fundamental resemblance-facts. This counts against Moderate Resemblance Nominalism, since it is a failure of quantitative economy (according to Ockham's Razor).

If one attempts to avoid some of these difficulties by replacing Lewis's definition of resemblance classes with a class of binary relations of resemblance to a degree (as in Rodriguez-Pereyra 2002), there are still a number of necessary connections left unexplained. For example, Rodriguez-Pereyra cannot explain why resemblance to a degree is symmetrical, that is, why, if $x$ resembles $y$ to a certain degree, $y$ resembles $x$ to that same degree. Realists can easily explain this, since to resemble to a certain degree can be analyzed in terms of the number of universals instantiated in common. In addition, Rodriguez-Pereyra cannot explain why, if the pair $\{a,b\}$ resembles the pair $\{c, d\}$ to a certain degree, the particular $a$ must resemble the particular $c$ to at least that same degree. In addition, since Rodriguez-Pereyra has an infinite number of degrees of resemblance, his theory will require an infinite number of fundamental resemblance-facts.

SIMILARITY BETWEEN PROPERTIES: A PROBLEM FOR REALISTS? When evaluating the relative simplicity of Resemblance Nominalism and UP-Realism, we have to take into account one more factor, namely, degrees of similarity between properties themselves. The property of *being red* resembles the property of *being orange* more than it resembles the property of *being green*. A cubical shape resembles the shape of a rectangular prism more than it does the shape of a cone. Some Moderate Resemblance Nominalists already have a three-place relation of comparative similarity ($x$ resembles $y$ more than it does $z$) in place. They could use this same relation to define comparative similarity among properties: property $A$ resembles property $B$ more than property $C$ if and only if the members of the resemblance class of $A$-tropes resemble the members of the resemblance class of $B$-tropes more than they do the members of the resemblance class of $C$-tropes. (Note, however, that this solution is not available to Resemblance Nominalists employing Lewis's definition of classes in terms of collective resemblance, since the members of the class containing both $A$-tropes and $B$-tropes will not resemble each other *collectively* in a way that they do not likewise resemble the $C$-tropes.)

Realists, in contrast, would seem to be forced to add a new fundamental relation, comparative similarity among universals, to their theory: the universal of REDNESS resembles the universal of ORANGENESS more than it does the universal of GREENNESS. If Realists take this route, then they increase their stock of qualitative primitives from two to three (*comparative resemblance* is added to *being a universal* and *being an instantiation pair*). This is a significant cost.

David M. Armstrong proposes that Realists provide some sort of reduction of comparative similarity to other fundamental terms. Realists might use a very sparse theory of universals, arguing that no two universals resemble one another at all. When two particulars resemble each other to some degree, this is always to be understood in terms of the two particulars (or their parts) instantiating or not instantiating the same universals. This would force Realists to deny that there are universals corresponding to the different colors or hues (like red, orange or green) or to the various shapes (cubical, conical, spherical). To be red must be to instantiate several, incommensurable color-universals which, when co-instantiated, constitute the color redness. Such a view does not square well with a common-sense view of our experience of these properties. In addition, it remains to be seen whether such a reduction can really be carried out.

As a second option, Realists could explain comparative similarity among universals by appealing to higher-order universals. If universals can instantiate other universals, then the similarity of two ground-level universals could be explained in terms of their sharing some third, higher-order universal. Perhaps two determinate shades, like REDNESS and GREENNESS, share a determinable universal, like SHADE OF COLOR. This would explain why REDNESS is more similar to GREENNESS than either is to CUBICALITY or any other non-color. To explain why REDNESS is more similar to ORANGENESS than to GREENNESS, we would have to introduce further higher-order universals, like the universal REDDISH COLOR, a universal shared by both REDNESS and ORANGENESS but not GREENNESS.

Third, Realists could move to a more abundant theory of universals, including generic universals (such as COLOR or REDDISHNESS) as well as specific universals (like PURE REDNESS) in their ontology. Comparative similarity could be a matter of how many universals two particulars share. An orange and a red sphere would be more similar to each other than either is to a green sphere because they share three universals (SPHERICALITY, COLOR, and REDDISHNESS), while each has only two in common with the green sphere (SPHERICALITY and COLOR).

Fourth, Realists could use a system of part-whole relations between universals to capture these similarity relations (this is an idea we will explore more fully in Chapter 10). If two universals are similar, we could suppose that one is a proper part of the other. Two universals are very similar when one contains nearly all of the parts of the other. This approach is very plausible when dealing with determinable quantities or qualities that vary in magnitude or intensity along a single dimension. So, the universal 2 GRAMS IN MASS could be supposed to have the universal 1 GRAM IN MASS as a part, with intermediate quantities containing the one and being contained by the other. In the case of qualities that vary continuously in more than one dimension, like color, we could suppose that each color property involves instantiating one universal for each dimension, such as hue, intensity, and brightness.

Let's summarize the problems for Resemblance Nominalism, along with the possible solutions to those problems. We've discussed five major problems: contingent predication, co-extensive properties, companionship, imperfect community, and the Hochberg-Armstrong objection. We've also considered four versions or repairs to Resemblance Nominalism: Concretism (with Worldbound Individuals), tropes (Moderate Resemblance Nominalism), degrees of resemblance, and a variably polyadic resemblance relation. Here is a table with the results:

**Table 8.1** Repairing Resemblance Nominalism

| | *Concretism with Worldbound Individuals* | *Tropes* | *Degrees* | *Polyadic Relation* |
|---|---|---|---|---|
| Contingent Predication | Yes | No | – | – |
| Co-extensive Properties | Yes, but with exceptions | Yes | – | – |
| Companionship | No | Only with generic tropes | Yes | Yes |
| Imperfect Community | No | Only with generic tropes | No | Yes |
| Hochberg-Armstrong | No | Only with bare particulars | No | No |

These solutions can be combined—no one excludes any of the others. For example, David Lewis's view incorporates Concretism, degrees of resemblance, and a variably polyadic relation, but rejects tropes. Gonzalo Rodriguez-Pereyra's theory has the same combination of solutions, although he replaces the variably polyadic relation with a binary relation between sets constructed ultimately from pairs of particulars. These theories dissolve all of the problems except the Hochberg-Armstrong objection, but at a significant ontological cost of adding merely possible things to our ontology and adding a large number of basic facts and of brute necessities.

Moderate Resemblance Nominalism with Concretism but without degrees of resemblance or a variably polyadic resemblance relation could handle all of the problems, but only at the cost of adding generic or determinable tropes and bare particulars. Generic tropes also bring a large number of brute necessities with them. For instance, every *color* trope has to be accompanied with a trope of some specific shade (hue and saturation), and every *triangularity* trope has to be accompanied with a trope of a specific triangular shape, like *equilaterality*.

## 8.2 Trope Theory

Let's return briefly to the extensionality problems for Class Nominalism. Suppose that Class Nominalists asserted that an ordinary object like the celestial body Pluto is actually a complex thing, a bundle of tropes. In other words, suppose Pluto is "built" from a number of tropes, each of which supplies Pluto itself with some sort of character. We could then identify shareable properties with classes of tropes. The property of *being spherical*, for example, would be a all the *sphericality* tropes. This is Trope Theory:

**8.2T Trope Theory.** Tropes directly ground the character of ordinary objects.

Those who reject tropes fall into two familiar camps, Nominalists and Realists:

**8.2A.1 Extreme Nominalism.** The character of ordinary objects is grounded neither in universals nor in tropes.

**8.2A.2 Classical UP-Realism.** The character of ordinary objects is grounded in universals and not in tropes.

To see how Trope Theory works as an account of attributes and similarity, recall RedSquare, RedCircle, and BlueCircle. Trope Theorists will say that RedSquare and RedCircle are similar because they each have a *redness* trope and that RedCircle and BlueCircle are similar because they each have a *circularity* trope. Further, Trope Theory offers a more serious account of character-grounding than does Extreme Nominalism, an account that mirrors Realists' in many ways. RedCircle is red *because* it is related to a *redness* trope. BlueCircle is circular because it is related to a *circularity* trope. It isn't difficult to see how these accounts of similarity and character-grounding can be generalized.

Trope Theory comes in two varieties depending on whether one plumps for universals. The difference concerns, as we have already seen, the sense in which distinct tropes are particularizations of the same shareable property. Trope Realists believe that tropes are related to universals in some way, and that these relations determine when tropes are particularizations of the same shareable property. Resemblance among tropes is a consequence of relations to universals. Moderate Resemblance Nominalists, on the other hand, take resemblance among tropes to be metaphysically fundamental and think of shareable properties as classes of exactly resembling tropes. (As we have seen, Trope or Moderate Resemblance Nominalism thus faces some of the extensionality problems for Class Nominalism, as well as the Hochberg-Armstrong objection to Resemblance Nominalism.)

More importantly, though, there are two different views of the nature of tropes themselves, and thus there are two types of Trope Theory that arise given different understandings of the nature of tropes.

Before we get to these distinctions, it is important to say a bit about why Trope Theorists think we need tropes in the first place. We have already seen one motivation, at least for Trope Nominalism: it is a way to do without universals while being more Realist about character-grounding. In this way, Trope Theory holds out hope for avoiding the difficult questions facing the postulation of universals while accruing some of Realism's benefits. Further, because of this more serious account of character-grounding, Trope Theory avoids the problem of co-extensive properties (the second extensionality problem discussed above). More generally, though, Trope Theorists have argued that we need tropes to make sense of causation, since tropes are plausible candidates to serve as the relata of causal relations. Relatedly, tropes can serve as the immediate objects of perception. What we see when we look, for example, at an apple, is *this particular red*, *this particular shape*, not *redness* in general or *apple-shapedness* in general. Finally, Trope Theory suggests an account of substances as bundles of tropes. We will return to these putative advantages in later chapters, but this should suffice to help make clear why believing in tropes might be initially attractive.

### 8.2.1 Two varieties of Trope Theory

There are two ways of thinking about tropes, ways that have not until recently been distinguished in the literature.[2] We can get at the distinction by considering this question: Do tropes have the character they ground? If they do, then tropes are modular tropes. If they do not, then tropes are modifying tropes.

**Def D8.5 Modifying Trope.** A *modifying trope* is a trope that does not have the character it grounds.

**Def D8.6 Modular Trope.** A *modular trope* is a trope that does have the character it grounds.

An example may clarify the distinction. Consider again RedSquare. RedSquare, according to Trope Theory, is uniquely related to a *redness* trope and a *squareness* trope. Call these tropes, '$r_1$' and '$s_1$', respectively. Since tropes ground the character of ordinary objects like RedSquare, $r_1$ grounds the redness of RedSquare, and $s_1$ grounds the squareness of RedSquare. We might say, with many Trope Theorists, that $r_1$ *is* the redness of RedSquare and that $s_1$ *is* the squareness of RedSquare. If $r_1$ and $s_1$ are modifying tropes, then $r_1$ is not itself red, and $s_1$ is not itself square. *Modifier tropes therefore ground character they do not themselves have.* If, on the other hand, $r_1$ and $s_1$ are modular tropes, then $r_1$ is itself red, and $s_1$ is itself square. That is, *modular tropes ground character they themselves have.*

With this distinction between modifying and modular tropes in place, we get a twofold distinction among types of Trope Theory:

**8.2T.1T Modifying Trope Theory.** Trope Theory is true, and tropes are modifying tropes.

**8.2T.1A Modular Trope Theory.** Trope Theory is true, and tropes are modular tropes.

As it turns out, the distinction between modifying and modular tropes affects Trope Theory in some interesting and important ways. In particular, modifying and modular tropes have different sorts of intrinsic character and also ground character in ordinary objects differently. These differences mean that Modifier Trope Theory and Modular Trope Theory face different obstacles. We turn to some of these obstacles presently, and others will emerge in the next chapter. Along the way, we'll examine in more detail the differences between modifying and modular tropes.

### 8.2.2 Some troubles for Trope Theory

Trope Theory is thought by many to be a plausible theory of properties, but shedding light on the distinction between modifying tropes and modular tropes exposes new issues. For example, suppose tropes are modifying tropes. Then it isn't clear that tropes are able to be the relata of causal relations or the immediate objects of perception. To play those roles, a thing must actually *exemplify* the relevant property. Modifying tropes are no more fit to be causal relata or objects of sense perception than are universals. What is needed are tropes that have the properties they ground; that is, to play these roles Trope Theorists must go in for modular tropes.

But there is an even more serious worry, one that constitutes a sort of dilemma for Trope Theory. We can put the dilemma in roughly the following way. If tropes are modifying tropes, then Trope Theory is explanatorily identical to Realism but quantitatively less economical. Therefore, one ought to be a Realist. On the other hand, if tropes are modular tropes, then on pain of violating very plausible metaphysical principles, tropes are just the ordinary objects of Extreme Nominalism. Therefore, one ought to just be a Nominalist without tropes, like an Extreme Resemblance Nominalist, for example. The upshot of the dilemma is that Trope Theory is in danger of collapsing into Realism on the one hand or Extreme Nominalism on the other.

We turn presently to the modular trope horn of the dilemma; the modifying trope horn will be developed in the next chapter. Suppose, then, that Modular Trope Theory is true. Modular tropes, recall, have the feature they ground, but the idea nonetheless is that modular tropes have *just one* type of character. A *redness* modular trope is *just* red; a *sphericality* modular trope is *just* spherical. Tropes, on this conception, are simply thinly charactered individuals. They are like the Extreme Resemblance Nominalists' ordinary objects except for having only one dimension of character.

To see why this conception of tropes is troubling, consider the following examples of "thickening principles" (see Garcia 2016):

**Color Thickening.** Every colored object has a definite shape.

**Shape Thickening.** Every shaped object has a definite size.

These principles are exceedingly plausible. But these principles entail that the *redness* modular trope of a red apple cannot be *just* red. That modular trope would, by Color Thickening, have to have some definite shape as well, presumably the shape of the apple itself. But then Shape Thickening would entail that our *redness* modular trope has a definite size as well, presumably again the size of the apple itself. Our *redness* modular trope seems to be transforming into the apple as it is conceived by Extreme Nominalists! That is, the *redness* modular trope is something that resembles things in various respects, has various features, and so on. These facts will need to be explained. Presumably the explanation will be in non-Realist terms, since Trope Theorists reject the Realist account of the redness of the apple. (Plausibly, even if one is a Modular Trope Theorist who plumps for universals, one doesn't think that the redness of the *redness* modular trope is a result of the trope's exemplifying the universal REDNESS. But if one did think that, then the *redness* modular trope begins to look, metaphysically speaking, like the apple as well!) The upshot of this is that it looks as though, given Modular Trope Theory, each object will have but one modular trope, in which case one wonders why one needs the trope *and* the object, rather than just the object itself.

The proponent of Modular Trope Theory could reply by saying that thickening principles apply only at the level of ordinary objects. In other words, they require only that every object with a color modular trope have a shape modular trope as well, that every bundle with a shape modular trope has a size modular trope, and so on. Intuitions about thickening were made for ordinary objects, not for tropes. While it may be hard to imagine an object that is colored but not shaped, it is equally hard to imagine certain things one finds in contemporary physics, like the 11-dimensional strings of string theory.

However, the reason why our cognitive powers give out in the string theory case is that the world of our experience is a three-dimensional one (four-dimensional if one counts time). We don't experience these other seven dimensions, even given that they exist. Since our conception of the spatiotemporal universe is plausibly derived from our experience of that universe, we have an easy explanation of why we cannot conceive of things that exist in more than three dimensions. But modular tropes are not this way. Our experience is sufficient to supply a conception of colored, shaped, and sized things. And those concepts seem irrevocably tied together, a tie that is captured by thickening principles.

The upshot of all this is that if one thinks that modifying tropes are unmotivated because they are unfit to be the relata of causal relations and the immediate objects of sense perception, then Trope Theory is already in serious jeopardy. For the very idea of modular tropes is, at the very least, seriously problematic. (We will consider a possible exception in Section 9.3.) Nonetheless, in the next chapter we will develop the other horn of the dilemma for Trope Theory, the one to do with modifying tropes.

## 8.3 Conclusion

The most viable form of Nominalism would seem to be Resemblance Nominalism. We have seen the competition reduced in effect to two competitors: Resemblance Nominalism and UP-Realism. Deciding between these two is not an easy matter. Each theory has one basic relation: *instantiation* in the case of Realism, and *resemblance* in the case of Resemblance Nominalism. Nominalists have a slight advantage in terms of quantitative economy, since they do not have to acknowledge the existence of universals. At the same time, Realists have an account of the essential laws of resemblance that Nominalists must take as brute facts. As we have seen, in order to meet the problems of imperfect community, co-extensive properties, and companionship, Nominalists must replace the simple, binary relation of resemblance with a variably polyadic relation of collective, comparative similarity that seems a great deal more cumbersome than the Realist's binary instantiation relation.

The next chapter, in which we consider theories of the constitution of particulars, may provide further considerations both for and against Realism to add to the balance.

### Notes

1 Gonzalo Rodriguez-Pereyra (2002) solves the problem while keeping resemblance as a binary relation. However, he allows the relation to connect not just two particulars, but also two pairs of particulars, or two pairs of pairs of particulars, or two pairs of pairs of pairs, and so on. However, the relation's holding between two pairs is never logically entailed by its holding between

the members of those pairs, so it is still (metaphysically speaking) a relation holding between variable pluralities of particulars.

2 The material about these two varieties of Trope Theory, both in this chapter and the next, derives from work by Robert Garcia (2015, 2016). We thank him for permission to use his arguments here and for many helpful conversations.

# 9

# Particulars and the Problem of Individuation

## 9.1 Introduction

In the last two chapters we have been considering some important questions about the nature of properties, including their relations to the powers of things. In this chapter, we turn our attention to the question of the relation of properties to particulars. We consider, first, facts (the shorter Section 9.2) and, second, substances (the longer Section 9.3). In Section 9.2, we will consider three theories of facts: as tropes, as states of affairs, and as nexuses between particulars and universals. We note that in each case, facts turn out to be particulars of a kind. This raises a question about how to understand the ordinary particulars of everyday life and science: are they constituents of facts or facts of a certain kind or some third thing? These ordinary particulars are richly charactered and potentially changeable things—things known as 'substances' in ancient and medieval philosophy.

Investigating the question of substances will take up Section 9.3, where we consider two accounts about the relationship between substances and properties, namely, Relational and Constituent Ontology. According to Constituent Ontology, properties are literally parts of substances they characterize. According to Relational Ontology, in contrast, properties are separate things, extrinsically related to their instances. After considering the Extrinsicality Objection to Relational Ontology, we will examine two versions of Constituent Ontology, Bundle Theory (according to which substances are mere "bundles" of properties) and Substrate Theory (according to which substances have both properties and some fundamental particular as parts).

The questions that we will be investigating in this chapter presuppose a set of possible answers to the questions about properties that we investigated in Chapters 7 and 8. Those who embrace Extreme Nominalism, whether in its Ostrich or Reductive (e.g., Resemblance) forms, need not worry about the internal metaphysical structure of things. On those views, all substances and facts (if there are such things as facts) are metaphysically

*The Atlas of Reality: A Comprehensive Guide to Metaphysics*, First Edition.
Robert C. Koons and Timothy H. Pickavance.

simple. In contrast, those who believe either in tropes or universals (or both) will have to confront the issue about the relationship between those properties and particulars.

## 9.2 Facts

The notion of a fact has already played an important role in our discussion. The most natural view of truthmakers is that they are facts. Simple subject-predicate sentences seem to be made true by a fact that some particular instantiates some property. Further, tropes seem closely allied to facts, since many facts seem to be simply a particularization or an instance of a property. This can be seen most clearly in causal statements. The fact that the vase was fragile caused it to break when it hit the ground. Trope Theorists, recall, say that tropes are the fundamental relata of causal relations. In our example, 'the fact that the vase was fragile' must then denote a trope, namely the glass's fragility trope. But Realists seem to have space for property instances as well, though on that view property instances will be complex objects composed of a universal and some particular. If facts, or anyway some facts, are just property instances, can we make any headway toward settling the dispute between Trope Theorists and Realists by considering the nature of facts?

A question thus arises: exactly what sort of thing is a fact? There are three popular answers, two of which have already emerged in our brief discussion.

1 A fact is a trope or individual accident like the particular whiteness of a particular snowball. This must be a *non-transferable* trope or accident, one that could not possibly be the accident of any other particular. If the *whiteness* trope of this snowball could be transferred to another object, like an icicle, then it wouldn't be a classical truthmaker for the proposition that this snowball is white, since its mere existence would not be sufficient for the truth of that proposition. If tropes are non-transferable, then the *whiteness* trope of this snowball exists if and only if this snowball is white. The greenness of the snowball, for example, simply doesn't exist, and as a consequence, it is not true that the snowball is green.

2 A fact is a complex state of affairs, something made up of at least one particular (the snowball) and at least one property or universal (the property of *being white* or WHITENESS). The combination of the particular and the property exists only if the corresponding sentence (the one predicating the property of the particular) is true. If the sentence is false, then the individual parts (particular and property) may exist, but their combination or union into the relevant state of affairs does not. If we adopt Totality Fact Maximalism (2.1T.3), then the basic facts are complexes built up from properties and those individuals that make up the totality of its instances.

   States of affairs of this kind can be thought of as the constituents of *thick particulars*. We could, if we wished, identify the snowball (the thick particular) with the combination of all the states of affairs that connect some underlying thing (a *thin* or *bare* particular) with some property.

3 A fact is a nexus or tie between one or more particulars and a property or universal. If a snowball is white, then there exists a single, simple thing (the tie or nexus) that connects the snowball (the particular) with *whiteness* (the property). This very nexus could not exist unless the snowball were in fact white. Hence, the mere existence of

the nexus is sufficient to ground the truth of the sentence 'The snowball is white.' (A Totality Fact Maximalist version would posit nexuses that simultaneously connect some property with all of its instances.)

These different accounts of facts will naturally combine with the different accounts of particulars that we will take up in the next section. As a quick preview: Trope Theorists will naturally prefer to think of facts as tropes (option 1). Relational Ontologists will prefer to think of facts as nexuses (option 3), and Constituent Ontologists will think of them as states of affairs or thick particulars (option 2).

REALISM AND NEXUSES If we were to combine UP-Realism (7.1T.1T) with Classical Truthmaker Theory (2.1T), then we would have to suppose that whenever a particular *P* instantiates a universal *U*, there must be a truthmaker for the corresponding truth. If the *U* is the *essence* of *P* (something *P* could not exist without), then *P* itself could be the truthmaker, since the existence of *P* in that case would be sufficient for the truth that *P* instantiates *U*. However, if it is only an *accidental* truth that *P* instantiates *U*, that is, if it is possible that *P* exist but not instantiate *U*, then we would have to posit some further truthmaker for the instantiation relation's actually holding in this case. Such a truthmaker could be what has been called an 'instantial tie' or a 'nexus', an entity that connects the particular to the universal. Alternatively, the truthmaker could be a *state of affairs*, a whole that is composed of the *P* and *U* as parts.

There is one obvious drawback to positing the existence of nexuses or states of affairs. Adding a new sort of thing counts as a demerit of a theory, according to Ockham's Razor (**PMeth 1.4**). But there are three advantages to positing nexuses or states of affairs:

1 Positing nexuses or states of affairs makes UP-Realism consistent with Classical Truthmaker Theory in the case of accidental predications.
2 Positing nexuses or states of affairs provides us with entities that can serve as aspects of particular things, which in turn can serve as the relata of causation (as we will discuss in Chapter 27), as well as objects of perception and thought. (This advantage is less clear for Totality Fact Maximalists, since they may lack the local, atomic facts needed for causal connection.)
3 Nexuses or states of affair can also serve as the reference of possessive gerunds. For example, if Socrates is bald, then the gerund phrase 'Socrates' baldness' can be taken as a name of the nexus connecting the particular Socrates with the universal BALDNESS. However, Atomic Truthmaker Theory (2.1T.4) would provide such nexuses only for atomic gerunds, not for ontologically complex ones, like negations ('Socrates' not participating in the Assembly'), disjunctions ('Socrates' walking or running') or gerunds involving non-fundamental entities or properties.

In fact, nexuses or states of affairs play the very same three roles for Realists that modular tropes play for the Trope Nominalists (8.1T.4.1A). Nexuses have also been called 'moments', 'modes', or 'individual accidents'.

ARE NEXUSES PARTICULARS OR UNIVERSALS? There are several reasons for thinking that nexuses must be particulars. First, it seems right to suppose that each nexus is intrinsically indiscernible from any other nexus. Hence, nexuses violate the Identity of Indiscernibles, the characteristic feature of particulars according to definition Def D5.1. Second, nexuses are not instantiated by anything. If a nexus were instantiated by the particular that it is tied to, then there would have to be a second nexus acting as the truthmaker of this instantiation, leading to an infinite regress of truthmakers. Finally, each nexus inherits its particularity from the particulars that it ties to universals, since otherwise it could not play the three roles mentioned above.

Since a nexus is a particular that does not itself instantiate any universals, it is what is called a 'bare' particular. Its being a bare particular does not mean that we cannot attribute any properties to it. We can certainly say that each nexus is a particular, that it is identical to itself and distinct from other things, that it is a truthmaker for some truth about instantiation, and that it figures in certain causal, perceptual, and epistemic facts. None of these properties require that the nexus instantiate any universals. Nexuses have no intrinsic character or intrinsic similarity to other nexuses that requires the existence of a universal as an explanation.

ARE STATES OF AFFAIRS PARTICULARS? Similarly, there are good reasons to think that states of affairs would have to be particulars. They are not instantiated by anything, and they inherit their particularity from their particular parts. If they didn't, they couldn't serve as the truthmaker for the claim that this particular instantiates that universal. In fact, some philosophers, such as David Armstrong, have suggested that all ordinary thick particulars are simply states of affairs.

IDENTITY CONDITIONS FOR NEXUSES AND STATES OF AFFAIRS If a nexus $N$ is a tie between particular $P$ and universal $U$, is it essentially so? Could the same nexus $N$ become a tie between a different particular $P_2$ and $U$ or between $P$ and a different universal $U_2$? It would seem not, since if the nexus were transferable in this way, we would need to find a truthmaker for its connection with this or that particular or universal, leading again to an infinite regress of nexuses. If nexus $N$ is a nexus between $P$ and $U$, then it must be a matter of $N$'s very essence that it be so.

For similar reasons, the constituents of a state of affairs must be essential to it. Otherwise, it could exist even though its actual constituents did not stand in the instantiation relation.

NEXUSES VS. STATES OF AFFAIRS If we do suppose that there are truthmakers for the connection between particulars and the universals they instantiate, should we suppose that the truthmakers are nexuses or states of affairs? The difference between the two lies in the fact that states of affairs are supposed to have particulars and universals as parts, while nexuses are simple entities that are supposed to be the connection between particulars and universals. Should we suppose that the particulars and universals are *part*s of the truthmaker, or is the truthmaker something disjoint from both of them that somehow *ties* one to the other?

There is a problem with the state-of-affairs model. If we suppose that for every pair of things, there is a whole that contains just the two of them as parts, then we will end up with far too many states of affairs. If some horse $H$ exists, and the universal $G$ exists that is instantiated by all and only galaxies, then there will be a whole composed of $H$+$G$, and this whole will be the state of affairs of *H*'s being a galaxy! We clearly don't want that. We have two options. (1) Suppose that in such cases, there is no whole composed of *H*+*G* (i.e., that wholes composed of particulars and universals exist only when the particular instantiates the universal). (2) Suppose that the whole *H*+*G* exists, but it isn't an actual state of affairs. In the second case, we would need something to be the truthmaker for the truth that some particular-universal composite is an actual state of affairs, which would force us to introduce nexuses. So, believers in states of affairs must take the first option.[1]

## 9.3 Substances

In this section, we will develop a taxonomy of theories of substance. To make our way toward this taxonomy, we must first get on the table an intuitive understanding of what we're talking about when we talk about substances, an understanding toward which we've merely gestured to this point. And the best way to do this is by example. Substances include naturally occurring inanimate objects like rocks and clouds, living things from viruses to trees to lizards to cats to human beings to angels to God, and artifacts like tables, stereos, and skyscrapers. So by 'substances' we here have in mind, more or less, the ordinary objects of everyday experience. What we're up to here is trying to understand the *metaphysical makeup*, if you will, of these ordinary objects.

Substances, unlike other particulars such as simple states of affairs and (if there are any) tropes, are *thickly* charactered. That is, substances have multiple features, multiple dimensions of character. A table is not just made of wood—it also has a definite size, mass, and shape.

One note before we get started. For the remainder of this section, we will be setting the trope-free or Extreme varieties of Nominalism (like Extreme Resemblance Nominalism) aside. The reason for this is simple: the fundamental question that will drive us concerns how substances relate to the things that ground their character. Since Extreme Nominalists don't think that *anything* grounds the character of substances, there's really no issue here to deal with.

So we assume that there are properties that ground character and begin with a question. Are properties *constituents* of substances? This question results in a bifurcation of theories of substance into Relational and Constituent Ontologies:

**9.1A Relational Ontology.** Instantiation is a fundamental relation between substances and properties, and instantiation is not a case of the *part-whole* relation.

**9.1T Constituent Ontology.** When a substance instantiates a property, the instantiation relation between the two consists in the fact that the property is a part of the substance.

The first thing to notice about the two approaches is that Constituent Ontology has a significant advantage in terms of qualitative simplicity over Relational Ontology (see **PMeth 1.4**). Relational Ontologists must posit a special relation of instantiation holding between substances and properties in addition to the *part-whole* or *parthood* relation. Constituent Ontologists, in contrast, can make do with just one relation, *parthood*. For

Constituent Ontologists, instantiation can be defined in terms of parthood. For a substance to have a property is for it to have the property as a part. Relational Ontologists must argue that instantiation is irreducible and primitive.

We examine these two approaches to substance in turn.

### 9.3.1 Relational Ontology

Relational Ontology requires that substances and properties stand in the fundamental relation of *instantiation*. As a result, Relational Ontology cannot analyze properties in terms of substances and thus demands that Extreme Nominalism is false. Similarly, Relational Ontology precludes analyzing substances in terms of properties and thus entails that there are fundamental properties and fundamental particulars. Therefore, if Relational Ontology is true, then so is either UP-Realism or something like UP-Realism with tropes instead of universals.

If one is a Relational Ontologist who embraces UP-Realism and rejects Trope Theory, endorsing Classical UP-Realism (8.2A.2), then one is a Classical Relational Realist:

**9.1A.1T Classical Relational Realism:** Relational Ontology and Classical UP-Realism are both true.

On the other hand, if one embraces Trope Theory, then one is a Trope Relational Ontologist:

**9.1A.1A Trope Relational Ontology:** Relational Ontology and Trope Theory are both true.

As we have seen, tropes can be thought of as modifying tropes or modular tropes. Trope Relational Ontologists must choose between these two conceptions. Thinking of tropes as modular tropes (recall that modular tropes have the character they ground) is uncomfortable for the Relational Ontologist because Relational Ontology requires that *instantiation* be a fundamental relation between substances and properties. Regardless whether tropes instantiate the character they ground, Relational Ontologists must insist that substances instantiate properties. This means that if tropes are modular, whenever a substance has some type of character, *two* things turn out to have the character: the substance and the modular trope that grounds the character in the substance. There is duplication of character if both Relational Ontology and Modular Trope Theory (8.2T.1A) are true.

Therefore, Trope Relational Ontologists ought to think of tropes as modifying tropes. Here we arrive at the central problem for Modifying Trope Theory (8.2T.1T), namely that given the presence of fundamental particulars, modifying tropes wind up seeming an awful lot like universals. Indeed, Modifying Trope Relational Ontology turns out to be a quantitatively bloated Classical Relational Realism. Why is this the case?

Modifying tropes, recall, do not have the character they ground, and this means that modifying tropes only have what we might call 'formal' character. Modifier tropes only have features like *being a trope*, *being a property*, and *being self-identical*. If tropes only have formal character, though, it is difficult to see why they can be instantiated by just one thing.

To press this point, consider this question. What underwrites the claim that tropes are numerically distinct? A standard answer is tropes are distinct in virtue of occupying different spatiotemporal locations. That tropes are spatiotemporally isolated then goes to explain why it is that a trope can only be instantiated by one thing. For if a trope were instantiated twice or more, it would have to be in more than one place at once. This is impossible on the view that spatiotemporal location individuates tropes. The trouble with the spatiotemporal individuation of tropes is that it runs afoul of a plausible thickening principle (see Section 8.2.2 for the initial discussion of these principles):

**Location Thickening.** If something is located, then it has a definite size and shape.

Location Thickening creates a problem for modifying tropes, since they are meant to have only formal character. Location Thickening, given spatiotemporal individuation, entails that modifying tropes have some size or other, and size properties are not formal properties. Likewise for shape. (Importantly, Location Thickening does not require that the size of a thing be more than point-sized. For example, a trope might have a size of zero cubic meters.) This tension puts pressure on Trope Theorists to adopt a different view of the individuation of tropes, and it is hard to see how there is space for anything but insisting that tropes are primitively distinct. This, however, looks to be no more than merely stipulating that tropes are not like universals in being shareable. What is being gained by this stipulation?

If modifying tropes are potentially shareable, then they are no different from universals. Even if there were multiple *rednesses*, as Trope Theory claims, this would just be like the Realist claiming that there are multiple REDNESS universals. Modifier Trope Theory, therefore, just posits many things of exactly the same sort that the Realist posits. This quantitative cost comes with no added explanatory power, and so by Ockham's Razor (**PMeth 1**), should not be embraced. Trope Relational Ontology, if tropes are modifiers, is only quantitatively bloated Classical Relational Realism.

Trope Relational Ontologists have one line of defense that is worth considering, however. The task they face is that of differentiating modifying tropes from universals. Here is one way in which they might differ: some modifying tropes might be modified by other modifying tropes. For example, suppose that the property of *being a living human being* is a modifying trope (something like Aristotle's conception of a soul, pneuma in Greek). This soul-trope might itself be modified by other properties, such as properties ascribing various sensory and intellectual states to the soul. Similarly, there might be spatial quantity tropes, that correspond to volume and shape, and these spatial tropes might themselves be modified by various quality tropes, like color, mass density, charge, or temperature. Schlolastic philosophers, following Aristotle's lead, took spatial (extensive) quantity to be the first of a substance's accidents (changeable, non-essential properties), with other accidents inhering in quantity. This sort of account is not available to the Realists, since they cannot suppose that universals (like HUMANITY) are modified when their instances take on new qualities.

Nonetheless, Classical Relational Realism is a quite plausible understanding of Relational Ontology. On this view, universals ground the character of substances, and *instantiation* is a fundamental relation between substances and universals. In fact, this is the only type of Relational Ontology that has been defended in the literature. However, it faces a serious and well-known challenge, the Extrinsicality Objection.

*9.3.1.1 The Extrinsicality Objection.* Classical Relational Realists propose that a substance's having some character consists in its standing in a certain fundamental relation to a universal. However, a substance's having character is an intrinsic feature of that thing (something to do with that thing, and that thing alone), whereas standing in a relation to some further entity (like a universal) is a paradigm of an extrinsic feature (something to do not just with that thing and its parts, but with that thing and something else). Classical Relational Realism, therefore, mistakes something intrinsic to a substance for something extrinsic to it.

What does it mean for a property to be *intrinsic* to a substance? Here is an answer that we introduced in Chapter 2:

**Def D2.3 Intrinsicality.** *x* is *intrinsically F* if and only if nothing that is not *x* or a part of *x* is part of the ground of *x*'s being *F*.

According to Relational Ontology, *x*'s being *F* is always a matter of *x*'s standing in the *instantiation* relation to the universal *F*, which is *not* part of *x*. Thus, it is hard to see how any substance can have a property intrinsically on this account.

Here is an example. Suppose that we have a substance, *S*, that is both red and spherical. According to Classical Relational Realists, this fact consists in the fact that *S* stands in the *instantiation* relation to the universals REDNESS and SPHERICALITY. But since these universals aren't parts of *S*, *S*'s being both red and spherical is not intrinsic to it. The objection can be put in the form of an argument based on grounding or truthmaking:

1 Assume that *a*'s being *F* is intrinsic to *a*.
2 Assume, for *reductio ad absurdum*, that *a*'s being *F* is grounded in *a*'s having some relation *R* to *F*-ness, but this *R*-relation is compatible with *F*-ness's not being part of *a*. (Relational Realism) Assume, then, that in this case *F*-ness is not part of *a*.
3 The *grounding-in* relation is asymmetric and transitive. (There are no cycles or loops of grounding.)
4 If *a*'s being *R*-related to *F*-ness is intrinsic to *a*, then *a*'s being *R*-related to *F*-ness is grounded in *a*'s being *F*.
5 If *a*'s being *R*-related to *F*-ness is intrinsic to *a*, then *a*'s being *F* is grounded in itself. (2, 3—transitivity, 4)
6 So, *a*'s being *R*-related to *F*-ness is not intrinsic to *a*. (5, 3—asymmetry)
7 If *a*'s being *F* is grounded in something that is not intrinsic to *a*, then *a*'s being *F* is not intrinsic to *a*. (Definition D2.3, transitivity of grounding)
8 So, if *a*'s being *F* is grounded in *a*'s having some relation *R* to *F*-ness, then this *R*-relation entails *F*-ness's being part of *a*. (Discharging assumption of 2)

Premise 4 seems plausible. If *a*'s being *R*-related to *F*-ness is intrinsic to *a*, it must be grounded in (made true by) some intrinsic feature of *a*'s, and what other intrinsic feature of *a*'s could that be besides *a*'s being *F*?

This seems like a solid argument. What reply could Classical Relational Realists make?

Classical Relational Realists can reply that there are no truly fundamental properties of substances (except perhaps that of *being a particular and not a universal*). All properties are reducible to the *instantiation* relation plus the class of universals. Consequently, REDNESS and SPHERICALITY are not really fundamental, and so not really intrinsic, to red things and spheres. This is a surprising result of Classical Relational Realism! It is contrary perhaps to common sense, to the impression that we have of many apparently intrinsic properties.

There is another way for Classical Relational Realists to try to blunt the force of the Extrinsicality Objection. They could suppose that universals are really present in space and time, by being co-located with the substances that instantiate them. If a substance instantiates REDNESS, then REDNESS is present wherever that substance is. But this move doesn't make REDNESS a fundamental property of the particular, and so it doesn't really defuse the Extrinsicality Objection. However, it might make the extrinsicality of a thing's redness more palatable, since it enables us to suppose that REDNESS and the particular that instantiates it share a location.

If one doesn't think that sharing a location is enough to account for the impression of the intrinsicality of these other properties, then one may want to give Constituent Ontology a go.

### 9.3.2 Constituent Ontology

Here again is the definition of Constituent Ontology:

**9.1T Constituent Ontology.** When a substance instantiates a property, the instantiation relation between the two consists in the fact that the property is a part of the substance.

Since substances are thickly charactered, the Constituent Ontologist thinks that among the parts of a substance are each of the properties it instantiates. Properties are not, of course, parts in the way that ordinary physical parts are. Properties aren't like branches or pen caps or hands or tails. We might say that properties are *metaphysical* parts, or *constituents*. Just as some things have physical structure, defined by the nature of and relations among its physical parts, substances have metaphysical structure, defined by the nature of and relations among its metaphysical parts, its constituents. (Incidentally, we will use constituent-talk here, to avoid any confusion between physical and metaphysical parts.) One of the ways this metaphysical structure manifests itself is in the character of a thing. A thing's character seems to be a result of the way it is, the way it is structured metaphysically. Sometimes this thought is expressed by saying that a thing's properties are *immanent in* it, rather than separate from it as in Relational Ontology.

Many Constituent Ontologists accept what Michael Loux has called the 'Principle of Constituent Identity':

**PMeta 4.1 Principle of Constituent Identity (PCI) for Substances.** If $A$ and $B$ are substances, and every constituent of $A$ is a constituent of $B$ and vice versa, then $A$ is identical to $B$.

The reason many Constituent Ontologists accept PCI is quite simple: Constituent Ontologists tend to think of a substance as a complex object composed out of its constituents and their interrelations. Thus, given that the interrelations among a substance's constituents are like the interrelations among other substances' constituents—which we will see is a common commitment of Constituent Ontology as well—it is hard to see how things with identical constituents could *fail* to be identical. If substances are made out of their constituents, and if *this* substance and *that* one have the same constituents, *this* and *that* must be the same substance.

PCI is very much like a corresponding principle about physical parts. Suppose there was a thing A that had just the same physical parts as a thing B. That is, every part of A is a part of B and vice versa. It seems to follow from this that A *just is* B, that the "two" are really just one. So the intuition underlying PCI is very much like the intuition underlying this claim about the connection between physical parts and identity. Whether we accept PCI should depend on exactly how we think of parts and wholes.[2] If a whole is "nothing over and above" its parts, that is, if it represents no "addition to being," then PCI seems undeniable. In contrast, if the basic relation is the relation of composition, the relation by which simples $x_1, x_2, \ldots, x_n$ compose a new, composite entity $y$, then composite things are just as fundamental as their parts. If so, it is hard to see why two different composite things might not share all of the same proper parts, in contradiction to PCI (see Hochberg 1965 and Rodriguez-Pereyra 2004).

Nonetheless, even if PCI were not true of all composite objects, there might be a good reason to think that PCI should apply to bundles of properties. The first issue to consider is whether bundles have their parts essentially or can gain or lose parts. If *B* is a mere bundle of various parts, then it seems reasonable to assume that *B* could not survive the addition or subtraction of parts from the bundling relation. Thus, it does seem that all bundles of properties have those property-parts essentially because the bundle is nothing more than the combination of those properties. If this is true, then it also seems plausible to think that sameness of constituents is sufficient for the identity of property-bundles. Consider the following, weaker version of PCI:

**PMeta 4.2 Weak PCI.** If $x$ and $y$ are necessarily composite and necessarily have the same proper parts, then $x = y$.

Weak PCI is plausible for bundles. Bundles of properties are intrinsically featureless—all of the features of the bundle are derived from the features of its constituents. So, bundles have no holistic character, and so they couldn't have any essence above and beyond the properties they contain. Weak PCI then applies to property-bundles, since each bundle is essentially composite and essentially has the components it does. If Weak PCI applies to bundles, then it is reasonable to infer that PCI also applies to them, since bundles have their parts essentially.

Further, suppose contrary to Weak PCI that two distinct bundles had the same constituents. If the constituents are essential to each, neither can gain or lose constituents. They are not only indistinguishable in fact—they are essentially indistinguishable. What then could possibly make the two bundles *two*?

Let's briefly consider some of the cases involving material objects where it seems that PCI might be false. One classic example involves a statue, Goliath, and the lump of clay of which the statue is composed, Lump. The two, Goliath and Lump, seem to be composed of exactly the same material parts (the same clay granules, the same atoms and molecules), and yet there seem to be two things rather than one, since Lump can survive if it is squashed into a ball, but Goliath cannot survive such squashing.

If we reflect on a case like this, we see that the statue is not a mere bundle of clay granules. It has, in addition, a certain essential structure or shape. If this shape is destroyed, the statue ceases to exist, even if the clay granules remain bundled together. However, bundles of properties have no internal structure or internal feature of any kind except insofar as that structure or feature is represented or constituted by the presence of an appropriate property in the bundle. Bundles of properties are *mere* bundles, and so their identity is constituted by the particular set of properties that are bundled together. Same properties, same bundle, same substance.

Suppose, though, that a property-bundle had some internal structure. Since we are working with Constituent Ontology, the bundle's having that structure must consist in its containing some appropriate property or properties. This is just an application of the general Constituent Ontology strategy of explaining a thing's attributes in terms of its constituent properties. Consequently, any bundle having exactly the same constituents would have exactly the same internal structure. Unlike the case of Goliath and Lump, there would be nothing to distinguish one bundle from the other.

Here is a further argument for PCI's applicability to all property-bundles. In every case, the distinctness of two things should be intrinsic to the pair. Nothing outside of the two things could be responsible for their being distinct from each other. But what could be intrinsic to a pair of bundles except their constituents? Hence, two bundles cannot be distinct except by containing distinct constituents. In contrast, two simple things (like two universals or two bare particulars) can be distinct despite the fact that neither contains any proper part. Simple things can carry their own distinct identity within them, while mere bundles must derive their distinct identities from the distinctness of their constituents.

In conclusion, we have good reason to accept PCI if we limit its scope to property-bundles:

**PMeta 4.3 PCI for Property-Bundles.** If $x$ and $y$ are bundles of properties, and $x$ and $y$ have the same proper parts, then $x = y$.

Since Constituent Ontologists maintain that a substance's properties are among its constituents, an obvious question is whether substances have constituents other than those properties. If one says that the only constituents a thing has are its properties, then one is a Bundle Theorist. If one says that a substance has a constituent or constituents other than its properties, then one is a Substrate Theorist.

**9.1T.1T Bundle Theory:** The only constituents of each substance are its characterizing properties.

**9.1T.1A Substrate Theory:** Each substance has a constituent other than its characterizing properties, a substrate.

We will consider these views in turn.

*9.3.2.1 Bundle Theories.* The idea behind Bundle Theory is that substances are nothing but bundles of properties. As we will see in detail below, this contrasts with Substrate Theory since substrates are in a different fundamental category from properties. Bundle Theorists thus have a putative qualitative advantage over Substrate Theorists because they maintain that the category of substance can be wholly reduced to the category of property. Bundle Theory has need of just one fundamental category in their account of substance, while Substrate Theory requires two. Bundle Theory thus constitutes a natural, simple starting point for working out one's Constituent Ontology. The question is whether Bundle Theory can pull off its promised reduction.

Since properties come in two varieties, tropes and universals, so does Bundle Theory:

**9.1T.1T.1T Trope Bundle Theory:** Bundle Theory is true, and tropes ground character.

**9.1T.1T.1A Classical Bundle Theory:** Bundle Theory is true, and universals ground character.

Trope Bundle Theory thinks of substances as bundles of tropes, whereas Classical Bundle Theory thinks of substances as bundles of universals. Therefore, Trope Bundle Theory reduces one category of particular, substance, to another category of particular, trope. Classical Bundle Theory, on the other hand, reduces the category of substance to a category of non-particular, universal. Each of these Bundle Theories comes in many varieties, distinguished by the ways in which they variously construct bundles. We will, however, only discuss those varieties with respect to Classical Bundle Theory, since the troubles for Trope Bundle Theory have more to do with the proposed reduction to tropes and less to do with the way the bundles of tropes are constructed.

Consider first Trope Bundle Theory. Tropes are either modifying or modular. There are serious problems for Trope Bundle Theory on both views.

Suppose tropes are modular and, therefore, that each trope has the character it grounds. As we saw with the Modular Trope variety of Relational Ontology above, there is a danger of character duplication if one thinks that substances themselves are charactered. The Trope Bundle Theorist, though, has a strategy unavailable to the Relational Ontologist: Trope Theorists can deny that modular tropes ground character by making something else, distinct from the trope itself, a possessor of that characteristic. The idea is to emphasize the bundle theoretic *reduction* of substances to tropes and to therefore deny that the substance itself has character. The sense in which a substance has some dimension of character, on this way of thinking, is just by having a constituent with that character. But this kind of character-having should be thought of in an attenuated sense. The substance doesn't really *have* character; only modular tropes do that. (Otherwise you get duplication!) This move is problematic, though, for it runs afoul of the following, plausible principle:

**Existence of Thick-Characteredness.** There are thickly charactered things.

If modular tropes are the only truly charactered things, then the Existence of Thick-Characteredness is false. It certainly seems, though, that there are things that are thickly charactered.

Notice that the trouble is created by the threat of character duplication, and this threat is in turn created by the fact that modular tropes have the character they ground. Modifier tropes, though, do not have the character they ground. Modifier Trope Bundle Theory does not, therefore, have to deny that substances really do have thick character, at least not for the reason Modular Trope Bundle Theory does.

Modifying Trope Bundle Theory, however, faces a different problem with character grounding. Modifying tropes do not have the character they ground. So consider the following question: What are the modifying tropes grounding the character of? What, that is, are modifying tropes *modifying* if Trope Bundle Theory is true? Given that Bundle Theorists reduce substances to a collection of properties (which are the only constituents of substances), the only answer Modifying Trope Bundle Theorists can give is that tropes are characterizing other tropes. There is just nothing else there to be characterized, on this view. But modifying tropes are usually supposed to have only formal character. If modifying tropes are characterizing other modifying tropes, then some modifying tropes will have to have non-formal character. This is now a second reason for defenders of modifying tropes to take seriously the idea that some modifying tropes modify other modifying tropes. Now we have a reason for thinking that modifying tropes might contain other modifying tropes as parts. However, we will still face the problem of saying what the biggest (most inclusive) modifying tropes modify.

The best solution for the Trope Bundle Theorist might be to adopt a hybrid theory—a theory according to which there are both modular and modifying tropes. We could now identify a substance with a single modular trope that contains many modifying tropes as parts. Now, however, the result looks less and less like a version of a bundle theory. The substance isn't merely a bundle—it is a modular trope (perhaps a spatial quantity module) filled with modifying tropes of various kinds.

What we have run up against is the need for an additional constituent, if Modifying Trope Bundle Theory is true, a non-property (or anyway a non-modifying trope) constituent. The only plausible way to work this out is by including some type of substrate. This, of course, simply abandons Modifying Trope Bundle Theory in favor of a Trope Substrate Theory. We will take up the latter view in Section 9.3.2.2.

We turn, therefore, to Classical Bundle Theory. Classical Bundle Theory purports to reduce substances to bundles of universals. Historically, Classical Bundle Theories have gone in for a more or less sparse theory of properties. This is not surprising, since if one has an abundant theory of properties, then each substance will have innumerable universals as constituents. Thus, substances would have a surprising degree of metaphysical complexity. We will assume a sparse theory here according to which the only universals there are are natural properties (in the sense discussed in Section 7.2.1.1).

The most obvious way to work out one's Classical Bundle Theory is to identify bundles with sets of universals. This is Constructive Bundle Theory:

**9.1T.1T.1A.1 Constructive Bundle Theory.** Substances correspond one-to-one to sets of universals.

Do *all* sets of universals correspond to particulars? It would seem not. There seem to be in fact no particulars that are green elephants. If that were so, then the set consisting of the universals GREEN and ELEPHANT would not be a particular. At the same time, there are many gray elephants. Consequently the set consisting of the universals GRAY and ELEPHANT does not constitute a single particular but a whole host of particulars.

Let's start with some ordinary particular, Jumbo the elephant, for example. Let's say that Jumbo is some actual elephant. We would like to say that there are some universals (like GRAY and ELEPHANT) that Jumbo exemplifies, and others (like GREEN or ANT) that it does not. A natural move for Classical Bundle Theorists to make would be to identify Jumbo with the set containing exactly those universals that Jumbo intuitively exemplifies. Since Classical Bundle Theorists deny that particulars are fundamental entities, it will also follow that the exemplification of universals by particulars is not a fundamental relation. Instead, Classical Bundle Realists will have to introduce a primitive relation of *co-instantiation* among universals. *Co-instantiation* is a property of sets of universals. Intuitively, a set of universals has the property of *co-instantiation* just in case the set corresponds exactly to the universals that are exemplified by some one particular. Jumbo would then correspond to a *co-instantiated* set of universals.

This is Simple Bundle Theory:

**9.1T.1T.1A.2 Simple Bundle Theory.** Substances correspond one-to-one to *co-instantiated* sets of universals; that is, a substance *S* exists if and only if some set of universals *U* is co-instantiated, and *S exemplifies* a universal *F* if and only if *F* is a member of *U*.

BUNDLE THEORY AND INTRINSIC CHANGE The central problem with Simple Bundle Theory is accounting for change and the potentiality for change. Suppose that Jumbo changes from being placid to being angry. At one point in time, we have ELEPHANT, GRAY, and PLACID co-instantiated, and in the next moment it is ELEPHANT, GRAY, and ANGRY that are co-instantiated. If Jumbo simply corresponds to a single set of universals, then it cannot be identical to both of these sets, since the two sets are not identical to each other (one contains PLACID and not ANGRY, the other ANGRY and not PLACID). We have to say either that Jumbo is identical to the first set or to the second. If Jumbo is the first set, then Jumbo was annihilated when the change from PLACID to ANGRY occurred; if Jumbo is the second set, then Jumbo first began to exist when the change occurred. We seem to have lost the obvious and important fact that the change was a change in *Jumbo*, that Jumbo went from being placid to being angry.

Thus, correlating particulars with sets of co-instantiated universals is too simple an account. It leaves out some crucially important facts about change and persistence. In response, Bundle Theorists can move in one of three directions:

1 Nuclear Bundle Theory
2 Four-Dimensional Bundle Theory
3 Evolving Bundle Theory

NUCLEAR BUNDLE THEORY First, Bundle Theorists could distinguish between *nuclear* co-instantiation and *peripheral* co-instantiation. Bundle Theorists could then identify particulars with nuclearly co-instantiated sets of universals, or 'nuclei' for short (see Simons 1994). A nucleus would correspond to those properties that some particular has at all times and cannot possibly lose, so long as it exists. In contrast to nuclei, a *complete bundle* is a peripherally co-instantiated set of universals. We will have to assume that each complete bundle contains exactly one nucleus (as a subset).

**Existence and Uniqueness of Nuclei.** If $x$ is a peripherally co-instantiated set of universals (a bundle), then there is exactly one subset of $x$, namely, $y$, such that $y$ is a nucleus (a nuclearly co-instantiated set of universals).

We thus get a third type of Classical Bundle Theory:

**9.1T.1T.1A.3 Nuclear Bundle Theory.** Substances correspond one-to-one to nuclei of universals. That is, a substance $S$ exists if and only if some set of universals $U$ has the property of being nuclearly co-instantiated, and $S$ exemplifies a universal $F$ if and only if $F$ belongs to some peripherally co-instantiated set (a bundle) that contains $U$.

In addition to Nuclear Bundle Theory, which introduces a distinction between nuclear and peripheral co-instantiation, there are three close cousins of Nuclear Bundle Theory:

1 One could claim that there are two kinds of universals, essential and accidental. The nucleus of a bundle would then just consist of all the essential universals in the bundle.
2 One could claim that bundles have internal compositional structure. On that view, bundles can have sub-bundles. A nucleus could be any sub-bundle that contains no proper sub-bundle of universals. On this account, a bundle could have more than one nucleus.
3 One could claim that every subset of the bundle that contains at least two universals is a nucleus. This would entail that any bundle with more than two universals will have multiple nuclei.

On options 2 and 3, it would be possible for a bundle to have more than one nucleus. In such cases, we would have multiple *coincident* objects, objects that coincide at a moment by having exactly the same constituents and (consequently) exactly the same properties. However, the various coincident objects would have different persistence conditions. If a bundle had two nuclei, $N_1$ and $N_2$, then it would correspond to two distinct substances, $S_1$ and $S_2$, one with $N_1$ as its nucleus and the other with $N_2$ as its nucleus. If $N_1$ survives but $N_2$ does not, then $S_1$ persists in existence, but $S_2$ does not. If $N_1$ and $N_2$ come to be nuclei of two different bundles, then $S_1$ and $S_2$ will both have survived but will have separated from each other, no longer being coincident in constitution.

On any version, Nuclear Bundle Theory and its cousins have a straightforward account of change and potentiality. When a particular changes, the nucleus changes from being the nucleus of one complete bundle to being the nucleus of a second, distinct bundle. To return to our simple example, let's suppose that Jumbo is the nuclear set {*I*, GRAY, ELEPHANT}, where '*I*' represents some universal that uniquely picks out Jumbo among other gray elephants and that Jumbo could not have failed to exemplify (perhaps '*I*' corresponds to Jumbo's having a particular origin in space and time). When Jumbo changes from being placid to angry, the set {*I*, GRAY, ELEPHANT, PLACID} loses its property of *being peripherally co-instantiated* (and so ceases to be a bundle) and the set {*I*, GRAY, ELEPHANT, ANGRY} gains the property of *being peripherally co-instantiated*. Meanwhile, the set {*I*, GRAY, ELEPHANT} continues to have the property of *being a nucleus*, and so Jumbo continues to exist, first as placid and then as angry.

While Nuclear Bundle Theory seems to be able to handle the reality of persistence and change, it introduces a new primitive property or relation of *nuclear instantiation*. Further, Nuclear Bundle Theory must make some additional postulates involving this new relation. Finally, Nuclear Bundle Theorists must suppose that every bundle of universals contains exactly one nucleus. Thus, the increase in explanatory power that Nuclear Bundle Theory gains over Simple Bundle Theory comes at a not insignificant cost of decreased simplicity.

Nuclear Bundle Theory's cousins also come at a cost. Option 1, for example, requires a distinction between two kinds of universals, essential and accidental. This option may simply be equivalent to Nuclear Bundle Theory, since the most plausible way to make the distinction would be to stipulate that essential universals are the sort of universals that tend to cohere together in the nuclear way.

Options 2 and 3 give up on the uniqueness of nuclei: one bundle can have many nuclei at the same time. This means that we can no longer simply identify substances and bundles, since the same bundle (if it has multiple nuclei) will correspond with multiple, coincident substances. Since we are working with a Constituent Ontology, this would seem to force us to identify substances with undivided nuclei, and not with the whole bundle. But then we can no longer attribute accidental properties to substances, since those accidental properties will no longer be *constituents* of the substances. In other words, options 2 and 3 seem inconsistent with the spirit of Constituent Ontology. At the end of the day, Nuclear Bundle Theory may be the only viable option in this family of views.

In addition, Nuclear Bundle Theory faces an especially virulent form of the Problem of Individuation that we will examine below. It would seem natural to think that the essential properties of a thing (the properties it cannot lose) are those properties that make up its species or natural kind. If so, two members of the same species would have exactly the same nucleus—any two elephants, for example, would have the same essential properties and so the same nucleus. But this would mean that each species or natural kind can contain just one instance—there could be only one elephant, one molecule of water, and so on. In order to avoid this disastrous result, the Nuclear Bundle Theorists will have to include in each nucleus individuating properties sufficient to distinguish one member of the species from others. But what sort of properties could these be? Why couldn't two members of the same species be intrinsically exactly alike? It would seem that the Nuclear Bundle Theory would either have to add certain tensed properties (detailing the unique history of each member of the species) or it would have to include special properties or thisnesses whose only job is to distinguish individuals from one another. Either approach would mean abandoning Nuclear Bundle Theory for either Four-Dimensional Bundle Theory or Scotism (to be discussed below).

FOUR-DIMENSIONAL BUNDLE THEORY AND EVOLVING BUNDLE THEORY A second way that Bundle Theorists can cope with change is to deny that bundles change intrinsically at all. Instead, a bundle is an eternal object, extended in time, the fourth dimension. If the bundle constitutes an ordinary particular that changes over time, this is represented in the bundle by the fact that each of the properties contained in the bundle is somehow *indexed* or relativized to an instant of time. So, suppose we have an iron rod that is cold at time $t_0$ and hot at a later time $t_1$. This rod could correspond to a single, unchanging bundle that contains two universals, one corresponding to *being cold-at-*$t_0$ and the other to *being hot-at-*$t_1$. Clearly, this will involve some complexity in our theory of universals, as well as an infinite number of instants or moments of time. We will examine this sort of question in more detail in future chapters, including Chapters 19–21 (on time) and 24–25 (on change).

The final way for Bundle Theory to cope with change is to deny the principle of Mereological Essentialism for bundles, allowing that a bundle can gain or lose parts (universals) while remaining one and the same bundle that it was before. What PCI forbids (on this interpretation) is the existence of two distinct bundles containing exactly the same universals at the same time.

How can bundles evolve over time, gaining and losing universals? If a rod exemplifies COLDNESS at one moment and then exemplifies HEAT in the next moment, what can account for the identity of this bundle through time?

Evolving Bundle Theorists might push back against this objection by postulating that each evolving bundle contains a single nature or essence universal, which governs the possible changes and non-changes which the bundle can undertake under various possible circumstances. This would correspond to Aristotle's idea of the *form* of a substance as the ultimate principle of rest and change for that substance. We will devote an entire chapter, Chapter 25, to the question of the persistence through time of composite objects (like bundles).

THE PROBLEM OF INDIVIDUATION We have canvassed five different types of Classical Bundle Theory, and now move to a classic, powerful argument against Classical Bundle Theory as such. The argument's classic statement is found in Max Black (1952) but was extended by Robert Adams (1979), among others. Given our assumption that Classical Bundle Theory is committed to a sparse theory of properties according to which all universals are natural properties, it follows that Classical Bundle Theory, when combined with the Principle of Constituent Identity, entails that substances satisfy Restricted Identity of Indiscernibles:

**9.3T.1 Restricted Identity of Indiscernibles.** If (thing *a* exemplifies *F* if and only if thing *b* exemplifies *F*), where *F* is a natural property, then *a* is identical to *b*.

Black argued, however, that in fact substances *can* be indiscernible without being identical, contrary to Restricted Identity of Indiscernibles. His argument relies crucially on a thought-experiment. Black asks us to consider a possible world in which only two substances exist, perfectly homogeneous and symmetrical spheres of exactly the same shape, size, and composition. Imagine that the two spheres are co-eternal, and that they are forever revolving around their common center of gravity through a perfect vacuum. The world is characterized by perfect bilateral symmetry: whatever can be said truthfully about one sphere (using only general terms) can also be said truthfully about the others. The two spheres are thus perfectly indiscernible, unless we are allowed to assign names to each—let's call them 'A' and 'B'—and make use of non-natural properties like *being exactly such-and-such distance from A* or *being exactly such-and-such distance from B*. If such a world is possible, then it is possible for two distinct particulars to be indiscernible. This possibility, if genuine, vindicates one half of Wise UP (7.2T.3), the half that says that particulars can be distinct yet indiscernible. But more important presently is the fact that if such a world is possible, Classical Bundle Theory must be false.

One crucial question is whether a Black world really is possible. It certainly seems to be. In particular, it is easy to imagine this sort of scenario, and there doesn't seem to be anything inherently contradictory about such a situation. To emphasize the plausibility, imagine a long series of possible worlds, where the first is exactly like ours, and each successive world in the series is more and more like the Black world. You can imagine, for example, that each successive world has one less object than our world, and makes one small change to some actually existing ball such that it becomes more like the spheres in the Black world. As we move through the sequence, then, we come to worlds containing two very but not exactly similar spheres. The last world in the series is just the Black world itself. Surely all of the worlds (except the last, so as not to beg questions!) is genuinely possible. What reason could we have for claiming that at some point in this sequence we move from a possible world to an impossible one? If there is no such reason, then it is hard to believe that the Black world is impossible, since there are perfectly possible worlds arbitrarily close to it in structure.

As an alternative to Black's spheres, we can also imagine a temporally symmetrical world. For example, consider a world in which exactly the same cycle of events recurs over and over again without beginning or end. This is the sort of world imagined by Friedrich Nietzsche as the "myth of eternal recurrence". Each event in each cycle could be distinct from the corresponding events in other cycles, even though all such events are qualitatively indistinguishable and bear exactly the same temporal relations to other events (qualitatively described). There would, in such a world, be indiscernible yet distinct particulars.

Finally, it might be possible for there to be indiscernible thinkers. Can we imagine a Black world that contains, instead of two symmetrical spheres or eternally recurrent epochs, only two indistinguishable and disembodied minds? It is not clear whether such a scenario is really possible because we don't have direct experience of co-existing, disembodied minds. It is thus hard for us to judge what is or is not possible for these sorts of beings. Though we are inclined to think that such a situation is possible, we are less sure of it than in the other cases, so we don't rest anything on such a case.

Further, we do have experience of our own conscious life, and we can, as the French philosopher René Descartes did, consider that conscious life as abstracted from the career of our physical body. Thus, we can imagine a Nietzschean world of eternal recurrence consisting of a single mind, eternally experiencing over and over (without beginning or end) the same cycle of experiences. In such a world, there would be distinct but indistinguishable mental events and actions, and these would be fundamentally real things. Given that these events and actions are particulars, we would here have another sort of Black-type world.

Our defense of the possibility of a Black-type world appeals to two epistemological principles, Imagination as a Guide to Possibility (**PEpist 1**), along with its corollary, The Limit of Possibles Itself Possible:

**PEpist 1.1 The Limit of Possibles Itself Possible.** If a series of scenarios is such that each represents a metaphysical possibility, and the series converges at the limit on a further scenario, then we have good reason to think that the latter scenario represents a metaphysical possibility.

Both Imagination a Guide to Possibility and The Limit of Possibles Itself Possible are, as we intend them, *defeasible* principles. That is, they are true despite being subject to counterexamples. One might be able to imagine that Mark Twain and Samuel Clemens are different people, but this doesn't establish the possibility of their distinctness. Similarly, it is possible for a massive object to approach arbitrarily close to the speed of light, but it is impossible for one to reach the speed of light. Nonetheless, despite the possibility of exceptions, both principles are *reliable guides* to possibility. If one wants to reject an inference made on the basis of one of these principles, one must supply a reason why that particular inference should not be made.

There are, however, several ways Classical Bundle Theorists could handle the possibility of Black worlds. First, they could suppose that each spatial location corresponds to a different universal. On this view, the two spheres are not really indiscernible after all. They do correspond to different bundles of universals. This is a version of the Theory of Spatial Qualities (17.1T.1T), a theory we discuss at greater length in Chapter 17. We will see there that there are serious problems with this view of spatial location. It is inadvisable to rest one's theory of substance on such a controversial view of location.

Second, as John Hawthorne (1996) has argued, Classical Bundle Theorists can embrace the possibility of a Black world, even if they reject the Theory of Spatial Qualities. Classical Bundle Theorists can suppose that it's possible for one and the same bundle of properties to be located simultaneously at two remote locations. Such a bundle could be *x* miles distant from itself. This isn't a problem for Classical Bundle Theorists, since they are already comfortable with the idea of universals being simultaneously located in many places at once. If individual universals can do so, why can't whole bundles? Notice that this strategy really amounts to embracing the claim that there cannot be distinct but indiscernible objects! Hawthorne's strategy simply rejects that possibility by insisting that there is really just one object with two locations. Thus, Hawthorne has not actually squared Classical Bundle Theory with the possibility of distinct yet indiscernible objects. Moreover, Hawthorne's suggestion would be of no help to those thought-experiments that don't use location, such as a world containing two indiscernible souls or two indiscernible particles in the same place at the same time.

What we are butting up against here is a problem that is not unique to Classical Bundle Theory. Indeed, it is a problem for any Constituent Ontology that goes in for both universals without tropes and the Principle of Constituent Identity. To see the problem, recall the Hochberg-Armstrong objection to Resemblance Nominalism. Their objection was that Resemblance Nominalists have no way to account for the dual facts of resemblance and distinctness. That is, Resemblance Nominalists have trouble finding distinct truthmakers for (10) and (11) (from Chapter 8):

(10) *A* and *B* are exactly similar fundamental particulars.
(11) *A* and *B* are distinct fundamental particulars.

What we have discovered is that Classical Bundle Theory faces a similar trouble, given PCI. Since only universals are constituents of substances, the facts explaining similarity are the same as those meant to explain distinctness. Since Constituent Ontologies with universals and without tropes provide a powerful account of similarity that is motivated independently of these theories of substance, what is apparently needed is a different account of what makes two things two, rather than one. This is the Problem of Individuation.

RESPONSES TO THE PROBLEM OF INDIVIDUATION There are, broadly speaking, four responses to the Problem of Individuation: (i) primitive identity, (ii) fundamental relations of distinctness, (iii) Scotism, and (iv) Substrate Theory. We consider these in turn.

PRIMITIVE IDENTITY The first response to the Problem of Individuation is to opt for primitive identity. Proponents of primitive identity assert that (11) is true simply because *A* and *B* exist. In truthmaker language, the truthmaker for (11) is just the pair, $\{A,B\}$. In this sense, *A* and *B* (and every other substance) are 'self-individuating':

**Def D9.1 Self-individuation.** A pair of distinct things *x* and *y* is *self-individuating* if and only if the truthmaker for *x*'s distinctness from *y* is the pair itself.

We can thus characterize the primitive identity view in the following way:

**9.2T Primitive Identity.** All pairs of substances are self-individuating.

The trouble with Primitive Identity is that it seems to constitute a rejection of PCI, given Classical Bundle Theory. The reason is simple: the whole idea behind Constituent Ontology is to identify a substance with some combination of constituents. Given the Classical Bundle Theorist's commitment to universals, Primitive Identity offers no space for the Classical Bundle Theorist to identify a constituent that is unique to every substance. We have noted, however, that PCI is a constraint that most Constituent Ontologists accept. Thus it is no surprise that Robert Adams (1979) opts for Primitive Identity while explicitly denying Constituent Ontology.

Moreover, Classical Bundle Theory (even without PCI) is still subject to a version of the Hochberg-Armstrong objection, namely, an appeal to One Truthmaker per Fundamental Property (see Sections 2.5.2, 3.4.3, and 8.1.3).

**PTruth 1 One Truthmaker per Fundamental Property.** If *p* is the true predication of a fundamental property *P* to $x_1$ through $x_n$, and *q* is the true predication of a different fundamental property *Q* to the same things $x_1$ through $x_n$, then *p* and *q* have distinct truthmakers.

If we have two bundles with the same universals as constituents, then the bundles are truthmakers for two, distinct natural relations: that of the co-instantiation of certain universals and that of the numerical distinctness of the bundles. Suppose, for example, that we have two bundles, $B_1$ and $B_2$, each comprising the same three universals (*F*, *G*, and *H*). $B_1$ and $B_2$ are individually and jointly a truthmaker for the proposition that *F*, *G*, and $H$ are co-instantiated, but they are also jointly the truthmaker of the proposition that $B_1$ and $B_2$ are numerically distinct. Given One Truthmaker per Fundamental Property, we should look for some additional truthmaker for their distinctness.

FUNDAMENTAL RELATIONS OF DISTINCTNESS The second response to the Problem of Individuation is to opt for primitive relations of distinctness, or fundamental distinctness nexuses. The idea is that the ground for the truth of (11) is a fundamental relation of distinctness that obtains between $A$ and $B$, and more generally, between any two distinct substances. We can put the view in this way:

**9.2A.1 Fundamental Relations of Distinctness.** For any distinct substances $x$ and $y$, the truthmaker for $x$'s distinctness from $y$ is a fundamental distinctness nexus between $x$ and $y$.

There are at least three troubles for this account. First, insofar as this is a solution to the Problem of Individuation, it seems to follow that what makes a substance the substance that it is are these relations of distinctness in which that substance stands. But this means that it is part of the identity of that substance that it stands in those relations. For example, consider THP's dachshund, Elsie. A fundamental distinctness nexus obtains between Elsie and THP, between Elsie and THP's daughter Gretchen, and so on. If these nexuses jointly individuate Elsie, then part of what it is to be Elsie is to stand in these relations. This means that nothing that *doesn't* stand in these relations can be Elsie. However, Elsie existed prior to the beginning of Gretchen's life. Thus it is possible for Elsie to exist while the distinctness nexus between Gretchen and her does not. This contradicts the claim that this nexus individuates Elsie.

Second, and relatedly, this account seems to get the order of explanation backwards. A relation is binary when it always holds between two things or between one thing and itself. A binary relation $R$ is *irreflexive* if and only if the relation never holds between any thing and itself. If $R$ is irreflexive and $x$ stands in $R$ to $y$, then $x$ and $y$ must be distinct. Any binary, irreflexive relation requires the existence of two distinct things in order for it to be instantiated. If so, then the distinctness of two things must be prior to their standing in any relation whatsoever, including the distinctness relation. This argument relies on the following metaphysical principle:

**Relata more Fundamental than Relations.** The existence and distinctness of the relata of any relation cannot be grounded in or made true by the holding of the relation itself.

Relata more Fundamental than Relations is a plausible principle. Ordinarily, we would take the grounding of the holding of a relation to include the existence and mutual distinctness of the relata, not the other way around. This is especially true when these relations hold only contingently, as we have seen that relations of distinctness do. Advocates of Fundamental Relations of Distinctness must reject the principle, however. This is a serious problem.

Third, it seems reasonable to suppose that any relation of distinctness is intrinsic to any two distinct things. That is, if two things are distinct (and distinct in a fundamental way), then the two things are in and of themselves the truthmaker for the truth that they are distinct. We don't need to add a separate, third thing in such cases. However, that is just what the theory of fundamental distinctness nexuses does: it assumes that the distinctness of *A* and *B* is grounded in the existence of a separate, third thing, the distinctness nexus between *A* and *B*.

There is, perhaps, a fourth objection to this account, a variant of Bradley's Regress. If we assume that it is never possible for two things to distinguish themselves, then the distinctness nexus *N* that is responsible for making *A* and *B* distinct can do so only by virtue of being distinct from both *A* and *B*. If *N* were identical to either one, then it would (by hypothesis) be incapable of grounding their distinctness. But this means that we have to posit two more distinctness nexuses $N_1$ and $N_2$ (each distinct from *N*) in order to distinguish *N* from *A* and from *B*. This would seem to lead to an infinite regress of such nexuses.

SCOTISM The third response to the Problem of Individuation comes from John Duns Scotus. He proposed that we introduce universals which are uniquely instantiated by particulars and which ground the distinctness of each particular from all the others. Scotus called these universals 'haecceities', which is Latin for 'thisnesses' ('haec', pronounced 'hike', is the Latin word for 'this'). A haecceity ('hike-say-ity') is a property that can, as a matter of metaphysical necessity, be instantiated by one and only one possible thing. It is impossible for two distinct things to instantiate the same haecceity, either in the same or in different possible worlds.

**Def D9.2 Haecceity.** If $x$ is a substance, then the *haecceity* of $x$ is a universal $H(x)$ that exists necessarily, is instantiated by $x$, is necessarily instantiated by $x$ if $x$ exists, and is necessarily instantiated by nothing other than $x$.

Scotism, then, is the view that haecceities exist and ground the distinctness of distinct objects, like Black's spheres.

**9.2A.2 Scotism.** Substances have haecceities, and haecceities are natural properties.

The trouble with Scotism is that haecceities do not seem to be natural properties because they do not ground either resemblance or causal powers. They cannot ground resemblance by stipulation, since haecceities in principle cannot be shared. But they also do not seem to ground causal powers, since the causal powers of a thing seem to be a consequence only of shared properties. In other words, it does not seem that there are causal powers unique to individual substances.[3] Unfortunately, unless haecceities are natural properties, they cannot do the work needed by Classical Bundle Theorists, since only natural properties are constituents of substances.

The fourth response to the Problem of Individuation is to adopt Substrate Theory.

*9.3.2.2 Substrate Theories.* Substrate Theories are Constituent Ontologies according to which substances have a constituent other than their properties. This additional constituent is called a 'substrate'.

**9.1A.1A Substrate Theory.** Each substance has a constituent other than its characterizing properties, a substrate.

A property is *characterizing* if and only if it is not a substrate. Substances, then, are metaphysically complex things containing some properties together with a substrate that is in a different fundamental category than those properties (see below for an explanation of why we don't just say that substrates are not properties at all).

One motivation for Substrate Theory is to supply a solution to the Problem of Individuation besetting Classical Bundle Theory. As we have seen, Trope Bundle Theory has significant troubles in both its modifying and modular versions, while Classical Bundle Theory lacks the resources to account for the individuation of substances. Further, Modifying Trope Bundle Theory faced a problem of accounting for what, exactly, was being characterized by a thing's modifying tropes. Classical Bundle Theory faces a similar worry, since universals, like modifying tropes, have only formal character. That is, universals do not have the character they ground.[4] Thus the Classical Bundle Theorist faces the challenge of saying what, exactly, universals are grounding the character of. Substrate Theories hold out hope for being able to solve that problem. Not only can substrates serve as an individuating constituent, they can also be the fundamental bearer of a substance's properties. Substrate Theory is, then, underwritten by this dual motivation.

As with Relational Ontology and Bundle Theory, Substrate Theory comes in both a trope and a classical variety. Trope Substrate Theory insists that tropes ground character, while Classical Substrate Theory deploys universals in that regard:

**9.1T.1A.1T Trope Substrate Theory.** Substrate Theory is true, and tropes ground character.

**9.1T.1A.1A Classical Substrate Theory.** Substrate Theory is true, and universals ground character.

Trope Substrate Theory faces troubles that have already, for the most part, been articulated. Modifying Trope Substrate Theory faces just the problem that Modifying Trope Relational Ontology faced (see Section 9.3.1 above, and note that the argument there didn't depend on the unique commitments of Relational Ontology). Thus, Modifying Trope Substrate Theory appears to be a quantitatively bloated version of Classical Substrate Theory. On the other hand, Modular Trope Substrate Theory, like Modular Trope Bundle Theory (see Section 9.3.2), faces a dilemma between character duplication and a violation of Existence of Thick-Characteredness. Classical Substrate Theory, therefore, seems the most promising option, at least at this stage.

There is a second issue that Substrate Theorists must wrestle with, and though we will focus on Classical Substrate Theory in discussing it, Trope Substrate Theorists are faced with it as well. The question concerns many properties a substrate has *in itself*. That is, considered independently of the constitutive properties of the substance of which it is a part, does a substrate have character? If so, is that character merely one-dimensional or not?

Before we name the views according to which substrates have zero properties or one property in themselves, we must dispatch with the possibility that substrates have two or more properties. The argument is a dilemma, one horn of which is a further dilemma. Suppose that substrates have two properties (the case of three or more properties goes in just the same way). The dilemma goes as follows. Either some properties ground the character of the substrate or they do not. If they do not, then substrates appear to be thickly charactered in just the way that Extreme Nominalists assert ordinary objects are. If that is right, then it's not clear why one shouldn't just adopt a more thoroughgoing version of Extreme Nominalism and avoid the discussion in this section entirely. On the other hand, if properties do ground the character of the substrate, then (and here's the further dilemma) either those properties are constituents of the substrate or they are not. If they are constituents, then one faces the problem of individuation for substrates. Given that one embraced Substrate Theory in the first place, one will need a substrate for one's substrate, and so on. A vicious infinite regress results. So a substrate's properties cannot be constituents of it. Suppose then that a substrate's properties are not constituents of it. Then one has abandoned one's commitment to Constituent Ontology. But Substrate Theory is explicitly a version of Constituent Ontology, and so one's view is incoherent.

Therefore, substrates must have either zero properties or one property in themselves. If one says substrates have one property, then one is a Modular Substance Theorist, while if one says they have no properties, then one is a Bare Particular Theorist:

**9.1T.1A.2T Modular Substance Theory.** Substrate Theory is true, and substrates have one property in themselves.

**9.1T.1A.2A Bare Particular Theory.** Substrate Theory is true, and substrates have no properties in themselves.

According to Modular Substance Theory, a substance's substrate is simply a modular trope. We've called the view Modular *Substance* Theory, though, in order to evince the idea that the substrate is thinly charactered in an irreducibly substance-kind oriented way. That is, you can think of a Modular Substance Theory substrate as thinly charactered in a dog-like way, a human-like way, a oak-tree-like way, and so on. Substrates, on this view, are *merely* dogs, humans, oak trees, and so on.

Why couldn't substrates be thinly charactered in the way that modifying tropes are? The reason is just that which plagued Modifying Trope Bundle Theory. Substrates are meant to be ultimate bearers of character, but modifying tropes are meant to be only formally charactered. Substrates must *be* people, turtles, electrons or whatever else in themselves. Modifying substrates cannot have such intrinsic character. They are thus unfit to play one of the roles substrates are meant to play. Modular substances must be modular tropes.

Modular Substance Theory, in order to avoid troubles with character duplication and the general problem of thickening principles, ought to think of characterizing properties as either modifying tropes or universals. Thus we see that substrates will be of a different fundamental kind from characterizing properties, even if substrates *are* properties. On this view, a substance is a bundle consisting of one modular trope (the substrate) and many modifying tropes or universals. The result is a weird hybrid of Trope Theories: modular tropes for the substance-kind properties, and modifying tropes or universals for the accidental properties.

The Realist version of Modular Substance Theory is also a weird hybrid of two incompatible theories: Relational Ontology and Constituent Ontology. The modular substance does not contain any property as a part. Consequently, it does not contain its substance-kind universal as a part. Either there is no substance-kind universal at all, in which case the Modular Substance Realists would have to give some Nominalistic account of substance-kinds (by, for example, appealing to fundamental resemblance relations) or there is such a universal but the modular substance stands in an extrinsic instantiation relation to that universal. However, if such an extrinsic instantiation relation is fine in this case, why not abandon Constituent Ontology altogether and suppose that substances stand in that same instantiation relation both to substance-kind and to accidental universals?

Let's suppose, then, that Modular Substance Theorists think of substances as composed of a modular substance-trope and a number of modifying accident-tropes. What holds this bundle of tropes together? It must be a relation between the modifying tropes and the modular substrate. In fact, this relation would seem to be exactly the relation that Relational Ontologists call 'instantiation'. It would be simpler to simply replace the modifying tropes with universals, resulting in Classical Relational Realism. What good is achieved by multiplying accidental properties into a large number of indistinguishable modifying tropes?

There is one plausible response to this objection, but it is a response that will get Modular Substance Theory into further trouble. Modular Substance Theorists could argue that what unifies modifying tropes and modular substrates is not the instantiation relation but spatial location. A substance is unified by the fact that its modular substrate and its modifying tropes are all located in the same place at the same time.

This solution generates two difficulties. First, it makes it impossible for two substances with different properties to share the same location at the same time. If they are in the same place at the same time, then the modifying tropes present there would belong equally to both substances. However, it seems possible for two things with different properties to exist in the same place at the same time. Photons and neutrinos, for example, can do this.

Second, this solution means that Modular Substance Theorists cannot provide a metaphysical account of spatial location. Spatial location cannot correspond to a modifying trope, since this would result in an infinite regress or a vicious circularity. Location trope *L* would have to belong to substrate *S* only by being located in the same place as *S*, but this can happen only by *L*'s belonging to *S* already. This version of Modular Substance Theory would have to be combined with an Extreme Nominalist account of spatial properties, resulting in another weird hybrid. If Extreme Nominalism is good enough for spatial properties, why not all properties?

In addition, Modular Substance Theory may also have a problem with thickening principles. Modular Substance Theory is asking us to imagine that substrates have substance-kind characteristics without having any sort of other characteristics. But here is another plausible thickening principle:

**Substance-Kind Thickening.** Every object that instantiates a substance-kind (like *doghood*) has some other characteristic (like a definite size and shape).

If Substance-Kind Thickening is true, then Modular Substance Theory must be false. However, the plausibility of this principle is not nearly as striking as that of the Color Thickening and Shape Thickening principles, especially in light of the reply considered in Section 8.2. There, we noted that Modular Trope Theorists ought to insist that these principles apply only at the level of ordinary objects, not at the level of tropes. We can add a further rejoinder to this reply at this stage. Color, for example, seems like the kind of thing that must spread over a certain physical area in order for it to be present at all, and such an area must be bounded in a certain way. The resultant boundaries will result in a color patch having a definite shape and size. This is one thing that lends enormous plausibility to the principle of Color Thickening, even at the level of tropes. Likewise the principle of Shape Thickening. Substance-kinds, on the other hand, do not seem to be connected to other features in this same way. At least the connections aren't quite so palpable and obvious. While it is indeed difficult to conceive of an unshaped object that exemplifies the substance-kind *doghood*, the connection seems much less necessary when one considers just the level of tropes. This is not the case with the connection between color and shape, for example. We leave it to the reader to consider this area in more detail.

Given the problem of Substance-Kind Thickening, defenders of Modular Substance Theory might consider going in a completely different direction. Perhaps the modular substrate is not an instance of substance-kind properties like *being a human being* or *being an oak tree*. We could instead think of them as instances of spatial extension properties, the property of having a certain definite shape and volume. The substance-kind property would then consist in a modifying trope that is contained by the modular substrate. We could then suppose that the modular substrate has one property in itself (its spatial extension) and many other properties (including its substance-kind) by virtue of containing (and being modified by) the appropriate modifying tropes.

This theory, the Quantitative Modular Substance Theory, might still face some thickening problems. We might worry that anything that has size and shape must have some qualitative properties to fill or occupy its volume. However, this isn't obviously true. The idea of an empty, qualitatively neutral spatial region doesn't seem self-contradictory, in the way that an oak tree without shape or a colored thing without size does. The Quantitative Modular Substrate Theory seems to be the most defensible version of Modular Substance Theory.

There is one more objection to any version Modular Substance Theory. Like other versions of Trope Nominalism, it is subject to the Hochberg-Armstrong objection. Two modular substrates will be both numerically distinct and exactly similar (with respect to their natural kind), and the only truthmaker for both truths will be the pair of modular substrates. In contrast, Bare Particular Theory is immune to this objection, since bare particulars are numerically distinct but not similar in any substantive way (see below). Similarity is grounded in the universals or tropes associated with the two bare particulars.

There is one view left to consider, namely, Bare Particular Theory:

**9.1T.1A.2A Bare Particular Theory.** Substrate Theory is true, and substrates have no properties in themselves.

Bare particulars are meant to be property-less, particular constituents of substances that both individuate and serve as the fundamental bearers of character. They are character-less character-havers, pure particulars without any features in themselves. In this way, they are analogous to Aristotelian *prime matter*, which for Aristotle was formless when considered in itself, was the principle of individuation of substances, and was the ultimate bearer of properties ('forms', in his terminology).

Wilfrid Sellars (1963) objected to the idea of bare particulars as self-contradictory.[5] As we have seen, a bare particular is supposed to be something that bears properties (since it is a substrate modified by a modifying trope or universal), but it is also supposed to be completely without properties (since that is what makes it *bare*). How can something both have properties and lack them? We can put Sellars's objection in the present context in this way, given Classical Substrate Theory: one and the same universal may modify distinct bare particulars in such a way that the bare particulars are themselves made similar, and yet bare particulars are not supposed to be the sort of thing that can be similar to one another. If bare particulars are not similar to one another, then what exactly is the universal doing? In what is the universal grounding character?

In response, Bare Particular Theorists must insist that we distinguish between the substrate as bare and the substrate as modified by one or more universals or tropes. It is only the former that cannot be substantively similar to other (bare) particulars. Substrates cannot exist without exemplifying at least one universal, but we can still ask whether a pair of bare particulars are (by themselves and without its connections to those universals) collectively the truthmaker of a similarity proposition (like (10) above). Two bare particulars are similar only by virtue of their connections to one and the same universal. Hence, the truthmaker for the similarity claim must include that universal, while the truthmaker for the distinctness claim consists in the pair of bare particulars alone.

It is worth noting that Relational Ontologists also believe in bare particulars, in this sense. For Relational Ontologists, all ordinary particulars are bare in themselves. They take on qualities only by virtue of an external relation to universals or modifying tropes.

Bare particulars might perform as many as four distinct functions simultaneously:

(1) Bare particulars of two distinct substances are jointly the truthmaker for the distinctness of those substances. This is the individuating or distinctness-grounding function.
(2) A bare particular is the truthmaker for the particularity (the concreteness and unrepeatability) of the substance to which it belongs. This is the particularizing function.
(3) A bare particular unifies the various properties that its substance possesses. This is the unifying or bundling function.
(4) The fact that several bundles contain (in succession) the same bare particular could be what makes those bundles successive stages in the history of one enduring thing. This is the persistence-grounding function.

In some metaphysical systems, including that of many medieval scholastic philosophers, these four functions are performed by three or four distinct entities. For example, for Thomas Aquinas, it is the prime matter that particularizes and grounds persistence (in some cases), the individual essence that unifies and grounds persistence (in other cases), and the spatial dimensions and location that individuate. If we accept One Truthmaker per Fundamental Property, it seems that we should look for distinct truthmakers for the four kinds of truths, rather than pressing a single category of things, bare particulars, to do all at once.

We have reached bare particulars through function 1, the distinctness-grounding function. If we adopt Wise UP, then whatever individuates will also be responsible for particularization (function 2).

What about unification (function 3)? In particular, is it necessary that each particular contain just one bare particular? This is far from obvious if the only function of bare particulars is individuation. If particular *P* contains five bare particulars, and so does particular *Q*, and if *P* and *Q* contain the same universals, then it is the distinctness of the five bare particulars in *P* from the five bare particulars in *Q* that is responsible for the distinctness of *P* and *Q*. All that is required is that, if *P* and *Q* are distinct particulars, then at least one of the bare particulars contained in *P* be distinct from all of the bare particulars contained in *Q* or vice versa.

If a particular contains more than one bare particular, it would be natural to look for some correspondence between those bare particulars and the parts of the ordinary particular. Suppose that a particular *P* contains a set *S* of bare particulars $\{B_1, B_2,\ldots, B_n\}$. If *Q* were a distinct particular that contains a subset of *S*, then it would be natural to suppose that *Q* was a part of *P* (in the language of formal mereology, a 'proper' part of *P*, since *Q* is not identical to the whole of *P*).

In fact, it could be that some bare particulars are proper parts of other bare particulars. Suppose that bare particular *B* contains parts $B_1$, $B_2$, and $B_3$. If particular *P* contained bare particular *B*, it would also contain bare particulars $B_1$, $B_2$, and $B_3$, and, if there are particulars $Q_1$, $Q_2$, and $Q_3$ that contain (respectively) only $B_1$, $B_2$, and $B_3$, it would be natural to suppose that $Q_1$, $Q_2$, and $Q_2$ are proper parts of *P*.

In fact, it could even be the case that every ordinary particular contains an infinite number of bare particulars. Perhaps each bare particular contains an infinite number of bare particulars as proper parts. Bare particulars might turn out to be *gunky* (in the language of David Lewis). An entity is gunky if it contains proper parts, and all of its parts contain proper parts. A gunky thing contains no simple or atomic parts (parts without further proper parts). On one interpretation of Aristotle's metaphysics, prime matter might be composed of just such gunky bare particulars.

What about bare particulars and change (function 4)? Could an enduring, changing particular gain or lose bare particulars, just as it can gain or lose universals? There is a historical analogy for this possibility, namely, Locke's theory of personal identity through soul-substitution (in his *Essay Concerning Human Understanding* 1689, Book IV, Chapter 3, Section 6). Locke created a thought-experiment in which two people (e.g., a prince and a pauper) exchanged not only their bodies but also their *souls* (whatever that might be like). Locke argued that, so long as the new *inhabitant* of the pauper's soul and body had all of the memories, values, and intentions of the prince, the new inhabitant is in fact the same person (now) as the prince was (then). We could think of a changing particular substance as the analogue of Locke's persisting prince and the two bare particulars as the analogues of the souls.

Of course, bare particulars cannot change intrinsically, since they have no intrinsic features to change. Suppose that bare particulars do persist through time. Then it seems very natural to suppose that any substance that contains just one bare particular at one time must be identical to any substance that contains just that same bare particular at another time. That is, it seems natural to suppose that the inclusion of the same bare

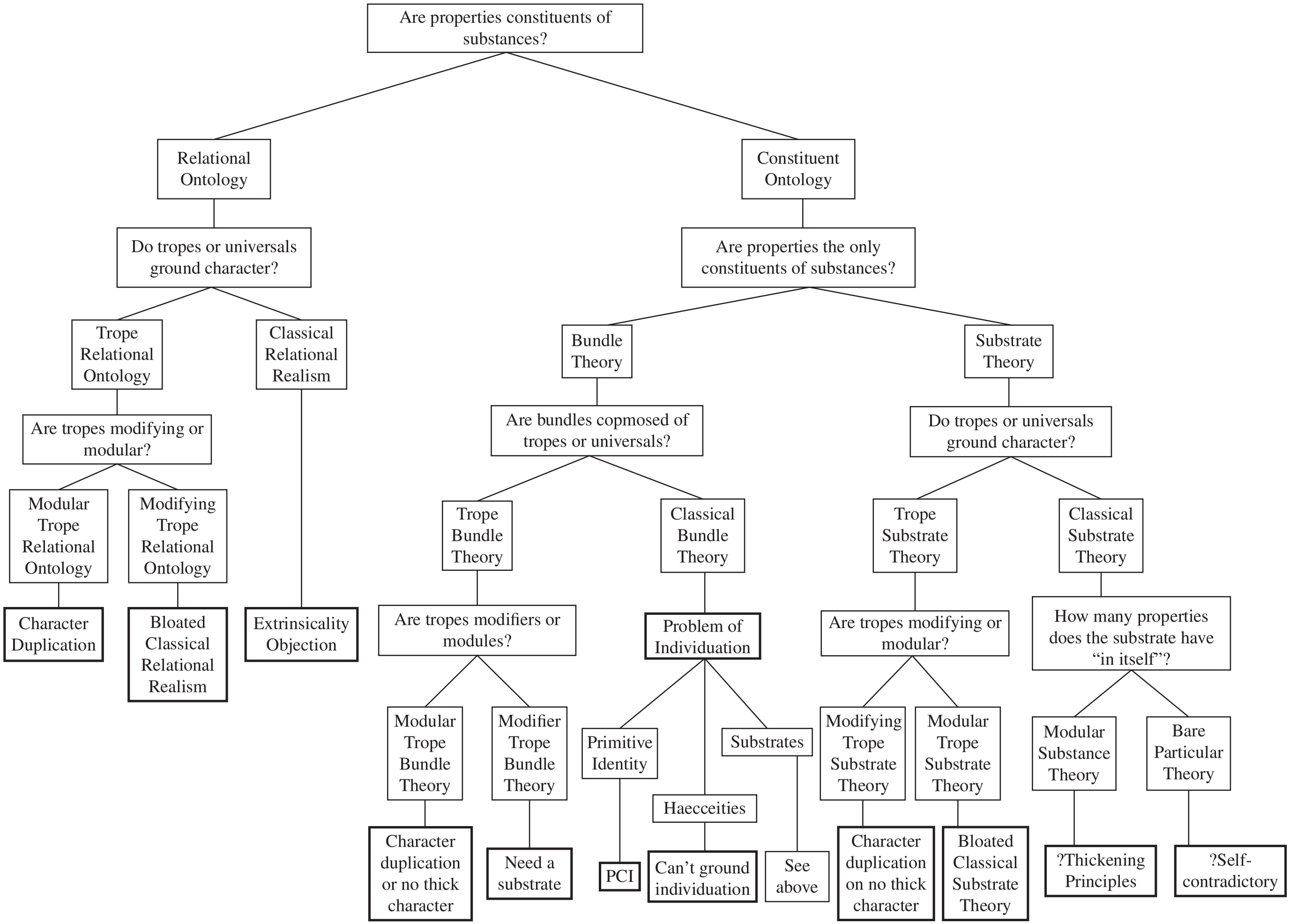


**Figure 9.1** Theories of Substances

particular (or bare particulars) over time should be *sufficient* to ground the persistence of a substance.

However, it does not seem obvious that inclusion of the same bare particulars is necessary for the persistence of a substance. Return to Locke's story of personal identity despite soul-substitution. In much the same way, it seems that it would be possible, in some cases, for a substance to survive the substitution of its bare particular by another. We will take up these questions of persistence again in Chapters 24–25.

The results of this section are summarized in the Figure 9.1.

As we have seen, Realism can be divided into Relational and Constituent Ontologies. Relational Ontology has two principal drawbacks: it must posit a primitive *instantiation* relation, while Constituent Ontologists can make do with the *part-of* relation, and Relational Ontology faces the Extrinsicality Objection. Neither is decisive. Much depends on whether the Constituent Ontologists can make good on the promises of their program.

Constituent Ontology divides into Bundle Theory and Substrate Theory. The most serious problem for Bundle Theory is that of individuation, making room for the possibility of distinct but indiscernible particulars. Substrate Theory solves that problem by positing substrates whose ontological job is to ground distinctness. Bare Particular Theory has the added advantage of immunity to the Hochberg-Armstrong objection, as well as avoiding the weird combinations of modular tropes with modifying tropes or universals that are required by the alternative, Modular Substance Theory. In the next chapter, we consider how Bare Particular Theory, as well as Ostrich Nominalism and Resemblance Nominalism, fare as accounts of relational facts.

## Notes

1 Somewhat surprisingly, David Armstrong is a Realist who believes in states of affairs and not nexuses, and yet who takes the second option. He does so by distinguishing two kinds of *part-whole* relations: *mereological* composition (resulting in wholes composed by every collection of things) and *non-mereological* composition (resulting in the special entities, states of affairs). He never explains what makes a whole of the first kind into a whole of the second kind if not the presence of a nexus.

2 PCI corresponds closely to the assumption of Strong Supplementation in the theory of classical mereology (see Section 23.1).

3 Cf. Pickavance (forthcoming).

4 This is a matter of some dispute, but we have not taken the question up here. Aristotle's Third Man argument is, famously, an argument against Plato's claim that universals are self-exemplifying.

5 For discussion, see e.g., Alston (1954), Anscombe (1964), Moreland (1998), Pickavance (2014), Sider (2006).

# 10

# Relations, Structures, and Quantities

In this chapter, we will examine four special problems involving properties (whether universals or tropes). First, in Section 10.1, we will look at various accounts of relational facts, facts that involve properties relating two or more particulars. Then, in Section 10.2, we will examine an important special case of relational facts: those that involve non-symmetric or ordering relations. In Section 10.3, we turn to structural properties, those relational properties that enable many things to form a single structure, like a group or a team. Finally, we take up the problem of measurable quantities in Section 10.4.

## 10.1 Accounts of Relational Facts

All Realists (7.1T) must answer this question: are there *relational* universals, in addition to *qualitative* ones? Trope Nominalists (8.1T.4.1A) have to answer a similar question: are there relational modifying tropes or modular tropes? Each of the following propositions attributes a relation to a pair of things:

(1) Dallas *is to the north of* Austin.
(2) John *loves* Mary.
(3) Texas is *next to* Oklahoma.
(4) Particular *P instantiates* universal *U*.
(5) Universals $U_1$ and $U_2$ are *co-instantiated* (or bundled).
(6) Modifying trope *A resemble*s modifying trope *B* exactly.
(7) Modular trope *C resembles* modular trope *D* exactly.

Are there universals corresponding to the relational property of *being to the north of* or *loving*? If so, must UP-Realists (7.1T.1T) recognize the universal INSTANTIATION?

*The Atlas of Reality: A Comprehensive Guide to Metaphysics*, First Edition.
Robert C. Koons and Timothy H. Pickavance.

Must Classical Bundle Theorists (9.1T.1T.1A) treat BEING BUNDLED itself as a universal?

Similarly, Modifying Trope Theorists (8.2T.1T) must consider whether there is a relational modifying trope connecting Dallas and Austin, a modifying trope that makes (1) true? For Modular Trope Theorists (8.2T.1A), is there a single modular trope, *something to the north of the other*, and if so, is this modular trope a part of both Dallas and Austin or a part of some whole that contains Dallas and Austin as parts? Are there modifying tropes of *resemblance* or merely-resembling modular tropes?

In each case, to answer Yes is to run the risk of a version of Bradley's Regress. Consequently, many UP-Realists, Classical Bundle Theorists, and Trope Nominalists have restricted their accounts to monadic (non-relational or one-place) properties. Doing so still provides all of these theories with an advantage in terms of economy over Ostrich Nominalism (7.1A.1A), since Ostrich Nominalists will have to recognize the same number of fundamental relational kinds as do the others, while the others will have far fewer fundamental monadic kinds.

One extreme position that such anti-relational theorists could take is simply to deny that there are any fundamentally relational truths at all. Such a theory was proposed by Leibniz, who called his unrelated substances 'monads'. Hence, we say that Monadism is the view that all relational truths are grounded in the intrinsic character of substances taken individually:

**10.1T Monadism.** There are no fundamental relational truths.

**10.1A Anti-Monadism.** There are some fundamental relational truths.

Monadists must reject the existence of all external or non-internal relations. We defined internal relations in Chapter 2:

**Def D2.2 Internal Relation.** *R* is an *internal relation* if and only if, necessarily, for every *x* and *y*, whether *R* holds between *x* and *y* depends only on the intrinsic properties of *x* and of *y*.

The expression 'depends on' as we used it in this definition was intentionally ambiguous. It could simply mean that the holding or non-holding of the relation weakly supervenes on the intrinsic properties of the two relata (so that it is impossible for two situations to be the same with respect to those intrinsic properties and yet different with respect to the holding of the relation), or it could mean that the relation's holding is wholly grounded in the intrinsic properties of the two relata. Let's call these two more precise notions 'weakly internal' and 'strongly internal', respectively:

**Def. D10.1.1 Weakly internal.** *R* is a *weakly internal* relation if and only if, necessarily, for every *x* and *y*, whether *R* holds between *x* and *y* weakly supervenes on the set of facts about the intrinsic properties of *x* and *y*.

**Def. D10.1.2 Strongly internal.** *R* is a *strongly internal* relation if and only if, necessarily, for every *x* and *y*, whenever *R* holds between *x* and *y*, the fact that *R* holds between *x* and *y* is wholly grounded by the facts about the intrinsic properties of *x* and *y*.

Monadists must not only deny that there are any external relations: they must also deny that there are any weakly internal relations. If Monadism is true, all relations must be strongly internal.

What about the relation between parts and wholes? Is this inclusion relation strongly or weakly internal? That is, if *x* is a part of *y*, is this part-of relation external or internal? If internal, is it weakly or strongly internal? This question will turn on whether the relations of identity and distinctness are internal or external. If *x* is part of *y*, and if *being identical to x* is an intrinsic property of *x*, then *x*'s being part of *y* will count as in intrinsic property of *y*, since, for any *F* that is an intrinsic property of a part of *x*, there will be a corresponding intrinsic property of *y*, namely, the property of *having a part that is F*. So, if *being identical to x* is an intrinsic property of *x*, and *x* is a part of *y*, then *having a part that is identical to x* will be an intrinsic property of *y*. But is *being identical to x* an intrinsic property of *x*? This is precisely the question we examined in some detail in Chapter 8. As we saw there, all Bundle Theorists, including those who believe in fundamental haeacceities, hold that identity is an intrinsic property. All others must treat *being identical to x* as an extrinsic property of *x*, and so they must count the part-whole relation as an external relation.

One thing to notice immediately is that only Ostrich Nominalists and Bundle Theorists can be Monadists, since all Substrate Theorists and all Reductive Nominalists accept the existence of some fundamental relational truths: either truths about *instantiation* (Classical Relational Ontologists 9.1A.1T) or *composition* (Constituent Ontologists 9.1A who are also Substrate Theorists) or *resemblance* or *class membership* (Class or Resemblance Nominalists 8.1T.3/8.1T.4). All of these relations will be either external or only weakly internal. For example, *instantiation* (for Classical Relational Ontologists) and *resemblance* (for Resemblance Nominalists) must be only weakly internal relations, since whether a particular instantiates a universal (for Classical Relational Ontologists) or resembles another particular (for Resemblance Nominalists) cannot be grounded in the intrinsic character of the particular, since it is the instantiation or resemblance relation that grounds that intrinsic character (not vice versa).

In addition, spatial relations pose a problem for Monadists. Being spatially contiguous or being a certain distance apart appear to be fundamental relational truths about material objects. Leibniz believed that this was not so, that each monad carried within it a spatial representation of the entire universe, with some element in that representation standing for that monad's own unique perspective. However, Leibniz's account seems to require some fundamental spatial or quasi-spatial relations between the parts of these spatial representations. In addition, his account seems to require representation itself as a fundamental relation.

An initially more promising direction for Monadists to take is to adopt the Theory of Spatial Qualities (see Chapter 17.1T.1T). On this view, to be located somewhere in space is simply to instantiate some intrinsic quality that is uniquely associated with that region. Distance between regions corresponds to the degree of dissimilarity between two such *spatial qualities*. On this view, *spatial contiguity* and *spatial distance* are internal relations, fully determined by the intrinsic character of the two material objects.

However, this merely pushes the problem of accounting for spatial relations back a step, since we still have to consider what makes one spatial location close to another one. We will need to posit spatial relations between spatial qualities, but the existence of such fundamental relational facts, even facts between universals, is inconsistent with Monadism.

Leibniz proposed a version of Monadism in which everything was conscious. The world consisted of an infinite number of souls, each of which represented a spatially-ordered world centered upon itself by a simple and non-relational act of representation. The public space of the physical world was, for Leibniz, a kind of projection pieced together from the internal representations of each of the conscious monads.

The most difficult problem for Leibniz's version of Monadism is giving a full and adequate theory of monadic representation. A monad's representing the world as being a certain way must consist of a set of intrinsic mental acts. These acts cannot be related to each other in the way that pixels in a picture relate to each other, or in the way words in a sentence relate to each other, since these would sneak fundamental external relations back into the picture. Since Leibniz was a Bundle Theorist, he can help himself to the part-whole relation, and to the relations of identity and distinctness (since these will all be strongly internal relations, given his commitments). Nonetheless, it's going to be a very difficult task to construct an adequate theory of mental representations of space without anything like spatial relations to work with.

Given these problems, few philosophers have embraced Monadism. So we turn to varieties of Anti-Monadism.

When we turn to relational properties, we find that some Realists about monadic universals and some Moderate (or Trope) Nominalists embrace Separatism (an Extreme Nominalistic attitude toward relational universals), in contrast to Connectionism:

**10.1A.1T Connectionism.** Some relational universals or relational tropes are fundamental entities.

**10.1A.1A Separatism.** There are no fundamental non-symmetric relational universals or relational tropes.[1]

Connectionism comes in several varieties. First, there is the distinction between Universal Connectionists, who believe in non-symmetric relational universals, and Trope Connectionists, who believe in non-symmetric relational tropes (either modular or modifying).

**10.1A.1T.1T Classical Connectionism.** There are relational universals.

**10.1A.1T.1A Trope Connectionism.** There are relational tropes that are fundamental, but no relational universals.

Second, among Classical Connectionists we can distinguish between those who embrace relational nexuses and those who instead rely on relational states of affairs as truthmakers. (We'll ignore for the sake of simplicity Classical Connectionists who reject

Truthmaker Theory entirely.) A Nexus Connectionist adopts what is essentially a version of a Relational Ontology (9.1T) for non-symmetric relations, treating *instantiation* itself as a weakly internal relation, while State of Affairs Connectionists adopt a version of Constituent Ontology, with the pair (corresponding to the state of affairs) actually containing the relation as a proper part.

**10.1A.1T.1T.1T Relational (Nexus) Connectionism.** When some particulars instantiate some relational universal, the truthmaker for the corresponding proposition is a relational nexus that ties those things to the universal. The nexus does not contain either the relation or the relata as parts.

**10.1A.1T.1T.1T.1A Constituent (State of Affairs) Connectionism.** When some particulars instantiate some relational universal, the truthmaker for the corresponding proposition is a state of affairs that contains those things and that universal as proper parts.

We could draw a similar distinction between two kinds of Trope Connectionists: Relational Trope Connectionists, who posit modifying relational tropes that modify each of the two relata, and Constituent Trope Connectionists, who posit states of affairs that contain a relational trope (either modular or modifying).

Connectionists have typically made a distinction between *real* relations and *merely logical* relations. The real relations are universals that are instantiated by pairs of things, while logical relations (including *instantiation* itself) are irreducible relational kinds, with no universal involved. This enables Realists to claim a further advantage in economy over Ostrich Nominalists, reducing the number of primitive relational kinds to one. We must avoid making the mistake of assuming that the "real" relations are somehow more real than the "merely logical" ones. In fact, the opposite is the case: it is the merely logical relations (like *instantiation* or *resemblance*) that are metaphysically more fundamental than the real relations. The use of the term 'real' here reflects an older, Latin tradition. It might be better to translate think of *real* relations as *things-ish* relations. A *real* relation is a relation that can be *reified*, that is, thought of as a *thing* (the Latin word for 'thing' is 'res', the root of 'real').

A similar move is available to Trope Nominalists. They can also distinguish between real relations, corresponding to relational tropes (either modifying or modular), and logical relations (such as *resemblance* and *being bundled*).

Separatists also come in two varieties. Just as there are Ostrich Nominalists and Reductive Nominalists, so are there Ostrich Separatists and Reductive Separatists. Ostrich Separatists take relational truths that involve external relations, like 'Oklahoma is north of Texas', to be metaphysically fundamental, with no need for explanation in terms of the instantiation of universals, resemblance of tropes or any other metaphysical device. (Internal relations, like *resemblance*, can be explained in terms of the monadic properties of the two relata.)

**10.1A.1A.1T Ostrich Separatism.** Separatism is true, and there is no general explanation of why some things taken in a certain order resemble other things taken in a certain order.

**Table 10.1** Combining Theories about Monadic Properties and Relations

| | *Ostrich Separatism* | *Resemblance Separatism* | *Relational Connect-ionism* | *Trope Constituent Connect-ionism* | *Classical Constituent Connect-ionism* |
|---|---|---|---|---|---|
| Ostrich Nominalism | Yes | **No** | **No** | **No** | **No** |
| Resemblance Nominalism | Yes | Yes | **No** | **No** | **No** |
| Classical Relational Ontology | Yes | Yes | Yes | **No** | **No** |
| Constituent Trope Theory | Yes | Yes | Yes | Yes | **No** |
| Classical Constituent Ontology | Yes | Yes | Yes | **No** | Yes |

**10.1A.1A.1A Reductive Separatism.** Separatism is true, and there is some general explanation of why some things taken in a certain order resemble other things taken in a certain order.

As in the case of Reductive Nominalism, the most plausible version of Reductive Separatism would be Resemblance Separatism, positing some fundamental relation of *resemblance* between the things involved in a certain kind of relational fact.

There is an obvious parallel between the accounts of monadic properties that we examined in Chapters 7–9 and the various theories of relations. In general, those who adopt a more nominalistic account of monadic properties should adopt an equally or more nominalistic account of relations. It would be very weird, for example, to combine Ostrich Nominalism for monadic properties with a Classical Connectionist account of relations. Among Realists and Trope Theorists, it makes sense to take a position on relations that exactly mirrors their views about monadic properties. Here is a table indicating which theory combinations are plausible.

Now that we have surveyed the various metaphysical options for Anti-Monadism, we shall next turn to the most difficult test case: that of non-symmetric relations (or orderings).

## 10.2 Non-Symmetrical Relations and the Problem of Order

A *symmetric* relation $R$ is one of such a kind that $Rxy$ always entails $Ryx$.

**Def D10.2 Symmetric Relation.** A relation $R$ is *symmetric* if and only if, necessarily, for all $x$ and $y$, if $x$ stands in $R$ to $y$, then $y$ stands in $R$ to $x$.[2]

A non-symmetric relation involves no such entailment. *Being next to* is symmetric, while *being to the north of* is non-symmetric. *Belonging to the same club* is symmetric, but *loving* and *hating* are non-symmetric.

A symmetric relation can always be thought of as a simple property of two things taken collectively. Universal Connectionists could suppose that Texas and Oklahoma jointly instantiate the relational universal BEING NEXT TO. However, non-symmetric relations are different. Connectionists cannot simply say that Texas and Oklahoma jointly instantiate BEING TO THE NORTH OF. Instead, it is Oklahoma and Texas *in that order* that instantiate the relation. The same two states, taken in the reverse order, do not instantiate the BEING TO THE NORTH OF relation, since Texas is not to the north of Oklahoma. In a same way, Modifying Trope Connectionists cannot suppose that Oklahoma and Texas are jointly modified by a relational modifying trope of *being to the north of*.

A similar problem besets Modular Trope Connectionists who take particulars to be bundles of modular tropes. They cannot simply say that there is a *being to the north of* modular trope that belongs to both the Oklahoma-bundle and the Texas-bundle, since this would fail to distinguish the case where Oklahoma is to the north of Texas from the case where Texas is to the north of Oklahoma.

Similar problems beset Reductive Separatists' account of non-symmetric relations. For example, the Resemblance Separatist cannot simply suppose that Cyrano's loving Roxanne consists simply in the fact that the pair of Cyrano and Roxanne resemble other lover-beloved pairs, since this would fail to distinguish Cyrano's loving Roxanne from Roxanne's loving Cyrano, which are obviously two distinct and independent truths.

There is a standard solution to all of these difficulties, one which involves introducing a new kind of set, the *ordered pair*. An ordered pair is a certain kind of set that has sets as its members (a second-order set). In the most popular version (the Kuratowski construction), the ordered pair $<A, B>$ is identified with the set $\{\{A\}, \{A, B\}\}$, that is, with a set with two other sets as its members: the unit set $\{A\}$ and the pair set $\{A, B\}$. In such an ordered pair, the pair set always represents the two objects being ordered, and the unit set represents the first member. A Standard or Ordered-Pair Connectionist can then reduce $A$'s standing in relation $R$ to $B$ in terms of the ordered pair $<A, B>$'s belonging to a set of ordered pairs that is a natural class or a resemblance class containing the paradigm cases of the $R$-relation. For instance, Oklahoma is north of Texas if and only if the ordered pair <Oklahoma, Texas> is in the set of ordered pairs that constitutes the *north-of* relation. Similarly, a Standard or Ordered-Pair Connectionist can suppose that $A$'s standing in relation $R$ to $B$ consists in the ordered pair $<A, B>$'s instantiating the $R$ universal or in some modifying $R$-trope's modifying the ordered pair $<A, B>$.

### 10.2.1 The twin problems of converse relations

As Kit Fine (2000, 2007) has argued, the standard or ordered-pair solution, in both its Reductive Separatist and Connectionist versions, encounters two problems involving converse relations. If $R$ is a non-symmetric relation, then a relation $S$ is the converse of $R$ if and only if whenever $Rxy$ then $Syx$, and vice versa.

**Def D10.3 Converse Relation.** *R* and *S* are *converse relations* if and only if, necessarily and for all *x* and *y*, *Rxy* if and only if *Syx*.

Here are some examples of relations and their converses:

| | |
|---|---|
| *North of* | *South of* |
| *Parent of* | *Child of* |
| *Loves* | *Is loved by* |
| *Greater than* | *Less than* |
| *Above* | *Below* |

Suppose that at least one of each of these pairs is a real and natural relation (i.e., a relation corresponding to a universal). If so, which is it? Is *north of* natural and *south of* unnatural or vice versa? Is *loves* natural and *is loved by* unnatural or the other way around? In each case, two converse relations seem so much alike that it is hard to believe that they could differ in their degree of metaphysical fundamentality. If so, then the Classical Connectionists are forced to say that both relations are fundamental.

**The Metaphysical Parity of Converse Relations.** If *R* and *S* are converse relations, then either both are natural or neither is (that is, either both correspond to relational universals or neither do).

Moreover, if Connectionists and Reductive Separatists were to deny the Metaphysical Parity of Converse Relations, they would have to suppose that some ordered pairs have natural properties other than mathematical ones. We'd have to assume that the ordered pairs <Cyrano, Roxanne> and <Romeo, Juliet> resemble each other (as lover-beloved pairs), but the ordered pairs <Roxanne, Cyrano> and <Juliet, Romeo> don't resemble each other at all! Taking this asymmetry seriously involves forgetting that the sets we use to represent ordered pairs are just that: representations. They shouldn't enter into the metaphysical foundations. As we have argued before, we should try to avoid making sets bear metaphysical relations relevant to the concrete world. *Membership* seems to be the only natural or fundamental relation into which sets are involved:

**PMeta 3 Membership the Only Fundamental Set Relation.** If *S* is a set, then the only fundamental relation involving *S* is the membership relation between *S*'s members and itself. (Principle of Metaphysics 3)

If Classical Connectionists accept the Metaphysical Parity of Converse Relations (as it seems they should), then they immediately face two further problems. First, they must suppose that there is some sort of brute metaphysical necessity that ensures that whenever a non-symmetric relational universal exists, there must also exist a second relational universal, corresponding to its converse. As we have seen, Ockham's Razor directs us to minimize such brute necessities.

Second, the Metaphysical Parity of Converse Relations forces Classical Connectionists to posit a massive duplication of truthmakers. The truth that Cyrano loves Roxanne will now be made true by both the fact that the ordered pair of Cyrano and Roxanne stand (in that order) in the *loves* relation and by the separate fact that the ordered pair of Roxanne and Cyrano stand (in that order) in the distinct *is loved by* relation. Such duplication of truthmakers seems indefensible.

Similar problems face both Trope Connectionists and Reductive Separatists. Trope Connectionists have to suppose that whenever a relational truth of the form *Rab* is true, then there are two relational tropes involved, one for *R* and one for *R*'s converse. Similarly, Natural Class Separatists have to suppose that there are two distinct natural classes involved in grounding the truth that *Rab*: the natural class corresponding to *R* (to which the ordered pair <*a*, *b*> belongs), and the natural class corresponding to the converse of *R* (to which the ordered pair <*b*, *a*> belongs). The Resemblance Separatists face exactly the same kind of duplication: the ordered pair <*a*, *b*> resembles certain paradigms of the *R*-relation, and the ordered pair <*b*, *a*> must resemble paradigms of the converse relation. All of these theories must posit a large, potentially infinite, number of necessary connections between distinct facts.

The only account that is immediately immune to these problems is Ostrich Separatism. Ostrich Separatists believe that all of these relational truths are metaphysically fundamental. Hence, they don't have to offer any explanation of what grounds their truth, and thus they don't have to introduce ordered pairs at all. Ostrich Separatists can suppose that the truth that Cyrano loves Roxanne and the truth that Roxanne is loved by Cyrano are the very same truth, merely expressed in different ways.

Resemblance Separatists and certain Connectionists could avoid this problem by denying that there are any natural non-symmetric relations. For Resemblance Separatists, this would mean that non-symmetric relational truths are never grounded directly in truths about the resemblances among ordered pairs, and, for Connectionists, it would be that there are no non-symmetrical relational universals or tropes. All non-symmetric relations would have to be unnatural and definable in terms of the more fundamental symmetric relations. For example, we could use the symmetric relations of *resemblance* and *co-bundling* to define the non-symmetric relation definable as: *x* resembles something that is co-bundled with *y*.

However, there are many non-symmetric relations that seem to be fully natural and not definable in terms of more fundamental symmetric relations. Consider, for example, the relations of *temporal priority* (*x* is earlier than *y*) or causal priority (*x* causes *y*). These seem to be natural and real non-symmetric relations, not artificial relations definable in terms of more fundamental symmetric relations. Similarly, the three-way relation of *betweenness* (*x* is between *y* and *z*) is non-symmetric and, in many popular developments of the foundations of geometry, more fundamental than such symmetric relations as *distance*.

Connectionists and Reductive Separatists have three options in trying to even the score with Ostrich Separatists by avoiding the use of ordered pairs. First, they can postulate the existence of ordered pluralities that are ordered by various relations (the *taking things in order* solution). Second, they can introduce a family of distinct instantiation relations (or modification relations), making the instantiation relation between a non-symmetric relation and its relata incorporate the non-symmetry within itself. Third, they can use a process of composition to build up complex states of affairs representing non-symmetric relational facts.

### 10.2.2 Taking things in order

The first solution is to take non-symmetric relations to be properties, not of ordered pairs, but of pairs of things *taken in a certain order*. Thus, it is Oklahoma and Texas that instantiate the *north of* relation, when taken in a certain order.[3]

This seems to be an attractive way to develop Relational Connectionism, Modifying Trope Connectionism, and Reductive Separatism.

In cases of non-symmetric relations like *north of* or *loves*, it would involve an undesirable duplication if Classical Connectionists had to posit both the universal NORTH-OF and the universal SOUTH-OF, as well as if we had to posit both the universal LOVES and the universal IS-LOVED-BY. It would be much better to suppose that there is a single universal in both cases. If so, we cannot say that it is Oklahoma and Texas taken in that numerical order (i.e., Oklahoma first and Texas second) that instantiate the relational universal NORTH-OF, since this would mean that the two states taken in that order would also instantiate the identical relation SOUTH-OF. Instead, we have to build an appropriate order into the pair: it is Oklahoma and Texas, *taken in the order of Oklahoma to the north of Texas*, that instantiate the NORTH-OF and the SOUTH-OF relation. However, this would seem to be an unsuccessful reduction of the primitive kind *being to the north of* to the kind *instantiating*, since we need the *being to the north of* relation itself to specify the ordered pluralities that are supposed to instantiate the supposed relational universal.

There are other cases, however, in which the reduction might be made to work. Consider the relation *loving* (and *being-loved-by*). We could say that Cyrano and Roxanne, taken in the order of Cyrano the *agent* to Roxanne the *patient*, stand together in the *loving* relation. We have to treat the *agent/patient* relation as a primitive, irreducible kind, in order to specify the ordered pluralities that then instantiate the various action relations: *loving*, *moving*, *creating*, *destroying*, *healing*, *harming*, and so on. The Classical Connectionists can now achieve a significant gain in qualitative ontological economy by positing action-universals in place of the Ostrich Separatists' primitive relational kinds of *loving*, *moving*, and so on. This would be true even if the Classical Connectionists had to posit two or three primitive kinds of *instantiation*: the instantiation of monadic universals by single things, the joint instantiation of symmetrical relational universals by pairs of things, and the joint instantiation of non-symmetric action-universals by pairs of things taken in an agent/patient order.

Modifying Trope Connectionists, assuming that they adopt Substrate Theory (9.1T.1A) rather than Bundle Theory (9.1T.1T), can claim a similar economy for their account. A single *love* modifying trope could stand in the *agent modifying* relation to Cyrano and the *patient modifying* relation to Roxanne.

### 10.2.3 A family of distinct, complex instantiation relations

We can accomplish much the same effect by replacing ordered pluralities with a family of complex instantiation relations. So, instead of supposing that there is a single instantiation relation that connects snow with WHITENESS and the ordered plurality of Cyrano and Roxanne (with Cyrano as agent and Roxanne as patient) with LOVE, we could instead suppose that there are at least two distinct instantation relations: a binary relation that connects individuals with monadic properties or qualities (like WHITENESS) and a three-place relation that combines two individuals with an action-universal (like LOVES). We will also need a distinct instantiation relation for each class of thematically structured non-symmetric relations: a patient-agent-action instantiation relation (for action relations like LOVES), a greater-less-comparison instantiation relation (for comparative relations like IS GREATER THAN), an extreme-extreme-intermediate-placement instantiation relation (for 3-place relations like BETWEEN), and so on. It is difficult to say which of the two theories, the ordered-plurality theory or the multiple-instantiation-relation theory, is most economical in terms of qualitative parsimony. We think it's plausible that the ordered-plurality view has a slight edge here, especially if the number of thematically structured relational families is relatively large. Each new instantiation relation introduces two or three new categories of things (e.g., agents and patients, things greater and lesser, things intermediate and extreme), while the corresponding ordered pluralities introduce just one new thing (agent-patient ordered pluralities, greater-lesser ordered pluralities, extreme-intermediate ordered pluralities).

In his 1913 manuscript on the theory of knowledge, Bertrand Russell (1984) offered what seems, at first glance, to be an alternative account of non-symmetric relations. This account is usually called 'positionalism'. On Russell's account, each binary relational universal has two positions or poles, with one relatum attached in some way to each of the two positions on the universal. So, the fact that Romeo loves Juliet involves Romeo's being attached to the *lover* position of LOVES and Juliet's being attached to the *beloved* position. However, Russell's account still leaves the fundamental problem unsolved: what is it for Romeo to be "attached" to a position within a particular fact? It's obviously not sufficient to say that the fact that Romeo loves Juliet consists simply in Romeo's instantiating the lover position and Juliet's instantiating the beloved position, since this leaves undecided whom Romeo loves or who loves Juliet. It provides no connection between Romeo as lover with Juliet as beloved. Whether the universal LOVES has two positions or not, we still need some account of what it is for the pair to occupy those two positions jointly, as part of a single fact.

Accounting for this joint occupation of the two positions requires something like associating each of the pair with one of the two roles (which is what the ordered-plurality theory proposes) or providing a single relation that simultaneously associates Romeo with one position and Juliet with the other (which is what the multiple-instantiation-relation theory proposes). In either case, the positing of different positions within the universal seems unnecessary. There is, however, a third option, one that may come closer to Russell's intention: we could think of particular facts as constructed from particulars and *parts* of universals. We will take this possibility up as option 2 in the next sub-section.

### 10.2.4 Constituent Ontologies and non-symmetric relations

With a little imagination, Constituent Ontologists can solve these problems, whether they take the constituents of states of affairs to be universals or modular tropes. For simplicity's sake, we will consider two options for Classical Connectionism. Trope Connectionism can also employ counterparts of these two options.

Both options make use of the non-symmetric *part-whole* relation, and thus are faced with an obvious objection: Doesn't appealing to *part-whole* make both accounts viciously circular?

In short, No. Constituent Ontologists can make use of the familiar distinction between real and merely logical relations. The *part-whole* relation is not a real relation, in the sense that there is no PART-WHOLE universal (nor are there any *part-whole* relational tropes). In a sense, the *part-whole* relation is so metaphysically fundamental that it cannot be analyzed in terms of instantiating or including any universal or trope.

Constituent Ontologists ask us to accept one non-symmetric relation, the *part-whole* relation, as a primitive, with no analysis of what makes it true. In effect, Constituent Ontologists treat the *part-whole* relation the way Ostrich Separatists treat all relations. By reducing all other relational facts to facts involving only the *part-whole* relation, Constituent Ontologists offer a much simpler theory of the world, in terms of the qualitative diversity of things.

In addition, a version of the Extrinsicality Objection (Section 9.3.1.1) can be wielded by Constituent Connectionists against Relational Connectionism. A natural relation is intrinsic to its relata, taken together. A binary relation is intrinsic to the pair, taken as a pair. We shouldn't have to introduce a separate, third thing to act as the truthmaker for a relational truth. However, this is just what Relational Connectionists have to do. Any *R*-related pair is related by *R* in virtue of the relata being extrinsically related (by the *instantiation* relation) to *R*.

In contrast, Constituent Connectionists can distinguish between *thin* and *thick* pairs. Thin pairs consist of two relata and their constituents. Thick pairs include two relata and any relational universals that belong to the pair as a pair. Relational facts are intrinsic on this view to the thick pair, which is as it should be: they are intrinsic to the pair *taken together*, as a pair.

Having dispatched the circularity worry, we turn to a consideration of the two options for solving the problems to with non-symmetric relations. The first option involves two-step constructions; the second involves divided universals.

OPTION 1: TWO-STEP CONSTRUCTIONS The first option borrows an idea from set theory and the representation of ordered pairs. Standard set theory represents the order of an ordered pair by building up the ordered pair in two steps. First, the first member of the pair is put into a unit set, and both members are put into a pair set. Second, the unit set (containing the first member) and the pair set (containing both members) are combined into a single set. Constituent Connectionists can do something similar, using the *part-whole* relation. If Cyrano loves Roxanne, they could suppose that there is a complex state of affairs, comprising exactly two parts. The first part combines Cyrano with the relation of LOVE, and the second part consists of Roxanne alone. The order of the *loving* relation corresponds to the fact that it is Cyrano, and not Roxanne, who forms a whole with LOVE. The fact that it is Roxanne who is the object of Cyrano's love corresponds to the fact that the Cyrano-LOVE complex combines with Roxanne to form a single state of affairs. (This solution presupposes the denial of Mereological Universalism (22.3T.1), the thesis that there is a whole composed of any set of members whatsoever.) Here is a tree-diagram (Figure 10.1) representation of two-step constructions (with parts arranged below the wholes that contain them):

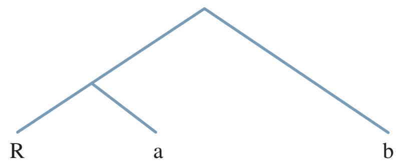


**Figure 10.1** A Two-Step Construction of a Binary Relational Fact

The same relationship can also be represented by means of circles, with parts inside their containing wholes (Figure 10.2):

The main drawback with this account is one that it shares with the various standard, ordered-pair accounts of non-symmetric relations: it requires that we deny the Metaphysical Parity of Converse Relations. We have to assume that, in reality, there is a metaphysical order that corresponds more closely to one of a pair of converse relations than to the other. If we think of the inner relatum as prior to the outer relatum, then we have to decide whether the truthmaker of the truth that Cyrano loves Roxanne puts Cyrano first or Roxanne. It would seem to be impossible for us to tell which is in fact the case.

However, this constituent account of order has one advantage over the standard account: it doesn't entail that any sets (such as ordered pairs) have non-mathematical natural properties. So it doesn't contradict Membership the Only Fundamental Set Relation (PMeta 3).

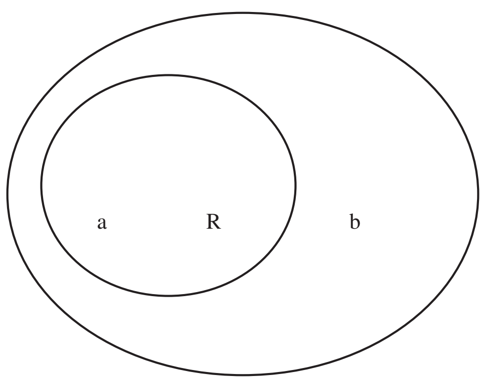


**Figure 10.2** A Pictorial Representation of the Two-Step Construction

OPTION 2: DIVIDED UNIVERSALS In option 2, Constituent Connectionists suppose that each relational universal consists of two parts, one representing one *pole* of the relation and the other the second pole. This is a version of what Kit Fine (2000) calls "positionalism". For example, in the case of the universal LOVE, it would contain two parts: a LOVER part and a BELOVED part. When Cyrano loves Roxanne, there will be a single state of affairs with two parts: the first part containing Cyrano and LOVER, and the second part containing Roxanne and BELOVED. Here is a diagram of divided universals (Figure 10.3):

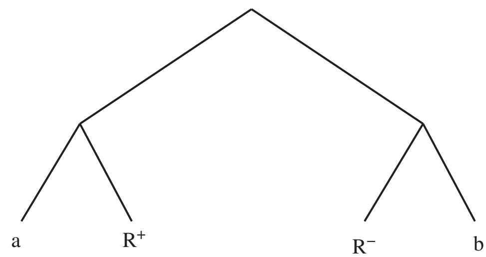


**Figure 10.3** A Divided-Universal Construction of a Binary Relational Fact

The top node of the diagram represents the total relational state of affairs, the state of *a's being R-related to b*. The left node just below the top combines the $R^+$ universal with $a$ (or some essential part of $a$), and the corresponding right node combines the other part of $R$, $R^-$, with $b$ (or some essential part of $b$). Here is the same relationship, in a circle (Venn-like) diagram (Figure 10.3):

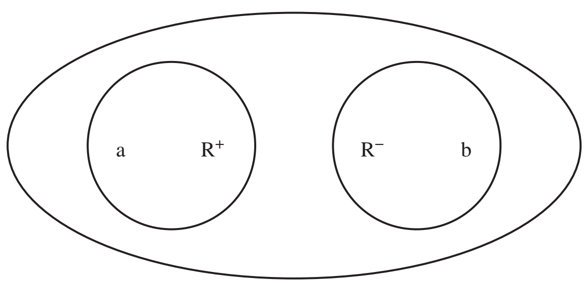


**Figure 10.4** A Pictorial Representation of the Divided-Universal Construction

PROS AND CONS OF THEORIES OF NON-SYMMETRIC RELATIONS We've seen that there are three solutions to the problem of non-symmetric relations: Ostrich Separatism, ordered pluralities (with either Reductive Separatism or Relational Connectionism), and Constituent Connectionism with divisible universals. Which of the three is best? There are two considerations. First, the Extrinsicality Objection tells against Relational Connectionism. Second, we can consult Ockham's Razor, in its qualitative economy version (PMeth 1.4 and 1.4.1). That is, we can try to minimize the number of metaphysically fundamental relations. To answer this question, we have to look at two sub-questions. First, how many natural relations are there? Second, how many types of basic ordering (e.g., *agent/patient*, *greater/lesser*, *medium/extreme*) are there?

1 If there are very few natural relations, then Ostrich Separatism is best. If there are very many, it is worst.
2 If there are very few basic orderings, then both ordered pluralities and Constituent Connectionism are viable theories. If there are very many, then Constituent Connectionism is the best theory.

## 10.3 Structural Universals and Constituent Ontology

David Armstrong (1983, 1989b) has argued for the existence of *structural universals*, universals that correspond to the realization of complex structures. For example, consider the property of *being a methane molecule*. This property can be defined in terms of having four hydrogen atoms and one carbon atom as parts, with each of the hydrogen atoms bound chemically to the central carbon atom. There is certainly such a property, but is there in addition a universal, BEING METHANE, shared by all individual methane molecules? A very simple case of structural universals is that of conjunctive universals, like the universal RED AND ROUND. If we have universals REDNESS and ROUNDNESS, do we need a third universal present in cases of things that are both? Ockham's Razor suggests not, so we need to consider whether there are any countervailing considerations.

David Lewis (1986b) collected six arguments for structural universals.

1 Providing meanings of complex predicates. A complex universal could serve as the meaning of compound predicates, like RED AND ROUND. However, this reason is incompatible with adopting a sparse theory of universals.

2 Accounting for resemblance. One might think that accounting for resemblance is not something that could justify structural universals. For example, if two things are both red and round, we can use the universals REDNESS and ROUNDNESS to explain their similarity. We seem to need a third universal to ground their similarity in having both properties. The sharing of the two simple universals seems to be sufficient. However, to push back against this, one might appeal to what the nineteenth-century German psychologist Christian von Ehrenfels (a student of Franz Brentano) called '*Gestalt* similarities'. A Gestalt quality is a feature of a structured whole that emerges from and supervenes on the instantiation of simpler universals by the parts. Gestalt properties are real and objective and are not in principle reducible to simple facts about the parts. Some examples of Gestalt qualities are artistic styles of individual artists or schools. Paintings can have a holistic quality that marks them as products of the Renaissance, or of Jan Brueghel the Younger. The similarity of two Brueghel paintings is not reducible to pixel-by-pixel comparisons of paint pigments, even though the quality does supervene (Def D2.6) on the details of both paintings, in that once you fix the distribution of colors and strokes on the painting, you have also determined whether it has that ineffable Brueghelish quality. Ethical qualities of actions might also be examples of Gestalt properties.

   To be clear, however, it is not the supervenient Gestalt qualities that are the structural universals. Rather, we need structural universals to be the bearers or supervenience bases for these Gestalt qualities. It is the resemblance between the structural universals that grounds or corresponds to the instantiation of the Gestalt quality by the particular whole. The simplest account of Gestalt qualities would be to take them to be simple, natural qualities of the structural universals themselves. Then it would follow that the Gestalt quality is realized in a particular whole whenever that whole instantiates one of the quality-bearing structural universals.

3 Accounting for possibilities. Peter Forrest (1986) has suggested that we identify possible worlds with structural universals (*world-properties*). We'll take up this suggestion again in Chapter 15, on Abstractionism (14.1T.1A = 15.1T). This requires a moderately abundant theory of structural universals. We don't need disjunctive structural universals (like RED OR SQUARE), but we will need structural universals that are conjunctions and amalgamations of properties and properties of parts if we are to have enough structural universals to supply at least one for each possible state of affairs.

We will also need something on the level of particulars (perhaps bare particulars?) to perform the job of making possible situations actual. We will need such *ultimate actualizers* as the things that instantiate those possible situations that are actual.

4 Providing resources for Strong Nomism (4.4A.2). If there are structural universals, then we could represent each causal law of nature as simply connecting one such structure (the initial situation) with a second structure (the resulting situation). There is, however, an alternative. We could instead posit a family of *lawmaking* relations, one for each logical form of a law (this was Tooley's original proposal in Tooley 1977).

5 Accounting for inclusion and incompatibility among universals. Using structural universals, we can easily explain why, for example, the universal RED AND ROUND includes the universal REDNESS and excludes the universal RED AND NOT ROUND. However, this sort of explanation doesn't apply to logically simple universals. Some philosophers, such as Wittgenstein at the time of writing the *Tractatus* (1921), have thought that all simple universals are mutually compatible and no set of simple universals ever entails the instantiation of other universals. But this is quite a controversial position. There do seem to be cases of incompatible simples, such as incompatible sensory qualities (color, smell, texture).

6 The possibility that there are no simple universals. At the very least, there might be some structural universals that are "structures all the way down", structural universals such that each part of the structure would be another complex structure, with still more parts. The world could contain infinitely complex properties. Lewis considered this to be the strongest argument for structural universals. If it is possible that some structures have no simple parts, then, if there are to be universals that are shareable in such cases, the universals must be complex.

There are two additional arguments (our seventh and eighth arguments) that Lewis did not consider. The seventh (emergent holistic powers) looms large in light of our consideration of causal powers (Chapters 4–6). As we will see in Chapters 22 and 25, there are good grounds for thinking that there may be *emergent* or holistic powers of composite things, powers that are not wholly grounded in the powers of their microscopic parts. This is most plausible in the case of sentient or rational beings. The whole organism in some cases transforms the nature of its parts, altering fundamentally the powers of those parts. Powerism entails that powers are internal to the properties that confer them. Thus, there must be structural universals that can confer these emergent or holistic powers on the composite particulars that instantiate them (or to the proper parts of those particulars).

The eighth argument also concerns composite material things. As we discussed in Chapter 9 and will examine again in Chapters 25, there are cases of composite things that seem to violate Mereological Constancy (25.1A). Such things are *mereologically inconstant*: they can continue to exist despite gaining or losing some of their constituent parts. Organisms, for example, are constantly exchanging material particles with their environment. Some artifacts can survive the loss or replacement of some of their parts (think of a car that survives the replacement of its spark plugs). If a composite thing is to survive such change, there must be a principle of unity that persists despite the substitution of some of the participants in that unity, and structural universals are plausible candidates to play such a unifying role. Katherine Ritchie has been investigating this use of structures in her recent work on social groups (Ritchie 2013, 2015).

Despite these advantages, structural universals are somewhat mysterious. Consider the structure of an $H_2O$ molecule. Since water molecules contain two hydrogen atoms, the universal corresponding to being an $H_2O$ molecule must have two "parts" each of which is the relatively simple universal HYDROGEN ATOM. How can this universal be present "twice over" in the $H_2O$ universal? There is only one HYDROGEN ATOM, which must either be a part or not a part of the $H_2O$ universal. It's not at all clear what being present "twice" could mean.

Even if we could solve this problem, a deeper mystery remains. Some structures have exactly the same parts assembled in different ways. For example, the hydrocarbon structures normal butane and isobutane both contain four carbon atoms, ten hydrogen atoms, and 13 chemical bonds. In normal butane, the four carbon atoms are in a chain, while in isobutane there is a central carbon atom with the other three carbon atoms and a hydrogen atom bound to it. Since there is only one universal CARBON ATOM, how can the two structures (normal butane and isobutane) put that one universal into two different relations to itself and to the hydrogen-atom universal?

In response, Lewis offers the theory of *amphibians*, on behalf of Constituent Ontologists who want to posit structural universals.[4] On this view, there are amphibious entities that are partly universal and partly particular, in addition to universals and particulars. The hydrogen structure contains one oxygen amphibian and two hydrogen amphibians. The two hydrogen amphibians are distinct duplicates, each of which instantiates the common universal HYDROGEN ATOM. However, hydrogen amphibians are not particulars, since they are themselves instantiated by parts of particular water molecules. They both instantiate and are instantiated.

Amphibians are like particulars in another way: they can stand in the same relations to one another that particulars do. Thus, the two hydrogen amphibians in the $H_2O$ structure are chemically bonded to the oxygen amphibian, in the very same way that particular hydrogen atoms are bonded to particular oxygen atoms in actual water. In fact, when a particular water molecule instantiates this structure, the chemical bonds between the particular atoms are identical to the chemical bonds between the instantiated atom-amphibians.

Constituent Connectionism would require *bare amphibians*, elements whose metaphysical function is to distinguish each amphibian from all of the qualitatively indiscernible amphibians. Bare amphibians are thus analogous to bare particulars.

Structures are often organized hierarchically, with some structures serving as components of still larger structures. Such hierarchical nesting of structures requires a similar nesting of structural universals. For example, just as there are atom-amphibians in the structure of the water molecule, there can also be water-molecule-amphibians in the structure of the ice crystal. Likewise, we should find atom-amphibians in the structure of amino acids, amino-acid-amphibians in protein structures, protein-amphibians in the structure of cellular organelles, organelle-amphibians in cell structures, cell-amphibians in organ structures, organ-amphibians in the structure of organisms, and organism-amphibians in social structures. A universal is a constituent of an amphibian; that amphibian is within another amphibian, and so on, finally reaching a particular. In other words, a particular hydrogen atom could realize an amphibian role $A_1$ within an organism, while $A_1$ realizes an amphibian role $A_2$ within a cell, which in turn realizes a third amphibian role $A_3$ within a protein, which in turn realizes a fourth amphibian role $A_4$ within an amino acid, which in turn realizes the universal of being a hydrogen atom. For Constituent Ontologists, this requires a nesting of amphibians within amphibians, with a universal at the bottom and a particular at the top.

Amphibian theory will come in two varieties, depending on the roles that we want structural universals to play. These two varieties are the *very sparse* and *moderately abundant* types. If we rely on Gestalt qualities, holistic powers or incontinent wholes (arguments 2, 7 or 8) as our basis for Realism about structural universals, then we will want to adopt a very sparse theory of structural universals. We should posit structural universals only when the structures are associated with emergent qualities or powers or with the kind of unity required for inconstant persistence.

In contrast, if our motivation for structural universals depends on Forrest's idea of possible worlds as structural universals (argument 3), then we will need to have a moderately abundant theory of structural universals. We will need one structural universal for every possible situation, including whole possible worlds.

If we adopt the sparse version of amphibian theory, then we will recognize three types of entities in the world: universals, amphibians, and particulars. Universals are instantiated (by amphibians and particulars), amphibians both instantiate universals and other amphibians and are instantiated by both particulars and other amphibians, and particulars instantiate universals and amphibians. If we adopt Substrate Theory, we will need both bare amphibians and either bare particulars or modular substrates as the ultimate ground of distinctness among amphibians and particulars. In addition, Constituent Ontology alone requires some entity to play the role of distinguishing particulars (that are essentially incapable of being instantitated by anything else) from amphibians (that are essentially capable of being instantiated by other things). We could give this role to bare particulars or modular substrates, or we could introduce some new kind of entity—an ultimate actualizer or particularizer—to do this job.

The moderately abundant amphibian theory would be especially attractive as a version of Bundle Realism: Amphibian Bundle Theory. On this view, the actual world consists entirely of universals and amphibians. A *particular* is simply a structure that is not in fact (in the actual world) duplicated. Duplication occurs only when a universal is instantiated by more than one amphibian within some actual or realized structure. Thus, a particular is either a maximal actual structure (one that isn't in fact a proper part of any other structure) or a unique (un-duplicated) part of such a maximal actual structure. We might imagine that there is a single structure—the cosmic structure—that includes all of the facts about the locations, properties, and relations of the whole physical universe. This might be the only maximal actual structure, which would give us a version of metaphysical Monism (11.2A). As we shall see in the next chapter, such a theory offers an attractive interpretation of the Monism of F.H. Bradley. Alternatively, we could suppose the cosmic structure is only one of many maximal actual structures, the other structures being partly physical (thus overlapping with the cosmic structure) and partly non-physical (including mental, ethical, or biological features not included in the physical world). This version of Bundle Theory would avoid the problem of individuation that plagued Classical Bundle Theory (9.1T.1T.1A), since bare amphibians would ground the distinctness of indistinguishable parts of a structure (as bare particulars do in Bare Particular Theory 9.1T.1A.2A). This account will require some way of distinguishing merely possible structures from actual ones, a problem we return to in Chapter 15.

## 10.4 Determinables, Quantities, and Real Numbers

As we have already seen, some properties stand in the *determinate-determinable* relation.[5] There is for example, the determinable property of *color*, and there are a large number of determinate colors (*scarlet*, *periwinkle*, *mauve*, and so on). In physics, we have such determinable properties as *mass*, *charge*, and *velocity*, and we also have determinate masses, quantities of charge, and velocities (e.g., *2 kilograms*, *5 coulombs*, or *one-half the speed of light*). There are four basic facts about determinables and determinates to be explained:

1 If anything instantiates a determinable property, it must also instantiate one of the determinates. Anything that is colored must have some specific color.
2 If anything instantiates a determinate property, it must also instantiate each of its determinables. Anything that is scarlet is colored and reddish.
3 If *P* and *Q* are both determinates of the same determinable, then it is impossible for something to instantiate both *P* and *Q* at the same time. It is impossible to instantiate two distinct, specific quantities of mass at the same time.[6]
4 The determinates of a single determinable can be ordered along one or more dimensions of similarity. For example, *1 kilogram* is more similar to *2 kilograms* than it is to *3 kilograms*, and *red* is more similar to *orange* than it is to *green*.

In many cases (physical quantities), these similarity orderings make possible a precise measurement of the determinates of a determinable. That is, the determinates can be assigned unique real numbers (or complex numbers) as measurements (e.g., 3.7 grams, 5 meters/second, etc.), relative to certain conventions about units of measurement and (possibly) a conventional coordinate system (such as latitude and longitude).

These facts can be explained by Realists and by Trope Nominalists in several ways. We will consider three popular theories: Determinate Universal Theory, Simple Intensity Theory, and Composite Intensity Theory. Determinate Universal Theory supposes that each determinate property corresponds to a single universal (or, for Trope Nominalists, to a class of exactly similar tropes). Real numbers on this view are second-order properties, natural properties that are instantiated by ordered pairs of determinate universals. Simple Intensity Theory makes do with a single class of quantitative intensities, while supposing that it is the determinables that are universals. On the Simple Intensity Theory, a substance instantiates a determinate property by jointly instantiating both a determinable (like mass) and a specific intensity. Composite Intensity Theory supposes that each degree of intensity contains all lesser degrees as proper parts. A determinate property of a substance corresponds to the largest or maximal intensity that is instantiated in conjunction with some determinable by the substance.

### 10.4.1 Determinate Universal Theory

Real numbers are used as measures of quantities, both extensive (distance, duration) and intensive (mass, charge, energy). A measure is something essentially relational. Consequently, we could propose, following John Bigelow (1988), that the fundamental role of real numbers is as relations between physical quantities. Thus, real numbers are relational properties. (More precisely, real numbers are relations between scalar quantities. Relations among vector quantities include the complex numbers.)

What are the quantities that real numbers relate? They could either be universals of a kind (exactly so-much mass, exactly so-much charge) or tropes (either modifying or modular).

**10.2T Real numbers are universals.**

**10.2A Real numbers are not universals.**

**10.2A.1T Real numbers are tropes.**

If we are Realists, we could suppose that each real number is a relational universal, one instantiated by pairs of quantities, taken in the *not-less-than/not-greater-than* order.

**10.2T.1 Determinate Universal Theory.** Real numbers are relational universals, each instantiated by ordered pairs of quantities.

For Modifying Trope Nominalists, when two quantity-tropes $q_1$ and $q_2$ stand in the *real-number* relation $r$, then $r$ is a modifying trope that modifies jointly $q_1$ and $q_2$ in the *not-less-than/not-greater-than* order.

**10.2A.1T.1 Determinate Trope Theory.** Real numbers are natural resemblance classes of relational modifying tropes, each of which modifies an ordered pair of quantities.

How does Determinate Universal Theory explain the four facts about determinates and determinables? First, determinable properties correspond to classes of determinate universals that stand in real-numbered relations to one another. Consequently, if a substance instantiates one of those determinate universals, then it will automatically possess the determinable property. Conversely, a substance cannot possess the determinable property without instantiating one of the associated determinate properties. Third, the mutual exclusion of pairs of determinates will have to be treated either as a brute metaphysical necessity or as the consequence of a special law of nature. Finally, the natural ordering of the real numbers corresponds to the similarity ordering of the determinates.

There are thus two major costs to Determinate Universal Theory. First, it requires laws or brute necessities corresponding to the mutual exclusion of distinct determinates of the same determinable (the mutual exclusion of *having exactly 1 gram of mass* and *having exactly 2 grams*, for example). Second, it requires a large class of brute facts. There are so many families of properties (mass, charge, volume, velocity, and so on) that are interrelated in the right way by the entire field of real (or complex) numbers. Both costs involve the postulation of a vast number of brute facts of a surprisingly convenient sort.

There is one further worry associated with the Determinate Universal Theory, namely, the problem of missing, uninstantiated determinates. This is especially a problem for those who embrace the Dretske-Armstrong-Tooley account of natural laws as relations among universals. A law like Newton's law of gravity specifies the force of attraction between any two bodies, given their masses and the distance between them. The law would seem to apply to any quantity of mass whatsoever, including quantities that are never in fact realized in the actual world (there must be such values, if there will have been only finitely many actual bodies, given that the number of possible mass-values is infinite). Many metaphysicians (including, for example, David Armstrong) are reluctant to posit uninstantiated universals, but if all universals are determinates, it seems that either there must exist some uninstantiated universals or the laws of nature are radically *gappy*, failing to apply to an infinite number of actually-unrealized values.

### 10.4.2 Simple Intensity Theory

As we have seen, Determinate Universal Theory has two major drawbacks. It can't explain why distinct determinates of the same determinable exclude each other, and it can't explain why the same set of real numbers can be used to order different kinds of quantities (like mass, charge, distance, and so on). It has to treat the applicability of the same real numbers to different quantitative determinables as a brute coincidence.

We could instead suppose that whenever a substrate $S$ instantiates a determinate quantity $q_1$, $S$ in fact jointly instantiates a *pair* of universals, a determinable $Q$ and an intensity $I$. On this account, the very same class of intensity-universals is instantiated by each determinate quantity, no matter what determinable it belongs to. In other words, the same set of intensity-universals would qualify instances of MASS, CHARGE, ENERGY, FREQUENCY, and so on. Real numbers would then be relations between pairs of intensities, with *intensities* forming a single ordered domain.

Suppose that two substrates $S_1$ and $S_2$ are characterized by determinate mass-quantities $m_1$ and $m_2$ and that the ratio between $m_1$ and $m_2$ corresponds to the real number $r$. According to the Simple Intensity Theory, in such a case $S_1$ is connected to both the mass-determinable universal $M$ and to some intensity-universal $I_1$, and $S_2$ is connected to $M$ and to some intensity-universal $I_2$. The pair of intensities $I_1$ and $I_2$ stand (in that order) in the real-number relation $r$. (A similar account can be given by Trope Nominalists, replacing the universals with tropes.)

**10.2T.2 Simple Intensity Theory (Nexus version).** Positive real numbers are relations between intensities, and intensities are universals that are jointly connected to particulars and determinable universals.

**10.2A.1T.2 Simple Intensity Theory (Trope-nexus version).** Positive real numbers are binary relations between intensity-tropes, and each intensity-trope is jointly attached to both a substrate and a trope of some determinable.

Realists who are Relational Ontologists (treating *instantiation* as an extrinsic relation, not involving *parthood*) will have to complicate their theories by introducing a new, three-place *instantiation* relation (*x jointly instantiates y and z*) in addition to the ordinary, two-place relation that connects particulars to single universals. Similarly, Modifying Trope Theorists will have to postulate a new kind of *modification* relation, one in which two tropes cooperate in modifying their associated concrete particular.

What about Realists and Trope Theorists who embrace Constituent Ontology? On these views, a particular instantiates a property by actually containing a universal or a trope as a part. What would it mean for two universals (or two tropes) to be a *joint part* of the particular? We will have to suppose that particulars have an internal compositional structure—that they are not simple bundles of properties. If a particular $x$ has a determinable quantity $Q$ with intensity $I$, then there must be a special part of $x$ that contains both $Q$ and $I$ as parts. In addition, suppose $x$ also has a different quantity $Q_2$ with a different intensity $I_2$. Then $x$ will have two complex parts, $y_1$ and $y_2$, with $y_1$ composed of $Q$ and $I$, and $y_2$ composed of $Q_2$ and $I_2$. At the same time, we will have to deny that any combination of parts of $x$ compose a further part. For example, we will have to say that $x$ has no part composed of just $Q$ and $I_2$, and no part composed of just $Q_2$ and $I$.

The situation is somewhat simpler if we have bare particulars in our ontology, in addition to universals or tropes. Then we could suppose that, if $x$ is a bare particular with two quantities at two different intensities, $x$ simply belongs to two different bundles, one containing the first quantity and its intensity, and the second bundle containing the second quantity and its intensity. The bare particular can tie all of these diverse bundles together into a single thick, ordinary particular.

THE SELF-APPLICABILITY OF REAL NUMBERS There is another simplifying modification that can be made to Simple Intensity Theory. Rather than assuming that there are two classes of universals, the class of intensities and the class of real numbers, the two classes can be identified. The identification is possible because pairs of real numbers stand in ratios that are themselves real numbers. If $r_1$ and $r_2$ are real numbers, then there is a real number $r_3 = (r_1/r_2)$, which can be thought of as the intensity of the *greater-than* relationship between $r_1$ and $r_2$. In other words, GREATER-THAN is itself a determinable universal that is instantiated by ordered pairs of numbers to varying intensities. We can identify the real number 2 with that intensity that is jointly instantiated (along with the GREATER-THAN determinable) by the ordered pair of real numbers 2 and 1. That is, given a set of real numbers and a set of intensities, there is a natural way of mapping each on the other. This gives us good grounds for supposing that there is in fact only one set of universals.

We can use facts about physical parts and whole to identify the intensity that is the number 2. Suppose that a massive object $x$ with determinate mass $m_2$ has two separate parts, each of which is equal in mass to $m_1$. We know then that $m_2$ is twice (in terms of mass) $m_1$, and so the intensity of the GREATER-THAN relation between $m_2$ and $m_1$ must be equal to the number 2. Once we have identified the number 2, we can identify other intensities in a principled way. For example, the number 4 is that intensity that stands in the GREATER-THAN relation to the number 2 with an intensity that is itself equal to 2.

The class $I$ of quantitative intensity-universals is ordered by the asymmetric relation GREATER-LESSER. This relation itself is modified by the very same class $I$ of quantitative intensities. That is, if intensity-universal $I_1$ is greater than intensity-universal $I_2$, then these universals (taken in the *not-less-than/not-greater-than* order) jointly instantiate the GREATER-LESSER universal and some intensity-universal $I_3$, where $I_3$ is another member of the class $I$.

If we assume that the intensity-universals are related to one another by a primitive additivity relation of the form $I_\mathrm{i} + I_\mathrm{j} = I_\mathrm{k}$ which satisfies the laws of commutativity and associativity, then we can identify each of the universals in $I$ with a unique real number. Suppose, for example, that $I' + I' = I''$. Then the intensity $I^*$ of the GREATER-THAN relation between $I''$ and $I'$ can be identified with the real number 2.[7]

How does Simple Intensity Theory explain the four basic facts about determinates and determinables? First, Simple Intensity Theorists must suppose that the nature of each determinable universal is such that it can only be instantiated jointly with some intensity. Second, it follows immediately that a substance cannot possess a determinate property without instantiating the corresponding universal, since possessing the determinate property consists in jointly instantiating the determinable and some intensity. Third, Simple Intensity Theorists must again suppose that it is either a brute metaphysical necessity or a law of nature that no substance can instantiate two different intensities in conjunction with the same determinable. Finally, the ordering of the real-number intensities explains the similarity ordering among determinates.

Thus, Simple Intensity Theory resolves one of the problems faced by Determinate Universal Theory. It can explain why the same system of real-numbered relations is realized in many different families of properties. However, it shares with Determinate Universal Theory the second cost since it cannot explain the mutual incompatibility of different intensities.

Simple Intensity Theory helps somewhat with the problem of uninstantiated determinates and gappy laws. So long as each intensity is realized in conjunction with some determinable quantity or other, the fact that specific masses or volumes are unrealized won't matter. However, given that there seem to be only finitely many determinable quantities, it seems almost certain that infinitely many intensities will be uninstantiated anyway.

There is one further cost to Simple Intensity Theory: it entails that there are metaphysically natural units of measurement, in addition to all of the familiar, conventional ones. For example, there is some specific mass that is the property of *having the determinable mass to degree 1* (the *metaphysical gram*). These metaphysical measures are empirically undetectable, and the positing of such empirically inaccessible facts comes at some cost.

### 10.4.3 Composite Intensity Theory

On the final theory, Composite Intensity Theory, there are once again both determinable universals and intensity-universals, but each intensity-universal is of the form: HAVING SOME QUANTITY TO AT LEAST DEGREE *D*. Thus, there is no universal corresponding to the having of exactly one degree of intensity, but there are intensity-universals of the following kinds:

($IU_1$) HAVING INTENSITY TO AT LEAST DEGREE 0.1.
($IU_2$) HAVING INTENSITY TO AT LEAST DEGREE 1.
($IU_3$) HAVING INTENSITY TO AT LEAST DEGREE 3.
etc.

Moreover, each determinate of the form HAVING INTENSITY TO AT LEAST DEGREE *D* contains as proper parts all of the determinates of the same form for any $D^* < D$. Thus, $IU_2$ above contains $IU_1$ as a proper part, and $IU_3$ contains both $IU_2$ and $IU_1$ as proper parts.

A substance's having a determinate quantity of mass, such as exactly 3 grams, corresponds to its instantiating the determinable MASS together with the intensity-universal $IU_3$ and *not* instantiating MASS together with any intensity-universal that contains $IU_3$ as a proper part. In other words, the determinate properties of a substance correspond to the maximal intensities of each quantity that the substance instantiates.

**10.2T.3 Composite Intensity Theory.** Positive real numbers are intensities, each positive real number contains all smaller positive real numbers as proper parts, and the having of a determinate property consists in the joint instantiation of a determinable universal and an intensity.

Composite Intensity Theory can explain why two determinates of the same determinable are logically incompatible and why the same system of real numbers applies to many different determinables. It is logically impossible to instantiate two different intensities of the same determinable, since one intensity must contain the other, and a determinate property corresponds to whichever intensity is maximal. Second, the same system of intensities is realized together with each of the various quantitative determinables—at least, for those scalar quantities having a natural zero: mass, charge, volume, momentum, temperature, entropy, and so on.

Composite Intensity Theory can also resolve, to a large extent, the problem of missing determinates and gappy laws. If an intensity of degree *D* is contained by any actual substance, then every intensity of degree less than *D* will also be contained, since they will all be proper parts of the universal DEGREE AT LEAST *D*. We will still be missing all of the very large intensities, intensities so large that nothing in fact realizes them in conjunction with any determinable. But the class of actual intensities will not be riddled with gaps or holes.

## 10.5 Conclusion and Preview

In Part III (Chapters 7–10), we have developed various views of the nature of facts and substances, with an eye toward understanding the relationship between properties and these different types of particulars. We turn next to further questions about substances. In particular, we will consider whether there really are any such things as substances, and if there are, how many there are. This is taken up in Chapter 11, on Nihilism and Monism. Later, we consider whether substances are anything more than collections of sensory properties, that is, whether there really is a physical world. This is Chapter 13, on Solipsism and Idealism.

## Notes

1 A *symmetric* relation *R* is one such that *Rxy* entails *Ryx*. The relation of *identity* and that of *being the same size as* are paradigm symmetric relations. Non-symmetric relations aren't symmetric. See below for more.

2 There are similar properties of symmetry that can be defined for ternary (three-term), quaternary (four-term), and other relations with still larger numbers of terms. Most of our subsequent discussion will apply to such relations as well.

3 This would involve extending the plural logic of George Boolos (1984) to cover not only unordered but also ordered pluralities. Boolos provides a logic that quantifies over, not just individual things (corresponding to singular pronouns like 'it'), but also over pluralities (corresponding to plural pronouns like 'they'). We can build on Boolos's theory by supposing that some plural variables and pronouns stand not just for some things but for some things in a particular order. If, for example, we say (concerning Texas and Oklahoma) that they stand in the *north to south* relation, we mean by 'they' not just the two states but the two states taken in the Texas-first and Oklahoma-second order. This seems to make sense, even if we don't suppose that we are talking about some further single entity, like an ordered pair (as defined in set theory).

4 Lewis rejected his own suggestion as "too bizarre to be taken seriously", but bizarreness, like beauty, is in the eye of the beholder. Amphibian theory is an attractive approach to structural universals for Constituent Ontologists.

5 This distinction was first introduced by W.E. Johnson (1921–1924). For further developments, see Prior (1949), Searle (1959), and Johansson (1989).

6 There is some controversy about whether sensory qualities like color or taste satisfy this condition. There are cases in which something can at least appear to have two different colors at the same time.

7 We can give a similar account for vector quantities, like velocity, acceleration or momentum (that is, quantities with both a scalar measure and a direction). We could suppose there to be, for example, a single universal of VELOCITY, instantiated by velocity tropes. Each velocity trope is attached to its substrate by a trope-nexus, which in turn instantiates some vector-intensity universal. The set of vector-intensity universals would form a three-dimensional space of vectors. Relations between vectors (comparing their scalar value and direction) would themselves have vector-intensities, which would enable us to identify a plane of vector-intensity universals that constitute the *complex numbers*. Each complex number is itself a vector, and so the intensities of the comparative relation between complex numbers are themselves complex numbers (vectors in the same plane).

# Part IV

# The Nature of Reality

# 11

# Nihilism and Monism

In this chapter, we look at two big questions: does anything exist, and, if so, does more than one thing exist? In Section 11.1, we consider the possibility of Nihilism, that nothing exists, ('nihil' is Latin for 'nothing') and its alternative, Aliquidism, that something exists ('aliquid' is Latin for 'something'). This will lead us into an investigation of the point of positing existing things. Then, in Section 11.2, we look at the debate between Monists, who believe in only one thing, and Pluralists, who believe in many. We consider both radical and more moderate forms of both Nihilism and Monism, including, for example, Priority Monism—the view that there exists only one *fundamental* thing.

## 11.1 Nihilism and Aliquidism

We have discussed the issue of universals and particulars, including views according to which the world contains both universals and particulars (UP-Realism 7.1T.1T), only particulars (almost all forms of Nominalism 7.1A) or only universals (Classical Bundle Theory 9.1T.1T.1A). Before moving on, we need to address a still more fundamental question: why think there is anything at all? We will use the terms Nihilism and Aliquidism to represent the two options:

**11.1T Aliquidism.** Something exists.

**11.1A Nihilism.** Nothing exists.

Why think that Aliquidism is true? Common sense, of course, suggests that many things exist, things like rocks, trees, clouds, rivers, people, artifacts, numbers, features and aspects, thoughts, propositions, facts, events, states of affairs, times, and regions of

*The Atlas of Reality: A Comprehensive Guide to Metaphysics*, First Edition.
Robert C. Koons and Timothy H. Pickavance.

space. In addition, René Descartes produced a celebrated argument, the *Cogito*, which purports to prove that at least "I" (the maker of the cogito argument, whoever that happens to be) exist. We could express the Cogito ('cogito' is Latin for 'I think') in the following form:

1 I think that some things exist.
2 Either I am right or I am mistaken.
3 If I am right, then some things exist.
4 If I am mistaken, then I am thinking something false.
5 If I am thinking anything at all (whether true or false), then I (the thinker) exist. (*Cogito ergo sum*, in Latin.)
6 Therefore, at least one thing exists.

Descartes's Cogito depends on two things. First, when I am thinking something, it seems to me that I am thinking something. Second, it is impossible for the skeptic to convince me that I am wrong about this, since if the skeptic were to succeed, I would come to think that I am mistaken in my thought, which still entails that I am thinking that very thought. Thus, the appearance to me of my own present thought is incapable of being defeated by any skeptical challenge.

Some skeptics, including Georg Lichtenberg and George Santayana (1955), have challenged premise 5. Lichtenberg argued that all we can conclude is something like 'Thinking is occurring.' Descartes has not established the necessity of an existing thinker as the subject of that thought.

What could it mean to embrace Nihilism? There are two ways of being a Nihilist, one radical and one more moderate. On the radical interpretation, Nihilists endorse a claim that entails that almost everything we ordinarily believe is simply false. There are no people (not even oneself), and no things, places, properties, facts, processes, and so on. The only truths there are are (a) the truth that nothing exists and (b) anything that follows logically from this (e.g., that no giraffes exist).

Nihilism can seem incredible. It invites the answer that G.E. Moore (1939) gave to those who denied the existence of the *external world*. His response was simply to point out the obviousness of (1):

(1) Here is a hand.

Moore argued that any philosophical reason we might have for denying (1) is obviously less credible to each of us than the fact that one's own hand exists. Likewise, we firmly believe that many things exist, and these beliefs seem to be obviously true. Any philosophical argument attempting to establish Nihilism will include at least one premise that is less obviously true than the obvious fact that things exist. Therefore, we could never be in a position to rationally deny Aliquidism.

However, we might interpret Nihilism in a less radical way, as the claim that no *things* exist. Such Nihilists could agree that most of what we ordinarily believe is true or could be counted as true if properly interpreted. So, it's okay to believe that there are human hands, so long as we don't infer that hands are *things* or *objects*.

However, for Nihilism to be a serious position, we have to take 'thing' and 'object' in the broadest possible sense. Anything of any kind, whether universal or particular, physical or mental, concrete or abstract, must count as a *thing*, in the relevant sense. If so, how can it make sense to say that is true that there are hands, if no *thing* of any kind exists? The only possible answer to this question would be one that makes use of some kind of paraphrase. That is, Nihilists must argue that a statement like 'Hands exist' is true but misleading in form. It should be taken as asserting something that can be better and more perspicuously expressed in some form other than 'Something exists and is a hand.'

This move is sometimes put this way: the ontologist "speaks with the vulgar but thinks with the wise". Nihilists can take many of the things we ordinarily say, like 'Here is a hand', as true, so long as they are understood in the proper way. In the context of expressing one's real ontological commitments, one must paraphrase such loose expressions of the facts into a form that perspicuously reveals what one really takes to be the case. Nihilists would then have to offer a paraphrase of statements like 'Here is a hand' that does not give the misleading appearance of entailing the existence of anything whatsoever.

This method of paraphrase could itself be understood in a number of ways. Nihilists might claim that we don't really mean what we seem to be saying when we assert that some things exist. In this case, that's an implausible account of our intended meanings. We really do believe in the existence of at least one thing. It would be more reasonable for moderate Nihilists to claim that our ordinary beliefs are approximately true, or close to being true, or "quasi-true" (in the terminology of Sider 2001). Nothing *really* exists, but our ordinary statements of belief in the existence of things correspond to genuine truths in a systematic way, a way that explains the practical usefulness of those ordinary beliefs.

Some, including process theorists like Alfred North Whitehead, have suggested that our belief in the existence of things is merely a product of the subject/predicate structure of our language, including the prominence that that natural language gives to nouns, both proper (e.g., 'Socrates', 'Paris') and common ('rock', 'cow', 'field'). If our natural language were quite different, as it seems it might have been, we might have made much greater use, perhaps even exclusive use, of verbs and adverbs. If so, we might have been less inclined to believe in a world that consists of things (the correlates of nouns).

More recently, John Hawthorne and Andrew Cortens have suggested three different versions of such moderate Nihilism (O'Leary-Hawthorne and Cortens 1995):

1 Nihilists might reject discrete objects in favor of a plurality of *stuffs* like water, blood, steel, and so on. There are no objects, in the sense of discrete, countable things. We should never assert that there are *N F*'s, for any number *N* and *count noun F*. Instead, there is just so much water, blood, steel, and so on. This approach appeals to the linguistic distinction between *count nouns*, nouns that can take the plural form and can be enumerated (like 'people', 'cars', 'pieces of metal', 'rocks'), and *mass nouns*, nouns that never take the plural form, cannot be combined with numerals, but which can instead be combined with phrases of quantity (e.g., 'so many gallons of milk', 'so many yards of fabric', 'so many tons of steel'). Nihilists could renounce all count nouns (in the context of perspicuous statements of ontology), replacing 'Here is a cat' with 'There is some cat-stuff here.'
2 As a further step, Nihilists might posit only one stuff, the "world stuff". Instead of saying 'Here is a cat', Nihilists could say, 'The world-stuff is feline here.'

3 Finally, Nihilists could make use exclusively of P.F. Strawson's (1959) "feature-placing sentences", sentences of the form 'It is *G*-ing *F*-ishly.' 'Here is a hungry cat' becomes 'It is felinizing hungrily here.'

Strawson's proposal is the most comprehensive and radical, since we might think that a stuff (like water or gold) is a kind of *thing*, which would result in, at most, a form of Monism (11.2A below), the belief in only one thing. The Strawsonian approach suggests that the subject-predicate (noun-verb) structure of ordinary sentences is misleading, since it suggests that the noun phrases refer to things. Nihilists who prefer the Strawsonian language will replace all nouns with verbs and adverbs in something like the following way:

(2) 'Socrates exists' becomes 'Socratizing happens.'
(3) 'Socrates is pale' becomes 'Being-pale happens Socrates-wise' or 'Socratizing happens palely.'

What about transitive verbs?

(4) 'Socrates teaches Plato' becomes 'Teaching happens Socratically and to-Plato-wise.'

On this view, things don't exist. Instead, processes happen or progress or unfold. But, don't *processes* then exist? No, they *happen*. Is this merely a verbal dispute? Is 'happening' just what we call *existence* when processes are involved? Aren't Nihilists merely proposing an odd reform of language without really changing our beliefs about what exists?

Perhaps not. In a way, the form 'Socratizing happens' is still misleading, since it suggests that something (a certain process of Socratizing) exists. The clearest form is purely verbal: 'It Socratizes', with the 'it' as a dummy subject. Compare (5) to our ordinary sentence (6):

(5) It Socratizes.
(6) It is raining.

It makes no sense to ask, *what* is raining? The 'it' in this sentence is not supposed to refer to anything at all. One might reply that the 'it' in (6) refers to the atmosphere (or some part of the atmosphere near a point of reference). However, we could imagine it raining meteorites on some planet with no atmosphere at all. One could then perhaps take 'it' to refer to some region of empty space, but surely such a sentence would not commit us to asserting the real existence of empty space.

Aliquidists could respond that shifting from nouns to verbs doesn't really help Nihilists, since 'exists' is simply the most generic of all verbs. If something runs or eats or is pale, then something exists. Similarly, if something Socratizes, then something exists. Thus, from 'It Socratizes', we should be able to infer 'It exists.'

Nihilists can agree with this. They can admit that 'It exists' is true, but still deny that *something* exists. It can be true that it is raining without its being true that something is raining. So, from 'It exists', it doesn't follow that something exists.

But won't Nihilists have to admit that something, namely Nihilism itself, is true? If so, then there must exist at least one proposition (or claim or thought or belief), and so Nihilism would be self-defeating. This is a difficult problem for Nihilists, but they might be able to claim something of the form 'It thinks truthfully-and-nihilistically', while denying that there are such *things* as propositions, truths, and so on, to say nothing of thinkers or believers.

There seems to be a Nihilist position that isn't completely crazy (with emphasis on 'completely'). However, why should we tempted to think that this view is *true*? What do we gain by denying that anything exists?

Nihilists can appeal to Ockham's Razor (**PMeth 1**). Nihilists could argue that their position is the most satisfactory from the point of view of Ockham's Razor, since Nihilists do not posit the existence of any entities at all. No theory could be less encumbered with the supposed existence of things than Nihilism.

We must ask, however, why we should think that Ockham's Razor is a good principle. A typical answer involves an appeal to simplicity: we should always prefer the simplest theory that explains all the facts. This is only rational, though, if simpler theories are more likely to be true, if simplicity by itself is an indicator of truth. But why should that be so? Couldn't reality be complex rather than simple?

These are difficult questions. What's required for a theory to be attractive is more than just its simplicity. What attracts us is the combination of great simplicity with a close fit to a large and diverse set of phenomena. The combination of these three (simplicity, closeness of fit, and breadth of explanatory success) does seem to be a reliable indicator of the truth in many cases. We certainly find such a combination convincing in science. Newton's theory is a classic case of a very simple theory that perfectly fits the data in a wide diversity of applications, from planetary orbits to traffic collisions. In the law and in practical affairs, a simple explanation that is confirmed by a rich diversity of evidence can attain certainty beyond any reasonable doubt. Where reality is very complex, or where reality is simple but a large body of precise data is inaccessible, we may simply be unable to find the truth with any confidence. However, we can always hope for the best. We can hope that we find simple metaphysical theories that fit closely to a wide range of facts drawn from many disparate domains. If we can find such combinations of theoretical virtues, we would have good reason to embrace the metaphysical theory possessing them.

Is Nihilism a simple theory? Simplicity involves much more than simply minimizing the number of entities in one's theory. By that standard, Nihilism is surely the simplest. However, a much more important measure of simplicity looks at the number of fundamental truths contained in a theory. In addition, we are keenly interested in qualitative economy (**PMeth 1.4** and **1.4.1**), and so the number of fundamental or primitive words and phrases in the standard formulation of the theory is even more important. Since Nihilists eschew the use of nouns, this means that we must look at the number of primitive verbs and adverbs required to express a moderate Nihilist view of the world.

By these new standards, Nihilism doesn't seem to be any simpler than our ordinary, common sense view of the world, since every true statement in ordinary English corresponds to a true statement in Nihilism's funny, Strawsonian language: 'Socrates teaches Plato' corresponds to 'It Socratizes teachingly to-Plato-ly.' At first glance, moderate Nihilists' account of the facts doesn't seem to be any simpler than Aliquidists' account.

Nihilists might argue that their account is simpler on the grounds that Nihilist accounts of the world need not address questions of identity and distinctness that Aliquidist approaches must settle. We might suppose that whenever Aliquidists posit some things, Aliquidism (if complete) must include propositions concerning the identity and distinctness of those things, both at a time and across a span of time. If, for example, Aliquidists suppose that the world contains things called 'electrons', it seems that we can always ask how many electrons exist at this or that time. Do any electrons that exist at one point of time persist until a later time? Nihilists, on the other hand, need not suppose that there are any propositions of this kind, since there are no things to stand in a relation of distinctness to one another.

However, when we describe the world by means of relational language, we need to know whether one relatum is identical or distinct from another, as in the following cases:

(7) Socrates teaches Plato.
(8) John likes Mary.
(9) Boston is north of New York City.

Take love for example: if *A* loves *B* and B loves *C*, it makes all the difference whether *C* is or is not identical to *A*. If *A* and *C* are identical, then we have the happy condition of requited love: *A* loves *B* and *B* loves *C* (=*A*) back. If instead *A* and *C* are distinct, then we have the reality of the dreaded love triangle where *A* loves *B*, but *B* loves some third person, *C*. If Nihilists are committed to being moderates, then, in order to avoid the burden of overturning all or most of our common sense opinions, they will want to treat such claims about the relational structure of the world as true, when properly interpreted.

In response, Nihilists will have to reverse themselves, admitting that we can, after all, speak truthfully about the identity and distinctness of the relata of relations. Nihilists might introduce peculiar adverbs, like 'distinctfully' and 'identically'. In this case, Nihilism would seem to be a mere verbal variant from Aliquidism. Alternatively, Nihilists could perhaps distinguish Nihilism from Aliquidism by insisting that claims about identity and distinctness are *reducible* to facts about the adverbial qualification of the corresponding verb phrases. To say that Socrates is distinct from Plato (i.e., that it Socratizes in a distinct-from-Platonizing way) is to say that there is some adverbial qualification ('wisely', 'palely', 'humbly', etc.) that applies to Socrates but not to Plato. For example, it Socratizes palely, but it does not Platonize palely. To say that Twain is identical to Clemens is just to say that there is nothing that distinguishes Twainizing (being Twain) from Clemensizing (being Clemens). This version of Nihilism leads to some version of Bundle Theory (9.1T.1T). Of course, Nihilism is quite an unusual kind of Bundle Theory, since Nihilists deny that there are such things as bundles or properties. In place of reifying properties, Bundle Nihilists will simply use verbs and adverbs, and in place of reifying bundles, they will talk in a *bundley* way, linking verbs and adverbs together by conjunctions like 'and' and 'while'.

Since we have already looked at the pros and cons of Bundle Theory in Chapter 9, let's assume that moderate Nihilists reject the Identity of Indiscernibles (7.3T.1) and, with it, Bundle Theory.

In the absence of Bundle Theory, the only alternative to talking about identity and distinctness among things would be to enrich our vocabulary with new, primitive adverbs designed to capture the relevant facts about the relational structure of the world. For example, if we have a case of reciprocal love, we could say that 'Love happens reciprocally', while if the love is unrequited, we could say instead that 'Love happens unrequitedly.' Similarly, if a relation $R$ holds among three distinct things in the fashion $Rab$ and $Rac$ (with $b \neq c$) we would have to have recourse to a special adverb expressing exactly this structure, something like: 'It $R$s in an open-angle-ish way.' However, at this point, we have good reason to suspect that that Nihilism is not going to turn out to be simpler than Aliquidism. In fact, it seems likely that Nihilism will be considerably more complex in its primitive elements.

Consider, for example, a world in which there are five objects, *A*, *B*, *C*, *D*, and *E*, which bear the following set of relations to one another:

(10) *A* loves *B*, *C*, and *E*, but neither *A* nor *D*.
(11) *B* loves *B* and *D*, but not *A*, *C* or *E*.
(12) *C* loves *A* only.
(13) *D* loves *A*, *B*, *C*, and *D*, but not *E*.
(14) *E* loves *A*, *B*, *C*, *D*, and *E*.
(15) Each of *A*, *B*, *C*, *D*, and *E* is distinct from the others.

Here is a diagram of the network of relations:

Let's call this network, LoveNetwork. Aliquidists can provide a complete description of LoveNetwork by means of 50 simple statements, each of the form '*X* loves *Y*', '*X* does not love *Y*', '*X* is identical to *Y*', or '*X* is not identical to *Y*.' Nihilists will instead require a single adverb, 'LoveNetwork-ly'. The whole situation can be expressed Nihilistically by saying this:

(16) It loves LoveNetwork-ly.

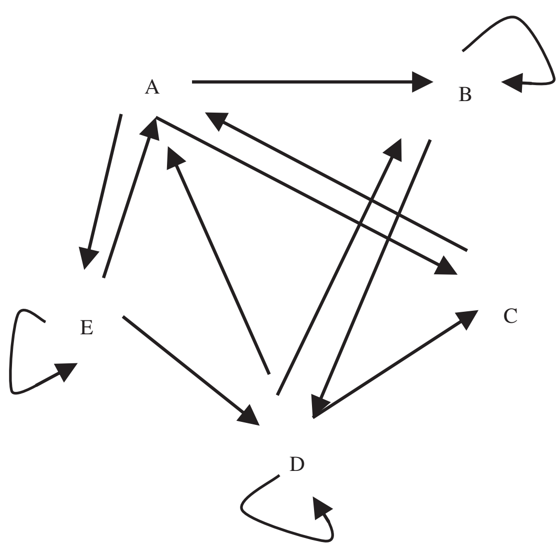


**Figure 11.1** The Love Network

At first glance, this seems to be a much simpler account of the facts than Aliquidists can give. However, appearances are deceiving. Aliquidists can state the facts about LoveNetwork by means of only two transitive verb phrases, 'loves' and 'is identical to'. In fact, Aliquidists can state all of the facts about any love network by means of these two elements. In contrast, Nihilists require a new adverb for every possible network. If we consider networks of just five members, there are a total of $2^{25}$ (more than 33,000,000) different network structures. Nihilists need distinct adverbs for each of these. As a result, Nihilism requires millions of fundamental laws of nature, each expressed by different adverbs, resulting in a horrendously complex theory.

In addition, even if Nihilism were the simplest theory of the world, we wouldn't have good reason to embrace it unless it were adequate to our data. We want a theory that accords with the facts as we know them pre-theoretically. Our ethical, moral, and legal practices presuppose the existence of things like persons and pieces of property. In addition, whenever we make choices in light of what we take to be our own self-interest and the interest of the people we care about, we presuppose that we and our loved ones exist. It seems reasonable to take the presuppositions of our established, natural practices as true, at least provisionally:

**PEpist 2.1 Ethical Practices Presumption.** It is prima facie plausible to suppose that all of the presuppositions of our fundamental ethical beliefs and practices are true.

The Ethical Practices Presumption is only a presumption. We're not begging the question against Nihilism. Given a compelling argument for Nihilism, the presumption in favor of those beliefs embedded in our practices could be overridden. However, it seems reasonable to treat these beliefs as true until proven false. Surely we have prima facie reason to believe such things as (17–19):

(17) Each person has a duty of gratitude toward those who have benefited him or her.
(18) One ought to compensate others for damage one has done to their property.
(19) Each person has a right to be treated with respect through impartial procedures.

We may be convinced, by overwhelmingly powerful philosophical arguments, that we are wrong to suppose that (17–19) correspond to reality, but, at the beginning of our inquiry, these things seem to be every bit as certain as the deliverances of our senses or our memories. They come close, at least, to what David Lewis has called 'Moorean facts' (after G.E. Moore), facts so apparently obvious that good philosophical theories should treat them as fixed points of data.

Even if moderate Nihilists offer a paraphrase of these statements into statements acceptable to Nihilism, the paraphrases would seem necessarily to leave out something crucial to the ethical and legal content of these ordinary beliefs. Rights presuppose rights-bearers, and responsibilities (including debts of guilt or of gratitude) presuppose the continued existence of the one incurring the debt. Paraphrasing these statements into Strawsonian feature-placing sentences would force us to revise many deeply held ethical beliefs.

## 11.2 Monism

Granting that Aliquidism is very likely to be true, we can now ask how many things there are. In particular, is there only one thing, or is there more than one?

**11.2T Pluralism.** More than one thing exists.

**11.2A Monism.** Exactly one thing exists.

By 'thing', do Monists mean anything at all or only particular things? We'll interpret Monism as the commitment to the existence of exactly one thing, whether that thing is conceived of as a universal or particular (or neither or both). Thus, Monists must either be Nominalists or Bundle Theorists of an odd sort, since the only possible bundle contains just the one universal.

Unlike Nihilism, Monism has been defended by a number of prominent philosophers. It can be difficult to tell about some historically important metaphysical theories (e.g., those of Parmenides, Baruch Spinoza or F.H. Bradley) whether they should be interpreted as cases of Nihilism or of Monism. As we saw above, Nihilism could be expressed by saying something like 'It Socratizes wisely', with the pronoun 'it' serving merely as a grammatical dummy, corresponding to nothing at all. Alternatively, the proposition that it Socratizes wisely could be interpreted as an expression of Monism: the one thing that actually exists Socratizes wisely.

The motivation for Monism could rely largely on Ockham's Razor (**PMeth 1**). If we can explain all of the phenomena with the postulation of a single thing, why postulate more? Of course, Ockham's Razor demands only that we avoid multiplying entities needlessly. We do have direct evidence for the existence of a plurality of things. Our sense perception and memory abound with examples of such evidence.

**PEpist 3 Appearance of Bodies and Minds.** Perception and memory present us with what are apparently distinct physical things, including some embodying apparently distinct minds.

**PEpist 4 Presumption of Reliable Perception.** It is prima facie plausible to suppose that human perception and memory are reliable.

In light of Appearance of Bodies and Minds and Presumption of Reliable Perception, we ought to accept that there really are distinct physical things and distinct minds, unless there is a metaphysical argument for rejecting the presentations of sense perception and memory. Thus, an appeal to Ockham's Razor alone is not a sufficient case for Monism.

As in the case of Nihilism, so too in this case we can distinguish between extreme and moderate forms. Extreme Monists endorse a strong claim, one that entails that all of the evidence of our senses and memory is simply and thoroughly wrong. For extreme Monists, the appearance of plurality is a mere illusion. In contrast, moderate Monists modify the thesis in such a way that they can accept that much of what we believe on the basis of our senses is true, if properly understood. Moderate Monists offer paraphrases of our ordinary beliefs in terms of the nature of a single particular, or they might claim that there are real facts about ordinary things but facts that are wholly grounded in facts about the one, uniquely fundamental thing.

We can also distinguish between two kinds of moderate Monists—a strong version for those who defend a conceptual grounding of the Many in the One, and a weaker version who defend an extra-conceptual grounding of the Many in the One (using the distinction we introduced in Section 3.4). Conceptual reductionists hold that the reduction of the Many to the One is licensed solely by facts about our *concepts* about the plurality of the things in our world. We are supposed to conclude that these concepts, although *ostensibly* true of many particulars, are *in fact* made true about facts about the One Real Thing. Ontological reductionists, in contrast, believe that the many things exist and have real essences, but these essences guarantee that all of the facts about the many things can be fully explained by reference to facts about the One Fundamental Thing.

Conceptually Reductionist Monists can claim that their view is superior, in terms of ontological economy, since they can deny that many things *really* exist, while Ontological Reductionist Monists cannot make this claim. Extra-Conceptual Reductionists can, however, claim that their theory has some advantage in economy over the thesis that there are many fundamental things, since the Monists reduce the number of fundamental things as well as the number of fundamental truths.

Moderate Monists, whether Conceptual or Extra-Conceptual Reductionists, must embrace some version of Ostrich Nominalism (7.1A.1A) if they believe that the One is a particular. Since Monism entails that only one thing exists, there cannot exist both universals and particulars. Thus, Monists must postulate either a single universal (and no particulars) or a single *fundamental* particular (and no universals). For simplicity's sake, let's assume that Monists' One Thing is a fundamental particular. Moderate Monists must suppose that many things are true of this one thing (corresponding to each of the many common sense beliefs we have). Moderate Monists must take each of these many things to be primitive, irreducible facts about the one thing. Hence, moderate Monists must embrace the Ostrich Nominalist account of predication.

In what follows, we examine four arguments for Monism: those of Parmenides, Spinoza, Bradley, and Jonathan Schaffer. It is likely that all four of these philosophers were moderate Monists. The latter three are explicitly so. In the case of Parmenides, the historical evidence is sketchy, but the latter half of Parmenides' philosophical poem suggests that a kind of truth can be attributed to beliefs about the physical world.

The distinction between conceptual and extra-conceptual reduction is a somewhat more subtle distinction, and it can be difficult at times to discern which sort of Monism a given philosopher is defending. Prima facie, Parmenides and Bradley seem to be making the stronger claim of Conceptual Reductionism, while it seems pretty clear that both Spinoza and Schaffer are defending the weaker claim of Extra-Conceptual Reduction.

### 11.2.1 Parmenidean Monism

Parmenides of Elea authored a single work, a poem traditionally titled *On Nature*. In the first part of this poem, Parmenides seems to be arguing for Monism. We reconstruct his argument as one designed to show that we cannot know or even believe that more than one thing exists, rather than to provide a positive case for Monism. The center of Parmenides's argument is the claim that we cannot think about *non-being*. He assumes that thought (and related things, like knowledge and reference) is a binary relation, a relation between the thinker and the object of thought. Whenever a relation holds, it seems to be necessary that both relata really exist. Hence, one can never think of anything non-existent. This is an issue that we will investigate in some detail in the next chapter, in Section 12.1. As we shall see, Meinongians will disagree with Parmenides on just this point. However, most contemporary philosophers reject the Meinongian solution.

Parmenides would then have to show that we could not think about a situation consisting of two or more things without thinking about something non-existent. It seems reasonable to suppose that, if two distinct things exist, then there must be something (some fact) that distinguishes one from the other. It could be that one has a quality, like *redness* or *roundness*, that the other lacks. At the very least, if there are two things *A* and *B*, it would seem that *B* would have to lack the property of *being identical to A*, and *A* would have to lack the property of *being identical to B*. So, to think of *A* and *B* as distinct, we would have to attribute negative properties to each of them. Parmenides argues that attributing a negative property to something necessarily involves thinking about the absence of the corresponding positive property. But to think about an absence is necessarily to think about the thing that is absent, which is to think of something non-existent (see Kirk, Raven, and Schofield 1983).

Observe, first of all, that even if this argument were successful, it wouldn't establish Monism. It doesn't even show that we couldn't *believe*, or even *know*, that there are two or more things. It only purports to show that we cannot *think* of two particular things as distinct. Just because I cannot think about particular cases of distinctness, it does not follow that I could not know that cases of distinctness exist. For example, I cannot think about a natural number so big that it cannot be thought of by me, but I can nonetheless believe that there are such numbers.

We can use the resources of Truthmaker Theory (2.1T) to recast Parmenides' argument into a form that demonstrates the impossibility of there *being* more than one thing (and not just of our *thinking* of them). This argument would depend on the assumption that every proposition ascribing distinctness to a pair of things must have a truthmaker, and this truthmaker must involve the attribution of contradictory properties (one positive, and one negative) to the two distinct things. The rest of the argument proceeds in a similar fashion:

1 If there exist two distinct things *A* and *B*, then the truth of their distinctness must have a truthmaker.
2 This truthmaker must make true the attribution of contradictory properties to *A* and *B*, one positive property and the other a negative property.
3 Such a truthmaker must include two truthmakers as parts. One must make it true that *A* has *F* and the other that *B* does *not* have *F*, for some positive property *F*. Let's call the truthmaker of *B*'s not having *F* '*N*'.
4 In order to make it true that *B* does not have *F*, *N* must make it true that the possible state of affairs of *B*'s having *F* is *non-existent*.
5 The truthmaker of the non-existence of any *x* must include *x* itself as a part (since that is the only way that it can specify that it is *x*, in particular, that doesn't exist).

6 To be part of an existing truthmaker, a thing must exist.
7 If the truthmaker of the non-existence of *x* exists, then *x* itself cannot exist (since the existence of a truthmaker for any proposition ensures the truth of that proposition.)
8 So, there is no truthmaker of the non-existence of *x* (for any *x*), since for the state of affairs to exist, *x* would have to both exist and not exist (from 5, 6, and 7).
9 Therefore, in particular, *N* cannot exist. (From 4 and 8)
10 Therefore, there cannot exist two distinct things. (From 1–4 and 9)

Every premise of this argument can be challenged. First of all, we might doubt whether a true proposition of distinctness needs a truthmaker at all. Second, as we discussed in Chapter 9, it is plausible to think that the only truthmaker that is needed for the distinctness of *A* and *B* is the pair *A* and *B*. This version of Parmenides's argument assumes that two indiscernible but distinct things are impossible, which the Max Black thought experiment of two indistinguishable spheres gave us reason to doubt.

Premise 3 is also doubtful, since it requires the assumption of Truthmaker Maximalism. Negative truths, like *B*'s not having *F*, may not require truthmakers: they might instead correspond to the *absence* of a truthmaker of *B*'s having *F*. Even if we did need truthmakers for negative truths, premise 4 assumes that such negative truthmakers must attribute non-existence to a possible state of affairs. Why couldn't *N* instead simply verify some negative relation (like *non-instantiation*) between *B* and *F*? Why couldn't *N* be a Totality Fact (see Section 2.4.2), a fact connecting *F*-ness with the totality of things that instantiate, which verifies *B*'s not having *F* simply by not including *B* in that totality?

### 11.2.2 Spinoza's Monism

Baruch Spinoza offered a set of metaphysical arguments that could be interpreted as attempting to show that only one thing exists. In fact, Spinoza argued that there is just one *substance*, where a substance is defined as something that exists *in itself* and that is conceived of *through itself*. Everything else, including people, stars, and motes of dust, are modes or attributes of that substance. The interpretation of Spinoza on this point is notoriously difficult. We might take the many finite modes to be modifying tropes (see Sections 8.2 and 9.3.2.2), all modifying a single, necessarily existing and necessarily unique substrate.

We build to an argument drawn from Spinoza's *Ethics* and his *Short Treatise on God, Man and His Well-Being* (Spinoza 1677a/1958a and 1677b/1958b), beginning with eight Spinozistic metaphysical principles (S1–S8):

S1. Something exists.

That is, Nihilism is false.

S2. Everything is either caused by something else or caused by itself.

Spinoza assumes that everything has a cause. His notion of a *self-caused* thing is puzzling, since it seems odd to say that something caused itself to exist. (We'll discuss this further in Chapter 26.) However, Spinoza's argument could easily be reformulated in such a way as to replace all talk of self-caused things with talk of *un*caused things.

S3. Any self-caused thing must have existence as its very nature.

Spinoza assumes that there are no unexplained, brute facts. He accepts the *Principle of Sufficient Reason*, that is, the claim that there is a sufficient reason or explanation for every truth. If something exists and is not caused to exist by something else, then it must contain within itself a sufficient explanation of its own existence. Spinoza supposes that this could happen only if it is of the very nature or essence of the thing to exist.

S4. Anything whose nature is to exist is infinite.

This is another controversial claim, but one that was widely held throughout the late ancient, medieval, and early modern periods. If a thing exists by its own nature, then it is a thing of pure or absolute existence, containing within itself all of the power and reality of all possible beings. Such a being could legitimately be described as unlimited or infinite.

S5. There cannot exist two infinite things with the same nature.

Spinoza supposes that whenever two things exist with the same nature (for example, two spatially extended things), then each is limited by the presence or existence of the other. Neither can exhaust the category that they share, since if one did, it would leave no "room" for the other.

S6. If one thing causes another, then the two things must have the same (fundamental) nature.

This is one of the most important principles of Spinoza's philosophy. When one thing causes another, it can give to the other only what it already contains in itself. According to Spinoza, something cannot cause a thing with a nature *N* without having the nature *N* within itself, in some fashion. Most philosophers in recent times have challenged the truth of S6 because there seem to be many exceptions to it. For example, there are cases of transmutation, like an atom of uranium becoming an atom of uranium plus some decay products. In some cases, one fundamental particle can decay into particles of new kinds. Energy of one kind can be converted into energy of another kind, like kinetic energy converting into heat or vice versa. According to Darwinian evolution, a population of one species can over time give rise to a population of a different species. Spinoza might respond that in all of these cases, there is some more fundamental kind of nature of the cause that is preserved in the effect. Natures involving a certain quantity of mass-energy and charge, for example, might be thought to involve some more fundamental common nature.

S7. A finite thing and an infinite thing cannot have the same (fundamental) nature.

The idea behind S7 seems to be that finite and infinite things are too different to share the same nature. In the *Ethics*, Spinoza makes the even stronger claim that no two things can have the same nature. He argues that two things cannot be really distinct from each other unless they are different in their most essential attributes, a mere difference in modes or inessential characteristics is not enough. Spinoza thought that one and the same thing could exist with many, contradictory qualities at once, by having those incompatible qualities at different times and places. He believed in the possibility of multiple, simultaneous locations of the same thing.

S8. If something is caused, then it is (ultimately) caused by a self-caused thing.

Like many philosophers, Spinoza rejected the possibility of an infinite causal regress. Any chain of causation must originate with an absolutely uncaused thing.

Here is a proof that exactly one thing exists, employing S1–S8:

1 Any self-caused thing is infinite. (S3, S4)
2 Any two self-caused things would have the same nature (S3).
3 There cannot be two self-caused things. There is at most one self-caused thing. (1, 2, S5)
4 If an infinite thing causes another infinite thing, then they have the same nature (S6).
5 No infinite thing can cause an infinite thing. (4, S5)
6 No finite thing can cause an infinite thing. (S5, S6)
7 Only finite things can be caused by another. (5, 6)
8 Everything that is caused is caused (ultimately) by a self-caused thing. (S8)
9 If one thing causes another, the second thing is a finite thing caused by an infinite thing. (7, 8, 1)
10 An infinite thing cannot cause a finite thing. (S6, S7)
11 Nothing is caused by something else. (9, 10)
12 Everything is self-caused. (S2, 11)
13 There is at least one thing. (S1)
14 There is exactly one thing. (3, 13)

The argument up to step 9 is one that could have been endorsed by many ancient, medieval, and early modern philosophers who engaged in what is known as natural theology. This would include Plato, Aristotle, Plotinus, Anselm, Thomas Aquinas, Duns Scotus, Leibniz, or Samuel Clarke. All would agree that we can establish the existence of a First Cause of the world that is infinite and unique. (We will consider such First Cause arguments in more detail in Section 26.2.) However, at step 10 Spinoza moves in a surprising and new direction: he attempts to show that this First Cause ("God") is incapable of causing the existence of anything else, since It could only cause things having the same nature (S6), but nothing else with the same infinite nature could possibly exist (S5).

Philosophers in the natural theology tradition would challenge step 10, which resulted from the combination of S6 and S7. Why can't an infinite thing cause a finite thing? Thomas Aquinas would have rejected S6, claiming that an infinite cause can have the nature of its finite effect "eminently". (Thomas Aquinas, *Summa Theologiae* I, Q4, A2) Concerning S7, we could ask. Why can't finite and infinite things have the same *nature*, in the relevant sense? John Duns Scotus, for example, would have rejected S7, arguing that God (an infinite thing) shares many natural features with many finite, created things (such as existence, goodness, life, knowledge, and consciousness) (*Ordinatio* 1.3.1, 1–2; see Cross 1999: 38–41).

Spinoza has a number of additional arguments for step 10. One such argument is based on the supposition that God, the infinite being, would not create a mere finite being, since the product would fall so far short of God's own excellence (see pages 51–52 of the *Short Treatise*, Spinoza 1677b/1958b).

### 11.2.3 Bradley's Monism

F.H. Bradley was a British Idealist, who was influenced by the work of G.W.F. Hegel. Bradley argued that only one thing existed: the Absolute. He had a number of arguments for this conclusion but the most influential one involved what has become known as *Bradley's Regress*.

Bradley connects two problems. First, there is the problem of grounding or explaining relational truths. Second, there is the problem of grounding or explaining the distinctness of distinct things. Bradley seems to suggest that these are one and the same problem. For two things to be distinct, they must be related somehow to each other. But, Bradley argues, there are hidden paradoxes lurking in the idea that one thing could be related to another thing.

Prima facie, however, these are separate problems. Couldn't there be relational truths that don't involve two or more distinct entities? Couldn't there be such a thing as *self*-relatedness? Couldn't there be two distinct but unrelated things? On the other hand, there are a couple of things to be said in favor of Bradley's combination of these two problems. First, in general, there couldn't be relatedness without some plurality in the world. It wouldn't make sense to have relations if it were impossible for a relation to hold between two things. If there were no plurality or multiplicity in the world, there would be no way to differentiate between relations and monadic qualities. Second, there couldn't be distinctness without relatedness. After all, *being non-identical* is itself a relation. It would be a contradiction to say that two things exist and yet are *absolutely* unrelated to each other.

In any case, even if Bradley is wrong to combine the two issues, his proof that there cannot be any relational truths is interesting in its own right.

Here's a reconstruction of Bradley's main argument (Bradley 1897/1930: 21–29):

1 If two distinct things exist, then their distinctness must be explained by (by being grounded in) their standing in some distinguishing relation to one another (at the very least, the relation of *distinctness*).
2 More generally: if $n$ distinct things (where $n \geq 2$) exist, then their distinctness must be grounded in their standing in some distinguishing relation or relations to each other (either by pairs or collectively).
3 If the distinctness of $n$ things is grounded in their standing in some distinguishing $n$-ary relation, then (i) that relation must exist, (ii) the relation must be distinct from each of the $n$ things, and (iii) the distinctness of the $n$ things must be partly grounded in the distinctness of that distinguishing $n$-ary relation from each of the $n$ things.

4 The partial grounding relation between distinctness facts is transitive and asymmetric—so no cycles of grounding.
5 If two distinct things exist, then there is an infinite regress of explanations of their distinctness (from 1–4).
6 Such an infinite regress cannot exist.
7 Therefore, there cannot exist two or more distinct things, and no two things are related.

Bradley's argument for clause (i) of premise 3 relies on the fact that a relation cannot be something like a quality that is predicated of one or both of the relata. Therefore, the distinguishing relation must exist *in its own right*. It cannot be something merely predicated of the relata.

Moreover, this relation must be distinct from its relata (clause (ii) of premise 3), for it is impossible for something to be both related to something else and to be the relation by which it is related to that other thing. A relation (if it exists at all) must always be distinct from its relata.

Moreover, if the two things could not ground their own distinctness but require a distinguishing relation to do so, that distinguishing relation had better be distinct from each of the two things. Thus, the distinctness of the relation from its relata is presupposed by the relation's doing its job of distinguishing them. Consequently, the distinctness of the relata depends upon the distinctness of the relation from the relata (clause (iii) of premise 3).

Here's how the regress works. Suppose that $a$ and $b$ are distinct. Then there must be a distinguishing relation $R_1$ that relates $a$ and $b$ and that grounds their distinctness. To do this job, $R_1$ must exist and be distinct from both $a$ and $b$. Thus, the distinctness of $a$ and $b$ is partly grounded in the distinctness of $R_1$ and $a$, and in the distinctness of $R_1$ and $b$. Now, since $R_1$ and $a$ are distinct, their distinctness must be grounded in their standing in some distinguishing relation $R_2$. But $R_2$ must be distinct from both $R_1$ and $a$, and so the distinctness of $R_1$ and $a$ is dependent on the distinctness of $R_2$ and $a$. In the same way we can generate further links in the dependency regress: the distinctness of $R_2$ and $a$ grounded in the distinctness of some $R_3$ and $a$, and so on.

The key step in the argument is premise 3, especially clause (i), which introduces a new distinguishing relation that holds between the first distinguishing relation and its relata. Some philosophers, following Plato's lead, have introduced the *instantiation*, *participation* or *exemplification* relation. (We called such philosophers 'Realists' in Chapter 7.[1]) This is the relation that is supposed to hold between a relation and two or more things whenever those things stand together in that relation. *Exemplification* might do the work of distinguishing a relation from its relata. But for the argument to work, we have to suppose both that such a relation exists, and that its holding between $R$, $a$, and $b$ is what explains the fact that $R$, $a$, and $b$ are mutually distinct. To take a concrete example, suppose that hill $A$ is due north of hill $B$. The two hills stand in the *one-north-of-the-other* relation, or the *north-of* relation, for short. Suppose that the *north-of* relation is itself a thing, no less than hills $A$ and $B$ are. If so, it does seem that the three things, the *north-of* relation, $A$, and $B$, all stand together in some relation: the ternary (3-place) relation of *the first thing's being a relation that the second thing stands in to the third thing.*

However, if this is so, and the holding of the *exemplification* relation is thought to be prior in terms of explanation to the holding of the *north-of* relation, then we seem to fall into an infinite regress. The holding of the *exemplification* relation by the *north-of* relation, *A*, and *B* would itself have to be explained by a prior holding of the four-place *exemplification* relation between the *exemplification* relation, the *north-of* relation, *A*, and *B*. And so on to infinity.

We have already encountered Bradley's Regress in Section 7.2.1.3, in our discussion of universals. As we saw, Ostrich Nominalists object to the introduction of relations as things. They suppose that *A*'s being north of *B* is *not* to be understood in terms of *A*'s and *B*'s standing with the *north-of* relation in the *exemplification* relation.

Realists (7.1T), who do believe in the existence of universal like NORTH-OF, will attempt to block the regress at the second step by denying the existence of the *exemplification* relation. Such Realists employ the slogan, Relations relate! In other words, a relation doesn't need some further entity to connect it to its relata. In the context of Bradley's Regress, such Realists would have to claim that a relational universal can be the ground for its own distinctness from its relata. In effect, such relations do relate themselves to their relata, contrary to step 3. The NORTH-OF universal could ground its own distinctness from its relata, pehaps by simply being a universal, rather than a particular. Consequently, there is no need to introduce an EXEMPLIFICATION universal, and thus no reason to posit another instance of *exemplification* linking it to the other three entities (the NORTH-OF universal and the two hills).

Bradley has supplementary arguments for premise 3. He explicitly considers an alternative, namely, that the relation is merely a property or "adjective" of the relata. We could take this in the spirit of Ostrich Nominalism. We could ask. Why not posit primitive, irreducibly relational facts, with no third entity involved?

In response, Bradley proposes a three-part dilemma (or *trilemma*). Either (i) this property is predicated of one relatum but not the other or (ii) it is predicated of each (distributively) or (iii) it is predicated of both taken together (collectively). (Presumably, we can ignore the other logical possibilities: that it is predicated of neither or that it is predicated of the relata plus some additional things.) Bradley argues that none of (i)–(iii) are workable:

> The relation is not the adjective of one term, for, if so, it does not relate. Nor for the same reason is it that adjective of each term taken apart, for then again there is no relation between them. Nor is the relation their common property, for then what keeps them apart? They are now not two terms at all, because not separate. (Bradley 1897/1930: 27n1)

There are two parts to Bradley's argument here, H1 and H2:

> H1. If the relation is predicated of just one of the relata, then this fact cannot ground the relatedness of the two. If it cannot ground their relatedness, then neither can it ground their distinctness (since distinctness is a relation).
> H2 (H2a and H2b). If the relation is predicated of both of the relata, either distributively (H2a) or collectively (H2b), then it marks no difference between the two, and so it cannot ground the distinctness of the two relata.

Let's consider each claim in turn.

*Reply to H1*. Why can't the ground of distinctness consist in the fact that some property is predicated of one of the two things and not of the other? For example, suppose that *a* is black and *b* is not. To avoid a regress, let's suppose that these facts are primitive and fundamental and so not to be analyzed in terms of a relation to some universal.

We might worry here about what the truthmaker for *b*'s *not* being black could be. Could it simply be *b*? Couldn't *b* be essentially and intrinsically non-black? Perhaps Bradley considered it to be impossible for such a fact to be fundamental. Perhaps it must be grounded in some further fact: either that *b* has a property that is *incompatible* with being black or in the fact that *b lies outside* the Totality Fact for BLACKNESS (thought of now as a shareable universal). Either of these would depend on some further relation: a relation of *incompatibility* between the two incompatible properties or a relation of *lying-outside* between *B* and the extension of the property it doesn't have. Even if *being non-black* were fundamental, one might think that it could ground *b*'s distinctness from *a* only if *being black* and *being non-black* stand in some kind of relation of logical incompatibility. Bradley could argue that, once again, the problem has only been shifted back a step: how are we to explain *those* relational facts? Another infinite regress threatens.

*Reply to H2a*. Bradley claims that distinctness cannot be grounded in a common property, a property possessed by each of the two relata. This seems reasonable. How could the fact that two things are *alike* in a certain way possibly ground their *distinctness* (their *non-identity*)? Or, to put the point another way, two monadic facts cannot add up to a relational fact. If Bob is in New York and Jones is in New York, this constitutes a relation between Bob and Jones *only on the assumption that Bob and Jones are distinct*, which is just what we are trying to ground.

*Reply to H2b*. Bradley objects that a relation that is predicated of two things collectively cannot account for the separation of *a* and *b*. Bradley's argument seems to be most clearly targeted at Fundamental Relations of Distinctness (9.2A.1). Bradley echoes our objection based on Relata more Fundamental than Relations (the claim that the existence and distinctness of the relata of any relation cannot be grounded in or made true by the holding of the relation itself; see the discussion of 9.2A.1). Two things cannot have the common property of *belonging together to a pair of distinct things* without being somehow independently distinct from each other. The ground of the distinctness of *a* and *b* cannot be the fact that they both belong as parts to the pair set $\{a, b\}$, since the existence of the pair *as a pair* depends on their independent distinctness.

What if *a* and *b* are two parts of a *structured* whole, like two amphibians in a structural universal (see Section 10.3)? Suppose, for example, that *a* is the pitcher position and *b* the catcher position in a baseball team *c*. In this case, we might suppose that the existence of *a* and of *b* is each partly grounded in the existence of the whole *c*, namely, that *c* is *ontologically prior* to both *a* and *b* and that *c*'s essence or nature provides the grounds for the distinctness of *a* and *b*. You can't have a baseball team without a pitcher and a catcher that are distinct from each other.

Bradley might respond that this just pushes the problem back a step. We will now need an account for the distinctness of *a* from *c* and for their relatedness (and also for that of *b*'s distinctness from and relatedness to *c*). However, we can ask Bradley, in turn: why can't *c* itself be the ground for these truths? Since *a* and *b* are both parts of *c*, we don't need an account of the separateness of *a* from *c*—for the simple reason that they aren't separate at all.

Perhaps this conclusion would not be unacceptable to Bradley. It would push us toward the view that the totality of all things (as a structured whole) is the metaphysically fundamental reality, grounding the nature and distinctness of all of its constituents. This might be merely a version of a moderate or *Priority* Monism, of the kind that we will examine in more detail next (in connection with Jonathan Schaffer's work).

Before we leave Bradley, we need to consider a second regress argument that occurs in the same passage (Bradley 1897/1930: 26–7). We can call this regress *the internal regress*, in contrast to the external regress that we have just considered.

Suppose *a* and *b* are related to each other. There must be two *moments* or aspects of *a*—one of which is *a qua related to b*, and the second *a qua existing on its own* (apart from its relation to *b*). Bradley claims that there cannot be a genuine case of relatedness without both of these aspects. These two aspects of *a* can't be simply identical to each other, since they involve contrary characteristics of *a* (as related to *b*, as not related to *b*). Consequently, for *a* to be related to *b*, there must be some unity (and hence some relatedness) between two distinct things: *a*-as-related-to-*b*, and *a*-as-not-related-to-*b*. This launches a second infinite regress.

In response, we could insist that these two aspects are not *really* distinct. They are distinct only in our thought or intention. Both aspects are really just *a*, after all. The aspects are both *abstracted* in thought from what is one and undivided in reality (to use Aristotle's term). Bradley's retort might be that this would entail that the act of abstraction generates a falsehood or fiction, which would still preclude the real existence of a relation.

Finally, let's consider the import of Bradley's Regress for *Constituent Realism*, UP-Realism combined with Constituent Ontology (9.1A). Bradley's Regress could be developed against combining Constituent Realism with any sort of self-individuating pairs of things, including bare particulars, haecceities, or even primitively distinct universals (see Section 9.3.2). On a pure version of Constituent Ontology, the one and only primitive relation is that of *being-part-of*. Primitive relational facts involve at least three things: *a* and *b* as distinct proper parts of a whole *c*. We could express this fact as: *a and b jointly compose c*.

There is no temptation to introduce a fourth entity. *c* is the truthmaker for this relational fact. It is by virtue of the existence of *c* that *a* and *b* jointly compose *c*. This may help with both regresses. The relationality of *a* and *b* to *c* (and to each other) is contained entirely in the being of *c*.

Does Constituent Realism work both when *c* is ontologically prior to *a* and *b* and when *a* and *b* are ontologically prior to *c*? (Let's assume that we can understand ontological priority in terms of grounding. If *x* is *ontologically prior* to *y*, then the existence and intrinsic character of *y* is wholly grounded in some facts that include the existence and intrinsic character of *x*.) In the first case, when the whole is prior to the parts, all three entities exist only *as-related* to the others. In the second case, when the parts are prior to the whole, *a* and *b* have independent existence but no internal relatedness to *c*, and *c* has internal relatedness to *a* and *b* but no existence independent of the relation, so in both cases the internal regress seems to be avoided.

What about the original external regress? In the case where *c* is ontologically prior to *a* and *b*, this difference in fundamentality by itself gives rise to a real difference between *a* and *c* (and between *b* and *c*). Given those real differences, *c* can play the role of relating and grounding the distinctness of *a* and *b*. But, what about the case in which *a* and *b* are ontologically prior to *c*? Now we do seem to have a problem, because now *c*'s existence seems to depend on the *prior* distinctness of *a* and *b*.

This could be turned into an argument that at least some wholes are ontologically prior to their parts, a holistic conclusion that would also be attractive to Bradley and other moderate Monists. This argument for the possibility of holism depends on just two assumptions: Constituent Ontology, and the thesis that there can be no distinctness without relatedness. Given Constituent Ontology, there are just two ways for two things to be distinct (on the assumption that distinctness requires relatedness). First, the two might differ intrinsically, in which case they contain distinct parts (since all the intrinsic properties of a thing are parts of it), and second, the two might inherit their distinctness from the different roles they play in some larger whole or structure. The first way cannot provide the ultimate ground for distinctness without an infinite regress. Thus, we must suppose there to be some cases of distinct things *a* and *b* that are distinct not because they contain different proper parts but because of the disparate roles they play in some larger structure. We thus arrive at the following argument:

A BRADLEYAN REFUTATION OF UNIVERSAL BOTTOM-UP GROUNDING (FOR CONSTITUENT ONTOLOGISTS)

1 Assume that there are some distinct things *a* and *b*.
2 The distinctness of *a* and *b* must be grounded in some relation (Bradley's assumption).
3 Given Constituent Ontology and premise 2, the distinctness of *a* and *b* must be grounded either in the fact that they have different parts, or in the fact that they play different roles in some larger whole, *c*.
4 It is impossible for the distinctness of all things to be grounded in some difference in their parts, since this would lead to an infinite regress. So, we can suppose that the distinctness of *a* and *b* is *not* grounded in any difference in their parts.
5 Consequently, the distinctness of *a* and *b* is grounded in the fact that they play different roles in some larger whole, *c*.
6 Assume for contradiction that all wholes are grounded in their parts.
7 So, *c* is *not* ontologically prior to *a* or to *b* (from 6).
8 However, the fact that *a* and *b* play different roles in *c* cannot even partially ground their distinctness, given 7, since the fact that *a* and *b* can play different roles in *c* presupposes their prior distinctness. Contradiction.
9 Consequently, the hypothesis in 6 must be false: if there are distinct things, their distinctness must be grounded in the different roles they play in some whole that is *not* grounded in its parts.
10 So, if there are distinct things, there must be some whole that is not grounded in its parts.

However, this conclusion does not prove that *all* wholes are prior to their parts, much less that the whole universe is prior to everything else. In fact, the argument seems to point to the conclusion that it is *universals* that derive their distinctness from ontologically prior structures in which those universals play different roles. Once we have established distinctness relations between universals, we can use those distinct universals to differentiate particulars. There is no refutation here of the thesis that all wholes composed entirely of particulars are grounded in their parts. Perhaps a Bradleyan Monist could turn to the theory of Amphibian Bundle Theory, which we discussed in Section 10.3, with a single, cosmic structural universal as the fundamental entity, and all ordinary particulars as amphibians within that structure. We will turn next to some more recent arguments for a similar conclusion in the next section.

### 11.2.4 Jonathan Schaffer's Cosmic Monism

Jonathan Schaffer (2010) has defended a moderate form of Monism, one involving the claim that there is only one *fundamental* thing, namely, the cosmos or the universe. Like Spinoza, Schaffer does not intend to contradict our ordinary beliefs about such things as "ships and shoes and sealing-wax, cabbages, and kings" (to quote Lewis Carroll's *Alice in Wonderland*). What we say and believe about these ordinary things are true, but they are ultimately made true by one thing: the whole universe. Schaffer's Monism should be thought of as an example of what we called ontological grounding in Section 3.4: parts of the universe exist and have essences that ensure that all of their properties and relations can be ultimately explained in terms of monadic properties of the whole universe.

Schaffer relies in part on modern physics in support of Monism, especially the phenomenon of *quantum entanglement* as described by quantum mechanics. When several particles become entangled together, they lose their individual identities, merging into a single, inseparable quantum system. On some, but not all, interpretations of quantum mechanics, the whole universe constitutes a single, entangled system, from the time of the Big Bang and ever thereafter.

However, this is true only on the so-called "no-collapse" interpretations of quantum mechanics. On the more standard Copenhagen interpretation, quantum systems *collapse* when interacting with observers. The collapse breaks the entanglement, giving rise to wholly separate, disentangled particles. Whether to accept the Copenhagen interpretation or the no-collapse interpretation of quantum mechanics is itself a decision involving metaphysics. To the extent that common sense and ordinary experience tell us that the world consists of a number of things, they give us reason to reject the no-collapse interpretation.

Importantly, Schaffer doesn't deny that the universe has parts, in a sense. He simply denies that the parts of the universe (like you and us) are fundamentally real. Instead, our existence simply consists in the universe's being a certain way. For example, your existence consists in the universe's being human-ish at some place and time.

Schaffer offers two principal arguments for Monism. First, he appeals to the plausible principle that all fundamental entities have precise, determinate, and non-arbitrary or objective boundaries. If something is fundamentally real, then there should be a definite and objective fact whether or not it is located at any spatial point and whether or not it includes as a part any given material body.

**Precise Boundaries of Fundamental Things.** If something is a fundamental entity, then both its spatial boundaries and its material composition are precise, determinate, and objective.

Next, Schaffer contends that all spatial and material entities other than the whole universe have vague boundaries, boundaries that are indeterminate or arbitrary. This certainly seems true of ordinary objects, like ships or shoes or cabbages. There are particles near or on the surface of the thing that could be counted as part of the object or not, with equal reasonableness. Thus, Schaffer's argument goes as follows:

1 If something is a fundamental entity, then its boundaries and material composition are not vague.
2 The whole universe is the only material thing without vague boundaries and composition.
3 Therefore, there is only one fundamental material thing, namely, the universe.

This argument cannot establish Monism all by itself. It would have to be supplemented by a thesis of Materialism, the claim that all fundamental entities are material. If, for example, there are immaterial souls or angels, they could not be fundamental entities.

In addition, both premises of the argument could be challenged. Perhaps some fundamental things are ontically vague (a possibility we will explore in Section 12.2.2). Moreover, it isn't obvious that every material thing is vague. Some vagueness might be merely a matter of the limits of our knowledge (the thesis of Epistemicism—discussed in Section 12.2). If there are point particles or if there are precisely defined regions of space occupied by some unique glob of stuff, both these particles and these precisely defined globs would have precise and determinate boundaries.

Schaffer's second argument begins with the proposal that we must choose between Monism and Atomism. That is, either there is only one fundamental thing, and everything else is merely a kind of virtual part of that One Big Thing, or else the only fundamental things are the metaphysical atoms, simple entities with no parts at all.

**11.2T.1 Fundamental Atomism.** Necessarily, every fundamental entity is atomic (lacking parts).

It would be redundant, Schaffer argues, to suppose that both wholes and their parts are equally fundamental. We should either choose the parts and treat the whole as nothing over and above its parts arranged in a certain way or we should choose the whole and treat the parts as merely virtual, as different aspects of the whole associated with different regions of space. Once we've made our choice, we should stick to it consistently, leading us ultimately to believe either in the One Big Thing or in the many small, indivisible things. That is, we should go for Monism or Fundamental Atomism.

This argument of Schaffer's turns on the possibility of *atomless gunk*, a notion that we explore in more detail in Section 22.4.2. Atomless gunk is a material body that is *gunky*.

**Def D11.1 Gunky Body.** A material body $x$ is *gunky* if and only if $x$ has parts but no ultimate parts, that is, for every part $y$ of $x$, there is a distinct, still smaller part $z$ of $y$.

Atomless gunk has parts but no ultimate parts. Each part of a bit of atomless gunk has parts of its own, and those parts have further parts, and so on, ad infinitum. Schaffer takes it as plausible that atomless gunk is a real possibility, that it is at least possible that some material bodies are gunky. In fact, there are live scientific theories that posit the gunkiness of matter. Schaffer therefore embraces the Possibility of Ubiquitous Gunk.

**Principle of Natural Philosophy (PNatPhil) 1 Possibility of Ubiquitous (Atomless) Gunk.** It is possible for nothing to exist but atomless gunk (things with parts but no atomic parts).

Schaffer also assumes that it is metaphysically necessary that something fundamental exists, a version of what we have called Aliquidism. We are now ready to give Schaffer's main argument for Monism:

1 If Monism is false, then Fundamental Atomism is true (i.e., every fundamental entity is atomic).
2 If Fundamental Atomism is true, then it is necessarily true.
3 It is possible for nothing to exist but atomless gunk. (Possibility of Ubiquitous Gunk)
4 Necessarily, if nothing exists but atomless gunk, then there are no atomic things. (From the definition of atomless gunk)
5 So, it is necessarily the case that if nothing exists but atomless gunk and Fundamental Atomism is true, there are no fundamental entities. (From 4 and the definition of Fundamental Atomism)
6 But, it is necessarily true that there are fundamental entities (since Aliquidism is necessarily true).
7 So, it is necessarily the case that if nothing exists but atomless gunk, then Fundamental Atomism is false. (From 5 and 6)
8 So, Fundamental Atomism is possibly false, since it is possible for nothing to exist but atomless gunk. (From 3 and 7)
9 So, Fundamental Atomism is in fact false (from 2 and 8).
10 So, Monism is true. (From 1 and 9)

The weakest part of Schaffer's argument is step 1: the forced choice between Monism and Fundamental Atomism. As we shall see in Chapter 22, there are other viable alternatives, alternatives that combine Pluralism with the thesis that some composite entities (like organisms and homogeneous continua) are fundamental.

In addition, Schaffer's position is vulnerable to a parallel argument against Monism. Let's say that *A* is an encompasser of *B* just in case *B* is a part of *A*. Then we can define 'junk' in a way parallel to our definition of 'gunk':

**Def D11.2 Encompass.** *x encompasses y* if and only if *y* is a proper part of *x*.

**Def D11.3 Junky Body.** A material body *x* is *junky* if and only if it has encompassers but no ultimate encompasser: that is, for every *y* that encompasses *x*, there is a distinct, still larger encompasser *z* of *y*.

If junky material bodies are possible, then it is possible for there to be no cosmos, since the cosmos is by definition something that has no further encompasser. This means that Monism is possibly false. But, if it is true, it is necessarily true. Hence, Monism is false.

Powerists have another reason for rejecting even a moderate form of Monism. Only fundamental things can bear powers, so Monism entails that there be just one power-bearing entity (namely, the world). This means that the only powers that could be exercised would be the immanent powers of the whole world. Active and passive powers cannot be exercised unless there exist at least two power-bearing entities, one to serve as the agent and one as the patient.

Consequently, if Monism were true, there would be no *interactions* at the fundamental level. We couldn't interact with the world since we are each part of the world, with no powers of our own. The world would be an extremely small world, a metaphysically lonely entity. This, in turn, would mean that Powerism would lack any special account of how we discover the powers of things through active experimentation (of the kind we described in Section 6.1.2). If Monism were true, it seems that Powerists would have to fall back on an essentially Neo-Humeist account of our knowledge of the laws of nature (as axioms in the best theory of the world), depriving Powerism of one of its most powerful motivations. This conclusion might be resisted if we could find a way for derivative entities to have derivative causal powers of the kind needed for the typical Powerist account of scientific knowledge. Perhaps we could think of the Cosmos as a kind of gigantic organism, whose state confers real but derivative powers on its integral parts.

## Note

1 A similar regress could arise for Trope Nominalists, if they imagine that relations are tropes (presumably modifying tropes). Now, instead of *exemplification* as a tie between universals and particulars, we would need a *modification* relation between tropes and relata.

# 12

# The Non-Existent and the Vaguely Existing

In this chapter, we take up two clusters of questions concerning existence. The first cluster (Section 12.1) concerns the scope of existence. How wide is the domain of existing things? Does it encompass absolutely everything, or can we speak meaningfully about non-existent things? The second cluster (Section 12.2) concerns vagueness and indeterminacy. Are there vague things and vague categories of things or is all vagueness a matter of referring indifferently to a large number of absolutely precise things? What is the ultimate source of vagueness? Is it ignorance, ambiguous language and concepts, or real indefiniteness in the things themselves?

## 12.1 Does Everything Exist?

In the last chapter, we considered whether anything exists, and, if so, how many things. Now we will turn to the converse question: does everything exist? Willard Van Orman Quine took it to be obvious that everything exists. When asked what there is, Quine replied that the obviously correct answer is simple: everything. This is Actualism:

**12.1T Actualism.** Everything (actually) exists.

Alexius Meinong was an early defender of Anti-Actualism (compare Chisholm 1982, Grossmann 1974, Paśniczek 1995, Perszyk 1993, Priest 2005, Routley 1979 and 1982). One of his statements of this position has been translated into the highly paradoxical sentence, "There are some things such that there are no such things." The suggestion of paradox is considerably lessened in the original German, where the phrase translated 'there are' is 'es gibt' (literally, 'it gives'). A better translation would be something like,

*The Atlas of Reality: A Comprehensive Guide to Metaphysics*, First Edition.
Robert C. Koons and Timothy H. Pickavance.

"Some things are not." This isn't self-contradictory or evidently absurd, in the way that "Some existing things do not exist" is. We thus formulate Anti-Actualism as follows:

**12.1A Anti-Actualism.** Some things don't (actually) exist.

If some things don't exist, must it at least be the case that everything *could* exist? Meinong and his followers deny even this, while Possibilists affirm that everything could exist.

**12.1A 1T Possibilism.** Some things don't exist, but everything could exist.

**12.1A.1A Meinongianism.** Some things couldn't exist.

Importantly, the debate between Actualists and Anti-Actualists does not turn on the supposition that there are different *modes of being* like existence and something else (like *mere subsistence*). Meinong himself made a distinction between existence and mere subsistence. Concrete things, like physical objects or people, exist, while abstract things, like numbers, states of affairs, or similarities merely subsist. However, we are interested in Meinong's claim that some things do not even subsist. This has nothing to do with the existence/subsistence distinction. To simplify things, we will use 'to exist' and 'to be' interchangeably, in a way that is equivalent to Meinong's broader term, 'to subsist'. Thus, we are concerned with the question of whether some things don't exist.

There are, at first glance, many examples of things that don't exist. Unicorns, golden mountains, and round squares don't exist. Mythical and fictional characters, like Santa Claus and Sherlock Holmes, don't exist. We can even say things like, "The innocence of Manson doesn't exist." In addition, many objects of thought, perception, and intention don't exist. Ponce de León searched for something that doesn't exist, namely, the Fountain of Youth. In hallucinations, one can "see" (in some sense) non-existent things, and one can believe in non-existent things, as physicists once believed in the ether or as some astronomers believed in Vulcan, a supposed planet inside the orbit of Mercury.

There are, therefore, plausible truths that seem to involve or presuppose the existence of non-existent objects. Unicorns don't exist; Conan Doyle wrote about Sherlock Holmes; Ponce de León searched for the Fountain of Youth; and Benjamin believes in his imaginary friend. Actualists must argue that, contrary to appearances, these truths do not entail that some things don't exist. The appearance of reference to non-existent things, either individually or en masse, must be dispelled through an appropriate *paraphrase*: a proposition that lacks any such apparent reference and that is plausibly equivalent in content to the original proposition.

### 12.1.1 Meinong's Characterization Principle and Russell's theory of descriptions

What does it mean to be *something*? It is natural to countenance certain inferences that involve the English word 'something'. For example, you might infer that something is four-legged from the claim that most dogs are four-legged. Similarly, you might infer that

Tom has something on his shoulder from the claim that Tom has a bird on his shoulder. But you wouldn't infer that Tom has something on his shoulder from the claim that Tom has a chip on his shoulder. We can then ask the general question: when are these sorts of *something*-inferences valid? That is, for what elements of a language *T* is the following inference logically valid?

(1) *T M*'s.
(2) Therefore, something *M*'s.

Many logicians and philosophers have suggested that this inference from (1) to (2) is valid whenever *T* is a referring expression or term. In general, proper names seem to be referring expressions. Thus, the following inference is valid.

(3) Houston grows rapidly.
(4) Something grows rapidly.

There are other expressions, however, that mimic the grammatical function of referring expressions without actually referring to any particular thing. Logicians call these expressions *quantifiers*. Among the quantifiers are 'nothing', 'no man', 'something', 'some city', and 'everything'. The fact that the expression 'nothing' (or 'no man') is both grammatically similar to and logically different from a proper name was noticed by the Greek poet Homer in a famous passage from the *Odyssey*. Odysseus and his men are captured by the monstrous Cyclops Polyphemus. When captured, Odysseus tells Polyphemus that his name is 'No Man'. Later, when Odysseus escapes from his bonds, he attacks Polyphemus, gouging out his single eye with a sharp stick. Polyphemus cries out to his fellow Cylcopses with the words, "No Man is attacking me." Naturally, they ignore him, and Odysseus escapes. Homer's story illustrates the logical invalidity of the inference from (5) to (6):

(5) No man is attacking me.
(6) Something is attacking me.

Another illustration of this distinction comes from Lewis Carroll's *Alice in Wonderland*. Alice explains to the Mad Hatter that she sees no one on the road, and the Mad Hatter replies that she must have sharper eyes than he does, treating the expression 'no one' as if it named a peculiar kind of entity, which Alice could see and the Hatter could not.

In many natural languages, including English, there are expressions that grammarians call *definite descriptions*. These are formed in English by placing the definite article 'the' in front of a noun phrase, as in 'the queen of England', 'the highest peak in the Rocky Mountains', 'the shortest spy', and so on. These definite descriptions seem in many cases to function exactly the way proper names function. When successful, they pick out exactly one thing, to which some attribute is then predicated. Consider, for example:

(7) The shortest spy works for the NSA.
(8) Samuel H. Ortcutt works for the NSA.

If (7) is true, this is because the person picked out by the phrase 'the shortest spy' works for the NSA. Similarly, (8) is true if and only if the person picked out by the name 'Samuel H. Ortcutt' works for the NSA.

What, then, are we to do with definite descriptions that do not refer to anything? Consider such expressions as 'the present reigning king of France' or 'the round square spot on my desk'. There is in fact no present reigning king of France, and there is not and could not be a spot on my desk that is both round and square.

Meinong argued that we should take all definite expressions as referring to something. The difference between 'the present queen of England' and 'the present king of France' is that the first expression refers to something that exists, while the second refers to something that doesn't exist. Similarly, 'the round square spot on my desk' refers to something that does not and could not exist. In fact, Meinong proposed what has been called the 'Characterization Principle' as an axiom of logic:

**Characterization Principle.** The *F* is an *F* (where '*F*' can be any meaningful noun phrase).

From the Characterization Principle, we can infer, with obvious correctness, that the present queen of England is a queen of England, and that the highest peak in the Rockies is a highest peak in the Rockies. Consequently, it follows logically that the highest peak in the Rockies is a peak and is in the Rockies. So far, so good.

What about the present king of France? The Characterization Principle entails that the present king of France is a present king of France, and so is a king. According to Meinong, this poses no problem so long as we remember that the present king of France is a *non-existent* king. Consequently, we should not be surprised when we don't see him on a throne in Paris or Versailles.

However, the Characterization Principle brings serious logical problems in its wake, especially when we consider logically contradictory descriptions. Consider again the round square spot on my desk. According to the Characterization Principle, this spot is both round and square. However, the laws of geometry dictate that whatever is round cannot be square. Hence, the spot is both square and not square, a logical contradiction. In fact, we can reach a contradiction even more quickly by considering the thing that is both square and not square.

Meinong's answer was to propose that the laws of logic apply only to things that are capable of existence. Things whose existence is impossible fall out of the range of logic. Hence, from the fact that the square and not square thing is square and not square, nothing follows. This object violates the Law of Non-Contradiction, but that is permissible since it is necessarily non-existent.

However, this solution seems to lead to further difficulties. Consider the *possibly existing* round square. According to the Characterization Principle, the possibly existing round square possibly exists, and so it should be subject to the laws of geometry and logic. However, it clearly is not.

In response to these difficulties, some Meinongians proposed a distinction between *nuclear* and *extra-nuclear* properties (see Parsons 1980 and Zalta 1983, 1988). Nuclear properties are ordinary properties, like *being round* or *being a king*. Extra-nuclear properties include special properties like *existing* or *being possible*. Meinongians restrict the

Characterization Principle to expressions that attribute nuclear properties to the designated object.

There is, however, a much simpler solution of all of these difficulties that is available to the Meinongians (as pointed out by Priest 2005): simply deny the Characterization Principle. Denying the Characterization Principle does not, by itself, entail that everything exists. In place of the Characterization Principle, we could use one or the other of these weaker principles:

**Modal Characterization Principle.** The *F*, if it existed, would be an *F* (or if it had existed, would have been an *F*).

**Representation Characterization Principle.** The *F* is normally represented as an *F*.

These principles do not generate contradictions, even if they are applied to logically contradictory descriptions. It is true that the round square, if it were to exist, would be round and square. Since it is impossible for it to exist, such a conditional is either trivially true or it could be taken to refer to some logically impossible situation. We are not forced to say that the round square *is* both square and not square, but only that it *would be* square and not square, which is not in violation of the Law of Non-Contradiction. Similarly, there is no contradiction in supposing that the round square is *represented as* both square and not square.

The round square (that RCK is thinking of) does not exist, but it is neither round nor square. It is *thought of by me* as being round and square, but that does not make it really round or square. Ponce de León's Fountain of Youth was not a fountain, nor did it have any magical, youth-restoring powers. It was, however, *believed by* de León to be a fountain with such powers. Similarly, Sherlock Holmes was not a detective. He is a fictional detective, but fictional detectives are no more detectives than imaginary mountains are mountains. A list of the world's detectives would include Eliot Ness but not Sherlock Holmes. Conversely, a list of the world's fictional detectives would include Holmes but not Ness. Similarly, golden mountains are neither golden nor mountainous; they are things that, were they to exist, would be golden and mountainous, and that are represented as golden and mountainous. Reformed Meinongians (like Graham Priest) hold that non-existent things have only intentional properties (like *being thought to be so-and-so*) and modal properties (like *being possible* or *impossible*).

RUSSELL'S THEORY OF DESCRIPTIONS Bertrand Russell proposed in his paper "On Denoting" (Russell 1905) a radically different approach to the understanding of definite descriptions. According to Russell, definite descriptions are *quantifiers* (like 'something' or 'nothing') and not referring expressions at all. Their semantic function is *not* to pick out a single entity. According to Russell's theory, the use of a definite description, such as 'the round square' or 'the man who lives on the North Pole and delivers gifts every Christmas', need not commit the user to the existence of anything. Everything depends on the context. For example, we can interpret the simple sentence:

'The *F* is a *G*.'

as expressing the following:

> 'There is one and only one *F*, and every *F* is *G*.'

This paraphrase is important, since one can then deny that the *F* is a *G* without being committed to the existence of a something called 'the *F*'. If one *negates* the sentence 'The *F* is a *G*', either one of two sentences may result, depending on the *scope* of the negation:

> (9) There is one and only one *F*, and not every *F* is *G*. (Narrow scope negation)
> (10) It is not the case that: there is one and only one *F*, and every *F* is *G*. (Wide scope negation)

(10) is logically equivalent to: there are no *F*'s or there are more than one *F* or some *F* is not *G*. Consequently, (10) is true whenever there is no unique *F*, while (9) would be false in that case.

This means that we can say that the round square does not exist without attributing a property (that of *non-existence*) to a thing (the round square). Instead, Russell argues that we can interpret that statement as saying, simply, that the expression 'the round square' does not signify anything at all. Russell thereby provides an account of definite descriptions that is consistent with Actualism.

Russellians have applied the same analysis to empty proper names like 'Santa Claus' and 'Zeus'. It is possible to interpret such names as *disguised* or *abbreviated* definite descriptions. Thus, we could interpret 'Santa Claus' as equivalent to 'the man that lives at the North Pole and delivers presents on Christmas Eve'. Under this interpretation, the statement that Santa Claus does not exist is equivalent to the statement that there is no unique man living at the North Pole who delivers presents on Christmas Eve. No non-existent objects are needed.

### 12.1.2 Quantifiers and plural expressions

This Russellian analysis of definite descriptions is quite plausible, and Possibilists and Meinongians should not rest their case on uses of definite descriptions. There are, however, plenty of other plausibly true statements that do seem to require reference to non-existent things. Consider, for example, the use of plural expressions, like 'round squares' or 'golden mountains'. We can say something like 'Golden mountains do not exist' or 'Golden mountains could exist, but round squares could not.' Russell's theory of descriptions does not apply to such statements. We do seem to be saying something like: some things (namely, golden mountains) have the property of *possible*, but not actual, *existence*.

Actualists can stick to their guns, however. They can insist that plurals ('golden mountains') and indefinite descriptions ('a golden mountain', 'some golden mountain') can also be interpreted as *quantifiers*, as expressions that simply indicate how many things there are of the described kind. To say that golden mountains do not exist is simply to say that nothing is a golden mountain, that there are absolutely no golden mountains. To say that round squares could not possibly exist is to say that it is impossible that anything could be a round square.

There are at least two kinds of quantifiers that could be used by Actualists in interpreting such plural expressions. They might take the phrase 'golden mountains' to be short for 'all golden mountains' or 'every golden mountain'; these are *universal* quantifiers. Following Gottlob Frege, modern logicians generally interpret statements like 'All *F*'s are *G*'s', which involve a universal quantifier, as equivalent to the conditional, 'for every *x*, if *x* is an *F*, then it is a *G*'. They interpret this conditional as a material conditional. It then is equivalent to, 'for every *x*, either *x* is not an *F* or *x* is a *G*'. Given this interpretation, if there are no *F*'s at all, then the universal statement 'All *F*'s are *G*'s' is true; indeed, modern logicians would say it is *vacuously* true. So, if there are no golden mountains, the statement 'All golden mountains are *G*' is vacuously true, no matter what adjective or predicative phrase we put in for *G*. 'All golden mountains are flat' is vacuously true, for example. Crucially, it is also true that all golden mountains are non-existent, not because the phrase 'golden mountains' refers to some things that have the property of non-existence, but simply because 'golden mountains' refers to nothing at all.

The other way of interpreting these bare plural expressions is as a form of *generic* quantification. For example, it seems to be true to say that mules are stubborn, even if there are a few exceptions to this rule. The sentence 'Mules are stubborn' could be taken as saying something about the property of *being a mule* and its causal or statistical connections to other properties, like the property of *stubbornness*. We could interpret 'Mules are stubborn' as saying something like the property of *being a mule* is normally or typically accompanied by the property of *being stubborn*.

Could Actualists make sense of 'Golden mountains don't exist' or 'unicorns don't exist' by means of generic quantification? Perhaps. They shouldn't interpret these sentences as saying that the properties of *being a golden mountain* or *being a unicorn* are normally or typically accompanied by the property of *non-existence*, since Actualists deny that anything has or could have the property of *non-existence*. However, they could take these sentences as saying something about the existing properties of *being a golden mountain* and of *being a unicorn*, namely, that absolutely nothing has those properties. Actualists can admit that the property of *being a unicorn* exists without admitting that anything has that property.

However, there are some sentences involving plural expressions that cannot be interpreted as quantifiers of any kind:

(11) Round squares are among my favorite things to think about.
(12) The explorers were seeking a golden mountain.

Proposition (11) is not vacuously true simply because there are no round squares, nor is (12) vacuously true simply because there are no golden mountains. I might like to think about round squares but not about cubical spheres, and the explorers could have been seeking a golden mountain without seeking a lake filled with lead.

In order to eliminate apparent reference to merely possible or impossible things in sentences like (11) and (12), Actualists can offer a different sort of paraphrase. They can propose that, at least in some cases, the use of plural expressions or indefinite descriptions refers to sets (the set of golden mountains or the set of round squares) or to properties (the property of *being a golden mountain* or *a round square*), or to concepts (the concept of a *golden mountain*) or to linguistic expressions (the description 'a golden mountain'). These other entities, whether sets, properties, concepts, or expressions, exist even if golden mountains and round squares do not. If there are no golden mountains, then the set of golden mountains is identical to the empty set (the set with no members). Similarly, if there are no golden mountains, then the property of *being a golden mountain* exists, it just doesn't have any instances.

However, Meinongians and some Possibilists can respond that these paraphrases are much less plausible than the simpler interpretation in terms of quantifiers, including Russell's theory of descriptions. If we say that golden mountains do not exist, we do not seem to be saying anything about sets or properties. We are talking about mountains of a certain kind. These are ordinary, concrete things, nothing so exotic as a set or a property. Moreover, there are many philosophers who are skeptical about the existence of sets or properties, especially sets with no members and properties without instances. If such a philosopher says that sets do not exist or that properties do not exist, Actualists must interpret them as making the obviously self-contradictory claims that the set of sets is empty or the property of *being a property* has no instances.

Consequently, Actualists (and Possibilists in cases involving impossible objects) must argue that these expressions (definite and indefinite descriptions, and plurals) are ambiguous. In some cases they are to be interpreted as mere quantifiers, and in other cases as names of sets, properties, concepts or expressions. When we speak of golden mountains or round squares as the object of thought or desire, as in (11) and (12), we must interpret the phrases as referring to something like the property or concept of *being a golden mountain* or *being a round square*. In contrast, when we say that round squares do not or cannot exist, we are simply using quantifiers, saying that there are or could be no round squares. This is a demerit for Actualism. Other things being equal, we should minimize the attribution of ambiguity to our language.

**PMeth 1.5 Fifth Corollary of Ockham's Razor.** Other things being equal, adopt the theory that attributes the fewest ambiguities to natural language.

There is no evidence, independent of the theory of Actualism and Possibilism themselves, that these expressions are ambiguous. This is a clear advantage to Meinongianism, which can treat all such expressions, whether referring to possible or impossible things, in a perfectly uniform matter. Nonetheless, although this is certainly an advantage, it is not an absolutely decisive one.

### 12.1.3 Exportation and the Barcan Formula

As we have seen, Actualists take definite descriptions to be quantifiers rather than referring expressions. They can also interpret ordinary proper names in much the same way. For example, if we consider the sentence 'Austin is a city', we could interpret 'Austin' as a quantifier. 'Austin is a city' then says that there is one and only one thing named 'Austin', and everything named 'Austin' is a city. (Importantly, other things might have names that look and sound just like this one, but those are in fact different names.) In this way, Actualists can handle empty proper names. To assert the sentence 'Santa Claus does not exist' is to say that the name 'Santa Claus' does not name anything.

If definite descriptions and proper names are quantifiers, then every use of these expressions has to be assigned a definite scope within the sentence in which it occurs. Suppose, for example, that a sentence contains an *operator*, an expression that operates on one sentence to produce another. Some common operators are 'It is necessary that…', 'It is possible that…', 'S believes that…', and 'It is not the case that…'. If we combine operators with quantifiers (including Russellian definite descriptions and names), we can produce two kinds of sentences, those in which the quantifier has wide scope over the operator, and those in which the quantifier has narrow scope within the operator. Consider, for example, (13–18):

(13) It is possible that the *F* is *G*.
(14) The *F* is possibly *G*.
(15) Something is possibly *G*.

(16) Sam believes that the *F* is *G*.
(17) The *F* is something Sam believes to be *G*.
(18) There is something that Sam believes to be *G*.

The transition from (13) to (14) and from (16) to (17) is *exportation*. Actualists deny that exportation is generally valid. They will insist that (13) can be true, although (14) is false, and that (16) can be true even if (17) is false. For example, Actualists might say that it is possible that something is a golden mountain, but that there is nothing that is possibly a golden mountain. Similarly, it may be that Sam believes that an elf lives at the North Pole, even though there is nothing about which Sam believes that it is an elf living at the North Pole. It is invalid, in general, to export a quantifier from a narrow scope position, like (13) and (16), to a wide scope position, like (14), (15), (17), and (18). The invalidity in these cases mirrors the invalidity of exporting the existential quantifier ('something') through these same operators, as in (19–22):

(19) It is possible that something be a golden mountain. (Narrow scope quantifier)
(20) There is something that is possibly a golden mountain. (Wide scope quantifier)

(21) Sam believes that something is a jolly man at the North Pole. (Narrow scope quantifier)
(22) There is something that Sam believes to be a jolly man at the North Pole. (Wide scope quantifier)

In contrast, Possibilists and Meinongians should insist that exportation of this kind is always valid. If (19) and (21) are true, then so must (20) and (22). They need not accept the exportation of all quantifiers, but they will accept the exportation of names, definite descriptions, and existential quantifiers through modal and intentional operators.

The logical transition from (19) to (20) corresponds to a formula of modal logic known as 'Barcan's formula' (after the logician Ruth Barcan Marcus). The Barcan formula states that we can always export an existential quantifier through the possibility operator. This formula is logically valid in those modal logics that assume a fixed domain of quantification in all possible worlds. That is, those logics that validate the Barcan formula are logics where the answer to the question—What exists?—doesn't vary from possible world to possible world. Such logics would be embraced by Possibilists and Meinongians but rejected by Actualists.

### 12.1.4 Arguments for Actualism

We will look briefly at five arguments for Actualism. The first appeals to Ockham's Razor (**PMeth 1**). The second raises problems for attributing properties to non-existent things. The third raises problems concerning the identity and distinctness of non-existent things. The first stems from the best interpretation of true statements in natural English. And the fifth stems from Truthmaker Theory (2.1T/2.1A.1T/2.1A.1A.1T).

1. Ockham's Razor. Possibilism and Meinongianism require more entities than Actualism, so Ockham's Razor gives us reason to prefer Actualism. But does Ockham's Razor apply to non-existent entities or does it merely enjoin us not to posit the *existence* of things needlessly? Possibilists and Meinongians do not do the latter. They posit exactly the same number of existing things as do Actualists, and they do not assign existing things to more fundamental kinds or categories than do Actualists.

However, it does seem that Possibilists, and *a fortiori* Meinongians, are doing something extravagant, ontologically speaking. They take the class of *things* (whether existing or not) to be much larger than do Actualists. Even if they resist saying paradoxical things like 'There exist some things that don't exist', they must agree that some things don't exist, and the class of possible values of the phrase 'some things' includes for Possibilists and Meinongians many more things, like golden mountains, than it does for Actualists. This should count, at least to some extent, against Possibilism, and even more so against Meinongianism, which includes impossible things).

This isn't a knockdown argument for Actualism, since Ockham's Razor only proscribes multiplying entities needlessly. Possibilists can argue that both common sense and an adequate interpretation of many ordinary assertions demand the additional entities. As we've seen, this claim is disputed by Actualists. We will return to this issue in Chapters 14 through 16, on modality.

2. Attributing properties to non-existent things. Possiblists and Meninongians claim that there are non-existent things. Presumably, these non-existent things will have certain properties. But it is difficult to say what properties they do and don't have. Did Sherlock Holmes live in London or in a fictional city called 'London'? Was Queen Victoria (the real Empress of India) acquainted with him? Was he left-handed? Did he ever live in Shropshire? Does he inherit all of the properties attributed to him in the stories? (What about *existence*? He exists in the Conan Doyle stories.) Is he an incomplete object, neither left-handed nor not left-handed? What happens if the Conan Doyle stories are logically inconsistent, if for example the stories imply both that Watson was and was not married in 1889? What then? These puzzles given us some reason to prefer Actualism, if Possibilists and Meinongians cannot supply principled answers.

Assuming that we have given up the Characterization Principle, these questions must be sorted into two kinds. First, there are questions about how a non-existent thing is conceived of, and second, there are questions about how a non-existent thing really is. If we ask whether Sherlock Holmes, as he is portrayed in the stories, is right- or left-handed, the answer is simply, Neither. We are not violating any principle of logic (like the Law of Excluded Middle, either $p$ or not-$p$, which will we discuss later in this chapter) in recognizing that stories and other mental pictures are often incomplete. In contrast, if we ask whether Sherlock Holmes is really right- or left-handed, the answer again is Neither, but for a different reason. Sherlock Holmes has in reality no hands at all—he is neither left-handed nor right-handed in the same way that a rock is neither. The only properties Sherlock Holmes really has are either modal (*possible* or *impossible*, *actual* or *not actual*) or intentional (*thought of by Conan Doyle*, *loved by many readers*, *widely imitated*). Further, inconsistent stories pose no special problems. We do not contradict ourselves by asserting that the Conan Doyle stories represent Watson as both married and not married during a certain period of time.

3. Identity and distinctness. Quine (1953/1980) argued that the merely possible entities of Possibilism behave very badly with respect to identity, distinctness, and number. He worried that Possibilists could provide no sensible answer to the question of how many possible fat men are standing in the doorway or whether the possible fat man in the doorway is identical to or distinct from the possible bald man standing there. Similar worries arise concerning fictional characters. Does the same Sherlock Holmes inhabit each of Conan Doyle's stories? Did he add properties as new stories were written? Can the very same individual appear in new stories?

Possibilists have offered answers to these questions, or have offered explanations of why we don't know the answers. The answers to many of these questions turn on details about the structure of possibility and of its relation to the actual world, as well as metaphysical doctrines concerning the grounds of identity and distinctness for individuals. Some of these questions are as hard for Actualists to answer as they are for Possibilists.

Where fictional or mythical individuals are concerned, there are cases in which it seems that we do have good grounds for making judgments about identity or distinctness. In fact, it seems pretty clear that one and the same Sherlock Holmes appears in all of Conan Doyle's stories. Peter Geach (1967) introduced a relevant thought experiment involving two peasants, Hob and Nob, who erroneously believe that a witch exists who is blighting the village's crops. We can imagine a scenario in which it is plausible that Hob and Nob are thinking about and fearing the very same non-existent witch. (We'll consider the Hob and Nob case again, when looking at arguments against Actualism in Section 12.1.5.)

4. Interpreting statements in natural language. We tend to think that claims like (23–25) are true:

(23) Everything exists.
(24) Nothing is a golden mountain.
(25) There are no golden mountains.

This provides some evidence in favor of Actualism, since Possibilists and Meinongians deny all three. Possibilists can respond that these sentences reflect a well-known linguistic phenomenon, namely, the contextual restriction of the class of objects relevant to the truth of a particular quantified sentence. For example, one might truthfully say something like 'All the students are here', meaning that all of the students who are registered for this course are here. Similarly, it would be natural to restrict the relevant class in many cases to the class of actually existing things. Consider the following amendments to (23–25):

(23') Everything that exists exists.
(24') Nothing that actually exists is a golden mountain.
(25') There are no golden mountains that actually exist.

It is, therefore, unclear how much support (23–25) provide for Actualism.

5. Truthmaker objections to Anti-Actualism. Anti-Actualism is fully consistent with the weakest form of Truthmaker Theory, Truth Supervenes on Being (2.1A.1A.1T). Statements about non-existent things won't require any existing truthmakers. Consider, for example, (26):

(26) Golden mountains don't exist.

The truth of (26) supervenes on being, since in order to make it false, we would have to add golden mountains to the class of existing things. Intentional properties also seem to supervene on the conditions of actual beings. If it is true that Sherlock Holmes is RCK's favorite character, then making this proposition false requires a change in some existing thing, namely, RCK. Modal truths about the non-existent, like the possibility of golden mountains, are plausibly considered to be necessary truths, and all necessary truths supervene on each contingent truth, given the definition of weak supervenience (Def 3.1): you cannot change a necessary truth to a falsehood without some change in being, since you cannot change a necessary truth to a falsehood in any case!

However, Atomic Truthmaker Theory (2.1T.4) poses a set of serious challenges to Anti-Actualism. Consider, for example, (27) and (28):

(27) Native Martian organisms could have existed, although they do not in fact.
(28) Sherlock Holmes is my favorite fictional detective.

If (27) attributes a modal property (*possibility*) to non-existent Martians, then its truthmaker must be somehow connected, by its very essence, to those non-existent Martians. Similarly, if (28) attributes an intentional relation between Sherlock Holmes and me, then Sherlock must be somehow implicated in its truthmaker.

It is natural for Truthmaker Theorists to assume that a truthmaker for an atomic predication *Fa* or *Rab* must contain the subject or subjects (the entities *a* and *b*) of the predication as parts. It is also natural to assume the following Principle of Actual Parts:

**Principle of Actual Parts.** If something actually exists, so do its actual parts.

If we combine Anti-Actualism with both the Principle of Actual Parts and Atomic Truthmaker Theory, then the truthmakers for (27) and (28) would have to contain an actual native Martian and an actually existing Sherlock Holmes. Since neither native Martians nor Sherlock Holmes actually exist, this is a serious problem. In contrast, Actualists can provide truthmakers for (27) and (28) that include such actual entities as the property of *being a native Martian* or the property of *being a brilliant detective named 'Sherlock Holmes'*.

Anti-Actualists have four options: (a) deny Truthmaker Theory altogether, (b) modify Truthmaker Theory in such a way that the *merely possible* existence of a truthmaker is sufficient for the truth of the corresponding proposition, (c) deny that the truthmakers for (27) or (28) need to contain the non-existent objects named by them or (d) deny or restrict the Principle of Actual Parts. Let's consider these one at a time.

### a. Deny Truthmaker Theory altogether

Denying Truthmaker Theory altogether opens the door to extreme cases of metaphysical cheating. Indeed, if modal and intentional statements like (27) and (28) require no truthmaker, why think that their truth even supervenes on being? Leaving behind the Truthmaker Principle (2.1T) would be welcome to Presentists (20.2T.4, see Section 20.4) and to Hypotheticalists (4.1T, which we discussed in Section 4.2).

### b. Modify Truthmaker Theory

Could Anti-Actualists respond by modifying Truthmaker Theory? Perhaps the attributions of possibility are made true by the *possible* existence of a truthmaker? There is a problem with this suggestion: it treats *possible existence* as a species of existence. That is controversial. It seems that *possible existence* is to *existence* as *fake diamond* is to *diamond*. To be merely possibly existing is not to exist at all, but to be something that could exist.

This option will help only with modal properties of non-existent entities, as in (27). It won't help with intentional relations to the non-existent, since non-existent objects need not have even possible existence, like RCK's favorite round square.

### c. Truthmakers that are disjoint from the objects referred to

Anti-Actualists could deny that the truthmakers for (27) and (28) contain the non-existent objects they refer to. But then what connects these truthmakers to the proposition's subject or subjects? This would require brute necessary connections between separate or disjoint existences (that is, connections between the truthmakers and the relevant non-existent objects). These connections would not be grounded in any internal relation between the connected things. This is contrary to the Third Corollary of Ockham's Razor (PMeth 1.3). Therefore, Anti-Actualists should either give up on Truthmaker Theory entirely or bite the bullet by admitting actual things with actual but non-existent parts.

### d. Denying the Actual Parts Principle

In the latter case, the truthmakers for (27) and (28) have non-existent objects as parts, even when the whole truthmaker exists. These would be weird entities, existing things with non-existent parts. However, Anti-Actualists can mitigate this drawback in two ways.

First, they could distinguish between more or less fundamental truthmakers, with less fundamental ones grounded in more fundamental ones (this would have to be *extra-conceptual grounding*, to use the term introduced in Section 3.4). Truthmakers containing non-actual particulars are always grounded (directly or indirectly) in truthmakers that include only actual ones. For example, the intentional relations between Sherlock Holmes and RCK are grounded in psychological facts exclusively about RCK, facts involving relations between actual things, internal psychological activities, and RCK. The truthmaker for true propositions about RCK's attitudes toward Sherlock Holmes is not fundamental but is rather wholly grounded in facts about RCK's psychology. The Principle of Actual Parts applies only to the most fundamental truthmakers.

Second, Anti-Actualists could say that non-existent objects, as parts of existing facts, stand only in external relations. They stand in no internal relations or and have no qualities. They are *super-thin*, much more abstract than normal abstract objects such as numbers and sets. For instance, each set has some internal nature that grounds the fact that it has certain members. Non-existent objects have no such internal nature. They have only intentional and modal properties, grounded in the natures of actual things.

### 12.1.5 Arguments against Actualism

On the other side of the ledger, we examine four arguments in favor of Anti-Actualism. The first appeals to common sense. The second appeals to the need for *intentional objects*, that is, objects of certain mental attitudes. The third appeals to the need for fictional and illusory entities. And the fourth appeals to the need for non-existent things as constituents of certain merely possible situations.

1. Common sense. As we've seen, it seems reasonable to say that many things, golden mountains and round squares among them, don't exist. Adopting a common-sense approach to the truth of these claims leads naturally to Anti-Actualism.

2. Intentional objects. Human beings can, apparently, bear *de re* attitudes like belief, though, desire, and intention toward merely possible or even impossible objects. An example of such an attitude would be that of *seeking the Fountain of Youth*. A *de re* attitude is an attitude directed toward a particular thing, in contrast to a *de dicto* attitude, which takes a property or a description as its object.

Saul Kripke, in his classic *Naming and Necessity* (Kripke 1980), builds a strong case for the existence of *de re* attitudes toward individual objects that are directed toward those objects *not* by way of some general description. Suppose, for example, that one is seeking O'Hare Airport in Chicago because one has a reservation on a flight that leaves from there. One has beliefs, thoughts, and desires that are directed toward a certain thing, namely O'Hare Airport. These attitudes cannot be identified with any corresponding attitudes toward a certain kind of thing. For example, in seeking O'Hare, one is not seeking the largest airport in the Midwest; at least, one isn't seeking it under that description. One might not know or care that O'Hare is the largest airport in the Midwest. Similarly, one is not seeking an airport that is commonly named 'O'Hare'. Again, one doesn't really care what it's called. If, unbeknownst to one, the airport's name had been changed to 'Daley Airport', one's seeking would have the same object: one would still be trying to get *there*, whatever it's called.

Meinongians point out that such *de re* attitudes can also be directed toward non-existent or even impossible objects. Ponce de León was seeking the Fountain of Youth, ancient Greeks worshipped Zeus, and many would-be detectives admire and emulate Sherlock Holmes.

At any rate, Anti-Actualists can claim that we ought to posit non-existent objects to function as the objects of these intentional attitudes.

A simple Actualist response to this problem is to accept that we do have *de re* attitudes in such cases, and such attitudes do involve our being related to something as the object of the attitude. However, Actualists can deny that we need non-existent things to serve as those objects. Instead, they can press into service abstract entities of various kinds: properties, states of affairs or propositions. So, when Ponce de León was seeking the Fountain of Youth, he was in the *seeking* attitude toward the state of affairs in which he finds the Fountain of Youth or the property of *being a finder of the Fountain of Youth* or the proposition with the content that *he himself finds the Fountain of Youth*. All of these abstract entities, Actualists might claim, actually exist even if there is no Fountain of Youth.

However, this view does involve some cost. From a naive point of view, it seems that Ponce de León did not seek a property or a centered proposition or anything of the kind. He sought something that, were it to exist, would be a fountain, something concrete. Properties and propositions, in contrast, are necessarily abstract.

In addition, Actualists face a dilemma. They must either suppose that all of our *de re* attitudes are actually directed toward abstract objects, even when there is a real, concrete object available (like O'Hare Airport) or they must give two separate accounts of *de re* attitudes, one for cases in which an appropriate object exists and another for cases in which no such object exists.

Both approaches come with some cost. We should always prefer a unified account to a combination of disparate accounts. This is another aspect of Ockham's Razor. If Actualists try to give a unified account in terms of abstract objects like properties or concepts, they will then have to answer the objections to such *descriptivist* accounts raised by Kripke in *Naming and Necessity*. When one seeks O'Hare, one is in a particular relation to a concrete object. Similarly, if one sees Bill Clinton, one sees *him*, the former President, and this perceptual relation cannot be cashed out in terms of a relation to any description, such as seeing the forty-second President or seeing the man named 'Bill Clinton'.

Probably the best Actualist account is one proposed by Alvin Plantinga in *The Nature of Necessity* (Plantinga 1974). Plantinga introduces the notion of a *haecceity* or *thisness* (Def 9.2). By way of reminder, a thisness or haecceity is a property that of necessity is possessed by at most one thing and could not be possessed by any other thing. The properties of *being O'Hare Airport* or *being Bill Clinton* are examples of haecceities. O'Hare Airport necessarily has the property of *being O'Hare Airport* (if it exists at all), and nothing other than O'Hare could have that property. Plantinga proposes that haecceities, like other properties, exist of necessity. O'Hare itself does not exist of necessity, but if it didn't exist, its haecceity, though uninstantiated, would still exist.

If there were such things as haecceities, then they could serve as the objects of all *de re* attitudes. When one seeks O'Hare Airport, one is really seeking to be in a place that instantiates the haecceity of *being O'Hare Airport*. When one sees Bill Clinton, one is seeing the object that instantiates the haecceity of *being Bill Clinton*. Similarly, when Ponce de León was seeking the Fountain of Youth, he was seeking to be in a place that instantiated the haecceity of *being the Fountain of Youth*. Although the Fountain of Youth does not exist, its haecceity might, and it might therefore serve as the object of Ponce de León's *seeking* attitude.

Other Actualists, like Robert M. Adams (1981), have expressed skepticism about whether there really are such things as uninstantiated haecceities. Adams has argued that the only haecceities that exist are the haecceities of actual objects, since Bill Clinton's haecceity contains Bill Clinton himself as a part or component. In Chapter 16, on *de re* modality, we take up again the issue of whether there are such things as haecceities, including uninstantiated ones.

Another Actualist response to this problem about intentional objects is to suppose that, strictly speaking, our intentional states do not relate us to things as their objects. Instead, we should re-interpret claims about intentional states using only adverbs. If Smith believes in unicorns, it is not the case that he stands in the *believing* relation to some unicorns or even to the set of unicorns or the property of *being a unicorn*. Instead, she believes *unicornly*. Similarly, Ponce de León was seeking in a *Fountain-of-Youthish* way.

Roderick Chisholm (1973a) pointed out a problem for this adverbial account. The following argument seems valid:

(29) Jones is thinking of unicorns.
(30) Jones thinking of something.
(31) Therefore, there are unicorns.

The adverbial account cannot explain the logical structure of these propositions in such a way as to give an account of the argument's validity that isn't entirely ad hoc. In other words, adverbialists must introduce a large number of new logical principles governing adverbial constructions. Possibilists, in contrast, can explain the argument's validity much more simply, using just the usual logical rules of the quantifiers. So, attributing Jones's thought to her should involve placing her in some relation to unicorns.

3. Fictional and illusory objects as persistent and available to many minds. We've already mentioned Peter Geach's imagined situation in which two medieval peasants, Hob and Nob, share a belief in a non-existent witch who has cursed their village (1967). It seems that there must exist in some sense an object that is believed by both Hob and Nob to exist and to be a witch in their village. There need not be a description or property that Hob and Nob agree upon that could be used to define the common object in their shared fantasy. Such shared objects seem to require exportation. Hob believes that there is a witch, Nob believes there is a witch, and the witch that Hob believes in is

identical to the witch that Nob believes in. In order to make sense of this, there must be a non-existent witch believed in by both villagers.

It is true that Hob could define the witch in terms that refer to Nob's beliefs. For example, Hob could think of the witch as *the witch that Nob believes cursed his cow last month*. If Hob believes that *that witch* also cursed his chickens today, there is a sense in which we can identify an object in Hob's imagined world with one in Nob's imagined world. However, it is quite possible that Hob and Nob could begin to refer to 'the witch' in their shared conversations after they have both forgotten who first thought of it. At that point, we cannot define 'the witch' in terms of any property or set of properties, whether ordinary or intentional.

A similar problem arises from the persistence through time and change of fictional and illusory objects. The same non-existent object can be the object of thought or perception at different times by the same person or by different people at the same time. These non-existent objects cannot simply be identified with sets of properties. Sherlock Holmes, for example, acquired more representational properties as Conan Doyle wrote new stories.

4. Non-existent objects as components of merely possible states of affairs. There could have been more things in existence than actually exist. There are possibilities according to which some things exist that do not exist in the actual world. If something exists according to a possibility, then it is in some sense a component or constituent of that possibility. The possibility *consists of* that very thing, standing in certain possible relations to other things. If that is right, then we have another reason to countenance non-existent objects.

Alvin Plantinga's (1974) Actualist response again involves haecceities. The possibilities apparently involving non-existent *things* really contain only uninstantiated haecceities. Again, critics of Plantinga, including Robert M. Adams, worry that haecceities involve non-existent things no less than possibilities do. Our haecceitiess (the properties of *being Robert C. Koons* and of *being Timothy H. Pickavance*) include RCK and THP themselves as constituents, say these objectors. Hence, uninstantiated haecceities must contain non-existent things. If Adams is right, then haecceities don't blunt this Anti-Actualist argument.

There is another problem with Plantinga's haecceity proposal. We can't make significant *de re* assertions about the relations between things and their haecceities. For example, we cannot say, in an informative way, that necessarily, if Socrates exists, his haecceity is instantiated. This becomes (on Plantinga's account) a mere logical tautology, instead of a substantive metaphysical proposition. It turns out to be nothing more than to say that necessarily, if Socrates' haecceity is instantiated, then it is instantiated, which isn't what we really want to say. We want to say something about the necessary relations between *things* and their haecceities, and yet this important feature of Plantinga's theory seems to be inexpressible if the theory is true.

Another Actualist response is what David Lewis (1986a) called the "ersatz" Actualist response. According to this response, a possible world is a kind of representation. So a possibility according to which non-actual things exist is not one that literally contains non-existent things or mysterious haecceities as constituents. Instead, it contains things that *represent* the possible existence of non-actual things.

This ersatz Actualist has a problem representing iterated *de re* possibilities, possibilities involving particular things. For example, there could exist a world undergoing eternal cycles without beginning or end, each cycle qualitatively like all the others, and each one containing a duplicate of Napoleon. If such a world, $w_0$, had been actual, then each "Napoleon" in $w_0$ could have been the very last "Napoleon" in another world. These further possibilities, or possible possibilities, should correspond to an infinite number of distinct worlds, each one with a different Napoleon from $w_0$ as its last.

The problem for ersatz Actualists is that they have no way to distinguish the many different *de re* possibilities from one another. For example, there should be one possibility, $w_t$, corresponding to *this* Napoleon's (picking out one of the infinite series from $w_0$) being the last and another, different possibility $w_{t+1}$ in which *the next* Napoleon is the last one. However, for the ersatz realist, these two possibilities would be the same. Both would contain an infinite series of qualitatively similar "Napoleon"-representations with an end but no beginning.

We can make the same point using a Max Black-style possible world $w_{MB}$. Suppose that the original Big Bang in our past had had the power to produce two exactly similar and symmetrically arranged black holes. Suppose further that the two black holes had had a kind of free will, or some other unpredictable and undetermined power to act. Let's say, for example, that each of the two black holes might have destroyed itself; $w_1$ is a world where one black hole destroys itself and $w_2$ is a world where the other black hole destroys itself. Ersatz Actualists have a problem explaining how it is possible for $w_1$ and $w_2$ to be distinct. What could distinguish them? They involve the same intrinsic properties and the same spatiotemporal relations to all of the inhabitants of the actual world. Could ersatz Actualists say that there is just a brute distinctness between the two possible worlds? This doesn't seem right. The two possible worlds are distinct *because* the black holes are (or would be) distinct.

There is really only one way out for Actualists: give up the modal logical principle S4, which says that what is necessary is necessarily necessary, and what is possibly possible is possible. Without S4, ersatz Actualists could argue that there *are not* in fact two such possibilities (one in which one black hole destroys itself, and another in which the other does), although there *could have been* two such possibilities. If $w_{MB}$ had been actual, there would have been $w_1$ and $w_2$, but because it is not, there are no such possibilities. In other worlds, these two further possibilities are themselves only possibly existent, and so the things that exist according to them are only possibly possible, not actually possible. This is probably the best response for Actualists, although it comes with some significant theoretical cost, namely, the loss of the simplest and most powerful modal logic, S5, which includes S4 (see Chapter 16 for a further discussion of de re possibilities).

### 12.1.6 Defining existence

Throughout this chapter we've been discussing existence, and yet we have never said what existence itself is. There's a good reason for this omission. For Actualists, it is arguably impossible to define existence, since existence is not a property that some things lack and other things have. Instead, it is trivially true that absolutely everything exists.

In contrast, Anti-Actualists can sensibly ask the question of whether existence is definable. For Possibilists, it is actual existence that must be defined, since they hold that it is trivially true that everything has at least possible existence. We will look in detail at actuality in Chapters 14 and 15. Consequently, we will focus here on Meinongian definitions of 'existence'.

Importantly, we cannot assume that absolutely every property or concept can be defined. Some notions are so fundamental that they cannot be defined. Existence might be one such notion. Nonetheless, some philosophers have proposed something like definitions of existence. For example, George Berkeley argued that *esse est percipi*, that to exist is to be perceived. To be completely accurate, we should say that for Berkeley there are two ways to exist: to perceive or to be perceived. If Berkeley were right, the very idea that something could exist while neither perceiving nor being perceived would be absurd. However, it does not seem absurd. It seems relatively easy to imagine such a thing, so by Imagination as Guide to Possibility (**PEpist 1**), it is reasonable to think it is possible to exist without perceiving or being perceived.

However, Berkeley had a response to this argument. He claimed that we cannot in fact imagine something that is unperceived, since the very act of imagining something involves imagining something to be perceived by one's self. Berkeley is clearly wrong here because he failed to distinguish between the act of imagining and the content of the act. Just as we see colors and do not see our own acts of seeing, so too we can imagine colored things without imagining our act of imagining them. I can imagine a tree falling in a lonely forest, making a great crashing sound, without imagining anyone's hearing the sound.

Samuel Alexander (1920) proposed a different definition of existence: to exist is *to act*, to have some effect on something else or at least to have the power to act upon something. This is Alexander's Dictum, sometimes called the Eleatic Principle, from a passage in Plato's *Sophist* in which a character, the Eleatic Stranger, suggests that having causal power is the essential mark of being (*Sophist* 247d3).

The trouble with Alexander's Dictum is that it seems possible that something exist without active causal powers. We can easily imagine something, like an invisible or ghostly observer, that is affected by other things but forever unable to have any effect. So, we might consider modifying Alexander's Dictum to the claim that to exist is to have some causal power, either active or passive. Things with causal powers have something further in common, namely, some intrinsic nature or quality. We should, therefore, entertain the possibility that to exist is to have some intrinsic nature.

This suggestion accords nicely with Anti-Actualism's problem with truthmaking. Recall that Meinongians could solve this puzzle by assuming that non-existent objects have only modal and intentional properties. Such entities would have no intrinsic natures, and so would lack existence in that final sense.

## 12.2 Ontic Vagueness

In our discussion of Monism (Section 11.2.4), we raised a question from Jonathan Schaffer (2010): can there exist things with vague boundaries, whether spatial, mereological, or temporal? Something with a vague spatial boundary would have a somewhat indeterminate location. We couldn't say (truthfully) precisely where the thing is at any point in time. A mereologically vague boundary would involve some vagueness in the composition of a thing. There would be material bits that are neither definitely a part of the thing nor definitely not a part. A thing would have a vague temporal boundary if there is no precise point in time at which it begins or ceases to exist.

What does it mean to say that something like a boundary is vague? Broadly, there are two theories of vagueness, epistemic and metaphysical (or semantic). According to Epistemicists, everything in fact has precise, determinate boundaries, but there are many cases in which we do not, and perhaps cannot, know exactly where those boundaries lie. Anti-Epistemicists maintain that in some cases there is no fact of the matter to be known. For Anti-Epistemicists, there are failures of the Law of Bivalence (the principle that True and False are the only truth-values a proposition can have) associated with the existence of vague boundaries. That is to say, sometimes a proposition about a boundary will be neither true nor false. Suppose Felix is a cat, and molecule *M* is vaguely a part of Felix. Anti-Epistemicists claim that (32) is neither true nor false:

(32) Molecule *M* is part of Felix.

In contrast, Epistemicists insist that (32) is either true or false, even if we cannot know which it is. Epistemicists claim that absolutely nothing in the world is vague. Every boundary of everything is precise and determinate.

**12.2T Epistemicism.** Every boundary is determinate. Any vagueness is merely epistemic, that is, merely a matter of our not knowing exactly where the boundary lies.

**12.2A Anti-Epistemicism.** Some boundaries are indeterminate, resulting in failures of the Law of Bivalence.

Epistemicism has a number of promising features. First, Epistemicists can hold to Bivalence. Bivalence provides a simple and convincing basis for the applicability of all of the principles of classical logic, including the Law of Non-Contradiction (no proposition is both true and false) and the Law of Excluded Middle (every proposition of the form *p or not-p* is a logical truth). Whatever we think about (32), the following propositions seem to be uncontroversially true:

(33) It is not the case that *M* both is and is not a part of Felix.
(34) Either *M* is a part of Felix or it is not.
(35) If *M* is a part of Felix, then *M* is a part of Felix.

However, if (32) is neither true nor false, it is not obvious how to explain the truth of (33)–(35).

Second, Epistemicists have no difficulty handling *higher-order vagueness*, vagueness about vague boundaries. For example, consider the question of whether the predicate 'is vague' is itself vague. For Epistemicists, this just amounts to the question of whether there are some cases in which we are ignorant about what we are ignorant about. It wouldn't be at all surprising or puzzling if we were ignorant about our ignorance.

Third, Timothy Williamson (2000) has provided an interesting and substantial theory that explains why we are ignorant about many boundaries. He does so by appealing to his *margin of error* principle. We can know a proposition *p* just in case we justifiably believe *p* now and we would still believe *p*, in any "nearby" counterfactual situation in which *p* remains true. So, I know that some molecule *M* in Felix's heart is part of Felix because I would still believe that *M* was part of Felix even if *M* were moved slightly to the left or right. This is not true of molecules at or very near the outer limits of Felix, since moving *M* slightly, even if it still remains part of Felix, could easily tempt me to error or at least to suspension of belief.

Finally, Epistemicists can appeal to the context-sensitivity of many vague terms and predicates (Kamp 1975). Terms are context-sensitive when what falls within the extension of a term depends on the conversational context within which the term is being used. Adjectives like 'bald' and 'tall' are paradigmatically context-sensitive. Someone might count as 'tall' in a conversation about college students but not in a conversation about college basketball players. These facts about contextual variation support Epistemicism in two ways. First, they make it harder for one to be sure about where the boundary lies in a particular context, since one may be unsure about some of the contextual factors. Second, they help to explain why what are actually sharp boundaries seem fuzzy or blurry: boundaries are in a constant state of flux.

Nonetheless, Epistemicism is difficult for many to swallow. It is hard to see how our reference to things and classes of things could be so much more precise than our ability to recognize their boundaries.

One way of motivating Anti-Epistemicism is by noting that it seems plausible that all vagueness is linguistic or conceptual in character. Some philosophers have asserted that vagueness exists only in language or in the mind, not "in the world". For example, David Lewis says,

> The only intelligible account of vagueness locates it in our thought and language. The reason why it's vague where The Outback begins is not that there's this thing, The Outback, with imprecise boundaries; rather there are many things, with different borders, and nobody has been fool enough to try to enforce a choice of one of them as the official referent of the word 'outback'. Vagueness is semantic indecision. (Lewis 1986a: 212)

Taken literally, this is nonsense, since language and the mind are just as much "in the world" as anything else. It is part of the very nature of a proper name (like 'The Outback') that it have a single referent, just as it is part of the very nature of a concept to have a single set as its extension. Lewis's talk of "semantic indecision" obscures the fact that he's positing real, worldly vagueness in the nature of the name 'The Outback' (see Merricks 2001 and Salmon 2010).

The most charitable interpretation of this Linguistic Theory of Vagueness is that vagueness is simply a matter of ambiguity. Vague terms refer ambiguously to a large number of precise entities, and vague predicates ambiguously express a large number of precise properties. We can call this the 'Multiple Meaning Theory of Vagueness'. If terms and predicates are not ambiguous in this way, and if vagueness is not merely epistemic, then these terms must refer to things that are really vague or indeterminate in their boundaries. Thus, we get two possible versions of Anti-Epistemicism.

**12.2A.1T Multiple Meaning Theory of Vagueness.** Some boundaries are indeterminate, but all such indeterminacy is merely a matter of ambiguous reference to fully determinate entities.

**12.2A.1A Real Ontological Vagueness.** Some entities lack determinate boundaries, independently of our knowledge or how we refer to them.

### 12.2.1 The multiple meaning theory of vagueness.

Multiple Meaning Theorists argue that *vague* things are really non-empty *classes* of *precise* things. Terms that seem to be singular, like the proper name 'Felix', really refer to a class of things, a class containing a large number of cat-like or nearly-*cattish* material bodies in a certain neighborhood. One way to work this out is to assume that vague terms are ambiguous, in something like the way the word 'bank' has at least two different meanings (the side of a river or a lending institution). Multiple Meaning Theorists should suppose that vague names like 'Felix' are massively ambiguous, with billions or trillions of distinct meanings and referents. (see Unger 1980 and Lewis 1993). (32) is indeterminate because molecule *M* belongs to some of these referents and fails to belong to others. The indeterminacy of (32) would be similar to the indeterminacy of (36):

(36) John and Mary went to a bank.

(36) could come out as true on one interpretation of 'bank' and false on the other, if John and Mary went to a side of a river, but not to a lending institution.

Such an ambiguity theory can provide an explanation for the truth of instances of the theorems of classical logic, including (33) through (35). The explanation takes the form of the device of *supervaluations* (van Fraassen 1969, Fine 1975, Kamp 1975). An ambiguous statement counts as *super-true* if it comes out true under every permissible interpretation or *precisification* of its vague/ambiguous terms. Sentences (33) through (35) are super-true, since no matter what precise referent we assign to the name 'Felix' (regardless of whether the referent contains or doesn't contain *M*), the statements come out true. However, although supervaluational semantics agrees with all of the theorems of classical logic, it does not endorse as valid all of the methods of classical logic, such as *reductio ad absurdum* (indirect proof) or conditional proof (Williamson 1994: 151–153).

Multiple Meaning Theory faces several serious problems. First, there is a legitimate question as to whether Multiple Meaning Theorists have pushed the concepts of *meaning* and *ambiguity* beyond their breaking points. We understand what it is for a word like 'bank' to be ambiguous: there are two semantical rules associated with the phonetical and grammatical symbol. However, in the case of a name like 'Felix', we have to imagine billions or trillions of distinct semantic interpretations, all of which are somehow encompassed as permissible by our linguistic practices.

In response, Multiple Meaning Theorists could point out that on Lewis's account there is a difference between lexical ambiguity and vagueness. Vagueness is the result of semantic incompleteness or "indecision", as Lewis puts it. The incomplete meaning is compatible with multiple, equally good completions. Lexical ambiguity, on the other hand, is the result of our decisions, rather than indecision. The community of English-speakers, for example, have chosen to use 'bank' to pick out two very different types of thing.

Second, there is the problem of higher-order vagueness. Just as it is can be unclear whether molecule *M* is or is not part of Felix, if 'Felix' is ambiguous and molecule *M* lies in the borderline region contained by some but not all the referents of 'Felix', so can it be unclear whether or not some molecule *M* is a borderline case. The Multiple Meaning Theorists seem to be committed to saying that there is a precise answer as to which molecules belong to some referent of 'Felix' and which belong to none. But this seems just as indeterminate as the first-order question of which molecules belong to Felix.

Finally, Multiple Meaning Theory, when combined with the view that human persons are vague material objects, destroys the unity of each person. RCK is surely aware of his own unity as a thinking being, and yet Multiple Meaning Theory would entail that 'I' in RCK's mind or mouth in fact refers to a vast multitude of overlapping material thinkers. This vast multiplication of human persons would be deeply inconsistent with many of our most central ethical and political beliefs and practices. If Smith is guilty of a crime or responsible for some act of heroism, which of the trillions of Smiths should be punished or rewarded, and how do we ensure that we interact with the right ones?

### 12.2.2 Real ontological vagueness

Let's turn finally to Real Ontological Vagueness, the view that there is real vagueness or indeterminacy in the world. If vague objects are non-fundamental or grounded, their grounding in precisely-bounded objects is a case of what we called extra-conceptual grounding (in Section 3.4). Such grounding does not enable us to deny that vague objects *really* exist.

How can we make sense of this view? One way would be to give up Bivalence. Suppose that Fred is borderline bald. We might say that is neither true nor false (speaking precisely) that Fred is bald. Or, suppose molecule *M* is a borderline case of a proper part of Felix. We could say that it is neither true nor false that Felix includes *M*. This is the *three-valued proposal* for ontological vagueness.

There are three major drawbacks to the three-valued proposal. First, giving up Bivalence gives us good reason to give up classical logic more generally, including the Law of Excluded Middle:

**Law of Excluded Middle.** For every proposition *p*, the proposition of the form '*p* or not-*p*' is a logical truth.

If some propositions are neither true nor false, then it seems that there should be cases in which neither *p* nor not-*p* are true. Given our standard understanding of the truth-conditions for 'or', these cases should give rise to exceptions to Excluded Middle, as well as many other principles of classical logic.

This is a heavy price to pay for a dubious metaphysical gain. First, we should surely want to say things like 'If Fred is bald, then Fred is bald', or 'If Fred is bald, then someone just like him with even fewer hairs is bald' to be logically or definitionally true, but this will be hard to do without the resources of classical logic.

Second, consider the principle known as 'Tarski's Schema' (after Alfred Tarski):

**PTruth 2 Tarski's Schema.** For any sentence *s*, if '*S*' is a name for *s*, then we should affirm the sentence of the form: *S* is true if and only if *s*.

The final clause, '*S* is true if and only if *s*', is sometimes called a 'Tarski biconditional'. On the left side of the biconditional ('if and only if') the name *S* is named or mentioned, and on the right side it is used. A classic example of a Tarski biconditional is (37):

(37) 'Snow is white' is true if and only if snow is white.

Not only do Tarski biconditionals seem to be logically or definitionally true, but they seem to have a further interesting characteristic: the left-hand and right-hand sides seem to be saying exactly equivalent things. These two sides should have the same semantic value. However, the three-valued proposal violates this rule. If Fred is a borderline case of baldness, then 'Fred is bald' is neither true nor false. If so, then (38) will be false:

(38) 'Fred is bald' is true.

(38) is false, since 'Fred is bald' is neither true nor false. Consider the relevant instance of Tarski's biconditional:

(39) 'Fred is bald' is true if and only if Fred is bald.

Since (38) is simply false, the left-hand of (39) is false and the right-hand side is neither true nor false, violating the semantic equivalence between the two.

Third, the three-valued proposal does not give us a workable account of higher-order vagueness. Suppose that Fred is neither definitely vaguely bald nor definitely definitely bald. That is, suppose he is a borderline case of a borderline case of baldness. What should we say about the sentence 'Fred is bald' in that case? Is it true, false or neither? If we say that it is true or that it is false, then we must conclude that Fred is definitely not a case of vagueness. If we say that it is neither true nor false, then we must conclude that Fred is definitely a case of vagueness. How can we get that he is a vague case of vagueness? Do we have to keep increasing the stock of non-standard truth-values, like vaguely true or vaguely false? Where, if ever, does this process end? And how do the various truth-values relate to one another? The three-valued proposal seems to be leading into a morass.

There is an alternative approach that we might consider, the Many Actual Worlds approach (see Elizabeth Barnes 2010). On this view, there might be more than one *world* that is actual. If Fred is a case of baldness, then both a world in which he is bald and a world in which he is not bald might be actual.

We will discuss the idea of possible worlds in more detail in Chapters 14–16, but our discussion here will require a bit more detail that we've needed to this point. For now, think of a possible world as a comprehensive state of affairs, a comprehensive way for things to be. Some states of affairs, like the Patriots' winning Super Bowl XLIX, are relatively non-comprehensive, leaving many facts undetermined. A possible world is a *maximal* state of affairs, one that settles all of the matters of fact in a possible scenario. More precisely, a state of affairs $A$ is maximal if and only if there is no state of affairs $A'$ such that it's possible both that $A$ obtains $A'$ obtains and that $A$ obtains and $A'$ doesn't obtain. Alternatively, we could think of worlds as maximal compossible sets of states of affairs. A set of states of affairs is compossible if it is possible that all the members of the set obtain together. But what makes such a set *maximal*? Just that no further state of affairs can be added to the set without producing a non-compossible set. (These definitions are from Plantinga 1974.)

However, if two or more mutually inconsistent worlds can all be actual together, as on the Many Actual Worlds approach, then none of them will qualify as maximal. We'll have to define 'possible world' in a different way: a possible world is a minimal set of states of affairs that could include everything that is actual. Each possible world is comprehensive enough that it could comprise the whole of reality, but no bigger.

If more than one world can be actual, how do we define 'truth in a world'? We can't define it in Plantinga's way (1974): namely, *p* is true in *w* if and only if *p* would be true if *w* were actual. A world *w* in which *p* is false might be actual, along with *another* world in which *p* is true. World *w*'s being actual isn't enough to ensure that *p* would be false, even if *p* is false 'in' or 'according to' *w*. We will have to appeal instead to what would be the case if *w* were the only actual world:

> *p* is true in *w* if and only if *p* would be true if *w* were the only actual world.

We can now say that a proposition *p* is *super-true* if and only if *p* is true in every actual world. Now there will be propositions that are neither super-true nor super-false, but all logical truths and necessary truths will be super-true, including Excluded Middle. We get to use all of classical logic, by van Fraassen's (1969) method of supervaluations, now interpreted ontologically instead of semantically.

However, there is a fly in the ointment: it will turn out to be super-true that there is only one actual world! Given our definition of truth in a world, the proposition that *w* is the only actual world will always be true in *w*, and so it will always be super-true, no matter how many worlds are actual. Our account of truth undermines the very theory we started with.

So, proponents of Many Actual Worlds need a more complicated theory of truth. The initial definition will work for all atomic propositions other than those that predicate actuality of individual worlds. Let's say that a set of worlds is a *cluster of worlds* if it is possible for all of them to be actual together. We can now define what it is for a proposition to be true relative to a pair <*w*,*S*>, where *w* is a world and *S* is a cluster containing *w*. A proposition of the form '$w_2$ is actual' will be true relative to <*w*,*S*> if and only if $w_2$ is a member of the cluster *S*. Other atomic propositions are true relative to <*w*,*S*> if and only if they would be true if *w* were the only actual world. A proposition *p* is *super-true* if and only if it is true relative to every pair <*w*,*S*>, where *w* is an actual world and *S* is the cluster of all actual worlds. All of the theorems of classical logic will be true relative to all worlds and clusters, and so these logical truths will always necessarily be super-true. So we preserve all of classical logic.

This notion of a cluster of worlds could also be spelled out in terms of totality facts (Def D2.5, Section 2.4.2). A totality fact is a fact connecting a property with a set of objects, a set representing the totality of things that instantiate or are characterized by that property. In the case of ontological vagueness, it could be that some properties participate in multiple totality facts, in such a way that those properties have extensions with multiple, inconsistent boundaries. Let's assume that each possible world assigns just a single totality fact to each property. If a property is capable of participating in a plurality of totality facts, then this plurality of facts corresponds to a cluster of possible worlds, all of which could be actual together. If several properties are capable of participating independently in multiple totality facts, then the size of the cluster multiplies accordingly, corresponding to what mathematicians call the "cross product" of the permissible property-totality pairings.

When is 'Definitely *p*' true? It would be best not to identify the proposition that definitely *p* with the proposition that *p* is super-true. Instead, we could say something like this: 'Definitely *p*' is true relative to $<w,S>$ if and only if *p* is true in every pair $<w_2,S_2>$, where $S_2$ is a possible cluster of worlds, and $w_2$ is a member of both *S* and $S_2$. Let's say that one world is a 'possible companion' of another just in case it is possible for both worlds to be actual together (in a single cluster). An ordinary, non-modal proposition is definitely true in one cluster if and only if it is true in every possible companion of every world in the cluster. Definite truth is a stronger condition than super-truth. Some super-true propositions are not definitely true, but every definitely true proposition is super-true.

For example, suppose there were only one actual world *w*, but suppose that it would have been possible for *w* to have been joined in actuality by other worlds, including $w_2$. If *p* is true in *w* but not in $w_2$, then *p* would be super-true but not *definitely* true relative to $<\mathrm{w},\{\mathrm{w}\}>$. However, all logical truths will be definitely true relative to every world-cluster pair, since logical truths are true in every world without exception.

Why is Many Actual Worlds preferable to the three-valued proposal? The superiority does not lie in the fact that on the three-valued proposal we have propositions that are neither true nor false, while on Many Actual Worlds some propositions are both true and false. In the Many Actual Worlds case, we do still have truth-value gaps of a kind: propositions can be super-true, super-false or neither. The crucial difference is that Many Actual Worlds can make use of the method of supervaluations, preserving both classical logic and all of the instances of Tarski's Schema.

Many Actual Worlds also enables us to mimic Epistemicism's account of higher-order vagueness. We have to assume that each world has an inherent tolerance for a certain degree of deviation. Each world has only a limited number of possible companions. This limited tolerance for deviant companions bears a close analogy to our finite capacity for knowledge and discrimination.

That is, we can have propositions that are definitely true but not definitely definitely true or propositions that are vague but not definitely vague.

Is Many Actual Worlds compatible with Truthmaker Theory, in particular with Atomic Truthmaker Theory? Prima facie, there are two problems. We want Not-*p* to be true in a world even if *p* is also true in another actual world, and we don't want (*p*&Not-*p*) to be true just because $p$ is true in one actual world and Not-$p$ in another. So, how do logically complex propositions find appropriate grounds in truthmakers?

We can use facts about worlds (if there are such) to group truthmakers into *worldly cohorts*. A set $S$ of truthmakers is a *worldly cohort* just in case there is an actual world $w$ such that, if $w$ were the *only* actual world, $S$ *would* contain all and only the truthmakers that exist. With this notion of *worldly cohorts* in place, we can construct a slightly modified version of Atomic Truthmaker Theory. For example, a negated atomic proposition Not-$p$ is true if and only if there is some worldly cohort that does not include a truthmaker for $p$. Similarly, a conjunction ($p\&q$) is true if and only if there is some worldly cohort that does include both a truthmaker for $p$ and one for $q$.

Here is a specific example. RCK is indefinitely $6' \, 1''$. There are actual worlds $w_1$ and $w_2$ such that RCK is $6' \, 1''$ in $w_1$ and exactly $6' \, 1'' + 1$ nanometer in $w_2$. The truthmakers for RCK's being exactly $6' \, 1''$ and for RCK's being $6' \, 1'' + 1$ nanometer must belong to two different worldly cohorts.

An alternative approach for the defender of Real Ontological Vagueness would be to embrace Truthmaker Maximalism with its totality facts (see Section 2.4.2). In a perfectly precise world, each property would have exactly one totality fact, specifying which things are (and by omission which things are not) in the property's extension. In the case of Real Ontological Vagueness, one or more properties would *tremble* a bit in their grip on totality facts, resulting in a plurality of totality facts associated with each vague property. When something belongs to the extension of a property according to one totality fact associated with that property but does not belong to its extension according to another totality fact, then that particular will be a vague or borderline instance of the property. If there are spatial location properties, then some entities will have vague spatial boundaries and, consequently, vague material composition.

### 12.2.3 Can identity or existence be vague?

Gareth Evans (1978) argued that there is at least one set of facts that cannot be vague, namely, facts about identity and distinctness. Evans argues that, for every entity $x$, it is always definitely the case that $x = x$. Now suppose, for contradiction, that there is an object $y$ that is vaguely identical to $x$. That is, it is neither definitely identical to $x$ nor definitely distinct from $x$.

Now suppose that $y$ is identical to $x$. Since $x$ has the property of *being definitely identical to* $x$, then $y$ must have that property. Thus, if $y$ is identical to $x$, then it is definitely identical to $x$.

Suppose instead that $y$ is not identical to $x$. If $y$ is not identical to $x$, then it surely cannot be definitely identical to $x$. In fact, if $y$ is not identical to $x$, it is definitely not definitely identical to $x$. However, $x$ is definitely definitely identical to $x$. So, $x$ and $y$ definitely differ in at least one property, namely, the property of *being definitely identical to* $x$. Therefore, $y$ is definitely not identical to $x$. Thus, if $y$ is not identical to $x$, then it is definitely not identical to $x$.

Since it follows from the Law of Excluded Middle that either $y$ is identical to $x$ or $y$ is not identical to $x$, we have proved that either $y$ is definitely identical to $x$ or $y$ is definitely not identical to $x$. There can be no vague identity (assuming Real Ontological Vagueness).

Actualists have a simple argument for thinking that there can be no vague cases of existence. For any entity $x$, the existence of $x$ can be defined (for Actualists) as the logical tautology $x = x$. Certainly, this logical tautology must be definitely true. Thus, every case of existence is a case of definite existence. Anti-Actualists, in contrast, treat *existence* as a property that does not belong, as a matter of logic, to every entity. So, Anti-Actualists can countenance vague existence.

## 12.3 Conclusion

We've covered two main topics in this chapter, both concerned with the scope of existence: are there non-existent or non-actual things, and are there vague objects? The No answers (Actualism and Epistemicism) seem the most conservative, from an ontological point of view. They certainly avoid multiplying objects. However, Anti-Actualists and defenders of Anti-Epistemicism (such as the Many Actual Worlds theory) insist that we have good reason for the multiplication.

The strongest case for Anti-Actualism appeals to the role of merely possible objects as objects of intentional attitudes like seeking, fearing or worshipping, and as constituents of the truthmakers of modal facts. To meet this case, Actualists must resort to ersatz entities, like haecceities or abstract representations. We will examine this debate in further detail in Chapters 14–16.

The positive case for Real Ontological Vagueness consists simply in the strong impression that many things lack definite boundaries or definite quantitative features. This positive case faces two challenges. One comes from Epistemicists, who insist that all apparently vague boundaries are merely projections of our own ignorance. The other comes from Multiple Meaning Theorists, who locate all vagueness in ambiguities in our language or our concepts.

# 13

# Solipsism, Idealism, and the Problem of Perception

We've examined reasons for thinking that the world includes more than one particular. Now we can turn to questions about what those particulars are like. In subsequent chapters, we take up a wide range of such questions, including:

- Do particulars change over time? Do they persist through time? Do they begin to exist or cease to exist?
- What is it for particulars to be located in space? Are regions of space themselves particulars? If so, of what kind?
- Do some particulars have other particulars as parts or are all particulars simple? When do some things compose a further, composite thing?
- Do some particulars cause others? Does everything have a cause? What sorts of things can be causes or effects?

However, before taking up those questions, we should ask a more fundamental set of questions. We must consider whether the world really includes a host of physical objects arranged in space and time, as our sense experience suggests. We do not consider the possibility of an all-encompassing form of skepticism, that kind of skepticism according to which we know absolutely nothing, as was defended by the ancient New Academic school of philosophy. We assume that we know something about the world. However, we must consider the question of the relation between our minds and the world as it appears to us.

Historically, many philosophers have adopted one version or another of *Idealism*, the view that everything in reality is in some sense fundamentally mental, consisting entirely of minds and their dependent contents. Philosophers who have embraced some form of Idealism include George Berkeley, G.W. Leibniz, Immanuel Kant, J.G. Fichte, F.W.G. von Schelling, G.W.F. Hegel, F.H. Bradley, and Alfred North Whitehead. In modern times

*The Atlas of Reality: A Comprehensive Guide to Metaphysics*, First Edition.
Robert C. Koons and Timothy H. Pickavance.

(since Descartes revived the problem of skepticism), many philosophers have been challenged by the need to prove the existence of *the external world*, of everything beyond one's own mind. Modern philosophy has thus been haunted by the specter of *Solipsism*, the view that only I and my experiences exist. At an even further extreme, George Santayana (1955) proposed (as a hypothesis) *Solipsism of the present moment*, the theory that everything that was, is, or will be is wholly contained in the contents of my own *present* state of consciousness. On this view, one must deny the reality even of one's own past or future.

At times, the threat of Solipsism has been taken as a challenge to the very possibility of metaphysics, as traditionally understood. How can one hope to describe reality in its most fundamental features if we can be certain of nothing beyond our own minds? It is better, however, to think instead of Idealism and Solipsism as metaphysical theories in their own right, to be evaluated just as we would evaluate other metaphysical theories.

We must make a sharp distinction between radical *skepticism*, which is an epistemological thesis, and Idealism or Solipsism, which are metaphysical theses. Radical skeptics deny that we can ever know or be certain about the existence of anything beyond our own minds. Solipsists adopt the metaphysical thesis that nothing exists beyond his or her own mind. One can be a skeptic about the external world without being a Solipsist. One might even think that the best metaphysical theory of the world includes the existence of other minds and of the physical world, while denying that we can know or be certain that this theory is true. In this chapter, we consider Solipsism as a theory about reality, not merely as a thesis about what we can know.

Of course, radical skepticism and Solipsism are connected. If one believes that one knows that the external world exists, then one cannot be a Solipsist. Conversely, if one thinks that one cannot know anything about the external world, then one will have to at least consider Solipsism as an initially attractive option, given its great economy.

## 13.1 Defining the Mental and the External

Idealism is the thesis that everything is mental. This raises the question of what it is to be *mental*.

Presumably, to count as mental in the appropriate sense is to be a thing all of whose intrinsic properties, or perhaps all of whose intrinsic and fundamental or natural properties, are mental. So, it would be helpful to have a clear conception of a *mental property*.

The mental should encompass two kinds of things: minds and the things that are parts, constituents, events or qualities of minds. The former are subjects of ideas, thoughts, experiences, beliefs, attitudes, feelings, and so on, while the latter are those ideas, thoughts, and so on, themselves.

We don't want to define mental properties in terms of those properties the having of which depends somehow on the existence of a mind, since this would make every variety of *theism* (the belief in the existence of a creator God) into a form of Idealism, on the assumption that God is a mind and that everything depends on God for its existence. Instead, we should focus on the nature of mental properties themselves.

We could proceed simply by listing some paradigmatic mental properties and then counting any property as mental if it sufficiently resembles those on the list. However, everything resembles everything, to some degree. To carry out this definitional strategy successfully, we should also have a list of paradigmatic *non*-mental properties, to serve as the contrast class. To be a mental property would be to resemble the first class more than the second. The latter task is difficult, since Idealists may claim that any property we could name is really a mental property. Certainly George Berkeley would have claimed this. Berkeley argued that what we might think of as paradigmatically physical properties, like mass, shape, and volume, are in fact only properties of our sensory ideas, and so covertly mental in character.

Another tack would be to attempt to define *physical property*, and then to identify the non-mental things as things having only physical intrinsic properties. Idealism could then be defined as the thesis that there are no non-mental things. However, this won't quite work, since some Idealists, Berkeley included, will insist that some ideas have only physical properties.

What we really need to do is to define first what it is to be a *mind*, a subject of thought and experience. We could then define a physical thing as something that is neither a mind nor a part, constituent, attribute, or event in the life of any mind.

Some mental properties are clearly *subject-entailing*, that is, they are the sort of properties that only minds or subjects can have. These include such properties as *thinking*, *believing*, *experiencing*, *intending*, and so on. Only minds or subjects can do those things. Now we can just use the sensible qualities themselves as the contrast class for the definition of 'subject-entailing property'. Even Idealists will admit that there is a difference between the subject-entailing properties of the mind itself (*thinking* and so on) and the properties of the ideas or experiences in the mind (including their supposed sensible or even physical qualities).[1]

So, here is a suggested path:

**Def D13.1 Sensible Property.** A *sensible property* is a quality that things appear to have in our sensory experience (like colors, tastes, shapes, textures, etc.).

**Def D13.2 Subject-Entailing Property.** A *subject-entailing property* is a property that is more like *believing*, *thinking*, *intending* or *experiencing* than it is like any sensible property.

**Def D13.3 Mind.** A *mind* is a thing with at least one subject-entailing property, or the property of having the power or potentiality for a subject-entailing property (like the property of *being able to think*) that is essential to it.

**Def D13.4 Mental Thing.** A *mental thing* is either a mind or a part, constituent, attribute or event intrinsic to the life of a mind.

**Def D13.5 Mental Property.** A *mental property* is a property the having of which entails being a mental thing.

**Def D13.6 Wholly Mental Thing.** A *wholly mental thing* is a thing whose only intrinsic or fully natural properties are mental properties.

Idealism can then be the thesis that everything is wholly mental.

**13.1T Idealism.** Every fundamental particular is wholly mental.

We can define a physical property as a perfectly natural *non*-mental property. Idealism entails that no physical property is instantiated, since it entails that all instantiated natural properties are mental properties. (An even stronger version of Idealism would entail that no physical property *could be* instantiated.) We also assume that if there are any instantiated physical properties, they are similar to the properties posited by our best physical scientific theories, including mass, charge, volume, baryon number, spin, shape, velocity, and so on.

## 13.2 Solipsism and Phenomenalism

As we said, Idealists suppose that everything in the world, including what appear to be ordinary physical objects, are in fact either minds or conscious aspects of minds. Solipsism is that form of Idealism that insists that only one mind exists.

**13.1T.1 Solipsism.** Every fundamental particular is wholly mental and part (or attribute or event intrinsic to the life of) of a single subject or mind.

Historically, Idealism and Solipsism were associated with *Phenomenalism*, the view that the things that appear to us in sense experience are in reality only mental phenomena.

**13.2T Phenomenalism.** Everything we are familiar with or that we have knowledge of, including apparently physical objects, is in reality wholly composed of or wholly grounded in wholly mental things.

Phenomenalism would have some plausibility for us if we accepted the Veil of Perception:

**13.3T Veil of Perception.** Wholly mental things are the only possible objects of sense perception.

What is the difference between Phenomenalism and the Veil of Perception? Phenomenalism entails that we know or are familiar with nothing but mental phenomena, while the Veil of Perception entails only that we can *perceive* only mental phenomena. To move from the Veil to Phenomenalism depends upon the assumption that our knowledge is limited to the objects of perception.

Let's try to reconstruct an argument from the Veil of Perception to Phenomenalism, from Phenomenalism to Idealism, and from Idealism to Solipsism. First, suppose we accept the Veil of Perception. In that case, all of our perceptual evidence provides direct evidence only for the Existence of Mental Phenomena, since both the subjects and the objects of sense perception are mental in nature (according to the Veil). If we perceive nothing but mental things, if there is a Veil, Ockham's Razor (PMeth 1) justifies the further claim that there are no non-mental things. Phenomenalism is, therefore, plausible given the Veil of Perception.

The move from Phenomenalism to Idealism also involves Ockham's Razor. If the world as we know it might consist of nothing but mental things, and we know that mental things exist, then Idealism would seem to be the simplest and most economical account of reality. This presupposes, of course, both that mental things exist and that they cannot be reduced to extra-mental things. In addition, we must assume that the existence of any mental thing is metaphysically independent of any non-mental thing. That is, we must assume that the simplest theory about what mental phenomena are does not by itself entail the existence of things other than minds and their experiences.

**Existence of Mental Phenomena.** We are justified in believing that there are mental entities.

**The Irreducibility and Self-Sufficiency of the Mental.** The simplest account of the truthmakers of true propositions about the existence of mental entities does not entail the existence of anything non-mental.

Given Phenomenalism, the Existence of Mental Phenomena, and the Irreducibility and Self-Sufficiency of the Mental, Idealism is the simplest account of the world as we know it. The Existence of Mental Phenomena is hard to deny, as René Descartes pointed out. Even if I am deceived about everything else, there must be something mental going on, namely, my false thoughts. The Irreducibility and Self-Sufficiency of the Mental is far from obvious, in contrast. At this point, the most we can say is that Idealism is prima facie simpler than other theories, since it posits both fewer entities and fewer fundamental kinds of entities. We'll consider in Sections 13.5 and 13.6 some arguments for doubting the Irreducibility and Self-Sufficiency of the Mental.

If we accept Idealism on the grounds of its greater simplicity, similar considerations should push us even further, toward Solipsism. It is obvious that each of us cannot directly perceive the experiences of others, nor is one subject of experience the direct object of the experience of another subject.[2] Hence, we can express the Veil of Perception more forcefully:

**13.3T.1 Veil of Perception (Solipsistic Version).** The only possible objects of sense perception for each subject are the wholly mental parts and properties of that very subject.

Given the Solipsistic Veil, we can make suitable adjustments to the rest of the argument.

**13.2T.1 Phenomenalism (Solipsistic Version).** For each subject of experience, the world as that subject knows it consists only in that very subject and its wholly mental parts and properties.

**Cartesian Justification.** Each subject of experience is justified in believing in its own existence and that of its experiences.

**The Irreducibility and Self-Sufficiency of the Isolated Mind.** The simplest account of the truthmakers of true propositions about the existence of a subject and its mental parts and properties does not entail the existence of anything apart from that subject.

Given these three principles, the most economical theory for each subject to adopt is the one according to which only that subject and its actual, conscious experiences exist. This simplest theory is just Solipsism. In fact, we can go even farther. The simplest theory of all, consistent with the data of Solipsistic Phenomenalism, would be that nothing exists at all except a *bare* subject of experience and the bare states of consciousness experienced by that subject. In other words, the simplest account of the world is one with no hidden facts or causes, one according to which there are no truths except those made true by aspects of the self and its experiences that are completely and explicitly "given" in that subject's conscious life.

In addition, as Santayana (1955) noted, we could press this argument all the way to the Solipsism of the present moment, the conclusion that nothing exists except the present state of one's own consciousness. How can one know that one's memory, even of one's own past states, is reliable, when it is impossible for one to perceive again the past events that one remembers having experienced?

On the other hand, surely one can know more than one's own present state of consciousness. Solipsism of the present moment is an unattractive theory, to say the least, even if we can recognize the initial plausibility of the argument on its behalf. We don't want to find ourselves in the position of the American who once wrote to Bertrand Russell, stating that she was a Solipsist herself, and that she was quite surprised that there weren't many more Solipsists!

Each of the steps that got us to Solipsism can be challenged. We will examine the Veil of Perception in the next section, and then turn in the following sections to a series of direct arguments against the Solipsistic Veil (Section 13.4), Phenomenalism (13.5), and Solipsism itself (13.5).

## 13.3 Theories of Perception

Why believe in the Veil of Perception? On its face it seems incredible, since we seem to perceive many physical, non-mental things, and events. We perceive lakes, trees, dogs, rocks, and flashes of lightning. None of these seems to be mere modifications of our minds.

Thomas Reid (1785) proposed a principle of common sense according to which we are justified in taking our experience and common sense more generally to be reliable. This is in the spirit of the Presumption of Reliable Perception:

**PEpist 4 Reliable Perception Presumption.** It is prima facie plausible to suppose that human perception and memory are reliable.

In addition, recall Appearance of Bodies and Minds:

**PEpist 3 Appearance of Bodies and Minds.** Perception and memory present us with what are apparently distinct physical things, including some embodying apparently distinct minds.

Given these Reidian facts, why believe that we really perceive only events and processes going on within our own minds? We look first at two obviously inadequate arguments for the Veil, namely, Berkeley's inconceivability argument and the argument from causal mediation. Then, we will look at two much more interesting and important arguments, namely, the argument from hallucination and illusion (Section 13.3.1) and the argument from color and other secondary qualities (Section 13.3.2).

TWO INADEQUATE ARGUMENTS FOR THE VEIL We already discussed (Section 12.1.5) Berkeley's argument for his claim that it is inconceivable that something sensible exist that is not being perceived. Berkeley gives a similar argument for the conclusion that it is inconceivable that anything should exist that is not being thought of. The argument begins with a challenge: try to conceive of something that is not being thought of. As soon as you succeed in thinking of something, you also make it true that the thing is being thought of. Hence, you can never conceive of anything that is not being thought of. Since conceivability, like imagination, is a guide to possibility, inconceivability should provide some evidence for impossibility. Therefore, we have some reason to think that nothing can exist without being thought of.

Now suppose we try to conceive of a sensible thing, that is, a thing with sensible qualities like color or shape or texture. Berkeley claimed that the only way to conceive of such a thing is to imagine it, and to imagine a sensible thing is to imagine sensing it. Hence, if one succeeds in conceiving of a sensible thing, one must conceive of it as being sensed by oneself. Therefore, one has good reason to think that it is impossible that any sensible object should exist unsensed. Note that Berkeley's style of argument should lead to an even stronger conclusion: the *necessity* of Solipsism. If one cannot imagine a sensible object without imagining that one is sensing it, then one cannot conceive of a sensible object that isn't being sensed by oneself. Therefore, one should conclude that it is impossible for a sensible object to exist that isn't being sensed *by oneself* (and right now)!

The fallacy of Berkeley's argument lies in his failure to distinguish between features of the *act* of thinking or representing and features of the *object* of that act. If we cannot conceive of an object with a certain feature *F*, that inconceivability does indeed provide evidence for the impossibility of *F*'s. However, we cannot project necessary features of the act of thinking onto the object of thought. Just because one cannot think of an object without thinking of it, this does not make one's thinking of it *a feature of the object* as one thinks of it. When one thinks of a right triangle, the object of one's thought is simply *a right triangle*, not *a right triangle being thought of now by oneself*.

The second inadequate argument for the Veil of Perception appeals to the fact of causal mediation between physical objects and our internal, mental ideas of them. This argument seems to be implicit in the early philosophical tradition of Idealism, beginning with Descartes.

1 We perceive sensible objects *directly*, not indirectly. That is, we do not perceive them by perceiving something else.
2 However, our sensory perception of external or physical objects is *causally mediated* by processes involving the transmission of light, sound waves or other processes, as well as internal neural processes.

3 If an act of perceiving $x$ is causally mediated, then we cannot perceive $x$ directly in that act.
4 Therefore, sensible objects are wholly mental (and not physical).

The error in this argument lies in step 3. The existence of a direct or immediate relation of perception between a subject and an object can be *causally* mediated, in the sense that the first fact can be grounded in or caused by a complex series of causal links. The series of causal links from the surface of the physical object to the mind of the perceiver simply extends the perceiver's power to perceive objects *directly*. This can be seen to be true in two ways.

The first way is phenomenological. When we perceive a causally distal object, like a distant star, we are not aware of the intervening causal links between the star and our mental state. If we were perceiving the star indirectly, we should be aware of our mental state or the state of our retina directly, and aware of the star only be being aware of some proximal stimulus. However, this is just not how our perceptual system works. When one sees a star, one is normally not even aware that one has a retina, much less that it is being stimulated in a certain way. It's even less plausible to suppose that one is aware of one's related brain states. So to perceive something directly does not, phenomenologically, require that there be no causal intermediaries between the object and the perceiver.

The second way is epistemological. As Fred Dretske (1988) makes clear, we are more reliable at tracking *distal* stimuli than we are at *proximate* stimuli. For example, we can reliably locate objects in our environment or identify their colors, but we are not at all reliable at identifying the patterns of stimulation of our retina or our optical nerve. There is a good biological or evolutionary explanation for this: it was essential to the survival of our ancestors that they should be sensitive to the existence and states of objects in their environment, but not at all important that they should know much about the state of their own sensory organs. So, epistemologically, perception seems to relate us directly to the external environment.

### 13.3.1 The argument from hallucination and illusion

One popular argument for the Veil of Perception draws on the phenomena of sensory illusions and hallucinations. When we are having a hallucination, as when Shakespeare's Macbeth hallucinates the presence of a bloody knife floating before him, we seem to be perceiving *something*. In fact, the thing we perceive when we are hallucinating seems to have the very same types of characteristics, like shape, volume, color, and location, that we attribute to the supposedly physical things that we perceive when not hallucinating.

There are also cases of sensory illusion that share this feature with hallucinations. For example, I might think that I am seeing a bear in the dark, when in fact I am reacting only to a pattern of light and shadows. Examples of illusory objects abound in everyday life. We regularly experience reflections, mirages, rainbows, and objects whose shapes and sizes are so distorted by refraction so as to constitute unreal entities. Even something as common as a rainbow could count as an illusory object, since there is in fact nothing located in the relevant part of the sky with any of the color-like features that we seem to see there. In all of these cases of hallucination and illusion, we seem to be perceiving something when there is nothing physical of the right sort to be an object of genuine perception.

**Hallucination.** There are cases in which no non-mental thing is perceived that are introspectively indistinguishable from cases in which some physical object is perceived.

**Def D13.7 Veridical Perception.** *S* perceives *x veridically* only if *x* is in fact as it appears to *S* to be.

Some philosophers have suggested that during hallucinations we perceive such things as properties (like colors and shapes), complexes of properties (like *redness-plus-triangularity*), or propositions or possible states of affairs (like the possible existence of a red triangle). In this context, we want to preserve the word 'perceive' for those cases where the thing perceived is something concrete, like a table or a flash of lightning. Let's use the word 'apprehend' for any relation to more abstract objects, like properties or states of affairs. Thus, we will insist that in hallucination we are not *perceiving* any non-mental thing, even if we are *apprehending* things like properties, propositions or possible states of affairs.

One might suggest that during a hallucination we are perceiving a region of space, which we misperceive as being occupied by a body of some kind. There are several problems with this suggestion. First, in a hallucination one might not be receiving any information from the region of space in which the hallucinated object seems to be occupied, so this would be a problematic case of perceptual knowledge. Second, the hallucinated object might not occupy any definite region of space—when for example, it does not occur with a perception of depth, or when we hallucinate an image in a mirror. Third, a hallucinated object can seem to move through space over time, which is something that a region of space cannot do.

Advocates of the Veil of Perception could argue that their view is the simplest and most elegant explanation of the phenomena of perception and hallucination. Nonetheless, this has to be weighed against the fact that it entails rejecting many common-sense views about what it is we perceive when we are not hallucinating.

Let's try to organize the alternatives to the Veil of Perception logically. *Perceptual Realism* is the denial of the Veil:

**13.3A Perceptual Realism.** It is possible to perceive something other than wholly mental things.

**13.3A.1T Indirect Realism.** It is possible to perceive something other than wholly mental things, but only indirectly, *by virtue of* perceiving wholly mental things.

**13.3A.1A Direct Realism.** It is possible to directly perceive things other than wholly mental things (i.e., not by virtue of perceiving wholly mental things).

Note that we are using the term 'Realism' in two quite different senses in this book. In Chapters 7–10, 'Realism' refers to views that accept universals as real and as grounds for similarity. In this chapter, we use 'Realism' to mean simply the negation of Idealism.

What are direct and indirect perception? It is important not to confuse this distinction with the distinction between direct (perceptual) and indirect (inferential) *evidence*. Direct and indirect perception both provide direct, perceptual evidence of some fact. Indirect perception does not involve any inference. Instead, indirect perception occurs whenever one perceives one thing by way of perceiving another. For example, a rabbit can perceive the approaching hawk by perceiving the hawk's shadow. The rabbit perceives the shadow directly (one might think), the hawk indirectly. Perception via reflection would be another example. A fisherman can perceive a fish by perceiving the disturbance of the water in its wake, as a physicist can perceive an electron by perceiving a trail in a cloud chamber. A child can perceive his mother's presence by hearing her voice. He perceives the sound directly, the source of the sound indirectly. We ordinarily take ourselves to perceive opaque objects by perceiving the colors on the object's facing surface.

Indirect Realists claim that we indirectly perceive all physical phenomena, by virtue of directly perceiving some mental phenomena. Direct Realists insist, to the contrary, that we can sometimes directly perceive physical objects and processes.

Indirect Realism implies Sense Datum Theory. *Sense data*, on this view, are the mental phenomena that we directly perceive. Sense Datum Theory came under heavy fire in the twentieth century (see, for example, Barnes 1945, Austin 1962), but many of the criticisms focused on features that are not essential to Indirect Realism, such as the so-called 'Phenomenal Principle' (see Crane and French 2015).

**Phenomenal Principle.** If something appears to someone to be *F*, then there must be something that really is *F*.

This implies that our perception of sense data is infallible. We can't perceive a sense datum to be *F* unless it really is *F*. Otherwise, we risk an infinite regress by making our perception of sense data indirect. However, Indirect Realists need not accept the Phenomenal Principle. To resolve the problem of hallucination, all the Indirect Realist needs is the much weaker Exportation Principle:

**Phenomenal Exportation Principle.** If it appears to someone that something is *F*, then there is something that *appears* to be *F* (to him or her).[3]

For similar reasons, Indirect Realists needn't assume that every feature of every sense datum is fully manifest to its subject. For example, a sense datum might have exactly 17 spots, even though the subject perceiving the sense datum isn't aware of the number of spots.

Is there any reason to accept even something as weak as the Phenomenal Exportation Principle? Why should we think that, just because I seem to be perceiving something, that there is anything at all that I am perceiving? Why couldn't my perceptual states sometimes be completely misleading? The most plausible response to this objection would be to appeal to the Principle of Sensory Error Minimization:

**PEpist 4.1 Sensory Error Minimization.** Other things being equal, prefer a theory that posits the fewest and least severe sensory errors to human subjects.

We have epistemological grounds for adopting such a principle. If we don't, we will have no reason to give the evidence of the senses weight in our choice of scientific or metaphysical theories. The Phenomenal Principle could be defended by an appeal to the Principle of Sensory Error Minimization, although this appeal would be complicated by the fact that we have to concede that our perceptual states are in some kind of error in the case of hallucinations and other illusions. However, there is no such complication in the case of the Phenomenal Exportation Principle. We don't have to concede that our perceptions are completely unreliable in the case of hallucination: only that they are unreliable as guides to the external, physical world. The Phenomenal Exportation Principle enables us to salvage some element of correctness within every case of hallucination, something that Sensory Error Minimization urges us to do. This provides some support for Indirect Realism.

The Perceptual Realist alternative to Indirect Realism is Direct Realism. There are two varieties of Direct Realism, namely, Perceptual Dualism and Unitary Direct Realism. They differ on what happens in hallucination. Perceptual Dualists claim that in a hallucination we perceive some real mental entity, while Unitary Direct Realists deny that we perceive any real thing in a hallucination.

**13.3A.1A.1T Perceptual Dualism.** It is possible to directly perceive non-mental (physical) objects, as in veridical perception. It is also possible to be in states (hallucinatory states) in which one directly perceives a wholly mental object but perceives no physical object, and some hallucinatory states are introspectively indistinguishable from some veridical perceptions.

**13.3A.1A.1A Unitary Direct Realism.** Cases in which existing physical things are directly perceived are introspectively indistinguishable from cases in which no existing thing is directly perceived.

Unitary Direct Realism in turn comes in two varieties, Meinongian and Non-Meinongian. The Non-Meinongian version of Unitary Direct Realism is *Intentionalism*.

**13.3A.1A.1A.1T Meinongian Direct Realism.** It is possible to directly perceive both existing and non-existing physical objects, and some cases of the former are introspectively indistinguishable from the latter.

**13.3A.1A.1A.1T Intentionalism.** It is possible to directly perceive existing physical objects, and there are possible states introspectively indistinguishable from these in which one perceives nothing whatsoever (whether mental or physical, existing or non-existing).

Intentionalists suppose that direct, veridical perception of physical objects can be analyzed into two factors. First, the perceiver must be in a sensory state that represents some physical object *O*'s being a certain way. Second, that physical object *O* must be that way in reality and *O*'s being that way must have caused the sensory state in the appropriate way. Thus, whether or not a given sensory state is a case of veridical perception depends ultimately on the causal connections between that sensory state and the external world.

Some versions of Intentionalism, including the adverbial theory of perception (Ducasse 1942, Chisholm 1957), the belief or propensity-to-believe theory (Anscombe 1965), and especially mental-representation Intentionalism (Tye 2000) became dominant in the last part of the twentieth century. A perceptual state possesses some kind of *content*, in much the way that a sentence or a map has content, that is, meaning or significance. What are present in the mind, what are introspectible, are symbols or representations of physical objects akin to proper names or mug shots. These internal symbols are ordinarily the representations of some unique physical object, that object whose characteristics cause the representation in the appropriate way. In the case of hallucination, this normal process goes awry, resulting in something like an *empty* or denotationless proper name. Suppose, for example, that unbeknownst to one there never was anyone named 'Homer' (the Greek poet). One could nonetheless use the name 'Homer' as though it signified some individual. One is not really referring to anyone in using the name, but it seems to oneself that one is. One can't tell the difference between using an empty and a non-empty name merely by introspecting the psychological processes involved in generating these uses. In the same way, a hallucination corresponds to the occurrence of an empty symbol within the sensory part of one's mind.

*13.3.1.1 The solipsistic veil of perception.* The argument from hallucination proceeds by process of elimination. We must first be convinced that we should be unhappy with Indirect Realism, Perceptual Dualism, Meinongian Realism, and Intentionalism. Then we are left with the Solipsistic Veil of Perception, by process of elimination, as the only plausible account of hallucination and non-veridical perception generally.

As we've said, the Solipsistic Veil of Perception runs strongly against common sense, as well as against contemporary science. Consequently, we should embrace the Veil only if there are irrefutable objections to the other four accounts. If any one of the alternatives is defensible, the Veil should be rejected.

*13.3.1.2 Indirect realism.* Indirect Realism was most famously defended by John Locke in his *Essay Concerning Human Understanding* (Locke 1690/1979). Locke's view subsequently came under severe criticism from Idealists like Berkeley. Here we will focus only on those criticisms that build on the problem of hallucination.

The strongest challenge to Indirect Realism in this context involves another appeal to Ockham's Razor. Why suppose that we even indirectly perceive physical objects when we never directly perceive them? The simplest account of perception is one in which all perception is direct. We could simply deny that indirect perception occurs, without any fear that this denial could be contradicted by an act of direct perception. After all, it would seem to be impossible to directly perceive that one is indirectly perceiving something.

There does seem to be some sense in which the fact being perceived is present in the very content of an act of direct perception, in contrast to cases of merely indirect perception. When one directly perceives a red triangle, *redness* and *triangularity* are somehow *given* or *present* to one, whereas when one indirectly perceives a hawk or electron, all that is given in experience is the thing (the shadow, the cloud trail) perceived directly. It seems at least initially plausible to think that only cases of direct perception license beliefs without the need for further evidential support.

This line of thought has been resisted, most notably by Thomas Reid and those who have followed him in embracing a kind of common-sense philosophy. Reid (1785) argued that any belief that we form naturally is prima facie justified, even if the belief is formed on the basis of indirect perception.

*13.3.1.3 Intentionalism.* Intentionalism is somewhat implausible on its face, since there seems to be something before us when we hallucinate. This *thing* that presents itself to us in hallucination can persist, move, and undergo other changes. Nonetheless, many philosophers have argued that this impression that we are seeing something when we hallucinate is itself illusory. Perhaps we are presented with an incomplete content, something like a sentence or diagram with an element that normally signifies a unique object but which, in abnormal cases like hallucination, fails to signify anything at all.

Critics of Intentionalism, from G.E. Moore (1903b/1993) to Frank Jackson (1977),[4] have argued that the theory overlooks the obvious act-object structure of perceptual experience. When Macbeth hallucinates the knife, there is *something* that is the object of his sensory awareness, something apparently knife-shaped and blood-red. The hallucinatory experience isn't anything like the experience of unwittingly using a name that happens to lack denotation. Whenever one uses a name, one is well aware of the fact that one's consciousness includes something that merely stands for the absent name-bearer. A hallucinated knife is not some merely hypothetical object supposedly related to a consciously apprehended symbol. Any such mere symbol would be something that is obviously not knife-shaped or blood-red, whereas there is in fact an object of awareness that itself *appears to be* knife-shaped (not merely to be something that is supposed *to represent* a knife-shaped object).

When one hallucinates visually, one seems to be perceiving something with a certain color and shape. This gives us good grounds, at least prima facie, for thinking that there really is something with that color and shape, or something close to them, in our environment. We may have good grounds for disbelieving that there is something physical with that color and shape, but this counter-evidence does not undermine our grounds, based on Sensory Error Minimization, for thinking that there is nonetheless *something* that we are perceiving, unless we assume that we cannot perceive anything non-physical. But hallucinatory and dream experiences give us some reason to think that we do perceive non-physical things.

In addition, it is possible via hallucination to gain real knowledge about perceivable objects, as Adam Pautz (2007) has noted. Suppose that one has never seen anything red or triangular, and that one has a hallucination of a red triangle. On the basis of the hallucination, one could learn that *red* is more similar to *yellow* than it is to *green* or that there is a plane figure with exactly three sides. One can learn that it is possible for something to be both red and triangular, and one can become acquainted with what such an object looks like. If such a case of hallucinatory learning is possible (as it seems it is), then the mere presence in one's mind of a symbol of *redness* or of *triangularity* was not in fact the ground of the knowledge one gained from hallucination, since one had such symbols without knowing what one learned about *redness* or *triangularity* by actually having the sensory hallucination. Nor was it enough for one to apprehend the properties of *triangularity* or *redness* in a purely intellectual way. In addition to that intellectual apprehension, the hallucination can add new information in an analogue and non-conceptual form.

Intentionalists can respond by appealing to the Ducasse-Chisholm adverbial theory of perception. When one perceives or hallucinates something red, one is in a state of *being-appeared-to-redly*. Nothing need actually be red, nor need there be an object that even appears to be red. Instead, it is one's sensory experiencing that is in a state somehow related to *redness*, perhaps by representing red things or by being the state that normally occurs when veridically perceiving red things or a state that justifies or prompts in a certain way the belief that something is red.

Of course, our sensory representations have a kind of *object structure*: We are not just appeared to *redly*; instead we are appeared to *red-triangularly*, as if we were seeing something red and triangular. In addition, our sensory appearances seem to have a kind of relative location in space. Macbeth's hallucination of the knife is just to the right of a table. Thus, our appearances have a kind of logical and quantificational structure: the appearance of *something* red *and* triangular *and* to the left of *something* green *and* circular. In the end, these object-like nodes in the structure of appearance must either be objects that appear to us in certain ways or mere symbols or signs that represent things as being certain ways in relation to us and to one another.

To make sense of this dispute, we will have to distinguish between conceptual and non-conceptual (or *digital* and *analogue*) representations. A conceptual or digital representation need not contain something the same or similar properties to the thing represented. Digital storage devices, for example, like computer memory, contain representations that share qualitatively very little in common with the things represented. On the other hand, a non-conceptual or analogue representation, like a map or picture, actually contains something with the same or similar properties to the thing being represented. A map represents a triangular region by containing something triangular or nearly so. It seems that one can learn from a hallucination precisely because it is an analogue representation. In a hallucination one perceives in a sensory way something that is itself red and triangular (or close to being so), in a way that no mere representation or content could be. If hallucinations do involve analogue representations, then while hallucinating we must be aware of something red and triangular. Since there is nothing physical that is red and triangular in the right location, we must be aware of some mental object.

However, Intentionalists have a plausible reply. Analogue representation does not require that the very same properties be realized in the representational medium and the represented object. All that is required is that there be a systematic, one-to-one mapping, an *isomorphism*, between the properties possessed by the representational objects and those possessed by the represented objects. For example, we can use colors to represent variations in temperature or length to represent acoustic volume. In the same way, it could be that there is nothing red or knife-shaped in our representational system when hallucinating a knife, but only a pattern of neural firings that belongs to a system of patterns isomorphic to the shape and color of external objects. It may be that in hallucination we are in some fashion *aware of* that pattern of stimulations (giving rise to the illusion of an object of perception), while in fact *seeing* absolutely nothing.

In fact, we could use the distinction between apprehending and perceiving that we introduced earlier. When we are having a hallucination or are subject to an illusion, we are not perceiving anything, but we may be *apprehending* a number of abstract objects, including properties and possibilities. The internal representations may fail to correspond to any existing physical object in the environment, but they may still correspond to certain possible situations. By having those internal representations we can gain knowledge of those possibilities. We can learn from what we apprehend even though we do not perceive anything.

Before leaving Intentionalism, we should note that John McDowell (1994) has objected to Intentionalism on the grounds that it, like the Veil of Perception, alienates our minds from the natural world around us. McDowell worries that by placing images or representations in the mind, Intentionalists concede too much to defenders of the Veil. When we perceive something, our minds stand in a direct relation to some external object, a relation not mediated by any supposed internal counterpart. Presently, we consider McDowell's favored solution to the problem of hallucination, Perceptual Dualism.

*13.3.1.4 Perceptual dualism.* The term 'disjunctivism' has been used for a variety of positions in the philosophy of perception since the 1980s. Our term, 'Perceptual Dualism', doesn't fit exactly to this history, but 'disjunctivist' seems apt as a description of Perceptual Dualism. Roughly, disjunctivism is the idea that hallucination is a fundamentally different kind of mental state than veridical perception. Perceptual Dualism represents a hybrid position, one agreeing with the Veil of Perception and Indirect Realism in the case of hallucination, and with Unitary Direct Realism in the case of veridical perception. In hallucination, on this view, we perceive only a mental object, and in veridical perception, we directly perceive some extra-mental, physical object.

Perceptual Dualists face a serious problem, however, concerning the location of the purely mental object that we are supposed to perceive when hallucinating. Macbeth's bloody knife seems to be located in space. It is hard to see how Macbeth could perceive the knife as located in front of him if it is a completely unlocated, extra-spatial object. If we are going to suppose that we ordinarily perceive something in hallucinations, then we should assume that our perceptions of these mental objects are mostly accurate.

Perceptual Dualists have several options here. First, they could simply concede that sense data are located "out there" in ordinary space. The experienced parts or properties of one's mind would then, or at least could, be scattered over a fairly large region of space. Suppose, for example, that one has a hallucination of a supernova in the sky. the sense data perceived in such a case would be scattered radically over a huge spatiotemporal region.

Alternatively, Perceptual Dualists could suppose that one perceives the location of objects in two different but coordinated spaces, a private mental space and a public physical or physical-cum-mental space (see Davis 2014). Thus, sense data and physical objects do in fact exemplify different sets of spatial properties, but we are unable to tell by introspection whether we are perceiving an object as being located in private or in public space. The two sets of properties are themselves indistinguishable by introspection. Nonetheless, in cases of veridical perception, we perceive the physical locations of physical objects directly.

To avoid falling into a simple Indirect Realism, Perceptual Dualists must suppose that physical things are located both in public space and in the private spaces of minds that are perceiving them. There is some plausibility for this. For example, the sofa is located to the left of the chair in RCK's private space, but not in objective space (where there is no left or right).

This sort of Perceptual Dualism seems to provide a viable alternative to the Veil of Perception as an account of hallucination. Its main drawback lies in its complexity, when compared with its main competitors, namely, Direct Realism and the Veil. Perceptual Dualists have to propose that two quite different processes are taking place in cases of hallucination and veridical perception, in spite of the fact that they seem, introspectively, to be quite similar.

*13.3.1.5 Meinongian direct realism.* Meinongian Direct Realists agree with Intentionalists that we don't perceive any real object in cases of hallucination. However, Meinongian Direct Realists insist that we do perceive something in hallucination: something unreal or non-existent. Thus, Meinongian Direct Realists do not need to deny the act-object structure of perceptual states. There is always something perceived in each perceptual state. Nor do they need to posit purely mental entities to serve as the objects of hallucination. Instead, they can insist that all perceptual states have physical, extra-mental objects. These objects exist in the case of veridical perception but do not exist in the case of hallucination.

We considered some of the pros and cons of Meinongianism (12.1A.1A) in Chapter 12. Here we note briefly a worry about the adequacy of this account as a theory of hallucination.

Assuming that we adopt Meinongianism without the Characterization Principle (as Graham Priest (2005) has advocated), we cannot say that perceived non-existent objects *really* have the sensible qualities that they seem to have. So, we cannot combine Meinongian Direct Realism with the Phenomenal Principle. However, we do get the Phenomenal Exportation Principle: if it appears that something is *F*, then something appears to be *F*. Consequently, a single non-existent thing can appear to move or change over time. So, there is some advantage, in terms of Sensory Error Minimization, of Meinongian Direct Realism over Intentionalism.

We want veridical perception to give us evidence for the existence of the things we are perceiving. We might even say that in veridical perception we perceive the *existence* of physical objects. What happens then in cases of hallucination? Do we perceive a non-existent existence? Does Macbeth perceive the existence of the bloody knife, an existence that happens to be unreal?

There is a danger here of falling into an infinite regress, moving from existence, existing existence, existing exististing-existence, and so on. Such an infinite regress would seem to lead to a global skepticism. I know that I exist, but do I know that my existence is an existing existence? Might it be a non-existing existence, like that of Macbeth's knife? Even if I know that my existence itself exists, can I know that the existence of my existence really exists? It seems that I would have to know an infinite number of things in order to know that any one thing really exists.

Suppose, then, that Meinongian Direct Realists reject all non-existent existences. Then there is a problem about supplying a fully adequate object for hallucinatory states. If Macbeth doesn't perceive the *existence* of the bloody knife (there being no such existence), won't he be able to notice by introspection its absence from his perceptual act? Meinongian Direct Realists should respond by denying that acts of existence are *ever* objects of perception. We perceive objects, but we don't perceive their existence. If they make this move, though, what is the connection between perceiving something and knowing that it exists? We seem to have lost our perceptual justification for believing that physical objects exist.

Meinongian Direct Realists could respond that every act of perceiving something *prima facie* justifies us in believing in the existence of the perceived object. It is normal for us to perceive an existing object. Perceiving the non-existent (as opposed to thinking about or hoping for the non-existent) is abnormal and exceptional.

Still, this leaves Meinongian Direct Realists with an odd dichotomy between existence and other properties.

UPSHOT OF THE ARGUMENT FROM HALLUCINATION As we argued at the outset, the Veil of Perception can only win by eliminating all of the alternative explanations of hallucination and illusion. This has proved difficult. Indirect Realism is the most vulnerable of the alternatives because it concedes so much to the Veil, by admitting that we always perceive wholly mental things. One is left wondering why one should assume that we sometimes also perceive physical things. A simple appeal to Reidian common sense would seem to be the best response on behalf of the Indirect Realists.

Intentionalism is the most popular alternative, but it does seem vulnerable to a challenge based on Sensory Error Minimization. Intentionalism implies that we get absolutely nothing right about the actual world when we experience a hallucination, while the alternative explanations attribute significantly less error to us in such cases.

We have, however, found little fault in either Perceptual Dualism or Meinongian Direct Realism. Of course, Actualists will object to Meinongian non-existent objects, but they could embrace the mental entities required by Perceptual Dualism.

### 13.3.2 The argument from colors

In the seventeenth century, scientists discovered that colors and other *secondary qualities*, like smells, tastes, sounds, and felt textures or temperatures, were not among the essential characteristics of physically fundamental entities, as ancient Greeks like Democritus had long suspected. Only the primary qualities, including position and velocity, volume, shape, and relative orientation, along with theoretical properties like mass and charge, are needed in a complete physical description of the world. The Idealist philosopher George Berkeley realized very quickly that these scientific discoveries provided a new basis for the Veil of Perception.

The argument from colors is an argument for the non-solipsistic version of the Veil of Perception, unlike the argument from hallucinations. The argument gives one no reason to suppose that the only sensible objects are parts of one's mind, although it may give one reason to suppose that they are all parts of some mind or other.

A simple version of Berkeley's argument for the Veil of Perception goes as follows:

1 Everything we can perceive has or can have color (or other secondary qualities).
2 Nothing physical has or can have color (or other secondary qualities).
3 Necessarily, everything we perceive is either physical or wholly mental.
4 Therefore, we cannot perceive physical objects, and everything perceivable must be wholly mental.

Anti-phenomenalists might resist this argument by supposing that although physical objects in fact lack color, they might have been colored, for example, had the laws of physics been different. A second argument for the Veil would appeal to the natural assumption that our sensory perception of the world is not radically and systematically in error:

1 We typically perceive objects as having color.
2 Nothing physical has color.
3 Our perceptions are not typically in error.
4 So, we don't typically perceive physical things. (From 1–3)
5 If we perceive physical things at all, then we typically perceive physical things.
6 So, we never perceive physical things.
7 Everything we perceive is either physical or wholly mental.
8 Therefore, we perceive only wholly mental things.

After all, if we discovered that our sense perception of color is constantly in error, this would provide us with grounds for undermining our confidence in any of the information deriving from the senses, leading to a global form of skepticism.

ARGUMENTS AGAINST PHYSICAL COLOR The crux of both of these arguments is premise 2, the claim that no physical entity has color. Many philosophers suppose that the only properties that physical objects have are physical properties. Let's assume for the moment that this is so. That is, assume the Pure Physicality of Physical Things:

**Pure Physicality of Physical Things (PPPT).** If a physical thing *x* has an intrinsic property *P*, then *P* is a physical property.

Assuming the Pure Physicality of Physical Things (PPPT), we can prove that no physical thing is colored simply by proving that color is not a physical property. We have assumed that physical properties are (1) non-mental (i.e., instantiating one does not entail being a mind or part of a mind), (2) natural, and (3) similar to the properties posited by modern physics. We could define a physical thing as something with at least one intrinsic physical property. PPPT insists that anything with a physical property must have only physical properties. As we shall see, there are some views that deny this, including Panpsychism, according to which every physical thing is conscious.

What is it to be a physical property? Why think color isn't physical? Let's start by defining a *core physical property*.

**Def D13.8 Core Physical Property.** *P* is a *core physical property* if and only if *P* is (1) fundamental or perfectly natural, (2) causally efficacious, and (3) non-mental.

We're going to assume that core physical properties can be mapped one-to-one onto the basic quantitative and qualitative predicates of our current physical theory, at least to a fairly good approximation. That is, we are going to assume that core physical properties are something like quantities of mass or charge, baryon number, volume, and so on. This is the Physics Carves Nature at the Joints Hypothesis:

**Physics Carves Nature at the Joints (Joints Hypothesis).** The core physical properties are something like the basic properties posited by modern science: mass, charge, baryon number, and so on.

We can now define a physical property in terms of core physical properties:

**Def D13.9 Physical Property.** *P* is a *physical property* if and only if *P* is either:

(1) a core physical property,
(2) a conjunction of core physical properties,
(3) a disjunction of conjunctions of core physical properties, or
(4) a power or disposition whose defining stimuli and responses satisfy one of conditions (1) through (3).

We will not assume that the conjunctions, disjunctions, or definitions of dispositions involved in cases (2)–(4) are finite. Infinitely long real definitions are permitted.

We are now in a position to consider whether colors and other secondary qualities are physical properties.

Here are six observations about color, developed in their most sophisticated form by Adam Pautz (2006), which can be used in an argument for the non-physicality of color.[5]

1. No one-to-one correspondence with fundamental physical properties. The fundamental properties of physical bodies are those identified by theoretical particle physics, properties having to do with mass, energy, charge, and so on. There is no simple or unitary physical property of this kind that corresponds one-to-one with any color. However, color is a simple and unitary property of colored things.

2. Unique patterns of similarity and dissimilarity. Our color perception, if it is at all accurate, reveals a system of similarities and dissimilarities among the things we perceive. We perceive, for example, all blue things as genuinely similar to one another in color. However, physics and the physiology of color perception reveal that the physical objects that we perceive as blue have no natural similarity with one another. A bewildering variety of physical surfaces can appear to have exactly the same hue.

3. Features of colors without physical counterparts. There are features of color, including the *binarity* of binary colors (like orange and purple) that correspond to the way in which color properties are processed by the brain, not to any way the perceived physical object is in itself. Orange things seem to have a property that is somehow a blend of red and yellow. Red and yellow things, in contrast, seem to have a pure or unitary quality, one that is constituted by no such blending. This apparent contrast between things with binary and things with unitary color corresponds to no consistent difference at the physical level. It is to be explained, instead, by the details of color-vision processing in the human brain, namely, the fact that signals from the optic nerve are directed into two separate channels, a yellow-blue channel and a red-green channel. Binary colors correspond to the simultaneous stimulation of both channels, and unitary colors to the stimulation

of one and not the other. The difference has nothing to do with any duality or lack of duality in the objects themselves.

4. Variation in the perception of color within a species. Perfectly normal observers in normal circumstances do not perfectly agree in their assignment of colors to perceived objects. The human sexes (male and female), for example, vary in their average perception of color. If color were a property of physical objects, then normal observers in normal circumstances would perceive exactly the same color, and yet they do not.

5. Variation in the perception of color across species. Perfectly normal organisms in different species do not perceive the same colors because their eyes are sensitive to different ranges of the electromagnetic spectrum and (more importantly) because their brains process visual information in radically different ways. Pigeons, for example, perceive objects as colored that human beings perceive as gray, lacking hue. Who is right? Physicalists about color have only three options for explaining this variation, assuming that our color perception is mostly accurate: (1) human beings are the only species with correct color perception, (2) physical objects really have a very large number (perhaps an infinitely large number) of distinct colors simultaneously (one corresponding to each possible species), or (3) colors are powers or dispositions to cause a variety of perceptual states in different species. We'll discuss option (3) below. Option (1) is problematic, since there is no reason to suppose that just one species gets colors right, and even less reason to think that our species is so favored.

Option (2) is contrary to the intuition that nothing can have two distinct colors at the same time, much less an infinite number. In addition, it seems plausible that colors can be identified with certain mixtures of "colorish" qualities. For example, orange is a roughly equal mixture of *reddishness* and *yellowishness*. Yellow, in contrast, involves the exclusive presence of *yellowishness*. If any other kind of "ishness" were present, the thing would no longer be yellow. Yet, if there are "ishnesses" corresponding to other species, these will almost always be present in cases of apparently yellow surfaces.[6] If we don't perceive these other ishnesses, we will be systematically misperceiving the actual colors of things.

6. Phenomenologically, color is natural, intrinsic, and categorical. If we consider how color appears to us, it seems to be a natural property, the basis of real similarities, and not a disjunctive or gerrymandered property (like *grue*). Moreover, the color of colored objects seems to be intrinsic to them. Colored objects seem to be colored in and of themselves. Color does not seem to be relational in nature. Finally, color seems to be categorical, a property that concerns simply how things actually are, without reference to how they would be or what they would do under merely hypothetical suppositions. Given Sensory Error Minimization, we should prefer a theory that respects these apparent facts about color.

CAN COLOR BE PHYSICAL? FOUR CASES In light of these six observations, can color be a physical property? There are four cases to consider. Colors might be core physical properties, conjunctions of core physical properties, disjunctions (of conjunctions) of core physical properties, or powers.

1. Are colors core physical properties? Colors don't stand in a one-to-one relationship to the basic vocabulary of physics, so they cannot be core physical properties, given the Joints Hypothesis. But why accept the Joints Hypothesis? Why not think that colors and other secondary qualities are causally powerful, non-physical properties? The main reason for this is an assumption about the causal closure of the physical world. There is some inductive, empirical support for the hypothesis that only core physical properties have fundamental causal powers. There is also a special reason in this case. It doesn't seem likely that ordinary physical objects, including inanimate objects, have any non-physical powers. The facts about light reflection and sound-wave generation, for example, seem to be completely explainable in terms of core physical properties. It just isn't plausible that colors and other secondary qualities have any additional causal *oomph* of their own.

2. Are colors conjunctions of core physical properties? This view is unlikely because colors can be realized in an extremely large number of ways, with no physical similarities among the realizers (see observation #2 above). In addition, this view cannot explain the appearance of binary colors, since there would be no common denominator between the conjunctions defining orange and its two constituent colors, red and yellow (see observation #3).

3. Are colors disjunctions? This view is also unlikely, because disjunctive properties are not causally relevant, and we have already seen that we have good reason for thinking that colors are powerful. In addition, disjunctive properties are not natural, and they cannot ground the kind of systematic similarities and dissimilarities that we observe among the colors (observation #2). Moreover, if colors were disjunctions, there would have to be some basis for deciding which physical-property conjunctions to include in the disjunction. If the list is not to be arbitrary, we would have to treat color as a *response-dependent* property. That is, we would have to include a conjunction on the list for color *C* if and only if such a property conjunction is perceived as having color *C* by *normal* observers in *standard* conditions. However, such a definition cannot accommodate the variability of color sensitivity (see observations #4 and #5), since the inclusion or exclusion of any disjunct would have to be determined by whether a physical surface of that kind produces the right color sensation in a normal observer in standard circumstances.

4. Are colors powers? This view has been a popular response by anti-Phenomenalists since John Locke. However, it is subject to two devastating objections. First, there is a phenomenological objection. Colored objects don't appear in our sensory experience to be powers or disposition. When I see something as red, I certainly don't see it as something having the power to cause red sensations in me and subjects like me. I see it as having some intrinsic, categorical, and non-relational property (observation #6). Further, it is plausible to suppose that 'red' is ambiguous, sometimes referring to the intrinsic property appearing to us in visual experience, sometimes to the power in things to produce visual experiences of that kind in standard circumstances. It is the first kind of color, the intrinsic property, that we perceive. We do not perceive the power. The defender of the Veil need only argue that this perceived, intrinsic property is non-physical. Second, there is a circularity objection. What appearance is it that the colored objects have the power or disposition to cause? If redness is just a property to cause certain sensations,

then the sensations they cause cannot themselves be red. One is left wondering what sort of sensations red objects cause essentially.

The idea of color as a power or disposition of physical objects only makes sense if there are, in addition to the physical colors, a set of *phenomenal* colors that characterize the sensations that colored objects are supposed to cause. Let's distinguish 'phy-red', the physical power to cause certain sensations, from 'phe-red', a certain intrinsic type of sensation. We could define 'phy-red' as the power to cause phe-red sensations in normal observers in normal circumstances without circularity. However, it seems clear that we do not directly perceive objects to be phy-red. Phy-red is not how objects appear to us to be in sense experience. If anything, objects appear to be phe-red.

This sort of response is also available to Indirect Realists. We could perceive mental objects to be phenomenally-red directly, thereby perceiving indirectly some physical objects to be phy-red. Indirect Realists can attribute such powers to physical objects, and we could call these a kind of "color," but they are not what are directly perceived, since they are not what is given in sense experience.

In addition, this power theory of color runs into some difficulties in accounting for the intra-species variability and the extreme context-sensitivity of color perception (see observations #4, #5, and #6). Physical colors will have to be somewhat vague and very context-sensitive.

In summary, we have seen good reasons for thinking that colors are natural (grounding similarities), causally powerful, categorical, and intrinsic. These facts are inconsistent with the physicality of color.

THE UPSHOT FROM THE ARGUMENT FROM COLOR If colors are not physical properties, what are the options? There are four possibilities.

1 Colors are not properties of physical things at all but rather properties of the perceiver's mind or some part of the mind. So, we perceive directly only mental things. This would entail either Indirect Realism or the Veil of Perception. This could be extended to a Double Color Theory. On this view, there are colors of mental objects (the things we directly perceive) and colors of physical objects (that we indirectly perceive or infer), with the physical colors being either unnatural disjunctions or relational dispositions or powers.
2 Color Irrealism. Color Irrealism, defended by Adam Pautz (2006, 2010), is the view that colors are uninstantiated physical properties. Color Irrealism is most naturally combined with a Meinongian version of the Veil of Perception, according to which we only perceive non-existent things. Or, we could adopt an error theory of color perception, according to which we do perceive physical objects, but we systematically misperceive them as colored. This option pays a heavy price in terms of violating Sensory Error Minimization. In addition, it is difficult for Color Irrealists to give an account of how our sensations have the intentional content they do have. They cannot provide a theory of intentional content in terms of causal connection or reliable covariation, since our color sensations are not causally connected to real cases of color, nor do they co-vary with any such cases. There simply aren't any real cases of color. It remains mysterious how our senses connect us to properties that are not, and probably could not, be instantiated in the sort of world we actually inhabit.

3 Emergent Properties. The Joints Hypothesis is false, and colors are core physical properties, after all. They just aren't at all like the properties posited by modern theoretical physics. Suppose that there are, at the macroscopic (or mesoscopic, medium-sized) scale of colored objects and surfaces, *emergent* core physical properties, fully natural properties that are caused by but not reducible to or wholly grounded in the properties of their microscopic parts. The phenomenon of quantum entanglement (which we discussed in connection with Jonathan Schaffer's argument for Monism 11.2A in Section 11.2.4) could provide some support for the possibility of such irreducibly macroscopic properties. We will examine emergent properties in more detail in Section 22.6.

On this view, color perception would give us direct access to some real sensible qualities of physical objects, an access that goes beyond the resources of theoretical science.

Emergent Color Theory is a version of Powerism (4.4A.3) with sicceities (Def 6.3, see Section 6.2).

Emergent Color Theory has the advantage that colors and other sensible qualities are real, intrinsic, categorical, natural properties of physical things. In addition, they would be causally relevant, since, they confer active and passive causal powers on their bearers.

One difficulty with Emergent Color Theory is the threat of redundant causation, with human color perception's being causally overdetermined by colored surfaces and by microphysical phenomena, such as photons and electron shells of atoms. Emergentists can ameliorate this by claiming that the emergent properties of the macroscopic surfaces absorb and co-opt the causal powers of the constituent micro-particles, with the result that the powers of the micro-particles are wholly grounded in the macroscopic qualities.

Another difficulty concerns the relativity of color perception, especially across species (fact #5). They could suppose that the emergent qualities correspond to the constituent colorishnesses (like yellowishness and greenishness), whose causal powers are exercised jointly. This would mean that human color perception is incomplete, insensitive to the colorish qualities perceived by other species. Nonetheless, what we do perceive would be really there (as natural, categorical, causally efficacious properties). Our two-channel perceptual system, for example, might have evolved precisely in order to detect the distinct colorishnesses that have emerged in the environment.

4 Color Neutralism. According to Color Neutralism, the Joints Hypothesis is true but PPPT is false: physical objects have non-physical properties, properties that are themselves core physical properties nor are definable in terms of the core properties of theoretical physics. However, these properties are not mental, either. On this view, colors are neither mental nor physical. That is, colors are not physical in the narrow sense of being definable in terms of the fundamental properties of modern physical theory (in any of the ways 1–4 listed above). They could still be *broadly* physical in the sense of being natural and non-mental (although not necessarily efficacious). We turn presently to a more detailed examination of Color Neutralism.

COLOR NEUTRALISM Color Neutralism comes in two varieties: Brutely Supervenient Qualities and Panqualityism.

*a. Brutely Supervenient Qualities* In Chapter 3, we introduced the idea of supervenience (Def 3.1). We offered this definition:

> **Def D2.6 Weak Supervenience.** A set of properties *A weakly supervenes on* a set of properties *B* if and only if it is impossible for any two possible worlds to agree on which things have which *B*-properties but to disagree about which things have which *A*-properties. That is, two situations which are indiscernible in respect of the *B*-properties *must* also be indiscernible in respect of the *A*-properties.

Let's say that one set of properties weakly supervenes on another *with metaphysical necessity* when the kind of impossibility mentioned in Def 3.1 is metaphysical impossibility (see Chapter 14). Let's say that a case of supervenience is *brute* when it holds with metaphysical necessity but it is a posteriori (not a priori or based on conceptual connections) and not derivable from facts about identity or distinctness.

> **Def D13.10 Conceptually Brute Supervenience.** A set of properties *A brutely supervenes* on a set of properties *B* if and only if *A* weakly supervenes on *B* with metaphysical necessity, and no predication of an *A*-property is a priori deducible from any conjunction of predications of *B*-properties, together with any set of true identities and non-identities.

According to Brutely Supervenient Qualities, colors brutely supervene on the physical properties of things. The connection between a thing's color and its physical constitution is metaphysically necessary, but we couldn't deduce a thing's color from its physical constitution, no matter how familiar one was with the intrinsic nature of that physical constitution.

There are two objections to such Brutely Supervenient Qualities. First, it's not clear that such supervenient qualities can be causally efficacious or even relevant. Jaegwon Kim (1992, 1998, 2001) has argued that the causal relevance of brutely supervenient properties is excluded by the causal completeness of the physical realm. Since physical properties are sufficient to ground all causal interactions, there is simply no room for the causal relevance of brutely supervenient properties. This is clearly true if we assume Powerism or Nomism (4.4A.2) and if we assume that Brutely Supervenient Qualities are less fundamental than their physical bases, since it is the fundamental properties that confer causal powers on things (according to Powerism) or are situated within causal laws (according to Nomism). Kim's exclusion argument is not compelling, however, if we assume Neo-Humeism (4.4T) or Hypotheticalism (4.4A.1), since on those views neither causal powers nor causal laws are fundamental, and so it is less clear that there is anything objectionable about attributing causal power to both physical properties and brutely supervenient ones.

Second, there is a strong modal argument against the supervenience thesis entailed by Brutely Supervenient Qualities. We seem to be able to conceive of a world where color and its physical bases come apart. We can imagine something's having the physical and chemical composition of a typical ripe tomato and yet being blue rather than red. In light of Imagination as Guide to Possibility (**PEpist 1**), we have good reason to think that colors do not supervene on the set of physical properties.

We can avoid this modal objection by supposing that colors did not supervene with *metaphysical* necessity on their physical bases. We might suppose instead that there are some *contingent* causal laws that link colors with physical properties, in such a way that the physical properties cause their bearers to have certain colors. Although this helps with the modal argument, it makes the problem of causal exclusion much worse. It would mean that colors were causally inert by-products of underlying physical processes. This would make it a mystery how we are able to get the actual colors of things right. It's hard to see why the colors we experience would correlate with the real colors of things if those real colors are epiphenomenal, that is, if they make no difference to the causal powers of things.

Thus, the defender of Brutely Supervenient Qualities faces a dilemma: the Scylla of causal exclusion versus the Charybdis of imaginability.

*b. Panqualityism* Panqualityism is a recent proposal by Brian Cutter (2013). Cutter looks back to Bertrand Russell's (1927) structuralist understanding of physical science. According to this type of structuralism, physical science tells us only about the causal powers of physical properties. It doesn't tell us anything about the intrinsic character or quality that corresponds to those properties. Yet we have good reason to suppose that there must be such intrinsic qualities, unknown to us. The physical sciences provide us with an abstract structure of properties, expressed in the form of causal laws. They do not tell us what qualities actually fill the roles specified by this abstract structure.

Structuralism corresponds to Powerism with sicceities. We don't have firsthand experience of or acquaintance with any of the sicceities of the fundamental physical properties. We know them only via a description of their causal and nomological profiles. Cutter, therefore, proposes that every core physical property has a qualitative sicceity, a *micro-quality*, of which we have no acquaintance. However, the colors we perceive are wholly *grounded in* these micro-qualities, in such a way that any subject that was acquainted with both the micro-qualities and the colors would be able to deduce (a priori) which combinations of micro-qualities ground which colors. Cutter's view is Panqualityism.

The possibility of such a strong, conceptual connection between the two rebuts the modal argument against supervenience that seemed to be decisive against Brutely Supervenient Qualities. Since we don't know what the micro-qualities are like, we can't in fact conceive of a case in which the micro-qualities are present and the color is lacking. According to Panqualityism, if we were able to conceive of the micro-qualities, we would be able to see a priori that it is impossible for those micro-qualities to be present and the color to be lacking.

In addition, such a close conceptual connection solves Kim's causal exclusion problem. Colors are intimately connected to grounding micro-qualities, and micro-qualities are causally efficacious, since they are strictly identical to core physical properties. We might plausibly say that an object's being colored *just is* its having the appropriate micro-quality. Since micro-qualities are causally effective, so are colors.

On Panqualityism, colors are *broadly* physical, since they are a priori deducible from core physical properties. However, we cannot so deduce them, since we are not acquainted with the core physical properties in their qualitative aspect.

Panqualityism has a number of advantages. Colors are intrinsic, categorical, natural, and causally relevant, just as they seem to be in our experience. But Panqualityism does face one daunting challenge: it isn't clearly plausible that colors could be a priori deducible from other qualities, even qualities with which we are completely unacquainted. Colors seem to be simple properties, neither conjunctive nor disjunctive, so the a priori deducibility of colors from micro-qualities cannot be grounded in logical relations, the way disjunctions are deducible from a disjunct, or conjuncts from conjunctions. Colors are not identical to micro-qualities, so the deduction cannot proceed by way of the substitution of co-referring terms. We are left with something of a mystery in trying to understand how the a priori deducibility could be explained.

## 13.4 Arguments against Phenomenalism

### 13.4.1 What follows from the Veil of Perception?

As we have seen, the main argument for Idealism depends on the Veil of Perception, which also provides a case for Solipsism. The two arguments for the Veil, from hallucination and from color, are inconclusive. Intentionalism, Meinongian Direct Realism, and Perceptual Dualism, in particular, have not been ruled out by the argument from hallucination, and several versions of Direct Realism, including Color Irrealism, Hidden Conjunct Theory, and Color Neutralism, are able to stand up against the pressure of the argument from color.

Nonetheless, the two arguments we have considered are not without force. Let's consider what would follow from accepting the Veil of Perception. As we mentioned above, there are two versions of the Veil of Perception: the Solipsistic Veil (SVP) and the non-Solipsistic Veil (NSVP). SVP is logically stronger than NSVP, since if each subject cannot perceive anything but parts of that subject's own mind, then no subject can perceive anything that isn't a part of some mind.

Both versions gain some support from the argument from hallucination, while the argument from color supports only the weaker NSVP.

Assuming NSVP, we have three options with regard to the non-mental and the external, namely, Inferred Anti-Idealism about the non-mental (of which more below), Phenomenalism (non-mental objects are wholly grounded in mental objects), and Eliminative Idealism (there are no physical objects at all). Phenomenalism is more moderate than Idealism, since it permits its adherents to admit that physical and external things do exist. Historically, there have been three versions of Phenomenalism, namely, Berkeleyan (Theistic) Phenomenalism, Possibilistic Phenomenalism, and Holistic (Constructive) Phenomenalism.

Berkeleyan Phenomenalism supposes that there is a God who shares our sensory experiences of the world. According to Berkeleyan Phenomenalism, God experiences every sensible object in every sensory mode and from every perspective. In fact, *physical* objects are really nothing but comprehensive bundles of ideas in God's mind. Our sensory perception is a kind of sampling by participation in God's comprehensive sensing of the world. The natures we attribute to physical objects in common sense and in science are merely useful fictions, constituting a kind of instrument for predicting and controlling our future sensations. Berkeleyan Phenomenalism has been summarized in a limerick by Ronald Knox (Reed 1924):

*There was a young man who said, "God*
*Must think it exceedingly odd*
*If he finds that this tree*
*Continues to be*
*When there's no one about in the Quad."*
*REPLY*
*Dear Sir:*
*Your astonishment's odd:*
*I am always about in the Quad.*
*And that's why the tree*
*Will continue to be,*
*Since observed by*
*Yours faithfully,*
*GOD.*

Possibilistic Phenomenalism was proposed by John Stuart Mill (1865/1963) as an attempt to reproduce the advantages of Berkeleyan Phenomenalism without positing God as the universal perceiver. Mill proposed that physical objects are what he called 'permanent possibilities of perception'. We could still think of each physical object as a bundle of sensations, but now not actual sensations in God's mind but merely possible sensations in anyone's mind. Possibilistic Phenomenalism requires Hypotheticalism. There must be metaphysically fundamental conditionals of the form, if perceiver *P* were in condition *C*, *P* would perceive sensation *S*. Since there are no non-mental entities and no actual perceiver of these sensations, such conditionals lack categorical truthmakers.

Holistic (or Constructive) Phenomenalism is best exemplified by Rudolf Carnap's *The Logical Structure of the World* (Carnap 1928/1967) or *Aufbau* (in German). In Carnap's system, individual physical objects are not replaced by bundles of possible sensations. Instead, the whole physical world is replaced in one fell swoop by a single system of sensory conditionals. Carnap supposes that we posit the physical world as the simplest system of objects and states that can, hypothetically, explain, predict, and unify the sensory experiences of minds. On this view, no physical fact is fundamental.

Holistic Phenomenalism runs into some danger of indeterminacy and relativism, since we may disagree about what counts as the best or simplest theory. It is then incompatible with a scientific realism that takes the statements of science to be true, when interpreted at face value.

### 13.4.2 Common-sense and Scientific Realism

Both common sense and modern science attribute natures to physical things. These natures, in turn, give rise to powers and dispositions if we embrace either Powerism or Nomism. However, for Phenomenalists physical objects are mere bundles of ideas or mere bundles of possible ideas or mere fictions. Neither bundles nor fictions are the sort of thing to realize powerful natures. Berkeley was quite willing to embrace this conclusion. On his view, only "spirits", that is God and human souls, are powerful and capable of real action. Therefore, one cost of Phenomenalism is that it puts two attractive views of powers, Powerism and Nomism, beyond the pale. Phenomenalists have to embrace either Strong Hypotheticalism or Neo-Humeism about powers.[7]

Much of modern science involves the positing of unobservable entities, such as microscopic particles, distant astronomical formations or fields of force like gravity and electric charge. Phenomenalists must be anti-realists about all such unobservable entities, since they do not correspond to actual or even possible sensations. It might be possible for Phenomenalists to suppose that such entities are unobservable only to us but are directly observable by God or by some other actual or merely hypothetical creature. However, there is little or no reason for Phenomenalists to make that assumption, since we have not encountered any creatures with the relevant sort of sensory apparatus. Further, it is doubtful that such an apparatus could actually exist. Given the nature of things like quarks, quasars, and magnetic fields, it is plausible that they are not even possibly directly observable.

Phenomenalists ought, then, take an *instrumentalist* view of scientific theories. On this view, scientific theories are all literally false, or at least we have no good reason to suppose them to be true. But they are useful fictions. By pretending they are true, we can better make predictions and obtain control every observable phenomenon.

Such an instrumentalist position faces two major challenges. First, there is the no-miracle argument for scientific realism. Defenders of this argument (first clearly proposed by Hilary Putnam in Putnam 1975: 73) claim that the broad instrumental success of scientific theories in a wide range of applications would be an incredible "miracle" unless those theories were true (or at least approximately true). Scientific realism offers a kind of explanation for the instrumental success of science that instrumentalists cannot offer.

Second, scientific realists (such as Quine 1957: 229, 233) have argued that theoretical science is wholly continuous with common sense. We believe in the unobservable entities proposed by science in exactly the same way and with the same justification that we believe in the many observable but not yet observed entities proposed by the exercise of ordinary common sense. Just as the empirical evidence makes it reasonable to believe in an unobserved burglar, so too can empirical evidence make it reasonable to believe in an unobservable field of force. In fact, if the evidence is strong enough, it is unreasonable not to believe in such unobservable entities. Consequently, the fact that Phenomenalism entails the unreasonableness of such beliefs counts against it.

### 13.4.3 The apparent impossibility of phenomenal reduction

If Phenomenalism is to be plausible, Phenomenalists must provide some sort of account of how facts about perceptions can make true all or nearly all of our ordinary beliefs about physical objects. Carnap attempted just such a reduction of physical theory to truths expressible in a purely phenomenal language (Carnap 1934/1937).

One crucial problem for Phenomenalists, which was noted by Mill, is that there simply aren't enough actual experiences to ground all of the physical truths. As far as we know, there is no one presently perceiving the far side of Pluto, and yet we believe that there are many facts about physical objects in that region of space. Phenomenalists like Mill have attempted to fill this gap by appealing to merely possible or *counterfactual* perceptions, facts about what observers would perceive if they were located in the right place. Carnap attempted to carry out this project. In doing so, he encountered a problem that he was never fully able to overcome: how to specify the counterfactual situations under which the possible perceptions are supposed to happen. If we say that there would be certain sense perceptions if one of us were looking at Pluto from the other side, we are making reference to a matrix of physical facts, including the possible location of an observer in physical space. In order to carry out the reduction successfully, the antecedents of the relevant conditionals (the *if-and-when* parts) would have to specify the counterfactual situations in language that is free of all reference to physical objects (including our bodies or sense organs) or physical space. This is a tall order.

In particular, Phenomenalists would have to reduce the spatiotemporal primary qualities of place, distance, and orientation to the intrinsic perceptual states of possible observers. Instead of saying something like 'If you were on the opposite side of Neptune and looking this way, …' we would have to say instead something like 'If you were to have the sensations that would accompany a trip to the opposite side of Neptune, along with the sensation of turning around by 180°, …' However, it is very far from clear that these two if-clauses would pick out the same possible situations. After all, one might have all of the sensations associated with a trip to Neptune by taking LSD or peyote! What one would then experience would have nothing whatsoever to do with the physical attributes of the farthest planet. The failure of would-be Phenomenalists like Mill and Carnap to solve this sort of problem gives us some reason to think it impossible.

## 13.5 Arguments against Solipsism

Besides the Veil of Perception, arguments for Idealism and Solipsism depend on an appeal to Ockham's Razor. This appeal can succeed only if Idealism or Solipsism provide the simplest overall account of the world, given our evidence. There is a battery of arguments in the history of philosophy that challenge this assumption. These arguments suggest that mental phenomena are dependent on or interdependent with a world of extra-mental physical objects. These arguments can be taken either as providing additional support for Perceptual Realism (and against the Veil of Perception) or as a defense of Inferred Anti-Idealism:

**13.1A.1 Inferred Anti-Idealism.** Non-mental (physical) things exist but cannot be perceived (either directly or indirectly).

### 13.5.1 Appeals to common sense and to perception

To believe in both physical objects and other minds is a matter of common sense. It is also an inseparable presupposition of the methods of modern science, from particle physics to sociology. Many of our ethical, legal, and political practices are grounded on the supposition of the reality of the external world. We can therefore justify the rationality of belief in the external world by appealing both to common sense and to our rational practices.

In addition, since the arguments for the Veil of Perception were unsuccessful, we can justify the appeal to physical objects on the grounds that we seem to perceive them via the senses. The Reliable Perception Presumption (PEpist 4) supports the idea that it is reasonable to assume that things are as they appear to us to be, unless we have strong grounds to think otherwise.

Complex thought depends on language, and language is a social phenomenon. Wittgenstein (1953) argued that, without a real society as a source of linguistic norms, we have no criterion for the correct use of language. Whatever *seems* right will *be* right, especially if we are "speaking" an idiolect of the present moment. We can't know what we are thinking if we have no criterion for the correct use of our language. Thus, Solipsism is self-refuting. The argument goes something like this (following the so-called "communitarian" interpretation):[8]

1 I would not know what I was thinking unless there were a society that provided me with linguistic norms.
2 I do know what I am thinking.
3 Therefore, there exists a society that provides me with linguistic norms.

The main weakness in Wittgenstein's argument is his assumption that thought depends upon language. In particular, Wittgenstein assumes (on the this communitarian interpretation of his argument) that the norms of thought (the correct application of our concepts) always depend on the social norms of language. Wittgenstein spends much of *The Philosophical Investigations* (1953) trying to substantiate this claim, but he doesn't demonstrate that a concept cannot somehow carry within itself some principle of application that determines which uses are correct and incorrect, apart from the judgments of others. In addition, Wittgenstein's argument seems to assume that the social norms embedded in language are necessarily reliable, and so are the final court of appeal for the correctness and incorrectness of our judgments. This seems wrong. It is natural to think that everyone in a language community could err systematically in the application of some concept. For example, it seems possible that everyone wrongly believes that a particular tree is older than every other tree.

### 13.5.2 The Paradigm Case Argument against ubiquitous illusion

Solipsists argue that all of our perceptions are in fact illusions, insofar as they present to us a world of things beyond our minds. However, as J.L. Austin argued (1962), the concept of *illusion* makes sense only if we can also grasp its contrary, *real (veridical) perception*. In order to know that some of our representations are illusory we must be also be able to know that they are *not* genuine perceptions of physical things. However, it is possible to grasp a concept in only two ways: by being able to define it in terms of simpler concepts or by knowing paradigm cases of the concept and its complement. The concepts *perception* and *physical* are not definable. Any attempt to define them would make use of a circle of synonymous concepts, such as *knowledge*, *seeing*, and *hearing* in the case of *perception*, and *material*, *external*, *colored*, and *located* in the case of *physical*. On Solipsism, we have no knowledge of any paradigms of any of these concepts. Hence, we cannot grasp the concept *illusory*, and Solipsism is self-refuting.

Austinians are appealing here to something like the following principle of concept possession:

**PEpist 5 Conceptual Acquaintance.** It is possible to know the truth of a proposition that involves the use of a concept *C* only if either (i) one knows a definition of *C* in terms of other concepts or (ii) one knows a non-empty class of instances of *C* and of instances of non-*C*.

**Definitional Foundationalism.** A system of definitions known by a single person at a time forms a finite, acyclic directed graph (that is, there are no definitional circles or infinite regresses—all definitional "trees" terminate in undefinable "leaves").

The combination of Conceptual Acquaintance and Definitional Foundationalism entails that all of our fundamental, undefinable concepts are concepts we know to apply in certain concrete cases, the *paradigm* cases. Solipsists believe that all concepts of physical things are uninstantiated, that there are in fact no instances of these concepts. This would contradict the combination of Conceptual Acquaintance and Definitional Foundationalism. Solipsists can't have the concepts that they claim are uninstantiated. Similarly, Solipsists believe that they have knowledge of only a single thing, and that thing is mental. This would also contradict the combination of Conceptual Acquaintance and Definitional Foundationalism, since we can't grasp the concept *being mental* without knowing some non-mental things.

Solipsists' claim that there appears to be a physical world in space but there is in fact no such world couldn't be true, since they cannot even conceive of the possibility of such a physical world without knowing some actual paradigms of their physical concepts.

There are other variations on Conceptual Acquaintance that would lead to the same conclusion. For example, consider a second version:

**PEpist 5.1 Conceptual Acquaintance (Second Version).** It is possible to know the truth of a proposition that involves the use of a concept *C* only if either (i) one knows a definition of *C* in terms of other concepts or (ii) one's idea of *C* is connected causally with at least one actual instance of *C* and one actual instance of non-*C*.

These Conceptual Acquaintance principles are far from obviously true, however. It would seem to be possible, for example, to acquire the concept *green* entirely from colorful hallucinations without ever seeing anything green or non-green. Suppose, for example, that we stimulate the optical nerve of a blind person in such a way as to produce a variety of colorful experiences. We could tell the person, 'This is red', 'Now this is green', and so on. It would seem that the blind person might acquire the relevant concepts in advance of any veridical experience of colored surfaces. Also, imagine someone born with color-altering lenses who learns to use color-words correctly (as terms for physical surfaces) but who mislearns what each of the colors really looks like. Suppose, for example, that orange things look red to such a person. He would have the concept *red*, even though he calls the color 'orange' (matching the behavior of normally-sighted people) and even though what he takes to be paradigms of that color and of its complements are not really so.

In response, anti-Solipsists could weaken Conceptual Acquaintance so that it only requires that, in order to have some concept *C*, either oneself *or some other member of one's species* must have been acquainted with both instances and counter-instances of *C*. Even if one has never encountered anything really red but base one's concept *red* on hallucinatory experience, it still must be the case that other members of the human species (including many of one's ancestors) must have used their perceptual capacities to distinguish red things from non-red things How else could our concepts and sensory appearances be connected with real properties?

This argument thus depends ultimately on a kind of causal theory of conceptual content, the theory that concepts have the content they do because of causal interactions between the concept-holder or his ancestors and the relevant properties. Support for such a theory would seem to rely heavily on empirical evidence, especially from cognitive psychology and evolutionary biology, evidence whose validity Solipsists would reject. However, do Solipsists have an equally plausible theory of conceptual content to offer?

Solipsists might attempt to appeal to the Realist-Aristotelian account of intentionality sketched briefly in Section 7.3. On this view, we can think about properties simply by incorporating the corresponding universals into our thoughts. A universal like SQUARENESS would be literally a part of each of our thoughts or apparent perceptions of square things. Solipsists would still owe us some account of why those universals are present in our thoughts and perceptual states. What is responsible for putting them there? Aristotle thought (in *De Anima* Book III) that it was the fact that these universals are instantiated in our natural environment, together with the fact that we are naturally equipped to perceive them by means of our senses, that are jointly responsible for the presence of the universals in our minds. Obviously, Solipsists would not have access to such an account.

## 13.6 Conclusion and Preview

Solipsism and Idealism run up against the obvious fact that we perceive many things beyond our own minds, things that co-exist with us in a world of space and time. The burden of proof is on Solipsists and Idealists, and historically they have sought to carry this burden by appeal to the Veil of Perception that is supposed to lie between our perceptual capacities and the world beyond our minds. We've found plenty of reason to doubt the existence of this Veil. In addition, critics of Solipsism have marshaled a battery of arguments in defense of the common-sense conviction that we ordinarily perceive physical things and other people.

In light of these conclusions, we can now turn to those metaphysical questions that arise from the world as we find it, questions about physical parts and wholes, space and time, possibility and necessity, and cause and effect.

## Notes

1 Even a no-subject Idealist like David Hume will recognize a difference between mental *actions* (like believing or intending) and mental *objects* (like colors and smells). For Hume, a mind is a bundle of mental events and acts. His conception of the self will fit our definition of 'mind' so long as some bundles include actions essentially.

2 This overlooks one possible complication. It could be that there are mental entities (like *sense data*) that are perceived by more than one subject at the same time, even if one subject's act of perceiving is not the object of any other subject's awareness. Nonetheless, Solipsism remains a very economical theory if one has no reliable evidence for the existence or perceptual activity of other subjects.

3 The reader may have noticed that the Phenomenal Exportation Principle is formally very similar to the Barcan formula, a formula that licenses the exportation of the existential quantifier through the possibility operator. The Phenomenal Exportation Principle licenses the same exportation of the existential quantifier through the 'It appears to S that' operator.

4 Jackson later recanted this objection to Intentionalism.

5 For details about the processing of color information in the brain, see Hunt 1982, Hardin 1988/1993, and de Valois and de Valois 1993.

6 Thanks to Richard Lawton Davis for this argument.

7 It might be possible for Berkeleyan Phenomenalists to embrace Nomism. The laws of nature could connect bundles of divine-ideas of various kinds. However, such a position does not seem to be available for Holistic Phenomenalists, since there are no actual entities to play the role of physical objects on those views.

8 See Kenny 1973, Kripke 1982, Malcolm 1989, Boghossian 1989, and Canfield 1996.

# Part V

# Modality

# 14

# Possibility, Necessity, and Actuality: Concretism

## 14.1 Introduction

In the next three chapters, we consider a subject that has been central to philosophy from the time of Parmenides on, namely, modality. The study of modality is the study of possibility, necessity, contingency, and actuality. Ordinary English contains a variety of modal vocabulary. Consider, for example, the following:

(1) THP had oatmeal for breakfast this morning but could have had Cheerios instead.
(2) Had THP's spouse made pancakes and eggs this morning, he would have eaten them instead of oatmeal.
(3) It's not possible to travel from the Earth to Mars in 15 minutes.
(4) Texas's offense will probably improve in the upcoming football season, but it might not.

Each of these truths makes a claim about a way the world could, would, can't or might (and hopefully will!) be. Words like 'could', 'would', 'possible', 'might', 'must', 'will', 'necessary', and so on are modal. They concern not just what is, but what could or will or must be.

When we study modality, we are studying the nature of truths about what could or will or must be, especially as that nature contrasts with the nature of truths about what simply is. When we study the metaphysics of truths like 'The cat is on the mat', we are concerned with the metaphysics of the cat and the mat and the relation between them. But things aren't so straightforward with truths like 'The cat *might be* on the mat.' The actual nature of the cat and the mat, and the actual relations between them don't seem to be the only things that matter. The cat needn't be on the mat for it to be true that the cat might be on the mat. But then, how must the cat and the mat *be* for this to be true?

*The Atlas of Reality: A Comprehensive Guide to Metaphysics*, First Edition.
Robert C. Koons and Timothy H. Pickavance.

Do these facts fall within the scope of metaphysics? One might argue that they are really epistemological in nature (concerned merely with what we know or don't know). Without a doubt, there are uses of 'might', 'must', 'possibly', and 'necessarily' that are epistemic in character. For example, if one observes that one's lawn is wet in the morning, one might utter the following:

(5) It might have rained last night or the sprinklers might have gone on.

If one's spouse assures one that she turned the sprinklers off last evening, one might then conclude:

(6) It must have rained last night.

The 'might' of (5) seems to express something like 'for all I know'. One doesn't know that it didn't rain last night, so it might have. Conversely the 'must' of (6) expresses knowledge: one now knows that it rained last night, so it 'must' have rained.

However, there are other uses of these words that seem to have nothing to do with knowledge. For example, one might conclude, after years of study, that Lee could have won the battle at Gettysburg if he had avoided certain strategic errors, like ordering Pickett's charge. By saying that Lee 'could' have won the battle, one is not expressing any doubt or ignorance as to whether he did in fact lose the battle. We all know that is true. The metaphysical sense of 'could' concerns whether there exists an alternative scenario of the right sort according to which Lee did win the battle, even if we know that all such scenarios are counterfactual.

This in turn leads to the problem of what exactly is a mere possibility. Philosophers of modality since the time of Leibniz have found it convenient to think in terms of *possible worlds*. But what is a possible world? Roughly, a possibility is a maximal way the world might be. A way the world might be is a situation or scenario that could occur or could have occurred (in that metaphysical sense of 'could'). THP could have had oatmeal for breakfast, RCK might have worn a green shirt yesterday, Lyle could have colored with crayons rather than markers; these are ways the world might be. But *two plus two's being five*, *substance S's being both human and non-human*, and *that ball's being both red and blue all over at the same time*, these are ways the world could *not* be. They are *ways*, but not ways the world might be. They are not *possible* ways. A way the world might be is *maximal* if and only if every proposition is either true or false according to it. Consider a possibility like *THP's having oatmeal for breakfast*. This possibility represents THP's having oatmeal, but represents nothing about the color of RCK's shirt. But a possibility like *THP's having oatmeal for breakfast and RCK's wearing a green shirt yesterday* represents THP's having oatmeal and RCK's shirt. It is, we might say, more representationally rich. A maximal possibility is maximally representationally rich. One cannot pick a proposition that a maximal proposition represents nothing about. Maximal possibilities represent THP's breakfast choice, RCK's shirt color, the facts of math, the existence of God, and so on. We can thus think of a possible world in the following way:

**Def D14.1 Possible World.** A *possible world* is a possibility that is maximal in that every proposition is either true or false according to it.

As we move forward, we will consider various views about the precise nature of possible worlds, but each view is compatible with this initial characterization.

The vast majority of maximal possibilities somehow misrepresents the world. Indeed, all but one must do so. One maximal possibility represents THP's having oatmeal for breakfast, one that he had yogurt, one that he had an egg sandwich, and so on. Only one of these possibilities matches the actual state of the world. This special maximal possible way the world might be, this special possible world, is the *actual* world.

We can use possible worlds to explain certain modal notions. Here are a few examples. The sentence 'The cat might be on the mat' is true if and only if there is a possible world of a certain sort according to which the cat is on the mat. 'It's not possible to travel from Earth to Mars in 14 minutes' is true if and only if there is no possible world of a certain sort according to which someone travels from Earth to Mars in 15 minutes. And 'You must help old ladies cross the street' is true if and only if every world of a certain sort is such that you help old ladies cross the street. Patterns emerge here. 'Necessarily *p*' is true if and only if *p* is true according to all worlds of a certain sort; 'Possibly *p*' is true if and only if *p* is true according to (at least) one world of a certain sort; and 'Impossibly *p*' (that is, 'Necessarily not *p*') is true if and only if *p* is true according to no worlds of a certain sort. Further, we can say that *p* is contingent if and only if *p* is true according to (at least) one world of a certain sort and is false according to (at least) one world of that sort.

The reader would be right to wonder why the phrase 'of a certain sort' shows up in the previous paragraph. The reason is simple: there are different types of possibility, necessity, and impossibility. For example, consider (3) again, that it is not possible to travel from Earth to Mars in 25 minutes. This is true in one sense, in that *given the state of existing technology*, such a fast trip cannot happen. But there is another sense in which it is false, since it's possible for technology to improve quite dramatically. We might say that travelling from Earth to Mars in 25 minutes is *technologically* impossible but *metaphysically* possible, possible in the broadest sense. There are other varieties of possibility as well. For example, things that are possible *given the actual laws of nature* are *nomologically* possible. And, as we have already seen, there are *epistemological* possibilities. This list could be extended further. We use 'of a certain sort' to accommodate these distinctions. Just as there are nomological possibilities, there are nomologically possible worlds. Just as there are technological possibilities, there are technologically possible worlds.

A corollary of this is that modal vocabulary in natural languages like English are flexible. The same modal terms can be used to designate different species of modality. It's important to mark, in particular cases, what species is in view. Our discussion of modality will focus on *metaphysical* possibility, necessity, and impossibility, that broadest kind of modality.

Before we move on, it might be worth offering an example of why one might care about this issue, an example of why the study of modality matters to philosophy more generally. It is plausible that modality is importantly connected to understanding. Understanding a contingent claim, a claim that might be true and might be false, seems to require the capacity to conceive of possibilities according to which the claim is true and possibilities according to which the claim is false. Additionally, if one embraces Possible Worlds, then there is a question about just how many there are. One answer is quite simple and principled: one, and it just is the actual world. This is Necessitarianism:

**Necessitarianism.** The only possible world is the actual one.

If Necessitarianism were true, then there would be nothing that is true only contingently.[1] Anything that is actually true would be necessarily true. That is, all truths would be incapable of being false. Clearly, Necessitarianism entails metaphysical fatalism, the view that there are no contingent truths about the future. If there are any future truths, they are also necessarily true.

If fatalism were true, then all deliberation about future actions would be in some sense illusory or ill-founded. All of our deliberations presuppose that certain matters are up to us and so not yet fully decided or fated. As Aristotle noted, it is impossible to deliberate about the past, given that past facts are now necessary and fixed. Anything that follows of necessity from such fixed past facts is also beyond the scope of rational deliberation. No one, for example, would deliberate about whether the law of gravity should continue to hold or whether the sun should rise (assuming that the rotation of the earth is already fully determined by the condition of the solar system). We deliberate only about matters that we take to be ones over which we hold some power of deciding between real alternatives. The Rational Practices Presumption (**PEpist 2.1**) gives us a reason to reject Necessitarianism.

It might be argued that rational deliberation requires only the epistemic possibility of alternative futures. All we must assume is that we do not know with certainty how the future is fated to unfold. However, this is simply false. Whenever we find out, after the fact, that we were wrong in thinking that the future was really open with respect to certain alternatives, we conclude that our deliberation about which alternative to bring about was empty and pointless. For example, suppose that one believes that one has two alternatives open to one: to leave or to stay in one's living room. If it turns out that one is locked in one's living room in such a way that one cannot possibly leave, then one's deliberation about whether or not to leave is, objectively speaking, pointless. This is so even if it is, subjectively speaking, a reasonable thing for one to deliberate. A metaphysical fatalist must believe that there are never real alternatives open to anyone. Thus, she must believe that all deliberation is objectively pointless. Believing this, she cannot rationally engage in deliberation at all.

Clearly, there is an immediate decision that must be made regarding possible worlds. Just one of the following must be true:

**14.1T Possible Worlds.** There are possible worlds.

**14.1A No Possible Worlds.** There are no possible worlds.

If we plump for Possible Worlds, there are two broad issues to examine. The first issue concerns what sorts of things possible worlds are. There are two primary questions in this area. First, what exactly *are* possible worlds? Some philosophers think possible worlds are something like a parallel universe "out there" somewhere, while others think they are more like propositions or mathematical structures or other abstract objects. Second, how are things true or false *according to* possible worlds? That is, how is that possible worlds *represent* things?

The second issue is whether modal facts are reducible to facts about possible worlds. This breaks into two sub-issues as well. The first is whether facts about possibility and necessity are reducible to facts about possible worlds, and the second is whether facts about actuality are so reducible. And if modal facts are not reducible to facts about possible worlds, what other options are there for explicating modal truth, and how then do possible worlds fit into the mix?

We begin with two contrasting views about the nature of possible worlds, Concretism and Abstractionism:

**14.1T.1T Concretism.** Possible worlds are maximal concrete objects.

**14.1T.1A Abstractionism.** Possible worlds are maximal possible abstract objects.

The fundamental difference between Concretism and Abstractionism is whether possible worlds are concrete or abstract. The distinction between abstract and concrete objects isn't easy to characterize, but there are at least two features that helpfully distinguish them. First, concrete objects, but not abstract objects, typically have a definite location in space and time. Concrete objects, at least in typical cases, are not wholly located at more than one spatiotemporal location. Abstract objects, on the other hand, tend to have no spatiotemporal location at all, or are at least capable of being wholly in many spatiotemporal locations. Second, concrete objects, but not abstract objects, have active and passive powers. Concrete objects can change other concrete objects, and are able to undergo change themselves. Abstract objects do neither. Concretism, then, is roughly the view that possible worlds are like parallel universes, universes like our own, populated with donkeys and buildings and people and stars and so on (more on this below). Abstractionism is roughly the view that possible worlds are maximal possible propositions. As will emerge, these differences manifest in still further, possibly more fundamental disagreements between Concretists and Abstractionists, both with respect to how possible worlds represent and with respect to whether modal facts are reducible to facts about possible worlds.

## 14.2 Concretism: Worlds as Universes

Concretism is the view that possible worlds are like parallel universes. The thought behind Concretism is, therefore, deceptively simple. But one important motivation for it is deceptively simple as well. Consider (7):

(7) Grass is green.

No doubt (7) is true. We can further note that (7) describes the way the world is: grass's being green is a way the world is. As we have seen, though, the world might have been a different way. Grass might have been purple. Grass's being purple is a way the world might have been. Notice that in both cases, we have spoken of *ways*. In the first case, we spoke of a way the world *is*; in the second, a way the world *might be*. But what are these *ways*? In the case of grass's being green, it's tempting to think that that *way* consists in there being certain things with certain features in certain relations. The fact that grass is green is a way the universe is. So it is also tempting to think that the actual world, the maximal way the world in fact is, is just the universe (where that is taken as including even non-physical things, if such there be). But if *ways* are just facts, then there must exist the fact that grass is purple. Grass's being purple is a *way*, and *ways* consist in there being certain things with certain features, in certain relations. Thus it follows that there are certain things with certain features, in certain relations, such that grass is purple. Counterintuitively, then, we have arrived at a commitment to the claim that there is, in some sense, purple grass. Indeed, there are humans that can fly without mechanical aid, there are objects unaffected by gravity, and so on. Possible worlds are, in this sense, just parallel universes. And these other *ways*, the possible but non-actual ones, are merely possible, rather than actual, because the relevant facts are not part of our world. David Lewis, the most prominent defender of Concretism (see especially Lewis (1986a)), offers this Concretist slogan: every way the world might be is a way some world is.

It is important to forestall a common confusion at this point. Concretists do *not* say that there is actually existing purple grass. They simply say that there is purple grass, purple grass that really does exist. But this grass is non-actual. It is merely possible. The difference between actual and merely possible objects, however, does not consist in a special way that merely possible objects exist. Rather, it is a matter of proximity. In particular, it is a matter of proximity to *you and us*. What it is to be actual, according to Concretism, is to be part of *our* world. We fill out these points about possible and actual truth in what follows.

There are two crucial parts of Concretism that allow it to make sense of modal truths. First, recall the connection we noted above between possible worlds and modal vocabulary. 'Possibly *p*' is true if and only if there is at least one world (of a certain sort) according to which *p* is true; 'Necessarily *p*' is true if and only if *p* is true according to every world (of a certain sort); 'Impossibly *p*' is true if and only if *p* is false according to every world (of a certain sort); and so on. Concretists embrace these connections. Indeed, Concretism is a simple-minded extension of these connections. To see why this is so, we need to consider the second part of Concretism.

Second, then, Concretists say that non-modal propositions are true according to possible worlds in just the way that they are true according to our world, the actual world. Whatever it takes for a non-modal proposition to be true in the actual world, that's what it takes for a non-modal proposition to be true in any other possible world. For example, consider (8):

(8) There are talking donkeys.

What would it take for (8) to be true? That is, what would it take for (8) to be true in, or true according to, the actual world? The answer seems pretty obvious: Talking donkeys would have to be part of our world. Concretists say that other possible worlds are just like ours, so the account of the *according to* relation goes in just the same way. In particular, it's true according to some world *w* that there are talking donkeys if and only if there are some talking donkeys that are a part of *w*. On this view, what it is for something to be true according to a merely possible world is just what it is for something to be true according to the actual world. The Concretist account of the *according to* relation is not, therefore, a modal account (at least insofar as one's story about truth in our world is non-modal).

Putting these two items together, we can see how Concretists account for the truth of modal claims. Consider (8) again. The proposition that there are talking donkeys is actually false but possibly true. It is actually false because the proposition that there are talking donkeys is false according to our world, which is to say, no part of our world is a talking donkey. Actual truth is truth in our world. Actual falsity is falsity in our world. On the other hand, (8) is possibly true because there is a possible world according to which there are talking donkeys, which is to say, there are talking donkeys that are part of some world or other. We can generalize these accounts. With respect to possibility, necessity, and so on, we can say that a proposition *p* is possibly true if and only if there is a world *w* according to which *p* is true, where for Concretists that means that the parts of *w* have the right properties and stand in the right relations so that *p* is true. The account for necessity and so on works similarly. It's possible that grass is purple if and only if there is a world with purple grass as a part; it's necessary that the God of Christianity exists if and only if the God of Christianity is a part of every world, and so on. With respect to actuality, a proposition *p* is actually true if and only if my world is such that *p*. So, it's actually the case that grass is purple if and only if our world has purple grass in it, and it's actually the case that the God of Christianity exists if and only if the God of Christianity is a part of our world.

More technically, to say that something is possible is simply to say that it is true, when its quantificational phrases and terms are interpreted in the widest possible way, taking into account the whole of reality (i.e., taking into account all of the parallel universes). Consider (9):

(9) Possibly, there are talking donkeys.

Both (8) and (9) contain the quantificational phrase 'there are'. Other quantificational phrases include 'something is', 'nothing is', 'everything is', 'no one is', 'some dogs are', 'all trees are', and so on. Quantificational phrases have a *domain*, which is the class of objects relevant to determining the truth of the claim involving that quantifier. For example (and simplifying somewhat), the domain of 'all trees are' is just the class of trees. If one said 'All trees are *F*', the way to check the truth of that claim is to see whether everything in the class of trees is *F*. If so, then 'All trees are *F*' is true. If not, then that sentence is false. Ordinarily, the largest domain that is relevant to claims containing a quantificational phrase is the class of all actually existing things. For example, the way to check whether (8) is true is to look around in the class of all actually existing things to see whether that class includes any talking donkeys. If so, then (8) is true; if not, then (8) is false.

Concretists, though, think that modal vocabulary expands the domain of quantification even further, out to the class of all existing things, rather than just the class of all actually existing things. Concretists, of course, take these to be different classes, since the actually existing things are just the things that exist in our world. Things in other worlds do not actually exist. Back to (9): Concretists say (9) is true just in case there are talking donkeys in the largest domain, that is, just in case they exist in some possible world or other. In fact, (9) is equivalent to one reading of (8). (8) is false, we have seen, if the domain of 'there are…' is the class of actually existing things. If we interpret 'there are' as speaking of anything, without restriction, then (8) is true, so long as there is at least one possible world in which some donkey talks.

Having sketched the nature of Concretism, and seen how Concretists might supply as many worlds as there are possibilities, we now face an obvious question: Why in the world would one ever embrace this crazy view?!? It's just so *implausible* that there are loads of other concrete universes that aren't parts of our world, and even if there were, how could we ever know that there are such things at all? We'll return to the second half of this question in a little while. Right now, we will focus on the first half.

At the beginning of this chapter, we noted something odd about modal truths, namely, that they seem to have little if anything to do with the way the world actually is. What, then, one might wonder, could ground the truth of modal claims? One way to deal with this oddity is to find some existent things in which one can ground modal truths, and then find a way to reduce modal locutions to locutions having to do with those existent things. In other words, if we can find a way to reduce modal truths to truths of some other, more easily understood sort, then we can minimize the oddity that is modal truth.

Concretism offers such a reduction of possibility, impossibility, necessity, and related matters. David Lewis (1986a) argues that Concretism allows one to claim that no modal truth is fundamental, whether about possibility, contingency, necessity, or impossibility. Further, Lewis claims this reduction as a reason to believe that Concretism is true. In fact, Lewis claims that Concretism offers a way to reduce a number of challenging philosophical items, from possibility and necessity to counterfactuals, properties, propositions, mental content, supervenience, and more. We focus on the reduction of modality here. (But we encourage the reader to think about, for example, how Concretism helps Reductive Nominalists (7.1A.1T/8.1T) deal with the extensionality problems we considered in Chapter 8.) Suppose presently that Concretism can successfully reduce modal facts to facts about possible worlds. (This supposition is challenged by the problem of isolation, considered below.) Then Concretism enjoys an advantage in terms of qualitative economy over any view of modality that requires fundamental modal truths, and still more so over any view that requires fundamental truths and that can't also offer reductive accounts of those other things (**PMeth 1.4**).

This advantage in qualitative simplicity, however, exacts a steep quantitative cost. On top of all the green grass that exists in our world, Concretists must commit to the existence of all the possible purple, pink, and magenta grass inhabiting other worlds. Concretism includes not just actual humans, but all the merely possible ones as well; not just any actual extraterrestrial persons, but all possible extraterrestrial persons; not just the actual coffee beans, but all the possible ones as well. This is not to mention the objects whose kinds we have never even conceived, the true *aliens*. If something is even possible, the according to Concretism, it exists. Abstractionists don't have to commit to all this grass and coffee or humans and aliens. It is true that Ockham's Razor implores us to prioritize qualitative economy over quantitative economy (**PMeth 1.4.1**), but quantitative economy counts for something. It's vital, then, that one consider whether Concretism's reductive accounts of these various notions are successful.

At any rate, in Chapter 15, we will see that Abstractionism cannot supply a reductive account of modal truths, and this gives some reason to prefer Concretism. For now, we focus on Concretism's reduction strategy.

Concretism offers a reductionistic account of possibility and actuality because there are no modal notions deployed in the final analysis. We have seen that Concretism's account of the way possible worlds represent is non-modal. *Parthood* is not a modal notion. And what makes something a part of our world is not modal either. Concretists reduce modal notions like *possibility*, *necessity*, and so on to facts about the intrinsic makeup of the very many existing concrete worlds. To exploit the more technical bits about quantification above, we can see that Concretist accounts are reductionistic because quantification is not modal, and Concretists say that modal vocabulary simply specifies a domain of quantification. Further, Concretism reduces facts about actuality to facts about the intrinsic makeup of *our* world. Thus, Concretism requires Modal Indexicalism and denies Modal Anti-Indexicalism:

**14.2T Modal Indexicalism.** All attributions of actuality are indexical in character.

**14.2A Modal Anti-Indexicalism.** Some attributions of actuality are not indexical in character.

Concretism is Indexicalist in that it specifically deploys the notion of something's being related in some way to *us*. This is one way, but not the only way, for an attribution to be indexical.

**Def D14.2 Indexicality.** An attribution is *indexical* in character if the content of a thought making the attribution depends essentially on some extrinsic feature of that thought or its thinker (such as location in space, time, or "logical space").

Consider the following thoughts:

(10) Austin is here.
(11) I am hungry.
(12) It is now noon, EDT on 17 July 2016.
(13) A human being has actually walked on the moon.

It is fairly uncontroversial that thoughts (10) and (11) are indexical in character. What is thought by means of (10) and (11) depends upon who is thinking them and where he or she is located while thinking them.

There is some controversy about whether (12) is indexical, a controversy we will explore in detail in Chapters 20 and 21. But here is a foretaste: so-called "B-Theorists" or "Anti-Tensers" (20.2A) hold that (12) is indexical in character. On these views, (12) expresses a different thought, depending on when it is (tenselessly) being thought. "A-Theorists" or "Tensers" (20.2T), in contrast, believe that (12) is non-indexical. Its content is to ascribe a metaphysically primitive property of *present-ness* to a moment of time.

Modal Indexicalism's reduction of actuality parallels the reduction of time in Anti-Tensism. Modal Indexicalists think that (13) is indexical in character. It is either true or false, depending on "where" (or "in which world") it is being thought. In the world that we call the actual world, our world here, (13) is true, since a successful moonwalk is part of our world. However, there are other possible worlds, worlds without moonwalks, and when (13) is thought in those worlds, it is false. Which world is actual for one is relative, on this view, to the perspective that one occupies.

Modal Anti-Indexicalists, by contrast, think that all this is confused. The proposition expressed by (13) attributes a special, metaphysically primitive property of *actuality* to some world containing a moonwalk. In our case, (13) attributes the property of *actuality* to the actual world. Since a human being did indeed walk on the moon, (13) is true *simpliciter* (that is, true absolutely and without qualification), and not merely true here in our possible world.

The distinction between Modal Indexicalism and Anti-Indexicalism seems to be a consequence of a metaphysical disagreement about the nature of possible worlds themselves. Indexicalists must be Concretists. Otherwise, the indexicality does no work. In contrast, Anti-Indexicalism situates much more nicely with Abstractionism. At any rate, whether or not our universe is part of a multiverse of parallel universes is, for Anti-Indexicalists, completely irrelevant to evaluating propositions about actuality. If dragons existed in some parallel universe, then the Anti-Indexicalists would conclude that (14) is true:

(14) Dragons actually exist.

Parallel *universes* would, for Anti-Indexicalists, be every bit as actual as our own, local universe. A *possible world* is something altogether different. A non-actual possible world is a way for *all* of reality to be, other than the way it actually is.

Now consider Possible Actuality:

**Possible Actuality.** If there is more than one possible world, then which world is actual is a contingent matter. That is, a *possible* world is a world that *could have been* actual.

Concretists, if they reject Necessitarianism, must not claim, given Possible Actuality, that the actuality of the actual world is due to some essential feature of that world. Each world has its essential features as a matter of necessity, not contingently. And all of the facts about what would be true if the world were actual—the *content* of a possible world—are essential to that world.

**Def D14.3 Content of Worlds.** The *content* of a possible world is the class of propositions that would be true if that world were actual.

**Essentiality of World-Content.** The content of each possible world is essential to it.

However, possible worlds have intrinsic properties of only two kinds. First, there are properties determined by a world's content, and second, there is (possibly) the primitive property of *actuality*. Hence, Concretists must deny that which world is actual is determined by the intrinsic properties of that world. Thus, the actuality of the actual world must consist in its extrinsic relation to something. The only plausible "somethings" that might be relevant are *us*. Hence, Concretists must embrace Modal Indexicalism.

## 14.3 Problems for Concretism

So far, we have canvassed the nature of Concretism and described the ways that it carries out its reduction of possibility and actuality. It is time to turn to problems for the view. (Another problem, having to do with the way Concretism handles *de re* modality, is explored in Chapter 16.)

1. The problem of irrelevance. Concretism seems to give an obviously incorrect account of the truth of modal claims (Plantinga 1974: 116; van Inwagen 2001: 222, 226). Whether there are parallel universes is relevant to what is actually the case, not to what is possibly or necessarily the case. Any universes running in parallel to ours are just as much part of actual reality as our own universe is. The distinction between local and remote universes has nothing to do with the distinction between actual and non-actual possibilities. Appealing to facts about parallel universes to ground modal truths is like appealing to facts about a thing's shape in order to ground truths about its color.

Concretists would insist that what goes on in parallel universes has *everything* to do with what is possibly and necessarily the case. *What it is* for a proposition to be possibly true is for there to be a world according to which that proposition is true. And *what it is* for a proposition to be necessarily true is for it be true according to every world. And Concretism's parallel universes *just are* possible worlds. What possible worlds are like has everything to do with modality, and so what Concretism's universes are like has everything to do with modality as well. Importantly, though, Concretists would also claim that here in the actual world, these proposition really do have the property of being possibly or necessarily true.

This dispute raises an important methodological point. Concretists defend their theory on the grounds of its simplicity, theoretical power, and consistency with linguistic data. The modal sentences that we believe to be true can be supplied simple, straightforward truthmakers by Concretists, in the form of concrete universes of the right kind. However, their opponents will point out that there is one important bit of data that contradicts Concretism, namely, our naive, intuitive sense that the sort of truthmakers supplied by Concretists are of the wrong sort. Facts in other universes just couldn't make our ordinary assertions about possibility true, given our understanding of what those assertions mean.

Concretists could reply in one of two ways. First, they could simply dispute the supposed semantic data. They could claim that, for all we could know as competent speakers of English, these assertions of possibility might be made true by facts about isolated, concrete universes. Alternatively, they might concede that their opponents have access to some genuinely recalcitrant data, while insisting that the other virtues of Concretism are so great that we should dismiss these wayward intuitions as mistaken. If our semantic intuitions are fallible, then they may fall as "spoils to the victor".

2. The problem of ethical fatalism. There are a variety of ethical absurdities that seem to follow from Concretism. In particular, Concretism seems to entail a kind of moral fatalism. Every possible choice is made by someone in some world. For example, if THP

actually chooses to wear a green shirt rather than an orange shirt, then in some other world THP (or someone very like him, one of his *counterparts*, see below) chooses an orange shirt rather than a green one. The same goes for our moral choices. For every bad choice made by someone in our world, the corresponding good choice is made by someone like us in another world, and vice versa. Further, nothing any of us can do can possibly change the total quantity of happiness and unhappiness in reality, since that is inexorably fixed by the collection of possible worlds. At most, we can affect the amount of happiness contained by the actual world, that is, by our own local world (see Adams 1974: 215–216).

Bizarrely, if there are no duplicate possible worlds (if each possible world is qualitatively and internally unique in some way), then acts of masochism and self-denial would be morally heroic. By harming myself, I am sparing one of my counterparts that same harm, by making the harm happen in my world instead of his (see Pruss 2011: 105–106 for this and other ethical absurdities).

Concretists reply that ethics must be agent-centered. We have no obligation to reduce the total amount of evil and suffering and increase the amount of good and pleasure in existence overall. We only have an obligation to reduce the total amount of evil and suffering and increase the amount of good and pleasure near us, in our own possible world.

3. The problem of scientific induction. David Lewis (1986a) couples Concretism to Neo-Humeism (4.4T). Given that the main argument for Concretism is its theoretical usefulness for supplying reductive accounts of various notions (see above), it is no surprise that Lewis would hope to reduce facts about conditionals, laws, and powers in a Neo-Humeist way, exploiting the resources of Concretism. But as we discussed in Chapter 5, if Neo-Humeism is true, then there are far more worlds (by any reasonable measure of 'more') that agree with all of our past observations but which fail to obey any simple generalizations and laws hereafter (and anywhere beyond our actual, past observations). We would seem then to have no reason for believing that the actual world is one of the few regular, well-behaved worlds, and so any confidence in scientific induction would be unwarranted.

As we saw, Neo-Humeists must respond by asserting that a preference for simplicity and inductive uniformity is simply a primitive, underivable postulate of reason, unaffected by considerations of what is true in the majority of worlds empirically similar to ours. We have already discussed the main objection to this reply, namely, that this additional postulate of reason is ad hoc, and so counts against the simplicity of the Concretists' account.

However, there may be a further problem once Concretism is added to Neo-Humeism. It's one thing to say that it is reasonable to believe that the one and only, metaphysically unique actual world is simple and regular, and quite another to believe that I just happen to be located in one of the few simple and regular concrete worlds. The first has some plausibility as a fundamental principle of rationality, even if it is somewhat ad hoc. The second seems hopelessly optimistic beyond all reason (see Pruss 2011: 117–119 for a formalization of this argument).

4. The problem of non-indexical uses of 'actual.' There are linguistic data that Modal Indexicalists cannot account for. Consider (15):

(15) The actual world might not have been actual.

The first use of 'actual' in (15) does seem to be indexical, but the second use cannot be, since (15) seems obviously true. If world Alpha is the actual world, Modal Indexicalists must read (15) as equivalent to (16):

(16) Alpha might not have been Alpha.

But (16) is obviously false. However, (17) seems both true and equivalent to (15):

(17) Alpha might not have been actual.

Contrast, for example:

(18) The equator is here, but it might not have been here.

(18) hardly makes any sense. It seems to be saying that the equator might not have been the equator. (17), in contrast, is asserting that there is some property, *actuality*, which Alpha has but which it might have lacked. The trouble is that Indexicalists don't have the option of reading (15) as (17).

Modal Indexicalists might respond that (17) can be understood as true, since 'might have been' is to be interpreted as 'is so in some other world'. This reading is suggested by (19):

(19) There is a possible world in which Alpha is not actual.

The problem is that 'actual' in (19) is not functioning indexically. If it were, (19) would be asserting the same thing as (20):

(20) There is a possible world in which Alpha is not Alpha.

But (20) is obviously false. How then, can we understand (19) as true? Modal Anti-Indexicalists will argue that (19) should be taken as a somewhat non-standard way of saying what is expressed by (21):

(21) There is a possible world *w* such that, were *w* actual, Alpha would not be actual.

According to Anti-Indexicalists, by talking about what is 'actual', we are talking about a special property of *being actual* that Alpha has and other worlds lack but could have had. We are not merely using an indexical to refer to our local world.

The best that that Concretists can do to make sense of (19) is to suppose that it is really making use of 'actuality' in 'scare quotes':

(19′) The actual world might not have been 'actual'.

(19′) could be understood as pointing out that, when people in other possible worlds use the word 'actual', they are not referring to Alpha:

(22) There are other possible worlds in which the word 'actual' does not refer to Alpha.

(22) is certainly true, but Anti-Indexicalists can plausibly claim that the true meaning of (19) is given by the non-indexical (21), and not by (22).

5. The problem of isolation. If all of Concretism's possible worlds really do exist, then it seems that there is just One Really Big World, containing myriad parallel universes, rather than many possible worlds, as Concretism insists. We might put the point in the form of a question. What makes something a world, rather than part of a world? What, that is, *isolates* worlds from one another? Put another way, what makes something part of the same world as something else? What, that is, makes two things *worldmates*?

There are three plausible answers to these questions. First, one might say that the *worldmate* relation is fundamental, that it is just a primitive of the Concretists' theory. On this view, there is no informative answer to the question of what makes two things worldmates. This undermines the Concretists' reduction strategy, since the worldmate relation appears to be primitively modal if it's primitive at all. Why? Because one now has fundamental facts about which possible world something is a part of. Given that Concretism becomes quite implausible if it doesn't at least offer a reductive account of modality, this situation would be rather unhappy for Concretism.

Second, one might think that two things are worldmates if and only if they are causally interrelated. On this view, worldmate relations are causal relations. This view couples poorly with Neo-Humeism, since it appears to require fundamental causal relations. That is not, of course, a principled objection to the view, but it is worth noting given the standard combination of Concretism with Neo-Humeism.

Third, one might think that two things are worldmates if and only if they are spatiotemporally interrelated. On this view, worldmate relations are spatiotemporal relations. Both the causal and spatiotemporal reductions of the worldmate relation face the following problem. It seems that there might have been several parallel universes which are spatially, temporally and causally isolated from each other. If the worldmate relation is reduced either to causal relations, spatiotemporal relations or a combination of these, this is just not possible. Such parallel universes would count as different worlds, and so could not be universe-size parts of the same world. To avoid this problem, Concretists must give up the claim that a possible world is a maximal aggregation of spatiotemporally and/or causally connected things or deny that parallel universes are possible. Both of these moves are problematic. The first strips Concretists of any reductive account of the worldmate relation. The second suggests that Concretists cannot have as many worlds as there are possibilities. Both are seriously problematic for Concretism's reduction of modal facts to facts about possible worlds.

## 14.4 Conclusion

Concretism provides a simple and reductive account of possibility according to which proposition is possibly true just in case it is true in or with exclusive reference to some concrete, spatiotemporally connected universe. A proposition is actually true (for us) when it is true in our own local universe. Although attractively simple, Concretism faces some serious challenges, some of which we have canvassed above.

## Note

1 Strictly speaking, this is only true if the number of possible worlds mirrors the number of possibilities, something we will assume here. One could reject this, but we won't get into the reasons for doing so.

# 15

# Abstractionism: Worlds as Representations

Given the troubles with Concretism and its immense implausibility even on top of those problems, many philosophers have sought to give an account of possible worlds that doesn't commit one to a plenitude of universes. These philosophers tend to think that possible worlds, rather than being concrete, are abstract. That is, they are more like properties, propositions, and sets than they are like rocks, donkeys, and angels. This view about possible worlds is Abstractionism:

**15.1T Abstractionism.** Possible worlds are maximal possible abstract objects. (=14.1T.1A)

Some Abstractionists think that worlds are maximal possible states of affairs. States of affairs are much like *ways*, though understood as *abstracta* that can obtain or fail to obtain in virtue of the way things are. So, for example, the state of affairs of grass's being green obtains because grass is green. One might think of the state of affairs of grass's being green as obtaining because of the fact that grass is green. For Concretists, *ways* are facts. For Abstractionists, *ways* are states of affairs. Other states of affairs include THP's being taller than RCK, the United States having 50 states, grass's being purple, a square's having five sides, and so on. A state of affairs is possible if and only if it is possible for it to obtain. A square's having five sides is not possible: it simply could not obtain. But grass's being purple is possible, despite that it does not obtain in fact. A state of affairs $s$ is maximal if and only if for any state of affairs $s'$, $s$ either includes or precludes $s'$. State of affairs $s$ includes state of affairs $s'$ if and only if necessarily, if $s$ obtains then $s'$ obtains; $s$ precludes $s'$ if and only if, necessarily, if $s$ obtains then $s'$ does not obtain. States of affairs are maximal possible states of affairs if they are, in a sense, complete descriptions of the way the world could be, such that no more information is needed to fully characterize the world. Such states of affairs are, on this view, possible worlds. Some states of affairs

*The Atlas of Reality: A Comprehensive Guide to Metaphysics*, First Edition.
Robert C. Koons and Timothy H. Pickavance.

are too "small" to be worlds because they aren't complete descriptions of the way things might be.

Other Abstractionists think that worlds are maximal possible propositions. This view is structurally similar to the view that possible worlds are maximal possible states of affairs. *Obtaining* and *failing to obtain* are replaced with *truth* and *falsity*, and maximality is characterized using *entailment* rather than inclusion and preclusion: proposition $p$ is maximal if and only if for any proposition $p'$, $p$ entails either $p'$ or not- $p'$. Still other Abstractionsists think that worlds are maximal possible classes of propositions. On this view, maximality is understood in terms of *membership*. Class of propositions $S$ is maximal if and only if for any proposition $p$, either $p$ is a member of $S$ or not-$p$ is a member of $S$.

A third group of Abstractionists take worlds to be maximal structural universals (see Section 10.3) that can only be properties of reality as a whole. The actual world is the one world that has an instance; all other worlds are uninstantiated maximal structural universals. A possible world is a property that has the second-order property of *possibly having an instance*. A world-universal $u$ is maximal in the sense that, for every other structural universal $u^*$, $u$'s being instantiated either entails that $u^*$ is also instantiated or else entails that $u^*$ is not instantiated.

The distinctions among these types of Abstractionism won't matter for our purposes, since the most useful taxonomy of Abstractionist views categorizes them using the ways they make sense of how possible worlds represent possibilities, rather than in terms of the categorial nature of worlds. So far, we have not addressed this question. In what follows, we consider three ways that Abstractionists might account for how possible worlds represent that such-and-such. There is Magical Abstractionism (Section 15.1), according to which that question has no informative answer. Then there are two views that answer the question by appealing to structural features of worlds. There is Linguistic Abstractionism (Section 15.2.1), according to which possible worlds represent in the way that languages do. And there is Pictorial Abstractionism (Section 15.2.2), according to which possible worlds represent in the way that pictures do.

## 15.1 Magical Abstractionism

Magical Abstractionism is the view that possible worlds represent in a primitive way. On this view, all we can say about the way worlds represent is that a world $w$ represents that $p$ if and only if it is necessarily true that if $w$ were actual, then $p$:

**15.1T.1T Magical Abstractionism.** A world $w$ represents that $p$ if and only if it is necessarily true that if $w$ were actual, then $p$, and there is no substantive account to give about why possible worlds represent what they do.

Alvin Plantinga (1974) is most plausibly read as a Magical Abstractionist. He says nothing about why worlds represent as they do, and offers the right-hand side of the first conjunct of the above definition as his account of representation. Importantly, according to Magical Abstractionism, worlds do not represent in virtue of having some type of internal structure. Any parts, and any arrangements of parts, that worlds might have are simply irrelevant to what they represent. Worlds just have representational features, and that is that.

It is easy to see that Magical Abstractionism embraces the existence of fundamental modal truths. If Magical Abstractionism is true, there are metaphysically fundamental modal truths even in the characterizations of possible worlds suggested above. For ease, focus on worlds as maximal possible propositions. (Similar remarks apply to the other views.) The notion of a *possible* proposition is explicitly modal: *p* is possible if and only if it is possible for *p* to be true, or *p* could be true. Entailment is modal as well, and so maximality is modal. Magical Abstractionism has no resources to reduce these appeals to modality, given that the view cannot offer a reductive account of representation. Matters will become complicated when we consider Structural Abstractionism below. Structural Abstractionists attempt to supply a substantive account.

### 15.1.1 Modal Anti-Indexicalism

It is clear that Magical Abstractionists require fundamental truths about possibility, but what about actuality? Magical Abstractionists, and indeed all Abstractionists, are agreed in rejecting one type of reductive account of actuality, namely Modal Indexicalism (14.2T). If the term 'actual' functions as an indexical, then every possible world is a concrete entity or sum of concrete entities, like a place or a time. This is obviously incompatible with Abstractionism. Abstractionism thus requires Modal Anti-Indexicalism:

**15.2T Modal Anti-Indexicalism.** Some attributions of actuality are not indexical. (=14.2A)

There are several varieties of Anti-Indexicalism. The most straightforward is Simple Anti-Indexicalism, according to which the property of *being actual* is a simple, fundamental property.

**15.2T.1 Simple Anti-Indexicalism.** *Actuality* is a simple, fundamental property of possible worlds.

There are a number of problems with Simple Anti-Indexicalism. First, the following inference should be logically valid:

(1) Lions (actually) exist.

Therefore,

(2) The actual world contains lions.

Indeed, the two statements seem to be logically equivalent. But how can (1) and (2) entail one another if *actuality* is a simple property of possible worlds? (1) makes no reference to worlds at all. But to say that a world contains lions is to say that the existence of lions is part of its content. That is, it is the claim that if that world were actual, then it would be true that lions exist. To make the logical connection between (1) and (2) work, we need the following principle:

**Uniqueness of the Actual World.** Necessarily, one and only one world (namely, the actual world) is actual.

If the Uniqueness of the Actual World were necessarily true, then (1) and (2) would entail one another. If lions exist, then the actual world must contain the fact that they do because all worlds are maximal. The actual world, and every world for that matter, must either contain the existence or non-existence of lions, and given that there are lions, no world can be actual and contain their non-existence. However, if *actuality* is a fundamental property of worlds, Uniqueness of the Actual World would have to be a brute necessity.[1] Given that *actuality* is a fundamental property, there is no immediate reason that it might not characterize more than one world. Clearly, *actuality* can be exemplified by many things, since many propositions are actual. Since worlds are just propositions, without Uniqueness of the Actual World, there is no reason multiple worlds might not be actual. Therefore, we cannot explain Uniqueness of the Actual World in terms of the impossibility of two worlds being actual, and its truth must be a brute necessity. This is a cost of Simple Anti-Indexicalism because we should minimize the postulation of brute necessities, according to Ockham's Razor and Structuralism:

**PMeth 1.6 Sixth Corollary of Ockham's Razor:** Other things being equal, prefer the theory that posits the smallest class of metaphysical possibilities.

**PMeth 3 Structuralism.** Other things being equal, adopt the theory that explains metaphysical impossibilities in terms of the essential structure of things.

It is worth looking for an alternative to Simple Anti-Indexicalism.

If Simple Anti-Indexicalism is false, what could *actuality* be? Suppose that we assume two primitive necessities: (1) there is exactly one possible world corresponding to each way the world could be, and (2) things *correspond to* a way the world could be if and only if things *would be* that way if that world were actual. Things are in fact some way or other, and necessarily so. Consequently, there must be exactly one possible world that corresponds to the way things in fact are. We could identify a total way things could be with a set of propositions, namely, the propositions that would be true if things (as a whole) were that way.

Consequently, we could assume the following three theses:

**15.2T.2 Actual-Truth-Defined Anti-Indexicalism.** A possible world is actual if and only if it corresponds to the class of (actually) true propositions.

**Completeness of Actual Truth.** Necessarily, there is a unique, maximal class that contains exactly the propositions that are (actually) true.

**Existence and Uniqueness of Worlds.** Every maximal class of propositions corresponds to exactly one world (either possible or impossible).

These three theses jointly entail the necessity of the existence of a unique actual world. Completeness of Actual Truth requires that there necessarily is a class containing the true propositions, and Existence and Uniqueness of Worlds implies that exactly one world corresponds to that class. Since Actual-Truth-Defined Anti-Indexicalism defines truth in terms of correspondence with truth, it follows that there is necessarily one and only one actual world. One consequence of this: we won't be able to make use of the Many Worlds account of ontological vagueness, discussed in Section 12.2.2.

This account of actuality involves two primitive notions, namely, the actual truth of propositions and the *correspondence* relation between worlds and classes of propositions. What are we to do with this notion of 'actual truth'? We could take it to be a special case of the actual possession of a property by a thing. A proposition is actually true if the proposition is in actual possession of the property of truth. So far, this is a theory on which *actuality* is a fundamental property of worlds, since we are treating the *actual* possession of properties as a primitive, metaphysically fundamental fact.

### 15.1.2 Non-indexical reductions of actuality

One could argue that actual truth is not a variety of truth, or a way of being true, but simply truth simpliciter, truth without qualification. Potential or possible truth, in contrast, is not a kind of truth, but only the possibility of or potentiality for truth. So, one might argue that the thought behind Actual-Truth-Defined Anti-Indexicalism is a kind of reductionism that reduces actuality to the non-modal notion of truth (simpliciter). That is, the view is better expressed as Truth-Defined Anti-Indexicalism:

**15.2T.3 Truth-Defined Anti-Indexicalism.** A possible world is actual if and only if it corresponds to the class of true (simpliciter) propositions.

Though Truth-Defined Anti-Indexicalism is a form of reductionism about actuality, it is not Modal Indexicalism because it denies that all possible worlds are metaphysically equal. It is a form of Anti-Indexical reductionism. Plantinga (1974), for instance, seems committed to something like Truth-Defined Anti-Indexicalism, though he formulates his view in terms of states of affairs rather than propositions. He says that *obtaining* and *truth* are fundamental notions. *Possible obtaining* and *possible truth*, and *actual obtaining* and *actual truth* are to be understood in terms of *obtaining* and *truth* simpliciter (Plantinga 1974: 48–49). Adams (1974) identifies the possible worlds with classes of propositions and takes the actual world to be the class containing all and only true propositions. At least two Magical Abstractionists, then, are committed to a form of reductionism about actuality that looks very much like Truth-Defined Anti-Indexicalism.

Those who accept some version of the correspondence theory of truth, in particular, those who embrace Truthmaker Theory (2.1T/2.1A.1T/2.1A.1A.1T) in one form or another, would have good reason to be unsatisfied with taking *truth* as a fundamental, monadic property of propositions. They would instead want to think of truth as a property possessed by a proposition in virtue of the obtaining of the binary relation of *truthmaking* between something that exists and the true proposition.

We could then define the actual world as the one that contains the *existence* of absolutely everything, and according to which nothing further exists. A non-actual world can fail to be actual either by not including the existence of everything or by including the existence of something that doesn't exist, or both. Worlds that don't include everything are *quantificationally deficient*. Worlds that contain something extra are not *reality-bounded*.

**Def D15.1 Quantificational Deficiency.** A world *w* is *quantificationally deficient* if and only if something would not have existed, had *w* been actual.

**Def D15.2 Reality-Boundedness.** A world *w* is *reality-bounded* if and only if there are some things that would have been (collectively) the totality of all that would have existed, had *w* been actual.

To be clear about what we're suggesting here, we need to look back to the discussion of Actualism (12.1T) and Anti-Actualism (12.1A). We're not saying that a quantificationally deficient world *w* contains things that both would have and would not have existed according to *w*. Rather, we're saying that *something* (in effect, quantifying over the things in the actual world) would not have existed had *w* been actual. Similarly, we are not saying that in a non-reality-bounded world *w* there is a class that contains more than everything (in world *w*). Rather, we're saying that, had *w* been actual, there would have been things that would not be identical to anything (in the actual world). We have to understand the 'something' and 'no things' in Def 15.1 and Def 15.2 as quantifying over all of reality.

Using these definitions, we can define actuality in terms of existence:

**15.2T.4 Existence-Defined Anti-Indexicalism.** The actual world is that unique possible world that is reality-bounded and not quantificationally deficient.

Actualists should favor Existence-Defined Anti-Indexicalism, while Anti-Actualists must reject it, since Anti-Actualists believe that some things don't exist (see Section 12.1). Existence-Defined Anti-Indexicalism also presupposes that there are no merely possible worlds agreeing exactly with the actual world with respect to what exists. This will be true only if we adopt Classical Truthmaker Theory (2.1T), rather than Non-Classical Truthmaker Theory (2.1A.1T) or Truth Supervenes on Being (2.1A.1A.1T).

What can Actualists say about worlds that are not reality-bounded? Again, these are worlds that contain the existence of everything, but that, in addition, contain the existence of things that don't actually exist, like twenty-first-century dodos or children of Pope John Paul II. Actualists cannot say that *some things* exist according to those worlds that do not exist in the actual world. For in saying that, Actualists would commit themselves to the existence of things that do not exist in the actual world, and this is exactly what they refuse to do! Instead, they must say that, if those worlds were actual, there *would have been things* in existence that do not in fact exist.

Anti-Actualists, it turns out, can use existence as the basis for actuality. They could suppose that actual existence is existence *simpliciter*, while merely possible existence is merely the possibility of existence. The actual world could be that world according to which exactly those things exist that exist simpliciter. This is Existence-Simpliciter-Defined Anti-Indexicalism

**15.2T.5 Existence-Simpliciter-Defined Anti-Indexicalism.** The actual world is that unique world $w$ such that, for every $x$, $x$ exists according to $w$ if and only if $x$ exists simpliciter.

Even if some things don't exist, we can still identify the actual world by finding the world that ascribes existence to exactly the right things (the things that really exist, that exist simpliciter). On this view, to exist possibly but not actually is not a kind of existence. The only way to exist is to exist actually. Merely possibly existing isn't a kind of existing, it's merely the attribute of the sort of thing that *might have* existed but doesn't. However, there is some plausibility to the idea that merely possible existence is a kind of existence. If so, actual existence would also be a kind of existence, and so we could not simply identify existence simpliciter with actual existence.

Nonetheless, we might take actual existence to be the *focal* meaning of 'existence'. On this view, the word 'exists' has a number of different but analogous meanings, just as the word 'health' does. We describe many things as healthy, from people to diets to urine. These different uses have different but interrelated meanings, and they are interrelated because they are anchored in a single, focal meaning, namely, that of being a healthy human being or organism of some kind. A diet is healthy because it promotes health in this focal sense; urine is healthy because it indicates health in the focal sense. Similarly, we might suppose that only actually existing things exist in the focal sense of the word. Merely possible things exist in some perfectly appropriate sense of 'exist', but not in the focal meaning of 'exist'.

What makes actual existence more fundamental than merely possible existence? Many philosophers, beginning with Aristotle, who introduced the distinction between the actual and the possible, have taken it to be obvious that the merely potential is always dependent on or grounded in what is actual, including the powers and tendencies of actual things.

**Priority of the Actual.** There are no metaphysically fundamental truths about what is merely potentially the case. All truths about mere potentiality depend metaphysically upon truths about what is actually the case.

If the Priority of the Actual is correct, then a better approach for the Anti-Indexicalists looking for a reductive account of actuality would be to appeal to the class of *fundamental* truths. Given the Priority of the Actual, we could define the actual world as that world containing all fundamental truths.

**15.2T.6 Fundamental-Truth-Defined Anti-Indexicalism.** The actual world is that unique world $w$ such that every fundamental truth is true in $w$.

This view of actuality has the advantage that it will work for Abstractionists whether or not they are Actualists. Further, given that Fundamental-Truth-Defined Anti-Indexicalism is motivated by the Priority of the Actual, it is plausible that Fundamental-Truth-Defined Anti-Indexicalists will want to be reductionists about not just actuality, but modality more generally. If we want to ground facts about actuality in fundamental truths, and facts about possibility (a kind of potentiality) in what is actually the case, then we should ground facts about possibility and necessity in fundamental truths as well.

Whether Fundamental-Truth-Defined Anti-Indexicalism is plausible depends on the account it can give of the truth of propositions of the form 'Possibly *p*'. Suppose that it is actually the case that *p* is possibly true. Fundamental-Truth-Defined Anti-Indexicalists must then argue that the possible truth of *p* is somehow grounded in the actual existence of something, something whose existence is contained by the actual world. In fact, reductionists about actuality who are committed to Truthmaker Theory must say that the possibility of every possible world, other than the actual one, is grounded in some fundamental truth contained within the actual world, and that the actual world is unique in this respect. No other world contains every truth that is fundamental.

Here again, Fundamental-Truth-Defined Anti-Indexicalists must appeal to what is fundamental *simpliciter*. Any possible world could have been actual, and so any world could have been fundamental. However, only one world is in fact fundamental, in the sense that every fact about other worlds is grounded in some truthmaker that exists according to the actual world. The actual world is (in fact) unique in this respect.

The most plausible approach for Fundamental-Truth-Defined Anti-Indexicalists to take grounds all propositions about mere possibilities in truths about the *powers and tendencies* of actual things. We'll take this issue up in Section 15.3 below, in which we develop an Aristotelian theory of modality.

### 15.1.3 A dilemma for Magical Abstractionism

David Lewis (1986a) raised a powerful objection to Magical Abstractionism. If propositions represent ways that things are, then the way things are "selects" certain propositions. The proposition that grass is green represents that grass is green, and since grass really is green, the way things are (which includes the fact that grass is green) selects the proposition that grass is green. Magical Abstractionists, given the Uniqueness of the Actual World, maintain that one possible world is uniquely selected by the universe (where the universe is here understood very broadly so as to include not just us and our physical surroundings, but any immaterial and abstract objects as well). The world that is selected in this way is thereby the actual world. Selection occurs because the actual world represents just what the universe is like. Had things been different, the universe would have selected a different world. So far, so good. Now, consider whether *selection* an internal or an external relation. That is, is which world is selected determined by the intrinsic characters of the universe and the actual world (in which case selection is internal, see Def 2.2) or is it not so determined (in which case selection is external)? Lewis believed both answers are problematic and, therefore, that Magical Abstractionism faces a dilemma.

Suppose that selection is an external relation. Thus, whether it holds between the universe and a world is not a function of the intrinsic features of the way things are and of possible worlds. Lewis thinks that, if this view of selection is true, then the selection relation must be "magical". What did he mean by this? Everyone agrees that worlds, and propositions more generally, have their representational features essentially. The actual world, Alpha, could not represent things as being a different way than it in fact represents them as being any more than the proposition that grass is green could fail to represent that grass is green. (This is just Essentiality of World-Content from Section 14.2.) This point is almost too obvious to state. It would be strange indeed to think of the proposition that grass is green representing that grass is blue, or worse, that there are exactly 17 electrons in the universe. One would be right to wonder what a person meant by "represent" if they were to make such a suggestion! So, what a world represents is essential to it. But selection is just the dual of representation; the two go hand in hand. Propositions are selected if and only if what they represent in fact occurs. Thus, a certain world *must* be selected, given that things are are a certain way, and given the representational features of worlds. What is the problem? According to Lewis, the problem is that every uncontroversially external relation is contingent. Consider spatial relations, a paradigmatically external relation. Just given the fact that two things stand in a certain spatial relation does not *guarantee* that they stand in that spatial relation. A necessary connection that isn't anchored in the intrinsic natures of the two relata would be a brute, inexplicable necessity. Given that there are literally infinitely many worlds, not to mention propositions generally, the view that selection is external requires infinitely many brute, inexplicable necessities. This is a wild violation of Ockham's Razor (**PMeth 1.2** and **1.3**). Lewis goes a bit further, and suggests that if selection is both necessary and external, then the relation would be magical, because there is nothing to explain why it obtains between *this* pair of things and any other pair. Lewis's is insisting that no relation can be both external and necessary, in the above way.

Suppose, then, that selection is internal. And it seems that selection *is* an internal relation. Indeed, that is how we introduced it. Given the way things are, given the intrinsic character of the universe, Alpha could not fail to be selected. And had the universe been different, Alpha couldn't have been selected. We might think of it this way. Given that worlds must represent what they in fact represent, selection must be an internal relation. But if selection is internal, then it's a relation that holds in virtue of the intrinsic characteristics of worlds and the intrinsic characteristics of the universe. So worlds must exemplify a rich variety of intrinsic features. Presumably, these intrinsic features are just representational features. To simplify, consider the proposition that grass is green. This proposition represents that grass is green, which is to say, it has the representational property of *representing that grass is green*. In virtue of this representational feature, this proposition is selected if and only if grass is green. Magical Abstractionists, given that they say nothing more about these representational features, have here "danced around a tiny circle", according to Lewis (1986a: 178). We thought we were getting an account of the selection relation, and Magical Abstractionists have offered us representational features. But in order to understand what these representational features are, Magical Abstractionists point us back to selection. We cannot, therefore, come to understand what these representational features are without understanding the selection relation, and we cannot understand the selection relation without understanding what these representational features are. Thus, we have no way to truly understand the selection relation. If we are able to grasp it, Lewis says, *we* must have magical powers. Whether the selection relation is internal or external, Magical Abstractionism involves magic. And philosophical views oughtn't involve magic.

Lewis's objection presupposes a certain view about how it is possible to think about properties and relations: the Acquaintance Model. According to the Acquaintance Model, it is possible to think about a property or relation only if we are acquainted with the property in our experience, only if our experience presents things to us as having this property or relation.

**Acquaintance Model.** It is possible to have thoughts about a property *P* only if one's experience sometimes includes the appearance that some things have *P*.

Lewis takes it as obvious that our experience never includes any appearances involving abstract possible worlds.

Even if we were to accept the Acquaintance Model, its relevance to this case depends on how we are to understand the notion of experience. If experience is limited to something like sensory experience, then it is certainly true that we never experience mere possibilities as such. However, it seems plausible to say that we also have intellectual experiences and that in those experiences certain abstract entities, like numbers, pure sets or possible worlds, do appear to be a certain way to us. If so, Magical Abstractionists can argue that a possible world could appear to us to represent the existence of golden mountains or talking donkeys, either by virtue of its intrinsic character or by virtue of an external representation relation.

In addition, Peter van Inwagen (2001: 238–242) offers a *tu quoque* response to Lewis (a response that shows that Lewis's overall theory is subject to the same objection Lewis lodges against Magical Abstractionism). According to van Inwagen, Lewis is also committed to the reality of "magical" representation. The problem is that Lewis accepts standard set theory. Consider the unit set or singleton whose one and only member is the Eiffel Tower. This set *represents* the Eiffel Tower, since by its very intrinsic nature it is such as to "select" the Eiffel Tower and only the Eiffel Tower as a member. However, it is highly implausible to think that sets represent their members in anything but a primitive way. Sets are supposed to be simple, structureless things. The set membership relation is either magical, or we require magical powers to grasp it. Either way, Lewis is subject to the same dilemma that he puts to Magical Abstractionists.

## 15.2 Structural Abstractionism

If one is dissatisfied with Magical Abstractionism's silence about the way that worlds represent but wants to avoid Concretism, it seems that the best account of how worlds represent will have something to do with the way that worlds are "put together". In other words, one will want to attribute to worlds a certain structure, where that structure will explain how it is that different worlds represent different things. This is Structural Abstractionism:

**15.1T.1A Structural Abstractionism.** Possible worlds represent what they do in virtue of their internal structure.

What sorts of structures might do the trick? An initial strategy is to identify common objects that represent the world somehow, and discern how it is that those objects do their representing. One then builds worlds on analogy with those common representational objects. Two sorts of things spring to mind: linguistic items like sentences and stories, and pictures of various sorts like paintings and photographs. Linguistic and Pictorial Abstractionism are respectively attempts to model the way worlds represent on analogies with linguistic items and pictures.

**15.1T.1A.1 Linguistic Abstractionism.** Structural Abstractionism is true, and worlds represent in the way linguistic things do, namely, by having components which represent things and which are arranged in according to a kind of "grammar".

**15.1T.1A.2 Pictorial Abstractionism.** Structural Abstractionism is true, and worlds represent in the way pictures do, namely, by having components that literally share features with what they represent.

We examine these views in turn, after a comment about whether Structural Abstractionism can supply a reductive account of modality.

We have seen that Magical Abstractionism is committed to the irreducibility of modal facts, though there are ways of being a Magical Abstractionist while holding that actuality in particular is reducible. The question that will occupy us here is whether Structural Abstractionists might embrace the reducibility of all modal facts. There is some reason to think that it might. Representation was modal according to Magical Abstractionism precisely because no account of it was on offer. Maybe a structural account of representation can supply the resources to reduce all the other recalcitrant modality as well. The key, it seems, is to find a reductive account of the notion of a maximal possible proposition. Recall that proposition $p$ is maximal if and only if for any proposition $p'$, either $p$ entails $p'$ or $p$ entails not-$p'$. There isn't anything *obviously* modal about that. Further, we might try to give the following reductive account of what it is for a proposition to be possible: proposition $p$ is possible if and only if $p$ does not entail a contradiction (that is, if and only if $p$ does not entail some proposition and its negation). Again, nothing *obviously* modal there.

However, entailment is modal. Proposition $p$ entails proposition $q$ if and only if *necessarily*, if $p$ is true then $q$ is true. One might try to get around this by saying that $p$ entails $q$ if and only if $p$ and not-$q$ are inconsistent. But again, inconsistency is an implicitly modal notion. Two propositions are inconsistent if and only if it is *possible* to derive a contradiction from them, or if and only if it is not *possible* for the two to be true together.

It seems, therefore, that Structural Abstractionism, like Magical Abstractionism, is best understood as embracing fundamental modal truths.

### 15.2.1 Linguistic Abstractionism

Linguistic Abstractionism is the view that worlds represent in the way linguistic things do. That is, worlds are something like sentences or very long and detailed stories. But how do linguistic items represent? Consider a sentence like 'Elsie is a dog.' This sentence represents Elsie's being a dog, or represents Elsie's exemplifying the property of *being a dog*. It does this by having a word, 'Elsie', which represents Elsie, and a string of words, 'is a dog', which represents the property of *being a dog*, and arranges those parts according to a grammar which represents that the thing represented by the one thing exemplifies the thing represented by the other thing. Sentences, then, contain parts (words and phrases) that represent certain things in the world, and are arranged according to a grammar that represents a certain arrangement of those things represented by those parts. Worlds, then, according to Linguistic Abstractionism, are built up out of parts that stand for things like words do, and which are put in certain quasi-grammatical arrangements, a propositional "grammar", that represent the arrangements of those represented things. *That* is how worlds represent.

Importantly, one needn't subscribe to the view sketched above about how sentences represent in order to embrace Linguistic Abstractionism. *Absolutely everyone* has to say something about how words and phrases and grammars contribute to the meanings of sentences, and Linguistic Abstractionists simply want to say that the way propositional elements contribute to the representational features of propositions, and therefore of worlds, need be no different from one's favored view of linguistic representation.

What sorts of parts one opts for needn't detain us much at this point. But David Lewis recommends that Linguistic Abstractionists opt for a "Lagadonian" language (an allusion to Swift's *Gulliver's Travels*). In a Lagadonian language, the name for each thing is that thing itself. So, the proposition that Obama is a Democrat contains both Barack Obama himself (as the proposition's name for Obama) and the property of *being a Democrat* (as the proposition's predicate for having that property). Importantly, Linguistic Abstractionists have traditionally been Actualists, and so will use only actually existing things to build their worlds.

*15.2.1.1 Combinatorialism.* A very simple variety of Linguistic Abstractionism takes possible worlds to be mathematical combinations of basic individuals and natural properties. (To see why this can plausibly be considered a version of Linguistic Abstractionism, recall the Lagadonian language noted above.) This sort of view corresponds to an account of propositions developed by Ludwig Wittgenstein in the *Tractatus Logico-Philosophicus* (1922/1961). Consequently, this account of possibility is often called 'Tractarian'.

**Def D15.3 Tractarian Propositions.** A *Tractarian proposition* is an ordered *n+1*-tuple consisting of an *n*-ary natural (fundamental) property and *n* basic individuals.

**Def D15.4 Tractarian Worlds.** A *Tractarian world* is a set of Tractarian propositions (intuitively, precisely those Tractarian propositions that would be true if the corresponding possible world were actual).

**15.3T Combinatorialism.** Every Tractarian world corresponds to a possible world.

**15.3A Anti-Combinatorialism.** Some Tractarian worlds correspond to no possible world.

Combinatorialists believe that every Tractarian world represents a real possibility. That is, every mathematically possible assignment of basic *n*-ary relations to *n* particular things could correspond to a way things could be. If I'm a particular thing and *being* *an electron* is a basic property, then there is a possible world in which I'm an electron. If *being red* and *being green* are basic properties, then it's possible for something to be both red and green.

Combinatorialists owe us an account of what to do with Tractarian worlds that contain no proposition involving some actual individual or some actual property. Suppose there is a Tractarian world that includes no proposition involving Socrates. Should we understand the Tractarian world as corresponding to the possibility in which Socrates does not exist or in which Socrates exists but has no natural properties? If the Tractarian world contains no proposition involving some natural property (like *being negatively* charged), should we understand it as corresponding to a world in which negative charge does not exist, or one in which it exists but has no instances? Can individuals exist without natural properties? Can natural properties exist without instances? The most natural way of taking the Tractarian picture would be the one Wittgenstein himself endorsed, according to which each basic individual and each natural property exists in all possible worlds. Every basic thing is a necessary being (see Skyrms 1981).

Presently, we sketch very briefly arguments for and against Combinatorialism.

### ARGUMENTS FOR COMBINATORIALISM

1 Combinatorialism offer a simple, reductive account of possibility. Whether or not something is possible is just a mathematical question: does a structure of the right kind exist?
2 Combinatorialism minimizes the class of metaphysical necessities. There are no necessary connections between distinct individuals or natural properties. This argument appeals, of course, to Ockham's Razor (**PMeth 1.4**).

### ARGUMENTS AGAINST COMBINATORIALISM

1 Combinatorialism cannot rule out certain impossible states of affairs. For example, consider color or charge. It is impossible to be red and green all over at once. Similarly, every charged particle must have one of either positive or negative charge, and none can have two charges at once. Combinatorialists have no basis for any of these necessities because they countenance no necessary connections among distinct natural properties.
2 It's implausible to suppose that everything exists necessarily. We can imagine many things not existing at all. For example, we can imagine a Big Bang that results in far fewer fundamental particles than the actual Big Bang produced. Combinatorialism is committed to counting this as absolutely impossible.

   Furthermore, it certainly seems possible that there might have existed things and properties that don't actually exist. The Big Bang could have produced more particles and even more basic kinds of particles than it actually did, but Combinatorialists must deny this.
3 The motivation for Combinatorialism, as illustrated by David Armstrong, involves a strong aversion to brute necessities of any kind. All necessities are supposed to be purely logical or conceptual in character. However, the attempt to avoid all brute

necessities seems to be quixotic. Even the Tractarian Combinatorialists admit some inexplicable necessities. For example, why do two-place relations, like *greater-than* or *part-of*, require two relata in order to make a complete possible fact? Why couldn't there be just one relatum, or any number of relata? Why does every possible fact require both one natural property and one or more basic individuals? Why couldn't a complete fact contain just properties or just individuals? Once we admit that there are some brute necessities, it seems reasonable to be open to further necessities, so long as these are needed to fit the theory to the rest of the data.

4 Like other versions of Linguistic Abstractionism, Combinatorialism faces the problem of alien possibilities, to which we turn presently.

*15.2.1.2 The problem of alien possibilities.* Linguistic Abstractionists face the problem of alien possibilities. They have a problem providing the sentences needed to represent certain possibilities involving non-actual items. Consider a Max Black world, consisting of two indiscernible steel balls, neither of which actually exists. The sentences representing this world cannot be in the pure Lagadonian language, since there are in fact no steel balls to act as names of themselves. We'll have to introduce some non-Lagadonian names, by pressing some abstract objects into service as names. However, the connection between any such name and its referent will be merely conventional or arbitrary. How then can we make each name refer to a different ball, when there are in fact no balls to refer to? How can we ensure, in other words, that if the sentence were true, that it would then designate one determinate possibility rather than another? In technical terms, Linguistic Abstractionists *conflate* what seem to be distinct possibilities into a single class of sentences, since there is no way for such sentences to distinguish between things that don't actually exist but would have been distinct, had they existed.

This conflation problem arises in other cases as well. Consider a second thought-experiment involving a Nietzschean world in which the same history endlessly repeats itself, but with different entities playing the historical roles in each epoch. In such a world, there would be a different person in the Nietzsche-role on each repetition. We can use an infinite class of sentences to represent such a possibility, but we cannot distinguish certain permutations of such a world. For example, suppose the Nietzsche-role occupier from the first epoch changed places with the Nietzsche-role occupier from the second epoch. This seems to be a different possibility from the one in which they are in their original epochs, but Linguistic Abstractionists cannot distinguish the possibilities. Since neither Nietzsche-role occupier actually exists, a Lagadonian language is inadequate, and there is no way to ensure that the "name" in a non-Lagadonian language which is putatively about the first Nietzsche-role occupier uniquely picks out that thing rather than the second, and vice versa.

Similarly, suppose you think that there could have been *alien* properties, properties that don't exist and are not instantiated in the actual world, like a third charge, in addition to *positive* and *negative*. These alien properties are like the Max Black spheres in that a Lagadonian language is too limited to name them, and a non-Lagadonian language will conflate possibilities involving such alien properties.

There are four solutions to the problem of alien possibilities that have been offered by Abstractionists. First, one could postulate the existence of uninstantiated haecceities or thisnesses (Def 9.2). These are properties that would, if they were instantiated, ensure that some object is identical to a certain thing. Haecceities are meant to exist necessarily, even though the objects (if any) to which they belong do not. Plantinga (1974) defends the existence of haecceities. Second, one could assume that the very same things exist in every possible world. On this view, everything exists necessarily, and nothing could exist except what actually does exist. As we have seen, this is the position of all Combinatorialists.

Third, one could assume that the actual world is special in that every possibility involving things—every *de re* possibility (see Section 16.1)—involves things that exist in the actual world, but that there are possible worlds that are ontologically impoverished in the sense that they do not contain enough objects to represent all the *de re* possibilities. This is the view of Kit Fine (1985), Robert M. Adams (1979) and others. Plantinga has labeled this position, with tongue in cheek, as 'Existentialism'. (No connection with Sartre or Camus.) Fourth, one could suppose that many things do not exist. That is, one could go for Anti-Actualism, whether Possibilist (12.1A.1T) or Meinongian (12.1A.1A). These non-existing things can be parts of possible worlds in which they would exist (if those worlds were actual) (see Section 12.1).

Each of these solutions has its drawbacks. Existentialists have complained that uninstantiated haecceities (the first solution) are a strange or even incomprehensible kind of property. Robert Adams believes in thisnesses, but he supposes that the property of *being identical to A* can exist only if A itself actually exists. So, according to Adams, there is a thisness for Pope Francis, but no thisnesses for any of his possible but non-actual grandchildren.

A second problem for the first solution is that it seems to make it impossible to have genuinely *de re* propositions. Every apparently *de re* proposition is really about, not the thing itself, but its haecceity. If one thinks that Socrates could have avoided the death penalty, then one is really thinking that Socrates' haecceity (*Socrateity*) could have been co-instantiated with the property of *avoiding the death penalty*. In effect, advocates of haecceities turn all *de re* thoughts into *de haecceitate* thoughts, thoughts about a haecceity. This has some severe consequences about what thoughts we can think. Suppose, for example, that we want to assert the following:

(3) Socrates doesn't instantiate any haecceity.

According to Plantinga, this sentence must express the same proposition as the absurd (4):

(4) Socrates's haecceity is co-instantiated with the property of *not instantiating any haecceity*.

Similarly, when Plantinga asserts (5), he seems to be asserting something substantive, not the trivial proposition expressed by (6):

(5) Whatever instantiates Socrates' haecceity is necessarily identical to Socrates.
(6) Whatever instantiates Socrates' haecceity necessarily instantiates Socrates' haecceity.

Thus, the first solution comes at a significant cost in terms of what genuinely *de re* propositions can be formulated.

As we have seen in discussing Combinatorialism, the second solution is problematic because it seems that some things exist only contingently, and that there could have been other things than those that actually exist.

The main difficulty for the third solution is that it is hard to believe that we are so lucky as to exist in one of the relatively few special worlds that contain everything that enters into every *de re* possibility. Christopher Menzel (1991) has proposed a way of modeling alien possibilities that involves using abstract objects to act as the substitutes or surrogates for the missing entities that exist in other possible worlds but not in the actual world. This can certainly be done, since there are plenty of such abstract objects around. However, Abstractionists must believe that possible worlds are identical to, not merely represented by, Menzelian representations. So Menzel's theory must postulate a deep, metaphysical difference between *de re* possibilities involving only actually existing things and alien *de re* possibilities.

With respect to the fourth solution, we've already discussed the pros and cons of Anti-Actualism in Section 12.1, so we won't repeat them here.

Finally, Combinatorialists must simply deny that alien possibilities are really possible. No individual could exist other than the individuals that do exist, and no property could exist other than the properties that have actual instances.

### 15.2.2 Pictorial Abstractionism

Pictorial Abstractionism is the view that worlds represent in the way pictures do. That is, worlds are like photographs or maps. But how do photographs and maps represent? Consider a picture of a family of four, with a dad, a mom, a son, and a daughter. To make things easier, give them names: Tim, Jamie, Lyle, and Gretchen, respectively. How does the picture represent Tim? By having a part, the Tim-part, with features like Tim's, a red part that represents Tim's beard, a blue part that represents his shirt, and a white part that represents his shoes. It represents Tim as taller than Jamie by having the Tim-part be larger than the Jamie-part. It represents Lyle's blue eyes by having two blue parts in the head part of the Lyle-part. And it represents Gretchen's impish smile by having an impish-smile part of the head part of the Gretchen-part. Pictures, then, represent by having parts whose features match the features of the things represented, where those parts are arranged in ways that match the arrangements of the things represented.

Worlds, then, according to Pictorial Abstractionism, represent by being "built up" out of parts which have certain features and which are arranged in a certain way, where those parts and their features and arrangements match the features and arrangements of the things represented. *That* is how worlds represent according to Pictorial Abstractionism:

**15.1T.1A.2 Pictorial Abstractionism.** Structural Abstractionism is true, and worlds represent in the way pictures do, namely, by having components that literally share features with what they represent.

*15.2.2.1 The problem of representational detail.* The problem of alien possibilities plagues Pictorial Abstractionism no less than Linguistic Abstractionism, and for the same reasons. Pictorial Abstractionists build up propositions, in a picture-like way, from actually existing items. That these items are combined in a picture-like way, rather than a language-like way, is inessential to developing the problem of alien possibilities.

But there is a further problem for Pictorial Abstractinoism: the problem of representational detail. The problem is that, in order for the Pictorial Abstractionists to have their pictures represent every feature of a world in maximal detail, worlds must turn out to just be Concretist worlds. Recall the picture of Tim, Jamie, Lyle, and Gretchen, and the impish-smile part in the head part of the Gretchen-part of the picture, which represents Gretchen's impish smile. What is it for a picture to have an impish-smile part? In the case of representing the color of a thing, it's not hard to see how the relevant part of two-dimensional picture can literally just *have* the property it represents. Likewise with relations among the sizes of things. A picture can represent that Jamie is twice the height of Lyle by having the Jamie-part be twice the size of the Lyle-part. But it's odd to think of a part of a picture as literally just *impish*. It is, anyway, unless we're talking about four-dimensional pictures or models that involve mental properties as well. Pictorial Abstractionists must therefore think of their worlds as four-dimensional, in some sense. And this will be required more generally, though, to represent temporal relations. And now we're getting ourselves into trouble. Worlds, on this view, are four-dimensional objects literally exemplifying all of the properties of the objects they represent. How could these worlds be abstract? They are colored, shaped, spatially and temporally extended objects. That's just a Concretist world!

Pictorial Abstractionists might respond by trying to identify a feature that Concretist worlds have that their worlds don't, or vice versa. The feature is going to have to simply be the property of *being abstract*, but it's not clear what content we can give to that property at this point. Abstract things like numbers and sets aren't colored or shaped in any meaningful sense. Only concrete things like people and rocks have those sorts of properties! Once we give worlds all of the representational detail they need, they morph into Concretist worlds. We think this is a serious, potentially lethal worry for Pictorial Abstractionism.

## 15.3 Aristotelian Theories of Possibility

If we reject the theories we've encountered to this point, what are we left with? There is one other reductive account to consider, an account which grounds possibility in the actual possession of causal powers. Alexander Pruss (2002, 2011) has recently defended such a view, Aristotelian Modality. The basic idea is that a proposition is possible if there is some actually existing thing that has the power to bring about the truth of that proposition. More precisely:

**15.2T.7 Aristotelian Modality.**

(i) It (tenselessly) is possible for $p$ to be the case if and only if there is, was or will be a time at which it is, was or will be possible for $p$ to be the case;
(ii) it is now *simply* possible for something to exist or to fail to exist if and only if something now has the power to cause or to prevent its coming into existence;

(iii) it is now *simply* possible for some things to stand in some natural relation if and only if something now has the power to make those things to stand in that relation;

(iv) it is now *possible* for $p$ to be the case iff either it is now simply possible for $p$ to be the case or it is simply possible for it to be simply possible for $p$ to be the case or it is simply possible that it is simply possible that it is simply possible that $p$, and so on.

On the Aristotelian account, the fundamental modal notion is that of *being possible at a time*. *Tenseless possibility* is derivative. Thus, what is possible changes as time passes. In fact, fewer things are possible over time, as the opportunities for things to exercise their powers pass. If we want Aristotelian Modality to line up with the standard practice in contemporary modal logic, according to which everything that is or was actually the case is possible, then we will have to modify clause (i) of Aristotelian Modality to include present and past actuality as cases of tenseless possibility:

> (i) It (tenselessly) is possible for $p$ to be the case if and only if there is, was or will be a time at which it is, was or will be possible for $p$ to be the case, *or it is now or once was the case that p*.

*Possibility* for the Aristotelian is an *iterative* or *recursive* concept, as clause (iv) of Aristotelian Modality makes clear. That is, we define first a base case of possibility, *simple possibility*, and then permit something to be possible simpliciter if it can be reached by repeated applications of clause (iv). This is a standard way of defining operations in mathematics (for example, addition and multiplication can be defined this way).

One way for something to be possible is for actual things to have the power to make it so. Another way would be for actual things to have the power to give something the power to make it so. A still further way would be for actual things to have the power to give something the power to give something the power to make it so, and so on. Ultimately, all possibilities are grounded in the actual powers of actual things at some point in history. Aristotelian Modality, therefore, validates the Branch principle that Shoemaker used to defend Causal Structuralism (a variety of Strong Powerism 4.4A.3):

**Branch.** For every possible world $w$, there is a time $t$ such that $w$ and the actual world are exactly alike up until time $t$.

Since the notion of power is central to Aristotelian Modality, we get sub-varieties of the view depending on one's metaphysical account of powers. It is possible to combine Aristotelian Modality with Strong Nomism (4.4A.2), Strong Powerism, and Neo-Humeism (4.4T). You can't be an Aristotelian and a Strong Hypotheticalist (4.4A.1), since Strong Hypotheticalism entails that *possibility* is a fundamental notion. Strong Hypotheticalists are committed to fundamental modal truths, since conditionals are modal. The most natural combination would be to join an Aristotelian Modality with Strong Powerism, especially with a Dual-Aspect Theory (Section 6.2.2). This would seem to have been Aristotle's own position. It is natural because there is a connection between conditionals and laws, and modality more generally. Strong Powerism reduces conditionals and laws to powers, and Aristotelian Modality reduces modality to powers. The two views, therefore, couple nicely. Further, by combining the two, one makes gains in the realm of simplicity (PMeth 1), especially with respect to qualitative economy (PMeth 1.4.1), since one can use the fundamental notion of a power to account for conditionals, laws, and modality.

ARISTOTELIAN POSSIBILITY AND NEO-HUMEISM We should consider, however, the option of combining Aristotelian Modality with Neo-Humeism. Neo-Humeists, unlike Hume himself, believe in causal powers; they just believe that truths about causal powers are grounded in more fundamental truths about the patterns and regularities in the spatiotemporal distribution of qualities. There is one serious point of tension between Neo-Humeism and Aristotelian Modality, a point that concerns the Branch principle. Suppose the past is finite, with a very first moment of time. Then the Branch principle entails that all modality is ultimately grounded in the powers of the things that existed in the first moment of time. However, there were, as yet, no interesting diachronic patterns of qualities in place at that first instant. The simplest scientific account of the world at that moment would have included no dynamic laws at all. Consequently, Neo-Humeists must deny that anything had any power in the first instant, which means that no possible world could branch off from the actual world then. It would only be after enough time has passed to bring about scientifically relevant regularities that anything could be said to have any power.

This result is quite bizarre. If we take the finitude of the past to be a real possibility, then Neo-Humeists must confront the small worlds objection that we considered (Section 5.2.2). At the very beginning of history, the universe would have been exactly the sort of small world that puts such pressure on Neo-Humeism. It seems that there should be a fact of the matter about what would or would not have been possible from the very beginning, but Neo-Humeists can give no account of this.

Neo-Humeists could respond by embracing Eternalism (to be discussed in Chapters 20 and 21 as 20.2A.1T), according to which all of the world's future history exists (tenselessly) at all times, including the very beginning. However, this would mean that bodies existing at the beginning of time would have different and contrary sets of powers, depending on which future history we take as actual. In addition, trying to combine Neo-Humeism, Eternalism, and Aristotelian Modality would lead to a kind of vicious circularity: what is possible in the future could not be grounded without first determining what happens in the future, but what happens in the future must be at least possible from the point of view of the past.

## 15.4 Conclusion

In this chapter, we have canvassed a number of Abstractionist views of modality. As we have seen, there are plausible Abstractionist reductions of actuality, but Abstractionist entails that there are fundamental truths about possibility and necessity. Different Abstractionist views are usefully categorized on the basis of whether and how possible worlds represent. But each type of Abstractionism faces serious hurdles. In the next chapter, we turn to questions of *de re* modality, which concerns what is possible and necessary for objects.

## Note

1 Even if we thought, for reasons considered in Section 12.2.2, that there could be more than one actual world (due to ontological vagueness), there would have to be some severe limits on how varied the actual worlds could be, and this would again have to be a brute necessity for Simple Anti-Indexicalism.

# 16

# De Re Modality and Modal Knowledge

## 16.1 Modality *De Re*: Transworld Identity and Counterpart Theory

To this point, we have focused mainly on how possible worlds relate to the truth and falsity of modal claims (or propositions), and therefore to whether claims are necessarily true, necessarily false, possibly true, possibly false, and so on. This issue is that of modality *de dicto*, modality concerning propositions. But there is another type of modality, one toward which we gestured above, namely modality *de re*. This has to do with the modal status of relations between things and their properties, with whether things possess properties necessarily, contingently or not even possibly. Consider the following:

(1) If there is a largest solar planet, then, necessarily, it revolves around the sun.

Read in one way, this sentence expresses a truth. On this reading, (1) expresses the idea that the proposition expressed by the sentence, 'If there is a largest solar planet, then whatever is the largest solar planet revolves around the sun', is necessarily true. This seems to be the case, since it is part of the definition of 'solar planet' that any solar planet revolves around the sun. If a possible world contains no largest planet, then this proposition is vacuously true, and if it does contain a largest solar planet, then that planet must (in that world) revolve around the sun. Hence, the proposition cannot be false, and so (1) is true. This is the *de dicto* reading of (1), on which (1) attributes necessity to something that is said ('dicto' in Latin).

However, there is another way of reading (1), on which (1) says about *the thing* that is in fact the largest planet, namely, Jupiter, that *it* necessarily revolves around the sun. This is the *de re* ('about a thing') reading, a reading which attributes the necessary possession of a property (that of *revolving around the sun*) to a thing (Jupiter). Read this way, (1) is

*The Atlas of Reality: A Comprehensive Guide to Metaphysics*, First Edition.
Robert C. Koons and Timothy H. Pickavance.

false. If the sun were to have a close encounter with another star, Jupiter could be flung out of our solar system. Or, we could build a big enough rocket to shoot Jupiter into deep space. There's nothing about Jupiter the thing that necessitates its revolving around the sun.

Quine (1953/1980) argued that attributions of *de re* necessity are just nonsensical. According to Quine, the only sort of necessity is *de dicto*. For example, suppose that every brilliant mathematician in Austin is also an Olympic cyclist and vice versa. It is necessarily the case that every Olympic cyclist in Austin has two legs (let's assume that the kind of cycling involved in the Olympics requires the use of two legs), but it is not necessarily the case that every Olympic cyclist in Austin can prove theorems, even though the very same things belong to the two classes.

However, many contemporary metaphysicians, and most metaphysicians of the past, would disagree with Quine. There seem to be some properties that are necessarily possessed by certain things, independently of how they are described or designated. For example, Jupiter has the property of *being identical to Jupiter*, and it seems that it has that property in any possible world, no matter how it is described or what other properties it has or doesn't have. Some philosophers, such as Saul Kripke, have argued that the origin of a thing is an *essential* property of that thing, where an essential property of a thing is a property whose possession by it is a *de re* necessity. If RCK's actual parents are Bruce and Margaret, it seems plausible to think that he would have Bruce and Margaret as his parents in any possible world. Anything in any world with different parents wouldn't be *RCK himself*. For similar reasons, it seems plausible to think that it is *de re* necessary that if THP exists, then THP is human. Nothing could have been THP without belonging to his actual species. The question thus arises as to how it is that worlds represent things as having properties, since that question will be connected to questions about the essences of things.

The question that has dominated discussions of modality *de re* in recent philosophical work is whether one and the same object can exist in many possible worlds. Those who think that objects can exist in many possible worlds believe in Transworld Identity. Those who deny Transworld Identity believe in Worldbound Individuals, which is the idea that everything is worldbound, existing in only one possible world.

**16.1T Transworld Identity.** Some things exist in more than one possible world.

**16.1A Worldbound Individuals.** Each possible individual exists, as a matter of necessity, in one and only one possible world.

The most common version of Worldbound Individuals is Counterpart Theory (16.1A.1), which says that although individuals are worldbound, they nonetheless have *counterparts* in other worlds that represent them in those alternative possibilities. We discuss this view in Section 16.1.2 below. At any rate, there are disputes even about what it is for a single object to exist in more than one possible world. The differences to do with *de re* modality are driven in large measure by differences about the nature of possible worlds more generally, which drive differences about what it is for something to be *in* a world, or alternatively, differences in what it is for a world to *involve* some thing. In particular, Abstractionists tend to go for Transworld Identity, while Concretists tend to go for

Counterpart Theory. We examine these views in turn, and display how a commitment to Abstractionism naturally drives one toward Transworld Identity, while a commitment to Concretism naturally drives one toward Counterpart Theory.

### 16.1.1 From Abstractionism to Transworld Identity

Abstractionism naturally drives one toward Transworld Identity. To see that this is so, let's consider a natural Abstractionist characterization of modality *de re*. Consider the claim that Lyle is essentially human. This claim is true if and only if Lyle has the property of *being essentially human*. Being a modal claim, there is a connection between its truth and facts about possible worlds, even if the connection is not a reductive one. So, despite being anti-reductionist about modality, Abstractionists can agree that Lyle has the property of *being essentially human* if and only if every possible world that represents Lyle at all represents him as having the property of *being human*. But what is it for a possible world to represent Lyle as being a certain way? We considered three Abstractionist views of representation above, but there is something that they have in common in this area: almost every Abstractionist view agrees that, however Lyle gets represented, a world does *not* represent him by having him as a part.[1] Lyle is a part of no possible world (not even the actual one), in a strict sense. The only sense in which Lyle is part of a world is that he is represented by it. Lyle is literally part of the actual world if Concretism is true, but he is not literally a part of any world if Abstractionism is true. He is simply represented. The actual world represents him as he is; the others as he could have been.

There are differences, though, in *how* the Abstractionist views represent Lyle as being some way or other. If Magical Abstractionism is true, we can't say anything meaningful at all about how this happens. However, it will be the case that the worlds that represent Lyle have certain representational features in common. For example, every world that represents Lyle have the property of *representing Lyle*. (Uninformative, but that's to be expected!) Consider (2) and (3):

(2) Elsie is necessarily a dog.
(3) Joe Biden might have been a plumber.

Linguistic Abstractionists will say that (2) comes out true if and only if every world containing a "name" for Elsie contains a "sentence" which contains a "name" for Elsie "grammatically" associated in the right way with a "predicate" designating the property of *being a dog*. Similarly, (3) is true if and only if there is a world containing a sentence containing a name of Joe Biden grammatically associated in the right way with a predicate designating the property of *being a plumber*. Worlds represent Lyle by having a "name" of him as a constituent. Maybe, for example, a world represents Lyle, Elsie, and Joe Biden by having their respective haecceities as parts. If Pictorial Abstractionism is true, then (2) is true if and only if every world-picture has no part that is both an Elsie-part and a non-dog-part, and (3) is true if and only of some world-picture has a part that is both a Joe-Biden-part and a plumber-part. Worlds represent Lyle by having a highly detailed, four-dimensional, abstract simulacrum of Lyle as a part.

Notice, though, that on all Abstractionist views, Lyle is represented by two worlds $w_1$ and $w_2$ if and only if $w_1$ and $w_2$ literally have something in common. In the case of Magical Abstractionism, there is sharing of a property; in the case of Linguistic and Pictorial Abstractionism, there is sharing of a part, whether a name or an abstract simulacrum. Because of this, it is right to say that, had a world representing Lyle been actual, Lyle himself would have existed. That is, he would have been part of our spatiotemporal surroundings. In this sense, these views are committed to, or at least are compatible with, Transworld Identity, the claim that it is possible that actual objects would have existed even if a different world had been actual. One way to express the idea of Transworld Identity is as the thought that a single object can exist *in* many worlds. Given that the word 'in' can be understood either representationally or mereologically, we must be careful to say to what exactly Abstractionists are committed. Concretists no less than Abstractionists agree that many worlds represent a single individual. The distinctive commitment of Transworld Identity is the claim that it is possible for some actually existing thing to be a mereological part of the concrete universe even were a different world actual. The view may crystallize by considering its chief rival, Counterpart Theory.

### 16.1.2 From Concretism to Counterpart Theory

Concretists take possible worlds to be like parallel universes. If that it right, though, it would be very odd to say that a single thing could exist in multiple worlds, where 'in' is understood to mean 'be a part of'. The only way for this to work, given Concretism, would be for possible worlds to overlap one another, that is, for possible worlds to share a part. Here's an example. Suppose Joe Biden could have been a plumber. This is a possibility involving Biden himself, not someone else. So this very individual (namely, Biden) exists in a possible state of affairs in which he is a plumber. But he is not a plumber in the actual world. Therefore, Biden must exist in another possible world, too, one where he is a plumber. Call that world $w$. The actual world and $w$ must share a part, they must overlap, and the part they share is Biden. If this is the story the Concretist tells about *de re* modality, then we get Overlapping Concrete Worlds:

**16.1T.1 Overlapping Concrete Worlds.** It is possible for something to fail to exist if and only if there is a concrete universe in which *it* (and not just a counterpart of it) does not exist.

The idea is that we keep Lewis's thought that something exists in some world just in case that thing is a part of that world, and we add the thought that something can literally exist as a part of more than one world. Thus it would follow that you are a part of many worlds, and thus that worlds share parts in common. Worlds overlap, which is just to say that worlds share parts.

David Lewis pointed out a critical problem for the Overlapping Concrete Worlds. Suppose that there is an iron meteor that in the actual world is spherical and always has been spherical, but which might have been cubical instead. Such a meteor would then be part of two different possible worlds. If the two worlds are concrete, overlapping universes, then the meteor would have to be both spherical and cubical at once. It is cubical in one world and spherical in the other, but there is only the one thing that is part of the two worlds!

Such overlap would be possible only if the things in the overlapping region had exactly the same intrinsic properties in both worlds.[2] Certain abstract objects, such as universals, numbers, and pure sets (like the empty set) would pose no problem, since they have all of their intrinsic properties of necessity. However, concrete, changeable things are another matter. If a rod is in fact bent but could have been straight, then it would seem that it could not exist in two worlds, in one of which it was straight and in the other of which it was bent, since being straight and being bent are incompatible properties.

Concretists might imagine that possible worlds can overlap at certain times and not at others. So, there could be two possible worlds that share the same history up to some point in time and that diverge thereafter. Perhaps there is a world in which Harold Godwinson defeated King William the Conqueror at the Battle of Hastings. Such a world might overlap our world up until that fateful day in 1066, diverging radically afterward. There might also be, in certain cases, converging worlds, worlds that share the same future but differ in the past, as well as worlds that repeatedly converge with and diverge from each other.

There is however, a decisive objection to such a branching-world version of Concretism: it is incompatible with Modal Indexicalism (14.2T). If worlds can converge or diverge, then there could be individual thoughts and speech acts that exist in multiple worlds. Consider again the thought expressed in Chapter 15's sentence (14):

(14) Dragons actually exist.

Suppose that, in fact, there never have been and never will be dragons, although there is a possible world diverging from actuality in the future in which geneticists construct dragons, or other possible worlds converging with the actual world some time in the past, in which dragons once existed. A use of (14) here and now could be neither determinately true nor determinately false, since it belongs simultaneously to worlds containing dragons and to other worlds not containing them. The word 'actually' would fail to act as a true indexical, since the thought has no definite location in logical space.

Hence, it seems that Concretists must embrace Worldbound Individuals. However, Worldbound Individuals pose an obvious problem for the Concretists: if each individual exists in only one world, then every individual in the actual world has all of its properties (both intrinsic and extrinsic) essentially. Nothing could have been different from the way it is, which seems wildly false.

Lewis offered Counterpart Theory as a solution to this problem. According to Counterpart Theory, individuals in one world can have one or more *counterparts* in other worlds. In most cases, a thing's counterpart in a world will be that thing in that world that most resembles it. So, although Socrates exists only in the actual world, other worlds contain philosophers whose biographies closely resemble the actual story of Socrates. Maybe they are named 'Socrates', live in a city called 'Athens', devote themselves to philosophical conversation, are forced to take hemlock for corrupting the youth, and so on. To say that Socrates *might have been* acquitted is to say that there exists a possible world containing a *counterpart* of Socrates who *is* acquitted. More generally, to say that some particular individual could have been otherwise is to say that it has a counterpart in some possible world that is (in that world) otherwise. This is Ludovician[3] Counterpart Theory:

**16.1A.1 Ludovician Counterpart Theory.** It is possible for something to fail to exist if and only if there is an isolated universe in which it has no counterpart, it is possible for something to have a property if and only if there is an isolated universe it which a counterpart of it has that property, and it is possible for some things to stand in some natural relation if and only if they have some counterparts that do stand in that relation in some isolated universe.

The *counterpart* relation needn't be one-to-one. Socrates has no counterparts in some worlds; these are worlds according to which Socrates doesn't exist. But he also might have several counterparts in others; these are worlds that make it true, say, that Socrates might have been twins. A thing might even have a counterpart in its own world. For example, imagine a world in which history endlessly repeats itself. Such a world would have, for example, an infinite series of "Nietzsches". Each of the infinitely many "Nietzsche"-like persons in such a world might well count as a counterpart of the others.

## The Humphrey Objection

Abstractionists have objected to Counterpart Theory on the ground that it makes *de re* modal facts have to do with things other than, for example, Socrates. But it's *Socrates* that might have been acquitted of corrupting the youth. What is true of the counterparts of a thing is irrelevant to what could be true of the thing itself. The fact that there exist beings similar to THP in certain ways who are in fact plumbers has nothing whatsoever to do with the question of whether or not *THP himself* might have been a plumber. Kripke (1980) famously pressed this objection against Ludovician Counterpart Theory, using an example about the possibility that Hubert Humphrey might have won the 1968 Presedential election. Here is Kripke:

> According to Counterpart Theory] if we say 'Humphrey might have won the election'…, we are not talking about something that might have happened to *Humphrey* but to someone else, a 'counterpart'. Probably, however, Humphrey could not care less whether someone *else*, no matter how much resembling him, would have been victorious in another possible world. (Kripke 1980: 45)

The worry has therefore become known as the 'Humphrey Objection' to Counterpart Theory.

Counterpart Theorists ought to say that, as stated by Kripke anyway, the objection is just misguided. Socrates himself really does have the property of *possibly being acquitted*, though he has this *because* a world represents him as having the property of *being acquitted* by having a counterpart of him who *is* acquitted. In other words, what is true of your counterparts has *everything* to do with what you yourself *might have been* like. *What it is* for you to have the property of *possibly being a plumber* is for you to have a counterpart that is a plumber. And *what it is* for you to have the property of *necessarily being a human* is for all of your counterparts to be human. But to emphasize, *you yourself* have the *modal* properties needed.

Importantly, there are at least two ways to read the Humphrey Objection in light of these observations. The first is that, if Counterpart Theory is true, *too many* objects are involved in accounting for a thing's modal properties. The second is that the *wrong* objects are involved in accounting for a thing's modal properties. On both readings, the Humphrey Objection accuses Counterpart Theory of problematically involving things that exist in other worlds in the truth-conditions for *de re* modal truths about actual objects. The first reading of the objection says that the problem is that Humphrey alone ought to matter. The second reading of the objection says that the problem is that other-worldly objects should not matter.

It isn't clear that Abstractionists can elude the objection when formulated in the first way, however. Abstractionists would agree that Socrates doesn't represent *himself*, since all of them (except for Lagadonian Linguistic Abstractionists) have worlds built up out of abstract representations or out of nothing at all. In this way, all parties agree that facts about the instantiation of modal properties involve or implicate representations that are not the objects themselves. Take Magical Abstractionism, for example. On this view, Humphrey might have won if and only if there is some world that represents Humphrey as winning. This account involves an abstract object which represents that Humphrey wins, and so in that sense there are objects other than Humphrey involved in the *de re* modal fact that Humphrey might have won, namely, possible worlds.

On some Abstractionist views, the similarities are even more acute. For example, suppose one thought that a world represents Humphrey by having Humphrey's haecceity, *H*, as a part. On this view, it is plausible that one ought to count *H* as a counterpart of Humphrey. *H* represents Humphrey in other worlds without being identical to Humphrey. Humphrey might have won if and only if there is some possible world which contains *H* as an element and the property of *winning* as an element in whatever way is required to represent that were the world actual, *H* would be co-instantiated with the property of *winning*. This sort of Counterpart Theory is not *Ludovician*, since it does not commit to the claim that counterparts are concrete objects that are very similar to the objects they represent. However, it could plausibly be construed as an Abstractionist-friendly sort of Counterpart Theory, since the structure of the view is very similar to Ludovician Counterpart Theory. There are objects in the universe with certain modal properties, and an object has its modal properties if and only if there is something non-identical to it that represents its existence and its having certain properties. The difference concerns the nature of the counterparts rather than whether there are such counterparts.[4]

The second reading of the objection, on the other hand, according to which the problem is the involvement of the *wrong* things, is the *de re* version of the problem of irrelevance for Concretism discussed in Section 14.3. To the degree that one finds that objection compelling, one ought to find the Humphrey Objection compelling as well. However, to the degree that one finds the problem of irrelevance unproblematic, one ought not be worried by the Humphrey Objection either.

Part of the problem arises from the fact that we can mean two very different things by 'existing in a world'. The first way is simply be being a part of that world. For Concretists, the most basic way for a thing to exist in a world (a way that has nothing to do with counterparts) is for it to have a property like the property of *existing in Philadelphia*: it is simply a mattter of *where* something concrete is located. For Linguistic Abstractionists, to exist in a world in this literal way is to be something like a symbol within the propositional language used to build up the propositions that constitute possible worlds. For Magical Abstractionists, nothing literally exists in a world: worlds are simply entities with no parts. Let's call this being in a world *mereologically*.

The second way of *existing in a world* is to exist *according to* a world. To use Plantinga's formulation (Plantinga 1974: 46–48): $x$ exists *in* world $w$ if and only if, $x$ would exist if $w$ were actual. This kind of existing in has nothing to do with location or parthood. We could call this being in a world *representationally*.

One important takeaway from this discussion is that Ludovician Counterpart Theory seems compatible with Transworld Identity where 'in' is understood representationally, while also being a version of Worldbound Individuals where 'in' is understood mereologically. Humphrey is not a *part* of other worlds, nor is he spatially located there, so in that sense he is not *in* them. However, Humphrey is *represented* by other worlds, by having counterparts that are parts of those worlds, and in that sense he is "in" them.

Perhaps we can capture the difference between the two accounts by using Kit Fine's notion of *ontological dependence*, which we discussed in Section 3.3 (Fine 1994b):

**16.1T.2 Strong Transworld Identity.** There is an object $x$ and distinct worlds $w_1$ and $w_2$ such that both $w_1$ and $w_2$ represent $x$ in part because each is ontologically dependent on $x$.

Counterpart Theorists will deny Strong Transworld Identity, since they will hold that if $x$ is an inhabitant of world $w_1$, then any distinct world $w_2$ will not include $x$ or any facts about $x$ intrinsically but will merely represent $x$ by virtue of similarities between $x$ and its counterpart in $w_2$: something's being similar to something external to is a paradigm of an extrinsic fact about it. In contrast, those who believe in Strong Transworld Identity suppose that $x$ itself must enter into the very constitution of the essence of any world that represents $x$ as being a certain way. Even if $x$ is not literally a part of the possible world, the essential properties of that world, the properties that make it a possibility according to which $x$ has certain properties, do *involve* $x$, in the sense that a full definition of their essence would include $x$. If a possible world is something like a book of sentence-like or map-like propositions, it will include its representational properties in its very essence, and so its essence will include any object, like $x$, whose possible properties are represented *de re* in that world. Counterpart Theorists, in contrast, will suppose that any representational properties of a world (with respect to inhabitants of other worlds) are entirely extrinsic to that world. For Concretists, what makes a possible world a world has nothing to do with its representational properties. And that might be the issue that the Humphrey Objection is pointing to.

The proponent of the Humphrey Objection might argue in this way: suppose that Biden is not a plumber but might have been one, and we're looking for the truthmaker for this *de re* possibility about Biden. Finding a world $w$ in which something is a plumber can be relevant to this task only if that world is ontologically dependent on Biden himself. Otherwise the fact that something is a plumber is $w$ could not even be a *partial* ground for the fact about Biden that he could have been a plumber (contrary to Counterpart Theory). What are the grounds for this claim? Some sort of reflection on our concept of *de re* possibility, perhaps.

The strength of this argument against Counterpart Theory obviously turns on the likelihood that Strong Transworld Identity is true. Before we turn to a worry about Strong Transworld Identity, however, we must explore how Abstractionists respond to the Leibniz's Law argument sketched above. One might think that it doesn't matter whether one is committed to Counterpart Theory. So long as Biden exists in more than one world, it can seem that one is committed to the claim that Biden both exemplifies and fails to exemplify the property of *being a plumber*. That's bad whether or not Counterpart Theory is true!

Importantly, Abstractionists do not think that Biden *actually* exemplifies the property of *being a plumber*, and Abstractionists all agree that the only Biden that exists in any sense is the actual Biden. Biden fails to exemplify the property of *being a plumber* but does exemplify the properties of *being represented by w as a plumber* and of *possibly being a plumber*. Further, had $w$ been actual, Biden himself would have existed and would have exemplified the property of *being a plumber* and the property of *being represented by α as not being a plumber*. Nothing identical to Biden actually exemplifies the property of *being a plumber*. That claim is not, according to Abstractionists, the denial of the claim that $w$-Biden is identical to $\alpha$-Biden. Abstractionists commit to the claim that $w$-Biden is identical to $\alpha$-Biden. The trick is that say that $w$-Biden actually exists and deny that $w$-Biden exemplifies the property of *being a plumber*! $w$-Biden *just is* Biden, and Biden doesn't exemplify the property of *being a plumber*. $w$-Biden (= Biden) *would have been* a plumber, had $w$ been actual. It is Biden himself who is represented by both $w$ and $\alpha$. The problem, then, lies in the move from the fact that $w$ represents Biden as a plumber to the claim that $w$-Biden exemplifies the property of *being a plumber*. That move only works if one is thinking as Concretists do, that possible worlds have as parts objects exemplifying all their properties like Biden himself. Since Abstractionists deny that this is so, they can elude the Leibniz's Law argument.

### 16.1.3 Strong Transworld Identity and mere haecceitistic differences

Suppose one thinks, with Plantinga, that Counterpart Theory is problematic because it entails the falsity of Strong Transworld Identity. Some have objected that this sort of view entails that there are "haecceitistic differences" between worlds. Two worlds are *merely haecceitistically different* if and only if they differ only with respect to the identities of objects in the two worlds, and not in any qualitative way.[5] Consider, for example, a world $w$ exactly like the actual world but where RCK and THP have switched roles. In $w$, RCK has all of THP's characteristics, and THP has all of RCK's characteristics, whether physical or psychological or otherwise. The only difference between $w$ and the actual world concern identities. Qualitatively, the two worlds are exactly similar. (A commitment to mere haecceitistic differences between worlds does not entail a commitment to the existence of haecceities Def D9.2, nor does a commitment to the existence of haecceities entail a commitment to mere haecceitistic differences between worlds. We leave it as an exercise for the reader to work out why this is so. See Lewis 1986a.)

Some have argued that Strong Transworld Identity entails that some worlds are merely haecceitistically different (see Chisholm 1967 and Forbes 1985, for example; for a helpful overview of these arguments, see Mackie and Jago 2013). The arguments often rely on a series of thought experiments where small changes are strung together to make large changes, though no such series is necessary. Maybe, for example, one might initially think that a world where RCK and THP completely switch roles is not plausibly possible. However, it is quite plausible that RCK and THP could have exchanged a small subset of their qualitative features. For example, imagine a world $w'$ in which RCK is a couple inches taller and THP a couple inches shorter; that is, imagine that RCK and THP exchanged merely their respective heights. If that is possible, then certainly they could also exchange their respective hair colors; this is $w''$. And shoe sizes; $w'''$. And so on. Soon enough, they seem to have switched all of their physical features. Now continue this step-wise process with psychological changes, and so on. At no point in the sequence of worlds is there a place where it seems there is a principled way to deny that the switching is possible. (Imagination as Guide to Possibility **PEpist 1** and patchwork principles **PMeta 5** are at work here.) Therefore, Strong Transworld Identity, together with the claim that some ordinary objects have accidental properties, seems to entail the possibility of mere haecceitisitic differences between worlds.

Why are mere haecceitistic differences problematic? Return to $w$, in which RCK and THP have completely switched roles. One might puzzle at the idea that, in that world, THP is identical to the object that has all of the properties actually possessed by RCK. What sense can be made of the claim that it's *THP* that has those properties, according to that world? If that world were actual, the RCK-role filler would have none of the properties that THP actually has. It is, therefore, difficult to see how that object could be literally identical to THP. Mere haecceitistic differences, therefore, can seem incompatible with putative claims to Transworld Identity that they were meant to describe!

Defenders of Strong Transworld Identity ought to deny that mere haecceitistic differences between worlds are problematic in this way. To execute this strategy, they must blunt the force of the thought that radical role-switching is incompatible with Transworld Identity. There are two ways of reading the objection. On a metaphysical reading, the problem is supposed to be that there is nothing to ground the identity of THP *as THP* in $w$. But this is not obviously true. We built the role-switching thought experiment, and indeed the idea of mere haecceitistic differences, in terms of *qualitative* features. One might, therefore, think that there are non-qualitative features of objects that will not be switched in cases of mere haecceitistic differences. For example, Scotists (9.2A.2) could insist that had $w$ been actual, THP would have still exemplified the haecceity that he actually exemplifies, despite the radical qualitative role-switching. More generally, one's solution to the problem of individuation discussed in Section 9.3.2 will impact the way one will deal with the present problem. That is, for each proposed solution to the problem of individuation, there will be a corresponding solution to this problem, whether in terms of haecceities, primitive identities, bare particulars, or whatever.[6]

There is, however, an epistemological reading of the worry. The concern is not that there is nothing to ground transworld identities, but rather that we have no way to *know* that $w$ represents THP, rather than some other object, as occupying the RCK role. Kripke (1980) responds to this objection by insisting that, for Abstractionists, possible worlds are "not *discovered* by powerful telescopes" (p. 44, emphasis in original). Rather, by setting out to consider the possible world where THP and RCK switch roles, we thereby come to think about $w$, that world wherein THP and RCK have switched roles (if such there be). In that sense, according to Kripke (and others), there is really no epistemological problem here at all. Anyway, there is no epistemological problem beyond the more general problem of our knowledge of modality. We consider certain aspects of that problem in the Section 16.2.

### 16.1.4 Summary

Summarizing, there are six basic questions for every theory of modality:

1. What sorts of things are possibilities or possible worlds? What are they like? Are there impossible worlds as well?
2. What makes a world possible? What accounts for the existence and variety of possible worlds?
3. What makes the actual world actual? What is actuality?
4. How do worlds represent things being a certain way? What is it for something to be true according to a world?
5. How do worlds represent *de re* possibilities (possibilities of or for particular things)? Do particular things exist in more than one world?
6. How do we know that possible worlds of certain kinds do or do not exist?

Here's a table summarizing the possible answers to these questions on behalf of Concretists and Abstractionists:

**Table 16.1** Comparing Concretism and Abstractionism

| | *Concretism* | *Abstractionism* |
|---|---|---|
| 1. | Concrete, parallel universes | (a) Sets of propositions<br>(b) States of affairs<br>(c) Properties of the world (maximal structural universals) |
| 2. | Brute fact of existence | (a) Primitive property of *possibility*<br>(b) (Combinatorialists) Mathematical facts about combinations |
| 3. | Location (indexicality) | (a) Primitive property of actuality<br>(b) Facts about truth simpliciter<br>(c) Facts about existence simpliciter<br>(d) Facts about fundamentality |
| 4. | Literally contains a truthmaker for *p* | (a) Magical representation<br>(b) Linguistic representation<br>(c) Pictorial representation |
| 5. | (a) Overlapping worlds<br>(b) Counterpart relations between Worldbound Individuals | (a) Russellian propositions (literally containing the thing)<br>(b) Plantingan propositions (containing the haecceity of the thing) |
| 6. | (a) Pure reason<br>(b) Best theory for metaphysical data | (a) Platonic vision of abstracta (a priori conceivability)<br>(b) Empirical knowledge of the powers and potentialities of actual things |

## 16.2 Modality and Epistemology: Possibility and Conceivability

We have canvassed a number of views about the nature of possible worlds, and about how possible worlds relate to modal facts. We have set aside epistemological issues for the most part, but the time has come to turn in that direction. There are number of questions with which philosophers wrestle regarding the relationship between modality and epistemology. We touch on just one here, namely the relationship between what is conceivable and what is possible. As we have seen in many cases, there is some sort of connection between what is conceivable and what is possible. Conceivability seems to be a fairly reliable source of information about what is possible. This is Imagination as Guide to Possibility, introduced in Chapter 3 and mentioned in a number of places thereafter:

**PEpist 1 Imagination as Guide to Possibility.** If a scenario is imaginable in great detail without evident absurdity, then we have good reason to think that it represents a metaphysical possibility.

One explanation of the truth of Imagination as Guide to Possibility is to suppose that possibility just is conceivability. If so, then every conceivable scenario is also really possible.

**16.2T Conceivability Entails Possibility.** Every conceivable scenario is true in some possible world.

There's an obvious problem with the suggestion, stronger than Conceivability Entails Possibility, that possibility *just is* conceivability. If this is supposed to ground possibility in something non-modal, then it can't work, since conceivability is already modal in character (as the suffix '-ability' indicates). Something is conceivable just in case it is possibly conceived of, or more precisely, possible to conceive of when one's faculties are in good working order. As a reductive account of possibility, this would be viciously circular.

However, one might take Conceivability Entails Possibility to be, not a reduction or definition of possibility, but merely an assertion of the nature and scope of what possible worlds represent. On this reading, Conceivability Entails Possibility asserts that there are enough possible worlds and of sufficient variety to verify every conceivable scenario. Alternatively, it asserts that our ability to conceive things is somehow constrained by the possible worlds there are, so we can't conceive of the impossible. If this view is right, then the fact that a situation is conceivable is infallible evidence of possibility.

### 16.2.1 Various notions of conceivability

Before we evaluate Conceivability Entails Possibility, we must get a great deal clearer about the nature of conceivability. In fact, philosophers have identified a number of varieties of conceivability (see, especially, Chalmers 2002). Here are two:

**Def D16.1.1 Negative epistemic conceivability.** The hypothesis that $p$ is *negative-epistemic conceivable* if and only if it is not knowable a priori that $p$ is false.

**Def D16.1.2 Positive-modal conceivability.** The hypothesis that *p* is *positive-modal conceivable* if and only if it is knowable a priori that *p* is possible.

There are two cross-cutting distinctions here: negative vs. positive and epistemic vs. modal. Negative conceivability involves not being in a position to know that something is either false or impossible, while positive conceivability involves knowing that something is possible. Something is negative-epistemically conceivable if it's impossible to know that it is actually false.

It is trivial that positive-modal conceivability entails possibility, since knowledge entails truth. However, it is not trivial that possibility entails positive-modal conceivability. To accept the converse implication is to hold that what we can't know to be possible must be impossible.

The most plausible and interesting version of Conceivability Entails Possibility incorporates negative-epistemic conceivability:

**16.2T.1 Lack of A Priori Falsity Entails Possibility.** Every negative-epistemic conceivable scenario is true in some possible world.

Lack of A Priori Falsity Entails Possibility has been a popular thesis in the history of philosophy, especially in the modern period of Descartes, Locke, Hume, and Leibniz. Nonetheless, the thesis forces us to ask why we ought to suppose that there is any connection at all between a priori knowability and possibility. A priori knowability is an *epistemological* property of propositions; it concerns how we know that a proposition is actually true. Possibility is a *metaphysical* property; it concerns whether the proposition could be true, regardless of whether we know or could know that it is actually false. Lack of A Priori Falsity Entails Possibility is logically equivalent to its contraposition, Necessity Entails A Priori Knowability:

**16.3T Necessity Entails A Priori Knowability.** If *p* is necessarily true, then we can know a priori that *p* is true in fact.[7]

The converse of Necessity Entails A Priori Knowability, A Priori Knowability Entails Necessity, is perhaps more defensible:

**16.4T A Priori Knowability Entails Necessity.** If it is knowable a priori that *p* is false, then *p* is true in no possible world.

One might think that if we could know a priori that some proposition *p* is false, it must be impossible for *p* to be true. If *p* were possibly true, how could we know that it is false without consulting some kind of empirical data?

At any rate, if both of these principles are true, then necessity is equivalent to a priori knowability: a proposition would be necessary if and only if it is a priori knowable.

### 16.2.2 Objections to identifying the necessary and the a priori

Recent philosophy of language, beginning with the seminal work of Saul Kripke (1980), has raised powerful objections to the supposed correspondence between what is necessary and what can be known a priori. There are apparent counterexamples in both directions. There are claims that are contingent (not necessary) but knowable a priori, and other claims that are necessary but knowable only a posteriori.

THE CONTINGENT A PRIORI Propositions expressed by the following sentences are contingent:

(4) The standard meter stick is one meter long.
(5) I am here now.
(6) I exist.

Let's consider these in reverse order. If any modal claim is certain for us, then the claim that (6) is contingent is among them: I might never have existed. Simple as that. It would be the height of hubris for any of us to claim that we are necessary beings. Likewise for (5), uttered by THP at midday on 27 October 2015. THP certainly might have been located somewhere else at the time of that utterance. For example, had he decided to work on this chapter somewhere else than a particular coffee shop in Fullerton, California, (5) would have been false. Instead of being here, at Green Bliss in Fullerton, he could have been at home or in his office at Biola University.

What about (4), which seems a bit more controversial? There used to be a standard meter stick in Paris, and that particular piece of metal could have been shorter or longer than it was. Nothing guaranteed that it should be exactly as long as it was in fact. Had it been longer or shorter, though, it would have been longer or shorter than that length we now pick out by the phrase 'one meter long'. To say that the standard meter stick is necessarily one meter long is to say something true read one way and false read another. The sentence, 'The standard meter stick is one meter long', is necessarily true understood *de dicto*. But the stick itself, considered as a particular bar of metal, could be beaten into any number of lengths, longer or shorter than a meter. The claim is, therefore, contingent understood *de re*.

Despite that each of (4), (5), and (6) are contingent, it seems that we can know them in a purely a priori way, without doing any observing, experimentation, or collection of empirical data. One doesn't need to do any observing to know that one exists, and one doesn't have to observe any facts about one's present location in order to know that one is "here". Similarly, the concept of *the standard meter stick* seems to guarantee, apart from any a posteriori investigation, that the standard meter stick is in fact a meter long, even if that very stick might have been shorter or longer.

Immanuel Kant, who discovered the distinction between a priori and a posteriori knowledge as we now understand it, thought that (7) and (8) were also knowable a priori.

(7) There is a linear temporal order.
(8) Space is approximately Euclidean.

Kant thought that all possible human experience involved linear time and Euclidean space. Consequently, we don't have to consult empirical science to learn that the world we encounter in experience satisfies (7) and (8). Nonetheless, both (7) and (8) seem to be contingent. We could imagine a world in which they fail, even if such a world would be one that we human beings could not possibly experience through sense perception.

At any rate, if any of (4) through (8) are indeed examples of the contingent a priori, then A Priori Knowability Entails Necessity is false.

THE NECESSARY A POSTERIORI Even more pertinently, Kripke argued that many necessary truths cannot be known a priori. Consider the following three identity claims:

(9) Mark Twain is Samuel Clemens.
(10) Water is $H_2O$.
(11) Heat is mean kinetic energy.

Each of these identities were learned through empirical investigation. There is nothing about the concepts of *Mark Twain*, *water*, or *heat* that guarantee the truths of the propositions. Nonetheless, as we have seen, Kripke offered a powerful argument for thinking that all identities are necessarily true. There is no possible world in which Mark Twain is not Samuel Clemens, since that would have to be a world in which Mark Twain is not Mark Twain, since Mark Twain just is Samuel Clemens. Since there is just one thing here, everything true of the "one" is true of the "other". It's impossible for Mark Twain not to be Mark Twain, so it must similarly be impossible for Mark Twain not to be Samuel Clemens. Similar remarks apply to (10) and (11). If any of (9) through (11) are examples of the necessary a posteriori, then Necessity Entails A Priori Knowability is false.

The upshot is that we cannot simply identify necessity with what can be known a priori. There are powerful counterexamples to the claim that everything knowable a priori is necessary and to the claim that everything necessary is knowable a priori.

There is, however, a plausible move that could be made in defense of the necessity-apriority connection. We could distinguish between sentences or statements, on the one hand, and propositions, on the other (as we did in Chapter 2). It is clear that we cannot know a priori that the sentences (9) through (11) are true: that is, we cannot know a priori that they express true propositions. However, it might be that the propositions that these sentences express all have the same logical form, namely, $x = x$. If so, those propositions can all be known a priori, since they are all instances of a law of logic, namely, the reflexivity of identity.

There is, however, a counter-move that also has some plausibility. We could suppose that each of the sentences necessarily expresses the proposition that it does. For example, it is plausible to suppose that the names 'Mark Twain' and 'Samuel Clemens' (as used in our actual context) are *essentially* names of the very person (the famous American author) that they are in fact names of. Now, of course, it is not necessarily true that any name spelled 'M-a-r-k (space) T-w-a-i-n' is a name for that same author. That string of letters and spaces could be the name of anyone or anything. I could give my pet rock the name 'Mark Twain' if I wanted to. However, we should think twice about identifying names with strings of letters and spaces. A name is a peculiar kind of word, with its own history and use. When we use the string of letters to name both the author and my pet rock, we have introduced two different names, each with the same spelling as the other. Such names are merely homonyms, not strictly identical. With this picture of naming in mind, it now seems plausible that (9), if it and its constituent words and names exist, must be true. It must express a true proposition of the form '$x = x$', and yet this necessity cannot be known a priori.

This last point is a special case of another claim about necessity made by Saul Kripke (1972), namely, the necessity of origins. It seems plausible to suppose that each concrete thing, whether a person, an artifact or a name, has its origin essentially. It is impossible that *this* very person, say, Elizabeth II of England, could have existed with entirely different parents, that is, with parents other than George VI and Queen Elizabeth. In particular, Elizabeth II couldn't have been created, Frankenstein-style, in a laboratory from spare parts. However, we do not know Elizabeth II's actual origins a priori. So, here again we find necessities that can only be known a posteriori.

### 16.2.3 Knowledge of possibility: patchwork principles

We never perceive non-actual possibilities directly. How could we? They aren't actual! How, then, do we know anything about them at all? One plausible route to our knowledge of merely possible worlds is to think of them in terms of rearranging the contents of the actual world. We take things of a kind that have really happened in the actual world and construct a non-actual scenario by arranging them in a novel spatial and temporal order. If we do this according to certain rules, the thought goes, we can be confident that our rearranged scenarios represent real possibilities. Thus, our knowledge of at least some possibilities can be secured.

This process of construction is something like the way we make patchwork quilts: we cut up pieces of fabric from an already-existing cloth, and then we stitch the pieces together in new ways. Because there is something intuitively attractive about supposing that rearrangements of the contents of the actual world represent real possibilities, David Lewis formulated what he called 'patchwork principles' (see especially Lewis (1986a)). These principles are Lewis's attempt to codify the rules that govern our rearrangements of actual things to represent possible scenarios. Here are two versions of Lewis's patchwork principles, one involving finitely complex recombinations, and the other extending the principle to the infinite.

**PMeta 5.1 Finite Spatiotemporal Patchwork.** If it is possible for an event or process of (intrinsic) type $A$ to occur, and if it is possible for an event or process of type $B$ (distinct from $A$) to occur, and if there is enough room in the history of the world to locate in it instances of both events or processes without overlap in time and space, then it is possible for an event or process of type $A$ to be realized together with an event or process of type $B$.

**PMeta 5.2 Infinite Spatiotemporal Patchwork.** If $T$ is a class of types of events or processes, and for each member of $T$, it is possible for an event or process of type $T$ to occur, and if there is enough room in the spacetime expanse of the world of the world to locate within it instances of each of the types in $T$ without overlap in space and time between the instances, then it is possible for all the types in $T$ to be realized together.

For example, clearly football matches are possible, since some of them are actual. For example, in 2011 Tottenham and Real Madrid clashed in the UEFA Champions League. Let football matches be events of the type FOOTBALL. But it also seems that quidditch matches are possible, though none actually occur. Surely it's at least possible that something occur which is intrinsically exactly like the match between Ireland and Bulgaria which Harry Potter and friends witness at the 1994 Quidditch World Cup (in *Harry Potter and Goblet of Fire*). Let quidditch matches, or anyway things intrinsically identical to quidditch matches, be events of the type QUIDDITCH. If you think that events of the types FOOTBALL and QUIDDITCH are possible separately, then Finite Spatiotemporal Patchwork demands that one also think that there is a world in which a FOOTBALL event and a QUIDDITCH event both occur, which is just to say that it is possible that there be both a football match and a quidditch match. This is just what we would expect, since it is intuitively plausible to suppose that a single world could contain both types of events. Indeed, if the world were as the *Harry Potter* series depicts it, we would be in just such a world! Similarly, suppose you think that Light Cycle matches (of *Tron* fame) are possible, or anyway that events intrinsically like Light Cycle matches are possible. Let's say these matches are of the type LIGHTCYCLE. Finite Spatiotemporal Patchwork then demands that there be a single world in which events of the types FOOTBALL, QUIDDITCH, and LIGHTCYCLE all occur. (Iterations of this recipe for four types, five types, and so on are easy to construct.)

Infinite Spatiotemporal Patchwork simply extends this idea to arbitrarily large classes of types of events or processes. Again, this is intuitively plausible antecedently, and so our patchwork principles are tracking our intuitions about what is possible.

Patchwork principles ensure that there are enough worlds to match the number of possibilities, and also give us some insight into how we are able to know what other possible worlds are like. We know what they are like, in at least some cases, because they are just rearrangements of things like the ones in the actual world. We'll return to a different epistemological question shortly.

### 16.2.4 Aristotelian/Powerist conceptions of modal knowledge

As we discussed in Chapter 15, Powerists (4.4A.3) and defenders of Aristotelian Modality (15.2T.7) have a more empirical approach to our knowledge of modality. We discover what can happen by investigating the natures or essences of things, where the nature or essence of a thing consists of the most fundamental powers and potentialities of that thing. On this view, modal facts (facts about what is possible) are grounded in facts about the essences of actual things, and not vice versa (Fine 1994a, 1994c). Powers and essences are discovered experimentally, as we described in Section 6.1.2, on the epistemological consequences of Powerism.

Aristotelian Modality provides support for Kripke's thesis of the essentiality of origins. If something $X$ comes into existence, there must have been something $Y$ there beforehand with the power of generating $X$. Suppose it were possible for $X$ to have had a completely different origin, with a different source or originator. Then there must have been some other possible thing $Z$, not identical to $Y$, with the power of generating $X$. If two distinct things had the power of generating $X$, then each could have exercised its power independently of whether the other did. Thus, it would have been possible for both $Y$ and $Z$ to generate $X$ in the same world. But then $X$ would have come into existence twice, which is surely impossible. Thus, if $X$ has an origin in time, it is essential to $X$ that it have the source it actually has. For similar reasons, it must be essential to X that its originator's power to create *X* could only be exercised in unique and unrepeatable circumstances. Thus, the entire origin of *X* must be essential to it.

## 16.3 Conclusion

We introduced the distinction between *de dicto* and *de re* necessity, and then we explored different ways of making sense of *de re* necessity—necessities involving particular individuals. For Concretists, *de re* possibilities for one thing involve the states of concrete counterparts of that thing in other worlds. Concretists are Counterpart Theorists. For Abstractionists, in contrast, one and the same thing can participate in both actual and merely possible situations. Abstractionists go for Transworld Identity. This gives rise to an especially vivid version of the irrelevance objection to Concretism that we first discussed in Chapter 15, namely, the Humphrey Objection.

We then turned to questions involving our knowledge of the merely possible. Much of that discussion concerned the issue of the relationship between possibility and conceivability. Most philosophers accept Imagination as Guide to Possibility (**PEpist 1**), which suggests that conceivability is a good guide to metaphysical possibility (although perhaps not an infallible one). However, some metaphysicians go farther and identify what is possible with what is conceivable by us, or perhaps with what we are justified (all things considered) in believing to be possible. Saul Kripke has argued that this identification of the possible with the conceivable fails in both directions. Most importantly, there seem to be metaphysical necessities that can only be known a posteriori, and whose contradictories are fully conceivable.

On the positive side, we have not ruled out the use of conceivability as a *fallible* guide to possibility. In addition, there are plausible patchwork principles that enable us to extend our knowledge from simple scenarios to much larger ones. Finally, if we adopt Aristotelian Modality, there is room for empirical and scientific investigation of the essences of things.

## Notes

1 The exception is Lagadonian Linguistic Abstractionism, according to which things serve as their own "names". It will turn out that this nicety won't matter for what we're about below, and anyway, the contrast here is between how things are parts of Concretist worlds versus how they are represented by Abstractionist worlds. Such a contrast exists even on the Lagadonian Linguistic Abstractionist view.

2 But see McDaniel (2004) for a defense of Overlapping Concrete Worlds.

3 'Ludovicus' is Latin for 'Lewis'.

4 Sider (2002) develops an explicitly counterpart-theoretic variety of Abstractionism.

5 This terminology has been used in the literature in a few interrelated senses. The differences would not substantively affect the discussion to follow.

6 This is a bit of a simplification. For example, if one thought that bare particulars merely individuate without grounding identity, one cannot use bare particulars to solve the present problem. But one could nonetheless deploy a transworld analogue of, for example, Primitive Identity (9.2T).

7 Here is the proof that they are equivalent (using parentheses to mark scope). (If $q$ is negative epistemic conceivable, then $q$ is possible) entails (if not-($q$ is possible), then not-($q$ is negative epistemic conceivable)), by contraposition. Distributing the negations in the antecedent and consequent, and using Def D16.6.1, it follows that (if $q$ is impossible, then NOT(we cannot know a priori that $q$ is false)). Distributing the negation in the consequent and assuming that $q$ is false entails not-$q$ is true, we get (if not-$q$ is necessary, then we can know a priori that not-$q$ is true). This is just the thesis that Necessity Entails A Priori Knowability, where $p =$ not-$q$.

# Part VI

# Space and Time

# 17

# Is Space Merely Relational?

In Part VI (Chapters 17–21), we will examine the nature of space and time, the framework that somehow contains all of the world's physical and living things, along with their associated events, actions, and processes. These questions have constituted a large part of metaphysics and the philosophy of nature from the time of the ancient Greeks.

In Chapter 17, we focus on a question that both metaphysicians and theoretical physicists share: is space a thing or world of things in its own right or does it consist merely in relations among bodies or physical events? *Substantivalists* take space to be composed of places, real entities that exist in themselves, whether occupied or unoccupied. *Relationists*, in contrast, take the talk of spatial locations to be merely a way of keeping track of the spatial relations among physical things.

In Section 17.1, we give an overview of the various theories that we will consider here. There are three Substantivalist theories, namely, the Theory of Spatial Qualities, Spatial Monism, and Body-Space Dualism, and two Relationist theories, namely, Aristotelian Relationism and Modern Relationism. We examine the three Substantivalist theories in Section 17.2 and the two Relationist theories in Section 17.3. In the final section, Section 17.4, we turn to a serious metaphysical problem for all of the theories of space, that of accounting adequately for vacuums, holes, and other forms of absence.

## 17.1 The Nature of Location

It should be uncontroversial that at least some things are located in space and that different things occupy different locations. Further, as we discussed in connection with Solipsism in Chapter 13, we seem to perceive things in space and to remember things

*The Atlas of Reality: A Comprehensive Guide to Metaphysics*, First Edition.
Robert C. Koons and Timothy H. Pickavance.

as located in space. These perceptual seemings have a right to be taken seriously, in the absence of strong arguments to the contrary.

But what does it mean for something to be located somewhere? Are places things and location a relation between a place and something else? If places are things, what kinds of things are they? For example, are places individuals or a kind of property? We first must choose between Spatial Substantivalism and Spatial Relationism:

**17.1T Spatial Substantivalism.** Places exist and are G-fundamental.

**17.1A Spatial Relationism.** Places are not G-fundamental.

The concept of G-fundamentality was introduced in Chapter 3 (D3.5): an entity is G-fundamental if and only if its existence is ungrounded. We also introduced there the narrower category of O-fundamentality: an entity is O-fundamental if and only if its essence does not involve any other entity. We will focus exclusively here on the broader issue of G-fundamentality. When we use the term 'fundamental' throughout this chapter, we will always mean 'G-fundamental'.

One might think, given the more or less uncontroversial points about things having location, that Spatial Substantivalism will win the day. However, Substantivalism involves more than the mere reality of space and locations. It involves the *existence of places*. Substantivalists *reify* space itself, as *a thing* in which other things are located. According to Substantivalists, *places* really exist, and things are located by being *in* a place. In contrast, Spatial Relationists deny that places really exist. There are located things, but being located is not a matter of being in some *place*. Rather, *having location* is simply a matter of bearing various other relations, like *contiguity*, *distance*, *direction*, to other things.[1] Spatial Relationists favor the conceptual grounding of truths about locations in terms of the facts about spatial relations, leaving no room for real, much less fundamentally real, places.

One terminological complication: 'Substantivalism' is the traditional name for the view expressed in 17.1T, but this thesis does not assert that places or parts of space are substances in the sense discussed in Chapter 9. In particular, Substantivalism does not by itself entail that spatial locations persist through time. Substantivalists could suppose that every location exists for only a moment.

Spatial Substantivalism comes in two forms, depending on whether places are properties or not. Assuming that places are properties amounts to the Theory of Spatial Qualities; the alternative version of Substantivalism is Spatial Particularism. Spatial Particularism in turn comes in two forms, Body-Space Dualism and Spatial Monism. According to Body-Space Dualism, both places and located bodies are fundamental particular things. Spatial Monists maintain that only places are fundamental particulars. For Spatial Monists, "bodies" are simply places that have a peculiar kind of body-ish or enmattered quality.

Spatial Relationists also come in two forms, Aristotelian and Modern. Modern Relationists believe that the fundamental spatial relation is the *distance* between dimensionless point-masses, while Aristotelian Relationists take the fundamental spatial properties to be *volume* and *shape*, and the only fundamental spatial relation to be *contiguity* (or *contact*).

We consider these views in turn.

## 17.2 Spatial Substantivalism

Consider Spatial Substantivalism. The first question that arises for Substantivalists concerns the nature of places themselves. There are two plausible answers to this question:

**17.1T.1T Theory of Spatial Qualities.** Places are fundamental properties or qualities, and location is predication: a place is predicated of the things located there.

**17.1T.1A Spatial Particularism.** Places are ordinary particulars (not properties or qualities), and *location* is an external relation between fundamental particulars (between locations and the occupiers of those locations).

There is a second question, though, one concerning the nature of places' relations to one another. In particular, the question is whether places are related internally, externally, or neither. Clearly, this question requires an understanding of this threefold distinction. We have already discussed internal relations; here again is the definition:

**Def D2.2 Internal Relation.** R is an *internal relation* if and only if necessarily, for every $x$ and $y$, whether $R$ holds between $x$ and $y$ depends only on the intrinsic properties of $x$ and of $y$.[2]

An *external* relation will obviously not be internal. The holding of external relations depends on factors beyond the intrinsic character of the relata (taken individually). But David Lewis (1986a: 62–63) has suggested that we make a more fine-grained distinction. He takes an external relation to be one that is intrinsic to the pair of the relata, treated as a whole or composite.

**Def D17.1 External Relation.** $R$ is an *external relation* if and only if $R$ is not internal, and necessarily, for every $x$ and $y$, whether $R$ holds of $x$ and $y$ depends only on the intrinsic properties of the pair of $x$ and $y$ (taken together).

If we adopt this definition, there will be relations that are neither internal nor external. (We supply examples below.) We have, then, a three-way distinction between internal relations, external relations, and relations that are neither internal nor external. There are a number of examples of relations that are uncontroversially internal: for example, the relation of *being twice as massive as* would seem to be internal. Given the intrinsic properties (including the exact mass) of two bodies, we can always determine whether or not this relation holds between them. If one book has a mass of 200 grams, and the second book has a mass of 400 grams, then it necessarily follows that the second is twice as massive as the first. The relations of *being the same color as*, *being happier than*, and *being denser than* are also plausibly internal.

It is also easy to find examples of relations that are neither internal nor external. The third category of relations always involves some kind of reference to a third, independent standard or reference point. So, for example, the relation of *having an earlier number in the Dewey decimal library catalog* is neither internal nor external. Whether one book's numbers is earlier than another is not fixed by any property intrinsic to the books themselves, whether taken individually or as a pair. We can only tell if the relation holds by consulting a third thing, in this case the library catalog. The relations of *having a smaller Social Security number than*, *being better liked by most people than*, and *being better known than* are also neither internal nor external.

An external relation is intrinsic to each related pair, not taken separately and individually but as a pair. Consider the set of metal rings. Most rings are unconnected, in the sense that neither ring passes through the hole of the other. Some pairs of rings are directly interconnected, however, in the way that successive links in a chain are intertwined. Being connected in this way is not intrinsic to either ring taken individually, but it does seem to be an intrinsic property of the pair. It is something to do with the way in which the pair is in itself.

Are spatial relations like *distance* and *contiguity* (or *contact*) internal relations, external or neither? It seems clear that spatial relations are either internal or external. It is hard to see how spatial relations between places could depend on something outside of space (assuming that we have already ruled out Idealism, which makes *distance* and other spatial relations mind-dependent). So the question is whether *distance* and *contiguity* are internal or external. We thus get a distinction between Spatial Externalism and Spatial Internalism:

**17.1T.2T Spatial Externalism.** Spatial distance and contiguity are external relations between places.

**17.1T.2A Spatial Internalism.** Spatial distance and contiguity are internal relations between places.

**Spatial Homogeneity.** If places exist and are not qualities, then any two places are intrinsically indistinguishable.

Given Spatial Homogeneity, the Theory of Spatial Qualities and Spatial Internalism entail each other. Suppose the Theory of Spatial Qualities is true. Then places must form a natural "quality space", similar to the "space" of colors. The "distance" in hue between any two colors (such as *scarlet* and *purple*) is internal to those colors. Given the two colors and their natures, their *hue-distance* is fixed. Similarly, if places are a kind of quality, then any two places must be distinguishable from one another by their qualitative difference, a qualitative difference that fixes their mutual distance. The story might go like this. Two places are in the same neighborhood if and only if they strongly resemble each other intrinsically. Two places are far apart if they are very dissimilar. *Spatial distance* is measured by degree of similarity. In order to provide a basis for a metric of distance, we need the following sort of primitive relation between places:

Place $x$ is *more similar* to $y$ than it is to $z$.

We can then define 'closer to' in terms of 'more similar to':

Place $x$ is *closer* to $y$ than it is to $z$ if and only if $x$ is more similar to $y$ than it is to $z$.

We apparently have no experience of these qualitative differences—one place seems intrinsically exactly like any other—but the Spatial Qualities Theorists must hold that such internal differences nonetheless exist. Which is just to say that Spatial Internalism is true.

Conversely, if places are ordinary particulars—that is, if the Theory of Spatial Qualities is *false*—then Spatial Homogeneity requires that every place is intrinsically indistinguishable from every other place. Further, if places are not properties but particulars, then *location* must be an external relation between things and their places. Nothing about the intrinsic nature of individuals fixes where they are (if they are things that occupy space) or what they contain (if they are themselves regions of space). *Distance* and *contiguity* (the absence of distance) would then have to be *external* relations between those places because the indistinguishable intrinsic natures of places could not determine what relations hold between the places. Two regions of space with the same shape and volume would seem to be intrinsically indistinguishable from each other, no matter what their relative locations. (Importantly, we have already ruled out that these relations are neither internal nor external.) *Part/whole* relations are internal because whether place $A$ is or is not a part of place $B$ is fixed by their intrinsic natures. But we cannot reconstruct the topological and metrical features of space based solely on these *part/whole* (mereological) relations. If *distance* is an external relation between places, then Spatial Internalism is false. This establishes that given Spatial Homogeneity, if the Theory of Spatial Qualities is false, then Spatial Internalism is false. By contraposition, Spatial Internalism entails the Theory of Spatial Qualities (again, given Spatial Homogeneity).

Defenders of Substantivalism (whether the Theory of Spatial Qualities or Spatial Particularism) must immediately confront an important question, namely, how to square Substantivalism with special relativity. Substantivalism makes distance something absolute, not relative to one's frame of reference, and this seems to contradict special relativity. One could solve this by moving to spacetime points (rather than mere spatial points) and by making the basic relation not *distance* but spacetime *interval*, a distance-like quantity that is independent of frame of reference. In other words, Substantivalists could move from Spatial Substantivalism to *Spacetime* Substantivalism.

Spacetime Substantivalism has the further advantage that Einstein's general theory of relativity treats spacetime as a real thing with its own "curvature", corresponding to the gravitational force of massive bodies.

### 17.2.1 Theory of Spatial Qualities: Advantages and disadvantages

Bertrand Russell (1927) first proposed the Theory of Spatial Qualities. It has at least four potential advantages. First, it gives an account of what places are and what location is. Second, it gives an account of the nature of spatial relations like *neighborhood* and *distance*. *Relative distance* becomes a kind of internal relation among places. Third, it explains why something cannot be located in more than one place, since doing so would be to possess different, presumably incompatible qualities. Fourth, it doesn't multiply particular entities, and so is quantitatively simpler (PMeth 1.4).

But the Theory of Spatial Qualities also has some potential disadvantages. Here are three. First, it is not clear how degrees of similarity between spatial points are to be understood. There are some clear cases in the color case; yellow is more similar to red than it is to blue. But what does degree of similarity really amount to in the case of space? Second, how can the Theory of Spatial Qualities deal with vacuums? What possesses the spatial quality, when no physical body occupies the space? Do we need to posit uninstantiated qualities and merely possible bodies? (We will discuss this problem in more detail in Section 17.4.) Third, it brings with it an ontological commitment to properties. It is incompatible with Ostrich Nominalism (7.1A.1A).

### 17.2.2 Spatial Particularism: Distance as an essential but external relation

It seems that the distance between two places is essential to those places. It doesn't seem to make much sense to suppose that two places, $P_1$ and $P_2$, might have been closer together or farther apart. Is there a connection between internal relations and essential relations? Some philosophers have thought so, that external relations hold contingently while internal relations hold necessarily. If these philosophers are correct, one could argue to the falsity of Spatial Externalism on the basis of the necessity of relations of *spatial distance*. For if external relations hold contingently, and if Spatial Externalism is true, then relations of *distance* among places are contingent. But such relations are necessary, so Spatial Externalism is false.

However, the distinction between internal and external relations is independent of the distinction between necessary and contingent relations. The internality of a relation has to do with its being guaranteed to hold by the intrinsic natures of the relata, not with its being essential to the relata that they stand in that relation.

First, then, there may be essential relations that are not internal. For example, it might be essential to Queen Elizabeth II that she was the daughter of George VI (any daughter of anyone else wouldn't have been that very woman), but the intrinsic natures of Elizabeth II and George VI do not seem to guarantee that they stand in this relation. We can easily imagine intrinsic duplicates of these two that do not stand in a father-daughter relation. Imagine, for example, that George VI had an intrinsically identical twin; it is an external matter that one rather than the other is the father of Elizabeth II. The *father-daughter* relation is, thus, external, despite that the relation is plausibly essential to at least one of the relata.

Conversely, a relation might be internal without being essential. The relation of *being twice as large as* is an internal relation. It holds by virtue of the sizes of the two relata, and size is an intrinsic feature of a thing. But we need not assume that anything has its size (either relative or absolute) essentially. Thus, it is possible for something to stand in the relation of *being twice as large as* to something contingently. THP's wife Jamie, for example, is currently about twice as large as their three-year-old daughter, Gretchen. But Gretchen is still growing; she is only contingently 39 inches tall. Soon, therefore, Jamie will not be twice as large as Gretchen, and so she will not stand in this relation to her. Despite the fact that the relation is internal—that is, guaranteed to hold given the current intrinsic natures of Jamie and Gretchen—it is nonetheless contingent that they stand in it.

Although, as we have seen, it is possible to distinguish internality from essentiality and externality from accidentality, nonetheless, this separation comes at a cost. To suppose that some external relations are essential, then we must posit brute necessities of some kind, like the brute necessity of Elizabeth's having George IV as a father. If the *distance* between two places is indeed an external but essential relation between them (as Spatial Externalism requires), then this involves a necessary connection between two separate things: the two places cannot exist without standing in their essential *distance* relation to one another. To emphasize, though, this necessity is not grounded in the intrinsic natures of the two places. In general, we should prefer a theory that posits fewer such necessary connections, according to Ockham's Razor (**PMeth 1.2**). External but necessary spatial relations should, then, be counted a cost of Spatial Externalism.

TWO VERSIONS OF SPATIAL PARTICULARISM: BODY-SPACE DUALISM AND SPATIAL MONISM
There are two versions of Spatial Particularism: Body-Space Dualism and Spatial Monism.

**17.1T.1A.1T Body-Space Dualism.** Both bodies and places are fundamental particulars of different kinds, with location being an external relation between bodies and places.

**17.1T.1A.1A Spatial Monism.** The only concrete particulars are places (parts of space). A body is simply a special kind of place—one that is characterized by a quality of *being massive*, *body-ish* or *en-mattered*.

RELATIONS BETWEEN THINGS AND THEIR PLACES The Theory of Spatial Qualities has a relatively clear account of the relation of *location*, that is, the relation that a thing holds to the place it occupies. For Spatial Qualities Theorists, this relation is that of *predication*. Locations are properties of things. This makes *location* an internal relation, since it is the qualitative character of the located thing that fixes its location, and a thing's location is simply one of those qualities.

Spatial Monists have a similarly simple theory. A body is simply identical to a place that has a special kind of *body-ish* quality or associated quantity (like *mass* or *charge*). On Spatial Monism, places that are the fundamental particulars, and bodies are modifications of space, like waves or disturbances passing through a constant medium.

In contrast, Body-Space Dualists need some further account of *location*. They could hold that *location* is an additional, fundamental relation between things. This comes with some theoretical cost, since we should prefer theories that posit the smallest class of fundamental properties and relations (**PMeth 1.4**). But Body-Space Dualism also has two putative advantages. First, there is something intuitive about thinking of places as things, and in thinking of both *location* and *spatial distance* as external relations. Second, one can avoid any ontological commitment to properties or qualities. Everything that exists is a particular.

But Body-Space Dualism has at least two further disadvantages. First, places seem to be superfluous. Given that *distance* is an external relation, why not make it a relation between ordinary things, non-places (e.g., bodies or particles)? The resulting Relationist account seems simpler. Second, Body-Space Dualism has difficulty explaining why a material object and the space it occupies must both have the same shape. The fact that a thing and its place must always be of the same shape seems to be yet another brute metaphysical necessity. Body-Space Dualists could hold that a thing derives its shape entirely from the shape of the spatial region it occupies. But this makes shape an extrinsic rather than intrinsic property of material bodies, which seems wrong.

This objection can be mitigated to some extent by supposing that the fundamental sort of body is that of point-sized or dimensionless bodies, and the fundamental sort of place is that of points. On this Pointillist (18.1T) view, extended bodies are nothing but sets or pluralities of point-sized bodies, and regions of space are nothing but sets or pluralities of points. Neither points nor point-sized bodies have any shape in any interesting sense, so the fact that each point-sized body "fits" into a point is not problematic. The shape of an extended body would then simply consist in the shape of the totality of points that its point-sized bodies occupy. Strictly speaking, the shape of such extended bodies wouldn't be intrinsic to them, since it depends on the relations of the ultimate parts of those bodies to spatial points, but in this case the extrinsicality of shape doesn't seem troubling.

Let's compare Body-Space Dualism with Spatial Monism, the alternative form of Spatial Particularism. Spatial Monists can appeal to Ockham's Razor (**PMeth 1.4**), since they posit fewer entities and fewer fundamental kinds of entities. In addition, they need not posit any relation of location, nor do they need to worry about the coincidence between the shapes of bodies and the shapes of their places. There are only places (parts of space). Some places are vacuum-ish or empty, and other places are body-ish or en-mattered. What we call a moving body is nothing more than the successive loss of *body-ishness* by some places and the gain of *body-ishness* by other, neighboring places. A moving body is something like a moving wave on the ocean's surface or the movement of an illuminated spot on a wall. Nothing is literally moving in the direction of the body's movement. The water particles are moving up and down, and not laterally, and the wall is not moving at all. Similarly, space never moves. What "moves" is merely the pattern of realization of some quality by various parts of an immobile space. So the principal advantage of Spatial Monism is ontological and metaphysical simplicity.

Spatial Monism's major disadvantage lies in its account of the persistence of material bodies, including people and other organisms. Spatial Monists must deny that each of us is literally the same thing as any person existing an hour ago. An hour ago, the region of space that RCK occupies, which is the only thing that RCK could be identical to, was not a person at all but exhibited quite a different pattern of en-matterment. Each time one of us moves in space, one region of space ceases to be a person and another region of space becomes a person. The two persons are numerically distinct, since each is identical to a different part of space. Spatial Monists must adopt a version of Perdurantism (24.1T.1T.1A.1T) as their account of what it is for bodies to persist through time. We will take up questions to do with persistence in Chapter 24.

## 17.3 Spatial Relationism

**17.1A Spatial Relationism.** Places are not G-fundamental.

Since the kind of grounding involved here is conceptual grounding (Section 3.4), Spatial Relationists reject the real existence of places. There are only located entities, material bodies of some kind.[3] Spatial Relationism had some ancient defenders. Aristotle and many of his followers, for example, seem to have been Relationists. The view has also been popular in modern times, beginning with George Berkeley and Gottfried Leibniz in the seventeenth century.

There are two schools of Relationism, and the two schools coincide roughly with the two periods, ancient and modern. The schools differ on the question of which spatial relation or property is most fundamental. Modern Relationism has it that *distance* is most fundamental, while Aristotelian Relationism has it that *extension* (that is, *shape* and *volume*) and *contiguity* are most fundamental.

**17.1A.1T Modern Relationism.** Spatial Relationism is true, and *shape* and *volume* are properties of material bodies that consist entirely in the holding of certain *distances* between the pairs of proper parts of those bodies.

**17.1A.1A Aristotelian Relationism.** Spatial Relationism is true, and *shape* and *volume* are metaphysically fundamental properties of material bodies, and *contiguity* or *contact* is a fundamental relation.

There is a related distinction to consider. Are fundamental material bodies point-sized (without volume) or extended (voluminous)? The two distinctions seem to be tightly connected. If fundamental bodies are point-sized, then the fundamental spatial relation must be *distance*, since point-sized bodies have no shape or volume and are never contiguous. Conversely, if fundamental bodies are extended, then it seems natural to take the properties of *size* and *shape* and the relations of *contiguity* (touching) and *non-contiguity* (non-touching) as fundamental.

**17.2T Fundamentality of Distance.** The fundamental spatial property is *distance* between point-sized things.

**17.2A Fundamentality of Shape and Contiguity.** The fundamental spatial properties are *shape*, *volume*, and *contiguity* between extended (voluminous) things.

We are tacitly excluding the view that *distance* is the fundamental spatial relation but holds between extended things rather than points. The distance between two extended things could be defined in a variety of ways on this view. It could be the distance between their closest points, the distance between their farthest points, the distance between their geometric centers, or the distance between their centers of mass. However, in each case, it seems that the distance between the extended things is derivative, and the distance between two points fundamental.

**Distance a Relation Between Spatial Points.** The distance between two extended things is not fundamental but derives from the distance between pairs of points occupied by those things.

Given the thesis Distance a Relation Between Spatial Points, Modern Relationism and the Fundamentality of Distance entail each other, as do Aristotelian Relationism and the Fundamentality of Shape and Contiguity. We leave it to the reader to discern these entailments.

POINTS, CURVES, SURFACES, AND REGIONS Before moving on, we should clarify what we mean by *points* and *point-sized bodies*, as well as such contrasting terms as *extended regions* and *surfaces*. We'll start with the definitions of points, curves, surfaces, and extended or voluminous regions, as they would be defined for Substantivalists:

**Def D17.3.1 Point.** A *point* is a zero-dimensional place (lacking length, width, and depth, as well as area or volume).

**Def. D17.3.2. Curve.** A *curve* is a one-dimensional place, having length, but lacking width and depth, as well as area or volume. A curve can be either finite or infinite in length. Points can be located on curves. A curve is a *line* if and only if any three points located on the curve are such that one is exactly between the other two.

**Def D17.3.3 Surface.** A *surface* is a two-dimensional place, having length, width, and area, but no depth or volume. A surface can be either finite or infinite in area. Points and curves are located on surfaces. A surface is a *plane* if and only if any any two points located on the surface are also located on a line on the surface.

**Def D17.3.4 Region.** A *region* is a three-dimensional place, having length, width, depth, and volume. Regions can be either finite or infinite in volume. Points, curves, and surfaces are located in regions.

For simplicity's sake, we are going to ignore the possibility of fractal entities, entities with dimensionality other than 0, 1, 2, or 3, such as space-filling curves with dimensionality between 1 and 2. For Spatial Substantivalists, these definitions can be applied to define point-sized, curve-sized, surface-sized, and region-sized material bodies. For example, a point-sized material body, or *point-mass*, is a body that exactly occupies a spatial point and nothing else.

Spatial Relationists do not believe in places and so cannot use Defs D17.3.1 through D17.3.4. However, they can make do with very similar definitions, such as Def D17.3.1′:

**Def D17.3.1' Relationist Point.** A point-sized body is a zero-dimensional body, lacking length, width, and depth, as well as area or volume. The distance between any two parts of a point-sized body is zero.

ARGUMENTS FOR AND AGAINST SPATIAL RELATIONISM According to Spatial Relationism, there are no places, strictly speaking. There are only things (material bodies of some kind) bearing various spatial relations (like *distance* or *contiguity*) to each other.

The main argument for Relationism is simplicity (**PMeth 1**). Substantivalists, realists about places, must posit two kinds of entities, namely, places and things that occupy places. Relationists believe only in place-occupiers (material objects), together with facts about the spatial relations among those objects. This difference threatens to disappear if the Substantivalists adopt the Theory of Spatial Qualities, especially if spatial relations are internal relations between such qualities. Even in that case, however, there does remain this difference: Substantivalists will want to posit the existence of unoccupied places, which for Spatial-Quality Theorists will amount to the existence of uninstantiated location-qualities. Relationists, in contrast, have only place-occupiers, with nothing corresponding to unoccupied places.

This lack of unoccupied places poses an immediate problem for Relationism, however. Relationists seem forced to deny many, if not all, of the usual axioms of geometry. For example, given Relationism, it is false that there always exists a spatial point between any two distinct points. If two points were occupied by material objects, but nothing existed between them, then of course it would be false to assert that something (namely, a point) exists between them.

This problem creates another problem for Relationism (as Hartry Field (1984) argues). Without the usual axioms of geometry, Relationists cannot provide the usual reductionist account of spatial quantities like *distance*, *area*, *volume*, and *angle measure*. Substantivalists can suppose that the only fundamental spatial relation is *betweenness*, which is a relation among three points. Given the usual axioms for geometry, expressed entirely in terms of *betweenness*, spatial quantities can be defined in such a way that any permissible definitions are isomorphic to one another, differing only in the arbitrary unit of measure. However, without those standard axioms, such a definition of spatial quantity is impossible. This implicit definition of quantities is supported by a *representation theorem*, a mathematical proof in which it is shown that any system of objects that satisfies the standard axioms of geometry can be assigned a set of real-numbered values that satisfy certain basic postulates for the quantity in question (e.g., *distance*), and, moreover, that any permissible set of assigned values is essentially equivalent to any other such set, in the sense that the one set can be converted into the other by multiplying by some constant number (representing the conversion from one arbitrary unit of measurement to another).

In the absence of such a representation theorem, Relationists would have to adopt a position that Field calls "heavy-duty Platonism". Heavy-duty Platonism requires that there is a single, fully *natural* unit of distance (the *metaphysical meter*), and that real numbers enter directly and ineliminably into the constitution of physical facts about distance, area, volume, and so on. The distance between two points, for example, consists in those points' being jointly related to some real number. It's possible that Relationists could make do with a somewhat less heavy-duty version of Platonism. It might be enough, for example, to assume that any two pairs of points are related uniquely to some real number *r*, in the sense that the distance between the first pair of points is *r* times the distance between the second pair. This still involves some degree of heavy-dutiness, since these relations to real numbers must be among the fundamental relations in the world. Field argues that this is undesirable, since it involves treating numbers as though they were themselves physical objects, whereas in fact numbers seem utterly separate from the physical world, however useful they may be in describing that world.

**No Heavy-Duty Platonism.** There are no fundamental relations between real (and complex) numbers and material things.

One way out of this problem is for Relationists to verify the standard axioms of geometry by positing the existence of *possible* material objects located at each point in space or the possible and counterfactual location of actual material objects at each point in space. This approach would be quite attractive to those who adopt Concretism (14.1T.1T) with Counterpart Theory (16.1A.1). We can suppose that there exists a concrete possible world, the *Plenum world*, with no vacuums at all (all points in space are occupied) and with counterparts of all actual bodies, with those counterparts standing in the same spatial relations (in that world) that corresponding actual bodies stand in in our world. We could then take the points of geometry to be small, point-sized bodies in the Plenum world, with an actual body occupying one of those points just in case the point is part of the counterpart of that body in the Plenum world.

At this point it is unclear whether there is any advantage in terms of ontological economy for Relationism. Relationists on this alternative are committed to an infinite number of possible material objects in an infinite number of concrete possible worlds, a totality at least as large in number as the real places of Substantivalism. There is perhaps some advantage to Relationists in that they can claim to make do with only one basic kind of thing, namely, material objects, whether actual or merely possible, rather than two, namely, material objects and spatial locations. (We will discuss the problem of vacuums and empty space further in Section 17.4.)

Instead of positing concrete possible worlds without vacuums, Relationists could instead simply "bite the bullet" and embrace heavy-duty Platonism. In fact, we discussed a version of heavy-duty Platonism in our survey of theories of quantity: Composite Intensity Theory (10.2T.3). According to Composite Intensity Theory, intensities or magnitudes are simply real numbers, and a body or ordered plurality of bodies has a certain magnitude of some quantity (like *distance*) by jointly instantiating the determinable universal (DISTANCE) and some real number. This would be an attractive option for Relationists who are Realists about universals and Abstractionists about possibilities.

There is a second argument for Relationism, in addition to ontological economy. Leibniz argued that Substantivalism needlessly multiplies possibilities. For example, if Substantivalism is true, then the universe could have been located six feet over, in some direction or other, from where it is in fact. If Substantivalism were true, then each of the following would represent a real possibility, distinct from the actual world:

1 Everything is exactly as it in the actual world, except that everything (both near and far) has been, is, and always will be located exactly six feet from its actual location in

the direction to which the earth's axis was pointing at noon Greenwich Mean Time on 1 January 2001.

2 Everything is exactly as it is in the actual world, except that everything in the universe has been, is and always will be rotated in an angle of 43° to the west along the line defined by the earth's axis at noon Greenwich Mean Time on 1 January 1 2001.

3 Everything is exactly as it is in the actual world, except that the entire universe is reflected, as in a mirror about the plane that coincided with the earth's equator at noon Greenwich Mean Time on 1 January 1 2001.

4 Everything is exactly as it is in the actual world, except that every length and distance is doubled compared to its actual value.

Of course, there are, for each of these four transformations, an infinite number of alternatives with different values for the magnitude of the transformation and different axes or planes of reference. For Substantivalists, each of these transformations corresponds to a distinct possibility, a different way the whole world could have been. For Relationists, in contrast, these supposed alternative possibilities would correspond to different conventions for describing a single set of spatial relations among the bodies. For example, a mirror-reversal like (3) would correspond merely to a difference in the meaning of 'left-' and 'right-handedness', not to a real difference in the relation between things and space.

Field (1984) argues that this needless multiplication of possibilities is especially problematic for Spatial-Quality Theorists, since on that account such possibilities would be *qualitatively* different. The Sixth Corollary of Ockham's Razor (**PMeth 1.6**) enjoins us to minimize the class of possibilities. This is a corollary of Ockham's Razor because the Razor asks us to prefer the theory that posits the smallest class of fundamental properties. The fewer fundamental properties there are, the fewer the number of distinct possibilities involving different combinations and permutations of fundamental properties. Substantivalists posit both fundamental relations of *occupation* (between material things and locations) and of *spatial distance*, while the Relationists posit fundamental relations only of the second kind. Arguably, Leibniz's argument provides a tie-breaker in favor of a modal version of Relationism over Substantivalism.

Isaac Newton offered a famous thought-experiment against Relationism in the General Scholium of his *Principia* (1687), the bucket experiment. Newton imagines a bucket containing water that is being swung in a circular path on a rope. As Newton hypothesized, the surface of the water will become concave, rather than flat, due to what could be described as the 'centrifugal force' of the rotation. This curvature of the water would distinguish the bucket from any bucket that is at rest or that is moving in a straight, unaccelerated path (i.e, an inertial path). Newton suggested that it is only the accelerated motion (in this case, circular motion) of the bucket relative to Absolute Space that can ground the difference between a bucket that does and one that does not experience these centrifugal effects.

However, Relationists have a number of effective responses to Newton's argument. Ernst Mach suggested that centrifugal forces depend on the motion of the bucket in relation to the mass of the earth and other celestial bodies (a hypothesis known as 'Mach's Principle'—see Mach 1960.). Still more effectively, Relationists can argue that accelerated and non-accelerated paths differ from each other intrinsically, and not due to changes in velocity relative to Absolute Space. All that is needed is a geometrical description of spatiotemporal relations that distinguishes the two kinds of paths, something that is accomplished by the so-called 'Minkowski' representation of spacetime, which is consistent with the non-existence of places.

Finally, Einstein's theory of general relativity may be relevant to the Substantivalism-Relationism debate. As standardly interpreted, general relativity includes the claim that spatial relations are merely part of a comprehensive spacetime structure, which includes both temporal and spatial relations in a flexible relationship that varies to some extent from one frame of reference to another. More importantly for our purposes here, general relativity entails that this spacetime is curved or warped by the presence of massive bodies. A massive body, like a large star, warps the shape of spacetime in its vicinity, resulting in the curving of the paths of moving bodies toward the center of the mass. This curvature is the relativistic explanation of the phenomenon of gravity. Instead of thinking of gravity as a force acting on bodies in a flat or uncurved spacetime, Einstein encouraged us to think that bodies move toward massive bodies by simply following the shortest, laziest path through local spacetime. For this reason, even mass-less entities like photons are curved toward massive stars, as if they were being attracted by a force. This curvature of light is actually observed experimentally.

If spacetime can be curved by massive bodies, then it seems that it must be something more than a set of relations between those bodies. It seems natural, at least, to think of such spacetime as an entity in its own right, subject to influence (in respect of its pattern of local shapes or curvatures) by the action of massive bodies. In support of this picture, the equations of general relativity enable us to describe a spacetime world that is entirely empty, unoccupied by bodies of any kind.

However, this Substantivalist tendency can be and has been resisted. First, one can refuse to believe that every possible solution of the general relativity equations represents a possible world. Perhaps an empty spacetime is metaphysically impossible, even if describable by consistent mathematics. Second, one can make a distinction between the bare, spatiotemporal relatedness of the world, the *spacetime manifold*, and the measurable shape or curvature of spacetime at various places, the *metrical field*. The spacetime manifold could be understood entirely in relational terms, and the metrical field could be interpreted as matter's effects upon those relations, and not as its effect on a substantial spacetime.

## 17.4 Absences and Vacuums

As we have seen, a critical difference between Spatial Substantivalists and Spatial Relationists concerns their treatments of empty space. For Spatial Particularists, empty regions of space are wholly unproblematic. They are ordinary, first-class entities, differing from occupied regions only in the absence of an occupying body. For Spatial-Quality Theorists, empty regions of space correspond to uninstantiated properties. For Spatial Relationists, in contrast, there are strictly speaking no empty regions of space; there are (at most) merely possible bodies that would (if they existed) be in contact with or at certain distances from certain actual bodies.

An empty region of space or vacuum would seem to consist in a certain absence of matter or bodies. Thus, before we consider vacuums in particular, we should examine the question of absences in general. Do absences really exist? Can all truths about absences be conceptually grounded in truths about bodies and their locations?

**17.3T Absentism.** There are real absences.

**17.3A Anti-Absentism.** There are no real absences.

There are three sets of arguments for the existence of absences: appeals to ordinary language and common sense, appeals to perception, and appeals to negative causation. We will consider negative causation in Chapter 27. Presently, we turn to the first two sets of appeals.

ARGUMENTS FROM ORDINARY LANGUAGE AND COMMON SENSE We talk about absences all the time. Consider (1) and (2):

(1) John's absence was widely noticed.
(2) The absence of rain ruined the crop.

Words like 'absence' (and 'lack', 'privation', 'omission', and 'deficit') are nouns, and noun phrases are typically used to refer to things. However, the mere fact that we often use a noun phrase does not by itself indicate that we really believe that something exists to which the noun phrase refers. If we can find a paraphrase that communicates the same information without implying that an absence exists, then we can ask ourselves whether or not common sense is really committed to the existence of the absence. (1) and (2) can easily be paraphrased using negative statements, as in (3) and (4):

(3) It was widely noticed that John was not there.
(4) The crops failed because there was no rain.

Unless we accept Truthmaker Maximalism (2.1T.1), there is no reason to think that something exists corresponding to negative statements like 'John was not there' or 'There was no rain.' Moreover, common sense dictates that negative things, like shadows, holes, and privations, are not as real as positive things. One way to make sense of this principle of positivity would be to suppose that there are no absences or other negative things.

So, ordinary language doesn't provide any strong evidence for absences, and common sense seems to weigh in on the other side. Hence, a successful case for absences will have to rely on evidence from perception or causation.

ARGUMENTS FROM PERCEPTION OF ABSENCES We often see or hear absences. This isn't always a matter of seeing or hearing that something is absent (a phenomenon that Fred Dretske (1969) has labeled 'epistemic' seeing). Sometimes we simply see something without knowing that we are seeing it. Barwise and Perry (1983) describe *scenes* or *situations* that can be the object of such non-epistemic seeing. In English, such non-epistemic seeing of a scene is expressed by means of *naked infinitive* phrases, as in (5) and (6):

(5) I saw Mary run.
(6) I saw that Mary ran.

(6) reports epistemic seeing, using a complementary that-clause, while (5) reports the non-epistemic seeing of a scene (one consisting of Mary's running) by means of the naked infinitive form of the verb 'run'. (5) can be true even if the speaker did not recognize Mary as the person in the scene, and even if the speaker did not recognize that it was *running* that the person was doing (as opposed to dancing or skipping).

Scenes and situations can involve absences. One can see Mary be absent, that is, one can see Mary's absence without seeing that Mary is absent. If so, it would seem that we must acknowledge that absences exist, in order that they may serve as the object of such non-epistemic perceiving.

Absences that we can see or hear include silence, darkness, the absence of dark spot on a light background, the absence of an elephant in the room, and a hole in a doughnut. If absences and other privations, like holes, are really perceived (see Sorenson 1984, 2009, Wright 2012), then they must exist, given the Reliable Perception Presumption (**PEpist 4**). Moreover, it seems that they cannot be mere logical constructions, since it is implausible to suppose that we perceive sets and other abstract entities.

One strategy for Anti-Absentists is to claim that we always perceive positive entities. For example, in supposing that we see holes, we are thinking in a confused way about what we are really perceiving. What we really perceive are either perforated objects or hole linings. (Both possibilities are explored in Lewis and Lewis 1970. See also Casati and Varzi 1994.) Each proposal has its difficulties. The main problem with treating *perforation* as a primitive property of material objects is that it leaves us unable to count holes. We would have to introduce an infinite set of properties—*singly-perforated*, *doubly-perforated*, and so on—and then we would have to add postulates allowing us to say that an $n$-ly-perforated object has the same number of holes as a set of $n$ objects has members. Casati and Varzi (1994: 25–30) point out a number of problems with the Ludovician theory that holes are hole linings. When we enlarge a hole, we do not necessarily enlarge the hole-lining. In some cases, the hole-lining becomes smaller in area. In addition, the inside of a hole is not the same as the inside of a hole-lining, since the latter is inside the hole's host, not within the space we ordinarily associate with the hole. Finally, to create a hole is not to create the hole-lining. It is rather to hollow out the material contained by the lining. Similarly, to remove a hole-lining is not to remove the hole; it is in fact to make the hole larger.

The Lewises responded that each of the words and concepts involved—*inside/outside*, *enlarge/shrink*, *create/remove*—are systematically ambiguous, with one meaning when we think of the hole/hole-lining as an ordinary material object and another meaning when we think of the hole/hole-lining as a hole or absence. This is certainly a cost of the theory. A theory that predicts ambiguity where none seems apparent is to that extent less credible (**PMeth 1.5**). In addition, it is far from clear that the Ludovician account of holes can provide a systematic account of all of the ambiguities that would be needed to make the theory work.

## 17.5 Conclusion

We've examined in this chapter six theories of spatial location, three Substantivalist, two Relationist, and one hybrid. The Spatial-Quality version of Substantivalism combined a high degree of elegance with a good fit with common sense, despite its unorthodox postulation of unperceivable "qualities" associated with each possible location.

Relationism is even simpler than any Substantivalist account, requiring no spatial entities and no absolute distinction between motion and rest or between right- and left-handedness. However, the great stumbling block for all versions of Relationism comes in the consideration of empty space. (This may be why some of the most prominent Relationists of the past, including Aristotle, denied the very possibility of a vacuum.) As Hartry Field has argued, vacuums deprive Relationism of the straightforward verification of the standard axioms of geometry and so deprive it of the aid of representation theorems that could verify the natural emergence of *distance* and other quantities from the inherent structure of space. Relationists have to choose between heavy-duty Platonism with metaphysically fundamental units of *distance* and reliance on geometrical facts about merely possible bodies filling the empty spaces, with the consequent worries about finding actual truthmakers for the possible facts.

## Notes

1 We will consistently use the word 'place' for a kind of thing, a part of space, and we use the word 'location' for the relation between a body and its place or between one body and other bodies.
2 This is equivalent to our definition of a weakly internal relation in Chapter 10, D10.2.1.1.
3 We do not mean to be ruling out the existence of immaterial beings here, only that if there are such things, that they are located in space in some direct way. So, for example, if there are Cartesian egos, then they are located in space only by being related to a body that is located in space. Which is to say: immaterial things, if there are any, are not determiners of relations of spatial distance.

# 18

# Structure of Space: Points vs. Regions

In this chapter, we take up whether space and extended bodies are ultimately composed of points (and point-masses) or spatial regions (and voluminous bodies). Here there are three positions: Pointillism, according to which only points and point-sized bodies are fundamental; Voluminism, according to which the only fundamental things are regions and voluminous bodies; and Volume-Boundary Dualism, according to which both points and regions really exist and are equally fundamental. It is possible to combine any of these three positions with either Spatial Substantivalism (17.1T) or Spatial Relationism (17.1A). For Relationists, the issue concerns only bodies: are the fundamental units from which all extended bodies are composed point-sized, dimensionless bodies, or voluminous bodies? For Substantivalists, the issue concerns both spatial regions and bodies. Most Substantivalists will accept some principle of Body-Space Correspondence: the ultimate bodies are point-sized if and only if the ultimate parts of space are points.

Modern analytic geometry, especially since the work of Descartes, normally takes it for granted that points are the fundamental geometrical entities, with lines, surfaces, and spatial regions defined as certain sets or collections of points. However, this assumption, although natural and easily carried out, is by no means the only way to proceed. Other mathematicians, like Alfred Tarski and Alfred North Whitehead, have shown that it is possible to start instead with spheres or other three-dimensional objects as fundamental, defining points as sets of regions. One can also take both points and regions as fundamental, as was done by many nineteenth-century philosophers of mathematics, including Franz Brentano.

This issue is tightly bound up with the issue of the relative fundamentality of *distance*, on the one hand, and *shape*, *volume*, and *contiguity*, on the other (see 17.2T and 17.2A). If *distance* is the fundamental relation, then it is plausible to take the fundamental entities to be points. Conversely, if *shape*, *volume*, and *contiguity* are the fundamental properties and relations, then it is plausible to take extended regions as the fundamental entities.

*The Atlas of Reality: A Comprehensive Guide to Metaphysics*, First Edition.
Robert C. Koons and Timothy H. Pickavance.

It is possible to hold, as Alexander Pruss has suggested (in correspondence), that *exact spatial distance* between regions is the fundamental relation. Informally, we understand the claim that the distance between region $A$ and region $B$ is $x$ to mean that $x$ is the shortest distance between two points, one of which is from $A$ and the other from $B$. Nonetheless, it could be that, metaphysically speaking, the *distance* relation holds between regions, and that "points" are a mere fiction, to be constructed mathematically from regions.

In Section 18.1, we look at the idea of *constructing* points from regions. This idea takes a little getting used to, since we are accustomed to thinking of regions as constructed from points, that is, as wholes containing points as parts. We also must more generally introduce *logical constructions*, a theoretical tool developed by analytic philosophers in the early twentieth century. Then we introduce, in Section 18.2, the basic terms of the debate for this chapter, which concerns whether points or regions are fundamental. In Section 18.3, we lay out three arguments against Pointillism, the thesis that points are the only fundamental entities, namely, the arguments from Finitism, from mathematical paradoxes, and from the nature of contact. Finally, we turn in Section 18.4 to a consideration of the two versions of Anti-Pointillism: Voluminism, and Volume-Boundary Dualism. Voluminism takes only voluminous regions to be fundamental, while Volume-Boundary Dualism includes both voluminous regions and points, surfaces, and other boundaries as equally fundamental.

## 18.1 Constructing Points from Regions

The construction of points as sets of regions was first proposed by Alfred North Whitehead (1919) using his "method of extensive abstraction". Alfred Tarski developed an especially elegant version of it in his classic paper, "Foundations of the Geometry of Solids" (Tarski 1956). Given a domain of regions, we can define a point as a certain set of regions; intuitively, a point is the set of all of the regions containing that point. We require that any two regions belonging to a single point overlap each other, and we also require that for every region within a point, there is a second region in that point that is a proper part of the first region. This ensures that, in effect, the set of regions mutually intersect in a single, dimensionless point, even if the original domain contains no such unextended regions. Once we have constructed a domain of such points, we can then add the usual geometrical axioms as postulates.

### 18.1.1 Digression on logical constructions

SET-THEORETIC CONSTRUCTIONS Alfred North Whitehead and Bertrand Russell introduced into philosophy the technique of *logical construction* in the early twentieth century. Making use of the mathematical theory of sets, they constructed mathematical objects that could fill the role in thought that was played by various kinds of entities. Thus, they promised a dramatic simplification of our metaphysical theories of the world. For example, instead of believing that the world consists of three things, points, regions of space, and sets, Russell and Whitehead showed that we could get by with just two, either sets and points or sets and regions. We can then identify things in the missing third category with special kinds of sets. We could suppose that spatial regions are nothing but sets of points or that points are nothing but sets of extended regions. Metaphysicians have used the method of logical constructions to reduce our ontological commitments by eliminating many different kinds of things from the list of fundamental entities, including propositions, properties, relations, and events.

The method of logical construction provides a way of carrying out the program of conceptual grounding or conceptual reduction that we discussed in Section 3.4. The proponent of a logical construction can argue that there is in reality nothing that fits exactly to the conceptual profile of the entity to be replaced by the construction. Instead, we appeal to the essence of the relevant *concept*, arguing that that concept is such as to play a role in the formulation of propositions that are made true by the facts described in the logical construction. Thus, logical constructions provide a means for ontological deflation of the original theory, showing how that theory can be true in a world with both fewer things and fewer kinds of things. We don't have to suppose that the things replaced by logical constructions really exist, although we can affirm the truth of ordinary propositions made in terms of the replaced entities.

In addition to the style of logical construction favored by Russell and Whitehead, there is a competing strategy that uses *mereology* in place of set theory. Mereology, developed by Lesniewicz, Goodman, and Leonard and discussed in Chapter 23, is the calculus of parts and wholes. However, as we shall see in Chapters 22 and 23, it is not at all obvious that a whole is nothing more than its parts (taken collectively), so it is not obvious that a mereological construction enables us to reduce the number of fundamental categories of things. In addition, the mereological strategy couldn't be used to reduce points to regions, since a sum of regions is an even bigger region, not an approximation to a point.

Logical constructions in this sense—the replacement of apparently concrete things by sets within a particular theory—have both advantages and disadvantages. The main advantage to a logical construction is greater ontological economy, as demanded by Ockham's Razor (**PMeth 1.4.1**). By eliminating a whole category of things and replacing them with sets, we can reduce the primitive qualitative complexity of the world as we represent it.

The principal disadvantage to a logical construction is that it forces us to eliminate or reduce the apparently *fundamental* properties and relations that were borne by the entities being replaced by sets. If we show that entities of kind *K* (such as regions of space) are in reality nothing but sets, then we cannot suppose that entities of kind *K* enter into any natural relations other than the natural or fundamental relations that sets enter into.

What sorts of natural or fundamental relations do sets stand in relations to other things, including other sets? As we have discussed, it is plausible to suppose that sets are involved in only a single natural or fundamental relation: that of *membership*. If *x* is a member of set *S*, then *x*'s *membership* in *S* is a fundamental fact about *S*, not reducible or explainable in terms of other relations. However, that's it. If *S* stands in any other relation to a thing, the holding of that other relation must be explained in terms of *S*'s members, and no other facts about *S*. This is Membership the Only Fundamental Set-Relation (**PMeta 3**).

PLURALIST CONSTRUCTIONS Besides sets and mereological sums, there is a third possible approach to treating regions or points as something other than fundamental entities: we could think of regions as *pluralities* of points or of points as *pluralities* of regions. On this view, a *plurality* is fundamentally 'they' rather than an 'it'.

Pluralist constructions could be used to eliminate regions in favor of points. To say that an object occupies a region is simply to say that there are some points that it occupies. To say that a region is connected is to say that there are some points and that any two of them can be connected by a path consisting entirely of some others of them.

Pluralism could also be used in the opposite direction, in parallel with the Whitehead-Tarski reduction of points to regions. To say that a body occupies a point is merely to say that there are some nested, concentric spherical regions and the body occupies some of them. (For more information on the *plural quantification* involved in the use of these plural pronouns, see Boolos 1998.)

Pluralists who reduce regions to points must embrace the view that all natural and fundamental properties and relations take points and only points as their relata. Similarly, pluralists who reduce regions to points must suppose that all natural relations take regions and only regions as their relata. Therefore, pluralists will have to suppose that there exist what are known as natural *multigrade* relations, relations that relate an indeterminate number of things, even infinitely many. For example, suppose that the points-only pluralist wants to say that a certain region has a volume of one cubic meter. This will have to be understood as a relation among points: there are some points, and these points stand together in the *filling-one-cubic-meter* relation, a relation that relates an infinity of things. If we extend the pluralist conception to extended masses, we can say that *having a mass of one gram* is a relation borne jointly by an infinity of point-masses.

LUDOVICIAN CONSTRUCTIONS David Lewis (Lewis and Lewis 1970, Lewis 1980a, 1993, 2004a) introduced, in the second half of the twentieth century, a quite different approach to replacing problematic entities with constructions. The strategy involves *Ludovician* constructions. Lewis's idea is closely related to the notion of "loose and popular identity" introduced by Joseph Butler (in Flew 1964: 166–172) and further developed by Roderick Chisholm (1976: 92–113, 1989: 25–41, 124–128). A Ludovician construction of points could be one that identifies points with very small regions, rather than with sets of regions. Intuitively, a point is a very small, spherical region. How small? Small enough for whatever is one's present purpose. Which small regions count as points depends on the context. For example, when discussing geometrical results in school, we often draw "points" and "lines" with chalk. The points are not really zero-dimensional objects. They are just very thin and relatively small smears of chalk on the board. Nonetheless, they are often small enough for the purpose at hand. The key is that whatever we count as points must have negligible volume (in the relevant context).[1]

Of course, if we count one very small region as a point, there will be infinitely many other, equally small regions that largely overlap the region in question. In fact, there will be infinitely many tiny spheres that share exactly the same center, intuitively speaking. We don't want to say that each one of them counts as a distinct point or else we'll be forced to contradict some basic axioms of geometry, such as the one stating that two distinct lines can intersect in at most one point. David Lewis's answer was to introduce the *same-X* relation—in this case, the *same-point* relation. If two tiny spheres are concentric, then they stand in the *same-point* relation. Each is counted as the same point as the other. Thus, if one line (a narrow tube) intersects another line (another narrow tube), there will be just one point of intersection, even though there are many tiny spheres in the intersection, since the many spheres will together count as just "one" point.

The *same-point* relation, as well as the similar *same-curve* and *same-surface* relations, can do additional work for Ludovician constructors. When a material body moves, its surface, with all of its attendant points and edges, moves with it. We cannot identify the surface of the body with a region of space or spacetime, since neither of them move with the body. Think of a ball that is rotating. The surface of the ball is also rotating, but the space in which the ball is sitting is not moving at all. The surface is the two-dimensional boundary between the ball and the surrounding environment. A Ludovician construction can enable us to replace such two-dimensional surfaces with very thin, three-dimensional skins or peels. Each material body has infinitely many such skins of a given thickness or less, but all of these can be counted together as the "same" surface, since they stand in the *same-surface* relation. There is no problem with saying that the ball's surface is rotating with the ball, since every outer layer of the ball, no matter how thin, is rotating.

The same idea could be applied to certain immaterial bodies or surfaces, such as holes or shadows. A hole or shadow could be identified at each moment with a region of spacetime, namely, the empty space constituting the hole at a moment or the region from which light is excluded by the occluding object. Shadows and holes can move through space, while regions of spacetime cannot. We can accommodate this fact by making use of appropriate *same-hole* or *same-shadow* relations. We can say that a shadow is moving through space so long as a series of spacetime regions, each located at a different time, are connected by the *same-shadow* relation.

The Ludovician construction has one principal advantage over the more traditional Whitehead-Russell logical construction: it identifies points (holes, shadows, etc.) with concrete things, like parts of material objects or parts of spacetime, and not with abstract objects, like sets. This enables us to say with more confidence that we can indeed perceive these things, and it enables us to assign them some real causal role in explaining phenomena.

The main drawback of Ludovician constructions is that it leads to a number of counterintuitive results. For example, it would be literally true that some points are larger than others or that some points overlap other, distinct points, even though it would be inappropriate to *say* these things in any given, fixed context. Everything that one refers to in one context as a dimensionless point could be accurately described in another context as having finite volume and containing many points as parts.

## 18.2 Points vs. Regions

We take it as obvious that space has parts. These parts are three-dimensional regions with finite volume. It also seems reasonable to concede that points, which are zero-dimensional spatial objects, exist in some fashion if places exist at all. However, it is not at all obvious which are fundamental. Are the fundamental things volumes, points, or both?

**18.1T Spatial Pointillism (Extreme Indivisibilism).** Indivisible, dimensionless parts of space (points) are (of necessity) more fundamental than extended regions; extended regions can be wholly grounded in points.

**18.1A Spatial Anti-Pointillism.** Necessarily, finite regions or volumes are at least as fundamental as points.

If *locations* are properties of material things, then Spatial Pointillism means that *point-location* properties are more fundamental than *region-location* properties, while Spatial Anti-Pointillism implies that *region-location* properties are at least as fundamental as *point-location* properties. If Spatial Pointillism is true, then a material object has a class of fundamental properties, each corresponding to some point that it occupies, while if Spatial Anti-Pointillism is true, then each material object has at least one fundamental property that corresponds to the region it occupies, or perhaps some finite set of fundamental *location* properties, each corresponding to some part of the region it occupies. Pointillism is also known as "Extreme Indivisibilism" because points are indivisible, and Pointillism implies that indivisible points are the fundamental spatial reality.

Spatial Pointillists advocate the conceptual grounding of truths about regions by fundamental truths about points. That is, they do not suppose that regions really exist, as entities with an essence that demands that their existence and properties be grounded in that of points. Rather, they suppose that it is the essence of our concepts of *region* and *volume* that permits us to explain how propositions ostensibly about regions and volumes can really be true in a world consisting only of points (or sets or sums or pluralities of points).

We are going to assume that space has its structure of fundamental entities as a matter of metaphysical necessity, whether those entities are points, volumes, or both. It seems very unlikely that space might be contingently made up of fundamental points and nothing else, or contingently made up of regions, or whatever. If space is a certain way in any world, it is essentially and necessarily so.

One issue concerns the relationship between one's view about fundamental spatial entities and one's view about the fundamental occupiers of space, namely, material bodies. We'll begin an investigation of this issue by considering the following plausible correspondence principle:

**18.2T Fundamental Entity Space-Matter Correspondence.** Necessarily, the fundamental spatial entities are points (as opposed to regions) if and only if the fundamental occupiers of space are point-sized bodies (as opposed to voluminous bodies).

Could Fundamental Entity Space-Matter Correspondence be false? There are two kinds of things that would be counterexamples to it: extended material atoms occupying a fundamentally Pointillist space and point-masses without any proper location.

**18.2A.1 Extended Material Atoms in a Pointillist Space.** There could be extended material atoms that occupy fundamental spatial points without point-sized parts occupying those points.

**18.2A.2 Material Simples without Proper Location.** There could be an extended body composed of infinitely many indivisible, zero-dimensional, volume-less material bodies, each without a unique fundamental location. That is, each indivisible body occupies infinitely many spatial regions, each of which is also occupied by infinitely many other bodies.

The likelihood that the thesis Material Simples without Proper Location could be true is remote. We would have to imagine a body that consists of point-sized fundamental entities, point-masses, which share a common extended region as their joint location, even though each of the point-masses have no unique fundamental location. But if the mass consists of point-masses, and there is some natural relation of *location*, then the point-masses must each bear the *location* relation to a unique and unshareable location, that is, a spatial point. Every body should have a unique, unshareable or proper location.

The truth of Extended Material Atoms in a Pointillist Space is more controversial. The main argument against the possibility of Extended Material Atoms in a Pointillist Space is given by Dean Zimmerman (1996a), and it goes like this. If some material entity *A* fills an extended region *R*, then there must be a part of the entity located in each part of the region. For example, *R* must have both a left and a right half. It seems reasonable to think that entity *A* must also have a left and a right half, one occupying the left half of the region *R*, the other the right half. This seems especially clear if we think of *A* as a mass of stuff of some kind. But in order for extended material atoms to be *atoms*, they cannot have any parts. So extended material atoms cannot occupy Pointillist spatial regions. Neither sort of counterexample to Fundamental Entity Space-Matter Correspondence is possible; that principle would seem to be true. And we have further happened upon another correspondence principle:

**Divisibility Space-Matter Correspondence.** If a mass *M* occupies an extended region *R*, and *R* has proper parts, then the mass *M* has proper parts corresponding one-to-one to those spatial parts.

There is an influential argument for Divisibility Space-Matter Correspondence, namely the "supercutting" argument of John Hawthorne and Brian Weatherson (2004). If there were extended material bodies in a Pointillist space, then it seems plausible that the following infinitary supertask would be possible. Take the extended material object *E*, and cut it into two pieces, leaving a small spatial gap between the two pieces. Repeat the process in half the time, splitting each of the halves into half. Keep repeating the process over and over, doubling the number of pieces in progressively shorter and shorter periods of time. After a finite period of time has elapsed, an infinite number of cuttings will have taken place, resulting in an infinite number of pieces, each with zero width and each some finite distance from any other piece. The same procedure could then be undertaken along the other two spatial dimensions, resulting in an infinite collection of point-sized bodies. However, if the extended body were a metaphysical atom, with no real parts, then the procedure would be impossible. Contradiction. Space-Matter Correspondence must be true.

Consider Spatial Anti-Pointillism, in which it is extended spatial regions (in three dimensions) that are metaphysically fundamental. We can ask whether these metaphysically fundamental regions have smaller parts. That is, we can ask which of Fundamentally Gunky Space and Discrete Space is true:

**18.1A.1T Fundamentally Gunky Space.** All fundamental volumes of space have proper, extended parts.

**18.1A.1A Discrete Space.** There are extended spatial simples.

*Gunk* was introduced by David Lewis as the concept of something which has proper parts, but also such that every part of it also has proper parts. A gunky thing is not simple and it has no simple parts. Every part of a gunky thing is divisible into still smaller parts.

If space is fundamentally gunky, then all of the regions of space, whether fundamental or not, have still smaller regions as actual parts. In contrast, if space is discrete, then the fundamental regions of space are simple, lacking any actual parts. There are therefore two versions of Discrete Space:

1 Berkeley-Hume indivisibles. There are indivisible spatial minima, each with a finite volume. On the most plausible version of this theory, each finite volume of space consists of a finite grid in three dimensions of cubical or dodecagonal "pixels".
2 Aristotelian discrete space with infinite divisibility. Each region of space has only finitely many actual parts, but each part is further *divisible* without limit. Each region is potentially divisible without limit, but some regions are not actually divided.

If we accept Fundamental Entity Space-Matter Correspondence, then we can reduce Spatial Pointillism and Spatial Anti-Pointillism to the following:

**18.3T Material Pointillism.** Necessarily, the only fundamental bodies are point-sized (dimensionless). Truths about extended bodies are wholly grounded conceptually in truths about point-sized bodies.

**18.3A Material Anti-Pointillism.** Necessarily, if there are any extended bodies (bodies with finite volume), then there are fundamental bodies with finite volume.

## 18.3 Arguments against Points as Fundamental

There is a certain initial plausibility to Material Pointillism. It seems to be implicit in our modern approach to geometry, which takes planes and space to be composed of parts, and in the analytic, bottom-up approach of modern physics, especially in the search for fundamental particles. There are three sorts of arguments against it and for Material Anti-Pointillism: (i) arguments from Finitism, (ii) arguments from the nature of contact, and (iii) arguments from mathematical paradoxes like the Banach-Tarski theorem. In addition, we should bear in mind the arguments for Priority Monism that we considered in Chapter 11, especially those of Bradley (Section 11.2.3) and Schaffer (Section 11.2.4).

The dispute between Material Pointillism and Anti-Pointillism would not be settled if it turned out that certain material bodies (like photons, for example) are point-sized, so long as it was necessarily the case that extended bodies are not composed of an infinite number of such fundamental, point-sized bodies. Material Anti-Pointillism does not entail that there couldn't be fundamental bodies that are point-sized; it simply entails that, if there are extended bodies, then not all fundamental bodies are point-sized.

We are going to assume that extended bodies are metaphysically possible.

**The Possibility of Extended Bodies.** It is possible that material bodies exist with finite volume.

Consequently, we will take Material Anti-Pointillism to entail that metaphysically fundamental extended bodies are possible. In contrast, Material Pointillism clearly entails that such fundamental extended bodies are impossible.

We will now turn to the three arguments for Anti-Pointillism.

### 18.3.1 Finitism

One way of rejecting Material Pointillism is to be a spatial Finitist or a Finitist about the whole concrete world.

**18.4T Finitism.** There are (with the possible exception of sets and numbers) only finitely many actually existing things.

**18.4A Infinitism.** There are infinitely many actually existing things, other than sets and numbers.

If Finitism is true, and there are actually existing extended things, then Spatial Anti-Pointillism and Material Anti-Pointillism must be true, since any spatial region contains infinitely many points, as well as infinitely many subregions.

**Infinite Numbers of Points.** Any finite spatial extension contains infinitely many points.

How can it be the case that there are not infinitely many spatial entities, given the obviously infinitary nature of geometry? Finitists have two options. First, they could maintain that space is discrete, consisting of a finite number of extended but indivisible places, a kind of minimum unit of spatial volume. George Berkeley and David Hume accepted such a view, as did some ancient Greek philosophers. This is Discrete Finitism. Second, as an alternative, Finitists might place special emphasis on the idea that only finitely many things exist *in actuality*, as opposed to potentially. The infinitely many points, lines, and planes, and the infinitely many extended regions of standard geometry must be thought of as being merely potential in nature. This was the position of Aristotle and of his many followers, so this is Aristotelian Finitism.

Although Finitism entails that Pointillism is false, it doesn't follow that every Anti-Pointillist must embrace Finitism. It is possible to be an Anti-Pointillist and an Infinitist. This was in fact the position of Whitehead, who believed that there are infinitely many voluminous regions in any finite region, but also that there are no points. However, Anti-Pointillists who embrace Infinitism lose the finitary arguments against Pointillism that we will consider in this section.

The arguments for Finitism are *reductios*, attempts to show that granting the possibility of actually infinite entities in space leads to metaphysical absurdities. There are two kinds of infinity to consider: entities of infinite extension or volume and entities with finite extension that include an infinity of actual point-masses.

If there could exist an extended entity of infinite extent or volume, then such an entity could exist with a finite, homogeneous density. Such an infinite object would be infinite in mass. Consequently, its momentum and kinetic energy would be infinite. Moreover, if such an entity could exist, then it could move with a constant velocity. A moving entity of infinite mass would contain infinite energy and momentum.

That this is absurd is clear if we imagine a scenario in which there are two such entities. Suppose, for example, that there were two infinitely long rods, each with a circular face behind which the matter extends infinitely far into the distance. These two rods could be located along a common axis, and they could be moving toward one another at a constant velocity. At some point, the two circular faces will collide, and each will carry an infinite momentum behind it. What will happen at the collision? Each possible answer seems absurd. We might suppose that symmetry requires us to suppose that the two rods will bounce off each other, leaving the collision with an equal and opposite velocity. However, it is easy to imagine cases in which symmetry is no help. Suppose, for example, that one rod has a cross-section that has twice the area of a cross-section of the other. Then we would seem to be forced to say that the one infinite momentum is twice as great as the other infinite momentum, which seems to be utter nonsense. At least, when choosing between competing theories and models, physicists treat real infinities as something to be avoided.

Perhaps, however, physicists are guilty of a mere prejudice against the infinite. In addition, even if they are right to avoid infinities in describing the actual world, it doesn't follow that physical infinities are absolutely impossible. There is a consistent mathematical theory, non-standard analysis, which takes seriously the possibility of infinitely large and infinitely small (or infinitesimal) real numbers. Since such non-standard real numbers satisfy all of the usual laws of real analysis, we could presumably use them to build a theory of infinite forces and masses that would obey the laws of physics. Alternatively, if we were to suppose that the idea of *infinite real numbers* is a mere mathematical fancy, we might still suppose infinite magnitudes to be possible by supposing that they obey different, less quantitative laws of physics.

Perhaps the objection to infinite force is rooted in an objection to infinite acceleration and infinite velocity. Imagine an infinitely long pool stick striking a finite ball. The ball would experience an infinite force at the moment of collision. An object under an infinite force would, according to Newton's laws of motion, undergo infinite acceleration. Infinite acceleration would seem to bring an object instantaneously to an infinite velocity, but what sense can we make of an infinite velocity? Where is the object while it is moving infinitely fast? All over the place in a single instant? That seems to be impossible, assuming that objects cannot be wholly located in more than one place at a time. So, we seem to have good objections to infinite velocities and infinite acceleration.

However, given Einstein's special theory of relativity, ruling out infinite velocity will not enable us to rule out infinite forces, since Einstein's theory tells us that an infinite force would accelerate a massive body instantaneously only to the speed of light. Nothing can go faster than that, no matter how much force is applied. If the universe contained an infinitely great force of gravity, the result would be that everything would move toward the source at the speed of light, disappearing into a massive black hole— an undesirable result, no doubt, but not metaphysically impossible.

Still, we might conclude that the idea of infinite mass is dubious at best. If so, we have good grounds for denying the very possibility of an entity infinite in volume. If such an entity is impossible, there is also some reason to think that an entity finite in volume but infinite in mass is also impossible if the latter is composed of infinitely many point-masses. Suppose, for example, that we had an entity with finite volume but with varying density. If this entity is composed of an infinity of point-masses, there is no reason that the density can vary within it in any way that can be described mathematically. Let's suppose that the entity is a cube 1 meter long on each of its edges. The first half of the cube has a uniform density of 2 grams per cubic meter, the next fourth has a density of 4 grams per cubic meter, the next eighth a density of 8, the next sixteenth a density of 16, and so on ad infinitum. Each of these infinitely many parts of the cube has a mass of 1 gram. Hence, the cube as a whole has an infinite mass. If we want to rule out the very possibility of infinite mass, it seems that we need to rule out point-masses altogether. To do that, it seems we would have to rule out the possibility of an infinite number of actual places.

Thus, we have some reason to suppose that any extended entity has only finitely many actual parts. Hence, it is impossible for an entity to have an infinite volume or an infinitely varying density. Consequently, every possible entity has finite mass. If, as seems reasonable, we suppose that infinite velocity is also impossible, then every entity will have some finite kinetic energy and momentum, avoiding the dynamical absurdity of colliding infinities.

There is an obvious objection to this line of argument against infinitary composition. We have some reason to suppose that infinitely massive bodies are impossible. Why not stop there? Why go further and ban infinite composition? Why suppose that all entities with infinitely many actual parts are impossible? The answer to this objection involves an appeal to a familiar methodological maxim: avoid brute necessities (**PMeth 1.2**). Further, where possible, we should explain all impossibilities in terms of hypotheses about the essential structure of space and time. This is Structuralism (**PMeth 3**). The point of Structuralism can be brought out by considering patchwork principles (discussed in Chapter 16). Here again is Infinitary Patchwork:

**PMeta 5.2 Infinite Patchwork.** If *T* is a class of types of events or processes, and for each member of *T*, it is possible for an event or process of type *T* to occur, and there is enough room in the history of the world to locate within it instances of each of the types in *T* without overlap in space and time between the instances, then it is possible for all of the types in *T* to be realized together.

If a body can be composed of an infinite number of distinct, non-overlapping material parts, each with its own unique, unshared location, then that body and its parts provide a spatial framework that could be occupied (in other worlds) by bodies of steadily increasing density. So long as there is no upper bound to the possible density of a body, we can use Infinite Patchwork to find a counterpart of our infinitely divided body that has infinite mass in some possible world. If such an infinitely massive body is in fact impossible, then the Patchwork principle entails that there cannot be an infinitely subdivided body in any world, and so Finitism must be true.

Aristotelian Finitism, the view that space and space-occupiers can never have an infinite number of actual parts, together with the assumption that mass-density can only take finite values, provides a structural account of why finite bodies cannot have infinite mass. This is much to be preferred to the view that the impossibility of infinite mass is a brute, inexplicable necessity.

On this score, Discrete Finitists are in a somewhat stronger position than Aristotelian Finitists, since Discrete Finitists can plausibly claim that it is impossible for any region or any body to be smaller than some fixed lower bound, the minimum unit of space. Aristotelian Finitists want to place no fixed lower bound on the possible size of bodies. Any body, no matter how small, could be subdivided further. Instead, Aristotelian Finitists claim that it is impossible for a body to be actually infinite in its compositional complexity. This seems somewhat unprincipled. If we can subdivide bodies without limit, why suppose that it is metaphysically impossible for a body to be actually subdivided into an infinite number of parts?

Aristotelian Finitists could respond by arguing that infinite subdivision of a body is impossible because of the structure of time. No process could complete an infinite number of sub-processes in a finite period of time. We will consider this sort of argument in more detail in Chapter 19.

There is another route that Spatial and Material Pointillists could take. They can deny that extended bodies ever fill all of the points of any extended region. Instead, all extended things of necessity consist of a finite number of point-masses embedded within a continuous spatial region. The apparent impenetrability of the matter could then be explained in terms of forces of repulsion between the particles, forces that increase exponentially as the particles approach one another in space. This was proposed by Roger Boscovich in 1763 (1966), building on the atomism of Isaac Newton.

However, this hypothesis will explain the impossibility of infinite mass or momentum only if the existence of these repellent forces are themselves a matter of metaphysical necessity, that is, only if it is absolutely impossible for there to exist point-masses that do not repel each other in this way. This seems implausible. Even in the actual world, it is likely that some particles, such as photons, neutrons, and neutrinos, exist that do not repel one another in this way. If so, there would be nothing absurd in a situation in which an infinity of neutrinos exist, occupying every point within a region. It seems that we must instead posit some necessity inherent in the very nature of space that prohibits such a situation from happening. It appears that particles like protons are not in fact point-masses, since the concentration of their finite matter and charge into a dimensionless point would result in infinite density, with consequent problems concerning self-interaction. (For example, a charged particle will interact with its own electrical field. If the charge is concentrated in a dimenionless point, this results in a field of unbounded intensity in the neighborhood of the point. Interaction with this unbounded intensity results in paradoxical infinities of force and motion.)

Even if it were the case that every particle repels every other particle with a force inversely proportional to distance, this would at most explain why it is impossible to compress a finite body into a condition of infinite density. It wouldn't explain why it would be impossible for an infinity of such particles to exist right from the beginning, or for all eternity, in such a space-filling, infinitely dense condition. Such a condition might seem absurd, since it would involve infinite mass and infinite potential energy, but we want to avoid a merely ad hoc prohibition of such infinite quantities. The repellent particles model fails to do so. (A similar problem faces Aristotelian Finitists: if bodies are potentially divisible without limit, why couldn't a body be actually divided into infinitely many parts for all eternity?)

ARGUMENT AGAINST FINITISM The main argument against Finitism relies on the fact that our best scientific theories, from the time of Newton to the present, represent space as an infinitely divisible continuum through which bodies take differentiable paths. This argument is based on Scientific Realism:

**PMeth 2 Scientific Realism.** Other things being equal, adopt the theory that implies that our best scientific theories are straightforwardly true, as standardly represented.

Since our best physical theories make reference to infinitely many points, curves, and surfaces in space, Scientific Realism supports a metaphysical theory that entails that there really exists such an infinity of spatial entities. This scientific argument against Finitism presupposes that Finitists also embrace Actualism (12.1T), the claim that everything exists. If the Finitists are Possibilists (12.1A.1T) or Meinongians (12.1A.1A), they can argue that Finitism does imply that our best scientific theories are true, since they quantify over all potential spatial entities, of which there are indeed an infinite number. To refute Possibilist-Finitism, Infinitists will have to argue that our best scientific theories entail that infinitely many points, surfaces, and regions are actually in existence.

Infinitists could argue that field theories do have such an implication, since they attribute causal influence to each point of the field, with the field itself being merely the aggregate of an infinite number of point-intensities. However, quantum field theory suggests that at very small scales, in the vicinity of the Planck length, physical fields must take on a discrete, chunky character, in order to avoid certain dynamic infinities resulting from self-interaction (the problem of *re-normalization*) (see Georgi 1989).

### 18.3.2 Mathematical paradoxes, from Zeno to Tarski

Second, there are arguments against Pointillism from mathematical paradoxes. The ancient Greek philosopher Zeno (as reconstructed by Skyrms 1983) poses the following paradoxical argument:

1 A line segment consists of an infinite number of mutually disjoint parts.
2 The concept of magnitude (length, area, or volume) applies to each of the parts.
3 The parts of the line segment all have the same magnitude, either zero or positive.
4 There are no infinitesimal magnitudes (all positive magnitudes are finite). (Archimedes's Principle)

5 The magnitude of a whole is the sum of the magnitude of the parts, taken in sequence.
6 Any infinite collection can be ordered into a sequence. (Zorn's lemma, which follows from the Axiom of Choice in set theory)
7 The sum of an infinite series of zeros is zero, and the sum of an infinite series of equal finite magnitudes is infinite.
8 Therefore, any line segment is infinitely long or has zero length.

Modern measure theory rejects premise (5). It insists that the measure of a whole is equal to the sum of the measure of its parts taken in sequence only when the sequence is either finite or *countably* infinite (equinumerous with the counting or natural numbers, 0, 1, 2, etc.). Modern mathematical theory, following Georg Cantor, entails that there are *uncountably* many points in a line, a very "large" infinity to which the additivity assumed in premise (5) does not apply. However, it is not the case that we cannot make sense of the idea of the summation of an uncountable collection. Given the Axiom of Choice, the process of summing such a collection is straightforward, as Skyrms observes, and the resulting sums will in fact satisfy premise (7), which follows from (5) and (6). If we repeatedly add zero to itself, the sum will remain zero, no matter how long the sequence. The sum of uncountably many zeros is well defined and is, as Zeno argued, equal to zero. Consequently, there is some significant metaphysical cost to be paid for adopting modern measure theory, namely, the rejection of the very plausible premise (5).

What about giving up premise (4), Archimedes's Principle? In order to do so fruitfully, we would need to have a coherent account of infinite and infinitesimal quantities. The only real candidate for such an account is Abraham Robinson's theory of nonstandard analysis, developed in the 1960s. All previous theories of infinitesimals were beset with antinomies and contradictions. However, if we adopt Robinson's theory, we can still derive an unacceptable conclusion, since we can add premise (9), deriving a further absurd conclusion:

9 The sum of an infinite series of Robinson infinitesimals is infinitesimal.
10 Therefore, any line segment is infinitely long or has zero or infinitesimal length.

THE BANACH-TARSKI PARADOX According to modern mathematics, building on the work on transfinite set theory by Georg Cantor, the set of points making up a finite segment of a line can be put into a one-to-one correspondence with all of the points in the whole of space! In this sense, there are just as many points in the segment as in any region of space, whether finite or infinite. Consequently, we cannot suppose that the mass or volume of a region corresponds in any way with the number of points in that region.

However, we might suppose that mass and volume are a function of the number of points together with their arrangement in space. That is, it seems natural to suppose that any two congruent arrangements of points have the same volume. Similarly, two congruent arrangements of qualitatively identical material points should have the same mass and density.

In addition, it seems plausible that if a region of space can be decomposed into a finite number of subregions, then the volume of the whole should be the sum of the volume of the parts. Likewise, the mass of the whole should be the sum of the masses of the parts.

Surprisingly, if we put these two natural assumptions together, we get a conclusion that is demonstrably false. It is possible to disassemble a finite region into a finite number of parts (as few as five), move the parts through space without disturbing the mutual relations among the points involved (in other words, preserving congruency), and yet end up with a region twice as large as the original. Similarly, we could disassemble a material body into five parts, rearrange those parts through rigid motions, and produce a new body with twice the volume and twice the mass of the original. The result is called the Banach-Tarski theorem. There are three spheres, $A$, $B$, and $C$, each with a radius of one unit. Sphere $A$ is composed of five regions $E_1,\ldots, E_5$, sphere $B$ is composed of $F_1$ and $F_2$, and sphere $C$ is composed of $F_3$, $F_4$, and $F_5$. Banach and Tarsksi proved that if the decompositions are defined in a particular way, we can prove that $E_1$ is congruent with $F_1$, $E_2$ with $F_2$, and so on. Thus, sphere $A$ can be decomposed into five parts, and those five parts can be re-arranged in space (the first two parts in one group, and the other three in a second group), in such a way that two spheres are produced, each equal in volume to the original. The first sphere has "magically" doubled in volume!

The Banach-Tarski theorem gives rise to the following argument against Pointillism:

1 Any set of points in a specific geometrical arrangement (any "region") has a definite magnitude (volume).
2 If two regions are congruent, then they have the same measure. (Congruence-invariance)
3 Measure is finitely additive: the measure of the fusion of a finite collection of disjoint regions is the sum of the measures of the members.
4 There are regions $E_1,\ldots, E_5$ and $F_1,\ldots, F_5$ such that $E_1$ and $F_1$ are congruent, $E_2$ and $F_2$, and so on, $E_1,\ldots, E_5$ compose a unit sphere, $F_1$ and $F_2$ compose a unit sphere, and $F_3,\ldots, F_5$ compose a unit sphere. (Tarski-Banach theorem)
5 Therefore, one unit sphere has the same volume as the sum of the volumes of two unit spheres.

The standard mathematical solution of the paradox is to stipulate that the sets produced by the construction are *unmeasurable*. This constitutes a denial of premise (1). The paradox demonstrates that it is impossible to assign *any* value to any of the five parts. Mathematicians have developed an entire branch of mathematics, Lebesque measure theory, designed to cope with this phenomenon. Lebesque measure theory puts precise mathematical conditions on the class of sets that are measurable, requiring that all other sets of points be treated as lacking volume, and thus also lacking mass or charge. Common sense tells us that whenever an extended region of space can be divided into a finite number of subregions, the volume of the whole must be the sum of the volume of its parts. Measure theory denies this. The additivity of volume only makes sense when the regions correspond to measurable sets of points.

Brian Skyrms draws this conclusion:

> At this stage of the game, the elimination of non-measurable sets may appear a rather quixotic goal. Lebesque measure has extended measurability to a far richer domain than Zeno and Aristotle imagined possible. It meshes with an elegant and powerful theory of integration adequate to the needs of the physical sciences. Perhaps Lebesque measure should be taken as the theory of measure for physical space, and the existence of non-measurable sets

should be viewed as just a mildly surprising consequence of the theory rather than as a real difficulty. This is, I believe, the dominant view among mathematicians and mathematical physicists. ]source[(Skyrms 1983: 247)

Despite Skyrms's sanguine attitude, the upshot seems to be some support for Aristotelian Finitism, according to which extended regions are not fusions of sets of points. Three-dimensional volumes may have points as parts *of a sort* (i.e., as parts of the region's internal and external boundaries), but they are not constituted, without remainder, of points, even when the points fill the space. If regions were fusions of sets of points, then there would be no good reason to deny that every set of points constitutes a region. Aristotelian Finitists can explain why some sets of points have measurable volume and why some do not. Those sets of points that are interior to some region can be assigned a volume, and those that are not, cannot.

### 18.3.3 Spatial boundaries and physical contact

Our third and final argument against Pointillism concerns the nature of contact between bodies. These arguments have their roots in medieval and late scholastic philosophy, especially the philosophy of Francisco Suárez, and were further developed in the early twentieth century by Franz Brentano (Brentano 1976/1988—these manuscripts were published after his death by Brentano's students). They can be found in recent essays written by Dean Zimmerman (1996a and 1996b).

These arguments proceed, as we shall see, from a somewhat naive and pre-scientific view of the nature of matter. They assume that an extended body occupies a volume of space that is absolutely filled with some material substance, in such a way that two such bodies can come into direct contact with one another. Modern views of matter (since the sixteenth century) have been *corpuscular* in that they assume that matter consists of a finite number of particles, widely separated from each other and held together by electrostatic forces. Does the unrealistic or counterfactual nature of these scholastic and Brentanian arguments render them unpersuasive?

There are at least two reasons for taking these arguments seriously. First, even if they assume facts about matter that are not realized in the actual world, we can still consider them thought experiments about a metaphysically possible world. So long as it is possible for two material bodies that fill volumes of space to come into direct contact, the argument provides support for Anti-Pointillism. Second, there is a historically important way of interpreting modern quantum theory, namely, the Copenhagen interpretation, according to which reality must be divided into two separate and incommensurable realms, namely, the quantum realm of microscopic particles and the classical realm of observable bodies. Even if matter within the quantum realm consists of isolated particles, it doesn't follow that classical bodies may not come into direct contact.

Here's the thought experiment. Suppose that we have two spheres completely filled with some homogeneous stuff. It should be possible for the two spheres to come into contact with each other. When two perfect spheres touch, there is a single point of contact between the two—call it '*P*'. This spatial point is occupied by a point-mass, *M* (on the assumption that every filled location corresponds to a unique physical object). This point of contact will also be on the surface of the two spheres. Now suppose that we separate the two spheres. Each sphere will now have its own point on its surface, $P_1$ and $P_2$, corresponding to the former point of contact, $P$.

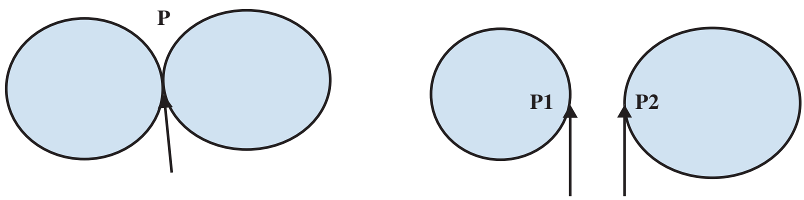


**Figure 18.1** Two Perfect Spheres Lose Contact.

The problem is whether points $P_1$ and $P_2$ filled with matter. Here are the possibilities:

1 Neither $P_1$ nor $P_2$ are occupied by matter.
  1a. No points on the surface of physical objects are occupied by matter (so, the point of contact $P$ is also unoccupied).
  1b(i). Some points on the surface of physical objects are occupied, and some are not. In this case, none of $P$, $P_1$, or $P_2$ are occupied.
  1b(ii). Some points on the surface of physical objects are occupied and some are not. In this particular case, $P$ is occupied, but neither $P_1$ nor $P_2$ are occupied.
2 Exactly one of $P_1$ and $P_2$ is occupied by matter.
3 Both $P_1$ and $P_2$ are occupied by matter.
  3a. One of these is occupied by $M$, the other by a new point-mass.
  3b. Both are occupied by new point-masses.

Two of these options, 1b(ii) and 3a, seem absurd immediately. Option 1b(ii) requires the point-mass $M$ that occupied $P$ to be annihilated simply by separating the two spheres. Options 3a and 3b requires that one or two point-masses are created *ex nihilo* (out of nothingness) by the separation. Such annihilation or creation cannot be a necessary consequence of simply moving the two spheres apart. In other words, the following principle seems highly plausible:

**PNatPhil 2 Independence of Motion and Substantial Change.** Neither the movement of two extended things nor the division of an extended thing into two parts can necessitate the creation or annihilation of material parts (not even point-masses).

Two of the options, 2 and 3a, require something asymmetric to happen as a result of a purely symmetric operation. But metaphysical necessities ought to be grounded in the natures of the things involved (PMeta 3), and there is nothing to distinguish the nature of one sphere from the other in order to ground such a difference in results. This is Symmetry:

**Symmetry.** Wherever possible, avoid theories that require violations of symmetry (e.g., that require symmetrical operations to result in asymmetric outcomes).

It might seem that option 3b is objectionable for the same reason 3a was, since 3b involves the creation *ex nihilo* of two new point-masses. However, this need not be a consequence of 3b: we could suppose instead that the original point-mass $M$ had two parts, $M_1$ and $M_2$, each with one-half the mass-density of $M$ and both co-located with $M$ at the same very point. When the spheres are pulled apart, the two halves are split. This does involve some violation of symmetry, nonetheless, since we cannot explain why one of the parts went one way and the other the other way.

But there is a second, weightier objection to 3b. The two parts $M_1$ and $M_2$ must each have half the mass-density of the original point-mass $M$ (where the density of a point-mass $x$ is equal to the average density of a region filled with point-masses intrinsically identical to $x$). This means that after the splitting of the two spheres, the density of each of the two spheres varies discontinuously, with a sharp discontinuity in density occurring at the location of $M_1$ and of $M_2$. Alternatively, the density of the two spheres-in-contact must have had a discontinuity at the location of $M$, which would have had to have twice the density of all of the surrounding matter. Each of these seems to be impossible. If the average density, and hence the total mass, of a body is to be well defined, the mass-density of its parts must vary continuously. Thus, option 3b is in conflict with Mass a Function of Density:

**Mass a Function of Density.** The mass of any extended body is a function of the density of its parts, and the mass of such a body is always well defined.

Thus, the alternatives left are 1a and 1b(i). Both require that the relevant points were never occupied. Option 1a means that the space occupied by matter is always *topologically open*, lacking its own external boundary. A physical sphere must fill all of those points that are less than a certain distance from the center, not the points that are located exactly that distance from the center on the dimensionless skin around the sphere's interior.

However, option 1a violates another very plausible principle:

**Unlimited Divisibility of Matter.** Any extended mass can be divided (separated) into parts along any two-dimensional surface passing through its interior.

Consider a single sphere $A$. Take a circular cross-section of $A$ that passes through its center, and call this circular surface $C$. $C$ divides $A$ into two hemispheres, $A_1$ and $A_2$. Suppose that sphere $A$ is completely filled with matter. Consequently, $C$, $A_1$, and $A_2$ are all fully occupied by this matter. Now suppose that we separate $A_1$ and $A_2$. The matter filling the cross-section $C$ must be annihilated, on the assumption that the two separated hemispheres must each fail to occupy their outer boundaries. So, either the mere division of the sphere necessitates the annihilation of some matter, violating the Independence of Motion and Substantial Change or the sphere is not divisible. Either result is absurd.

Therefore, we seem to be forced, by process of elimination, to adopt option 1b(i), the view that surfaces of bodies are partly closed and partly open. That is, that some of the points on the surface of bodies are occupied and others are not. This view has a number of odd metaphysical consequences. For example, this view requires that two bodies can be pushed into contact at a point only if one of the bodies has an open surface at that point and the other has a closed one. If both surfaces are closed, then the two occupied points will exclude each other, so that the two bodies must always be some finite distance apart. This seems absurd, as can be seen by considering the plausibility of this principle:

**Potentiality of Contact.** Any two extended masses can be brought into mutual contact.

Potentiality of Contact might be taken to be a consequence of the Principle of Epistemology 1: that imagination is a guide to possibility. We can certainly imagine (in some loose sense) the contact of any two extended bodies. But, if both surfaces are open, then once again the two bodies cannot be brought into contact, since what would be the point of contact between them must always be either unoccupied or occupied by a point-mass that belongs to neither body. This would also violate Potentiality of Contact.

## 18.4 Voluminism vs. Volume-Boundary Dualism

If we conclude that spatial regions and material bodies are not composed of points, that is, if we conclude that points are not metaphysically more fundamental than three-dimensional regions and bodies, then we still have a choice concerning points and other superficial entities. We might suppose that there are no points at all or that points are merely logical constructions, whether sets of regions, pluralities of regions, or Ludovician very small regions. Alternatively, we might believe that regions have boundaries, like surfaces, edges, and point boundaries, even though they do not consist entirely of points. That is, we might consider the possibility that both regions and points are fundamental entities.

**18.1A.2T Voluminism.** Entities of fewer than three dimensions (like boundaries that are points, curves, and surfaces) are not fundamental entities: either they do not exist at all or they are mere logical constructions from regions and extended entities.

**18.1A.2A Volume-Boundary Dualism.** Entities of fewer than three dimensions (like boundaries that are points, curves, and surfaces) exist and are not mere logical constructions, although they are not *more* fundamental than extended things.

Voluminists either eschew belief in points altogether or replace points with logical, pluralist, or Ludovician constructions. Volume-Boundary Dualists accept both points and regions as fundamental entities, as well as material point-masses and bodies. Moderate Volume-Boundary Dualists (following Aristotle) believe only in the actual, external surfaces of material bodies. However, we will assume that typical Volume-Boundary Dualists believe in all of the points, curves, and surfaces of classical geometry, including the ones that are interior to a material body. Volume-Boundary Dualists believe that both boundaries and material bodies are G-fundamental: that is, their existence is not wholly grounded in the existence of things of the other category. However, neither boundaries nor material bodies are O-fundamental (in the sense of definition D3.6): the essences of bodies include the existence of boundaries, and the essences of boundaries include the existence of bodies. The two categories of things are equally G-fundamental but ontologically interdependent (as we discussed in Section 3.3.2).

ADVANTAGES OF VOLUMINISM Voluminism has at least two advantages over Volume-Boundary Dualism.

1 Voluminists have a much smaller and more uniform ontology (**PMeth 1**), consisting entirely of three-dimensional regions and bodies, plus sets and other abstract objects. In contrast, Volume-Boundary Dualists posit the existence of two kinds of fundamental concrete objects, both three-dimensional regions and bodies and zero-, one-, and two-dimensional boundaries.
2 Volume-Boundary Dualism requires additional necessary connections between distinct entities, since material boundaries cannot exist without material bodies and bodies cannot exist without boundaries (see **PMeth 1.2**). In addition, we cannot have regions with points, curves, and surfaces, nor points or other surfaces without regions. Voluminism requires no such necessary connections, since points either don't exist at all, are identical with small bodies, or are mere sets or pluralities of regions.

There is one qualification to the last point. Voluminists who identify points with sets of regions do have to accept the necessary connections associated with set theory. For example, it is a necessary truth that if $x$ exists, then the set containing $x$ also exists and contains $x$ as a member. This fact could neutralize the second advantage if Volume-Boundary Dualists reject set theory.

ADVANTAGES OF VOLUME-BOUNDARY DUALISM Volume-Boundary Dualism has two advantages over Voluminism, however.

1 Suppose that there are natural, fundamental properties or relations of a physical or geometrical character that are borne by points, curves, and surfaces. If so, then these things cannot be mere sets, since we've assumed that the only natural relation that sets enter into is the *set-membership* relation. Modern physics does typically assign certain fundamental properties, like *density* or *gravitational* or *electromagnetic field strength*, to points in space, and not to whole regions. If physics is correct in doing this, it provides a powerful argument against taking points to be logical constructions. We'll take this up in Section 18.4.1, on continuous variation.

   Voluminists who use a Ludovician construction will also have problems with natural properties that are assigned to points, since physical theory would not assign determinate values of those properties, like *density*, to small volumes, no matter how small. A pluralist construction might work, however, taking a property of a point to be the collective property of a plurality of regions (one that converges on that point).
2 In our naive, pre-scientific thinking about physical objects, we need to talk about points, curves, and surfaces in order to describe various kinds of contact or fusion between those objects. For example, two perfect spheres would, if they were to come

into contact, meet in a single point. It is hard to believe that such a real point of contact could be some kind of set of overlapping bodies. The Ludovician construction is more reasonable, but it involves positing a primitive, unanalyzable relation of *contact*, while Volume-Boundary Dualists can analyze *contact* in terms of sharing a common point, edge, or surface. A pluralist construction could take care of this problem, since it would enable us to define *contact* in terms of sharing a common point, where the point itself is identified with a plurality of regions. We'll take this up in Section 18.4.2, on coincident boundaries.

### 18.4.1 Continuous variation in quantity and quality

The first prima facie problem for Voluminism concerns continuous variation. A quantity or quality varies continuously when each point in some region has a value that is different from all of its neighboring points. Since Voluminism denies that points are fundamental entities, in contrast to both Pointillism and Volume-Boundary Dualism, Voluminists must either deny that continuous variation is possible or else come up with some inventive account of how it is possible, despite the absence of fundamental points to act as bearers of the quantities or intensities.

A spatial region is uniform or *homogeneous* when every quality and quantity that is instantiated there takes a constant value at every subregion. A region is *continuously variable* with respect to some quality or quantity when it contains no homogeneous subregion. When a quantity or quality varies continuously within a region, the value of that quantity (like *density* or *temperature*) or quality (like *color*) at each point in the region is unique within some finite neighborhood around that point. If the quantity does not vary continuously throughout any subregion of a region *S*, then *S* is composed of a set of homogeneous subregions. Let's say that a region is *Aristotelian* if it is composed entirely by homogeneous subregions.

**Def. D18.1. Aristotelian Region.** A region *x* is *Aristotelian* if and only if there is a set *S* of homogeneous regions such that *x* is the sum of *S*.

**Def D18.2 Sum.** *x* is the *sum* of set *S* if and only if every member of *S* is a part of *x* and every part of *x* overlaps (has a part in common with) some member of *S*.

**18.5T Spatial Aristotelianism.** Every region of space is Aristotelian.

**18.5A Continuous Variation.** Some quality or quantity varies continuously throughout some region of space.

To simplify our discussion, let's suppose that the property that varies continuously throughout certain regions according to Continuous Variation is *density*. What then would be the fundamental properties? It would be natural to suppose that the fundamental properties are specific measures of actual densities. The bearers of these fundamental properties would have to be points in space or point-masses, since there are by hypothesis no subregions of constant density.

CONTINUOUS VARIATION AND VOLUME-BOUNDARY DUALISM For Volume-Boundary Dualists, the best solution to the problem of continuous variation is to adopt the idea of *distributional properties*, as proposed by Josh Parsons (2004). On this view, the property of *having a certain density at point P* is a fundamental property of some extended region or extended material object *E*, where *P* is a metaphysically dependent part of *E*. Points are real entities, but they cannot exist except as parts of extended things. On this view, it is hard to resist the conclusion that points are bearers of fundamental properties like *density*: the property of *an extended object E's having density d at point P* just consists in *P*'s having density *d* and in *P*'s being a part of *E*. This distributional properties option leaves us with a class of unexplained, brute necessities, however, which is a cost by Ockham's Razor (**PMeth 1.2**). Why can't points exist as independent entities?

The best response to this question would be to suppose that points have no independent spatial location. The *location* of a part *P* relative to other parts of the extended object *E* and to the whole of *E* is a fundamental property of *E* as a whole, not a property that *P* possesses on its own. In contrast, the *spatial location* of *E* is a simple quality of *E*, and its distance from other extended objects is a matter of the degree of similarity between the two intrinsic spatial qualities. Thus, the spatial location of an extended whole *E* and of a point-sized part *P* are radically different in character, explaining why *P* cannot exist on its own. Extended parts of *E* also derive their own location and internal organization from the internal organization of *E* itself, but such extended parts can potentially exist on their own, since they are the potential bearers of fundamental spatial qualities.

The metaphysical necessities involved with the distributional properties option are relatively few, and we can provide an explanation for them in terms of the essences or natures of extended objects and their parts. The only possible bearers of fundamental spatial locations are extended material objects. Internal points have only a derived location, based upon the location of the whole object and the point's role within the internal organization of the extended object, where that organization is a fundamental distributional property. Consequently, points are metaphysically dependent, since they are not the sort of thing capable of holding a spatial location on their own.

CONTINUOUS VARIATION AND VOLUMINISM Given Continuous Variation, Voluminists could suppose that the fundamental properties are properties of *average density*, borne by regions of space or extended material bodies. There would have to be necessary connections among these properties. The average density of matter in region *S* must be the volume-weighted mathematical average of the average densities of the subregions belonging to any *partition* of *S* (that is, to any set of mutually exclusive and jointly exhaustive subregions of *S*). Ockham's Razor (**PMeth 1.2**) enjoins us to minimize such brute metaphysical necessities, making such an account unattractive.

What are the alternatives? Voluminists seem to have only three further options:

1 Deny the possibility of continuous variation, insisting on Spatial Aristotelianism.
2 Claim that densities are properties of points, with points as logical constructions from regions.
3 Opt for fundamental properties of voluminous bodies that are simultaneously properties of both punctual (pointlike) spatial location and punctual density.

VOLUMINIST OPTION 1: SPATIAL ARISTOTELIANISM Voluminists' first option is simply to deny that continuous variation is possible. This is, scientifically speaking, still a live option, since we still do not understand how matter is organized at very small scales. However, it would be a serious defect of the view if it were inconsistent with the mere possibility of continuous variation, since such variation seems possible, and since it may be, for all we know, actual.

VOLUMINIST OPTION 2: DENSITIES AS PROPERTIES OF SET-THEORETIC OR LUDOVICIAN CONSTRUCTIONS OR OF PLURALITIES We can reject the set-theoretic version of option 2 on the basis of Membership the Only Fundamental Set-Relation (**PMeta 3**). Sets of regions cannot have fundamental physical properties like *density*.

The Ludovician version of option 2 is also problematic, since the very definition of continuity presupposes that points are dimensionless and volume-less. It doesn't make sense to talk about a distribution as *approximately* continuous. A distribution that is only approximately continuous is simply not continuous at all. Moreover, it will be simply false that the Ludovician points, which are in reality very small bodies, have precise density-values. The Ludovician could try to make use of average values here, but only at the cost of introducing a large number of brute necessary connections among the various average values. The various average values must satisfy the arithmetical law of averages, but the Ludovician has no metaphysical explanation of why this should invariably happen if each average value is a primitive fact.

However, a pluralist solution might be worth considering here. On this view, one treats *density* as a joint property of a plurality of nested regions (like Tarski's nested spheres). Pluralists can distinguish between distributive and collective predications of groups. For example, consider (1) and (2):

(1) The members of the glee club are male.
(2) The members of the glee club harmonize well.

The predication in (1) is distributive: each of the members of the glee club is a male. The predication in (2), in contrast, is collective. No member of the club harmonizes well all by himself. Proposition (2) is talking about how they harmonize with one another. Similarly, the pluralists could treat *density at a point* as a collective property possessed by those spheres that together make up the point, by virtue of being concentric. This has the advantage over option 1 of attributing *density* to something physical and concrete, namely, the spheres considered collectively, rather than to an abstract object, like a set.

Some spheres, the *x*'s, are *collectively pointlike* if and only if for all spheres *y* and *z*,

(1) if *y* and *z* are among the *x*'s, then the intersection of *y* and *z* is also among the *x*'s,
(2) if *y* intersects all of the *x*'s, then *y* is among the *x*'s, and
(3) if *y* is among the *x*'s, then *y* overlaps each of the *x*'s.

VOLUMINIST OPTION 3: FUNDAMENTAL SPATIAL-CUM-QUANTITATIVE PROPERTIES On option 3, an extended material object has a family of properties, each corresponding to an internal point. These properties are simultaneously spatial-location qualities and densities. Any two of these primitive punctual properties bears at least two different kinds of similarity to one another: similarity with respect to spatial location and similarity with respect to density. Two properties can be very similar in the spatial-location aspect and very different in the density aspect (representing two spatially close points with very different densities) or vice versa (representing two spatially distant points with similar densities). The space of colors provides an analogy. Two color-properties can be very similar with respect to hue but dissimilar with respect to brightness (like pink and maroon) or very similar with respect to brightness and very dissimilar with respect to hue (like pink and sky blue). Both *spatial distance* and *relative density* would then be internal relations between pairs of properties.

One advantage of option 3 is that it provides a uniform account of spatial distance: *spatial distance* is always an internal relation between spatial/quantitative properties.

However, option 3 is incompatible with both Classical and Non-Classical Truthmaker Theory (2.1T/2.1A.1T). In particular, it violates the One Truthmaker per Fundamental Property principle (**PTruth 1**), since a single entity, whether a trope or an ordinary particular, is supposed to serve as the truthmaker for the instantiation of determinates of both *spatial location* and *density*.

The main disadvantage of option 3 is that it provides no basis for thinking that point-sized material objects are impossible, and so no basis for denying that point-sized material objects are fundamental entities. If extended material objects each bear some infinite set of spatial-cum-quantitative properties, then why not suppose that there are infinitely many point-sized masses, each possessing exactly one of those properties? We can posit that, as a matter of metaphysical necessity, any material object must possess an infinite number of spatial-cum-quantitative properties, forming a continuous spatial region, but this necessity would seem to have to be a brute, unexplainable necessity.

Another objection to option 3 comes from the thought experiment of supercutting. If option 3 were true, it seems plausible that Hawthorne-Weatherson supercutting would be possible, resulting in a collection of isolated point-masses. On option 3, each of the resulting bodies would possess a single fundamental spatial-cum-quantitative property, so there is no simple explanation for the impossibility of the procedure. If the procedure is possible, it is arguable that the point-sized bodies were there all along, since there is no reason to think that merely cutting a body in half generates new entities, since there is no change in the fundamental properties that are instantiated. Thus, option 3 seems to be forced to give up the metaphysical priority of extended wholes, eliminating it as an option for making continuous variation compatible with that thesis.

### 18.4.2 Coincident boundaries: Actual vs. potential boundaries

If there could be several physical surfaces, edges, or points that coincide exactly in their spatial location, in their *coincident boundaries*, then this would pose a serious challenge to the combination of Voluminism with Spatial Monism (17.1T.1A.1A).

**18.1A.2A.1 Coincident Boundaries.** There are spatially coincident points (curves, surfaces).

Brentano (1976/1988) offered strong arguments for thinking that two distinct surfaces can co-exist in exactly the same place. Suppose the face of one cube is in direct contact with the face of another. Part of the surface of the one cube is then located in exactly the same place as part of the surface of the other, unless the two surfaces somehow fuse into a single entity. However, there seems to be little, if any reason, to believe that such fusion must take place. Each surface is dependent on a different entity, so why should something come into existence that is simultaneously dependent on both cubes? Why should mere contact necessitate the annihilation of the two distinct surfaces? If the two surfaces have natural properties, like *color*, it would seem that each could retain that distinct property, even when they coincide in space. It is hard to see why mere coincidence should necessitate that these properties be destroyed or altered.

Aristotelian philosophers have traditionally distinguished between actual and potential boundaries. A surface exists in actuality only when it is the outer boundary of some actual, extended thing. On this view, the surfaces that intersect the interior of extended things have merely potential existence. They could exist if the extended thing were actually divided along that surface. On this view, there can be actually coincident surfaces when two really distinct bodies are in contact, but there will be no actually coincident surfaces within the interior of an undivided body.

## 18.5 Conclusion

We have examined two major issues in the last two chapter. First, do places exist, and second, do material bodies (the things that have location) consist fundamentally of point-sized or voluminous parts? Our examination of the first issue was inconclusive, but we did find three theories that seemed to be the simplest: the Theory of Spatial Qualities (17.1T.1T), Spatial Monism, and Relationism (17.1A). We postponed our evaluation of Spatial (or Spatiotemporal) Monism until Chapter 24, where we will take up the question of how things persist through time, leaving us the Theory of Spatial Qualities and two versions of Relationism (Aristotelian 17.1A.1A and Modern 17.1A.1T). The gravest difficulty for both theories concerns the reality of empty space. How best to handle this problem turns on other metaphysical issues, like Actualism vs. Anti-Actualism (12.1A) and Realism (7.1T) vs. Trope Theory (8.2T). On the second major issue, we found a number of arguments that support Anti-Pointillism. The problem of continuous variation suggests that we take very seriously the mixed position of Volume-Boundary Dualism, despite the fact that it requires that we posit both points and regions as fundamental entities.

In the next chapter, we examine questions concerning the structure of time, including an issue that is similar to the debate between Pointillists and Anti-Pointillists, namely, the debate between Instantists and Intervalists.

## Note

1 A Ludovician reduction of regions to points would not be very plausible, since all points are equally small and equally lacking in shape or finite volume. We could, in some contexts, take a body to be some point located somewhere near the intuitive *center of gravity* of the body, but it would be hard to see what sense could be made of a body's having a certain shape or volume.

# 19

# The Structure of Time

In this chapter, we will examine three issues concerning the structure of time. First, are temporal intervals (extended periods of time) metaphysically fundamental entities or are they and their attributes wholly grounded in their constituent instants? Second, if there are fundamental intervals, are instants also fundamental or are they merely derived entities (such as sets of intervals)? Finally, is it metaphysically necessary that time have a beginning or is it possible that the past be infinitely long? We will take each of these up in successive sections.

## 19.1 Is Time Composed of Instants or Intervals?

In the case of time, we have a set of issues that closely parallel those concerning space. In the case of space, we considered whether it is spatial simples (points), extended spatial regions, or both that are metaphysically fundamental. There is a similar issue in the case of time. Does every finite temporal interval consist in a set of durationless instants in some set of mutual relations, or are temporal intervals metaphysically fundamental, with instants being fictions, logical constructions, or metaphysically dependent boundaries of intervals?

**19.1T Instantism.** Temporal intervals are not G-fundamental entities: intervals are wholly grounded in dimensionless instands.

**19.1A Intervalism.** There are extended temporal intervals that are G-fundamental.

G-fundamental entities are entities whose existence is not wholly grounded in the existence of other entities (as we defined in D3.5). In the rest of this chapter, when we use the term 'fundamenta', we will mean 'G-fundamental'.

*The Atlas of Reality: A Comprehensive Guide to Metaphysics*, First Edition.
Robert C. Koons and Timothy H. Pickavance.

Just as in the case of space, it is plausible to assume that there is some kind of metaphysical correspondence between time and the fillers or occupiers of time:

**Time-Process Correspondence.** The fundamental temporal entities are instants (as opposed to intervals) if and only if the fundamental occupiers of time are temporally unextended things.

What would it take for Time-Process Correspondence to be false? We would have to have either temporally extended processes occupying infinitely many fundamental instants without temporal parts or durationless parts of extended processes without any corresponding fundamental temporal location. The second option seems quite far-fetched. If a process consists of an infinity of durationless "time-slices", then surely there must be an infinite set of equally fundamental temporal locations—instants—for each of the instantaneous parts. Similarly, the first option is unattractive. If a process takes place through a series of real, fundamental temporal instants, then surely the process as a whole simply consists of a set of correspondingly instantaneous parts.

If Time-Process Correspondence is true, then Instantism is equivalent to Procedural Instantism, and Intervalism to Procedural Intervalism:

**19.2T Procedural Instantism.** No temporally extended process is fundamental: only their indivisible, dimensionless parts (time-slices) are fundamental.

**19.2A Procedural Intervalism.** Some extended processes are metaphysically fundamental.

Hereafter, we will use "Instantism" to refer to the combination of 19.1T and 19.2T and "Intervalism" to refer to the combination of 19.1A and 19.2A. Intervalism comes in several varieties, depending on whether we think that all intervals (and all extended occupiers of time) have proper temporal parts (parts that are finite but still shorter in duration).

**19.3T Temporal Finitism.** It is impossible for any temporal interval to have infinitely many actual temporal parts.

**19.3A Temporal Infinitism.** It is possible for temporal intervals to have infinitely many actual temporal parts.

Instantists must be Temporal Infinitists, since there are infinitely many instants contained in any interval of time. Intervalists can go either way. Alfred North Whitehead, for example, was both an Intervalist and a Temporal Infinitist. However, the clearest and strongest arguments for Intervalism are arguments for Temporal Finitism. Consequently, in the rest of the chapter, we will consider arguments for and against Temporal Finitism.

Temporal Discretism is the thesis that time consists of discrete time-atoms, each taking up some finite duration. Temporal Discretism is a kind of Finitism: any finitely extended interval is made up of only finitely many indivisible units of time.

**19.4A Temporal Discretism.** There are extended occupiers of time without proper temporally extended parts.

**19.4T Temporal Anti-Discretism (Infinite Divisibility).** All extended occupiers of time have proper temporally extended parts.

All Discretists are Finitists, but must all Finitists be Discretists as well? As we saw in the analogous case of space, there is some tension between Infinite Divisibility and Finitism. Anti-Discretists who are also Finitists must be Aristotelian Finitists: they must make a distinction between *actual* and *potential* temporal parts of an interval. There are intervals (and extended metaphysical processes) that are metaphysically fundamental, and these have proper temporal parts, but only in the sense that they are potentially divisible into sub-intervals. A given process might have begun earlier or ended sooner for instance, and it has potential temporal parts corresponding to these alternative scenarios, but the temporal parts are merely potential and are dependent upon the whole.

So, Aristotelian Temporal Finitists illustrate that it is possible to be both an Intervalist and to embrace Infinite Temporal Divisibility. What about those, like Whitehead (1919), who embrace both Intervalism and Temporal Infinitism? Such a position, Infinitary Intervalism, faces a serious problem in identifying what is metaphysically fundamental in the case of time. Each fundamental interval of time would have to be wholly composed of an infinite number of equally fundamental sub-intervals. This would generate a very high degree of metaphysical redundancy at the fundamental level. Each interval could be thought of as both fundamental in its own right and as a merely derivative entity, wholly grounded in the lengths and arrangements of some of its fundamental sub-intervals.

Such metaphysical redundancy would involve a large number of brute metaphysical necessities. For example, each interval $I$ would be wholly composed of a finite set of disjoint, contiguous sub-intervals $S$. (In fact, for each interval $I$, there would be an infinite number of such sets $S$.) It is necessarily the case that the temporal length (or duration) of $I$ is equal to the sum of the lengths of the members of $S$. Since both the length of $I$ and the length of each of the members of $S$ are, for the Infinitary Intervalists, fundamental truths, we have a necessary connection between equally fundamental, distinct sets of truths. Therefore, by the standard of the Second Corollary of Ockham's Razor (**PMeth 1.2**), Infinitary Intervalism is quite unattractive, relative to the three competing theories of Aristotelian Temporal Finitism, Temporal Discretism, and Instantism.

Consequently, we will assume that all Intervalists are either Discretists or Aristotelian Finitists, and that all Instantists are Temporal Infinitists. Thus, the issue of Finitism is crucial.

### 19.1.1 Argument for Instantism: Possible super-tasks

A *super-task* (as defined by James Thomson 1954) is a process consisting of infinitely many sub-tasks, completable in a finite period of time. The original super-task was proposed by the ancient philosopher Zeno of Elea in the paradox of Achilles and the tortoise. Achilles is much faster than the tortoise, one hundred times faster, we'll say, but the tortoise starts out with a head start of, say, 1000 meters. In order to catch up with the tortoise, Achilles must first cover the 1000 meters between his starting point and the tortoise's. By the time Achilles has done that, the tortoise will have moved 10 meters forward. To catch the tortoise, Achilles must first cover these 10 meters. By the time he has done that, the tortoise will have moved 0.1 meters ahead. Achilles can never catch the tortoise because whenever Achilles reaches the tortoise's starting point at the beginning of the most recent period, the tortoise will always have moved forward. To catch the tortoise, Achilles must complete an infinite number of tasks: running first 1000 meters, then 10 meters, then 0.1 meters, and so on *ad infinitum*.

There is a simpler version of Zeno's puzzle, that of Homer and the stadium. Suppose that Homer wants to walk across the stadium. To get to the other side, he must first walk halfway, then walk half the remaining distance, and so on *ad infinitum*. The single process of walking across the stadium can be described as consisting of an infinite number of smaller sub-tasks: walking half of the way, then a fourth of the way, then an eighth, and so on. If it takes one-half hour to walk halfway, one-quarter hour to walk a quarter of the way, and so on, Homer can complete these infinitely many tasks in just one hour.

According to Aristotle and Simplicius, Zeno offered the paradox as an argument against the reality of motion and in defense of Parmenides's theory that reality is unchanging. Zeno took it as obvious that such an infinitary task cannot be completed. However, we have a great deal of evidence, from sense perception and memory, for the reality of motion and, hence, of the possibility of such infinitary super-tasks (**PEpist 4**). Given Time-Process Correspondence, we have reason to believe that there exist finite intervals of time with infinitely many actual parts, one corresponding to each part of the infinitary sub-task. This provides a strong argument for the fundamentality of instants, since if it were intervals that were fundamental, we would seem to have in such infinitary processes a case of an infinite regress of metaphysical priority, with each interval being dependent on the many proper sub-intervals it is composed of, never reaching a level of absolutely fundamental instants.

However, Intervalists could respond by claiming, with Aristotle, that Homer's walk across the stadium is only potentially, and not actually, divisible into an infinite number of parts. It seems plausible to think that Homer's walk in fact consists of a finite number of metaphysically fundamental, temporally extended sub-processes, each of which could have occupied a shorter duration. On this picture, the intervals of time involved have potential, but not actual, parts.

Adolf Grünbaum (1967) argued that we can conceive of a super-task with infinitely many actual parts. Imagine that Homer walks across the stadium in the following *staccato* fashion. He walks halfway across, stops for a short period of time, and then walks a quarter of the way across, stops again for an even shorter period of time, and so on. Let's suppose that each period of rest is equal in length to the preceding period of motion. Homer could then complete the walk across the stadium in exactly two hours. In this case, the Aristotelian cannot argue that the walk consists in reality of only a finite number of parts. It is clear that Homer's staccato walk is divided actually into an infinity of smaller parts, each separated from the next by a period of rest.

There is a key difference between Homer's ordinary walk and his staccato walk. We have experience demonstrating that walks across stadiums are in fact possible. We do not have comparable evidence that staccato walks, of the kind described, ever happen in reality. We can, however, imagine such a walk, and we can, apparently, do so without contradicting ourselves or falling into any obvious absurdity. By Imagination as Guide to Possibility (**PEpist 1**), therefore, we have grounds for holding that Homer's staccato walk is possible, and so we have reason to believe that it is at least possible for an interval to have infinitely many actual parts. In such a case, the instants within the interval would be the metaphysically fundamental entities. It seems reasonable to assume that if instants (or points) are possibly fundamental, then they are necessarily fundamental, since metaphysical fundamentality would seem to be an essential feature of things.

### 19.1.2 Arguments for Intervalism: Impossible super-tasks

The main argument for Intervalism is that it follows from the metaphysical impossibility of certain infinitary scenarios involving super-tasks.

THOMSON'S SUPER-LAMP Thomson's (1954) original super-task involved a lamp with a power-switch having two settings: On and Off. Suppose that the lamp begins in the On position, is switched to Off after one-half hour, and then is switched to On after one-quarter hour, and so on. At the end of the hour, the lamp's switch has been moved from Off to On and back to Off infinitely many times. There is no last state of the switch. For any switch movement in the series, there is always a still later one within the hour period. The problem is this: what is the setting of the lamp after the hour is over? On or Off? Either outcome seems possible, without obviously contradicting any laws of motion.

So far, we have an unsolved problem, but no reason yet for thinking that the Thomson lamp involves any impossibility or absurdity. To get to a contradiction, we have to make some assumptions about the causal or explanatory structure of the situation.

First, we have to assume that the final setting of the lamp has some causal explanation, in terms of the lamp's prior history. It seems plausible to assume that any event or process that has a beginning in time must have a causal explanation.

**Def D19.1 Immediate Causation.** $x$ is an *immediate cause* of $y$ if and only if $x$ is a cause of $y$, and there is no $z$ such that $x$ is a cause of $z$ and $z$ is a cause of $y$.

**Causation of the Initiation of Processes.** Any event or process with a beginning in time has an immediate cause.

Next, we have to assume that the convoluted history of the lamp cannot provide a causal explanation of the final state of the lamp at the end of the hour. As we shall see in Chapter 28, causal explanations come in two varieties, namely, continuous and discrete. A continuous explanation provides a cause for an event in terms of a process that reaches the occurrence of the event through an infinity of prior states, smoothly or continuously varying over time. A discrete explanation refers to a single event as the cause.

**Causation: Continuous or Discrete.** If $x$ is a cause of $y$, then either (a) $x$ is a part of a simple (undivided) process including $y$ as a part, which begins before the beginning of *y*, or (b) *x* is a process or event that is separate from *y*, in the sense that *x* and *y* do not belong to any single, simple process.

**Partial Simultaneity of Discrete Causes.** If *x* is a cause of *y*, and *x* and *y* do not belong to any one simple (undivided) process, then *x*'s time of occurrence includes the time at which *y* begins.

Why must immediate continuous causes and their effects belong to a single simple or temporally undivided process? If the process isn't simple, if it contains many temporal parts that intervene between the cause and the effect, then the cause cannot be an immediate cause of the effect. This is because there can be no immediate action at a temporal distance: immediate causation cannot jump over a gap in time.

**Impossibility of Action at a Temporal Distance.** If *x* is an immediate cause of *y*, then either *x* and *y* overlap in time or *x* and *y* belong to a single, temporally undivided process (with no intervening temporal parts).

Given these three principles, we can conclude that the final setting of the lamp, whether Off or On, must have a cause, that the cause must be part of a simple process that persists until the end of the hour of switching, and that the cause contains no intervening temporal parts. However, the switching process clearly has infinitely many temporal parts, corresponding to the infinite number of distinct switch movements, from Off to On and from On to Off. Any of the earlier temporal parts is cut off from the final state of the lamp by infinitely many intervening temporal parts. Hence, there can be no *immediate* cause of the final state of the lamp, and so the lamp is metaphysically impossible.

Thus, we have a strong argument, based on causality, for thinking the Thomson lamp scenario is impossible. However, there is also a plausible argument for its possibility, based on Infinite Patchwork (**PMeta 5.2**). Appealing again to Imagination as a Guide to Possibility (**PEpist 1**) as well as to everyday experience with electric lamps, we know that it is possible for a lamp to be switched on and off again. It also seems possible that this switching process could be scaled downward to an arbitrary degree, making each switching event smaller and faster than its predecessor. If we assume that time is infinitely dense, that is, that there is an infinite series of smaller intervals that is completed before noon, then we can use patchwork principles to prove that the Thomson lamp scenario is possible, after all.

1 It is possible to switch the lamp on and off again.
2 The switching process described in premise (1) can be compressed without limit in time and space. That is, if it is possible to switch the lamp on and off again by moving the switch *x* millimeters in *y* seconds, then it is possible to switch the lamp on and off again by moving the switch *x*/2 millimeters in *y*/2 seconds.
3 There exists an infinite series of intervals, each $\frac{1}{2}$ the length of its predecessors, all of which are located before noon (with noon as the earliest moment not contained in any of the intervals).
4 By Infinite Patchwork (**PMeta 5.2**) and premises (1–3), it is possible for the lamp to be turned on and off again infinitely often before noon.

If we accept the causal argument for the impossibility of the Thomson lamp, then we must reject either one of the premises (1–3) or Infinite Patchwork. Premise (1) is known to be true, and the patchwork principles seem to be indispensable guides to possibility, so we shall focus on premises (2) and (3).

Temporal Finitism offers a simple solution to the paradox by denying premise (3). If time is discrete, with indivisible temporal "atoms", then premise (3) is clearly false. In addition, even if time is not discrete, if it is impossible for any finite interval of time to be actually divided into an infinite number of sub-intervals, then premise (3) will still be false, even if every finite interval is potentially divisible into smaller sub-intervals.

We might try to avoid Temporal Finitism by concentrating instead on premise (2). One worry about the Thomson lamp scenario is that it seems to require the switch to move faster and faster as the switching hour progresses. At some point, the switch will have to move faster than the speed of light. There would be no upper bound to the velocity of the switch, with the consequence that the path of the switch through spacetime is discontinuous at the final boundary of the hour. It is not implausible to suppose that the structure of spacetime guarantees (as the theory of relativity says that it does) that nothing can move faster than some fixed velocity (such as the velocity of light).

We can, however, re-describe the lamp scenario in order to avoid such super-luminal velocities. For example, imagine that the switch consists of a single particle. The lamp is On if the particle lies on a certain plane (or with its center on the plane, if the particle has finite volume), and it is Off if the particle is located off the plane (no matter how close it might be). Further, imagine that each successive switching Off involves the particle's moving a smaller distance from the On plane. Since the particle moves a shorter distance in each succeeding sub-interval, its velocity can remain constant throughout.

However, if we follow this version of the scenario, there is a definite answer as to the final state of the lamp's switch: it must end up on the plane, which means that the lamp will be On at the end of the switching hour. We can see that this must be the case by supposing that the particle ends up off the plane at the end. If it lies off the plane, then it must lie some finite distance from the plane, say *k*. However, the particle will stay closer to the plane than *k* for some final segment of the switching process. Supposing that the particle ends up *k* units away from the plane requires us to assume that the particle jumps instantaneously through space at the end of the process, which we could plausibly suppose to be impossible.

**PNatPhil 3 Continuity of Motion.** It is impossible for any material thing to move discontinuously through spacetime.

**Maximum Velocity.** There is some maximum velocity above which it is impossible for anything to move.

Continuity of Motion and Maximum Velocity offer an alternative explanation to the impossibility of the original Thomson lamp scenario, and they guarantee that the second scenario has a unique solution. However, we are still left with an argument from causation for the impossibility of the second scenario. We can conclude that the lamp must end up in the On position, but we cannot find a non-redundant cause of this fact that does not involve some action at a temporal distance. So, we still have some reason for preferring Temporal Finitism's solution to the puzzle.

FORREST'S SUPER-URN Peter Forrest (1999) has offered another paradox, similar in structure to Thomson's lamp scenario, that supports Temporal Finitism: the Super-Urn task. We are to imagine an infinite sequence of actions, taking shorter and shorter periods of time, and ending in a finite interval, say, an hour. We start with an urn containing a single particle. In the first half hour, we remove a particle from an urn and then move a particle into the urn. We do the same thing in the next quarter hour, and so on. After an hour has passed, we have completed infinitely many particle-replacements. The question is, at the end of the hour, is there a particle in the urn or is the urn empty?

The odd thing about the urn case is the fact that the answer to this question seems to depend on whether we move the same particle out of and then back into the urn or replace one particle with a new one in each sub-period. If we move the same particle out and in, then it seems clear that the particle must still be in the urn at the end of the period (assuming that the particle cannot jump discontinuously through space). Alternatively, if we move a new particle in during each period, and then out in the next period, then it will follow that the urn will be empty at the end, since all of the particles that were ever in the urn at any time will now be in some discard pile at the end of the hour. Therefore, it seems that the result of the super-task depends entirely on the identities of the particles moved around. However, it seems clear that the result of a physical process can only depend on the kinds of particles and material bodies involved, not in bare facts about the identities of those particles and bodies. Suppose that all of the particles involved in the task are perfectly indistinguishable from one another. Let's suppose that it is possible for many particles of this kind to occupy exactly the same place at the same time (as, for example, photons can). We could then describe two versions of the Super-Urn task that are qualitatively identical, differing only in the identity or distinctness of the particles involved. This, however, requires that mere differences in identity with no accompanying qualitative difference can make a difference to the result of a physical process. But this is implausible; Impotence of Identity seems true:

**Impotence of Identity.** Two qualitatively identical processes (processes indistinguishable except with respect to facts about the respective identity or distinctness of the particles involved) must have exactly the same result.

If we suppose Impotence of Identity, then we must conclude that the Super-Urn task is impossible.

Here's a more precise description of the Forrest super-task, using an infinite collection of qualitatively indistinguishable particles, any two of which can occupy exactly the same place at the same time, like photons or other bosons. The urn is now a plane in space. To begin with, the plane is occupied by particle 0. The other particles are arranged in space in the following way: particle 1 is one meter away, particle 2 is one-half meter away, and so on. In the first version of the super-task, particle 0 moves off the plane, coincides in position with particle 1 for a moment, and then returns to the plane. In the next stage, it moves off the plane and coincides with particle 2 for a moment, and then returns to the plane. At the end of this version of the task, particle 0 ends its journey on the plane.

In the second version, particle 0 moves off the plane and coincides in location with particle 1. Particle 0 stays there, and particle 1 moves to the plane and then off the plane, ending up in the starting-position of particle 2. Particle 2 then moves to and from the plane, and so on. In this version, every particle $n$ ends up in the starting-position of particle $n+1$, and in the end there is no particle on the plane. Yet, the two versions of the super-task are exactly the same except for the question of which of two indistinguishable particles moves away from their common position at each phase of the task. Intuitively, this kind of difference cannot make a substantive difference to the end-state, that is, whether or not there is any particle at all on the plane at the end of the process.

The most plausible explanation of the impossibility of the Super-Urn task is Temporal Finitism: it is impossible to carry out any super-task in a finite period of time, because no finite period of time can have infinitely many actual temporal parts.

PRUSS'S GRIM REAPER PARADOX Alexander Pruss (2009) has posed the following version of the Grim Reaper paradox (Benardete 1964, Hawthorne 2000) as an argument for the discrete character of time. We are to suppose that there are an infinite number of Grim Reaper mechanisms, each of which is engineered to do two things. First, each checks whether the victim, Fred, is still alive at the Grim Reaper's appointed time. Second, if he is still alive, each Reaper kills him instantaneously. The last Grim Reaper, Reaper 1, performs this dual task at exactly one minute after noon. The next-to-last Reaper, Reaper 2, is appointed to perform the task at exactly one-half minute after noon. In general, each Reaper number $n$ is assigned the moment $1/n$ minute after noon. There is no first Reaper: for each Reaper $n$, there are infinitely many Reapers who are assigned moments of time earlier than Reaper $n$'s appointment.

It is certain that Fred does not survive the ordeal. In order to survive, he must still be alive at one minute after 12 p.m., but we have stipulated that, if he survives until 12:01 p.m., then Reaper 1 will kill him. We can also prove that Fred will not survive until 12:01, since in order to do so, he must be alive at 30 seconds after 12, in which case Reaper 2 will have killed him. In the same way, we can prove that Fred cannot survive until $1/n$ minutes after 12, for every $n$. Thus, no Grim Reaper can have the opportunity to kill Fred. Thus, it is impossible that Fred survive, and also impossible that any Reaper kill him! However, it seems also to be impossible for Fred to die with certainty and yet without any cause.

If one worries that this paradox depends somehow on the vagueness of the *life/death* distinction, consider the following variant: the Grim Mover. In this case, we have a particle that is located exactly on a plane. At 12:01 p.m., Mover 1 will move the particle off the plane exactly one meter, if it hasn't already been moved. If it has already been moved then, he does nothing. Mover 2 is primed to perform the task of moving the particle one-half meter away from the plane at one-half minute after noon, if it hasn't already been moved. And so on. We can now prove that the particle is moved off the plane after noon, even though none of the Movers has moved it. Even worse, it must have been moved off the plane, even though there is no finite distance that it has been moved, since every particular distance corresponds to exactly one Mover!

The whole set-up must be metaphysically impossible. We can use Infinite Patchwork (**PMeta 5.2**) to turn this impossibility into a positive argument for Temporal Finitism (Koons 2014a). We can build a *reductio ad absurdum* of the hypothesis that any finite interval is divisible into an infinite number of sub-intervals. We will use an even simpler version of the paradox: the paradox of the Grim Signaler. Each Signaler has a unique natural number $n$, and each is capable of sending a signal representing that number $n$ to its successor. Each Signaler #$n$ is built in such a way that (i) if it receives an appropriate signal (of some number $m > n$) at its appointed time, then it simply transmits that signal to its successor, Signaler #($n+1$), and (ii) if it does not receive an appropriate signal at its appointed time, then it sends a signal representing the number $n$ to its successor. Here is the *reductio*:

1 Assume for contradiction that it is possible for there to exist a finite temporal interval that is divided into an infinite number of sub-intervals, with a last sub-interval but no first sub-interval.
2 It is possible to build a Grim Signaler with the disposition to respond to a signal from its predecessor and to send a signal to its successor in the specified manner.
3 The specification of each Signaler is intrinsic to its situation (that is, the passive and active powers that are attributed to it are intrinsic to it during its interval of activity).
4 Assume Infinite Patchwork (**PMeta 5.2**).
5 It is possible for there to exist an infinite series of Signalers, with a final Signaler but no first one.
6 At least one number $n$ is such that Signaler #$n$ initiated a signal representing $n$. If we assume that no Signaler sent such a signal, then no Signaler #$m$ with $m > 1$ would have done so. On this assumption, Signaler #1 would have sent a signal representing 1 to its successor, a contradiction.
7 If Signaler #$n$ sent a signal representing $n$, then there is no $m > n$ such that Signaler #$m$ sent a signal representing $m$.
8 So, there is no $m > n+1$ such that Signaler #($n+1$) sent a signal representing $n+1$.
9 So, Signaler #($n+1$) would have sent a signal representing $n+1$. But $n+1 > n$, contradicting (7).
10 Hence, there cannot be a beginningless infinite series of sub-intervals.

In fact, the Grim Signaler paradox suggests not only that no *finite* time period can be divided into infinitely many sub-periods but also that it is impossible that there should exist infinitely many time periods at all. It seems to provide grounds for thinking that time must be bounded at the beginning: there must be a first period of time.

Can we also show that time is bounded in the future, that there will be a last period of time? Apparently not. The only way to construct the Grim Signaler paradox in reverse would be to stipulate that each Signaler is able to check whether or not any future Signaler has initiated its characteristic signal. The apparent connections between time, knowledge, and action all seem to rule out the possibility of such a paradox, without requiring any limitations concerning the end of time.

One important proviso. The Grim Reaper/Signaler arguments depend on the assumption that the active and passive powers attributed to each Reaper or Signaler is intrinsic to that Reaper or Signaler during its assigned period of activity. This intrinsicality assumption is needed in order to apply the Infinite Patchwork (PMeta 6.2). That principle states only that an intrinsically described possible situation can be infinitely duplicated within a possible temporal structure. Situations that are not intrinsically described may be incapable of duplication. For example, consider a situation that is described in the following way:

(1) The First Reaper. Reaper *n* is the first Reaper to swing his scythe.

It seems possible that a situation of type (1) could occur: we can imagine a world in which Reaper *n* is the first one to swing his scythe. However, it would be fallacious to try to apply a patchwork principle to (1) in order to conclude that there could be an infinite series of Reapers each of which is the *first* one to swing his scythe. Why is this fallacious? Because to describe a Reaper as the first one to swing his scythe is not to describe the Reaper intrinsically. Instead, it is to describe the relation between Reaper *n* and a large number of extrinsic situations.

Thus, to apply Infinite Patchwork (**PMeta 5.2**) to the Grim Signaler story, we have to assume that the description we gave of the Signalers' dispositions to receive and send signals are intrinsic to each of them. This involves the Intrinsicality of Powers (**PMeta 2**), the claim that having a power is an intrinsic property of a thing. As we saw in Chapters 4 and 5, Powerists (4.4A.3) have good reason to accept this principle, while Neo-Humeists (4.4T) will reject it. For Neo-Humeists, whether one of the Grim Signalers has a certain power or not depends on the actual history of the whole world, including the actions and inactions of all of the other Grim Signalers. Neo-Humeists will conclude, not that an infinite regress of times is impossible, but merely that it is impossible for an infinite series of Signalers, arranged in the way described, to have the powers we ascribed to them. Neo-Humeists feel no pressure from Infinite Patchwork (**PMeta 5.2**) to think otherwise, since for them powers are not intrinsic.

## 19.2 Instants as Dependent Entities

In this section, we will assume for the sake of argument that Intervalism is true, that is, that some temporally extended intervals and processes are among the world's fundamental entities. Given that assumption, we can then ask if instants or instantaneous events are among the fundamental entities of the world, or if all fundamental entities are finite in duration. Strong Intervalism denies that instants are fundamental, while Moderate Intervalism, or Interval-Boundary Dualism, embraces the fundamentality of both instants and intervals.

**19.1A.1T Strong Intervalism.** Instants either don't exist at all or are derived entities—mere logical constructions from finite intervals.

**19.1A.1A Interval-Instant Dualism (Moderate Intervalism).** Both instants and intervals are fundamental entities.

The burden of proof is on Moderate Intervalists, since Strong Intervalism is a simpler, more economical theory.

Is the present, the "now", an instant or an interval? William James (1890) introduced into philosophy the psychological notion of the *specious present*: the present experienced as encompassing a short interval of time, some fraction of a second. Do we really have no experience of the present moment as an instant? What about the leading edge of the specious present? Do we have an experience of that edge as a dimensionless surface? Or, do we experience a present instant sweeping through the specious present? If we didn't, how would we be able to distinguish between earlier and later parts of that specious present?

Aristotle argued that we must suppose the present to be a dimensionless instant, since it is the boundary between the future and the past. It is impossible for a whole interval to be that boundary, since the earlier parts of such an interval would have to be already in the past. To resist Aristotle's argument, Strong Intervalists have to maintain that the specious-present interval has no temporal parts at all, no earlier or later. How, then, can it include motion and change? If a ball rolls down a ramp during the specious present, won't there have to be an earlier part of that interval during which it is on the top half of the ramp? This certainly seems to be the case. Denying the veridicality of such experiences is a theoretical cost to Strong Intervalists who appeal to the specious present.

A better tack for Strong Intervalists might be to locate the fundamental intervals at a very small scale, a matter of nanoseconds.

The principal advantage of Strong Intervalism is that it enables us to give answers to Zeno's paradox of the arrow and to the paradox of death. The paradox of the arrow points out that the arrow in flight does not move anywhere during each instant of flight. During each instant, it has exactly one precise location, and so is (in a sense) completely at "rest" "during" the instant itself. However, if the flight of the arrow merely consists in the sum of the arrow's state at each instant during the flight, then the arrow is motionless throughout its flight. How can the arrow move during this period if it is at rest during every instant of the period? The paradox of death starts with the simple question: is one alive or dead at the very instant *t* of one's death? If we say that one is dead at *t*, then it seems that *t* cannot be the instant of death, since one is already dead at *t*, and someone who is already dead cannot die. Alternatively, if we say that one is alive at *t*, then once again it seems that *t* cannot be the moment of death, since one is still alive at *t*, and so one has not yet died at *t*.

Strong Intervalists can respond to both paradoxes. The flight of the arrow is not made up of momentary (and stationary) instants: rather, each part of the arrow's flight occupies a finite interval of time, during which the arrow covers some distance. Similarly, there is no instant of death, and so no issue about whether one is then alive or dead. The "instant" of death for Strong Intervalists is merely a logical construction built up from finite intervals. For example, we could define the *instant* of death as the set containing all *final* intervals of the person's life, that is, as the set containing all intervals throughout which the person is alive, and which are not succeeded by a still later period of life.

However, there is a clear analogy between instants and spatial boundaries like surfaces, curves, and points. Just as it seems plausible that extended things have superficial (zero-, one-, and two-dimensional) boundaries, so it seems plausible that temporally extended events have instants as initial and final boundaries. Moderate Intervalists can embrace this view, Instants as Dependent Entities:

**19.1A.1A.1 Instants as Dependent Entities.** Instants exist only when they are the actual boundaries of extended processes or events.

What about Zeno's paradoxes? Can Moderate Intervalists dissolve them? Yes, because they do not claim that all intervals are ultimately composed of instants in a strong, metaphysical sense of composition, that is, one according to which the instants are supposed to be *actual and independent* parts. Instead, instants are dependent parts in that they are actual and potential boundaries of extended processes. Processes, like motion, do not derive their properties from their instantaneous parts. If anything, the dependency goes the other way around, from parts to wholes. Thus, Moderate Intervalists are not bothered by the fact that the arrow is not moving "within" any single instant. Instants aren't the fundamental units of motion; intervals are.

In the case of the paradox of death, the most plausible response for Moderate Intervalists may be to suppose that there can be more than one simultaneous instant, just as there can be more than one spatially coincident surface. In this case, there would be two, simultaneous instants of death. One is the final boundary of one's life, and the second is the initial boundary of one's death.

## 19.3 Does Time have a Beginning?

The Grim Reaper or Grim Signaler paradox that we considered in Section 19.1 gives us reason to believe that time must have a first period or interval. Does it follow then that time must have a beginning?

**19.5T Beginning of Time.** Time necessarily has a beginning.

The Beginning of Time could be understood in one of two ways:

**19.5T.1 Existence of a First Temporal Part.** Time necessarily has a single part that is earliest.

**19.5T.2 Metrical Finitude of the Past.** The past is necessarily finite in duration or measure.

The Grim Reaper paradox gives us a direct argument for the Existence of a First Temporal Part, but not for the Finitude of the Past. There is a gap between the two. Imagine a universe in which time begins with a single, undivided but infinitely long period of time. Let's call such a period an 'infinite simple past'. To get from the Existence of a First Temporal Part to the Metrical Finitude of the Past, we would need to assume that an infinite simple past is impossible:

**No Simple Infinite Past.** It is impossible for a temporal interval to extend infinitely far (in duration) into the past unless (i) it actually has infinitely many proper parts or (ii) it overlaps with an infinite number of disjoint intervals or events.

It seems reasonable to assume that for a simple region to have measurable temporal duration without parts, the simple region would have to either contain a process with a natural beginning and end or temporally overlap with one or more such processes. Thus, we can reasonably embrace the possibility of simple regions with *finite* duration, a duration corresponding to the natural distance between the two endpoints in processes of this kind. However, a simple region with an *infinite* duration in the past would have to contain only processes without a natural beginning, and we might well ask how any such process could have a temporal measure, without having actual proper parts or overlapping in time with other regions. Time is the measure of change, which seems to require both a *terminus a quo* and a *terminus ad quem*, a starting point and an end point for the change. This assumes, of course, that time has no intrinsic metric of its own.

Here's another argument for the principle No Simple Infinite Past. A simple region can have a temporal measure only if it is at least *potentially* divisible into parts. A region is divisible into temporal parts only if it contains one or more processes that can potentially be stopped or interrupted. A process $P$ is potentially stoppable only under certain conditions:

(i) $P$ itself has a natural, finite measure, based on the normal distance in time between its *terminus a quo* and *terminus ad quem*, a measure that can be shortened by accelerating $P$, or
(ii) There is another process $P_2$ that, when it reaches its terminus ad quem, has the power of terminating $P$, and $P_2$ is stoppable before the termination of $P$.

However, if the early history of the world consisted entirely of processes without finite measures, then none of those processes would be potentially stoppable, and hence none of the temporal intervals containing them would be even potentially divisible. Intervals that contain no temporal parts at all, whether actual or potential, and that temporally overlap only other regions without temporal parts cannot have a temporal measure. Hence, a simple past must be a quantitatively finite past.

Here's a third argument against a simple infinite past. Either there is an intrinsic metric to the pure passage of time or not. If there is, then the infinite past is actually divided, in and of itself, into an infinite number of actual periods. This contradicts the conclusion of the Grim Reaper paradox. If there is no intrinsic measure of time, then a period of time with no actual parts cannot have any measure unless it contains both the beginning and end of a process with an intrinsic duration.

Here is an apparently possible scenario that is more difficult for us to dismiss:

> Suppose that, before the Big Bang, there was a single, infinitely extended process of locomotion, by which two membranes were continuously moving close to each other. As we look farther into the past, we see the membranes farther and farther apart, with no limit. When the membranes collided, they produced the Big Bang, the beginning of complex time.

This scenario would be a problem for Aristotelian Finitists (although not for Temporal Discretists), since Aristotelian Finitists admit that locomotion can involve the covering of distance by means of a continuous, *undivided* process. However, they could question whether the imagined scenario is really possible, on the grounds that nothing in it licenses us to speak of measurable *spatial distances* before the collision. This is because distance, like temporal duration, derives its measure from finite processes, bounded at both their beginning and their end. Without real spatial distances, real locomotion would also be impossible. Once again, there is a dilemma. Either space is self-measuring or it is not. If it is self-measuring, then every region of space has an infinity of actual parts, and every process of locomotion through that region has a correspondingly dense infinity of temporal parts. Again, this contradicts the conclusion of the Grim Reaper paradox. If space is not self-measuring, then the imagined scenario is impossible.

## 19.4 Conclusion

If temporally extended intervals and processes are fundamental, this fact will have a profound effect on our understanding of change and motion, questions that we will take up again in Chapter 24. It may also provide insight into the problem of persistence or identity through time, as we will see in Chapter 24 as well. There is a natural affinity between Intervalism about time and the fundamentality of persisting things, including people and other organisms. In contrast, if Instantism is true, then we have good reason to adopt the At/At Theory (24.5A.1T) of motion and change, reducing facts about motion and change to facts about what properties or locations things have at one instant or another. This leads naturally to the view that all fundamental entities are instantaneous or durationless, and so that all persisting things are mere logical constructions. That, in turn, has profound ethical and legal consequences when applied to persons.

# 20

# Time's Passage

Time is, of course, an important part of life. We are always experiencing time and its passage. As the saying goes, change itself is the only constant. So time is the sort of thing that seems unproblematic, at least until one begins to think about it (as St. Augustine of Hippo noted in Book 11 of his *Confessions*). Some philosophers have questioned whether there is such a thing as time at all:

**20.1T Temporalism.** There are moments of time.

**20.1A Atemporalism.** There are no moments of time.

J.M.E. McTaggart is the most famous proponent of Atemporalism. In addition, some of the historic proponents of Monism (11.2A), specifically, Parmenides and F.H. Bradley, took their Monism to entail the unreality of time. Since we've already discussed Monism, we will focus here on McTaggart's challenge to Temporalism (McTaggart 1908).

McTaggart distinguished three "series" of moments: the A, B, and C series. The A series compares moments in relation to the present, dividing time into past, present, and future. The B series consists of relations of *earlier-than* and *later-than*, and the C series is limited to the relation of *betweenness*, without any temporal direction.

Let's suppose that the first manned mission to Mars occurs at some point in the late twenty-first century. Compare these three propositions:

(1) The Civil War occurred over one hundred years in the past, while the first manned mission to Mars is still in the future.
(2) The Civil War occurred at least two hundred years earlier than the first manned mission to Mars.
(3) The election of Barack Obama occurred between the Civil War and the first manned mission to Mars.

*The Atlas of Reality: A Comprehensive Guide to Metaphysics*, First Edition.
Robert C. Koons and Timothy H. Pickavance.

(1) locates the Civil War and the Mars mission in the A series, relating them to the present moment. (2) concerns these events' locations in the B series. (1) was false in 1850 and will be false again in 2100. (2), in contrast, will remain true forever and arguably has been true throughout all the past. Positions in the B series never change: they are eternal. (3) relates three events in the C series. Like (2), (3) is eternally true. The difference is that (3) implies nothing about the direction of time. (3) could be true, even if we thought that our experience of *before* and *after* were a mere illusion.

McTaggart's argument for the unreality of time proceeds in several steps:

1 Time is real only if the B series is real.
2 The B series is real only if change is real.
3 Change is real only if the A series is real.
4 The A series is contradictory and hence unreal.
5 Therefore, time is unreal.

McTaggart argues that the reality of the C series alone is not sufficient to make time real. If time is unreal, then there are no moments of time. Relations of *before* and *after* are essential to time. If these relations are unreal, then what we call "time" is really a dimension of reality without any intrinsic direction, like space.

In step 2, McTaggart connects the reality of temporal direction with the reality of change. If nothing changed, then there would be no sense to talking about the passage of time, and no way to distinguish earlier from later.

Next, McTaggart claims that the reality of change depends on the reality of the A series. He rejects Bertrand Russell's notion that real change is possible so long as objects have different properties at different times. Consider an iron poker that "changes" from being cold at one time to being hot at a later time. McTaggart argued that without the A series this case is no different from a case of an iron poker that is cold on one end and hot on the other. In neither case is there real change, unless the events involved are themselves changing in their position in time. If only the B and C series were real, then events would never change, since an event's B and C relations to other events are all unchanging. The only way that an event can change is to pass from being future to being present or from being present to being past. These changes involve the A series.

Finally, in step 4, McTaggart argues that we cannot coherently conceive of the A series. His argument for this has become known as "McTaggart's paradox". We will take up McTaggart's paradox in Section 21.5.

There is another sort of challenge to the reality of time to consider. McTaggart accepts that there are many moments but argues that these are not moments *of time* because they cannot really stand in a temporal order. We could instead consider a challenge to the existence of more than one moment of time. This would be like George Santayana's Solipsism of the Present Moment (13.1T.1), discussed in Chapter 13.

**20.1T.1T Temporal Pluralism.** There are, have been, and will be more moments of time than one.

**20.1T.1A Solipsism of the Present Moment.** There are, have been, and will be only one moment of time (the present).

The argument for Solipsism of the Present Moment involves an appeal to Ockham's Razor (**PMeth 1**): don't multiply entities needlessly. If we suppose all of our beliefs about the past and future to be simply false, then we could make do with only the present moment.

However, there are powerful arguments for the existence of past and future moments.

1 Memory. We have apparent acquaintance with past moments. And given the Reliable Perception Presumption (**PEpist 4**), this gives us good reason to believe that those moments are real.
2 The specious present. Sensory perception and introspection present us with an apparently moving and changing world, entailing the existence of multiple times.
3 Intelligibility of language and of complex thoughts. If Solipsism of the Present Moment were true, then we would have good grounds for doubting that we really understand any language, since language learning takes time. If we do not understand the sentences of what we take to be our native language, then we do not understand most of our complex thoughts, which take linguistic form. This includes all of our philosophical thoughts. Thus, if Solipsism of the Present Moment were true, we could not understand or believe it, or even consider it as a possibility. Therefore, by engaging in philosophical discourse, we tacitly presuppose the falsity of Solipsism of the Present Moment. More generally, Solipsism of the Present Moment runs afoul the following principle of epistemology:

**PEpist 2.2.1 Presumption of Philosophical Discourse.** It is prima facie plausible to presume the truth of all of the pragmatic presuppositions of philosophical discourse.

## 20.1 Tensers and Anti-Tensers

As we have seen, McTaggart introduced into philosophy the distinction between the A and B series of moments. Later philosophers of time have extended McTaggart's labels to theories of time. In particular, many philosophers talk about the dispute between the 'A-Theory' and 'B-Theory' of time. We say this is the debate between 'Tensers' (A-Theorists) and 'Anti-Tensers' (B-Theorists).

**20.2T Tensism (A-Theory).** Some tensed truths are metaphysically fundamental.

**20.2A Anti-Tensism (B-Theory).** There are no metaphysically fundamental tensed truths.

A *tensed* truth is a proposition whose content contains a tense (past, present or future) or that predicates a tensed property. Tensed properties include the property of *being past*, *being present*, and *being future*. By virtue of containing some reference to the *pastness*, *presentness*, or *futurity* of some time, tensed propositions can be true at some times and false at others. They do not have fixed truth-values.

**Def D20.1 Tensed Proposition.** A *tensed proposition* is a proposition that can have different truth-values (true or false) at different times.

Tensers maintain that some true tensed propositions are fundamentally true: that is, that their truth is not conceptually grounded in the truth of untensed propositions. Anti-Tensers think that no tensed proposition is fundamental in this sense.

Anti-Tensers come in two varieties: Eliminative and Reductive:

**20.2A.1T Eliminative Anti-Tensism.** There are no tensed truths.

**20.2A.1A Reductive Anti-Tensism.** There are tensed truths, but all tensed truths are wholly grounded in tensed truths.

Eliminative Anti-Tensers either deny that there are any tensed propositions at all, though they may acknowledge that there are tensed *statements*, or insist that all tensed propositions are false, thereby committing to an error theory of our common sense judgments about tensed propositions. Reductive Anti-Tensers concede that there are true tensed propositions but insist that they are all made true by tenseless propositions. Either way, there are no fundamental tensed truths.

McTaggart argues that Eliminative Anti-Tensism is false, since the lack of tensed truths entails Atemporalism. For McTaggart, it is obviously true that if event 1 is earlier than event 2, then it was once, is now or will be true that event 1 is present and event 2 future, and it was once, is now or will be true that event 2 is present and event 1 past. Hence, if there are never any tensed truths, there can be no temporal order whatsoever.

The other apparent drawback to Eliminative Anti-Tensism is that it forces us to reject as false many sentences that we are all naturally inclined to accept as true. We can appeal to common sense and the viewpoint embedded in our ordinary linguistic practices to support the claim that many tensed sentences or sentences predicating tensed properties (such as *pastness*, *presentness*, and *futurity*) are sometimes true.

Eliminative Anti-Tensers have a response to this objection, however. They can suppose that each sentence that seems to express a tensed truth does not really do so. Instead, it expresses a tenseless truth, but a different one at different times. In other words, Eliminative Anti-Tensers can suppose that one and the same sentence-type can express different propositions at different times. Consider the following sentence-type:

(4) It is now 9 a.m. 10 July 2010.

This sentence seems to express the tensed truth of the presence of a particular date and time. However, Eliminative Anti-Tensers could suppose instead that utterances of (4) express different untensed propositions at different times. When (4) is expressed at time *t*, then it predicates the property of *being identical to t* to the time and date (of 9 a.m. 10 July 2010). For each time *t*, the property that is predicated of 9 a.m. 10 July 2010 is a B-property.

Reductive Anti-Tensers must also account for the truth of statement (4). If we can ground the truth of such statements in truths about the B series, it will be easy to ground the truth of propositions about what is past and future, since the past is obviously earlier than the present, and the future later.

There is a strong argument for thinking that any such reduction of tensed to untensed truths must fail, since untensed truths are true permanently, while tensed truths are (by definition) true only temporarily. How can what is true temporarily be derived from what is true permanently? If a truth $p$ is grounded in another truth $q$, we have assumed that the truth of $q$ must entail the truth of $p$. If $q$ is permanently true, and the truth of $p$ is grounded in the truth of $q$, then $p$ must be permanently true as well. So, it is impossible to ground temporary truths in permanent truths, and Reductive Anti-Tensism must be false.

In order to give Reductive Anti-Tensism a chance, we must relax somewhat our standards for what counts as grounding one truth in another (in the sense of the logical or conceptual grounding discussed in Section 3.4). Reductive Anti-Tensers must argue that predications of tensed properties are not true absolutely or *simpliciter*. Instead, they are true or false *relative to this or that time*. Reductive Anti-Tensers can then argue that the truth of some tensed proposition $p$ at time $t$ is grounded in the permanent truth of some untensed proposition $q$, and the falsity of $p$ at some other time $t_2$ is grounded in the permanent falsity of some untensed proposition $q_2$. Consider, for example, the relation of (4) to (5) and (6):

(5) 9 a.m. 10 July 2010 is the same time as 9 a.m. 10 July 2010.
(6) 9:01 a.m. 10 July 2010 is the same time as 9 a.m. 10 July 2010.

(4) is true at 9 a.m. and false at 9:01 a.m. We could say that its truth at 9 a.m. is grounded in the trivial truth of the identity (5), and that its falsity at 9:01 a.m. is grounded in the trivial falsity of the identity (6).

We can clarify Reductive Anti-Tensism further if we assume Classical Truthmaker Theory (2.1T). Reductive Anti-Tensers propose that the only truthmakers for tensed truths are facts about B-relations, like fixed *earlier-than* and *later-than* relations among events. What is it that makes (4) true at 9 a.m.? Nothing except the truthmaker of (5), that is, nothing but the self-identity of the relevant moment of time. What makes (4) false at 9 a.m.? Nothing but the absence of a truthmaker for proposition (6), that is, nothing but the distinction between the two times 9 a.m. and 9:01 a.m. Consequently, one way of bringing out the difference between Reductive Anti-Tensers and Tensers is this: for Tensers, statements like (4) are true *simpliciter*, while for Reductive Anti-Tensers, they are true only at some time or another.

Here's another way of capturing the difference between Tensers and Reductive Anti-Tensers. Tensers believe that there is a unique present that is constantly in motion. There is always a unique point in time that is present, but different moments have been present in the past and will be present in the future. For Tensers, this movement of the property of *presentness* is one of the fundamental facts of the world.

In contrast, Anti-Tensers deny the existence of any such pure flow of time. Of course, they do not deny the obvious truth that there is a unique present moment, nor the obvious truth that what time counts as present is different at different times. Anti-Tensers will give an account of the truth of the propositions expressing the flow of time that grounds their truth in eternal truths, making use of a kind of *conceptual grounding* (in the sense described in Section 3.4). Consequently, Anti-Tensers can consistently deny that the flow of time is *real*. Consider (7) and (8):

(7) There is exactly one moment that is present, but in the past other moments were uniquely present.
(8) There is exactly one moment that is identical to 9 a.m. 10 July 2010, but there are earlier times, each of which is distinct from 9 a.m. 10 July 2010, and each of which is uniquely identical to itself.

Tensers believe that (7) is true, and that the truth of (7) is in no sense grounded in the truth of (8). The *presentness* of the present moment has nothing to do with its being identical to itself, and the fact that earlier moments were once uniquely present is not grounded merely in the fact that those moments stand in the *earlier-than* relation to the moment that happens to be present. In fact, Tensers might argue that the truth of (8) is grounded in that of (7): what makes certain times earlier than the present is the fact that they once were but are no longer uniquely present.

For Anti-Tensers, in contrast, the priority between (7) and (8) runs the other way. When (7) was expressed at 9a.m. on 10 July 2010, the truth it expressed was exactly the truth expressed by (8) and no other. (7) was true then because (8) was then, and still is, true.

We can see this difference in explanatory direction in other cases. Consider the following pairs of statements:

(9a) The Civil War ended 145 years ago.
(9b) The Civil War ended in 1865.
(10a) The Civil War has already ended, and human beings will but have not yet colonized Mars.
(10b) The Civil War ended before human beings colonized Mars.

Assuming that the Civil War did indeed end in 1865, that human beings will colonize Mars in some year after 2010, and that all four statements were made in 2010, all four statements were true at their time of utterance. Tensers take (9a) and (10a) to be fundamental, with the truth of (9b) and (10b) to be explained in terms of them. Anti-Tensers take (9b) and (10b) to be fundamental, with (9a) and (10a) being merely ways of expressing the very same truths in a way that exploits facts about when the statements are being made.

## 20.2 Varieties of Anti-Tensism

### 20.2.1 Reflexive thoughts and Old Anti-Tensism

One attempt to reduce tensed to untensed facts, what we'll call "Old" Anti-Tensism, makes use of *reflexive thoughts*. A reflexive thought is a thought that is partly about itself.

Consider (11a), (11b), and (11r):

(11a) I ate breakfast three hours ago.
(11b) My eating of breakfast occurred at 9 a.m. on 5 July 2010.
(11r) My eating of breakfast occurred three hours earlier than the occurrence of this very thought.

(11a) is a tensed truth, and (11b) is a simple untensed truth. But (11r) is reflexive. (11a) and (11b) are clearly not equivalent: (11a) changes its truth-value from false to true and back to false, while (11b) remains true forever. Old Anti-Tensism proposes that (11a) is really equivalent to (11r), that (11a) is covertly a self-reflexive thought. However, (11a) is not even covertly about itself. It is not about thoughts or the times of their occurrence. Moreover, (11r) is an untensed truth. Its truth-value never changes. Hence, it cannot be logically equivalent to (11a). For these sorts of reasons, Old Anti-Tensism has been almost universally rejected.

### 20.2.2 Indexical propositions and New Anti-Tensism

New Anti-Tensism eliminates tensed propositions altogether and makes use of the theory of *indexical* sentences as a way of understanding the function of the tenses. An *indexical sentence* is one that contains an element, like 'I' or 'here', that indicates some fact about the context of use of a statement of the sentence. (The reader may want to remind themselves of Def 14.2, and of the discussion of Modal Indexicalism 14.2T.) The word 'I' picks out the speaker of the statement, 'here' the location of that speaker at the time of the utterance, and 'now' and present tense verbs pick out the time of the utterance. Similarly, the past tense picks out the range of times earlier than the utterance, and the future tense those times that are later. Compare (11a), (11b), and (11r) with (12a), (12b), and (12r):

(12a) I am warm.
(12b) RCK is warm.
(12r) The thinker of this very thought is warm.

The truth of (12a) (even when thought by RCK) is not a priori equivalent to the truth of (12b), nor to the truth of reflexive thought (12r). Nonetheless, it is plausible to think that whatever it is in the world that makes (12a) true, the truthmaker of (12a), when (12a) is entertained by RCK is the very same fact which makes (12b) true. Thus, whether we can reduce the truth of statements like (12a) to the truth of those like (12b) depends on what we are willing to count as a reduction.

In order to claim that the truth of (12a), when thought by RCK, reduces to that of (12b) or that the truth of (11a), when uttered on noon of the relevant day, reduces to that of (11b), we must embrace some version of Truthmaker Theory. Then, whenever something like (11a) or (12a) is true, the truth of these indexical statements is grounded in the truth of the propositions expressed fully and clearly by (11b) and (12b). That is, these truths are made true by non-indexical facts. This position has become more popular in light of the problems for Old Anti-Tensism noted above.

## 20.3 Varieties of Tensism

There are four ways of fleshing out the idea that *being present* is a property that times have temporarily. First, we could hypothesize that there is a simple, primitive, unanalyzable property of *presentness* that is possessed by one and only one moment at a time. This is often called "the Moving Spotlight Theory" because we can imagine the present as something like a spot of light that shines on a single moment and is constantly moving in the direction of the future. (This image was first employed by C.D. Broad 1923/1959.) The special "light" consists in the present moment's possessing some unique, irreducible quality of *presentness*.

**20.2T.1 Simple Tensism (Moving Spotlight Theory).** There is a simple, primitive property of *presentness* that is uniquely and temporarily possessed by a moment of time.

We have to add that this same property has been possessed by past moments and will be possessed by future ones. If we could analyze tensed propositions into tenseless ones, this claim would collapse into the triviality that each moment is (tenselessly) present when it is present. Thus, Simple Tensers have to embrace Tensism, refusing to reduce tensed propositions to tenseless ones.

Second, we have Falling Branches Tensism, proposed by Storrs McCall (1976, 1994).

**20.2T.2 Falling Branches (Changing Possibility) Tensism.** The present moment is the earliest moment such that all later times involve at least one contingent fact.

Falling Branches Tensism has a critical focus on a particular feature of modality, namely, the changing status of possibility or contingency. The future is characterized by contingency. Every future moment is such that there are some still contingent propositions concerning that moment, and these future contingent propositions represent alternative possible scenarios. The present and past, in contrast, are closed and fixed. The present moment is the latest of the closed moments and the earliest of the moments that are succeeded entirely by open moments. In other words, the present moment is the earliest branch point. As time progresses, possible branches, which represent alternative scenarios, "fall off" the tree of time. They cease to be possible once their moment of potential actualization has passed.

Third, there is another variety of *Modal* Tensism, Tensism that appeals to a modal notion. Aristotelian or Changing Actuality Tensism is the claim that there is one moment that is uniquely *actual* or real:

**20.2T.3 Aristotelian (Changing Actuality) Tensism.** The present moment is uniquely actual.

Changing Actuality Tensism itself comes in a variety of forms, depending upon the account of actuality that we embrace (see Chapters 14 and 15). If, for example, we adopt Existence-Simpliciter-Defined Anti-Indexicalism (15.2T.5), then Aristotelian Tensism is equivalent to Presentism, the fourth form of Tensism, introduced below.

The most interesting version of Aristotelian Tensism is one that builds upon Hybrid Anti-Actualism:

**Hybrid Anti-Actualism** Some things don't exist, but all concrete states of affairs (or all nexuses between properties and substances or all bundles of properties) exist.

On the combination of Aristotelian Tensism and Hybrid Anti-Actualism, all concrete states of affairs (nexuses, bundles) exist actually, and so all states of affairs (nexuses, bundles) exist in the present moment. This is because, according to Aristotelian Tensism, only the present moment is actual. However, there are many things, such as ordinary particulars, events, tropes, or bare particulars, that do not actually exist and that *can* therefore be located entirely in the past or future. These merely-past or merely-future things can be parts of the truthmakers for past-tensed and future-tensed truths.

Fourth, there is Existential Tensism, usually called 'Presentism'.

**20.2T.4 Presentism (Existential Tensism).** The present moment is the unique moment such that everything exists then. That is, absolutely everything exists in the present moment and only then.

According to Presentism, the uniqueness of the present moment consists in the fact that everything exists in the present: that is, nothing exists wholly in the past or the future.

Before we consider Presentism more thoroughly, it is worth contrasting it with the view of what exists that is most naturally combined with Anti-Tensism, namely Eternalism:

**20.2A.2T Eternalism.** Everything that exists in the past, present, or future exists simpliciter.

Eternalism is the claim that past, present, and future moments are all metaphysically on a par, at least with respect to existence. The past and future exist no less than the present does. We can gain clarity about Eternalism by considering the indexical Anti-Tenser strategy discussed above. On that view, it is very natural to think that the present moment is much like one's present location in space: it's just one among many moments of time, and not metaphysically special in any deep way. To say that something is "here" is simply to locate it where you are. Likewise, it's natural for Anti-Tensers to think that to say that something exists "now" is simply to locate it at a particular moment of time. That something isn't *here* is no bar against its existence. Likewise, that something is past or future is no bar against its existence. Dinosaurs and Martian colonies may not exist *now*, but they nonetheless *exist*. Simple Tensism is also helpful in understanding Eternalism: the spotlight moves along the four-dimensional, Eternalist universe.

(Strictly speaking, Anti-Tensism does not entail Eternalism. However, non-Eternalist Anti-Tensism is, for the reasons outlined here, a strange hybrid. In what follows, we treat Anti-Tensers as committed to Eternalism.)

## 20.4 Presentism

We have to distinguish Presentism from the following triviality: everything that exists in the present exists in the present. (13) is obviously trivial:

(13) Everything that exists in the present exists then.

(13) simply says that the present moment is the present moment. Everyone, even Anti-Tensers, agree with that. The simplest way to distinguish Presentism from the triviality of (13) is to adopt Actualism (12.1T), the thesis that everything exists. Anti-Actualists (12.1A), whether Possibilist (12.1A.1T) or Meinongian (12.1A.1A), cannot be Presentists (in the sense we have articulated), since they insist that some things do not exist at all, whether in the past, present, or future. Anti-Actualists would need to be Possibilistic Presentists:

**20.2T.4.1 Possibilistic Presentism.** The present moment is the unique moment at which everything that actually exists (simpliciter) actually exists *then*.

There is one further issue of interpretation with respect to these two varieties of Presentism: how are we to understand the verb 'exists'? Obviously, we cannot understand it as a simple, present-tense use of 'exists', or else we would again produce the trivial (13) or its Possibilist counterpart (14), which is no less trivial:

(14) Everything that actually exists in the present actually exists then.

So to make sense of Presentism, we must understand the verb 'exists' tenselessly. Consider, for example, (15) and (16):

(15) Two plus two is four.
(16) The date of the beginning of the Civil War is earlier than the date of the beginning of World War I.

It's plausible to think that the verb 'is' in (15) and (16) is tenseless. Two plus two didn't become four at some point in time nor could they ever cease to add up to four. Similarly, 1865 is eternally earlier than 1914. Such facts couldn't possibly change with the flow of time.

Nonetheless, some Tensers (like Arthur Prior 1957, 1967, 1968 and Richard Gale 1968) have denied that there is such a thing as tenseless verbs. These are Serious Tensers:

**20.2T.5 Serious Tensism.** Only tensed truths are metaphysically fundamental.

For Serious Tensers, the right way to understand the verb 'is' in (15) and (16) is as an abbreviation for the conjunction 'always was, is, and always will be'. The past, present, and future tenses are fundamental and cannot be eliminated from a perspicuous statement of the facts.

For Presentists who are Serious Tensers, the verb 'exists' as it occurs in the statement of Presentism or Possibilistic Presentism is either a present-tense version of 'exists' or the disjunction 'has existed, exists, or will exist'. The first interpretation results in the trivial (17) and the second in the apparently absurd (18).

(17) Everything that exists in the present exists in the present.
(18) Everything that has existed, exists, or will exist exists in the present.

Actualists who are Presentists can, however, embrace (18), since they insist that absolutely everything exists. Consequently, Actualistic Presentists can (surprisingly) affirm (19) and (20):

(19) *Everything* that has existed still exists.
(20) *Everything* that will exist exists already.

Actualistic Presentists deny that some dinosaurs once existed. And they hold that absolutely nothing was once a dinosaur. But they can affirm that there *once were* some dinosaurs. In other words, Actualistic Presentists accept (21) and deny (22):

(21) In the past: (There exists an $x$: $x$ is a dinosaur).
(22) There exists an $x$: (In the past: $x$ is a dinosaur).

Since the existential quantifier ('There exists an $x$') occurs only within the scope of the past-tense operator ('in the past') in (21), Actualistic Presentists can assert (21) without accepting that there exist any dinosaurs. The quantifier cannot be exported from within the past-tense operator. Consequently, for Actualistic Presentists, it is a trivial truth that everything that exists in the past, present, or future exists in the present, since *there is* nothing that exists merely in the past or merely in the future. For Actualistic Presentists to assert (18), (19), or (20) is no more shocking than for them to assert something like (23) or (23*):

(23) Everything that exists in actuality or is a unicorn exists in actuality.
(23*) Everything that exists in the present or is a dinosaur or a Martian colony exists in the present.

For Actualists, (23) is trivially or vacuously true, since there are no unicorns. For Actualistic Presentists, (23*) is also vacuously true, since there aren't any dinosaurs or Martian colonies. In modern logic (that is, since Frege), any statement of the form 'All $A$'s are $B$'s' is *vacuously* true whenever there are no $A$'s, since to say that all $A$'s are $B$'s is simply to say that no $A$ is not a $B$. If there are no $A$'s, then it's impossible for any $A$ to fail to be a $B$. In such cases, logicians say that the statement 'All $A$'s are $B$'s' is vacuously true.

**Def D20.2 Vacuous Generalization.** A generalization of the form 'All $A$'s are $B$'s' is *vacuously true* if and only if there are no $A$'s.

Therefore, for Actualists, (24) is vacuously true:

(24) Every unicorn exists in actuality (that is, there are no unicorns that do not exist in actuality).

There are no unicorns that do not exist in actuality because there simply are no unicorns *period*. For the same reason, it is true that all unicorns are three-legged or ugly or have round square horns. These are all vacuously true for Actualists, since there are no unicorns at all.

In contrast, Possibilists and Meinongians reject (23) and (24). For similar reasons, they must reject (18), since, for them (18) would entail that dinosaurs still exist and Martian colonies exist already.

We will assume, then, that Presentists who are Serious Tensers are also Actualists. In contrast, Simple Tensers and Modal Tensers, whether Falling Branches or Aristotelian, in contrast, could be Possibilists or even Meinongians. So, too, could Presentists who are not Serious Tensers, since such Presentists can insist that (14) be interpreted as using a tenseless 'to be' in the subject term:

(14*) Everything that actually exists (tenselessly) actually exists in the present.

There is a variation on Presentism that has been proposed by C.D. Broad (1923) and others: Growing Block Tensism. On this variation, we are to suppose that the present is the latest moment at which anything that exists exists. That is, the Growing Block Tensists suppose that everything exists either in the past or the present. Nothing exists solely in the future. The present moment is, therefore, the "leading edge of reality". The future is wholly unreal, but once something becomes real, it stays real forever after.

One last point about Presentists. It is important to recognize the difference between Presentists and Solipsists of the Present Moment. Both agree that only the present moment exists. However, Solipsists also deny that there ever were or ever will be any other moments. In fact, they deny all tensed propositions: nothing was or will be. In contrast, Presentists are Tensers. They not only affirm past-tensed and future-tensed propositions; they also believe that some of these are among the world's *fundamental* truths.

THE PRESENTISTS' SOLUTION TO THE PARADOX OF INTRINSIC CHANGE Put simply, the paradox on intrinsic change consists in the problem of accounting for how it is possible that one and the same entity should have one intrinsic property at one time and a contrary property at another time. For example, how could one and the same bit of clay go from being hot to being cold, or from being cubical to being spherical. Leibniz's Law tells us that if a thing $A$ is identical to a thing $B$, then $A$ and $B$ must have all the same properties. If the clay goes from exemplifying the property of *being hot* to exemplifying the property of *being cold*, then the clay before the change, $C_1$, exemplifies a property that the clay after the change, $C_2$, does not. Therefore, $C_1$ cannot be identical to $C_2$. This is paradoxical because it is intuitively plausible for things to change properties (see Chapters 24 and 25 for more on this and related problems).

Presentists have the option of *taking tense seriously*. They can claim that the clay must now be either cubical or not cubical, hot or not hot. However, it is possible for it to be hot now but to *have been* cold (in the past), or to be *going to be* cold (in the future). The contradiction between being hot (now) and not hot (then) is dissolved, since the latter assertions occur with a past or future tense. As linguists and logicians put it, these uses occur "within the scope of a tense operator". We can put the assertions in these forms:

(25) Clay is hot.
(26) WAS(Clay is not hot).
(27) WILL(Clay is not hot).

There is no need for a present-tense operator. To say that something is (present-tense) hot is simply to ascribe the property of *being hot* to it, without qualification.

Tense operators are non-factive. From WAS($p$) or WILL($p$), nothing follows about the present. In particular, from WAS(There exists an $F$), nothing follows about the existence of $F$'s simpliciter. So, a commitment to (26) and (27) does not commit Presentists to there being something identical to the clay that lacks the property of *being hot*.

The main objection to this Presentist account of change involves an appeal to Truthmaker Theory (2.1T/2.1A.1T/2.1A.1A.1T). If we accept some form of Truthmaker Theory, then we have to ask what makes past- and future-tensed truths true. We will consider several arguments of this form in Chapter 21.

## 20.5 Arguments for Tensism

In the remainder of this chapter, we consider six arguments for Tensism.

### 20.5.1 The "Thank Goodness" Problem

We'll start with Arthur Prior's "Thank Goodness That's Over" argument (Prior 1959). Prior was a pioneer in the formal study of time and tense. Prior (1959) points out that we have very different attitudes towards events in the near future and the near past. If we know that we must undergo some painful surgery, then, upon learning that the surgery is over, we naturally express relief: "Thank goodness that's over!" Prior thought that this fact could be turned into an argument for Tensism.

In order to reconstruct Prior's argument, we need the notion of a Priorean proposition:

**Def D20.3 Priorean Proposition.** $x$ is a *Priorean proposition* if and only if (i) $x$ can be expressed by a sentence employed in some context, (ii) $x$ is a fundamental bearer of truth and falsity, and it can be true or false simpliciter (absolutely, without qualification), and (iii) $x$ can be the object of belief or other intentional attitudes (such as *hope*, *fear*, *relief*, etc.).

A proposition is a tensed proposition if it necessarily has its truth or falsity temporarily, while an untensed proposition necessarily has its truth or falsity permanently. A proposition that predicates a tensed property of something is a tensed proposition. Prior's argument can be taken as a proof that there are tensed propositions. If tensed propositions exist, then Tensism is true, since the truth of a tensed proposition is an A-property.

Let's try to reconstruct Prior's argument for the truth of Tensism:

1 Mental attitudes, like *relief* or *dread*, are relations between persons and Priorean propositions. (Premise)
2 Priorean propositions are truth-bearers: they are true or false simpliciter. (Def 20.3)
3 When a person is relieved that an experience is just over, there is no untensed proposition that could be the object of that attitude (and similarly for other temporal attitudes). (Premise)
4 Hence, mental attitudes, like *relief* or *dread*, are relations between persons and tensed propositions. (From 1 and 3)
5 Tensed propositions exist and are true or false simpliciter. (From 4 and 2)
6 Tensed propositions change their truth-values over time: they are sometimes true and sometimes false. (Definition of tensed propositions)
7 Hence, Tensism is true. (From 5 and 6)

There is now almost universal consensus that premise 3 is true. Premise 2 is a stipulative definition, and premise 6 seems to follow from the fact that A-truths have truth only temporarily. Hence, if the argument fails, it must fail in the very first step: mental attitudes don't take Priorean propositions as their objects.

We shouldn't overlook the fact that there is a lot to be said for premise 1. As David Lewis points out in "Attitudes De Dicto and De Se" (Lewis 1979a), there is a great theoretical advantage to assigning the objects of the attitudes to a uniform ontological category. Beliefs, desires, and other attitudes interact in ways that reflect the logical relations between their objects, and it is propositions that most naturally and generally possess the required network of logical relations. In addition, in the case of belief, the pull toward thinking of the object as something true or false simpliciter is, as we shall see, very strong.

OLD ANTI-TENSIST RESPONSES How Anti-Tensers respond to Prior's argument depends on whether they embrace Old or New Anti-Tensism. Old Anti-Tensers have two options. First, they can deny premise 1, claiming that untensed propositions can be adequate objects of relief and dread. Second, they can acknowledge that these attitudes do have tensed propositions as objects but deny that any of these tensed propositions are ever true. This second option involves an error theory about our judgments about the truth-values of tensed propositions.

Nathan Oaklander (1994) has taken the second option and defended an error theory about A-beliefs and A-attitudes. In many ways this is an elegant solution, accepting Prior's claim that only tensed propositions are adequate objects of many attitudes. Oaklander accepts the first three premises of the Priorean argument, disagreeing instead with premise 6: that tensed propositions have varying truth-values. Instead, Oaklander claims that all tensed propositions are necessarily false, since the required A-facts are necessarily absent. Such an approach is not unprecedented. There are error theories of morality (J.L. Mackie 1977), mathematics (Hartry Field 1980), and color judgments (Adam Pautz 2014), among others. It's a proposal that is well worth having on the table, but, in our view, should be considered only as a last resort.

What's wrong with untensed propositions as objects of attitudes? As Prior (1959) pointed out, the two obvious candidates are hopelessly inadequate in the case of the

Thank Goodness attitude. (28) and (29) are applications of these two candidates to a particular case:

(28) Thank goodness that the pain ended on February 7, 2007 at noon CST.
(29) Thank goodness that the pain ended just before the occurrence of this token.

As Prior remarks, it is inconceivable that one would thank goodness for either of these facts, nor do we have any explanation of why we would be inclined to feel thankfulness toward these B-facts at some times rather than at others, since we might be aware of these facts long before or long after the indicated time has passed.

In his 2006 dissertation, Max Goss proposed that we each have changing *now* concepts that pick out for us the present moment via a description that ascribes to that moment the total contents of our immediate experience at that moment. These concepts are, according to Goss, "*n*-senses" for particular times. Goss's account of a thank goodness attitude would go something like what is represented by (30):

(30) Thank goodness that the pain ended just before *I had a total experience state of type N.*

Goss argued that, although it is clear that this is a contingent rather than a necessary fact, it is nonetheless quite plausible to suppose that each moment of time during which one is conscious is associated for each person with a unique total experience type, so the various definite descriptions formed in this way will succeed in picking out a unique time. Moreover, it is plausible to suppose that there is a kind of contingent inaccessibility of the experience type for one moment at any other moment. Only when one is in an experience state is one in a position to be able to refer to that experience state in thought. At other times, the state is too complex and precise to be reproduced by the limited resources of one's general concepts or imaginative memory. Thus, Goss claims that his account can explain the de facto inaccessibility of these temporal thoughts at any time but the appropriate one.

However, even if it is a fact that one can think (30) only at a time while one is experiencing a state of type *N*, the complement of (30) still fails to capture what it is that one is grateful for. The only reason that one cares that the pain ended before one experiences a state of a certain type is that one knows that one is experiencing that type (uniquely) in the present. The problem is that there is only a contingent, accidental association between Goss's *n*-senses and their referents. Consequently, we cannot suppose that the *n*-sense delivers the right intentional element to the object of belief. Hence, Goss's analysis hasn't really avoided the need for a tensed proposition as content.

NEW ANTI-TENSIST RESPONSES: *DE SE* ATTITUDES To deal with the thank goodness problem, many New Anti-Tensers have invoked the idea of *de se* attitudes. The phrase '*de se*' is Latin for 'about or pertaining to oneself'. A *de se* attitude is a belief or other intentional relation directed *towards oneself as oneself.* If, for example, one is conceited, then one has a very favorable opinion of oneself. To be conceited, the opinion can't just happen to be about oneself. Suppose RCK thought that the tallest philosopher in the department was very smart, thinking that Professor Juhl is the tallest member, while in fact it is the case that RCK is the tallest philosopher in the department. Such an opinion doesn't count as making RCK conceited, since he is not attributing intelligence to himself *as himself*, but rather as the tallest philosopher in the department.

It is also possible to have an attitude that is about oneself *de re* ('about the thing') that happens to be oneself but not about oneself *de se*. *De re* attitudes have *Millian* or *Russellian* propositions as their objects. A *Millian (Russellian) proposition* is a proposition that contains some concrete thing as a part or component and not just some concept or identifying property of that thing. The Millian proposition 'Mt. Blanc is snowy' is an entity that contains Mt. Blanc itself, "with all its snow fields", as Gottlob Frege put it when objecting to such propositions to Bertrand Russell. *Fregean propositions*, in contrast, are wholly abstract. The corresponding Fregean proposition about Mt. Blanc contains some uniquely identifying concept or property of Mt. Blanc, not the actual mountain. Millian philosophers, including Bertrand Russell, have argued that Millian propositions are needed as the object of *de re* attitudes, including beliefs, desires, and other attitudes that are directly about some thing with which we are acquainted. Suppose, for example, RCK catches sight of a very handsome man in what he take to be a window, and RCK thinks that he is quite handsome (where the 'he' is directed *de re* to the thing RCK is seeing). If it turns out that RCK is really seeing a reflection of himself in a mirror, the opinion expressed is not conceited, since it is about a thing that happens to RCK himself, but it is not about RCK *as himself*. The reality of *de se* attitudes towards oneself provides substantial support for Anti-Tensers' claim that our attitudes take something other than propositions as their objects. When one thinks *I am tired*, one seems to be either in a special *de se* attitude toward a property (the property of *being tired*) or toward an incomplete proposition ('*x* is tired, or ___ is tired') or else one is thinking about the proposition that RCK is tired in a special, self-reflexive, *de se* way.

Anti-Tensers can plausibly argue that we can also have *de se* beliefs about time. There is a difference between believing that the war is *now* over and believing that the war is over in April 1865, even if "now" is April 1865. Sentences using tense and temporal adverbs like 'now' are *de se* sentences about the present moment as the present moment, not in the sense that the present moment has some special property, but only in the sense that the person thinking and speaking is wholly located in that time. It's as if the time in question were thinking of itself through the thinker of the now-thought.

One way out for New Anti-Tensers is the one taken by John Perry (1979) in accounting for the *essential indexical* as the object of a *de se* attitude. Perry argues that a belief state must be characterized both by its object, which is an untensed proposition, and by the *way* in which that untensed proposition is grasped. This "way" is an associated "character", which is a function which determines a unique proposition given a particular context. The notion of the *character* of a thought or proposition was introduced by David Kaplan in his seminal work "Demonstratives" (Kaplan 1989). To apply this to Prior's example, my thankfulness stands in a relation to one of the untensed propositions (28) and (29) above, but in addition it is related to the character corresponding to the sentence, 'The pain ended just before now'.

Alternatively, Anti-Tensers could instead suppose that some mental attitudes, like *dread* and *relief*, take objects other than complete untensed propositions. Here are some possibilities that have been proposed:

(a) Propositional functions or Kaplanian characters: functions from times to untensed propositions (Kaplan 1989, Theodore Sider 2001: 19–21)

(b) Properties (David K. Lewis 1986a: 28–30, 54–55; Roderick Chisholm 1989, 129–138)
(c) Sets of centered worlds (Lewis1979a)

A final option, briefly mentioned by Perry (1979), is still more radical. It involves giving up any uniformity to the objects of belief and the other attitudes. Instead of a single attitude of belief, there would a large family of belief-like attitudes, each taking a different ensemble of entities as its object. There would be a *believing* relation that took pairs of particulars and properties as its objects, another that took particulars, properties, and times, and so on. There would be a family of *now*-believings: if RCK now-believes <*a*,*F*>, then he believes that *a* is now *F*. Similarly, there would be an attitude of *now-relief* that RCK bears to the pain and the property of *ending when one is relieved that the pain has just ended*.

We will not pursue this most radical option. It seems to introduce an unacceptable degree of complexity and heterogeneity within our theory of the attitudes. This already runs afoul Ockham's Razor (**PMeth 1**). Further, in order to get many obvious inferences to come out as correct, we would have to postulate a large number of self-evident axioms relating the various versions of belief and the other attitudes to one another. In fact, it is not at all obvious that, once begun, it would be possible to place any limit at all on the complexity introduced. This seems like a blatant violation of the First Corollary of Ockham's Razor (**PMeth 1.1**).

Thus, the most plausible Anti-Tenser solution takes belief and other attitudes have something other than propositions as their objects. Philosophers have often assumed that it is one and the same kind of thing—namely, propositions—that play two crucial roles. First, propositions are the fundamental bearers of truth and falsity. Second, they are the object of belief and other attitudes. Anti-Tensers who adopt this solution must give up this assumption. Propositions are the bearers of truth or falsity, but what we *believe* belongs to a different category. Mental attitudes are relations to something like functions from times and persons (the 'now' and the 'me' of the thinker's world) to propositions. Hence, we cannot ask whether what a person believes is true or false simpliciter. We can only ask whether what the person believes yields a true proposition when *applied* to the thinker and the time of the thought.

THANK GOODNESS AS AN INFERENCE TO THE BEST EXPLANATION Prior's Thank Goodness observation could be used to support Tensism in an entirely different way, as part of an inference to the best explanation. Tensers can offer a plausible explanation for the rationality of our relief at the recent passage into the past of something painful and of the rationality of our joy at the imminent arrival of something pleasant from the anticipated future. Anti-Tensers have no comparable explanation.

Of course, Anti-Tensers can always posit that it is simply, as a matter of brute fact, rational to take differing attitudes toward the past, present, and future. However, this additional postulation counts against Anti-Tensism by way of Ockham's Razor (**PMeth 1**). Tensers have a simpler, less ad hoc explanation. The near future has the potential of being (fully) real, while the near past does not.

Tensers' explanation of the rationality of the thank goodness reaction appeals to the following principle of practical rationality:

**Value of Movement.** It is a bad thing for a bad thing to be approaching us; it is a good thing for a bad thing to be moving away from us.

The explanation must also appeal to the Vehicle Thesis:

**Vehicle Thesis.** As the present moves through time, those of us who continue to exist move with the present.

The Vehicle Thesis explains the thank goodness phenomenon, and it clearly entails the truth of Tensism. Does it favor one type of Tensism over the others? Simple Tensism seems inconsistent with the Vehicle Thesis. As the spotlight of actuality moves forward, it would not seem to be carrying anything forward with it. Presentism, however, is the version of Tensism that is most obviously consistent with the Vehicle Thesis. In fact, Presentism might even entail the Vehicle Thesis. If so, then Presentism gains credibility by virtue of its greater explanatory power.

### 20.5.2 Evidence against indexicality: Parallels between 'now' and 'actually'

As we have seen, the strongest version of Anti-Tensism is one that takes A-thoughts to be indexical in character, as something like functions from the thinker and the moment of thought to some untensed proposition. Anti-Tensers rely heavily on relatively uncontroversial examples of such indexicality, like sentences involving 'I' and 'here', to provide grounds for thinking that tense is just another indexical element. Unless one is a Solipsist, one doesn't think that the irreducibility of the first-person pronoun is any evidence for the claim that one is uniquely real or existent. Consider the egocentric, *de se* proposition expressed by (31) and its ordinary *de re* counterparts, (32) and (33):

(31) I am now leaking sugar from my grocery cart.
(32) John Perry is now leaking sugar from his grocery cart.
(33) *That* man is now leaking sugar from his grocery cart.

Clearly, there is a difference in significance between (31) and (32), even when both are being thought by John Perry. Perry may not realize that he is John Perry (or the John Perry who figures in proposition (31)). Maybe he is suffering from amnesia, for example. Similarly, there is a difference in significance between (31) and (33), even when (33) is being thought by John Perry and 'that man' picks out John Perry himself. Perry might see himself in a mirror without realizing that it is he himself whom he is seeing.

Anti-Tensers use these facts to argue that we have no reason to believe in special egocentric or 'I' facts in order to ground the truth of (31). (31) is made true by the same thing (whatever it is) that makes (32) and (33) true.

Anti-Tensers can then argue that, like egocentric propositions, tensed propositions give us no reason for positing metaphysically real or fundamental tensed truths. Instead, we should take tensed truths as indexical, as made true by whatever makes true the corresponding untensed proposition, as determined by the time of utterance or thought. An indexical sentence like 'It is now raining' is simply a function from moments of usage to untensed propositions. The same is true of past- and future-tensed sentences. The tensed sentence 'It will be raining' is a function that, when applied to any given moment of time, yields an untensed proposition stating that some time later than that time is one at which rain occurs.

Tensers will respond that there is no good reason to lump tense in with such indexicals as 'I' or 'here'. Instead, tense is much closer to grammatical "mood". We use the indicative mood to say what *simply is the case*, and we use the subjunctive mood to talk about what *isn't but might have been*. Similarly, Tensers can insist that we use the present tense to speak about what simply is the case, and we use past and future tense to speak about what *isn't but was or will be*.

The indicative mood (and related adverbs like 'actually') does not seem to be at all indexical. When one says that the earth *actually* has only one moon, one is not speaking about the location of one's utterance in some kind of "logical space". One is not saying that in our local reality, it is the case that the earth has only one moon. One is saying that the earth has only one moon *simpliciter*. If one says that the earth might have had more than one moon or no moon at all, one is not speaking of something true in some other, more remote part of reality. One is saying that our reality had the potential at one point to produce a three-moon or no-moon system around the earth. If someone were to say that the earth has actually had three moons, we would suppose that what he said was simply false. We wouldn't have to check first to see who said this and where in logical space the assertion was made.

Tensers can appeal to many parallels between tense and mood. For example, we sometimes talk about timeless entities such as numbers. When we do, we almost always use the present tense. We say that two plus two *is* four, not that it *was* four or *will be* four. This suggests that the appropriate use of the present tense is to speak of how things are *simpliciter*, without qualification about temporal location. This is why it is appropriate to use it in speaking of things entirely outside of time.

Some Anti-Tensers, in particular those like David Lewis (1986a) who are Modal Indexicalists, object to this line of argument on the grounds that the indicative mood and related adverbs like 'actually' or 'in actuality' are indexicals after all! Concretists, for example, say that when we speak of how things are in actuality, we are only referring to our local universe. Somewhere out there are speakers who can truthfully say, 'There are actually three moons around the earth', just as there are speakers in other times who can truthfully say, 'The Roman Empire is now intact'[1] or 'There are human colonists on Mars.' If these Modal Indexicalists are right, then Tensers lose an important argument for their position.

### 20.5.3 The experience of the flow of time

Most arguments for Tensism from experience have focused on aspects of experience that appear in a single moment, but it seems that the most interesting challenges to Anti-Tensism come from considering experience holistically or diachronically.

There is a methodological assumption lurking here. We take it for granted that a metaphysical theory gains support whenever it enables us to take the world of phenomenal appearance at face value, as accurately representing the world as it is. This is in the neighborhood of Reidian Common Sense (**PEpist 2**). Obviously, this evidence is defeasible, since there are cases where it seems plausible at least to take appearances to be misleading (in the case of secondary qualities, for example). Defeasible evidence is still evidence, however. If Anti-Tensism is able to explain, in the sense of explain *away*, the phenomena of experience, this lightens somewhat the evidentiary strength of that phenomenon for Tensism, but it doesn't entirely neutralize it.

*20.5.3.1 Experience considered atomistically* There is one datum concerning our temporal experience considered at a moment that seems to tell against Anti-Tensism, namely, the palpable fleetingness of the present moment. The sense we have of the present's having just succeeded a past moment and its being about to be succeeded by a new present. This fleetingness seems to represent moments as coming into and out of existence. (An excellent source for the phenomenology of time is Augustine's *Confessions*, Book 11. We don't know of anything better.) Anti-Tensers must attribute this sense of fleetingness to a kind of illusion. Past moments don't pass out of existence nor do future moments emerge from nothingness. They are all eternally existent. Some are merely "out of view", so to speak.

Anti-Tensers might claim that our sense of fleetingness is simply a perception of the brevity of the present moment, but this seems to miss the mark. Even when the present moment is specious. That is, even when we seem to experience an extended duration as present, the sense of fleetingness is not attenuated. So, it seems that the sense of fleetingness is independent of the length of the present moment.

*20.5.3.2 Experience considered holistically* As we saw when we considered Zeno's paradoxes in Chapter 19, there are reasons to deny that dimensionless instants of time are ontologically prior to finite periods of time. That is, there are good reasons to embrace Intervalism (19.1A). Untensed facts about time do not supervene on facts about dimensionless time-slices. In addition, there are irreducible metrical facts about the temporal distances between such slices. These metrical facts don't supervene on intrinsic facts about instants, including for example their number or ordinal arrangement. So the existence of periods of time with some sort of intrinsic, even if frame-relative, measure must be ontologically on a par with the existence of the instantaneous slices.

This metaphysical fact seems to support a phenomenological claim that Tensers could put forward: we have a genuinely *diachronic* experience of time. Our experience of time is not limited to our experience of the present but includes the experience of an ordered sequence of moments spread out in time. (We think the same metaphysical fact puts some pressure on a theory of persistence favored by Anti-Tensers, namely, Ramsey-Lewis-Sider Worms (24.3T.3). But we ignore this point in this context.)

Tensers could argue that our experience of time is diachronic in two senses. First, each moment is experienced as a specious present that includes some movement and change. Second, we experience time as a *sequence* of such specious presents. The idea of the specious present seems to be pretty well accepted by those who reflect on the phenomenology of experience, going back at least to William James's *Principles of Psychology* (1890). Let's focus here on the second, more contentious claim that our experience of time is as of an inexorably and really ordered succession of specious presents.

Here are two possible points of tension between that observation and Anti-Tensism:

- the *inexorability* of the temporal succession
- the appearance of real *change* in one's intrinsic mental state

First, the inexorability of succession. The succession moments, with earlier moments succeeded by later moments, appears real, categorical, and inexorable, as opposed to notional, hypothetical, or optional. The standard Anti-Tenser response appeals to some theory about the intrinsic directionality of time, either in terms of causal order (causes precede their effects) or in terms of increasing entropy or information. The intrinsic directionality of time is important in accounting for the fact that we remember the past and not the future and the fact that we deliberate about the future and not the past. There is, however, a very important gap between *intrinsic direction* and *real succession*. For example, the words in a well-written, meaningful book define a real direction within the text (standardly in English, this is left-to-right, top-down). This direction is intrinsic to the text as meaningful, and not merely the product of extrinsic conventions. However, this fact about the text in no way determines a real, as opposed to a notional, succession of *earlier* words by *later* ones. It is perfectly possible to read the words in such a book in reverse order or in any other order one might choose. In the same way, even if the moments of time are intrinsically directed, this by itself is not sufficient to explain the inexorable experience of real succession.

In *The Time Machine*, H.G. Wells toys with the idea that a well-trained mind could choose to read the moments of time in a non-standard order or with a non-standard velocity, shifting experience dramatically into the past or the future. This seems impossible, but it is unclear how Anti-Tensers could explain the impossibility.

What Tensers must claim is that we experience a unique, *sui generis* sort of succession in the succession of time. All other "order" on this view involves a kind of metaphorical or indirect reference to temporal succession. Without the real succession of times in the A series, we would have no concept of *earlier* or *later*, of *prior* or *posterior*. Our concept of ordinal numbers (first, second, third, etc.) are also dependent on the reality of temporal passage, since without real temporal passage, there is no real succession or direction. Mathematicians can model the ordinal numbers and the relation of succession in the intrinsically static world of set theory, with 0 modeled as the empty set, and the successor $n$ as the union of the members of $n$ with the set $\{n\}$. However, we can understand this construction as a model of the ordinal numbers only because we have a prior understanding of order as such, an understanding anchored in the real direction of temporal passage.

Merely identifying some static but asymmetric relation between world-states is not enough, not even if we label this relation as one of 'causation', 'memory', or what-have-you. We can always take the members of the domain of an asymmetric relation in any order: we can move up or down the set of integers, and we can always extend such a set to include members "earlier" than its original "origin". In contrast, there could be no such thing as an ordinal number "earlier" than *first* (the use of 'zeroth' is a kind of joke), and there is something essentially *backward* about counting from *third* to *second* to *first*. Only Tensism provides a metaphysical basis for these distinctions.

Could Anti-Tensers take a Kantian line (as in Kant's Analogies of Experience in the *The Critique of Pure Reason*) and argue that it is impossible for a rational person to do otherwise than order his memories in accordance with irreversible causal laws, such as increasing entropy? Here we need to distinguish between long-term and short-term memory. The Kantian line seems plausible as an account of our ordering of long-term memories, which do seem something like a stock of snapshots stored in some box in the mind. Presently, however, we are talking about very short-term memory, of our experience of the passage of time as an aspect of our present, lived experience. Here it seems the Kantian line would simply be wrong. We can easily imagine experiencing counter-causal and counter-entropic sequences in short-term memory in which broken eggs become whole, circular waves converge on a point from which a pebble jumps up, and so on. Again, there seems to be a clear gap between claims about the intrinsic directionality of time and our experience of the inexorability of temporal passage.

Second, the experience of real change in one's mental states. In *Logic Matters*, Peter Geach appeals to the experience of real change to support some form of Tensism:

> But now, quite similarly, even if my distinction between past, present and future aspects of physical things is a fragmentary mis-perception of changeless realities, it remains true that I have various and un-combinable illusions as to what realities are present. I must therefore have these illusions not simultaneously but one after another; and then there is after all real time and real change. (Geach 1972: 305)

According to Geach, our experience of time presents the self as numerically one, as an enduring entity, with mutually incompatible intrinsic states. This isn't merely notional change, as characterizes McTaggart's iron rod that "changes" from hot on one end and cold on the other. We don't experience ourselves, when we go from seeing something red to not, as having a part that sees red and a part that doesn't. Still less do we think of ourselves as a thing that sees red with a counterpart that doesn't. It is one and same thing that sees (first) red and (then) not red.

Some Anti-Tensers seem simply to deny that our experience presents reality in such a metaphysically laden way. They might plausibly claim that it is merely acculturation in a Tenser worldview that has colored our phenomenology in this way. We claim no expertise in the anthropology of time experience, but our impression is that this Tenser coloring is universal in human experience. (Some linguists in the mid-twentieth century claimed the contrary, but most later commentators seem to have dismissed these claims.)

As we said above, these phenomenological considerations are not decisive, but they must be given considerable weight. If we don't treat the appearances as innocent until proven guilty, we will quickly be driven into global skepticism. Further, the same observations may create problems for Tensism no less than for Anti-Tensism. This question is one that one must consider in order to evaluate the relative merits of Tensism and Anti-Tensism. In any case, one thing that Anti-Tensers cannot do is admit that real succession occurs in our experience, but not in the extra-mental world. If there are A-facts anywhere, whether in the mind or without, then Anti-Tensism is refuted. A hybrid view according to which there are A-facts with respect to the mind, but only B-facts with respect to the physical world, gives rise to a bizarre kind of dualism.

### 20.5.4 The present as a uniquely actual boundary between past and future

If we accept Aristotelian Finitism (Temporal Finitism 19.3T plus Temporal Anti-Discretism 19.4A), we have an additional argument for Tensism. Aristotelian Finitists argue that instants of time exist only as boundaries, dependent on the existence of temporally extended processes. Any process is indefinitely divisible in time, but any finite interval in the actual world can contain only finitely many actual boundaries. We could say that time is *indefinitely* divisible (there is no lowest bound to the temporal extent of any process), but it is not *infinitely* divisible: only a finite number of actual boundaries can coexist. An infinite division is metaphysically impossible.

If we accept such a view of time, and we think of the present as an actual boundary between past and future parts of occurring processes (which is how Aristotle himself suggested that we think of it), then we can construct the following argument for Tensism and for a strongly Tensist view of reality.

1 The present moment is a fully actual (not merely potential) boundary of all processes. (Premise)
2 If Aristotelian Tensism is false, then every moment contained by any process is a boundary of that process that is as actual as the present moment. (Premise: definition of Aristotelian Tensism)
3 So, if Aristotelian Tensism is false, then every moment contained by any process is an actual boundary of that process. (From 1 and 2)
4 Time is dense: any process contains infinitely many moments. (Premise: Anti-Discretism)
5 Therefore, if Aristotelian Tensism is false, then any process is divided by a temporally dense set of actual boundaries. (From 3 and 4)
6 Boundaries of potential parts are themselves potential: consequently, actual boundaries are boundaries of actual parts. (Premise: metaphysical dependency of boundaries on extended things)
7 Therefore, if Aristotelian Tensism is false, then every process contains temporally extended parts that are actual. That is, there are no temporally extended but actually undivided processes. (From 5 and 6)
8 There exists a fundamental level of temporal things, and this consists entirely of temporally extended processes. (Premise: Intervalism, 19.1A)
9 Fundamental intervals have only potential parts. (Premise: Temporal Finitism, 19.3T)
10 Therefore, Aristotelian Tensism is true. (From 7, 8, and 9)

This argument provides an argument for Aristotelian Tensism, for the view that the present moment is uniquely actual. It seems to be at least consistent with Simple Tensism and with Falling Branches and Growing Block Tensism. However, there is some tension between Aristotle's argument and Presentism. Aristotle's position entails that processes exist as fundamental, actual entities, even though their duration stretches into the past, and possibly even into the future. Presentists would have to argue that only that instantaneous part of the process that occurs now can really exist. How, then, can that instantaneous boundary be *dependent* on the temporally extended process of which it is a boundary? How can the existent be dependent on what is mostly non-existent?

We might consider a modification of standard Presentism, which we could call 'Aristotelian Presentism'. Aristotelian Presentists hold that the only things that exist are fundamentally processes that are going on at present, along with those entities that are participants in those processes. However, any process going on at present exists as a whole, including those parts that stretch into the past. This would be something like Growing Block Tensism, except that the past is thought to exist only insofar as it involves parts of presently occurring processes. More remote parts of the past no longer exist in any sense.

### 20.5.5 The threat of fatalism

We defined untensed propositions as those that possess their truth-value permanently. There is an important qualification to add. We may need to treat untensed propositions about contingent events in the future as a special case. Some Tensers hold that such future contingent propositions have no truth-value. Aristotle suggested such a position in his *Posterior Analytics*, using the example of a sea battle that may or may not take place tomorrow. If tomorrow is 30 September 2016, then the untensed proposition that such a sea battle takes place on 30 September would have no truth-value today (i.e., on 29 September 2016). The idea that future contingent untensed propositions have no truth-value, while all untensed propositions about the present or past are determinately true or false, may be behind the attractiveness of Growing Block Tensism.

Since Anti-Tensers hold that the differences between past, present, and future are merely perspectival, they would have no reason to treat the past or future differently in this respect. Some Tensers, such as Simple Tensers would agree. Modal Tensers and Presentists could take either position, since there could be primitive, simple facts about what will in fact happen (contingently) in the future. Thus, the truth-valuelessness of future contingent propositions is an option for Tensers but not for Anti-Tensers. If we have good reason to think that future contingent untensed propositions do lack a truth-value, this would provide another argument for Tensism.

**20.3T Indeterminacy of Future Contingents.** Some future contingent untensed propositions lack any truth-value.

**20.3A Determinacy of Future Contingents.** All future contingent untensed propositions are either true or false.

The main argument for the Indeterminacy of Future Contingents appeals to the existence of free will, in a metaphysically serious sense. Here is a standard version of the argument against what is called 'logical fatalism':

1 If future contingents are determinate, then any untensed proposition concerning what one will freely choose to do tomorrow has already been either true or false in ages past.
2 Any fact about ages past is beyond one's control, both now and in the future. (The fixity of the past)
3 Therefore, if future contingents are determinate, the truth or falsity of any untensed proposition concerning what one will freely choose to do tomorrow is beyond one's control, both now and in the future. (From 1 and 2)

4 The truth or falsity of any untensed proposition concerning what one will freely choose tomorrow logically entails that one will choose or not choose accordingly, and vice versa. (That is, the truth of *p* logically entails that *p*, and vice versa.)
5 If one cannot control whether it is the case that *p*, and *p* and *q* logically entail one another, then one cannot control whether it is the case that *q*. (The Transfer Principle)
6 Therefore, if future contingents are determinate, then one cannot control whether it is the case that one freely choose to take some definite course of action tomorrow. (From 3, 4, and 5)
7 One can control whether it is the case that one freely choose to take some definite course of action tomorrow. (Freedom entails self-control.)
8 Therefore, some future contingents are indeterminate. (*Reductio* from 6 and 7)

Anti-Tensers have the option of denying the reality of free will, in the strong sense involved in step 7. However, most Anti-Tensers will object already to step 2. They might deny that the past is already fixed. For example, some Anti-Tensers accept the possibility of time travel. Even if they admit that the past is fixed in some way, they might deny that it is fixed in all respects. In the Middle Ages, scholastic philosophers such as William of Ockham proposed a distinction between *hard* and *soft* facts about the past. The fact that the Battle of Hastings occurred in 1066 AD is a hard fact about the past: it is clearly beyond our powers of control. However, the fact that it is now true that a sea battle will take place tomorrow might well be a soft fact, one that is still under our control.

On Ockham's view, the future is genuinely open relative to the present, in the sense that multiple futures are now possible, even though only one of them is the actual future. In effect, Ockhamists can accept McCall's Falling-Branches picture of time, while holding that there is a unique *stem* that leads from the present through all future branch-points, a stem that represents the unique, actual future. Hence, for Ockhamists, branches don't *fall off* the tree of time: whether a branch *B* is possible at a time *t* depends simply on whether *t* is located before the root of *B*.

Tensers could respond by insisting that this distinction is forced and ad hoc. Either the past is fixed *in toto* or it is totally open. In addition, Tensers could argue that Anti-Tensers have no account of why the past is in any respect fixed or closed or beyond our control, while the Tensers can explain this fact in terms of the determinacy of the past, in contrast to the indeterminacy of the future. We will take this issue up again when we consider the problem of the direction of time and of causal control in Sections 27.1.1.2 and 28.5.

### 20.5.6 Circular time and time travel

Is time travel possible? If Tensism were true, then it would seem to be impossible. For example, if Presentism is true, then there is no future or past to travel into. If we adopt instead Simple Tensism, time travel still seems impossible, since it would involve some "spark" of actuality breaking away from the unique present and wandering off into a region of near-total non-actuality. What kind of time travel would that be? The realms of past or futurity we reach would be devoid of any reality. Finally, time travel in a world in which Modal Tensism is true seems equally absurd. Travel to the future would be travel to a realm lacking any closure or definition, while travel to the past would involve encasing oneself in the inexorable grip of fixed fate.

In the case of Anti-Tensism, it is not obvious that time travel is possible, but neither is there any obvious barrier to its possibility.

So, is time travel possible? Could one travel into the past or the future?

There are at least two reasons for thinking that it might be possible. First, it seems to be conceivable, as evidenced by the many time travel stories written in the last hundred years. Given Imagination as Guide to Possibility (**PEpist 1**), this is good reason to think time travel is possible. Of course, this reason is defeasible, so arguments for Tensism might count as evidence that defeats the evidence supplied by the conceivability of time travel. Second, there are solutions to the equations of Einstein's general theory of relativity discovered by mathematician Kurt Gödel that describe a world in which time is partly circular. This would mean that some objects could, by traveling far enough through the universe, arrive back at their destination before they departed. This would seem to involve real time travel, and it is reasonable to take our best scientific theories at face value:

**PMeth 2 Scientific Realism.** Other things being equal, adopt the theory that implies that our best scientific theories are straightforwardly true, as standardly represented.

However, general relativity by itself cannot establish the real possibility of time travel, since even assuming that it is true, it still might be incomplete as a description of the physical constraints of the world. Something outside the scope of general relativity might render such circular time impossible.

A well-known argument for the impossibility of time travel is the grandfather paradox.

1. If time travel were possible, it would be possible for someone, Mr. X Jr., to travel back in time and kill his own paternal grandfather before Mr. X's father, Mr. X Sr., was conceived.
2. It is impossible for Mr. X Jr. to exist if his paternal grandfather died before his father was conceived.
3. It is impossible for Mr. X Jr. to do anything, including travel back in time, if he doesn't exist.
4. Hence, it is impossible for Mr. X Jr. to travel back in time and kill his own paternal grandfather before Mr. X's father was conceived.
5. Hence, time travel is impossible.

The defender of time travel might reject step 2, with its implicit assumption that the genealogical origin of both Mr. X Jr. and of Mr. X Sr. are essential to their identities. Still, it seems we could change the story by enabling Mr. X to annihilate the earth, or even the entire cosmos, before his own conception. Thus, the most vulnerable premise is the very first one. Why think that just because time travel in general is possible, it must be possible to kill one's own grandfather as a toddler?

The response to this challenge is to resort again to patchwork principles like Finite Patchwork:

**PMeta 5.1 Finite Patchwork.** If an event or process of (intrinsic) type *A* is possible, as is an event or process of intrinsic type *B*, and if there is enough room in the history of the world to locate in it instances of both events (or processes) without overlap in time and space, then it is possible for events (or processes) of both types to occur together.

Let's suppose that it is possible for Mr. X Jr. to travel into the past, so long as he doesn't disturb anything essential to his own existence. Call this process *A*: Mr. X Jr. travels into the past and passively observes events, without disturbing anything. It is surely possible for Mr. X Jr. to kill a toddler, so long as the toddler is not his own grandfather or other ancestor. Call this toddler-killing scenario process *B*. Finite Patchwork justifies us in assuming that there must then be a possible world in which processes *A* and *B* are spliced together, resulting in the grandfather paradox. The only solution seems to be to deny that any time travel is possible at all.

The cogency of this line of reasoning depends on the Intrinsicality of Powers (**PMeta 2**). If powers are intrinsic, as Powerists (4.4A.3) and Nomists (4.4A.2) believe, then Finite Patchwork (**PMeta 5.1**) is applicable to the toddler scenario, with the result that time travel is impossible. However, if powers are not intrinsic to their bearers, as Neo-Humeists (4.4T) contend, then Finite Patchwork does not apply, since we have not limited ourselves to intrinsic facts in describing the toddler-killing scenario (Lewis 1976). Whether one is able in a certain situation to kill a toddler depends on the pattern of qualitative facts that pervades the actual world.

The cogency of the argument also depends on the question of Transworld Identity (16.1T) versus Worldbound Individuals (16.1A). According to the Transworld Identity, in order for Mr. X Jr. to be able to kill Mr. X Sr.'s father as a toddler, there must be a world in which they both exist and in which the first kills the second. This seems to be impossible, since Mr. X Jr. cannot exist in any world in which Mr. X Sr.'s father has died as a toddler. However, if we adopt Worldbound Individuals, and in particular Counterpart Theory (16.1A.1), instead, all we need is for Mr. X Jr. to have a counterpart in a world who kills a counterpart of Mr. X Sr.'s father. This seems possible, since there doesn't seem to be any bar to Mr. X Jr.'s having a counterpart that isn't a grandson of Mr. X Sr.'s father's counterpart (in that world).

## 20.6 Conclusion

In this chapter, we have been examining arguments in favor of Tensism. These arguments fall into two main categories. First, there are those arguments, like the Thank Goodness argument, our experience of the flow of time, and the reality of intrinsic change, that appeal to how time appears to us in our ordinary, everyday experience. Metaphysicians must take such data seriously if they are not to embrace a global skepticism about the world of appearances. However, such considerations are in principle defeasible. Second, there are arguments that presuppose a particular conception of time and of the nature of power. Those who embrace Powerism and an Aristotelian views about time have several good arguments for embracing Tensism, namely, the threat of fatalism, the impossibility of time travel, and the present as an actual boundary between past and present. In the next chapter, we consider arguments against Tensism.

## Note

1 Well, we don't suppose that they would have said it then in modern English. Better: 'Imperium Romanum intactum est,' or something like that.

# 21

# Arguments for Anti-Tensism

In this chapter, we look at six arguments against Tensism (20.2T). They are, equivalently, arguments for Anti-Tensism (20.2A). The arguments are of three basic kinds: those that argue that Tensism is incoherent or mysterious, those that argue that it is in irresolvable conflict with modern science, and those that fault Tensism for its unexplainable or brute necessities. In Section 21.1, we consider the objection that Tensism cannot sensibly account for the rate of the flow of time. This is followed by the long Section 21.2, in which a variety of objections based on Truthmaker Theory (2.1T/2.1A.1T/2.1A.1A.1T) are lodged against various versions of Tensism. In Section 21.3, we look to science, especially Einstein's theory of relativity, as the basis for an argument against Tensism. Relativity theory seems to deny the possibility of a cosmic "now", since it denies the existence of absolute, or observer-independent, *simultaneity* relations. The final three sections concern additional philosophical objections to Tensism. Section 21.4 considers epistemological problems for Tensism; Section 21.5 considers McTaggart's argument that Tensism is self-contradictory; and Section 21.6 explores the number of brute necessities for which Tensism can provide no simple explanation.

## 21.1 How Fast Does Time Flow?

Both C.D. Broad (1938) and J.C.C. Smart (1949) raised the following sort of argument against Tensism:

1 Tensism implies that time itself passes, that events and times continually change their position in the A series.
2 Any process of passage or change takes place at some sort of rate.

*The Atlas of Reality: A Comprehensive Guide to Metaphysics*, First Edition.
Robert C. Koons and Timothy H. Pickavance.

3 The pure passage of time is a process that cannot take place at any rate, since (i) if its rate is measured against some kind of meta-time, then Tensists face an infinite regress, and (ii) the trivial rate of one second per second is absurd.
4 Therefore, Tensism is false.

Some Tensers (for example, George Schlesinger 1980) are willing to embrace the reality of meta-times but deny that an infinite regress results. (Schlesinger proposes that there are two A series, each flowing at a rate to be measured against the other.) One could go further and embrace an infinite series of meta-times while denying that this infinite regress is vicious. This isn't a very popular line, however. Therefore, we ignore it here.

Ned Markosian (1993) seems to have the right Tensist response to this argument. Markosian distinguishes two possibilities. First, it may be that all talk of rates of processes is irreducibly relative. The most we can do is compare the progression of one process with that of another. Second, it may be coherent to talk about the absolute rate of a process, that is, its rate when its progression is compared with the passage of time itself.

On the first possibility, there is no difficulty in specifying the rate of the passage of time. Time passes, for example, at the rate of one year per 365.25 revolutions of the Earth. We can pick any process whose regular progression is constant when compared with the flow of time itself.

Markosian seems to suggest that the second possibility entails Tensism. If there were no pure passage of time, there could be no absolute rates or velocities, since absolute rates are simply the comparison of the progression of one process with the progression of time itself (p. 841). However, this isn't right. Anti-Tensers could insist that the time dimension has an intrinsic measure, so that it would make sense to measure the absolute rate of process by using equally spaced benchmarks along the B series.

Markosian divides the second possibility into two sub-cases: (a) it makes sense to determine the rate of time's passing by comparing its progress with itself or (b) it does not make sense to do so. If (a), we should embrace the trivial answer that time passes at the rate of one second per second. If (b), we should say that the rate of time's passage is undefined because all absolute rates are defined by reference to it. One might defend (b) by saying that we are not allowed to use the rate of time's passage to define its own rate, for the sort of reason that Wittgenstein proposes for saying that it makes no sense to ask for the length of the standard meter stick. If we do so, one might think we are committing what Gilbert Ryle called a "category mistake", applying a description to something in a category of the wrong sort. Of the two, (a) seems more plausible. Wittgenstein is wrong about the standard meter stick—it seems to be straightforwardly true that it is one meter long!

So, what exactly is supposed to be wrong with saying that time passes at the trivial rate of one second per second? Robin Le Poidevin (2005) offers a suggestion:

1 If a process progresses at a rate of $x$, then it is metaphysically possible for it to progress at some rate $y \neq x$. In fact, it is metaphysically possible for it to progress (regress) at the rate $-x$.
2 It is metaphysically impossible for time to progress at any rate other than one second per second, and in the standard direction (with times moving into the past).
3 Hence, time does not pass at the rate of one second per second.

Markosian might respond that step 1 derives its credibility from the first possibility above: it should be metaphysically possible for time to progress at different rates, when those rates are defined by reference to the progression of other processes. We could imagine a world in which all physical and mental processes progressed at exactly twice their current rate. In such a world, time would pass at a rate exactly one-half as fast (as defined with reference to any of those other processes). Once we realize that we are considering the second possibility, though, it may not be so plausible to assume that time could progress at a different rate. When that rate is measured by comparing the passage of time to itself, it's not obvious that we can imagine it passing at a different rate.

The problem of the rate of pure passage is a problem for anyone who believes in absolute units of time, whether Tenser or Anti-Tenser. Anti-Tensers should concede that time passes in exactly the same sense that a child grows, by using an At-At Theory of change (compare 24.5A.1T). The child changes its height in the sense that we find, as we pass mentally along the time-axis, the child having progressively greater heights. In exactly the same way, as we pass mentally along the time-axis, we find events having progressively more *past-ish* B-times. If time has an intrinsic measure, then Anti-Tensers must also embrace the one second per second account of its rate. Tensers could avoid the problem of explaining how fast time flows by denying that time has an absolute, intrinsic measure. Thus, the problem seems to be orthogonal to the debate between Tensers and Anti-Tensers.

What about the possibility of time changing direction? This would seem to make sense only for those who thought that events had a position in the B series that was metaphysically independent of the A series. That is, if, like McTaggart, we thought that the B-times of events supervened on their A-times, then time flowing backwards would make no sense. The B-times of events necessarily correspond to the order in which they had become present.

But couldn't we imagine the A-times of events moving in the opposite direction, such that events become more future and less past? Probably not, since past and future are, on this view, distinguished by the different ways in which they "react" to the passage of time. Could we imagine time moving first in one direction and then in the opposite one (whichever direction we label as 'past' or 'future')? This seems to be a real problem for Tensers, one calling for some sort of asymmetry between past and future that explains the impossibility of such a reversal.

Would Growing Block Tensism (see Section 20.4) solve this problem? It could only if the growing block could not become a shrinking block. We see no principled way to rule out the shrinking block.

Dynamic modalities, of the kind proposed by Falling Branches Tensism (20.2T.2), are more promising. It seems plausible to think that what is impossible cannot become possible, and the necessary cannot become contingent, but that the contrary transformations are possible. If so, and if the future is associated with alternative possibilities and the past with a clear division into the impossible and the necessary, then we would have grounds for the impossibility of time reversal. This, however, merely raises the further question of why the impossible cannot become possible.

If we think of the B series as independent of the passage of time, Simple Tensism (20.2T.1), then Le Poidevin's refutation seems conclusive.

## 21.2 Truthmakers for Truths about the Past

Presentism (20.2T.4), although attractively common-sensical in many ways, has a number of problematic consequences concerning truthmakers for propositions about the past. According to Presentism, everything that exists exists in the present. There is nothing existing entirely in the past. Alexander the Great's horse Bucephalus, the Hanging Gardens of Babylon, dinosaurs: none of these exist according to Presentism. If Actualism (12.1T) is true, then no existing thing can make true such propositions as that Alexander once had a horse, that Babylon was once graced with beautiful hanging gardens or that dinosaurs once roamed the earth. Presentists have offered four accounts: (i) reductionist Presentism or Occurentism, (ii) *sui generis* past-tensed properties of present things, (iii) primitive past-tensed facts or states of affairs, and (iv) the wholesale rejection of Truthmaker Theory. Each has its difficulties.

### 21.2.1 Reductionistic Presentism or Occurentism

Occurentism is the thesis of the reducibility of the past and future to present-tensed facts. On this view, the past is wholly constituted by present memories and traces of it, and the future is wholly constituted by presently existing pre-determining conditions. This is the view of the logician Jan Łukasiewicz (1967: 38–39), Michael Dummett (1969), and Peter Ludlow (1999: 150–152). This view implies that aspects of the past literally cease to exist once all traces of them have disappeared. This seems a pretty crazy view and one that is dialectically unstable, since the argument for Tensism relies so heavily on a common-sense epistemology. Indeed, we can imagine situations in which all traces of some past event have disappeared. If this happens, the present underdetermines the past. But the past seems fully determinate, and so Occurantism is at least at odds with our common-sense view of what is possible. This violates Reidian Common Sense (**PEpist 2**).

There is a further intentionality objection. Statements about the past are simply not about present-day traces. Statements about dinosaurs are about dinosaurs, not about fossils and other vestiges of dinosaurs. Occurentism, therefore, seems to get facts about the contents of our language and mental states wrong.

### 21.2.2 *Sui generis* properties of present-day objects

Tensers might instead appeal to primitively tensed properties of eternal things like particles or parts of absolute space in order to ground truths about the past. This is John Bigelow's (1996) account, which follows the suggestions of ancient Stoics and Epicureans. The proposition that George Washington slept in this very bed is made true by the fact that this bed (or this part of space) has the property of *having been slept in by George Washington*. If one is a Presentist, one will need to appeal to something like the property of *having been slept in by someone named 'George Washington'*, since George Washington is a past object. That there were once hanging gardens in Babylon is made true by some presently existing particles exemplifying the property of *having composed some hanging gardens in Babylon*. And so on.

If there is no matter or region of space handy, we could make the world the bearer of the property. It is true that dinosaurs once existed, and that there have never been unicorns, because of (1) and (2), respectively:

(1) The world has the property of *having once contained dinosaurs*.
(2) The world has the property of *never having contained unicorns*.

There are at least two objections to this view. First, these are just the wrong truthmakers. For example, some matter could be annihilated without annihilating all the facts about what it once composed. Second, there is again an intentionality objection. Statements about specific, localized past events are not about the whole world. To select anything short of the whole world would seem arbitrary.

### 21.2.3 Primitive past-tensed facts or states of affairs.

Arthur Prior (1968: 1–14) seemed to suggest that there were facts or states of affairs that are primitively past-tensed, and that these could be Presentist-friendly truthmakers for statements about the past. Theodore Sider (2001) argues that this amounts to metaphysical "cheating", as it runs afoul of truthmaker principles. What Sider means is that Prior's primitive tensed-facts are somehow *hypothetical*, and the hypothetical ought to supervene on the non-hypothetical, the categorical. Prior would object to both premises, and Sider offers little by way of argument for his contentions.

The answer to these worries is given by careful attention to David Lewis (2002). Lewis's article is primarily a criticism of those (like Mark Johnston 1983, 1987) who would try to combine Eternalism (20.2A.1T) with enduring things, things that are wholly present at more than one time. Nonetheless, Lewis's central objection to Johnston can be turned to a different purpose by Presentists.

Johnston's solution to the paradox of intrinsic change was to suggest that the copula linking subject and property involves essentially a relation to a time. The stick is-at-*t* bent, but the stick is-at-$t'$ straight. Lewis objects that such a copula "alienates" the stick from its properties. The stick is never straight or bent simpliciter but only bears some relation or other to *straightness* or *bentness*. If one retorts that there is, in any case, a relation of *instantiation* between a thing and its properties, then one faces the onslaught of Bradley's Regress (see Section 7.2.1.3): if *a's being F* depends on a real relation of *instantiation* holding between *a* and *F*, then it follows that *a*'s instantiating *F*-ness depends on a real relation of *instantiation* holding between the pair <*a, F*> and the relation of *instantiation*. And so on, ad infinitum.

Lewis appears to be right. A tensed-copula is inherently alienating. Ironically, this fact provides the Presentists with a solution to Sider's truthmaker objection: a past-tensed fact consists in something's bearing a tensed-copula (*was* or *will-be*) relation to some property or cluster of properties, while the corresponding present-tensed fact consists in that thing's having that property *simpliciter* (without any alienating copula). Thus, the stick is straight simpliciter (no alienation), but it has the *is-at-t* relation to the property of *bentness* (alienation) for some past time *t*. The alienation of objects from their past-tensed properties is exactly what Presentists want! The contrast between alienation toward past-tensed properties and no alienation toward present-tensed ones secures the uniqueness of the present.

This solution, like the rejection of Truthmaker Theory we consider below, has a problematic feature, namely, necessary connections between distinct facts. If *a is now F*, then at all future times, *a* will bear the *is-at-t* relation to *F-ness*, where *t* is the time that will hereafter correspond to the present moment. (We're imagining Presentists who believe that there are only past times. We can speak about the present moment, but only loosely, since strictly speaking by the time has uttered such a phrase there is no longer the relevant moment, although another will immediately begin to exist that corresponds to the facts that presently obtain.) This connection between present and future past facts is a necessary one, even though the two facts involve different relations. The *past-in-the-future* fact involves a time-relativized copula that does not yet exist. Such necessary connections violate the Second Corollary of Ockham's Razor (**PMeth 1.2**).

This account is different from Growing Block Tensism, since the present isn't simply the leading edge of the past according to Presentism. There is a clear metaphysical difference between the present and any past moment. Only present predications are predications simpliciter. Past predications, in contrast, introduce an intermediate relation between a thing and its quondam properties.

There's an alternative approach that works but that seems much less attractive. One could postulate that corresponding to every *n*-place relation, there is an *n+1*-place relation that takes a time as its last relatum. So, corresponding to the one-place property *bentness*, there is a two-place relation *bent-at-ness* that holds between things and times. If *a* is now bent, then in the future it will be the case that *a* bears the *bent-at-t* relation to the time *t* that will correspond to the present moment.

Michael Tooley (1997) argues that all these views are burdened by a large number of unexplainable brute necessities, such as the incompatibility between *pastness* and *futurity* (or between one degree of *pastness* and another) and the systematicity of time flow (the synchronicity with which the relevant modalities are stepped up as time passes). Again, such necessities are costs of these views (**PMeth 1.2**).

Here are two responses to Tooley. First, it's not easy to say which theory of time postulates the greatest number of brute facts. All of them seem to require many, and so it may be that Ockham's Razor doesn't tell in favor of one over the others. Second, Ockham's Razor has a proviso: we don't multiply entities *without necessity*. If alternative accounts fail or are incredible or have decisive disadvantages, we could well be justified in positing the brute necessities Tooley lists.

### 21.2.4 Rejection of Truthmaker Theory

Presentists can circumvent these problems by rejecting Truthmaker Theory *tout court*. This would leave Presentists without any account of the metaphysical and logical independence of atomic facts about the past. But it is not clear how much this matters.

### 21.2.5 Truthmaker problems for Growing Block Tensism

Growing Block Tensism avoids most of the truthmaker objections lodged against Presentism, since Growing Block Tensers acknowledge the reality of the past, as well as that of the present. There is, however, a truthmaker problem for Growing Block Tensism. Consider the proposition that the battle of Hastings is at the leading edge of reality. This proposition was once true (in 1066), so, presumably, it had then a truthmaker. According to Growing Block Tensism, however, once a fact comes into existence, it stays in existence. Therefore, the truthmaker of this proposition must still exist. And yet the proposition is now clearly false! Either reality is incoherent because it includes truthmakers for both the proposition and its negation, or the size of the growing block of time is irreducibly a matter of one's perspective.

To avoid this conclusion, the defender of Growing Block Tensism must distinguish between those facts that endure and those that do not. B-facts endure, but total-block-relational facts, like how far a given event is from the leading edge, do not. Thus, Growing Block Tensism seems committed to at least a moderate analogue of Presentism about the total-block-relational facts. It thus faces truthmaker problems pertaining to past-tensed growing-block propositions of exactly the sort that Presentism faces. It's unclear what has been gained. Growing Block Tensism appears, therefore, to be dialectically unstable.

Tooley (1997) argues that there is a great difference between an event's having different intrinsic A-properties at different times and an event's having different A-relational properties (such as the *distance to the edge of the growing block of time*) at different times. We're unclear as to how this difference is supposed to be decisive. Presumably, relational propositions require truthmakers just as much as intrinsic-property facts do. Perhaps this is not the case, but more needs to be said by way of explanation, and Tooley offers no help on that score.

### 21.2.6 Further truthmaker problems: Cross-temporal spatial relations (Motion)

Sider (2001) argues that Presentists cannot account for cross-temporal spatial relations, like *being in the same place at different times* or *being in a different place at different times*, and so cannot account for such features of the world as velocity, acceleration, and continuous and discontinuous motion. On the Presentist picture, we have two kinds of facts. First, there are those that make up the present moment; these obtain simpliciter. Second, there are those that once made up past moments; these can be spoken of only via instantaneous temporal operators like 'once upon a time' or '*n* units of time ago'. Thus, Presentism's world consists of an infinite collection of *snapshots*. The problem of cross-time relations is that of defining the correct 4-dimensional geometry of kinematics, of characterizing locomotion (accelerated vs. un-accelerated, continuous vs. discontinuous). We could line up the snapshots in such a way that each enduring thing jumps discontinuously from one location to another at each moment. Presentists need an account of why only certain arrangements of these snapshots are appropriate.

The first thing to notice is that if Presentists follow postulate absolute, substantival space, then the problem disappears. We can use the standard, Russellian At/At Theory of Motion (24.5A.1T), according to which something moves in so far as it is located at different places at different times. This allows us to construct a unique function, assigning each object's absolute spatial location at each moment of time, which in turn will supply the needed kinematic geometry. So, let's assume that Presentists don't go that way, instead adopting some version of the minimalist reinterpretation of special relativity described in the next section.

This is certainly a non-trivial problem for Presentists, but we shouldn't lose sight of the fact that there are comparable problems for Eternalists who reject the existence of absolute space. Without absolute space, we cannot ask whether a particle's absolute motion is continuous or discontinuous (since there is no such thing) but only whether the relative motions of two or more particles are continuous or discontinuous. It seems that Presentists and Eternalists are in the same situation here: both can assert that relative motions are discontinuous just in case the relative spatial positions of the particles change discontinuously over time. This answer is available to Presentists as much as to Eternalists, since Presentists have access to all of the facts about the relative spatial positions of particles at times (the synchronic spatial relations).

However, this objection to Sider overlooks the affine and topological structure of spacetime, as discussed by Lawrence Sklar in *Space, Time and Spacetime* (Sklar 1974). In neo-Newtonian and Minkowski spacetime, there is a level of structure that is intermediate between old-fashioned Newtonian absolute motion and the merely relative motions recognized by Leibnizians. We can distinguish between continuous and discontinuous paths of particles, even when the relative positions of the two paths vary continuously over time. Similarly, we can distinguish between accelerated and un-accelerated paths, even when all the relative velocities are constant (as in the case of distinguishing Newton's rotating bucket from a stationary bucket, which we discussed in Section 17.3). The distinction could be based (for Eternalists) on a primitive relation of *linear betweenness* holding among triples of spacetime points (including cases in which the three points occur at different times).

For Eternalists, this affine structure can be fixed by metaphysically primitive features of four-dimensional paths. This solution would seem to be unavailable to Presentists, since it involves the obtaining of primitive *diachronic* facts, that is, facts that span an interval of time.

Sider considers two families of solutions for Presentists, holistic and atomistic. Holistic solutions involve assigning the properties of *continuity* and of *constant-velocity* to particle-paths in a way that maximizes some holistic feature of the assignment. For example, we could maximize the number of paths that are continuous or un-accelerated. As Sider points out, this has counterintuitive consequences, especially with respect to small-world thought experiments. Suppose Newton's rotating bucket constituted the entire world. It is still possible for it to follow an accelerated path through space, even though assigning it such an accelerated path would violate the proposed holistic constraint.

A better version of the holistic strategy would be to maximize the applicability of simple laws of motion, including force laws, like gravity and electromagnetism. If no force is acting on the bucket, then the preferred assignment would give it an un-accelerated path; in contrast, if the bucket is under the influence of some force, then the preferred assignment would give it an appropriately accelerated path. This still doesn't eliminate all plausible counterexamples. We could imagine a world that accidentally satisfies some very simple laws of motion. The holistic assignment would then assign accelerations inappropriately, in accordance with merely apparent laws of motion.

Perhaps, then, Presentists should take the holding of the laws of motion and the force laws to be a metaphysically brute fact, as Nomists (4.4A.2), Powerists (4.4A.3), and

Hypotheticalists (4.4A.1) might believe. If the laws of nature dictate continuous motion, then the correct assignment of affine structure should assign continuous paths to all particles, and it should assign the appropriately accelerated paths, given the configuration of forces, to each particle. This will work, however, only if a violation of a world's laws of motion is metaphysically impossible. If there are worlds containing miracles, then Presentists will be unable to find a suitable assignment to that world of affine structure.

Alternatively, Presentists could take an atomistic, rather than holistic, approach. There would have to be a system of acceleration vectors possessed by some particles at each moment of time. If a particle has had an acceleration vector throughout some period of time, then it must be assigned a continuous path, and if it has had a zero-valued acceleration vector throughout that period, then it must be assigned an un-accelerated path. This enables Presentists to acknowledge the metaphysical possibility of a miracle in which a particle's atomistic acceleration vector doesn't conform to the forces acting on it.

Now, what is supposed to be wrong with this? The greatest problem for the atomistic approach is to explain the connection between acceleration vectors and the actual changes in relative positions of particles. As Tooley (1988) suggests, we could use a kind of Ramsification strategy (see Chapter 6 for more on Ramsification). Atomistic acceleration properties of particles are those properties that play a specified causal role with respect to the effects of the force laws and the causes of the changes in relative position over time.

Sider argues that Presentists face some sort of vicious circularity here, since they cannot merely help themselves to a function relating position in the affine structure to time. However, Presentists do have full access to the set of facts about the configuration of forces and to the set of facts about changes in relative positions of particles over time. We don't see why these two sets of facts aren't sufficient to provide a Ramsified definition of acceleration properties. The change in relative velocity between two particles should be a function of the vector sum of the two particles' acceleration vectors, and the acceleration vectors should, together with the intrinsic forces, satisfy Newton's $F = ma$ force law. It seems, in fact, that Eternalists are in much the same boat. How else do we gain cognitive access to the affine structure of the physical world except as the product of forces and as the explanation for changes in relative position?

Once again, however, the possibility of miracles seems to be a persistent problem for Presentists. It seems possible for a set of particles to change their relative positions over a period of time in a way that diverges from the atomistic vectors possessed by the particles at each moment. Presentists must invoke some kind of metaphysical necessity linking the intrinsic acceleration vectors and changes in relative position over time, a necessity so strict as to rule out all miraculous deviations. Eternalists need no such necessary connections. This is a clear advantage for Eternalism.

### 21.2.7 Truthmaker problems: Other cross-temporal relations between past entities

Presentists also face problems to do with quantification and existence. It seems that there are things that once existed but no longer do so, and Presentists have no truthmakers for such truths, since there aren't things that exist wholly in the past. We focus first on an example from Lewis (2004b):

(3) There have been two English kings named 'James'.

James I and James II had non-overlapping lifetimes. Thus, Presentists' analysis of this sentence must take this form:

(3*) WAS$(\exists x)$($x$ is an English king named 'James' & WAS $(\exists y)$($y$ is an English king named 'James' & $x \neq y$))

The difficulty arises in giving the truth-conditions for the final clause, '$x \neq y$'. Since James II (our '$x$') did not exist at the same time as James I (our '$y$'), he does not belong within the domain of quantification associated with the second, embedded 'WAS' operator. Yet (3*) seems to entail that James II did exist then, so that we can meaningfully assert, within the double embedding, that James I ≠ James II.

Lewis argues that there is a solution available to Presentists. They can make use of not only past-tensed sentences, but also past-tensed predications. In other words, we can suppose that James II had, while he was alive, the past-tensed property of *having been distinct from James I* (or, more precisely, the property of *having been distinct from an English king names 'James'*). This solution has two puzzling consequences. First, it entails that people can have properties that predate their own existence. When RCK came into existence, he immediately acquired a host of past-tensed properties, namely the properties of *not having existed at any previous point in time* (and, consequently, the properties of *not being identical to any of things existing then*). But there seems to be something paradoxical about saying that RCK has *any* property (even the property of *not having existed*) indexed to any time before his existence commenced. *Serious* Presentists (it would seem) should deny that RCK has any property indexed to any time before he began to exist. But this would make it impossible to provide an adequate interpretation of (3).

Second, Lewis's solution entails that the past changes with the generation or annihilation of things. When RCK began to exist, his existence added a new layer of fact to each past moment, namely, the fact that he did not exist then. Analogously, when RCK ceases to exist, these layers of fact are scrubbed away, leaving behind only the purely general fact that there once existed a person like RCK who didn't exist then. This seems to conflict with a pretty strong intuition of the absolute fixity of the past.

Presentists have two possible alternatives. The first embraces Anti-Actualism (12.1A), whether Possibilism (12.1A.1T) or Meinongianism (12.1A.1A); that is, it appeals to something like Possibilistic Presentism (20.2T.4.1). The second invokes surrogates of past things, like haecceities, bare particulars, or some kind of abstracta.

POSSIBILISTIC PRESENTISM Anti-Actualism gives Presentists non-existent things to which present things can stand in cross-temporal relations. All the while, one can maintain the claim that merely past or merely future things do not exist. The drawbacks are just those of Anti-Actualism generally. We discuss those in Chapter 12, and do not repeat that discussion here.

SURROGATE PRESENTISM Surrogate Presentism appeals to haecceities, bare particulars, or some kind of abstracta as surrogates for wholly past or future objects. The idea is that present things stand in cross-temporal relations to past or future things by standing in present relations to their surrogates. There are at least three versions of surrogate Presentism.

1 Zalta-Williamson mutationism. Merely past and future objects exist but only as abstracta (Zalta 1988, Williamson 1998). George Washington was once an abstract object, then he became concrete (a human being), and then upon death, he went back to being an abstract object (setting aside for simplicity's sake the possibility of an afterlife as a soul).
2 Bare particulars. Past and future objects exist, but they lack all qualities and relations, other than the minimal logical and metaphysical properties, such as *self-identity* and *particularity*. This may be equivalent to (1), depending on our conception of abstract objects.
3 Thisnesses or haecceities. The property of *being George Washington* may have existed before George Washington himself and may have continued to exist after he ceased to do so. If so, the fact that George Washington was a great President could consist in the fact that the property of *being George Washington* was once co-instantiated with the property of *being a great President*. Suppose we accept the existence of uninstantiated haecceities corresponding to no-longer or not-yet existing things. We can use haecceities to provide truthmakers for the non-existence or distinctness of things. Consider (4) and (5):

(4) James II no longer exists.
(5) James I ≠ James II.

(4) is made true by the fact that the haecceity of James II is uninstantiated but has the property of *having been instantiated*. (5), is made true by the fact that the haecceity of James I is possibly instantiated without being co-instantiated with the haecceity of James II, and vice versa. Similarly, we could suppose that both James I and James II exist eternally but are concrete or thick particulars only for limited periods. Contrast this with mutationism, according to which (5) receives its usual interpretation, and (4) is interpreted as the claim that James II is no longer concrete (alive, causally active, spatially located, having occurrent physical or mental properties).

Here is another example of a cross-time relation:

(6) Some American philosophers admire some ancient Greek philosophers.

The trouble is that there is no time at which both American and ancient Greek philosophers are both alive. The haecceity account would have to introduce a new kind of relation, that of *admiration**, which holds between someone and the haecceity of someone whom the first person admires. (6) becomes (6*):

(6*) Some American philosophers admire* some haecceities that were in ancient times co-instantiated with the properties of *being Greek* and of *being a philosopher*.

### 21.2.8 A final truthmaker problem for Presentism: Open-ended generalizations

Presentists have at least one more difficult phenomenon to deal with, namely, universal quantification over all of time involving accidental generalization. Consider (7):

(7) There will never be a golden mountain.

(7) seems possibly true. But it cannot be true at any time. Of necessity, there never will come a time at which (7) has a truthmaker. This is because there will necessarily always be still later times lying in the future, and so there may always come a time when there are golden mountains. (We assume that the future existence of a golden mountain is never ruled out by the laws of nature. If this assumption is false, merely modify the example, or consider an indeterministic world in which (7) is true but in which a golden mountain is never nomologically excluded.) This seems to put considerable pressure on the combination of Presentism and Truthmaker Theory. However, most Truthmaker Theorists reject the strong position of Truthmaker Maximalism (2.1T.1), which requires every truth, even negative ones, to have a truthmaker. (7) is clearly a negative statement, so it may be enough for its truth that it never has a falsity-maker.

Still, there's something paradoxical about saying that (7) is possibly true. At least, it seems paradoxical for Presentists to say this, since they identify *true at the present* with *true simpliciter*. How can (7) be possibly true if it is impossible for it ever to be true simpliciter?

Perhaps Presentists can add a real omega-moment, at the infinite horizon of time. This would have to be a moment (or quasi-moment) without successor or immediate predecessor. Such generalizations as (7) will have truthmakers then (namely, the totality of all time-indexed facts that will ever obtain). The greatest problem with such a proposal is the necessary lack of moments later than the omega-moment. Why and how would time freeze then? The answer might lie in the fact that the omega-moment consists entirely of past-tensed facts, all of which are intrinsically static.

## 21.3 The Theory of Relativity

The standard interpretation of Einstein's special theory of relativity makes trouble for Tensism because it challenges the idea that there is a unique present moment. Tensism, in all of its forms, is committed to the claim that there is a unique present moment. But according to the standard interpretation, physical reality contains nothing that corresponds to a relation of *absolute simultaneity*. There is no fact of the matter as to whether two space-like separated events (events such that no light signal or other direct causal chain could possibly connect one to the other) are or are not simultaneous in an absolute sense. Questions of simultaneity between such events have real answers only relative to an *inertial frame of reference*. That is, whether one such event is earlier than, later than or simultaneous with another varies, depending on one's point of view, where each velocity-vector corresponds to a different point of view or reference frame.

The conflict between Tensism and special relativity is pretty straightforward, since the standard interpretation of special relativity deprives us of a unique present moment. Two events belong to the same moment of time just in case they are simultaneous. Hence, according to special relativity (by which we mean the standard interpretation of special relativity), whether two events do or do not belong to the same moment is relative to one's reference frame. If there is no metaphysical fact of the matter as to whether any set of events constitutes a single moment at all, there could hardly be a metaphysical fact of the matter that one such moment is uniquely present. As simple as this argument is, it seems decisive.

Indeed, if we combine the standard interpretation of relativity theory with Presentism, we would be forced to embrace the uncomfortable (maybe absurd) conclusion that existence itself is frame-dependent. Presentism, after all, is the view that all that there is exists at the present moment.

If there is an incompatibility between special relativity and Tensism, then Tensers are forced to reject either special relativity itself or at least the standard interpretation of the theory. Sider (2001) considers either of these options to be equally revisionary of scientific theory and practice. But this is not obviously true. There is clearly a gap between the claim that special relativity requires no relation of *absolute simultaneity* and the claim that it excludes the existence of such a relation. Put another way, there is a gap between the fact that it is impossible to identify any privileged frame of reference and the hypothesis that there is no such privileged frame. Only if we assume the sort of verificationism that was (due to the physicist Ernst Mach) quite common among physicists when Einstein formulated his theory do these gaps disappear. Since verificationism is now almost universally thought to be indefensible, it behooves us to reconsider the matter carefully.

Both Michael Tooley (1997: 340–370) and William Lane Craig (2008) have argued for a non-standard interpretation of special relativity that is compatible with the existence of *absolute simultaneity*. Tooley's approach, in particular, seems defensible. To get a privileged frame of reference by which one can measure absolute simultaneity, one needs something like absolute space. (This requires Spatial Substantivalism 17.1T.) It is hard to see how one can get absolute space, especially assuming Presentism, without supposing that space is a kind of substance. That is, one must reify the points (or regions) of space. One can then assume either that points of space endure eternally throughout time or that, as Tooley assumes, each point in spacetime causally generates a unique spatial successor at the next moment of time. (It's a good question whether Tooley's model requires discrete time. This will depend on how Tensers handle diachronic causation, a matter we take up in Chapter 28.)

If absolute space exists, then moments can be defined as those sets of events that are simultaneous relative to a privileged frame, namely, the frame that is at rest with respect to absolute space. On such an interpretation, special relativity poses no problems for Tensism.

Tooley describes how to reach the non-standard interpretation in *Time, Tense and Causation* (Tooley 1997: 341ff.). (What follows assumes some knowledge of relativity theory.) First, we replace the Lorentz transformations with a more general scheme of *e*-Lorentz transformations, which were first described by Hans Reichenbach and used by John Winnie in his formulation of special relativity (Winnie 1970). In this formulation, we don't assume (as Einstein did) that the speed of light is the same in all directions in all frames of reference, but only that the average round-trip speed of light is the same in all directions in all frames of reference. We allow that the one-way speed of light in one direction might differ (in some frames) from its speed in the opposite direction. The *e* parameter, which takes values between 0 and 1, allows us to express a weaker assumption: if the speed of light in one direction is $c/2e$, then its speed in the opposite direction must be $c/(2\text{-}2e)$. If *e* is always assumed to be $^1\!/_2$, which was Einstein's assumption or convention, then the one-way direction of time in any direction and in any frame is always *c*. However, in the *e*-Lorentz transformation, we make no such assumption. Second, we add the assumption that absolute space exists, which gives a privileged frame of reference. The result is consistent with the experimental results of special relativity, since it can be proved that the *e*-Lorentz transformations guarantee that the round-trip average velocity of light is always *c* in every frame of reference.

It's not clear that this involves a substantial revision of current scientific theory, as opposed to an alternative interpretation of it. It clearly involves metaphysical assumptions that differ from the usual interpretation, like the existence of substantival space, but this does not require a change in scientific theory or practice.

Tooley argues that his account has several advantages. First, it can account for the continued existence of spatial relations by way of a causal mechanism, since on this view each momentary spatial point causes the existence of its location-successor. This is a fact that Spatial Relationists (17.1A) have to treat as a brute fact. This putative advantage is hard to evaluate. One might argue that the *spatial relatedness* of the new points generated at each moment remains an unexplained brute fact on Tooley's account, just as the continued *spatial relatedness* of things is unexplained on the standard account. Tooley's account explains the continued existence of the substance of space, but this involves an explanandum that simply doesn't exist on Relationism.

The second putative advantage Tooley cites involves the compatibility of his account with wave-collapse interpretations of quantum mechanics. On wave-collapse interpretations, a measurement somehow precipitates the collapse of a quantum wave into a set of discrete *eigenvalues*. If the quantum system involves mutually "entangled" particles that are, at the time of measurement, space-like separated, then the measurement of one particle has an instantaneous and therefore superluminal influence on the other particle. Such collapses can, from a God's-eye point of view, be used to synchronize distant events. This requires frame-independent, absolute simultaneity relations.

The question of how to interpret quantum mechanics and how, if at all, to make relativity and quantum mechanics mutually compatible, is a thorny and deeply contested ones. There are no-collapse versions of quantum mechanics, such as the Everett many-worlds or Albers many-minds interpretations, that may not require superluminal influences. In addition, the minimalist reinterpretation of special relativity described below shares this advantage with Tooley's postulation of absolute space.

The strongest argument against Tooley's version of special relativity is the conspiracy-of-silence objection raised by Elie Zahar in an exchange with J.L. Mackie (Zahar 1983). Zahar argues that the non-standard interpretation requires an unlikely coincidence, namely, that the laws of nature are fine-tuned to ensure that the actually privileged frame of reference is empirically undetectable. This sort of claim introduces some thorny issues, in particular, the question of whether it makes sense to assign some sort of prior probability measure to alternative laws of nature, a measure that can be used to assign a very low-probability to conspiracy-of-silence coincidences. On some views, the laws of nature are metaphysically necessary, in which case it can be difficult to assign the actual laws a probability other than 1. However, the appeal to coincidence is a compelling one, as it is in the case of anthropic coincidences, so this is merely a challenge to the probability theorist to come up with a model of the cogency of the reasoning.

A better response is the one offered by Tooley, who argues that it is special relativity itself, and not the non-standard interpretation, that is the true source of the conspiracy-of-silence coincidences. On the standard interpretation of special relativity, there is no way to measure the one-way velocity of light. It is merely a stipulation, and not an observable result, that the speed of light in one direction is the same as its speed in the opposite direction. Thus, the Lorentz transformations themselves seem to support a kind of conspiracy, even without the assumption of a privileged frame.

Our knowledge of physics isn't sufficient to adjudicate this dispute. We're inclined to guess that the undetectability of any privileged frame does count against the existence of such a privileged frame, even without relying on any dubious kind of verificationism.

In addition, there is another problem with the view that there is an empirically undetectable privileged frame, one pointed out (in conversation) by Brian Cutter. If there is a metaphysically privileged frame of reference, we are probably moving relative to it. A body's true shape would be its shape and size relative to the privileged frame. A high velocity relative to that frame would cause a contraction of length along the axis of absolute motion, according to the Lorentzian transformations. So, if there is a hidden frame and we are moving at a significant velocity in relation to that frame, then we are systematically misperceiving the true shapes of things. This argument assumes, plausibly, that shape is something intrinsic to bodies, and not a relation to frames of reference, together with the Reliable Perception Presumption (**PEpist 4**).

There is another tack Tensers can take: a minimalist accommodation of special relativity to the demands of Tensism. Tensers might simply stipulate that, in addition to the frame-relative relations of *physical simultaneity*, there is a frame-independent relation of *metaphysical simultaneity* representing those events that belong or have belonged to a common present. The metaphysical relation partitions spacetime into a set of three-dimensional hypersurfaces; each represents a single moment of real, metaphysical time. *Physical simultaneity* relations do impose a constraint on this partition: each hypersurface must consist of events that are pairwise space-like separated. That is, for each pair, there must be some frame relative to which they are physically simultaneous. However, we need not assume that this gives a privileged frame of reference. It may be that none of the hypersurfaces consists of events that are all physically simultaneous relative to a single, privileged frame of reference. (Note the shift in quantifier scope: for every pair, there is a frame relative to which they are physically simultaneous, but there may be no frame relative to which all such pairs are physically simultaneous.)

The minimalist view avoids the conspiracy-of-silence objection, since it does not privilege any particular frame. It is ontologically more economical, since it posits only a new binary relation and not the existence of a substantival space. It remains true that the relation of *metaphysical simultaneity* is empirically transcendent. There is no way to verify empirically that two space-like separated events are or are not metaphysically simultaneous. But this isn't surprising, since unless we're verificationists, there is no reason to suppose a priori that every fact must be empirically discoverable. In fact, there's every reason to suppose that a great many facts are empirically inaccessible to us.

Before leaving the topic, let's consider briefly the possibility of embracing both Tensism and the standard interpretation of relativity (without any kind of absolute simultaneity). Storrs McCall (1976) recommends doing so. On McCall's view, there is a unique, metaphysically privileged present moment, but which facts belong to that unique moment is a matter relative to one's frame of reference at that time. Let's suppose that our current frame of reference makes a particular event $E$ on Alpha Centauri present. Event $E_1$ occurs on the worldline of a particle $p$ that is moving rapidly relative to our frame of reference. Let's call our frame $F_1$ and the frame of $p$ $F_2$. Relative to $F_2$, some event on $E_2$ Earth in 2009 is part of the present moment. So, whether reality includes $E_2$ is relative to one's frame of reference: it is real relative to $F_2$ and unreal relative to $F_1$. This position seems dialectically unstable. If we are willing to embrace such radical perspectivalism, why not embrace a view according to which which moment is present is also a matter of one's perspective? That is, why not simply embrace the Anti-Tensism?

This is bad enough, but things get worse. Suppose that we now accelerate from $F_1$ to $F_2$. Since $E_2$ lies in our absolute future, it is impossible for it to become part of our present. Instead, what happens is that our present shifts in such a way that it no longer includes $E$ in Alpha Centauri, instead including events there that occurred years before $E_1$. In other words, by accelerating, we have changed the status of $E_1$ from real to unreal. (McCall admits this on p. 352.) On McCall's model, branches that had fallen off are grafted back on again (or more precisely, the substance of the intact branches shifts in such a way that events that had fallen off are pushed back onto extant branches). This strikes us as crazy and as incompatible with the motivation behind Tensism.

## 21.4 Epistemological Problems for Tensism

We turn now to three epistemological problems for Tensism.

### 21.4.1 Does anybody really know what time it is?

The first epistemological objection is easiest to appreciate for Growing Block Tensism. Growing Block Tensers maintain that the past and present are real, while the future is not. The epistemological problem is that this view seems incompatible with knowing that one is experiencing the present moment. Past moments, and hence past experiences and thoughts, are just as real on this view as the present. How can one be sure that one is not living in the past, possibly billions of years in the past, at a great temporal distance from the real present moment? All of one's experiences are just as real and just as vivid, even if one were living entirely in the distant past. There are far more people living in the past than on the cutting edge of reality. Hence, we are all far more likely to be deluded than to be well informed about our temporal location. A similar problem faces Simple Tensism, with its moving spotlight of presentness. What does presentness look or feel like? How do I know that this moment, the moment of this very experience or thought, really has this mysterious property of *presentness*?

This objection doesn't work against Presentism, since there are no past people to be deluded on that view. If one knows that one exists at all, one knows that one exists in the present.

The situation is not so clear with respect to the Modal versions of Tensism, such as Falling Branches Tensism. We might suppose that we have some direct knowledge, through our awareness of our own free will, for example, that times immediately after this time are open, and that all earlier times are fixed. If so, this knowledge would provide us with the information we need to locate ourselves at the juncture of the present.

### 21.4.2 Everything we see is unreal

A second objection is directed specifically at Presentism. If Presentism is true, then we never perceive a real event, and many of the objects we see are non-existent, since sensory perception takes time. What I perceive now is in fact the state of the world a few milliseconds in the past. The events and many of the entities (photons and so on) that existed then no longer exist now. Hence, much if not most of what we perceive is a kind of illusion. This would seem to undermine the reliability and trustworthiness of our senses, since they present their objects as really existing.

### 21.4.3 The problem of induction

As Alexander Pruss (2010) has pointed out, the problem of using past and present evidence to predict the future is more acute for Tensers, since the future represents a fundamentally different domain of reality. We know of many cases of sulfur samples burning with yellow flame, but all of these instances are located in either the past or the present. If the future is fundamentally different in nature, what grounds do we have for thinking that samples of sulfur will burn yellow in the future? Consider two hypotheses:

(8) Every sample of sulfur has burned yellow in the past.
(9) Every sample of sulfur has burned yellow in the past, is burning yellow now, and will be burning yellow in the future.

For Tensers, (9) is more complex than (8). Hence, we should prefer (8), but that's obviously a crazy result since we want induction to extend into the future. This problem largely disappears for Anti-Tensers, since they believe the past, present, and future are intrinsically indistinguishable, differing only in their relations to a present observer.

## 21.5 McTaggart's Paradox

Here is an attempt to reconstruct McTaggart's famous (or infamous) argument for the self-contradictory nature of the A series:

1 The A series is real. (Hypothesis for *reductio ad absurdum*)
2 If the A series is real, then there are three temporal statuses: *past*, *present*, and *future*.

3 If these three statuses exist, then it is necessarily the case that every event has* all three statuses. [What sort of 'having' is involved here is the crucial question.]
4 It is impossible to have* all three statuses, unless the compossibility of doing so can be explained by and so is dependent on the fact that it is possible to have* the three statuses at different times.
5 The fact that it is possible to have* the three statuses at different times is identical to the fact that it is possible to have* each of the following three meta-statuses:
   (i) having* the status of *being past in the present and future* (*and very near past*), having* the status of *being present in the more distant past*, and having* the status of *being future in the still more distant past*,
   (ii) having* the status of *being past in the future*, having* the status of *being present in the present*, and having* the status of *being future in the past*,
   (iii) having* the status of *being future in the past and present* (*and very near future*), having* the status of *being present in the more distant future*, and having* the status of *being past in the still more distant future*.
6 The three meta-statuses are identical to the three statuses, i.e., having* (i) = having* the status of *being past*, having* (ii) = having* the status of *being present*, and having* (iii) = having* the status of *being future*.
7 Hence, if the three statuses exist, then it is necessarily the case that every event has* all three meta-statuses. (From 3, 6, and Leibniz's law)
8 It is impossible to have* all three meta-statuses, unless the compossibility of doing so can be explained by the fact that it is possible to have the three statuses at different times. (From 4, 6, and Leibniz's law)
9 *Explanation* is a transitive and irreflexive relation of facts.
10 If it is possible to have* all three meta-statuses, then the fact that it is compossible to have* all three meta-statuses is explained by the fact that it is compossible to have* all three meta-statuses. (From 8, 5, and Leibniz's law)
11 The fact that it is compossible to have* all three meta-statuses is not explained by the fact that it is compossible to have* all three meta-statuses. (From 9)
12 It is not possible to have* all three meta-statuses. (From 10 and 11)
13 It is not possible to have* all three statuses. (From 12, 6, and Leibniz's law)
14 The A series is not real. (From 13, 2, and 3)

From the point of view of Tensers, the weak points in the argument are steps 3 and 4. On Tensism, there is nothing problematic about one moment of time *having* each of the three statuses, once we distinguish *having* from *having had* and from *going to have*. The use of past and future tense insulates the predication of the three statuses from contradicting one another, and no further explanation in terms of meta-statuses is needed.

McTaggart would respond by claiming that he cannot understand how tense performs this logical magic, except in terms of the three meta-statuses introduced in step 5, and that this explanation ultimately fails, as the rest of the argument demonstrates. We leave it to the reader to discern whether he or she can understand what McTaggart could not.

## 21.6 Brute Necessities of Time

However the truthmakers of past-tensed and future-tensed truths are supposed to be, there are certain necessary connections among times that need to be explained. For example, consider the following:

(10) If *p*, then it once was the case that it would be (or at least, might be) the case that *p*. (If *p*, then PF*p*.)
(11) If *p*, then it will be the case that it once was the case that *p*. (Symbolically: if *p*, then FP*p*.)
(12) If it will be the case that it will be the case that *p*, then it will be the case that *p*. (If FF*p*, then F*p*.)
(13) If it was the case that it was the case that *p*, then it was the case that *p*. (If PP*p*, then P*p*.)

Anti-Tensers have relatively little difficulty in explaining these. They have to suppose that spacetime is so structured that events can be put in linear *time-like* relations of *before* and *after*. If so, to say that *p* entails FP*p* simply means that there is a time index later than the one by which one's assertion is being evaluated: from the perspective of that later index, the current index is in the past. Similarly, to say that *p* entails PF*p* is simply to say that there is an index earlier than the one used to evaluate the assertion. The necessity of (12) and (13) can be explained so long as *earlier-than* and *later-than* are transitive.

Anti-Tensers will need some explanation of the direction of time. What makes some events later than others, rather than earlier? Why are causal relations lined up, at least for the most part, with time, in such a way that causes are typically earlier than. or at least no later than, their effects? We will take up some of these questions again in Chapter 27.

Explaining these necessities seems to be something of a challenge for Tensers, however. Let's take them in reverse order. Tensers might be able to argue that (12) and (13) are merely verbal necessities, and so are true by definition. Perhaps the fundamental properties of events in terms of which we define past and future are not such as to make (12) and (13) true, but we make use of defined notions of *pastness* and *futurity* such that (12) and (13) are trivially true. In other words, if we define 'future' in terms of being future* or future* in the future*, or future* in the future* in the future*, and so on (where '*future**' is the metaphysically primitive property), then (12) will come out as true by definition.

Let's turn then to (10) and (11). First of all, it is not obvious that these propositions are universally and necessarily true. Suppose, for example, that time had a beginning (something for which we found some support in Chapter 19). In the first moment of time, (10) would be false, since it would have been false then that anything had already been the case. Similarly, if there should be a last moment of time, (11) would be false. Let's take care of this possibility by introducing the sentences 'PT' and 'FT', where 'T' is supposed to represent some tautology (like '0=0'). Let's suppose that 'PT' is true just in case there has been some past time, and 'FT' is true just in case there will be future times. Consider (10*) and (11*):

(10*) If *p* and PT, then PF*p*.
(11*) If *p* and FT, then FP*p*.

(10*) and (11*) would be true, even in a first or last moment of time. Tensers must provide some account of the necessity of the truth of (10*) and (11*), even if just the minimal account that posits a set of brute necessities corresponding to each of the two. In fact, (10*) and (11*) don't give the whole story. There are also converse implications to consider. For example, if it was the case *n* units of time ago that it would be the case *n* units of time in the future that *p*, then *p* would be the case now. Let's look, then, at the following two necessities instead:

(14) If $P_n T$, then: *p* if and only if $P_n F_n p$.
(15) If $F_n T$, then: *p* if and only if $F_n P_n p$.

However, there is one more complication. If the future is open, then '$F_n p$' should be interpreted as saying something like 'It *might* be the case that *p*, *n* units of time hence.' If we do interpret '$F_n p$' that way, then (14) will come out false, since $P_n T$ and $P_n F_n p$ do not entail that *p*. It could be that it was *n* units of time ago the case that *p* might be the case *n* units hence, even though *p* did not in fact happen. All that's needed is for *p* to have been an open possibility *n* units ago. Thus, we will have to replace 'F' with something stronger, an operator that means that something must happen in the future. Let's use '$G_n p$' to represent 'It is definitely going to be the case that *p n* units of time hence.' We can define 'G' in terms of 'F':

**Definition of G.** $G_n p$ if and only if $F_n T$ and $\sim F_n \sim p$.

We can now replace (14) and (15) with our final pair of necessities:

(16) If $P_n T$, then: (i) if *p*, then $P_n F_n p$, and (ii) if $P_n G_n p$, then *p*.
(17) If $F_n T$, then: (i) if *p*, then $G_n P_n p$, and (ii) if $F_n P_n p$, then *p*.

We will assume that we can limit our attention to that case in which '*p*' corresponds to some simple, atomic (positive) proposition. If we can explain the universal and necessary truth of (16) and (17) in that case, it seems likely that the other cases (negations, disjunctions, quantificational propositions) can also be handled.

The necessity of (16) could be grounded in the temporal structure of the process or processes that link the present time with the past of *n* units ago. We have to assume that Tensers embrace an Aristotelian Modality (14.2A.5), and we have to assume that it is somehow necessary that only possible events occur. That is, if the truthmaker of *p* is part of a process *P*, it must be the case that *P* had *p*'s truth among its possibilities from the very beginning. That is, if *P* was initiated *m* units of time ago, *P* must have included a truthmaker of $F_m p$ at that time. All that we need is for the process also to have contained $F_m P_n F_n p$ at that time. Since we have $F_m p$, we will also have $F_{(m-n)} F_n p$ trivially. If we can appeal to the universal necessity of (17), we can replace '$F_n p$' with '$G_n P_n F_n p$', resulting in $F_{(m-n)} G_n P_n F_n p$. We can now collapse the '$F_{(m-n)} G_n$' back to '$F_m$', resulting in $F_m P_n F_n p$. Thus, the fact that *p* and $P_n F_n p$ would be simultaneously the case was already in the cards of process *P* from the very beginning.

If this is on the right track, then the necessity of (16) can be explained by combining Aristotelian Modality with the necessity of (17), on the assumption that possibility satisfies Actuality Entails Possibility:

**Actuality Entails Possibility.** Necessarily, if it is (now) the case that *p*, then it has always been possible that *p* be true now.

In addition, the necessity of (16) could be explained by reference to the necessity of causation. If every event in time must be preceded by a cause, Tensers could hypothesize that the cause of an event consists in some prior process with the power of producing that event. If the present event corresponds to *p*, such a prior process would have to have included a truthmaker for F*p*, thus explaining PF*p* as a necessary consequence of *p*.

Let's turn then to the necessity of (17). Here we think Tensers must appeal to some fundamental, brute necessity. The passage of time just consists in the fact that actual, present nexuses successively ossify into higher-order nexuses involving some *pastness* trope. This is certainly a costly assumption about necessity (**PMeth 1.2**). Whether or not it is a decisive objection depends on whether Anti-Tensers can offer a simpler account of the flow of time.

## 21.7 Conclusion

Tensers have reasonably good responses to several of the standard objections lodged against them by Anti-Tensers, including the problem about the rate of the flow of time and McTaggart's paradox. Truthmaker objections are a serious challenge to Presentism and to Growing Block Tensism, but they pose no challenge to other versions of Tensism. In addition, Presentists have a variety of plausible ways of meeting this challenge, including past-tensed properties and tensed copulas.

There are a limited number of brute necessities that are inseparably part of the Tensers' account of the flow of time. Whether Anti-Tensism offers a simpler account is a matter that we will take up in Chapters 27 and 28, where we consider the problem of causal direction.

Special relativity offers the most straightforward objection to Tensism, at least as special relativity is usually interpreted by physicists and philosophers of physics. Tensers must suppose that there are metaphysical facts about absolute simultaneity that are empirically undetectable. This is a real cost. How significant it is depends on one's view of the weight of scientific practice as a criterion for metaphysical theory.

# Part VII

# Unity

# 22

# Material Composition: The Special Question

In Part VII, composed of Chapters 22–25, we examine the problem of unity. In particular, we consider how it is possible for one thing to exist in and through a plurality of parts or phases. The unity problem has two dimensions, one spatial and the other temporal. In the spatial case, one worries about how a plurality of things can compose a single whole at the same time. In the temporal case, the problem concerns how a series of intrinsically variegated phases can be stages in the existence of a single, persisting thing. The spatial dimension occupies us in this chapter and the next, and the temporal dimension occupies us in Chapters 24 and 25. In fact, Chapter 25 combines both dimensions of the problem of unity, since we consider there the persistence through time of spatially composite things.

This chapter begins with a general discussion of the existence of composite things (Section 22.1). Then we consider (in Section 22.2) the view that composite entities are always an "ontological free lunch", things that can be freely posited without incurring any cost in relation to ontological economy or Ockham's Razor. Next, we look at the issue of causal redundancy (Section 22.3), a consideration which suggests that positing composite entities as fundamental things is in fact quite costly, something to be undertaken only with strong reasons.

In the next three sections we examine three candidates for fundamental composite entities: heaps (Section 22.4), artifacts (Section 22.5), and organisms (Section 22.6). We finally turn to the question of which sorts of composite things exist at all, whether fundamentally or derivatively. This involves a search for an intelligible principle of composition (Section 22.7).

In the following chapter, Chapter 23, we will investigate the nature of the part-whole relation itself.

*The Atlas of Reality: A Comprehensive Guide to Metaphysics*, First Edition.
Robert C. Koons and Timothy H. Pickavance.

## 22.1 The Existence of Composite Things

Much of this chapter will be concerned with three questions. First, do any composite things (things having parts) exist at all, and, second, if so, which ones? These two questions should be distinguished from a third question: are any composite things metaphysically *fundamental* (as in, G-fundamental Def D3.5)? A fundamental entity is one whose existence and intrinsic character are not grounded in other things. We look at this third question Sections 2.3 through Section 2.6, returning to the first two in Section 2.7.

In *Material Beings* (van Inwagen 1990a), Peter van Inwagen introduced the distinction between the general and special composition questions:

**General Composition Question:** what is it for one thing to be a part of another?

**Special Composition Question:** when do some things compose a further thing?

We take up the General Composition Question in Chapter 23. This chapter focuses on the Special Composition Question. There are two versions of the Special Composition Question, one Narrow and one Broad. The Broad version asks, when do some things compose anything at all? The Narrow version asks, when do some things compose a fundamental thing?[1]

We consider three positive answers to the Narrow version: (1) things compose something fundamental when they are spatially connected or continuous, (2) things compose something fundamental when they are parts of an artifact, and (3) things compose something fundamental when they compose a living organism. And we consider a negative answer, Priority Atomism, according to which nothing ever composes anything fundamental. Before arriving at these answers, however, we consider two issues about composite things, namely, whether they are an "ontological free lunch" (Section 22.2), and whether they are redundant vis-à-vis their parts (Section 22.3).

## 22.2 Are Composite Things an "Ontological Free Lunch"?

David Armstrong (1997) argued that we should treat composite entities as an "ontological free lunch". In other words, the number and variety of composite entities that are entailed by a theory shouldn't count against its simplicity. We shouldn't apply Ockham's Razor (**PMeth 1**), which directs us to prefer the more economical or parsimonious theory, on the basis of a theory's commitment to composite entities. Only atomic entities, entities without parts (or 'proper parts', as modern mereologists prefer to put it), should be counted.

There are two possible reasons for embracing Armstrong's free lunch principle: (1) facts about wholes and facts about their parts are really the very same facts: differences between truths at different mereological levels are metaphysically equivalent, and only conceptually different, or (2) facts about wholes are always reducible to (wholly grounded in) facts about their parts. The first possibility requires something very like the theory of *composition as identity*, proposed by Donald Baxter (1988), according to which wholes just *are* their parts, taken collectively. Thus, if we have some things, then to say that something exists that contains them as parts is just another way of saying that those same things exist. We examine Composition as Identity (23.1T.1.1) in more detail in Chapter 23.

This first option also requires that there be no metaphysical priority of parts over wholes or of wholes over parts. Something cannot be prior to itself, so if wholes just are their parts, then neither level can be prior to the other. Here is a breakdown of the relevant theses:

**22.1T Universal Compositional Priority.** Whenever some parts compose a whole, either the parts are wholly grounded in the whole, or the whole is wholly grounded in the parts.

**22.1T.1T Universal Bottom-Up Priority.** Whenever some parts compose a whole, the whole is wholly grounded in the parts.

**22.1A No Universal Compositional Priority.** There are cases of composition in which neither the whole is wholly grounded in the parts, nor are the parts wholly grounded in the whole.

**22.1A.1T No Compositional Priority.** In cases of composition, the whole is never wholly grounded in its parts, nor are the parts wholly grounded in the whole.

**22.1A.1T.1 Compositional Equivalence.** In cases of composition, truths about the whole and truths about the parts are metaphysically equivalent.

Given these theses, we can identify the two ways of defending the free lunch principle: Compositional Equivalence (22.1A.1T.1) and Universal Bottom-Up Priority (22.1T.1T). Compositional Equivalence forces us to clarify what could be meant by 'metaphysical equivalence'. We could take it as a new primitive, conveying the idea that two true propositions *say the same thing* about the world. If we adopt some form of Truthmaker Theory (2.1T/2.1A.1T), metaphysical equivalence can be having the same truthmakers in every possible situation. Finally, if we adopt some form of Real Grounding (3.1T), we could suppose that two propositions are metaphysically equivalent just in case they are both wholly grounded in the truth of some third proposition. We presently take up this last possibility.

If all truths concerning composite entities and their parts are wholly grounded in other truths, then the fundamental truths must describe a composition-free world. What could such a world be like? It would seem that there are just two possibilities. It could be a Nihilist (11.1A) world, or it could be a world in which all the fundamental truths are truths about simple entities. If the fundamental truths of the world include truths about things at all (whether one or many), and these fundamental truths are composition-free, then those fundamental things will count as mereological atoms (as simple entities), in which case those atoms will necessarily be metaphysically prior to all composite entities. In order for the fundamental truths to be composition-free, it must be the case that the fundamental atoms never combine into fundamental wholes. In fact, it seems that they could not (consistent with Compositional Equivalence) combine into wholes at all. If they did, wholes would be wholly grounded in their parts. So, on this view, either there are no fundamental entities at all, or the fundamental entities are all uncombinable atoms. In either case, the fundamental level of reality wholly grounds all the facts about composition and parthood. Thus, there are two versions of Compositional Equivalence:

**22.1A.1T.1.1 Nihilistically Grounded Composition.** All truths about wholes and their parts are wholly grounded in a class of nihilistic truths (truths that do not entail the existence of anything).

**22.1A.1T.1.2 Atomistically Grounded Composition.** All truths about wholes and their parts are wholly grounded in a class of truths about atomic things that never compose anything.

Nihilistically Grounded Composition can be readily combined with a certain metaphor or picture: that of reality as a mass of undifferentiated *dough* that we divide (by means of our conceptual *cookie-cutters*) into discrete things, some of which can be parts of others. On this view, the fundamental level of reality consists in this thing-free dough, and all things are derivative in status.

For now, let's turn to the second idea, that wholes are free lunches because their existence and their properties are *reducible* to those of their parts. The idea is that composite entities are nothing "over and above" their parts, as Armstrong puts it.

What do we mean by 'reducibility'? As we saw in Section 3.1.5, it is not enough that the existence and the properties of the whole supervene (Def D2.6) on those of the parts. Supervenience merely tells us that we can't have any change in the property of the whole without some corresponding change in the properties of the parts. However, supervenience is not asymmetric, and reduction must be (see Section 3.1.5).

Cases of reduction are cases in which the reduced facts are wholly grounded in the reducing facts, as we argued in Section 3.1.5. Reduced facts exist *by virtue of* the existence of their reducing facts. If wholes were reducible to their parts, then it would follow that only atomic or partless things are fundamental. This would make wholes, if not a free lunch, at least an ontological cheap lunch, since Ockham's Razor (**PMeth 1**) applies most forcefully at the level of fundamental things. When a theory entails the existence of additional reducible things without entailing the existence of additional fundamental things, we do not usually count this against the theory (or, at least, not very much; see Section 3.9).

What about those, like Quine, who reject metaphysical fundamentality? Let's assume that they also reject Composition as Identity. Such philosophers must deny Armstrong's free lunch principle.

There are, however, two necessary qualifications to this otherwise happy result. First, we might find some grounds for believing that composite things of kind $K$ do or do not exist *simpliciter*, grounds that do not depend on their supposed fundamentality. So, we will have to distinguish hard and soft anti-realists. Hard anti-realists about some kind of thing deny that such things exist at all, while moderate or reductive anti-realists about some kind merely deny that such things are fundamental. We must track this difference.

Second, some philosophers reject both fundamentality and Ockham's Razor. It is hard to find anyone who is *completely* indifferent to Ockham's Razor. Most agree that it is problematic if a theory posits the existence of fairies, round squares, phlogiston, ordinary matter that travels faster than light, and so on. However, one might adopt a more limited version of Ockham's Razor, one that exempts kinds of things whose existence is a matter of common sense. We might value theories that agree with common sense or with ordinary experience, without insisting that such things are fundamental in some metaphysically serious way.

## 22.3 Redundancy

We often think that composites are not fundamental. One reason for thinking so is that the whole does no further work, above and beyond the work done by its parts. To use the language of Chapters 4 through 6, all of the powers of the whole consist in the possession of fundamental powers by its parts. Imagine, for example, a block of stone sliding along the surface of an icy lake. The stone possesses a certain kind of causal power to move things, based in its total momentum and kinetic energy. The momentum and energy of the whole stone, however, is nothing but the sum of the momenta and energies of its constituent atoms. The stone as such adds nothing new. We could describe the result of the block's striking another block on the lake by referring only to the fundamental particles making up the two blocks. Their inertia and mutual interactions could fully explain the results we observe. There is no scientific reason to introduce either block as a further agent, in addition to their constituent parts.

This observation suggests a methodological principle: a theory should posit fundamental entities of a certain kind only if they are needed in giving a complete inventory of the world's causal powers. This principle is defended by Trenton Merricks (2003). Here's a first draft of such a principle:

**Redundancy.** Reject any theory that posits fundamental entities whose causal powers are redundant, given the other fundamental entities posited by that theory.

As we saw in Chapter 6, causal powers come in four varieties: active, passive, and immanent powers, and tendencies. Are all four kinds of power relevant to Redundancy? The one kind that seems questionable in this respect is that of passive powers. Suppose that a kind of entity *K* is wholly redundant with respect to its active and immanent powers and its tendencies, but which had passive powers that could not be accounted for in terms of the passive powers of its parts. One possible example of this would be the human soul as conceived of by the theory of *epiphenomenalism*. According to epiphenomenalism, the soul has no active or immanent power, and so no tendency. Everything that happens happens because of the action of one or more material bodies. However, the soul is affected by the state of the body in ways that cannot be accounted for in physical terms. The soul has experiences and feelings that are irreducible, on this account.

One of the difficulties with epiphenomenalism is that it is hard to see how we could know that we were having experiences in our soul if these soulish experiences had no active causal power over anything, including other states of the soul. How can we explain introspective belief and knowledge if the experiences we're introspecting are causally inert? This point might be resisted by positing a form of knowledge, namely, *direct* *acquaintance*, that doesn't require any causal efficacy on the part of the object. Logical and mathematical knowledge is often supposed to be of this kind, on the assumption that logical and mathematical entities are causally inert. (For a different view, see Koons 2000.)

In addition, if epiphenomenalism were true, we would have to take seriously the possibility that everyone else is a *zombie*—a body with no attached, experiencing soul—since zombies would behave no differently in the absence of a soul than would en-souled bodies. Finally, we all have direct experience of acting upon the world by moving our bodies in intentional ways. All of this would have to be an illusion for epiphenomenalism to be true.

The evident failure of epiphenomenalism illustrates the difficulties with positing entities solely on the basis of those entities' supposed passive powers. Consequently, we can plausibly strengthen Redundancy, restricting the permissible entities to those with non-redundant active and immanent powers. In addition, if an entity is wholly redundant with respect to its active and immanent powers, then its tendencies to use those powers would also seem to be redundant, consisting entirely in the tendencies of the constituent parts to use their powers.

Finally, suppose that an entity's active powers are redundant. Could it have non-redundant immanent powers? It is hard to see how it could, since it would have no active powers of its own to affect. So, its immanent powers would have to be powers only over other immanent powers, keeping the effectiveness of the entity forever trapped in a self-referential circle. Let's restrict the Redundancy principle, then, to active powers alone:

**Active Power Redundancy.** Reject any theory that posits fundamental entities whose fundamental active causal powers are redundant, given the other fundamental entities posited by that theory.

When do the active powers of the parts of a composite thing make the active powers of the whole redundant? What would it be for a whole to have non-redundant active powers? Some philosophers, going back to the British Emergentists of the early twentieth century, have called such powers 'emergent'. A whole has non-redundant active powers when it has emergent powers.

A whole has emergent active powers when it is able to do things that cannot be explained in terms of the powers of its parts. By saying that this cannot be explained, we do not mean merely that it cannot be explained *by us right now*, due to the complexity or other practical obstacles involved in formulating such an explanation. We mean something like that the emergent power cannot be explained in principle, not even by God, in terms of the powers of the parts. We are interested in *ontological* emergence, not merely some sort of human-centered epistemological emergence.

When a whole has emergent powers, actions take place that are not merely the sum of the individual actions of the component parts. The whole is greater in active power than the sum of the active powers of the parts.

Could we ever be in a position to know, or at least to believe with good reason, that some whole has emergent powers? One worry is that whenever we see a whole seeming to do something over and above what could be done by its parts, we could always describe what's going on as the exercise of *hidden* powers of the parts, powers that those parts never exercise except when they are put together in such a way as to constitute a whole of this kind. Since that always seems to be a live option, we would never be able to tell whether it is really the whole exercising some emergent power or just the parts jointly exercising some hidden powers.

We could escape from this impasse if we could find certain powers that are *essentially unitary*, powers that could only be exercised by some one, unified entity, and that could never consist in the joint possession of individual powers by the members of a plurality. Here are several possible candidates for essentially unitary powers:

1 The power of *self-reproduction*, as exercised by living organisms.
2 The power of *growth*, *self-development*, and *self-repair* through the assimilation of new material.
3 The power of *self-determination* as exercised in free, conscious choices.

Each of these powers makes reference to the self that exercises it. This self-reference is what seems to make each essentially unitary. The particles making up an organism's body cannot exercise the power of *self-reproduction*, since it is the organism and not any of the particles that is reproduced. Similarly, the particles do not grow, either individually or collectively. If the organism contained a billion billion particles before the episode of growth, then those billion billion particles are no larger or more massive afterward than they were before. It is the organism that grows, not the particles. Finally, the particles making up a person's body do not make up their collective mind to do one thing rather than another. The particles don't have minds to be made up—only the person does. Therefore, we will propose that a theory may reasonably posit the existence of a fundamental composite entity only if it credits that entity with powers that are essentially unitary. It seems that these essentially unitary powers will be either immanent or passive powers, not active ones. Anything that can be done to some other thing could, it seems, by done either by a single agent or by a plurality of agents acting jointly.

A passive power, in contrast, might be essentially unitary. The power of *sensation*, for example, seems essentially unitary. Being affected with a particular sense-quality is something that can happen only to a single thing because sensory consciousness is essentially unified. Immanent powers can also be essentially unitary, like the power of *self-reproduction* or *organic growth*. Thus, in order for a theory's postulation of composite entities to be justified, the composite entities must be assigned both emergent active powers and essentially unitary passive or immanent powers.

**Def D22.1 Emergent Power.** A power of a composite entity $x$ is *emergent* if and only if the power of $x$ is not wholly grounded in the sum of the causal powers of $x$'s parts, together with the intrinsic qualities and mutual relations of those parts.

**Def D22.2 Essentially Unitary Power.** A power is *essentially unitary* if and only if it is a fundamental power that could by its very nature be possessed only by a single entity, not collectively possessed by a plurality of entities.

**PMeth 4 Redundancy for Composite Entities.** Reject any theory that posits any kind of fundamental composite entity without *both* emergent active powers and essentially unitary passive or immanent powers.

Redundancy for Composite Entities presupposes that we can distinguish between fundamental and derived powers. This distinction is implied by Strong Powerism (4.4A.3), the view that some causal powers are fundamental properties.

A similar distinction could be made on the basis of either Strong Nomism (4.4A.2) or Strong Hypotheticalism (4.4A.1). Suppose, first, that Nomism were true. This would entail that there are fundamental truths concerning the existence of certain laws of nature, where each law of nature involves some necessary connection among two or more properties. Each law of nature could be thought of as grounding a set of relatively fundamental causal powers. Suppose, for example, we had a law requiring that *F*'s change themselves into *G*'s. This would correspond to a fundamental immanent power possessed by all *F*'s, namely, the power to become a *G*. Other powers, involving properties not linked together directly by laws, would count as derivative powers. For instance, if *F*'s have the power to become *G*'s and also the fundamental power to become *H*'s, then they would have the derived power to become *G* and *H*.

Similarly, Strong Hypotheticalism posits certain counterfactual conditionals as fundamentally true. Each fundamentally true counterfactual would correspond to a set of relatively fundamental powers. Counterfactual conditionals that are only derivatively true would correspond to derivative causal powers.

Suppose, however, that we were to embrace Neo-Humeism (4.4T). According to Neo-Humeism, no causal powers are metaphysically fundamental, and neither are any causal laws or counterfactual conditionals. The only fundamental truths involve the distribution of fully categorical qualities (non-powers) in space and time. Powers, laws, and conditionals are determined by whatever set of global regularities is simplest and most powerful, that is, by the fact that some scientific theory is the best fit to the world's basic facts. On such a theory, there would be no basis for a distinction between fundamental and derived causal powers. All powers are equally derived. There is no one-to-one correspondence between fundamental facts and relatively fundamental powers either, unlike with Nomism and Hypotheticalism.

Without a distinction between fundamental and derivative causal powers, the very idea of redundant powers makes no sense. Both simple entities and composite entities derive their powers from the overall pattern of facts throughout the world. On Neo-Humeism, there is no real redundancy in assigning powers with equal status both to a whole and to each of its parts, even if we could explain the same behavioral regularities in terms of the powers at either level alone.

Neo-Humeists should, we think, embrace the thesis of Compositional Equivalence. They should deny that parts are prior to their wholes, and that wholes are prior to their parts. Instead, the existence of both parts and wholes is wholly grounded in facts about the Humean mosaic, the distribution of qualities in spacetime. They must think of this mosaic either nihilistically (as a thing-free world) or atomistically (as consisting of simple, point-like entities that cannot be combined into composites). As we saw in Chapter 11, Nihilism runs up against some serious problems with Ockham's Razor, with the attainment of qualitative economy. Thus, the Neo-Humeists' best bet would seem to be one of treating the pixels in the mosaic as simple and uncombinable. The entities that can be either parts or wholes should be treated as derived entities, wholly grounded in the facts about the Mosaic.

There is one further wrinkle to mention. It might be possible for Powerists and Nomists to reject Redundancy for Composite Entities and to embrace instead the first form of the Ontological Free Lunch principle, that is, Compositional Equivalence (22.1A.1T.1). They can do this if they can claim, with plausibility, that composite entities have causal powers, but that those powers are *identical to* the powers of their parts. This would require, in turn, both that the composite entities instantiate power-conferring universals, and that the instantiation of such universals by a composite entity is nothing over and above the instantiation by its proper parts of that or other power-conferring universals. For this to make sense, we would have to make sense of one case of instantiation's being identical to a number of other cases of instantiation.

In order to investigate whether we can make sense of such reducible or non-fundamental cases of instantiation, we will have to break the investigation up into two parts, one relying on Relational Ontology (9.1A) and the other on Constituent Ontology (9.1T). Consider first the case of Relational Ontology, specifically Classical Relational Realism (9.1A.1T). Since Relational Realism takes *instantiation* itself to be a fundamental relation, it would seem to be impossible for Relational Realists to count some cases of instantiation as identical or equivalent to *other* cases of instantiation. There may be an exception, however, in the case of those Relational Realists who believe in nexuses, as entities that connect particulars to universals. If *instantiation* consists in the existence of a nexus, then we could suppose that some nexuses are composite, that is, that they have proper parts. If so, it could be that the existence of a nexus $N$ connecting a composite substance $C$ to a universal $U$ just is the existence of the proper parts of $N$, each of which connects some proper part of $C$ to $U$.

This can work only in cases in which a composite entity instantiates some quality that is perfectly homogeneous, in the sense that every part of the entity necessarily instantiates the same quality whenever the whole does. *Having a certain uniform density* or *uniform temperature* would be examples of such homogeneous qualities. However, this strategy of composite nexuses will not work in the case of non-homogeneous properties, like *size*, *shape*, or *total mass*. The proper parts of a cubical whole do not have to be cubical themselves, and the proper parts of a body with mass of one gram must have a mass that is less than one gram. In order to deal with such non-homogeneous properties, the Relational Realists will have to suppose that the corresponding universals are themselves composite. So, when a composite body $B$ is connected by nexus $N$ to some quantity $Q$ (like BEING ONE GRAM), then if $B$ is composed of two discrete parts, $B_1$ and $B_2$, then $N$ must also be composed of two parts $N_1$ and $N_2$, and the universal $Q$ must consist of two parts $Q_1$ and $Q_2$, such that $N_1$ connects $B_1$ to $Q_1$, and $N_2$ connects $B_2$ to $Q_2$.

However, this raises an obvious problem. How can a universal being BEING ONE GRAM be composed of two parts, each of which is identical to the BEING ONE-HALF GRAM universal? What would it mean for one universal to be composed of two "copies" of another universal? We encountered exactly this problem in Section 10.3, when we investigated structural universals. Our solution was a resort to amphibian theory. So, we could suppose that the universal BEING ONE GRAM consisting of two amphibians of the universal BEING ONE-HALF GRAM.

Will amphibian theory enable Realists to embrace Compositional Equivalence? The answer is unclear. There is a difference between instantiating a universal and instantiating an amphibian of that universal. Therefore, it does not seem that a nexus connecting a whole to some structural universal could be merely the sum of the nexuses connecting the parts to the amphibians of that structure. It seems that instantiating a universal would be a more fundamental fact than merely instantiating an amphibian, in which case the whole that instantiates a structural *universal* would be metaphysically prior to its parts (which merely instantiate amphibians of parts of that structure). This would seem to exclude Compositional Equivalence.

Now, let's turn to Constituent Ontologists who are also Realists (7.1T), whether Classical Bundle Theorists (9.1T.1T.1A) or Classical Substrate Theorists (9.1T.1A.1A). Let's consider homogeneous qualities, like color or density first. It seems that a Constituent Realist cannot embrace Compositional Equivalence, since the instantiation of a universal by a particular just is (according to Constituent Realism) that universal's being an immediate proper part of the particular: that is, its being a proper part, and not by virtue of its being a part of some other proper part. Consequently, whenever we find a universal as an immediate proper part of some particular, that particular must be metaphysically fundamental, and the fact that its proper parts also instantiate that universal must be wholly grounded in the instantiation of that universal by the whole.

## 22.4 Fundamental Heaps

We conjecture that the only sort of composite things that could have essentially unitary powers are living things, artifacts, and their functional parts (like organs and so on). Any other sort of thing is a *heap*.

**Def D22.3 Heap.** *x* is a *heap* if and only if *x* has actual parts and *x* is not a living thing, an artifact or a functional part of a living thing or an artifact.

One faces a choice with respect to the fundamental existence of heaps, and a further choice if one denies that there are fundamental heaps:

**22.2T Fundamental Heapism.** Some heaps exist fundamentally.

**22.2A Anti-Heapism.** No heaps exist fundamentally.

Examples of heaps abound. All positive natural formations are heaps. These include rocks, clouds, grains of sand, mountains, rivers, lakes, oceans, planets, stars, galaxies, crystals, and literal heaps. Some very exotic things, if they exist at all, would also be heaps. If, for example, there existed something consisting of THP's left big toe and the Eiffel Tower, that thing would be a heap. So would the thing consisting of all of the earth's water if there is such a thing. Arbitrary parts of things, if they exist, are also heaps. RCK's left side is a heap of organic matter, as it is neither itself a living thing nor is it a functional part of my body, like my heart or liver.

### 22.4.1 Arguments against fundamental heaps

Do heaps really exist fundamentally? There are three objections to Fundamental Heapism: the appeal to Redundancy (**PMeth 4**), the problem of identifying the fundamental occupiers of space, the appeal to Ockham's Razor (**PMeth 1**). In addition, we will consider in Section 2.7 an objection to the existence of heaps of any kind, whether fundamental or not—namely, the difficulty of finding a principle of composition that matches our pretheoretical convictions.

1. Appeal to redundancy. Redundancy for Composite Entities (**PMeth 4**) makes trouble for heaps. Heaps seem to be causally redundant. There is no reason to think that they have either emergent active powers or essentially unitary passive or immanent powers. Anything that a heap does or causes is really done or caused by its constituent parts acting in concert. For example, to use Merricks's (2003) example, when a rock breaks a window, it is really the congeries of particles arranged in a baseball-pattern that shatter the window or, to be more precise, that scatter the particles arranged window-wise. The only agency rocks have is that of physical matter: energy, momentum, angular momentum, temperature, and so on. All of these powers are simply the aggregation of the powers of the rock's constituent particles. The rock itself is causally superfluous.

Similarly, rocks and other heaps have no essentially unitary passive powers either. Everything that can be done to a heap can always be re-described, apparently without loss of information, as something that is done to its constituent parts. Heaps can be changed in shape, scattered or re-assembled, but each of these processes consist in nothing more than the movement of the particles. Heaps can be painted or made radioactive, but these changes also involve nothing over and above certain changes to the simplest parts. Heaps can grow or diminish in size, but only as a result of particles' being spread out or pressed together. If more particles are added to a heap, we could always say that, strictly speaking, we have a new heap, composed of a new set of particles.

But don't mountains, rivers, and crystals grow, just as organisms do? When a mountain grows, its constituent particles don't grow (either individually or collectively). It grows by adding particles, just as organisms do. However, there is a crucial difference. When a mountain "grows", this is merely the accidental by-product of more fundamental processes, like continental uplift. In contrast, it is at least possible that organic growth is a fundamental process in its own right, with the growth and development of the organism as its proper effect. Organic growth is much more than the mere accidental accumulation of additional matter.

We think the appeal to Redundancy is a powerful argument against the fundamental existence of most heaps, and so constitutes a serious challenge to Heapism.

However, some inorganic structures, including molecules, may be able to escape this argument from Redundancy. Quantum chemistry provides some indications of the emergent nature of whole molecules (Hendry 2006, 2010, Bishop 2005). Take for example, a water molecule, a molecule of $H_2O$. Such molecules consist of a quantum system with 18 protons and electrons and (typically) eight neutrons. The quantum equation for such a system of particles is spherically symmetrical in shape. However, we always find water molecules to be asymmetrically organized in a V-shape, with approximately a $106^o$ angle between the two chemical bonds, and the global behavior of water gives us reason to believe that this is so prior to observation (before, that is, the "collapse of the wave packet"). This asymmetric V-shape of the molecule appears to be an emergent property of the whole that has to be determined empirically. It cannot be deduced from the physical properties of the particles, even taken collectively. (At least, it's not deducible given our current knowledge of quantum chemistry.)

2. Occupation of space. The second objection to Fundamental Heapism concerns the problem of explaining how things occupy space. The key question is why is the location of a part of a thing is a part of the location of the whole thing. In fact, there are two facts that need explaining (in terms of metaphysical grounding):

**Spatial Occupation Fact 1.** If $x$ is a material part of $y$, then the location of $x$ is a part of the location of $y$.

**Spatial Occupation Fact 2.** If material body $y$ is a sum of the $x$'s, then the location of $y$ is a sum of the location of the $x$'s.

There are three possible explanations of these two facts:

1 They are metaphysically brute necessities or necessities imposed by the laws of nature.
2 The existence and location of the parts are fundamental, and the existence and location of the whole are wholly grounded by them. The location of the whole is always to be explained metaphysically in terms of the parts, in such a way as to make true the two Spatial Occupation Facts.
3 The existence and location of the whole are fundamental, and the existence and location of the parts are wholly grounded by them. When an extended material body has a region as its location, the body simultaneously has all of the parts of that region as partial locations. A material part of a fundamental body consists in the fact that the body has some region as a sub-location.

The first option is unattractive from the point of view of Ockham's Razor (**PMeth 1.2**), since it imposes a vast system of brute necessities between the locations and location-occupiers. Option 2 corresponds to the Priority of Spatial Proper Parts, and option 3 to the Non-Priority of Spatial Proper Parts:

**22.3T Priority of Spatial Point-Parts.** If $x$ is a composite thing, then the location of $x$ is wholly grounded in the point-locations of point-sized parts of $x$.

**22.3A Non-Priority of Spatial Point-Parts.** If $x$ is a composite thing, then the location of $x$ is *not* wholly grounded in the location of $x$'s point-sized parts.

There is something very natural about the Priority of Spatial Point-Parts. It fits well with Spatial Pointillism (18.1T), the view that spatial points are the fundamental sort of location while spatial regions are something like constructions from spatial points. However, there are some powerful arguments against Spatial Pointillism (which we saw in detail in Chapter 18), and we have F.H. Bradley's argument against the universal priority of parts over wholes (Section 11.2.3), so we should take seriously the alternative position, the Non-Priority of Spatial Point-Parts. If spatial point-sized parts are not prior, then it is the region-filling location of extended bodies that must be metaphysically fundamental (assuming that we don't want both points and regions to be fundamental locations). If extended regions are the fundamental locations, what explains the fact that one region is a part of another? Given Spatial Anti-Pointillism (18.1A), we cannot give the simple answer that one region is a part of another just in case its points are a subset of the points of the other. We must instead either embrace composition as identity or take *parthood* to be a fundamental relation among regions.

The Non-Priority of Spatial Point-Parts requires a brute necessity linking fundamental material wholes with appropriate locations and sub-locations, in such a way that the derivative *parthood* relation among regional locations satisfies the axioms of mereology. In addition, it seems to require a brute necessity connecting dependent parts of fundamental wholes with corresponding sub-locations of those wholes. This is a theoretical cost of the view. We could lessen the cost somewhat by making the parts of fundamental bodies into derivative entities, entities whose very existence consists simply in certain facts about the location and nature of the whole body. The following picture emerges. The fundamental entities are certain extended, composite bodies. These fundamental entities have *regional locations* and *sub-locations* among their fundamental properties. The so-called 'parts' of these bodies are simply identical to the bodies' sub-locations. Each proper part of the location of a fundamental body is a potential proper part of the whole body. Thus, there is no need to explain the proper parts' locations. They simply *are* the locations, under a different guise.

This whole picture seems implausible if we apply it to heaps. How could the location of any heap be metaphysically fundamental and the location of its parts merely derivative? The heap has no being or nature over and above the being and nature of its parts. The dependency seems to run clearly in the other direction. Proper parts and their locations are fundamental, and the locations of heaps are dependent on them. If so, then we must embrace the Priority of Spatial Point-Parts, at least insofar as it applies to heaps:

**22.3T.1 Priority of Spatial Point-Parts for Heaps.** If $x$ is a heap, then the location of $x$ is wholly grounded in the locations of some proper parts of $x$.

If Priority of Spatial Point-Parts for Heaps is true, then we have a good reason to think that heaps are not metaphysically fundamental. If the location of the heap is grounded in the point-locations of its proper parts, then presumably the same thing is true of all of the properties of heaps, including their intrinsic qualities and their material composition. If so, then the only fundamental entities are the proper parts, and the heaps are merely derived entities. If a heap is heap-like all the way down to its point-sized parts, in the sense that all of its proper parts that are not point-sized are themselves mere heaps, then it is those point-sized parts that are metaphysically fundamental. Let's call such a heap a 'thorough heap'.

**Def D22.4** $x$ is a *thorough heap* if and only if $x$ is a heap and all of $x$'s proper parts whose locations are larger than points are also heaps.

**22.3T.2 Priority of Spatial Point-Parts for Thorough Heaps.** If $x$ is a thorough heap, then the location of $x$ is wholly grounded in the locations of its point-sized parts.

This argument presupposes that all thorough heaps have point-sized parts. Suppose that *all* of the parts of a thorough heap were spatially extended. Then the Priority of Spatial Point-Parts would be obviously false, since there would be no point-sized parts to be the fundamental entities. The possibility that some thorough heaps might have no point-sized parts is closely related to the thesis that there is *gunk*, which we introduced in Chapter 11 and will take up again in Section 22.4.2 below, when we discuss the arguments for heaps.

Even if we dismiss the possibility of gunk, there is still one kind of thorough heap that can survive this argument: point-sized heaps. If a heap is point-sized, then its location has no proper parts. This means there is no need to explain the connection between proper parts of locations and proper parts of location-occupiers.

But can heaps be point-sized? How can a heap be composed of proper parts if it is the size of a point? Suppose that there are point-sized particles that can share exactly the same location (like photons). Several such particles could come to occupy exactly the same point, resulting in a point-sized heap.

3. Ontological economy. A theory of the world that denies the fundamental existence of heaps is simpler, both quantitatively and qualitatively, then one that includes heaps among its fundamental objects. Other things being equal, this gives us good reason to prefer Anti-Heapism, according to Ockham's Razor (**PMeth 1.4**).

UPSHOT OF OBJECTIONS TO HEAPS To sum up, there is just one kind of fundamental heap that can withstand scrutiny: point-sized bodies. This kind of heap could be exempt from the principle of Redundancy. Suppose, for instance, that several point-sized masses are fused together into a point-sized mass. In such a case, we might well suppose that the powers of the constituent masses themselves are fused together into a new set of powers, possessed only by the heap.

### 22.4.2 An argument in favor of fundamental heaps: The possibility of gunk

A more serious argument for Heapism relies on the possibility of gunk. If gunk is possible, then Heapists have a very powerful rebuttal to both the Redundancy argument and the spatial occupation argument. In addition, in a world containing gunk, heaps would have to exist, unless absolutely every bit of matter were filled with living organisms and artifacts. This possibility was embraced by Leibniz, but one which has seemed to nearly everyone else has thought it wildly unlikely. At any rate, we here explore the argument for Heapism from the possibility of gunk in some detail, and we therefore step back a bit to get clearer about what, exactly, gunk is.

'Gunk' is a term introduced by David Lewis (1991) for a hypothetical kind of stuff that has no atomic part. Each part of gunk can be further subdivided, without limit.

**Def D22.5 Mereological Gunk.** $x$ is *mereologically gunky* if and only if $x$ has no atomic parts.

What do we mean by 'atomic parts' or 'atoms'? In a metaphysical context, we don't mean a complex entity like a hydrogen or oxygen atom. Chemical atoms have proper parts, namely, protons, neutrons, and electrons. By an 'atom', we mean something without any proper parts at all:

**Def D22.6 Atom.** *x* is an *atom* if and only if *x* has no proper parts.

**22.4T Universal Atomism.** Everything has an atomic part.

**22.4A Existence of Mereological Gunk.** Some things have no atomic part.

It is important, however, to distinguish between mereological gunk and spatial gunk. To be spatial gunk, a material body must have no point-sized masses as parts: each material part must have a proper part that is strictly smaller in extension.

**Def D22.7 Spatial Gunk.** *x* is *spatially gunky* if and only if *x* is mereological gunky and *x* has no point-sized parts.[2]

**22.4A.1 Existence of Spatial Gunk.** Something is spatially gunky.

Although there is a clear distinction between mereological and spatial gunk, spatial gunk seems to be the most plausible kind of gunk. If mereological gunk is possible, then it seems that spatial gunk would also be possible.

Suppose that a spatially gunky heap is possible. Then we cannot apply Redundancy for Composite Entities (**PMeth 4**) to eliminate all heaps from our theory. The active powers of a heap cannot be redundant if the heap has no atomic parts. Without atomic parts, there would be no final resting place, no ultimate bedrock for the fundamental powers of the heap. If we were tempted to say that the powers of a heap are always grounded in the powers of its smaller parts, then we would have to say that none of the powers of the heap or of any of its parts would be fundamental. All of the powers would ultimately be grounded in nothing. This is impossible.

If a spatially gunky heap were possible, what would be the fundamental bearers of the heap's powers? If we accept Strong Powerism, we must find some bearer for those powers. It would hardly be acceptable to suppose that the heap's powers are infinitely over-determined, with equivalent powers being fundamentally possessed by parts of the heap at every possible scale. Thus, Strong Powerism gives us good grounds for supposing gunk to be impossible. Similar arguments against spatial gunk could be made on the basis of Strong Hypotheticalism or Strong Nomism. Conversely, if spatial gunk is possible, then we have good grounds for rejecting all of these theories and embracing in their place Neo-Humeism.

In addition, the possibility of gunk provides a powerful response to the spatial occupation argument. If it were possible for a heap to be gunky, it would seem to be possible for a gunky heap to have no point-sized parts at all. If the heap has no point-sized parts, then it is obvious that the location of the heap could not be grounded in the location of its point-sized parts. What would then ground the location of the heap, and what could explain the fact that the location of any part of the heap must be a part of the location of the whole? Heapists could suppose that in each case it is the location of the part that is grounded in the location of the whole. Perhaps the only truly fundamental location-fact is the fact about the location of the entire material universe. Parts of the material universe then would derive their location somehow from the location of the whole. This seems counterintuitive, but the possibility of gunk seems to force us to some such conclusion.

If there is gunk, it seems virtually certain that that gunk or some part of it will be neither a living organism nor an artifact. Hence, the existence of gunk is incompatible with Anti-Heapism.

So, a lot turns on whether gunk is possible. It doesn't matter whether there is actually any gunky body. The mere metaphysical possibility of gunk would be enough to wreak havoc on the appeals to Redundancy and to the explanation of spatial occupation. What arguments are there for gunk's possibility?

ARGUMENTS FOR SPATIAL GUNK The strongest argument is a simple appeal to conceivability. We can imagine (in some broad sense) the Existence of Spatial Gunk, and no obvious contradiction appears in our conception of it. We might even feel a strong mental "pull" in the direction of believing it to be possible. These are not absolutely conclusive arguments, but they have to be given their due weight, as we have embodied in Imagination as Guide to Possibility (**PEpist 1**). In fact, we have already claimed that the possibility of gunk is a principle of natural philosophy (**PNatPhil 1**) on this basis.

ARGUMENTS AGAINST SPATIAL GUNK Besides the Redundancy and spatial occupation arguments, the best argument against spatial gunk is one that we have already encountered in another context: Hawthorne and Weatherson's (2004) supercutting argument.

This supercutting argument is based on a thought experiment of José Benardete's (1964: 184). We are to imagine that a block of material substance is subjected to an infinitary process of supercutting: after 1/2 second, it is cut into two pieces, in the next 1/4 second into four pieces, and so on, for each power of 2 (8, 16, 32, etc.) After one second has passed, the block will have been divided infinitely many times. The infinitely many slices will lack any depth, and the process can be described in such a way that, at the end of one second, the resulting set of slices are such that between any two slices, there is some completely unoccupied space, without requiring that any matter move at any point faster than the speed of light (Hawthorne and Weatherson 2004: 343). The distribution of slices will be dense (between any two slices there will always be others) and disconnected (between any two slices, there will always be some finitely wide, totally unoccupied space).

The supercutting process could be repeated along the two other dimensions of space, first dividing each of the slices into an infinite number of one-dimensional splinters, and then dividing each splinter into an infinite number of point-masses, each with zero volume. If matter is gunky and every piece of gunk occupies some finite volume, then the process of supercutting, when completed, will have annihilated all of the matter contained in the original block. It seems impossible for a process of mere dividing, even if repeated infinitely often, to annihilate any of the material substance being divided. The possibility of this supercutting is incompatible with the Existence of Spatial Gunk.

What are some of the assumptions of the Supercutting argument? First, we must assume that the process of supercutting does not simply annihilate the spatially gunky block:

**PNatPhil 2 Independence of Motion and Substantial Change.** Neither the movement of two extended things nor the division of an extended thing into two parts can necessitate the creation or annihilation of material parts (not even point-masses).

Next, we have to assume that the parts of the block do not jump discontinuously at the very last moment into a new position.

**PNatPhil 3 Continuity of Motion.** It is impossible for any material thing to move discontinuously through spacetime.

Finally, we have to assume that if some body is spatially gunky, then all of the parts of the body must occupy connected regions of space.

**Connected Occupation.** If no part of a body is point-sized, then no part of the body occupies an isolated point—that is, if some part of the body occupies a point *P*, then there is a region, including *P* and extended in three dimensions, that that part occupies.

The Supercutting thought experiment shows that if we make these three assumptions and we assume that the process of supercutting, followed immediately by a period of stasis (the absence of motion) is possible, then spatial gunk must be impossible.

Can the Supercutting argument also be used as an argument against Aristotelian atoms, as well as against spatial gunk? Yes, since nothing turns on the fact that spatial gunk has actual, as opposed to merely potential parts.

Defenders of spatial gunk are going to have to object to the very possibility of supercutting. The most plausible argument against supercutting would depend on an assumption about the structure of time, namely, that it is impossible for any finite period of time to be divided into an infinity of actual parts. This requires Temporal Finitism (19.3T).

Summing up, we have one argument in favor of gunk (the appeal to conceivability) and several arguments against it (redundancy, spatial occupation, and supercutting). In order to avoid the arguments against gunk, those believing in gunk will have to embrace three positions: a Neo-Humeist account of causal powers (to avoid the Redundancy argument), the fundamentality of some but not all gunky heaps (to resist the spatial occupation argument), and an Aristotelian prohibition of processes with an actual infinity of parts (to counter the supercutting argument). This would be a very odd combination. Neo-Humeists are generally opposed to all Aristotelian prohibitions of a finitary nature. Thus, the appeal to gunk does not seem to provide a strong argument for Fundamental Heapism.

## 22.5 Fundamental Artifacts

Suppose that we have decided to embrace Anti-Heapism. This leaves us with the following possible fundamental entities: metaphysical atoms, artifacts, living organisms, and functional parts of artifacts and living organisms. In this section, we will turn our attention to the existence of artifacts and their functional parts.

The issue at hand concerns the existence of *material* artifacts. We are not concerned here with immaterial ones, like theorems, works of fiction or musical compositions. We are also not concerned with artifacts that are identical to living organisms. If a bonsai tree counts as both an organism and an artifact, it is not a mere artifact. We are concerned here with material mere artifacts.

**22.5T Fundamental Artifactualism.** Some material mere artifacts exist fundamentally.

**22.5A Anti-Artifactualism.** No fundamental material mere artifacts exist.

There are few, if any reasons, to suppose artifacts to be fundamental entities. Let's consider some of the reasons to deny them this status.

### 22.5.1 Arguments against fundamental artifacts

We consider two metaphysical objections to Fundamental Artifactualism. There is an appeal, first, to the intrinsicality of existence, and, second, to the arbitrariness of the persistence conditions of artifacts. The second objection is taken up in Chapter 25.

In addition to these two objections, two of the arguments against heaps also apply to artifacts. First of all, Ockham's Razor (**PMeth 1**) favors a theory without artifacts. Second, it does not seem plausible to suppose that artifacts have emergent active powers. The whole point of building artifacts from certain pre-existing materials is based on our confidence that the resulting artifact won't have any emergent powers, that what it will do or not do can be predicted confidently on the basis of the powers of its non-artificial parts. Hence, an appeal to Redundancy for Composite Entities (**PMeth 4**) gives us a good reason not to believe in artifacts as fundamental entities.

The argument from the Priority of Spatial Point-Parts is not as strong as an objection to artifacts as it was as an argument against heaps. Artifacts typically have a structure and organization that heaps lack. This structure might make it more plausible that it is the whole artifact that is the fundamental occupier of space, with the locations of its parts somehow dependent on the structure of the whole.

1. Intrinsicality of composition. The first argument against artifacts appeals to the Intrinsicality of Composition:

**Intrinsicality of Composition.** Whether or not the *x*'s compose something fundamental depends only on the intrinsic nature and mutual relations of the *x*'s, not on facts extrinsic to them.

The composition of artifacts seems to be extrinsic in two ways. First, it is dependent on the attitudes and practices of their users and maintainers, and, second, it is dependent on their physical surroundings.

Whether or not some things compose an artifact seems to depend on human intentions and actions that are extrinsic to the artifact itself. For example, suppose some ancient hunters shaped and chipped some rocks in order to form some crude implements, like axes and hammers. Thousands of years later, the chipped rocks have been abandoned and their functions forgotten. Do the axes and hammers still exist? Maybe, but it also seems plausible to say that all that remains are the chipped pieces of rock, now no different from other rocks that have been chipped or shaped by purely natural, unintentional processes.

Here is another argument for the same conclusion. As we have seen, in certain cases, found objects can constitute an artifact, like a stump-chair or driftwood-art. Whether or not the wood in the stump constitutes a chair seems to depend wholly on whether or not it is used and maintained as a chair by external agents. Similarly, imagine that the exact duplicate of a watch were to form by chance in an asteroid field. The watch-duplicate wouldn't be a real watch, but the difference between it and the watch is entirely extrinsic.

The existence of artifacts is extrinsic in a second way. Consider a statue that has been cut and chipped from a natural block of marble. The existence of the statue depends entirely on the physical surroundings of the marble making up the statue. That marble existed, with the very same size and shape, before the sculptor has removed any rock. The statue comes to exist, not by virtue of what happens to its internal material parts, but by virtue of what was done to the marble surrounding it.

2. Arbitrariness of persistence conditions. Finally, when we try to determine under what conditions an artifact persists or fails to persist in existence, we seem to fall into a series of insoluble paradoxes and puzzles, suggesting that artifacts are not fundamentally real. We take up this issue in Section 25.1.2.

## 22.6 Living Organisms vs. Mereological Nihilism

We have seen that there are some good arguments against both Fundamental Heapism and Fundamental Artifactualism. This leaves us with just two possibilities. Either (1) living organisms (and their functional parts) exist and are fundamental composite things or (2) no fundamental composite thing exists. The first view is 'Organicism' and the second is 'Priority Atomism'.

**22.6T Organicism.** Some living organisms are fundamental composite material things.

**22.6T.1 Composite Persons.** Living organisms that are persons capable of free will are fundamental composite material things.

**22.6A Anti-Organicism.** Living organisms do not exist or are not fundamental composite, material things.

**22.7T Priority Atomism.** There are no fundamental composite material things. All fundamental material things are atoms.

Priority Atomism is simply the conjunction of Anti-Heapism, Anti-Artifactualism, and Anti-Organicism.

### 22.6.1 Arguments for fundamental living organisms

First, we examine two arguments against Extreme Anti-Organicism, namely, the appeal to common sense and Descartes' *cogito* argument. Then we move to three additional arguments against Moderate Anti-Organicism, each of which supports Composite Persons.

1. Appeal to common sense. We start again with an appeal to common sense. We talk and act as though living organisms existed. In fact, we treat them and many of their properties, whether biological, psychological, or social, as among the fundamental

entities of the world. This appeal is, like other appeals to common sense, subject to the familiar strategies of rebuttal. Moreover, in this case there are four other arguments that can be joined with common sense in support of living organisms, namely, the appeal to the Cartesian *cogito*, an appeal to our legal, ethical, and political practices, an appeal to free will and irreducible human agency, and the charge that Priority Atomism involves a vicious circularity.

2. The cartesian *cogito* argument.[3] If heaps, artifacts, and living organisms don't exist, then there are no composite entities at all. What about us? As Descartes famously argued (an argument anticipated by St. Augustine), it is impossible for me to doubt that I exist, since whatever I may be led to doubt, I will still be doubting something, and I cannot doubt anything unless I exist. This is Descartes' *cogito* argument (from the Latin for 'I think'): I think, therefore I must exist. The *cogito* argument may not be airtight as a refutation of extreme skepticism, but as a beginning point for a sober and sane approach to metaphysics it is hard to resist. As Descartes himself, argued, the cogito seems to establish, not only that I exist, but that I have certain essential properties (properties of a sentient and mental character) that are at least prima facie irreducible to the merely physical properties of my proper parts (see Koons and Bealer 2010 for contemporary support for these convictions). If Priority Atomism entails that I, as a composite material thing, do not exist as a fundamental entity, then I would seem to have an excellent reason for rejecting it.

However, it is possible to embrace both Priority Atomism and Descartes's *cogito*, so long as I deny that I am composite. I could exist, even if no composite things existed, so long as I were a perfectly simple thing, a thing without proper parts. I could, for instance, be a simple soul, an immaterial being with no parts. Alternatively, as Roderick Chisholm (1976) proposed (and as Leibniz seems to have believed), I could be a simple material entity, some elementary particle that is located in my brain, a thinking and willing particle that interacts somehow with the other particles in my brain. Indeed, according to *panpsychism*, all particles are thinking, conscious entities!

The view that I am such a simple being, whether material or immaterial, is in some conflict with our common-sense view of our selves, putting the case for Organicism into some internal conflict.

On our common-sense view, I am both physical and mental. I have mental properties and powers, like thinking, feeling, consciousness, and physical powers and properties like *volume*, *mass*, and *chemical composition*. Thus, on our ordinary view I am composite but possess essentially unitary passive powers, like the powers of sensory experience and volitional self-determination. In addition, I am able to move my body, which seems to require emergent active powers. I move my hand; the particles that make up my body don't move the hand by themselves.

In addition, even if I am a simple soul or simple particle, what about my body? Surely it exists and is composite. The *cogito* argument, therefore, supplies evidence for Organicism by supplying evidence for Composite Persons and for the fundamental existence of other minded organisms, both of which entail Organicism.

Avoiding arbitrary distinctions among similar pluralities It would surely be crazy to adopt a view we could call 'Priority Solipsism', the thesis that I'm the only fundamental composite thing. If I exist fundamentally and am composite, then if some particles are arranged in the same way that my parts are arranged, then they should compose something of the same kind as me, namely, a person.

This principle raises an obvious problem: where is the most natural place to draw the line? All persons? All sentient organisms or animals? All living things? Even bacteria, viruses, prions?

Van Inwagen proposes (1990a) that we draw the line at cells and living organisms. Van Inwagen's answer to the Special Composition Question is that the *x*'s compose something if and only if their activities constitute a single life. Trenton Merricks (2003), in contrast, takes seriously the possibility of limiting composition to persons only, that is, to entities with self-determining free will.[4] Merricks's more restrictive answer would be that the *x*'s compose something fundamental if and only if their activities constitute the life of a person, a being with emergent powers of will. Clearly, there are a great many alternative positions. We might limit composition to those organisms that are sentient or conscious. We might try to exclude viruses and prions by insisting that the lives in question be relatively independent of the activities of other living things, although obviously no living thing is completely independent of other living things, and we don't want to treat complex parasites like tapeworms in the same way as viruses and prions.

If, with Merricks, we believe that the strongest argument against heaps and artifacts is the appeal to the Redundancy principle, then the question of where to draw the line will be answered by locating the point at which we first encounter emergent active powers and essentially unitary passive and immanent powers. This, in turn, may require us to look both to empirical science and to philosophical reflection on human experience, broadly considered. Van Inwagen makes no claims about emergent active powers, either in the case of human persons or in the case of living organisms more broadly. In fact, van Inwagen and many other contemporary philosophers consider emergent active powers to be dubious and outré in the extreme, based on some sort of principle of naturalism or physicalism. We will take up this challenge to an emergence-based account of Organicism in the next section.

3. Appeal to ethics, law, and politics. In our practices of laws, politics, and morality, we presuppose the fundamental reality of persons. If persons did not really exist and were utterly lacking in active, passive, and immanent powers, then our practice of holding people accountable for their actions in certain cases would make no sense. If human beings and sentient animals lacked the essentially unitary passive power of sensation, we could not explain the moral wrongness of cruelty. This argument relies on the Ethical Practices Presumption:

**PEpist 2.1 Ethical Practices Presumption.** It is prima facie plausible to suppose that all of the presuppositions of our fundamental ethical beliefs and practices are true.

4. Appeal to free will and human agency. Van Inwagen developed an influential argument for the incompatibility of free will and determinism in his *An Essay on Free Will* (1986b). The argument turns on the idea that, if I have free will, then it must be in my power to make it either true or false that I perform certain actions. If determinism is true, and if I have no power to make any proposition about the past true or false, then it follows that I lack free will, since in a deterministic world those past truths (over which I have no power) entail all future truths. The argument depends on van Inwagen's rule

*Beta*: if I have no power over the truth of $p$, and $p$ entails $q$, then I have no power over the truth of $q$.

This argument for the incompatibility of free will and determinism can be modified into an argument for the incompatibility of free will and the non-fundamentality of persons:

1 If I am not a fundamental entity, then any personal action $A$ that I take is wholly grounded in some facts $X_1, \ldots, X_n$ (where these are propositions that describe the movements of my constituent micro-particles).
2 If I'm not a fundamental entity, then I was not able to prevent the truth of $X_1, \ldots, X_n$: it was not in my power to make these facts be other than they are.
3 Necessarily, if $X_1, \ldots, X_n$, then I do $A$ (since $A$ is wholly grounded in $X_1, \ldots, X_n$).
4 The Beta rule: if I was not able to prevent the truth of $X_1, \ldots, X_n$, and if the truth of $X_1, \ldots, X_n$ entails the truth of $A$, then I am not able to prevent the fact that I do $A$, that is, it was not in my power to make it that I act otherwise.
5 If it was not in my power to make it that I act otherwise, then I am not personally responsible for my doing $A$.
6 So, if I am not a fundamental entity, then I am not personally responsible for my doing $A$.
7 But I am personally responsible for my doing $A$. Hence, I am a fundamental entity.

The weakest link in the argument is probably premise 2. Neo-Humeists in particular will object, as they would object to the corresponding step in van Inwagen's original argument. Much turns on what is required for my having the *power* to prevent a truth. If it simply means that in the nearest world in which I act differently, my particles also act differently, then it is obvious that I have such power over truths about my particles, even if they are fundamental entities and I am not. In contrast, if we think of powers as fundamental properties or as consequences of fundamental laws, as per Powerism or Nomism, it seems that non-fundamental entities must be essentially powerless.

5. A vicious circularity in Priority Atomism? Suppose one thought that facts that are grounded in, derived from, or reducible to the fundamental facts are at least in part constituted by our *conventions* or the contingencies of our *conceptual scheme*. On this view, non-fundamental facts are projections of contingent features *of us* as thinkers or as language users. If one embraces this view, Priority Atomism, who suppose that *we* are merely derived entities, face a problem of vicious circularity. If people are among the derived entities, then they would have to be both the projectors and the projections. If we are merely derived entities, what fundamental entities could be responsible for our quasi-existence? This objection seems especially cogent to one who argues that human persons are logically or conceptually grounded in impersonal facts.

### 22.6.2 Argument against fundamental living organisms: physicalism and emergent powers

As we mentioned above, the main objection to Organicism and to Composite Persons is based on Physicalism. Physicalism is the view that only physical (indeed, only micro-physical) phenomena are fundamentally real. All fundamental truths, on this view, are truths about the positions, motions, and inherent powers of micro-physical particles and fields. Such Physicalism absolutely rules out the existence of emergent active powers at the level of macroscopic objects such as organisms or persons. Any action of an organism or person is simply the joint action of its constituent particles and fields.

**Physicalism.** All fundamental truths are truths about the positions, motions, and inherent powers of microscopic particles and fields.

Why believe that Physicalism is true? The principal argument for Physicalism is empirical: we have discovered how, in broad terms, to reduce the phenomena of the macroscopic world to the microscopic. We can explain much, and we think that in principle we could probably explain all, of the behavior of molecules, crystals, fluids, cells, organisms, mountains, oceans, planets, stars, and galaxies in terms of the four fundamental forces of microscopic physics, like gravity, electromagnetism, and the strong and weak nuclear forces. We have shown how nearly all of the familiar objects of the macroscopic scale are wholly composed of a small number of fundamental particles (electrons, protons, neutrons, photons, and a few others), and we find little or no reason to believe that these particles ever behave in ways that could not in principle be explained by the powers that they manifest in isolated, experimental conditions.

Someone who rejects Physicalism, like Trenton Merricks, is an Anti-Physicalist or Emergentist. Emergentists have two possible responses to the empirical case for Physicalism. First, Emergentists could point out that our knowledge of how fundamental particles act when they are situated in complex organisms or in the brains and bodies of persons is still very partial, to say the least. Even when dealing with relatively simple bodies like crystals or fluids, some solid-state physicists have expressed skepticism about Physicalism, supposing that composite material bodies sometimes exercise emergent powers that are more than the sum of the powers of the constituent particles. A fortiori, we might think that it is still more reasonable to suppose that organic bodies and bodies of persons possess emergent active powers. Bertrand Russell at one point in the early twentieth century took seriously the possibility that electrons and other particles might move differently when under the influence of thought or will within the human brain, and nothing we have discovered since makes that possibility any less serious today.

Alternatively, Emergentists might accept that the uniformity of behavior that we have observed gives us reason to believe that fundamental particles always behave in essentially the same way, following the same regularities whether inside or outside a living body. We'll take up this second option in the next section.

### 22.6.3 Mock Physicalism and the absorption of powers

To make good on the second option just noted, Emergentists might concede that it would be unreasonable to suppose that our best theory of micro-physics is empirically inadequate. A theory is *empirically inadequate* if its predictions are at variance with any possible observation or experiment, unless we have a good explanation, in terms of the theory itself, for the unreliability of the conflicting observation or empirical result.

**Methodological Physicalism.** Reject any theory that is inconsistent with the empirical adequacy of our best theory of fundamental micro-particles and fields.

Can one accept Methodological Physicalism while rejecting Physicalism proper? If our micro-physical theory is empirically adequate, even with respect to the behavior of human brains and human bodies, shouldn't we admit that it contains all of the metaphysically fundamental truths? Not necessarily, especially if the Cartesian argument or the appeal to our practices gives us reason to suppose that we ourselves must be fundamental but macroscopic entities. We could suppose that when some particles compose a human body, the fundamental active powers of the particles are absorbed by the human being as a whole. As a result, the particles then *behave in accordance with* the ordinary laws of physics, but they do not *obey* those laws! They act as they do because of the exercise of new active and immanent powers of the whole human being, not as a result of their ordinary micro-physical powers, which they have utterly lost. The particles *satisfy* the descriptions of their behavior given by physics, but they do not *obey* the laws of physics.

To say that the microscopic parts of an organism *satisfy* descriptions in the language of physics is to stretch somewhat the usual meaning of 'satisfaction'. Strictly speaking, the atomic particles of an organism do not have the same physical natures as do the corresponding particles in the inorganic world. Instead, they have natures qua functional parts of the organism, natures that may perfectly mimic the behavior of the natures of the corresponding *free* particles, given the parameters of the organic body in which they occur.

The resulting picture is one that we could call 'Mock Physicalism'. It is immune from empirical refutation by empirical observation, but it rejects the further claim that the micro-physical level is metaphysically fundamental. Instead, within persons and other organisms, it is the organic and personal level that is fundamental and the micro-physical level that is derived and reducible.

## 22.7 Finding an Intelligible Principle of Composition

In this chapter so far, we have examined the question of whether there are any *fundamental* composite entities, and if so, of what kind. In this last section, we want to turn to the question of what sorts of composite entities exist at all, whether fundamental or derived. We might call this the Broad version of van Inwagen's Special Composition. The first question to ask is whether there is any principle of composition at all. We will return to our earlier discussion of heaps, since we will need to determine when a heap exists as a derived entity.

### 22.7.1 Is there a principle of composition?

First, we have to consider whether there is a principle of composition, a finite answer to the Broad version of van Inwagen's Special Composition Question. If there is, the answer to the Special Composition Question is *intelligible*. If not, then the Special Composition Question has a *brutal* answer:

**22.8T Intelligible Composition.** There is a finite set of necessary and sufficient conditions *C*, such that for any set of simples *S*, the members of *S* compose something if and only if they jointly satisfy *C*.

**22.8A Brutal Composition.** There is no such finite set of conditions.

A metaphysical theory incorporating Brutal Composition (Markosian 1998) would be in some sense much more complex than one satisfying Intelligible Composition, since the Brutal theories have to include separate axioms for the existence of each composite thing or for an infinite number of families of composite things, while Intelligible theories could instead list a finite number of axioms specifying which kinds of sets of simples compose something composite. So long as there are more things in the world than there are conditions in the Intelligible Composition set, the Intelligible theory will be simpler. Presumably that would be so.

So, let's assume that there is an intelligible answer to the Broad version of the Special Composition Question. What are some possible answers? One extreme answer would be to deny that some things ever compose anything, whether fundamental or derived:

**26.8T.1T Mereological Nihilism.** No non-empty set of things ever composes anything (whether fundamental or not).

**26.8T.1A Mereological Aliquidism.** Some non-empty sets of things do compose something: there is some other principle of composition besides Mereological Nihilism.

Mereological Nihilists can appeal to Ockham's Razor in defense of their principle, although, since Mereological Nihilism rules out not only fundamental composite things but also non-fundamental composite things, the force of Ockham's Razor is much reduced, as we saw in Section 3.9. A slightly less radical view would admit organisms as composite entities:

**26.8T.1A.1 Extreme Organicism.** A set of entities composes something if and only if they participate in a single life.

Extreme Organicism is defended by Peter van Inwagen (1990a). It entails that there exist no heaps and no artifacts:

**26.9T Extreme Anti-Heapism.** No heaps exist.

**26.10T Extreme Anti-Artifactualism.** No composite material artifacts exist.

The opposite extreme from Mereological Nihilism is Mereological Universalism:

**26.8T.2 Mereological Universalism.** Any non-empty set of material things composes something.

Mereological Universalism will be true if organisms, artifacts, and arbitrary heaps exist. Since we have discussed the existence of organisms, we will now turn to heaps and to artifacts.

### 22.7.2 Do heaps exist?

If heaps exist at all, when do things compose a heap? One simple answer is simply to deny that heaps exist at all. On this view, the only composite things that exist at all are the fundamental composite things, like organisms (or perhaps artifacts).

**22.9T Extreme Anti-Heapism.** No heaps exist.

We certainly talk as though we believed in the existence of heaps of many kinds. We talk about rocks, clouds, lakes, planets, stars, and so on. We say things like 'There are many mountains in Colorado' or 'Nine planets rotate around the sun.' Statements like these involve *bound variables*, which were Quine's proposed criterion of ontological commitment. These statements imply that some things are mountains and some things are planets. Prima facie, our acceptance of these common-sense statements commits us to believing in the existence of such heaps.

Common sense, then, tells us that many heaps exist. Therefore, Extreme Anti-Heapism might seem initially rather implausible. Extreme Anti-Heapists have a number of strategies for explaining away this apparent conflict with common sense:

1 **Paraphrase.** We don't really mean it when we say heaps exist. What we really mean could be put more precisely by saying that there are simples arranged heap-wise. That is, (1) is simply a loose and popular way of saying (2):

> (1) Some mountains are tall.
> (2) Some atoms (metaphysical simples) arranged in a mountainish way are collectively tall.

2 **Useful fiction.** We mean it when we say that heaps exist, but we do so in an intentionally fictional mode of speech. On this view, (1) is true only when understood as something like (3):

> (3) According to the useful fiction of Heapism: some mountains are tall.

(3) is a good deal like (4):

> (4) According to the Conan Doyle stories: a great detective once lived on Baker Street in London.

(4) is obviously true even if in fact no great detective ever lived on Baker Street. Similarly, (3) could be true even if in fact there are no mountains. We merely pretend that things like mountains exist because it would be so tedious and inconvenient to be talking all the time about simples arranged mountain-wise.

3 **Quasi-truth.** A statement about heaps is *quasi-true* (according to Sider 2001) if it is something that would have been true if heaps had existed. It seems plausible that (1) is quasi-true. If there were heaps, mountains would probably be among them, and some of these mountain-heaps would be tall. All of this could be true even if, in fact, there are no heaps and thus no mountains. It seems reasonable that we would assert

quasi-truths, in part because they are so much easier to say and to understand than the corresponding real truths, like (2) or even (3). In addition, it would be understandable if we were even to believe the quasi-truths to be really true. Most people are not aware of the complex subtleties of metaphysics, and many who are couldn't care less about them. Such people might well believe false metaphysical theories and thus come to believe many quasi-truths. Since quasi-truths are close enough to the truth for nearly all practical purposes, such errors would be unlikely to be corrected.

Thus, there are a number of strategies available to Extreme Anti-Heapists to argue that we aren't really or seriously committed to the existence of such heaps (e.g., paraphrase, useful fiction, quasi-truth). And here is some additional evidence that confirms Extreme Anti-Heapism. If we ask ourselves whether one and the same rock or lake or cloud has persisted through time or whether it has been destroyed and replaced by a new one, we often feel that the question has no deep or principled answer. If a small stream has been dammed into a lake, has one and the same body of water survived and been changed or have we destroyed the stream and created a new lake? If a small chip has been struck from a stone, has the stone survived the deletion or has it been destroyed and replaced by a new, smaller stone? When we consider questions like these, we are inclined to think that there is no real fact of the matter—that we should feel free to legislate or stipulate any answer that is practically useful under the circumstances. This attitude can be easily explained if, strictly speaking, there simply are no heaps. (The attitude can also be explained by Mereological Universalism, but that's no aid to common sense!) If talk of heaps is not literally true or is a merely useful fiction or only quasi-true, it would make sense for us to be free to alter our way of talking about heaps without concern that we might be losing our grip on any truth about the world.[5]

In *Material Beings* (1990a), Peter van Inwagen proposes a large number of possible answers to the Special Composition Question that are compatible with the existence of heaps. He finds serious difficulties with or objections to each.

The first is the simplest principle of all: Mereological Universalism. Mereological Universalists believe that the members of any non-empty set of things, no matter how diverse or scattered, no matter how loosely related to one another, jointly compose some further thing, namely, their mereological sum.

**22.8T.1A.2 Mereological Universalism.** For every non-empty set *S*, the members of that set compose something. (Equivalent to Arbitrary Sums MA6.)

On the plus side, Mereological Universalism is a very simple principle. In addition, if Universalism were true, there would be no vagueness or indeterminacy in the world. We would never have to worry whether some things really met the conditions necessary for composition, since there are no such conditions beyond mere existence.

On the minus side, Universalism seems to multiply entities needlessly, contrary to Ockham's Razor (**PMeth 1**). Second, Universalism is contrary to our ordinary way of identifying and counting things. Suppose we found four wooden rocking chairs in a room and were asked, 'How many wooden things are in the room?' We would be unlikely to answer that there were an astronomical number of wooden things, one corresponding to every possible combination of woody bits contained by the four chairs. We don't ordinarily believe in scattered or arbitrarily gerrymandered objects.

A third objection to Universalism involves the *Problem of the Many*. Suppose we have a wooden chair, Chair Maximus. If we carved off a few small wood chips, the resulting object would still be a wooden chair, Chair Minus. According to Universalism, Chair Minus is there all along, even before we carve off the chips. If Chair Minus is there, isn't it already a chair? It doesn't seem to be missing anything required to be a chair.[6] So, we are forced to say that we really have two chairs, Chair Maximus and Chair Minus. In fact, we have billions and trillions of chairs, one corresponding to each possible way of subtracting some of the matter of Chair Maximus in such a way as to leave a chair in existence. This is the problem of the many.

The problem of the many obviously applies to heaps, likes stones or mountains. But the problem of the many can also be developed in more disturbing cases. Consider a person Susan and Susan Minus, where Susan Minus would be there if some trivial bits of Susan's matter, like a hair follicle or a skin cell, were deleted. Universalism presses us toward the bizarre conclusion that wherever Susan is there are billions and trillions of other people, the various Susan Minuses. Or, if you prefer to identify Susan with an immaterial mind or soul, we are pressed toward the conclusion that she has trillions of overlapping but distinct human bodies.

The fourth major drawback to Universalism concerns the question of how composite things persist through time, a subject we will take up in Chapters 24 and 25. This is in fact van Inwagen's principal objection to heaps. For the time being, let's set aside Universalism.

The next principle is one that van Inwagen calls 'Fastenation'. On this view, a composite entity is brought into being whenever a number of things are fastened tightly together.

**Def D22.8 Fastened.** The *x*'s are *fastened* to each other if and only if they are so arranged that there are very few ways of separating them without breaking or distorting some of them. (van Inwagen 1990a: 56)

**22.8T.1A.3 Fastenation.** The *x*'s compose something if and only if they are fastened to each other.

In the *Physics*, Aristotle considers an even weaker and broader principle that we could call 'Locomotive Unification'. Some things are locomotively unified if and only if they tend to move together as a unit. Aristotle suggested that things have a certain kind of oneness, and so count together as a single composite thing, if they have locomotive unity.

**Def D22.9 Locomotively Unified.** The *x*'s are *locomotively unified* if and only if they are joined together in such a way that the vast majority of ways of moving one of the *x*'s (i.e., causing its acceleration in some direction) would move the others in the same direction.

**22.8T.1A.4 Locomotive Unification.** The *x*'s compose something if and only if they are locomotively unified.

However, neither Fastenation nor Locomotive Unification are plausible principles of composition, since they lead to implausible results in imaginable cases. Suppose two people are shaking hands. They do not compose a third thing, since they are not fastened together or locomotively unified. Suppose, however, that while they are shaking hands, their hands lock together as a result of muscular paralysis. Now they are fastened together, and we are supposed to believe that this fact about paralysis is sufficient to bring into existence a new thing (van Inwagen 1990a: 58). Similar cases pose problems for Locomotive Unification.

The next two principles van Inwagen considers are Cohesion and Fusion.

**Def D22.10 Cohere.** The *x*'s *cohere* together if and only if they are so arranged that it is impossible to separate them without breaking some of them (van Inwagen 1990a: 58).

**22.8T.1A.5 Cohesion.** The *x*'s compose something if and only if they cohere together.

**Def D22.11 Fuse.** The *x*'s are *fused* together if and only if they cohere together in a way that leaves no discernible boundary.

**22.8T.1A.6 Fusion.** The *x*'s compose something if and only if they are fused together.

Van Inwagen argues that even if the hands of the two people discussed above were glued together, satisfying Cohesion, this wouldn't be sufficient to bring into existence a new composite thing having the two people as its parts. Even if a mad scientist were to fuse two people together, say, two identical twins, into a case of artificial Siamese twins, this would not produce a new, composite thing containing the two twins as parts. We still have only the two individual twins, even when there is no discernible boundary between them. If the two people do compose some third thing, it is not because of their cohesion or fusion.

So far, van Inwagen's objections to the bonding principles have relied on our common-sense intuition that when two people or two living organisms of any kind are bonded together, we do not thereby produce a new composite thing. Van Inwagen next considers some principles that sidestep these objections by ensuring that the only things that can compose anything are metaphysical atoms:

**Def D22.6 Atom.** *x* is an *atom* if and only if *x* has no proper parts.

**Def D22.12 Maximally Fastened.** The *x*'s are *maximally fastened* if and only if they are fastened, and there are no *y*'s such that the *y*'s are fastened and the *x*'s are among the *y*'s, unless the *x*'s just are the *y*'s (van Inwagen 1990a: 63).

**22.8T.1A.7 Democritean Composites.** The *x*'s compose something if and only if they are all atoms and the *x*'s are maximally fastened.

If two people are fastened together, Democritean Composites still gives the intuitively wrong answer that a new composite thing has come into existence. Still worse, it now entails that neither of the two people exist, since neither is composed of atoms that are maximally fastened to each other, since the atoms in one are fastened to the atoms in the other.

Perhaps the right solution is to build up the world of composite things in a step-by-step, hierarchical fashion, beginning with the atoms and building progressively larger and more complex entities. In this way, we might get chemical atoms, molecules, organisms, planets, star systems, and galaxies, without getting any of the bizarre entities we want to avoid, like the supposed thing that results when two people are fastened or fused together. This is Serial Bonding:

**22.8T.1A.8 Serial Bonding.** The *x*'s compose something if and only if the *x*'s are atoms bonded in ways *R1* or *R1′* or *R1″*, and so on, or the *x*'s are things composed of atoms bonded in way *R2* or *R2′*, and so on, or the *x*'s are things composed of things composed of atoms bonded in way *R3*, and so on, or ….

So, protons and neutrons have to be bonded in a certain way to make nuclei, and nuclei and electrons are bonded in a different way to make chemical atoms, and chemical atoms are bonded in a third way to make molecules, and so on up to rocks, clouds, organisms, planets and stars, solar systems, galaxies, galaxy clusters, and even the universe!

Van Inwagen offers a number of worries about and objections to Serial Bonding. First, since 'atom' is defined in mereological terms (i.e., as not having a proper part), van Inwagen rules out both Serial Bonding and Democritean Composites as satisfactory answers to the Special Composition Question on the grounds of circularity. However, it is not clear that Heapists must propose a non-circular answer, so long as there is a finite and intelligible principle of the composition of heaps.[7]

Second, Serial Bonding violates the transitivity of *parthood*, since an atom can be part of a composite of level 1, and a composite of level 1 be a part of a composite of level 2, but the atom will not, according to Serial Bonding, be a part of the composite of level 3, contrary to axiom MA3 of mereology (see Section 23.1). This doesn't seem to be a difficult objection to meet, since Heapists could take Serial Bonding to define a '*part**' relation, and then could define the '*part*' relation as what is called the 'transitive closure' of the *part** relation. Roughly, *x* is a part of *y* if and only if *x* is a part* of *y* or *x* is a part* of a part* of *y* or ….

Third, even if a version of Serial Bonding could be constructed that would work in our world, given the kinds of matter we actually have and the sorts of bonding permitted by the actual laws of nature, it would seem to tell us nothing about what sort of things might exist in worlds with different kinds of matter and different laws of nature. Serial Bonding is too parochial a principle to serve as a metaphysical account of composite entities.

Fourth, if Serial Bonding is going to avoid the counterexamples involving fused persons and organisms, it will have to explain in a principled way why fusing together some crystals produces a composite rock, but fusing together two or more organisms produces no composite entity at all. It is hard to see how this explanation could avoid being ad hoc and therefore implausible.

Finally, Serial Bonding, like Universalism, faces a serious challenge in accounting for the persistence and non-persistence of things through time. We take up this issue in Chapters 24 and 25.

We need to consider one final principle, one not mentioned by van Inwagen. This principle seems to be implicit in much of Aristotle's work:

**Def D22.13 Homogeneity.** *x* is *homogeneous* if and only if *x* has proper parts, and any two parts of *x* of the same size and shape are perfect duplicates (intrinsically indistinguishable).

**22.8T.1A.9 Homogeneous Continua.** The *x*'s compose something if and only if they form a homogeneous material continuum (a material thing with no sharp internal boundaries).

Homogeneous Continua avoids van Inwagen's test case, since it is impossible for a living organism to be a homogeneous continuum. Consequently, two organisms fused together will also fail to form such a continuum. However, Homogeneous Continua will entail that no composite organisms exist, since organisms are never perfectly homogeneous. We could modify Homogeneous Continua to guarantee that all organisms count as composite things:

**22.8T.1A.10 Homogeneous Continua Plus Organisms.** The *x's* compose something if and only if they form an organism or a homogeneous material continuum.

However, this principle involves a significant departure from common sense. Few of the world's heaps are homogeneous continua. Neither rocks nor clouds nor bodies of water live up to this standard. Still, we might point out that many of them are *approximately* homogeneous and *appear to the naked eye* to be homogeneous. Perhaps common sense simply ignores the deviation of many heaps from the standard of homogeneous continua, indulging in the useful fiction that they are homogeneous and continuous or being satisfied with the quasi-truth that they are.

Thus, Homogeneous Continua seems to avoid the usual challenges to answers to the Special Composition Question. However, would material continua of this kind really be heaps? Would they really be composite things at all? Do homogeneous continua have *actual* parts or only *potential* ones? Since no sharp boundaries articulate a material continuum into ready-made parts, it seems natural to think that the parts of such a continuum are only potential. If this is so, then material continua are really metaphysical atoms of a certain kind, which we could call 'Aristotelian atoms', to distinguish them from the absolutely indivisible Democritean atoms.

**Def D22.14 Aristotelian Atom.** *x* is an *Aristotelian atom* if and only if *x* is an atom (with no actual proper parts) and *x* has the passive power of bringing into existence one or more fission products.

**Def D22.14 Fission Product.** *y* is a *fission product* of *x* if and only if *x* has a passive power *P* such that, as a matter of necessity, when *P* is exercised, *x* ceases to exist, *y* begins to exist, and *y* occupies part of the former location of *x*.

**Def D22.16 Democritean Atom.** *x* is a *Democritean atom* if and only if *x* is an atom but not an Aristotelian atom.

If all homogeneous continua are Aristotelian atoms, then the two versions of Homogeneous Continua will be empty, since there will never be a number of actual distinct things that together make up a single continuum of this kind. Aristotelian atoms are not heaps at all, since they lack actual proper parts. If all homogeneous continua are Aristotelian atoms, then none of them are heaps. Once again, we have failed to come up with a principle of composition that gives us precisely the heaps that common sense wants, undermining the legitimacy of an appeal to common sense.

### 22.7.3 Do artifacts exist?

Consider again the thesis of Extreme Anti-Artifactualism:

**22.10T Extreme Anti-Artifactualism.** No composite material artifacts exist.

As in the case of heaps, the main argument against Extreme Anti-Artifactualism is an appeal to common sense. If we claim that there are no forks or spoons, automobiles or ships, tables or chairs, we are likely to be met with an incredulous stare. How could we possibly deny the existence of such familiar, everyday objects? Anti-Artifactualists can appeal to the same strategies here as Anti-Heapists: paraphrase, useful fiction, and quasi-truth. (5), for example, can be paraphrased as (6):

(5) THP is sitting on a chair.
(6) THP is sitting on some particles arranged chair-wise.

Finally, we might suppose that we can give a coherent answer to van Inwagen's Special Composition Question insofar as it applies to artifacts. The answer might go something like this:

**22.8T.1A.11 Composition of Artifacts (1).** Some material things compose a mere material artifact if and only if they have been altered or arranged by some intelligent agent (or group of agents) for a single purpose or interdependent set of purposes, and they do not compose a living thing.

However, is it really necessary for some things to be altered or arranged in order to constitute an artifact? There are such things as found artifacts. We could find a stump in the woods and make it into a chair simply by using it as such.

In addition, is arrangement sufficient to constitute an artifact? Consider a broken and abandoned watch. Is it still an artifact or is it now a mere heap of metal parts? The parts were altered and arranged for a purpose, but the watch no longer serves that purpose. To overcome these challenges, we might try an alternative principle of composition:

**22.8T.1A.12 Composition of Artifacts (2).** Some material things compose a mere material artifact if and only if they are used and maintained by some intelligent agent (or group of agents) for a single purpose or interdependent set of purposes, and they do not compose a living thing.

1. Immaterial artifacts. Let's suppose that holes, shadows, and spots of light on a wall, for example, are *not* material things. Artifacts can be made of such immaterial entities. For example, Alexander Pruss has pointed out that one could make a chess set simply by forming holes in a thick, viscous mound of jelly. One moves one's queen in this set by inserting a tool into the hole in the jelly that is the queen and slowly moving the hole to a

new position on the chessboard, and then removing the tool. The pieces of such a chess set cannot be material entities, since their parts are not material entities. But if this holey chess set is not a (composite) material thing, then we shouldn't suppose that an ordinary chess set is, simply by virtue of its being composed of different things.

There are even simpler examples of the same phenomenon. A trench is an artifact, but a trench consists simply in a long, narrow hole that is produced by digging. In addition, there are works of art that consist entirely of shadows or points of light.

Some ordinary artifacts, like works of literature (poems, plays, novels) and musical compositions, seem to be wholly abstract. A musical composition is just a sequence of notes, an abstract entity that is not located in space. Do we really believe that a musical composer alters the world of abstract objects, bringing into existence a new sequence? Surely not. But if the creation of a new thing is not required for such musical or literary creating, why think that an act of manual creation involves the coming into being of an entity, as opposed to the mere re-arrangement of existing things?

2. Making artifacts out of living things. Van Inwagen (1990a: 126–127) asks us to imagine an artifact made entirely of a living thing. For example, we could imagine making a very long snake into a hammock by tying it together into a network of knots. Doing this to the snake would not bring into existence a new thing nor would it destroy the snake. This isn't a case in which the living thing counts, all by itself, as a material artifact (as a bonsai tree or an artificial organism might). Since the hammock is neither an old nor a new material thing, it seems that it cannot be a material thing at all. If we appeal again to Functionality the Essence of Artifacts, no hammock could be a material thing, since any hammock is functionally equivalent to any other.

We might suppose that turning the snake into a hammock fails to produce a new material thing precisely because the snake already existed. However, this seems an entirely ad hoc solution. How can arranging a rope in a certain way bring into existence a new material entity if arranging the rope-like snake in precisely the same way for the same purpose fails to do so?

## Notes

1 As we saw in Section 3.4, there are two further sub-questions: When can the truths about composite things be conceptually grounded in facts about their parts? and, When can the facts about composite things be non-conceptually grounded in facts about their parts?

2 Compare Definition D11.1, gunky body. A body that is spatially gunk will satisfy D11.1.

3 We write this section, as well as the one one below about free will, in the first-person singular, for ease of exposition.

4 Merricks considers it an open empirical question whether there are emergent powers possessed by other composite things, including molecules and sub-personal organisms.

5 See Merricks (2003), Chapters 2 and 7.

6 To be precise, it is not missing anything that can be expressed in a non-mereological vocabulary.

7 Van Inwagen defines the Special Composition Question in such a way as to require a non-circular answer (one that doesn't include mereological vocabulary). The defender of Serial Bonding could challenge the appropriateness of that requirement.

# 23

# Composition: The General Question

In this chapter, we take up issues to do with van Inwagen's (1990a) General Composition Question: what is it for one thing to be a part of another? We begin in Section 23.1 with some background to do with formal mereology, the study of parts and wholes. We identify four answers to the General Composition Question in Section 23.2, though we discuss only three in detail, and in Section 23.3 we consider whether those answers can supply a ground for the correct principles of mereology. Finally, in Section 23.4, we briefly consider the connection between parthood and truthmaker, which in part prompts the discussion of Section 25.1, on whether wholes can change their parts.

## 23.1 Formal Mereology: Leśniewski, Goodman, and Leonard

In discussing the metaphysics of parts and wholes, it is helpful to have some specialized vocabulary, as well as a well thought-out mathematical model of a very broad, inclusive theory. The theory of mereology, proposed by the logician Stanislaw Leśniewski (1916) and introduced to the wider world by Goodman and Leonard in an article in the *Journal of Symbolic Logic* (1940), provides that vocabulary and such a model. These mereologists proposed some basic axioms for the *part-whole* relation. First, they assumed that the *part-whole* relation is reflexive, antisymmetric, and transitive. Reflexive relations are relations such that everything stands in the relation to itself; antisymmetric relations are relations such that if a thing *A* stands in the relation to a thing *B* and if *B* stands in the relation to *A*, then *A just is B*, *A* and *B* are identical; transitive relations are relations such that if a thing stands in the relation to a thing *B*, and if *B* stands in the relation to a thing *C*, then *A* stands in the relation to *C*. In the case of *parthood* then, Leśniewski, Goodman, and Leonard proposed these axioms:

**(MA1) Mereological Reflexivity.** Everything is a part of itself.

*The Atlas of Reality: A Comprehensive Guide to Metaphysics*, First Edition.
Robert C. Koons and Timothy H. Pickavance.

**(MA2) Mereological Anti-Symmetry.** If $x$ is a part of $y$, and $y$ is a part of $x$, then $x = y$.

**(MA3) Mereological Transitivity.** If $x$ is a part of $y$, and $y$ is a part of $z$, then $x$ is a part of $z$.

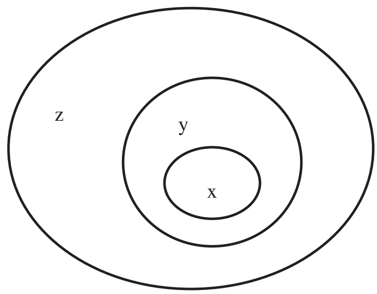


**Figure 23.1** Mereological Transitivity

Mereological Reflexivity seems an odd requirement, since we do not usually speak as though anything were a part of itself. However, it is convenient for technical reasons to use 'part' in this slightly odd way. (We needn't worry about these reasons at this stage.) Whenever we want to say that $x$ is a part of $y$ but not identical to $y$ itself, we shall say that $x$ is a 'proper part' of $y$. A 'proper part' is just defined as: a part of but not identical to the whole.

**Def D23.1 Proper Part.** $x$ is a *proper part* of $y$ if and only if $x$ is a part of $y$ and $x \neq y$.

Once we understand the distinction between part and proper part, the reflexivity of *parthood* is unproblematic. Reflexivity is simply a consequence of the reflexivity of identity. Everything is identical to itself, so everything is a part (in this sense) of itself.

Anti-Symmetry makes *parthood* like the subset relation and the *greater-than-or-equal-to* relation from mathematics. If $x$ is greater than or equal to $y$, and $y$ is greater than or equal to $x$, then $x$ and $y$ must be identical. It is impossible for two distinct numbers each to be greater than or equal to each other. Similarly, mereologists assume that it is impossible for two distinct things each to be a part of the other. There are no part-whole loops involving two different entities. This seems a reasonable principle. At least, it is hard to think of any clear counterexamples.

Transitivity is somewhat more problematic. There do seem to be cases in which one thing is part of a second, the second a part of a third, and yet we do not want to say that the first is a part of the third. For example, the quarterback's left earlobe is part of the quarterback, the quarterback is part of the football team, and yet it would be strange to say that the quarterback's left earlobe is part of the football team.

Mereologists have two possible responses to these apparent counterexamples to Mereological Transitivity. First, they could suppose that in certain contexts we use 'part' to refer to only special, salient parts of the whole. For example, when we talk about the 'parts' of a football team, we normally mean the team's members, and we don't intend to include all of the parts of those members, even though, strictly speaking, the parts of the members are really parts of the team.

Second, mereologists could admit that the metaphysically most basic relation is non-transitive or even intransitive. If we have a non-transitive relation $R$, it is easy to define a new relation $R^*$, the *ancestral* or *transitive closure* of $R$, that is transitive. For example, the relation of *being a parent* is obviously not transitive: the parent of a parent is not normally a parent. However, we can define the relation of *being an ancestor* as the transitive closure of *being a parent*: $x$ is an ancestor of $y$ if and only if $x$ is a parent of $y$ or there is some $z$ such that $x$ is a parent of $z$ and $z$ is a parent of $y$ or there are a $z$ and $z'$ such that … The relation of *being an ancestor* is necessarily transitive. Similarly, we could let the logic of mereology apply to the transitive closure of whatever *parthood* relation is metaphysically fundamental or natural. This option comes at a cost, however. If the fundamental relation is non-transitive, then each of the other axioms of mereology have to be strengthened in order to apply to the transitive closure of *parthood*.

Next, it is convenient to define the relation of *mereological overlap* or 'overlap' for short. Two things overlap just in case they have a part in common. *Overlap* is obviously a reflexive and symmetric relation, where $R$ is a symmetric relation if and only if, if a thing $A$ stands in the $R$ to a thing $B$, then $B$ stands in $R$ to $A$ as well. Everything overlaps itself, and $x$ overlaps $y$ if and only if $y$ overlaps $x$. *Overlap* is not transitive, however.

**Def D23.2 Overlap.** *$x$ overlaps $y$* if and only if there exists a $z$ such that $z$ is part of $x$ and $z$ is part of $y$.

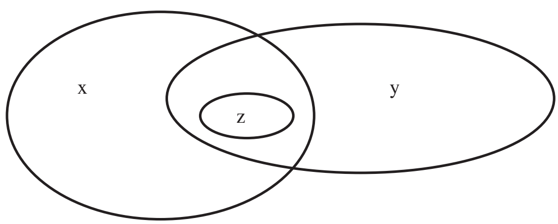


**Figure 23.2** Mereological Overlap

There are other plausible principles of mereology included by Leśniewski, Leonard, and Goodman as axioms of mereology. First, suppose that one thing is a proper part of another. For example, my hand is a proper part of my body. This seems to entail that my body has some other proper part, one that does not itself overlap my hand. This principle is Weak Supplementation:

**(MA4) Weak Supplementation.** If $x$ is a proper part of $y$, then $y$ has some part $z$ that does not overlap $x$.

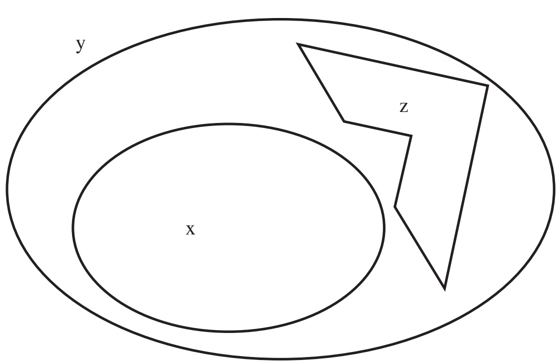


**Figure 23.3** Weak Supplementation

Second, suppose something isn't a part of RCK's body. Suppose RCK is standing in the middle of the room. The north half of the room is not a part of RCK's body, although it does overlap his body. In that case, the north half of the room must have some part that doesn't overlap RCK's body. Otherwise, his body must include the north half of the room. Thus, if *x* is *not* a part of *y*, then *x* must have some part that does not overlap *y*. This is Strong Supplementation:

**(MA5) Strong Supplementation.** If *x* is not a part of *y*, then *x* has some part *z* that does not overlap *y*.

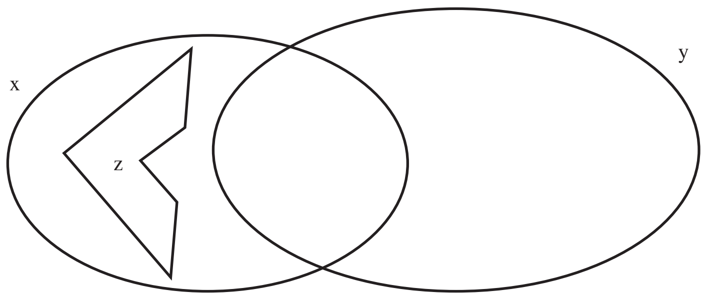


**Figure 23.4** Strong Supplementation

Third, and most controversially, Leśniewski, Leonard, and Goodman assume that if *S* is a non-empty set or collection of things (a *non-empty* set is one that has at least one member), then there exists something that every member of *S* is part of, and which has nothing more as parts. An object *o* is said to be a *sum* of the members of *S* just in case everything that overlaps a member of *S* overlaps *o*, and everything that overlaps *o* overlaps at least one of the members of *S*.

**Def D23.3 Sum.** *x* is a *sum* of set *S* if and only if *S* is a non-empty set, and for all *z*, *z* overlaps *x* if and only if *z* overlaps one of the members of *S*. (Equivalent to our definition D18.2 in Chapter 18.)

This controversial principle, that every set of things has a sum, is Arbitrary Sums:

**(MA6) Arbitrary Sums.** If *S* is a non-empty set, then there exists at least one sum of *S*.

Arbitrary Sums is controversial because it entails the existence of some very odd objects, including highly scattered objects. Consider the non-empty set where the only two members are your left big toe and the Eiffel Tower. Arbitrary Sums entails that there exists a scattered object having your left big toe and the Eiffel Tower as parts, but which overlaps nothing but things that overlap either your big toe or the Eiffel Tower. Such a scattered object has a part here and a part in Paris, but no parts in between the two objects. Arbitrary Sums also entails that there exists something we could call "Big Aluminum", an entity that includes as a part everything made wholly of aluminum and nothing that isn't made of aluminum. This object includes every aluminum can as a part, but none of the air or packaging that separates one can from another. Even worse, it entails the existence of sums that are both scattered and randomly heterogeneous, such as the northern half of every molecule on Earth. A common sense inventory of the world's furniture wouldn't include these bizarre amalgams!

There is a somewhat weaker version of Arbitrary Sums to consider, namely, Arbitrary Parts.

**(MA6′) Arbitrary Parts.** If there is a sum of the members of set $S$, and $S'$ is a subset of $S$, then there is a sum of the members of $S'$.

Arbitrary Parts is weaker than Arbitrary Sums because it is restricted to those sets of things that are already parts of some other thing. Suppose that there is no such thing as the World, a single object containing everything as a part. Then it might be that President Obama's left hand and the Eiffel Tower are not parts of any one thing. If so, then the axiom Arbitrary Parts does not entail that there is a sum of those two things. However, if Mr. Obama's left hand and his right foot are both parts of his body, then the axiom Arbitrary Parts would entail that there is something that is exactly the sum of his left hand and his right foot.

Finally, we can define what it is for some things to compose something else.

**Def D23.4 Composition.** The $x$'s *compose* $y$ if and only if $y$ is a sum of the set having the $x$'s as its only members.

With this terminology and the theory comprised of axioms (MA1–MA6) on the table as a starting point, we can consider the metaphysics of composition.

What sort of relation is *parthood*? Clearly it is binary and non-symmetric. For $A$ to be a part of $B$ does not entail that $B$ is a part of $A$. In the next section, we move beyond these initial, uncontroversial observations.

## 23.2 Three (or Four) Answers to the General Composition Question

Given our work on universals and relations in Chapters 7 and 10, we can identify a range of possible answers to the General Composition Question.

1. Fundamental or natural parthood. First, there is there is the view that the *parthood* relation is either fundamental or natural:

**23.1T Compositional Realism.** The *parthood* relation, or some closely related relation, like *joint-composition*, is fundamental or corresponds to a relational universal.

Compositional Realism comes in two varieties, depending on whether we take parthood to be absolutely fundamental and ungrounded, or whether we take it to be grounded in the instantiation of some universal (or grounded in a set of fundamental resemblance facts):

**23.1T.1 Fundamental Parthood.** The *parthood* relation, or some closely related relation, like *joint-composition*, is fundamental.

**23.1T.2 Natural Parthood.** The *x is-a-part-of y* relation, or some closely related relation, like the *x's jointly compose y*, corresponds to a relational universal or a natural relational resemblance class.

The idea behind these views is that *parthood* is like *instantiation* for UP-Realists (7.1T.1T) or *resemblance* for Resemblance Nominalists (8.1T.4), or else part-whole pairs instantiate some relational universal or belong to a perfect similarity class. For Constituent Ontologists (9.1T), *parthood* may be the only fundamental relation. For Ostrich Nominalists (7.1A.1A), the distinction between Fundamental Parthood and Natural Parthood collapses, since for them the natural properties just are the fundamental ones.

2. Compositional Anti-Realism. The second answer to the General Composition Question is Compositional Anti-Realism:

**23.1A Compositional Anti-Realism.** The *parthood* relation is neither fundamental nor natural.

If Compositional Anti-Realism is true, then no fundamental truth involves *parthood*. Compositional Anti-Realism comes in two versions, moderate and extreme.

**23.1A.1T Moderate or Reductive Compositional Anti-Realism.** There are truths about proper parthood, but all truths about *proper parthood* are grounded in more fundamental, non-mereological truths.

**23.1A.1A Compositional Nihilism.** Nothing is a proper part of anything else.

Compositional Nihilism implies that everything we ordinarily say and believe about wholes and parts is simply false. Moderate Compositional Anti-Realists instead might propose some sort of translation or paraphrase of talk using mereological vocabulary into a theory without *parthood* or *composition* relations. Alternatively, Moderate Compositional Anti-Realists can treat *wholes* as wholly derivative or grounded entities. On this view, fundamental reality consists only of simple things.[1] When we refer to composite things, we are merely referring to something wholly grounded in the simple parts and their arrangement.

All Compositional Anti-Realists must assume that, if anything exists at all, the fundamental things are all simple or atomic (lacking proper parts). Either composite things don't exist at all (Compositional Nihilism) or else they can simply be identified with derived entities that are wholly grounded by pluralities of atoms.

Among Moderate Compositional Anti-Realists, there is a further distinction to make, between those Moderate Compositional Anti-Realists who want to make Arbitrary Sums come out true, and those who do not. The members of the first group are Arbitrary Compositional Reductionists, and those of the second are Non-Arbitrary Arbitrary Compositional Reductionists.

**23.1A.1T.1T Arbitrary Compositional Reductionism.** All truths about *proper parthood* are grounded in more fundamental, non-mereological truths in such a way that Arbitrary Sums is true.

**23.1A.1T.1A Non-Arbitrary Compositional Reductionism.** All truths about *proper parthood* are grounded in more fundamental, non-mereological truths in such a way that Arbitrary Sums is false.

All Compositional Anti-Realists reject the existence of a metaphysically fundamental *composition* relation, the relation that holds between some *x*'s and *y* when the *x*'s compose *y*. However, not all Compositional Reductionists need suppose that any set of atoms constitute a composite thing. Non-Arbitrary Compositional Reductionists can suppose that whether something is a part of another depends upon special facts about those things and their non-compositional relations. The Arbitrary Compositional Reductionists, in contrast, think that we can talk truthfully about composites formed from every set of simples, since there are no relevant facts to distinguish some sets of simples from others.

Arbitrary Compositional Reductionists can offer something like the following explications or definitions of the part-whole and composition relations:

> (1) If *S* is a non-empty set of atoms, then the members of *S* compose *x* if and only if *x* is that unique entity whose existence and intrinsic character is wholly grounded in the existence, character, and arrangement of the members of *S*.
> (2) If *A* is composed of the atoms in *S*, and *B* is composed of the atoms in *S*', then *A* is a part of *B* if and only if *S* is a subset of *S*'.

Thus, it seems that Arbitrary Compositional Reductionists should think of the grounding of facts of composition and parthood in non-mereological facts as a case of conceptual grounding (Section 3.4), with the implication that composite objects do not really exist. Non-Arbitrary Compositional Reductionists, in contrast, could instead think of the grounding relation as extra-conceptual, dependent on the real essences of the kinds of composite things there are.

A major drawback with Arbitrary Compositional Reductionism is that it rules out the possibility of metaphysically fundamental composite things, in opposition to F.H. Bradley's argument against the universal priority of parts (Section 11.2.3), and in opposition to evidence for emergent substances (Section 22.6).

3. Other Forms of Compositional Reductionism. There are several versions of Non-Arbitrary Compositional Reductionism that deserve consideration. One specific proposal to consider is that one thing *x* is a part of *y* if and only if the location of *x* is part of the location of *y*. This proposal suffers from two problems. First, it seems that two things could be located in the same place without sharing any parts. For example, physics allows multiple bosons (like photons) to co-exist in the same location without sharing any physical parts. We might also think of various fields (like magnetic and gravitational fields) as mereologically separate entities that share the same location. Second, we shouldn't rule out the possibility of non-located objects, like souls or mathematical objects. It might turn out that some of these have parts. Finally, and most importantly, the proposal doesn't

really define the *parthood* relation in a non-circular way, since it presupposes that we have an account of what it is for one place to be a part of another.

Another proposal for defining composition is found in Koons (2014b), Aristotelian Reductionism. On this view, composition and the part-whole relationship (for substances) is wholly grounded in facts about causal powers and real processes.

**23.1A.1T.1A.1 Aristotelian Compositional Reductionism.** Some things, the *x*'s, compose *y* at time *t* if and only if the *x*'s are more than one, and there is some process *P* involving exactly the *x*'s as participants such that (i) the existence of *y* at time *t* is causally explained by the continuation of *P* during some *I* that is terminated by *t*, (ii) all of the power-conferring properties of each of the *x*'s are at *t* wholly grounded in the power-conferring properties of *y* at *t*, and (iii) every property of *y* that confers an active or passive power on *y* at *t* ontologically depends on one or more of the *x*'s.

We can use this model of composition to define the part-of relation:

**Aristotelian Definition of Proper Part.** A substance *z* is an immediate *proper part* of *y* at *t* if and only if *z* is among some *x*'s that compose *y* at *t*.

This account provides a two-way dependency between parts and wholes, one that respects the principle of Redundancy that we discussed in Chapter 22. The whole depends causally and diachronically (across time) on the activities of its parts, and the parts depend synchronically (at each moment) upon the whole for their occurrent causal powers. In addition, every active or passive power of the whole essentially involves one or more of the parts: that is, the whole cannot act or be acted upon except through some of its parts (thanks to clause iii). Clauses (ii) and (iii) provide grounds for thinking that a whole is spatially located exactly where its parts are located. In addition, the definition explicitly permits *mereological inconstancy*: wholes that gain or lose parts over time.

4. Composition as identity. The fourth answer to the General Composition Question was suggested by Donald Baxter (1988) and has been endorsed, with some important qualifications, by David Lewis (1991).

**23.1T.1.1 Composition as Identity (CAI).** Every whole is identical to its parts: to be a proper part is to be one of the things that are (collectively) identical to the whole.

Put simply, Composition as Identity is the view that wholes are simply *identical to* their parts (taken collectively, not distributively). If an encyclopedia is composed of ten volumes, then the encyclopedia simply *is* the ten volumes, and the ten volumes *are* the encyclopedia. The cloud is identical to its constituent droplets, and the droplets are identical to the cloud. Thus, to be a part of a whole *W* is simply to be one of the things that are (taken collectively) identical to *W*. On this view, *parthood* is simply a form of identity, an unproblematic, merely logical notion. Thus, this view is a form of Fundamental Parthood, since the *identity* relation is plausibly fundamental.

Unlike Compositional Reductionists, CAI theorists are not obviously committed to the view that only atoms are fundamental, or even to the view that there are atoms at all. CAI theorists are not saying that parts are more fundamental than the wholes they compose. In fact, they are equally fundamental, since the parts just *are* the whole.

Since according to CAI, the whole is nothing "over and above" its parts, a theory that posits the existence of a whole, given any collection of parts, doesn't commit its endorsers to any addition to reality, over and above the parts themselves. One might then think that if the droplets exist, so does the cloud; if the volumes exist, so does the encyclopedia; and, in general, for any non-empty set *S*, the members of *S* jointly form a whole, the whole being nothing more than the parts themselves. However, as Kris McDaniel (2010) has pointed out, CAI is compatible with Extreme Compositional Nihilism because it does not entail that the members of any set ever compose anything at all. Nonetheless, we might conjecture that a modified or stronger version of CAI would entail the axiom Arbitrary Sums (MA6) by requiring that the members of any non-empty set are collectively identical to something. (We expand this point in Section 23.3.)

Peter van Inwagen (1994) has complained that Baxter's theory can't be formulated without violating the rules of grammar. The identity predicate of standard logic is a binary predicate that combines two singular terms: *x* is identical to *y*. Baxter's theory requires us to use identity when referring to both a single thing (the whole) and a plurality (the parts): *x* is identical to the *y*'s. Moreover, it seems that linking singular and plural terms with identity violates one of the fundamental principles of logic, Leibniz's Law of the Indiscernibility of Identicals (not to be confused with the converse principle, the Identity of Indiscernibles 7.3T). Leibniz's Law says that if *A* is identical to *B*, then *A* and *B* must exemplify all the same properties. But a whole is one thing, and its parts are many (more than one). If the whole were identical to the parts, then the parts would also be one thing, and not more than one. So the whole cannot be identical to the parts.

In response, the defender of CAI could reply that grammatical convention is a poor guide to metaphysical truth. The fact that ordinary language and standard logic requires a sharp distinction between singular and plural terms shouldn't constrain our metaphysical theorizing.

In response to the charge of inconsistency concerning one-ness and many-ness, we should say that composite wholes are both one (in a way) and more than one (in another). The encyclopedia would be one *encyclopedia* and more than one *volume*, and the cloud would be one *cloud* and many *droplets*. It's not one cloud and more than one cloud or one droplet and more than one droplet, so no contradiction is involved.

CAI bears some similarity to Compositional Nihilism, even if its proponents are committed to the denial of Nihilism. If the world were ultimately composed entirely of simples, then both CAI and Compositional Nihilism would deny that there is anything fundamental *over and above* the simples. However, CAI does not force us to say that wholes are *less* fundamental than their parts; instead, we can say that the composite things exist and are not wholly grounded *in* any of their proper parts, since they are simply identical *to* their proper parts (taken collectively). Thus, CAI theory is consistent with Compositional Equivalence (22.1A.1T.1), as we discussed in the last chapter. Wholes are not posterior to their parts, but they are nothing over and above their parts, either. In addition, CAI is consistent, and Compositional Anti-Realism is not, with the Existence of Mereological Gunk (22.4A).

## 23.3 Accounting for the Correct Principles of Mereology

Can any of the proposed answers to the General Composition Question provide a basis for a plausible theory about the correct principles of mereology? Can any account for the essential features of the *parthood* relation? Five of the most interesting candidate features are Anti-Symmetry, Transitivity, Weak Supplementation, Strong Supplementation, and Arbitrary Sums, repeated here for convenience:

**(MA2) Mereological Anti-Symmetry.** If $x$ is a part of $y$, and $y$ is a part of $x$, then $x = y$.

**(MA3) Mereological Transitivity.** If $x$ is a part of $y$, and $y$ is a part of $z$, then $x$ is a part of $z$.

**(MA4) Weak Supplementation.** If $x$ is a proper part of $y$, then $y$ has some part $z$ that does not overlap $x$.

**(MA5) Strong Supplementation.** If $x$ is not a part of $y$, then $x$ has some part $z$ that does not overlap $y$.

**(MA6) Arbitrary Sums.** If $S$ is a non-empty set, then there exists at least one sum of $S$.

These principles are widely, but not universally, accepted by metaphysical mereologists. Metaphysicians have three ways to go here. (1) They could simply deny that these principles hold without exception. (2) They could admit that they hold, but suppose that their holding is a brute, inexplicable fact. Or (3) they could explain their holding by appealing to other principles and assumptions that are already part of their theory. Obviously, (3) would be the best response, although a theory's being able to endorse response (3) would not necessarily be a decisive reason for preferring that theory to its competitors.

TRANSITIVITY Transitivity can be explained by Moderate Compositional Anti-Realists, whether they embrace the Arbitrary or the Non-Arbitrary version of Compositional Reductionism. If $y$ and $z$ are composite, then in speaking of them we are really speaking of their composite simples. Let's say that the composite simples that make up $y$ are the $u$'s, and the simples that make up $z$ are the $v$'s. Then, to say that $x$ is part of $y$ is to say either that $x$ is one of the $u$'s, or that $x$ is composed of some of the $u$'s. To say that $y$ is part of $z$ is simply to say that all of the $u$'s are also $v$'s. Therefore, $x$ is one of the $v$'s or is composed entirely of $v$'s, and so $x$ is also part of $z$.

Somewhat surprisingly, Composition as Identity cannot provide a similarly elegant, logical explanation of Transitivity. The problem is that CAI theory does not entail that composite things are wholly composed of simples. All we can say is the following:

(A) $x$ is part of $y$ if and only if there are some $u$'s such that: $y$ is identical to the $u$'s, and $x$ is one of the $u$'s.
(B) $y$ is part of $z$ if and only if there are some $v$'s such that: $z$ is identical to the $v$'s, and $y$ is one of the $v$'s.
(C) $x$ is part of $z$ if and only if there are some $w$'s such that: $z$ is identical to the $w$'s, and $x$ is one of the $w$'s.

Let's suppose that $x$ is part of $y$ and $y$ is part of $z$. So, $y$ is among the $v$'s such that $z$ is the $v$'s, and $x$ is among the $u$'s such that $y$ is the $u$'s. To prove that $x$ is part of $z$, we would have to prove either that $x$ is one of the $v$'s or that there are some $w$'s such that $x$ is among the $w$'s and $z$ is identical to the $w$'s. However, neither of these follows by logic alone. Thus, CAI theorists must accept Transitivity as an additional law of metaphysics, as do those who accept Fundamental or Natural Parthood.

To be precise, CAI theorists must adopt an axiom somewhat weaker than Transitivity itself in order to be able to derive Transitivity as a result. In particular, they must assume Plural Expansion:

**Plural Expansion.** If the members of set $A$ are jointly (or plurally) identical to the members of set $B$, then the members of the union of $A$ and $C$ are jointly identical to the members of the union of $B$ and $C$.

Plural Expansion is a plausible principle, but it is not in any sense a theorem of logic, not even of Boolos's plural logic. At first glance, this claim is wrong: here's how you might try to derive Plural Expansion by logic alone. If the members of $A$ are identical to the members of $B$, then each member of $A$ is a member of $B$, and vice versa. So, by the extensionality axiom for sets, $A = B$. But then the union of $A$ and C is identical to the union of $B$ and $C$, and so their members must also be identical. However, this proof is fallacious, since it confuses two different forms of identity that can link pluralities. There is a definable form of identity, call it member-by-member identity or $=_{\text{M}}$, and it is true that whenever the $x$'s $=_{\text{M}}$ the $y$'s, then each of the $x$'s is one of the $y$'s, and vice versa. However, Composition as identity insists that primitive, indefinable identity holds, not only between individuals, but also between individuals and pluralities, and, by transitivity, between pluralities. Let's call this collective identity, whenever it holds between two pluralities. For CAI, it is possible for the members of a set $A$ to be collectively identical to the members of set $B$, even though set $A$ does not have the same members (considered individually) as $B$ does. For example, suppose $x$ has two proper parts, $y$ and $z$. Then, according to CAI, the members of the set $\{x\}$ are collectively identical to the members of the set $\{y, z\}$, even though $\{x\} \neq \{y, z\}$.

If we assume Plural Expansion, then Transitivity follows. Assume that $x$ is a part of $y$ and $y$ is a part of $z$. Then there is some set $A$ such that $x$ is a member of $A$ and $y$ is identical to the members of $A$ (taken collectively), and there is some set $B$ such that $y$ is a member of $B$ and $z$ is identical to the members of $B$, taken collectively. Obviously, $y$ is identical to the members of the set $\{y\}$, and so the members of $A$ are identical to the members of $\{y\}$. By Plural Expansion, the members of the union of $A$ and $B$ are jointly identical to the members of the union of $\{y\}$ and $B$, taken collectively. Since $y$ is a member of $B$, the union of $\{y\}$ and $B$ is identical to $B$. So, the members of the union of $A$ and $B$ are jointly identical to the members of $B$. Thus, the members of the union of $A$ and $B$ (taken collectively) are jointly identical to $z$. Since $x$ is a member of $A$, $x$ is a part of $z$.

Alternatively, the CAI theorist could try to build transitivity into their definition of the part-whole relation. For example, suppose that CAI theorists defined a 'completely plurality' in the following way:

**Definition of Completed plurality (1).** The $x$'s are a *completed plurality* if and only if for every $y$ such that $y$ is one of the $x$'s, if $y =$ the $z$'s, then each of the $z$'s are one of the $x$'s.

CAI theorists can now define the part-whole and the composition relation in this way:

**CAI 1.1 Composition.** The *x*'s compose *y* if and only if *y* = the *z*'s, the *z*'s are a completed plurality, each of the *x*'s is one of the *z*'s, and the *z*'s are the smallest completed plurality that includes each of the *x*'s.

**CAI 1.2 Part-whole.** *x* is a part of *y* if and only if *y* = the *z*'s, the z's are a completed plurality, and *x* is one of the *z*'s.

Now we can prove that if *x* is part of *y*, and *y* is part of *z*, then *x* is part of *z*. The definition of part-of now insures that *z* is identical to some *w*'s such that the *w*'s are a completed totality. This guarantees that if *y* is one of the *w*'s, and *x* is part of *y*, then *x* is also one of the *w*'s, and so also a part of *z*.

Turning finally to Aristotelian Compositional Reductionism, we see that such Aristotelians can explain the transitivity of parthood by appealing to the transitivity of the three components of their reduction: the grounding relation, the causal explanation relation, and the ontological dependence relation (see Chapter 3 on grounding, and especially 3.3 on ontological dependence).

ANTI-SYMMETRY Anti-Symmetry is also easy to explain for Compositional Reductionists, whether of the Arbitrary or Non-Arbitrary variety. Where we talk of two composite things, there are really just two pluralities of simples, the *x*'s and the *y*'s. One composite thing is part of another just in case the simples corresponding to the first are included in the simples corresponding to the second. If each composite is a part of the other, then we are really speaking about exactly the same simples in both cases. Hence, symmetrical *parthood* entails identity.

Advocates of Fundamental or Natural Parthood and of Composition as Identity can offer no interesting explanation for Anti-Symmetry. They must suppose it to be a metaphysically brute fact. It may be a little surprising that CAI cannot derive Anti-Symmetry from the laws of logic. If *x* is a part of *y*, and *y* is a part of *x*, then *x* must be identical to some things, the *z*'s, and *y* must be identical to some things, the *w*'s, such that *x* is one of the *z*'s and *y* is one of the *w*'s. Still, we cannot use logic alone to prove that all of the *w*'s are *z*'s, or that all the *z*'s are *w*'s, and so we cannot prove that *x* and *y* are identical.

However, Strong CAI (option 3) can explain Anti-Symmetry if it includes Plural Expansion. Suppose that *x* is a part of *y* and *y* is a part of *x*. Then *x* is a member of *A* and *y* is identical to the members of *A*, taken together, and *y* is a member of *B* and *x* is identical to the members of *B*, taken together. Since *x* is a member of $\{x\}$, we know that the members of *B* are collectively identical to the members of $\{x\}$. Consequently, by Plural Expansion, the members of the union of *A* and *B* are collectively identical to the union of the members of *A* and $\{x\}$. Since *x* is a member of *A*, the union of *A* and $\{x\}$ is simply *A*. So, the members of the union of *A* and *B* are collectively identical to the members of *A*, that is, to *y*. By a similar argument we can show that the members of the union of *A* and *B* are identical to *x*. By the transitivity of identity, *x* is identical to *y*.

Once again, however, we should note that Plural Expansion is a substantive assumption and not itself derivable by logic alone.

What if we use instead the definitions of composition and part-of in terms of completed totalities (CAI 1.1 and CAI 1.2)? If $x$ is a part of $y$, and $y$ is a part of $x$, then both $x$ and $y$ must be completed pluralities. So, let's say that $x =$ the $z$'s and $y =$ the $w$'s. To show that $x = y$, we must show that each of the $z$'s is one of the $w$'s, and each of the $w$'s is one of the $z$'s. But this follows immediately from transitivity.

For Aristotelian Compositional Reductionists, the anti-symmetry of the part-whole relation derives from the more fundamental asymmetry of the proper part relation, and that asymmetry in turn derives from the still more fundamental asymmetry of the grounding relation. It is impossible for all of the power-conferring properties of $x$ to be wholly grounded in the power-conferring properties of $y$, while the power-conferring properties of $y$ are also wholly grounded in the power-conferring properties of $x$, which is what would be required if $x$ and $y$ were each proper parts of the other.

ARBITRARY SUMS As we mentioned above, Arbitrary Compositional Reductionism is designed to validate Arbitrary Sums. In contrast, many mereologists who accept Fundamental or Natural Parthood simply deny Arbitrary Sums.

What about Composition as Identity? Some, like Lewis (1991) and Sider (2007), have assumed that CAI automatically validates Arbitrary Sums. It is true that if $S$ has some members, then there is something (a *whole*) that is identical to those members, and of which each of those members is a part. However, is this whole a *sum* of that set in the precise, Leśniewski sense? Recall the definition of sums:

**Def D23.3 Sum.** $x$ is a *sum* of set $S$ if and only if $S$ is a non-empty set, and for all $z$, $z$ overlaps $x$ if and only if $z$ overlaps one of the members of $S$.

Let W($S$) be the whole that is identical to the members of $S$. According to CAI, there is such a whole. The members of $S$ are all parts of W($S$). Is W($S$) a sum of $S$? Is it the case that, for all $z$, $z$ overlaps W($S$) if and only if $z$ overlaps one of the members of $S$? Given Transitivity, the right to left direction is no problem. If $z$ overlaps one of the members of $S$, namely $u$, then $z$ and $u$ have a common part, $v$. Since $u$ is a member of $S$, $u$ counts as a part of W($S$), and so $z$ has a part that is a part of one of the parts of W($S$), and thus $z$ overlaps W($S$).

However, the other direction is problematic. Suppose $z$ overlaps W($S$). Is there any way to prove, by logic and CAI alone, that $z$ overlaps some member of $S$? In a word, No. If $z$ overlaps W($S$), then they have some common part $u$. That is, $u$ is part of $z$ and a part of W($S$). However, just because $u$ is part of W($S$), it doesn't follow that $u$ is a member of the set $S$.

Could we solve this by re-introducing the notion of a completed plurality? Our first definition of completed plurality took care of Transitivity, but it won't help with the left-to-right direction of the proof that W($S$) is a sum of $S$. We can, however, fix this by replacing W($S$), the plurality of members of $S$, with the *completion* of $S$, C($S$).

**Definition of CAI completion.** The completion C($S$) = the $z$'s if and only if the $z$'s are the smallest completed plurality that includes all of the members of $S$.

We can now prove that the completion of any set $S$ is a sum of the members of that set, so long as each member of $S$ is either an atom or a completed plurality. Suppose that *z* overlaps C(*S*). Then they have some common part *u*. That is, *u* is part of *z* and a part of C(*S*). Since C(*S*) is the smallest completed plurality including the members of *S*, this means that u is either a member of *S*, or a member of a member of *S*, and so on. Since each member of *S* is either an atom or a completed totality, it follows that *u* will be a member of some member of *S*.

Therefore, CAI theorists can claim a restricted version of Arbitrary Sums: every set *S* of atoms and completed pluralities has a sum, namely, the CAI completion of *S*. However, as we shall see, CAI theory has difficulty validating the axioms of Weak and Strong Supplemenation, which calls into question whether these CAI "sums" are really sums at all.

Non-Arbitrary Compositional Reductionists, including Aristotelian Reductionists, will simply reject the axiom of Arbitrary Sums, opting for a much more restrictive principle of composition.

WEAK SUPPLEMENTATION To consider the grounds of Weak Supplementation (MA4) raises the question of why something cannot have only a single *proper* part. What is it about wholes that requires a *multiplicity* of disjoint proper parts? Arbitrary Compositional Reductionists have an obvious answer to this question, since they translate '*x* is a proper part of *y*' into 'The *x*'s are some of but not all of the *y*'s.' For example, to say that the giraffe's neck is a proper part of the giraffe is to say that the particles arranged giraffe's-neckwise are some but not all of the particles arranged giraffe-wise. From this translation, Weak Supplementation follows as a matter of simple logic: these (the *x*'s) can be some but not all of those (the *y*'s) only if there are some other things (the *z*'s) that are not among these (the *x*'s) but are among those (the *y*'s).

Non-Arbitrary Compositional Reductionists can make use of the same explanation, on the plausible assumption that if the *y*'s cohere together, and the *x*'s are among the *y*'s, then the *x*'s cohere together as well. In other words, in order to explain Weak Supplementation, Non-Arbitrary Compositional Reductionists might also want to accept Arbitrary Parts.

Aristotelian Compositional Reductionists can explain Weak Supplementation in a different way, however. Their definition of proper parthood requires multiple, metaphysically independent entities to cooperate in some process that generates and sustains the composite entity (see Koons 2014b for more detail).

Advocates of Fundamental or Natural Parthood and of Composition as Identity must appeal to some primitive and inexplicable fact about the *parthood* relation or the universal PARTHOOD or *parthood* tropes in order to ground Weak Supplementation, or else they must simply deny Weak Supplementation altogether. Some metaphysicians have in fact denied Weak Supplementation, supposing that in some cases a thing can have a proper part that includes (or at least) overlaps all of its other proper parts. For example, Alfred North Whitehead supposed that a solid sphere has its interior (the sphere minus its outer surface) as a proper part, but he denied that the sphere had any other proper parts that were not parts of the interior, since he refused to count the surface of the sphere as a real entity in its own right. Here's another example. Classical Bundle Theorists (9.1T.1T.1A) might suppose that each universal in a bundle is a proper part of a part of the bundle (a trope-like part) that contains only that universal as a part. The universal and its trope-like container would be different, the universal a part of the container, even though the universal is the container's only proper part.

What about Composition as Identity? Suppose that $x$ is a part of $y$, and $x \neq y$. Does it follow that $y$ has a part that has no part in common with $x$? This doesn't follow from CAI plus plural logic alone. If $x$ is part of $y$, then $y =$ the $z$'s, and $x$ is one of the $z$'s. If $x$ is not identical to $y$, then (given Anti-Symmetry), it must be the case that $y$ is not a part of $x$. So, either $x$ is an atom, or $x =$ some $w$'s, and $y$ is not one of the $w$'s. This tells us nothing about other parts of $y$, much less parts of $y$ that have no part in common with $x$.

Does it help if we assume Plural Expansion? We don't think so. What about our proposed definitions of composition and part-of in terms of completed pluralities? No, those won't help either. Thus, CAI theory fails to provide an independent account of the essential principles of mereology. We will consider one more attempted solution to this problem: a solution that, if it worked, would simultaneously support both Weak and Strong Supplementation.

STRONG SUPPLEMENTATION Strong Supplementation raises a different question: why can't two distinct wholes share the same proper parts? Strong Supplementation is equivalent to the General Principle of Constituent Identity, the most general version of the various forms of PCI that we considered in Chapter 9:

**PMeta 4 General Principle of Constituent Identity.** If $x$ and $y$ both have proper parts, and they have exactly the same proper parts, then they must be identical.

If we accept Strong Supplementation, then we don't have to worry about Weak Supplementation, since Strong Supplementation, Reflexivity, Anti-Symmetry, and Transitivity together entail Weak Supplementation.

Once again, Arbitrary Compositional Reductionsists have a ready answer. The $x$'s are different from the $y$'s only if something is among the $x$'s but not among the $y$'s or vice versa. The other options must either reject PCI for Composite Objects or accept it as a further, brute metaphysical fact.

Those who reject General PCI (and Strong Supplementation) face two serious problems, one metaphysical and one epistemological. These problems concern the number of wholes composed by a given set of constituents. If the members of set $S$ compose at least one thing, how many do they compose? If we reject General PCI, we cannot say, in all cases: *exactly one*. In some cases, the members of a set might compose two or three or a billion or even an infinite number of distinct composite entities, each having exactly the same proper parts as all the others. It would be like supposing that the atoms in my favorite tie compose not just one tie but exactly 17 ties. The metaphysical question is this: what accounts for the fact that the things in set $S$ compose exactly $n$ composites and not $n + 1$ or $n - 1$?

The epistemological puzzle is similar. If any number of distinct composites can have exactly the same parts at the same time, how could we ever tell how many of them there are? The most plausible answer to this question would be to appeal to relational structures as the ground for duplication (see Section 10.3 on structural universals).

As explained in Koons 2014b, Aristotelian Compositional Reductionists can account for Strong Supplementation by appealing to the fact that the compositional structure of an Aristotelian universe will be linear: if $x$ is not a part of $y$, and $y$ is not a part of $x$, then $x$ and $y$ will not overlap at all. If two wholes did partially overlap each other, the entity constituting the overlap would have causal powers that are wholly grounded in two separate, independent sets of facts (facts about the two distinct wholes), which would happen only very rarely, if ever.

Let's turn one last time to Composition as Identity. Is there a way to repair CAI in order to provide a foundation for Strong Supplementation (and, thereby, for Weak Supplementation as well)?

We could try modifying the definition of completed plurality in this way:

**Definition of Completed plurality (2).** The $x$'s are a *completed plurality* if and only if

(i) for every $y$ such that $y$ is one of the $x$'s, if $y =$ the $z$'s, then each of the $z$'s are one of the $x$'s, and
(ii) for every $y$ such that there are some $z$'s, $y =$ the $z$'s, the $z$'s are a completed plurality, and each of the $z$'s overlaps the $x$'s, then $y$ is one of the $x$'s.

If this were a workable definition, and if we defined composition and part-of in terms of completed pluralities (CAI 1.1 and 1.2), then we could show that Strong Supplementation would be valid. Clause (ii) would simply guarantee that any completed plurality $x$ all of whose parts overlap some completed plurality $y$ would have to be a part of $y$, which is exactly equivalent to Strong Supplementation.

However, there is an obvious problem with this definition: it is viciously circular. We have to include the phrase 'completed plurality' in the very definition of that phrase, since we are restricting the part-of relation to wholes that are themselves completed pluralities (by CAI 1.2).[2]

The only way around this problem is to make the definition *recursive*: which means that one must stipulate one or more *base cases* of completed plurality (such as: every pair of atoms is a completed plurality), and then use repeated applications of the definition above to generate additional cases. However, this will only work on the assumption that every composite thing is built up recursively from atomic parts, and so CAI theory would simply collapse into Reductive Compositional Anti-Realism.

## 23.4 Parthood and Truthmaking

Finally, let's turn to the question of the relationship between *parthood* and truthmaking. If $x$ is a part of $y$, what is the truthmaker for this truth? Much turns on the essences of $x$ and $y$. Is $x$ essentially a part of $y$ or is $y$ essentially an encompasser of $x$ or is neither the case? Let's define a 'constant' whole as one that contains all of its parts essentially and a 'rigid part' of something as something that is essentially a part of that whole:

**Def D23.5 Constant Wholes.** $x$ is a *constant* whole if and only if for all $y$, if $y$ is possibly a part of $x$, then necessarily, if $x$ exists then $y$ is a part of $x$.

**Def D23.6 Rigid Parthood.** $x$ is a *rigid part of* $y$ if and only if necessarily, if $x$ exists, then $x$ is a proper part of $y$.

In the case of either constancy or rigidity, the *parthood* truths have a simple truthmaker. If $x$ is a constant whole and $y$ is a proper part of $x$, then $x$ itself is the classical truthmaker for $y$'s being a part of $x$. Similarly, if $x$ is rigidly a part of $y$, then $x$ itself is the classical truthmaker for $x$'s being a proper part of $y$.

Suppose that $x$ is a proper part of $y$ but not a rigid part, and $y$ is not a constant whole. What then could the truthmaker be? It would seem that in such a case the *parthood* relation between $x$ and $y$ must be grounded in some further facts about $x$ and $y$. In other words, it seems that we cannot then have a case of fundamental *parthood*. It seems that the fundamentality of *parthood* requires constancy or rigidity of some kind:

**Fundamentality Entails Constancy/Rigidity.** If it is a fundamental fact that $x$ is a proper part of $y$, then either $y$ is a constant whole or $x$ is a rigid part of $y$.

We will return to the questions of mereological constancy and rigidity in Section 25.1, on persistence though a change of parts.

## Notes

1 Of course, it can't be a *fundamental fact* about these fundamental things that they are *simple* (i.e., lacking in parts), since all facts about parts (including negative facts, like being *simple* or *atomic*), must be merely derived facts. Still, it seems to be coherent to suppose that there are fundamental things that are only derivatively atomic, since the proposed reduction of the part-of relation assigns them no parts.

2 If we were to try to solve this problem by using a weaker restriction on the pluralities in clause (ii), by, for example, using the first definition of 'completed plurality', we would then run into problems of cardinality, i.e., problems with the absolute size of the pluralities involved. A weaker condition of that kind would in effect require a completed plurality to contain a distinct element for each of its non-empty, proper sub-pluralities, which is mathematically impossible for any plurality of greater than two elements (unless the members of the plurality form what mathematicians call a 'proper class', a fantastically large plurality).

# 24

# Change and Persistence

In the last couple chapters, we looked at the issues surrounding composite things: when do some things compose a whole, and what is composition? In this chapter and the next, we examine questions having to do with whether and how things persist through change. Is persistence possible? If things do persist, how do they do so? Are persisting things a kind of whole, made up of merely momentary pieces? If so, what unifies such instantaneous things into a single unified life? The issues in these chapters have profound ethical and existential implications, since we ourselves are among the world's persisting things. Is our identity over time something fundamental, or is it reducible to more basic facts? Is each of us truly one thing that exists through many moments, or a multitude of temporary things? Maybe more basically, is each of us really just a single thing or are we really a multitude of more or less coincident things?

Two principal issues have occupied philosophers in this area. First, can objects persist through a change of properties, and if so, how? Second, can objects persist through a change of parts, and if so, how? We will take up the second issue in Chapter 25. With respect to the first issue, we will first consider (in Section 24.1) the question of whether things can change intrinsically *at all*. This is the dispute between Staticism and Kineticism. Next, assuming that intrinsic change does take place, we will examine in Section 24.2 two principal views about how things persist through change of intrinsic *properties*, Substratism and Replacementism. Substratists think that persisting things are fundamental, that one and the same fundamental object can have different properties at different times. Replacementists think that persisting things are not fundamental, that they are composed of a series of instantaneous *temporal parts*, and that a persisting thing has different properties at different times by having different parts that have those properties once and for all. Our final section on change (Section 24.3) will focus on the specific but very important case of motion, or change of *location*. Here there are three major theories: Intrinsic Motion, according to which motion is intrinsic at each moment to the

*The Atlas of Reality: A Comprehensive Guide to Metaphysics*, First Edition.
Robert C. Koons and Timothy H. Pickavance.

moving thing; Bertrand Russell's At/At Theory, which defines motion simply as change in position over time; and an Aristotelian theory, Motion Intervalism, according to which motion is intrinsic to temporally extended processes.

## 24.1 Does Anything Change? Does Anything Persist?

Our experience of the world seems to present us with things that both change and persist, where a thing persists if it exists at more than one time. For example, objects in the external world, like people and dogs and computers and rocks, seem able to change and persist. When one reflects on oneself, one is also led to the view that one is capable of both change and persistence. For example, each of us seems to remember past actions and events in which he or she—the very person doing the remembering—once participated. Indeed, even our internal conscious states change and persist. Importantly, there are really two issues here, one having to do with change and another having to do with persistence through time. These issues are intimately interconnected, but for the most part we will focus on change. As we will see, views about change tend to drive views about persistence. At any rate, it is plausible to think that both Kineticism and Persistence are true.

**24.1T Kineticism.** Some things change.

**24.1A Staticism.** Nothing changes.

**24.2T Persistence.** Something persists (has existed or will exist at more than one instant).

Staticism commits us to counting the appearances noted above as illusory or misleading. One version of Staticism would be Solipsism of the Present Moment (20.1T.1A), according to which only one thing exists (one's own consciousness) and only for an instant. On this view, all memories of a so-called "past" and expectations for a "future" are simply illusory. Since we've already discussed this view in Chapters 13 and 20, we here assume that many things exist, and that there is some kind of real distinction among past, present, and future.

How then, could one deny the reality of change? Clearly, different things happened in the past from what is happening now; dinosaurs once roamed the earth but do so no longer. And many things will happen in the future that have not happened yet, including, perhaps, human colonies on Mars. This is a question that will occupy us.

To get at the real issue, we need to distinguish *intrinsic* change from mere *Cambridge* change. It is easier to give examples than to provide a clear and non-circular definition of these categories. When something changes its color or temperature, this seems to amount to some change in its intrinsic character or way-of-being. In contrast, when a woman's husband dies, her change in status from wife to widow need, in and of itself, involve no such change (of intrinsic character). The wife might not become aware of the husband's death for some time, and the same might be true of everyone with whom she interacts.

**Def D24.1. Mere Cambridge Change.** Something undergoes a *mere Cambridge change* when it changes without changing intrinsically.

**24.1T.1T Strict Kineticism.** Some things change intrinsically.

**24.1T.1A Moderate Staticism.** Some things undergo mere Cambridge change, but nothing changes intrinsically.

Democritus, Empedocles, and Descartes (if we focus on the physical world) advocated Moderate Staticism. All change consists of motion or, more precisely, locomotion: change in position. Locomotion, in turn, can be understood in two different ways: (1) as change in absolute position, that is, as moving from one place to another, or (2) as change in relative position, in a thing's distance from other things.

Moderate Staticists must deny that there are composite things, things composed of atoms (like organic bodies or solar systems), since if there were such things, and if their constituent atoms were to move with respect to one another, then there would be intrinsic change (at the level of composite entities) after all. That is, Moderate Staticists must embrace Mereological Nihilism (22.6T.1T, 23.1A). They must also accept Spatial Relationism (17.1A), since if places exist, then locomotion would seem to involve either an intrinsic change in the places or in the place-occupiers. However, if spatial position consists entirely of relations to other things, and there are no composite things, then there would be absolutely nothing for locomotion to be intrinsic to. If a change occurs but is not intrinsic to anything, it would seem to be utterly unreal. Hence, Moderate Staticism seems an incoherent position.

In addition, modern physics seems committed to the existence of intrinsic change. For example, some fundamental particles, like electrons and positrons, can undergo events of creation and annihilation. If an electron meets a positron, both particles are instantly annihilated, creating one or more new particles. Creation and annihilation are certainly cases of intrinsic change!

Shortly, we will move to a discussion of two types of intrinsic change, change of property and change of part. First, though, we want to touch on one other issue about change.

IS THERE AN ENDURING SUBJECT OF CHANGE? One important, and of late somewhat under-appreciated, dimension of change and persistence is whether change requires the existence of some enduring thing that *underlies* the change, something that exists both before and after the change, such that the intrinsic change is intrinsic to the enduring thing. To get at this issue, consider the following distinction between two kinds of change:

**Def D24.2 Qualitative Change.** A thing undergoes *qualitative change* if and only if it changes intrinsically and exists both before and after the change.

**Def D24.3 Substantial Change.** A thing undergoes *substantial change* if and only if it undergoes intrinsic change and either does not exist before the change (creation) or does not exist after the change (annihilation).

It may be possible that a single change is both qualitative and substantial. The qualitative change of one thing (the enduring subject) might just be the creation and destruction of two additional things (some modes or attributes of the subject). A change is a case of *purely substantial* change if and only if it does not consist in the qualitative change of anything whatsoever. Using this distinction, we can distinguish two views about whether all intrinsic change requires an enduring subject, two versions of Strict Kineticism:

**24.1T.1T.1T Enduring Substratism.** Fundamental things can undergo qualitative change, existing before, during and after the change.

**24.1T.1T.1A Replacementism.** It is possible for something fundamental to undergo substantial change but not qualitative change.

Replacementism asserts that the only kind of intrinsic change that is possible is purely substantial change, while Substratism affirms the possibility of qualitative change at the most fundamental level. If nothing *fundamental* persists through a change, then that change must consist in something's coming into existence (creation) or something's ceasing to exist (annihilation). Creation and annihilation are the two forms of substantial change. If a substantial change can occur even though nothing persists through the change, then it should be possible for everything in the world to be annihilated and replaced in an instant with a new set of things. Let's call such a global event 'Global Replacement'.

The possibility of Global Replacement follows from Replacementism assuming plausible patchwork principles. We have distinguished two patchwork principles:

**PMeta 5.1 Finite Patchwork.** If an event or process of (intrinsic) type $A$ is possible, as is an event or process of intrinsic type $B$, and if there is enough room in the history of the world to locate in it instances of both events (or processes) without overlap in time and space, then it is possible for events (or processes) of both types to occur together.

**PMeta 5.2 Infinite Patchwork.** If $T$ is a class of types of events or processes, and for each member of $T$, it is possible for an event or process of type $T$ to occur, and there is enough room in the history of the world to locate within it instances of each of the types in $T$ without overlap in space and time between the instances, then it is possible for all of the types in $T$ to be realized together.

The possibility of Global Replacement follows from an instance of the Infinite Patchwork Principle. If pure substantial change is possible in a single case, then it would be possible for such changes to fill all of space at the same time (since such changes wouldn't overlap in space and time).

Tensers (20.2T) can use patchwork principles to build a simple argument for the necessity of an enduring substrate of change:

1 Suppose that it was possible for some change to occur without any enduring thing.
2 If this could happen once somewhere, then it could happen always everywhere. (Finite/Infinite Patchwork)

3 So, if 1 and 2, then there could be a world containing time and change, but in which nothing persists for more than an instant.
4 But such a world would be nothing but a pointillist mosaic of qualities extended over four dimensions, that is, an Eternalist (20.2A.2T) "block" universe.
5 An Eternalist universe is one without any real change (as McTaggart argued) and so without time.
6 Consequently, 1 must be false: every change must be accompanied by something that persists through the period of change.

Anti-Tensers (20.2A) will reject premise 4. They could embrace a version of Russell's At/At Theory applied to change of properties, and they will deny that an Anti-Tensist, B-theoretic block universe is really devoid of change. So long as a thing has different properties at different times, they will say that there really is change in the world. Is there a version of the argument from Global Replacement that does not depend on Tensism?

Suppose the world did undergo Global Replacement. This possibility would pose certain difficult problems, even for Anti-Tensers. First, what would make it the case that what is being described is really a change in a single world? Why not say that one world, with its one time-line, has simply been annihilated, bringing its time to an end, while another, entirely separate world has been created *ex nihilo*, with its own time beginning to exist for the first time. What is it that makes the last moment of the first world identical to the first moment of the second world, in such a way that the two time series together compose a single time series? It's not obvious that one can answer these questions in a principled way, or even distinguish between the two possibilities.

Second, if we assume that there were many things existing before the Global Replacement event and many things existing afterward, how do we pair each entity that existed before the event with the entity which replaces it after the event? We have this *pairing problem*: which new entities have replaced which old entities?

One solution to both problems would be to introduce space and its parts as a further entity, with each entity susceptible to creation or annihilation occupying some part of space during each moment of its existence. The endurance of space itself would tie the two world histories together into a single history, and we could use the occupation of enduring places as the criterion for saying which new entities have replaced which old entities. New entities $x_1, \dots x_n$ have replaced old entities $y_1, \dots y_n$ just in case the place occupied collectively by the $x$'s after the change is identical to the place occupied by the $y$'s immediately before the change.

But now it seems that we have an enduring substrate after all, namely, space itself. What we thought was pure substantial change turns out to be identical to a kind of qualitative change in the parts of space, where that qualitative change is a change in what entities or sorts of entities occupy those places. Once again, it seems natural to think that an entity's occupying a place is an intrinsic feature of that place.

Even if one could resist the identification of the putative purely substantial changes with qualitative changes in the parts of space, one would still have to deny the possibility of Global Replacement. Space itself, on this view, is not created or annihilated. Indeed, this solution to our two problems *requires* the endurance of space, since only if space endures through the Global Replacement event can it provide the continuity needed to unify the two time-lines and pair the created entities with their respective annihilated ones.

Another solution to this pair of problems is to have recourse to relations of cause and effect. Suppose that the annihilation of entity *a* is a cause of the simultaneous creation of entity *b*. These causal links could account for the unity of the two halves of world history (before and after the Global Replacement threshold). They do this because there would be a time at which both the entities existing before the Global Replacement event and the entities existing after it exist. In addition, causal links could solve the pairing problem: entities $x_1, \ldots x_n$ are replaced by entities $y_1, \ldots y_n$ just in case the annihilation of the *x*'s is causally linked to the creation of the *y*'s.

However, the causal theory will have to survive another thought-experiment. If Global Replacement were possible, then it would be possible for there to be a stretch of world history during which global annihilation and re-creation occurs at every instant. Let's suppose that it is also possible that time is dense:

**Def D24.4 Density.** The time series is *dense* if and only if between any two moments of time there always exists a third.

The assumption that time is dense corresponds to Temporal Anti-Discretism (19.4T). If time is dense, then there is never a *next* moment. There will be an infinite number of intervening moments between any two moments. Let's suppose that the period of dense Global Replacements lasts for exactly one hour; call it the 'Interregnum'. We now face again the same two problems. (1) What unifies the history of the world before and after the Interregnum into a single history of a single world? (2) How are the things existing before the Interregnum paired with their post-Interregnum replacements? The sort of causal linkage appealed to above will no longer do the work, since we had to suppose that the causes and effects were always simultaneous—the new entity is created at the same moment that the old entity is annihilated, with the creation event linked causally to the simultaneous annihilation event. Simultaneous causal links, by definition, never bridge any span of time, so they cannot link an entity existing at one moment in the Interregnum with a successor existing at any later moment in the Interregnum. Simultaneous causal links can only link an entity annihilated at some time with an entity created *at that same time*, and during the Interregnum, no entity exists for more than a moment! We cannot solve the pairing problem. But we also cannot connect the world histories before and after the Interregnum, then, either. Simultaneous causal links will do nothing to connect one moment in the Interregnum with any other moment in the Interregnum so as to make them a continuous history of a single world because the way such causal links accomplished that in our earlier case was by linking entities that existed before and after the single Global Replacement event. Again, though, during the Interregnum, no entity exists for more than a moment, so this rationale is unavailable.

To deal with dense stretches of time, the causal theorist must introduce Wesley Salmon's notion (Salmon 1984, 1998) of a *real* or *causal process* (see Section 28.6.4). However, if the causal theory makes use of causal processes in place of causal links between discrete events, then once again we have something that endures through the supposed global substantial change, namely, the process itself. Therefore, we have good reason to believe that it is necessarily the case that something persists through all change.

However, there are two responses that we have not yet considered on behalf of Replacementism. First, Replacementists could appeal to irreducibly spatiotemporal relations, relations of the kind that are commonly used in relativistic physics. Physicists (when employing the theory of relativity) make use of such relations as *time-like separation*. Two events are time-like separated if it would have been possible to send a light signal from one event to the other. The events that are time-like separated from a given event from the forward and backward *light cones* of that event. If *time-like separation* is a fundamental relation between events, it could be used to solve the pairing problem through Global Replacement events, and even through the hypothesized Interregnum. Replacementists who rely on spatiotemporal relations to solve the pairing problem are Perdurantists.

If such spatiotemporal relations do not suffice, Replacementists have one last recourse: posit a primitive pairing relation, one that connects two events just in case they belong to the *same* enduring thing. Metaphysicians call this relation 'genidentity'. The event of Barack Obama's birth stands in the *genidentity* relation to the event of his inauguration as President. This version of Replacementism is Primitive Genidentity Theory.

How is Primitive Genidentity Theory different from Substratism? The difference is a subtle but important one. Substratists posit a new kind of fundamental entity, an entity that exists both before and after the intrinsic change. Primitive Genidentity Theorists deny that any such fundamental entity exists. All that exists are quantitively unchangeable fundamental entities, entities that can undergo creation and annihilation, but no other changes. Where Substratists posit a persisting fundamental thing, Primitive Genidentity Theorists posit a fundamental *relation* between things, the *genidentity* relation.

## 24.2 How Objects Change Properties: Substratism vs. Replacementism

We now turn to the question of how things change their intrinsic qualities. We start our discussion of intrinsic qualitative change by considering the paradox of intrinsic change. In his dissertation, Mark Johnston (1983, 1987) argued that there is something paradoxical about cases of intrinsic change. Suppose that some thing $x$ changes from having intrinsic property $P$ to having property *not-P* or vice versa. Let's say that $x$ is $P$ at $t_1$ and is not-$P$ at $t_2$. The paradox of instrinsic change is that one seems to be committed to the view that $x$ has incompatible properties. Since $x$ is $P$ at one time and not-$P$ at another, then $x$ is both $P$ and not-$P$ (albeit at different times). How could this be?

Some philosophers, reflecting on this issue, have inferred that objects have parts along the temporal dimension in something like the way they have parts along spatial dimensions. Just as, for example, a beach ball can be both red and green by having a (spatial) part that is red and a (spatial) part that is green, so a persisting apple can be wholly green at one time and wholly red at another by having a *temporal* part or a *time-slice* that is wholly red and that exists only at one time and another temporal part that is wholly green and that exists only at another time.

Temporal parts are relatively uncontroversial when it comes to processes and temporally extended events. A trip across the English Channel presumably has parts corresponding to the first and second half. An extended event, like the American Civil War, has parts corresponding to its temporal parts (1861, 1862, etc.). Likewise, baseball games have innings as temporal parts. However, some Substratists might deny that there really are such things as processes or events (at least, as fundamental entities). Talk about events could be replaced by talk of objects having properties at various instants or intervals. Anyway, temporal parts are an essential component in the theory of Replacementism.

**Def D24.5 Temporal Part.** A thing $x$ is a *temporal part* of $y$ if and only if $x$ is a part of $y$, and there is no $x'$ such that $x$ and $x'$ exist at exactly the same times, $x \neq x'$, and $x'$ is a part of both $y$ and $x$. That is, $x$ is a *maximal* part of $y$ occupying some temporal interval.

Usually, philosophers who commit to temporal parts at all go one step further by committing to Universal Instantaneous Temporal Parts:

**24.3T Universal Instantaneous Temporal Parts (UITP).** All persisting things have instantaneous temporal parts (time-slices), one corresponding to each instant during which they exist, and these time-slices are fundamental and not mere logical constructions.

Change can then be handled by claiming that it is fundamentally time-slices that instantiate intrinsic properties, and persisting things only have those properties derivatively, by having temporal parts (the time-slices) that have them non-derivatively.

Replacementists must believe in temporal parts, since anything that changes must have at least two temporal parts, one before and one after the change. Substratists have the option of either embracing or rejecting temporal parts, including time-slices. Historically, most Replacementists have affirmed not only the existence of some temporal parts but also UITP. These Replacementists are Classical Perdurantists and Classical Genidentity Theorists:

**24.1T.1T.1A.1T Classical Perdurantism.** The only fundamental things exist only for an instant, and the only fundamental relations between such time-slices are spatiotemporal relations.

**24.1T.1T.1A.1A Classical Genidentity Theory.** The only fundamental things exist only for an instant, and there is one fundamental relation (primitive *genidentity*) that is not spatiotemporal.

Substratists can also be divided into two classes, depending on whether they affirm the existence of temporal parts, including time-slices. Historically, most Substratists have denied that fundamental and changeable things have temporal parts at all, whether instantaneous time-slices or extended temporal parts. Since most philosophers who believe in temporal parts believe in instantaneous time-slices of persisting things, we can focus on whether Substratists deny the existence of instantaneous time-slices. Those who deny that changing things have any temporal parts are traditionally called Endurantists—we will label them 'Classical Endurantists'. Substratists who accept UITP are Time-Slice Substratists.

**24.1T.1T.1T.1T Classical Endurantism.** Fundamental things can undergo qualitative change, existing before, during, and after the change, but they do not have any fundamental temporal parts.

**24.1T.1T.1T.1A Time-Slice Substratism.** Fundamental things can undergo qualitative change, existing before, during, and after the change, and they have instantaneous temporal parts corresponding to each instant during which they exist.

Classical Endurantists can embrace the existence of temporal parts, so long as they do not maintain that those temporal parts are fundamental entities. (This runs against the traditional use of the label, but maintains, we think, the spirit of Endurantism.) Thus, Classical Endurantists, by definition, commit to the Absence of Temporal Parts:

**24.3A Absence of Temporal Parts.** Some persisting thing $x$ and some instant of time $t$ are such that $x$ exists at $t$ but has no instantaneous part at $t$ that is fundamental and not a mere logical construction.

However, Time-Slice Substratists do believe in fundamental instantaneous temporal parts. Time-Slice Substratists suppose that instantaneous parts and persisting wholes are both metaphysically fundamental and necessarily interdependent, in the same way that Volume-Boundary Dualists (18.1A.2A) believe that both three-dimensional regions and their zero-, one-, and two-dimensional boundaries are both fundamental and interdependent. However, such a view comes at some cost, from the point of view of Ockham's Razor (PMeth 1.4.1), since it posits two classes of fundamental entities, both persisting and instantaneous, instead of just one.

Incidentally, the issue between UITP and the Absence of Temporal Parts has a close parallel with the issue discussed in Chapter 19 concerning the existence of the instants themselves, namely, the dispute between Strong Intervalism (19.1A.1T) and Moderate Intervalism or Interval-Instant Dualism (19.1A.1A). If one accepts Strong Intervalism, then one would also have to accept the Absence of Temporal Parts and reject UITP. If we were to accept UITP, we could simply identify each instant with a whole composed of all of the instantaneous temporal parts that are simultaneous with any one such part.

However, one could embrace Moderate Intervalism, with its fundamental instants, while rejecting UITP. Just because a persisting thing exists at an instant, it isn't obvious that each thing has a real part that exists at and only at that instant. It seems natural to say that each persisting thing is *wholly present* at each moment during which it exists. Thus, Substratists who are Moderate Intervalists embrace either Endurantism or Time-Slice Substratism.

Now, there is an important question that Endurantists face at this point: how can they respond to the paradox of intrinsic change? First, Endurantists can opt for Presentism (20.2T.4). This does the trick because one will not be able to move from the fact that $x$ *was* $F$ and the fact that $x$ *is* (presently) not-$F$ to the fact that $x$ *is* (tenselessly) $F$ and not-$F$. This is because according to Presentism there is just no sense in which it follows from the fact that something was $F$ that it is true to say that it is (tenselessly) $F$.[1] As we have seen, Presentism is a controversial position, and the problems with taking this route to salvage Endurantism are just the problems with Presentism.

Second, Classical Endurantists might say that intrinsic properties are really relations between things and times or that the *exemplification* relation is time-indexed. In other words, one can say that when something $x$ is $F$ at a time $t$, what that means is that $x$ is (tenselessly) $F$-related to $t$ or that $x$ is-a-$t$ $F$. This does the trick because being $F$-related to $t$ and not-$F$-related to $t'$ is not paradoxical; likewise, being-at-$t$ $F$ and being-at-$t'$ not-$F$ is not paradoxical. What would be paradoxical is being $F$-related and not-$F$-related to the same time $t$ or being-at-$t$ to both $F$ and not-$F$.

Opponents of Classical Endurantism (like David Lewis 1986a) have argued that this view is dissatisfying because it follows that intrinsic properties aren't really intrinsic after all. Rather, intrinsic properties are relational or only *had-at-a-time*. For example, suppose that a bar is hot and straight at $t_1$ and cold and bent at $t_2$. *Heat*, *coldness*, *straightness*, and *bentness* all seem to be paradigms of intrinsic properties.[2] If these aren't intrinsic to the bar, all by itself, what is the bar like intrinsically? It would seem that the bar has *no* properties intrinsically. However, this option asks us to believe that *bentness* is a relation between the bar and various instants of time. The bar fundamentally bears the *bentness* relation to $t_2$ and not to $t_1$. How decisive is this objection? Some Classical Endurantist Anti-Tensers, like Peter van Inwagen (1990b), find it utterly untroubling. Perhaps metaphysics reveals that changing things have thinner intrinsic natures than we might originally have thought. Still, the result seems to come at some cost by way of overturning deeply held beliefs of common sense.

Time-Slice Substratists have a third option for explaining intrinsic change. They can appeal, just as Replacementists do, to the qualitative difference between two temporal parts of the changing thing. That is, Time-Slice Substratists can take the fundamental bearer of the changeable or *accidental* features of the changing thing to be temporal parts of that thing. To change from cold to hot is to have a cold part located at one period of time and a hot part located at an immediately later period of time. In order to distinguish such Time-Slice Substratism from Replacementism, we have to assume that according to the Time-Slice Substratists, some fundamental substrate exists both before and after the change and that two temporal parts are somehow dependent on this substrate for their existence. In order to distinguish Time-Slice Substratism from Classical Endurantism, we have to assume that the relation between a thing and its temporal parts is supposed by the Time-Slice Substratists to be timeless, eternal, and unchanging. The Time-Slice Substratists can agree with common sense in taking *size*, *shape*, *temperature*, and the other changeable properties to be one-place qualities (of temporal parts) and not as relations between objects and times.

What positive reasons are there for believing in either Enduring Substratism or Replacementism? The relative attractiveness of the two views depends a great deal on one's position on other issues, including the structure of space and time, whether tense is fundamental, and the fundamentality of powers, laws, and conditionals. We look briefly at some of arguments for each of the two positions.

SUPPORT FOR ENDURING SUBSTRATISM FROM OTHER METAPHYSICAL THESES Enduring Substratism seems rather intuitive as an account of persistence and change. Aside from this, though, here are three types of reasons one might have for embracing Substratism.

1 An argument from Intervalism (19.1A). If temporal intervals are metaphysically fundamental and not mere logical constructions from instants, then it seems natural to

assume that at least some temporally extended things are fundamental as well. Consequently, Replacementism, whether Classical Perdurantism or Classical Genidentity Theory, with its commitment to Universal Instantaneous Temporal Parts, is a poor fit with Intervalism (especially with Strong Intervalism). One could, however, combine Intervalism with a non-classical version of Replacementism, one in which all temporal parts are temporally extended but intrinsically unchangeable. Such a non-classical Replacementism will have a quite limited set of options for dealing with the possibility of continuous change. (For the corresponding problem involving space, see Section 18.4.1 on continuous variation.)

2 If we adopt Presentism, then we must believe that persisting things are metaphysically fundamental. A persisting thing could not be a logical construction from a collection of instantaneous things spread out over time, since Presentists deny the existence of anything that does not exist in the present. There cannot be any instantaneous things that do not exist in the present moment, and so there cannot be any logical constructions of such things. Similarly, Presentists cannot take there to be fundamental spatiotemporal or *genidentity* relations, since the relata of such relations do not exist at the same time, and so are never in existence together.

3 If we have embraced Strong Powerism (4.4A.3), then we have some reason to embrace Enduring Substratism as well. The possession of passive and immanent powers both presuppose that the bearer persists through time, since to have a passive power is to have the ability to be affected (now or in the future) in a certain way, and to have an immanent power is to have the power to evolve through time in accordance with a certain pattern. If there are fundamental truths about such powers, there must be fundamental truths about the persistence of their bearers. This is especially true if some of the fundamental powers involve continuous activity over intervals of time (such as the passive capacity to acquire inertial motion through absorbing kinetic energy or such biological activities as respiration, metabolism, sensation, and voluntary motion). For similar reasons, Strong Nomism (4.4A.2) may provide support for Substratism, so long as the fundamental laws are laws about the interactions of persisting things.

GENERAL ARGUMENTS FOR REPLACEMENTISM As we have noted, the paradox of intrinsic change is often taken by Replacementists to support their view. Indeed, it is sometimes touted as the main reason to be a Replacementist (cf. Lewis 1986a). Here are two other arguments for Replacementism.

1 An argument from Instantism (19.1T). If temporal instants are fundamental and all temporal intervals are mere sets or pluralities of instants, then it would be natural to conclude that persisting things are likewise mere sets or pluralities of instantaneous parts. However, this argument isn't airtight. Substratists could believe that both instants and persisting things are fundamentally real.

2 Replacementism is very attractive from the perspective of Ockham's Razor (**PMeth 1**), especially if one adopts Spatial Pointillism (18.1T), Anti-Tensism, and Neo-Humeism (4.4T). On this combination of views, there is only one kind of fundamental thing, namely, a thing that occupies a single point of time for just an instant. Each of these things has one or more pure qualities (non-powers) during its brief moment of existence. All of the other facts that make up the world are facts about the temporal and

spatial distances between these instantaneous, point-sized entities. The world is just a four-dimensional mosaic of pure qualities, and nothing more. (Couple this view to Concretism 14.1T.1T, and you have, in broad strokes, David Lewis's metaphysical system. See Lewis 1986a.)

### 24.2.1 Replacementism: reducing the persistent to the instantaneous

Replacementists, as we have noted, think of persistent and changing things as composed, in some way or another, of temporal parts. We turn now to a more detailed exploration of this view.

FROM PLENITUDE TO RAMSEY-LEWIS-SIDER PERDURANTISM Replacementists imagine a world that is literally filled with instantaneous objects, each fundamental and each existing only for a moment. Since we are assuming that there are persisting things, we must suppose that these instantaneous objects sometimes form persisting wholes that persist for some finite period of time. Replacementists must ask an analogue of the Special Composition Question (from Chapter 22) applied to temporal parts and wholes: when do some instantaneous things compose a persisting thing? One possible and very simple answer is: Always. This is Temporal Plenitude:

**24.3T.1 Temporal Plenitude.** For every set *S* of instantaneous temporal parts of things, containing exactly one instantaneous object for each instant in some finite interval *T*, there is a persisting thing that persists throughout *T* with exactly the members of *S* as its instantaneous temporal parts.

Temporal Plenitude can be expressed in the language of *spacetime worms*. We could say that Temporal Plenitude asserts the existence of arbitrary spacetime worms, one worm for each set of instantaneous temporal slices of things, containing exactly one time-slice for each instant in some interval *T*. Adopting Temporal Plenitude would force us to recognize all kinds of strange objects, such as things that are squirrels before midnight and battleships after midnight or car-like objects that go out of existence whenever a car is driven out of a garage and that are generated whenever a car is driven into a garage. Any successive chain of time-slices, however different and however discontinuously scattered about space, would constitute a persisting object, according to Temporal Plenitude.

A somewhat restricted version of this thesis would posit only the existence of arbitrary *continuous* worms, worms whose trajectory through time and space is continuous, without any jumps or discontinuities.

**Def D24.6 Spacetime Worms.** *S* is a *spacetime worm* over interval *T* if and only if *S* is a set of instantaneous things, containing exactly one instantaneous thing for each instant in *T*.

**24.3T.2 Continuous Arbitrary Worms.** For every spacetime worm *S* over interval *T*, if the spatial locations of the members of *S* form a continuous trajectory through spacetime, there is a persisting thing that persists throughout *T* with exactly the members of *S* as its instantaneous temporal parts.

Why do philosophers find it natural to talk of spacetime "worms" in this context? Imagine a dog walking a mile down the street. The path of the dog is worm-like: very narrow at each moment in time (less than a yard wide, tall or deep), but very long over the interval of the walk (several hundred yards long). In addition, if the dog takes 15 minutes to walk the mile, then the corresponding spacetime worm is 15 minutes long along the time dimension. We can convert time into space and vice versa by using the natural ratio given by the speed of light. A second of time would be equivalent to a light-second of space (186,000 miles). Consequently, a 15-minute walk is very long in the time dimension and narrow in the three spatial dimensions.

The main advantage of Continuous Arbitrary Worms over Temporal Plenitude is that the latter lends support for the impossibility of discontinuous motion.[3] That is, it provides support for Continuity of Motion:

**PNatPhil 3 Continuity of Motion.** It is impossible for any material thing to move discontinuously through spacetime.

There is, however, another serious problem with Continuous Arbitrary Worms, one explored by Hawthorne (2006: 111–144). In fact, Hawthorne isolates two related problems: the *restriction* problem and the *collapse* problem. To satisfy the restriction problem, Replacementists must somehow explain which continuous worms correspond to *quality* or *dynamically first-class* objects, that is, objects that obey the fundamental laws of physical dynamics. Many continuous worms obviously do not satisfy those laws. For example, suppose that two neutrinos with different velocities were to meet at a point in space and pass through one another without effect. The history of each particle corresponds to a continuous spacetime worm. However, there are also two other worms, one made up of the pre-collision history of particle 1 and the post-collision history of particle 2, and the other made up of the pre-collision history of particle 2 and the post-collision history of particle 1. These two spacetime worms correspond to fictional or *junk* objects. Call these junk objects particle 1-2 and particle 2-1. Particle 1-2 might suddenly speed up at the meeting point and simultaneously change directions, while particle 2-1 slows down and makes a mirror-image change in direction. The two junk particles might well violate the laws of conservation of energy and momentum. They might even violate other laws, such as the conservation of spin.

A slight variation in the same thought-experiment also illustrates the collapse problem. Suppose for example that two identical particles actually collide and ricochet off one another. In such a case, there are actually four continuous spacetime worms. Let's use *A* and *B* to represent the paths of the two particles before the collision, and *C* and *D* to represent the two paths after the collision. Particle 1 follows the trajectory *A*+*C* and particle 2 follows the course *B*+*D*. In addition to these two, there are two additional spacetime worms: *A*+*D* and *B*+*C*. These two worms represent the case in which the two particles pass right through each other without deflecting or being deflected from their original, straight-line trajectories. Common sense tells us that there are two distinct possibilities here, one in which there is collision and ricochet and a second in which the particles pass through one another without effect. The doctrine of Arbitrary Continuous Worms would force us to say that there is just one possibility here, one that can equally well be described in either way. There would be no real fact of the matter as to which happened, a result that seems quite paradoxical.

Another set of thought-experiments involves motion in absolutely continuous and homogeneous substances. Examples of this sort have been offered by C.D. Broad (1925: 36–7), Saul Kripke (discussed by Shoemaker 1984: 242–247), and David Armstrong (1980). We could imagine an infinitely long river of homogeneous stuff. Imagine that we've reduced friction to zero, and the river is not undulating or pulsating in any way. We can't tell whether the river is moving at all or how fast it is moving or in which direction, simply by looking at the changes in qualities at various spatial locations, since there are no qualitative changes of this kind. A similar thought-experiment involves a sphere of homogeneous stuff that is spinning (cf. Zimmerman 1999). Again, its spinning in a certain direction doesn't correspond to any pattern of qualitative or quantitative change. The sphere maintains its shape and position, and its interior remains homogeneous in quality, density, and chemical composition. How can we distinguish the quality objects from the junk objects if the only fundamental things are instantaneous?

Take a sphere of fluid in the homogeneous river. We could build a spacetime worm of such spheres corresponding to a northward movement at 10 mph. We could build an equally good worm corresponding to a perfectly stationary sphere. In fact, there are infinitely many such spacetime worms, all coinciding with a given sphere at a time, all moving in different directions and at different speeds, and all perfectly consistent with the uniform distribution of qualities. (As Hawthorne pointed out, the qualitative uniformity isn't essential to these thought-experiments. Even if the river varied in quality over time, we would still have to ask whether it is a stationary fluid that is changing in its qualities at various places or whether it is a qualitatively constant fluid that is moving around.)

The homogeneous movement thought-experiments also illustrate the *collapse* problem. Replacementists seem to be forced to say that there is no fact of the matter as to whether the fluid or the sphere is moving. There is just one possibility that can be described in infinitely many different ways, each equally true. Intuitively, this is simply wrong. There is all the difference in the world between a stationary disk and a spinning one, or a disk spinning in one direction and one spinning in the opposite direction.

Replacementists might hope that the laws of nature or facts about causal connections could help sort out both of these problems. Suppose, for example, that there is a law of nature to the effect that whenever the paths of two particles meet, they must collide and ricochet off one another. That would rule out one of the two possibilities in the case of the colliding particles thought-experiment. However, Replacementists had better be careful here. If the laws of nature are fundamental truths, and those laws make reference to persisting things, like particles, then it would seem that the persisting things are also fundamental and irreducible.

The best way (and perhaps the only way) out would be for the Replacementists to adopt Neo-Humeism, reducing laws to patterns of particular fact by means of the Ramsey/Lewis Theory of laws (see Section 5.2). As Sider (2001: 230–236) explains, we could then extend the Ramsey/Lewis Theory to an account of persisting things as well. The account would go roughly like this: a spacetime worm corresponds to the existence of a persisting thing if and only if the simplest and most powerful scientific theory of the actual world assigns the worm to a single persisting thing as its trajectory. This view is Ramsey-Lewis-Sider Worms:

**Def D24.7 Worm/Thing Correspondence.** If *x* is a spacetime worm and *y* is a persisting thing, then *x corresponds to y* if and only if *x* and *y* have exactly the same instantaneous parts.

**24.3T.3 Ramsey-Lewis-Sider Perdurantism.** A spacetime worm *S* over interval *T* corresponds to the existence of a derived thing persisting through *T* if and only if the simplest and most powerful scientific theory of the actual world assigns a persisting entity to *S*.

If Ramsey-Lewis-Sider (R-L-S) Perdurantism is combined with Neo-Humeism, we end up with a version of Replacementism according to which the existence of persisting things is reducible to the more fundamental facts about instantaneous things and their qualities and spatiotemporal relations. R-L-S Perdurantism is one way of providing an account of *genidentity* as a non-fundamental or derived relation:

**Def D24.8 Genidentity.** Two instantaneous things *x* and *y* stand in the *genidentity relation* if and only if there is a single persisting thing *z* (either fundamental or derived) such that *x* and *y* are time-slices of *z*.

For Substratists, the *genidentity* relation depends on the real, fundamental existence of persisting things. For Replacementists, this cannot be so. Replacementists have two options: they could add *genidentity* as a new, fundamental relation (resulting in Classical Genidentity Theory) or they could try to reduce *genidentity* to other relations, such as causal, nomic, and temporal ones. R-L-S Perdurantists take the second course. (We'll take up Classical Genidentity Theory in the next section.) For R-L-S Perdurantists, whether two instantaneous things stand in the *genidentity* relation to one another depends on whether the simplest theory of the world treats them as time-slices of the same thing.

How does R-L-S Perdurantism help with the problem of the spinning homogeneous sphere or the flowing infinite and homogeneous river? It can help if the sphere and the river have a more interesting history and if they are embedded in a world in which a variety of things happen. Suppose that someone started the sphere's spinning by striking it in a certain way, and suppose that the simplest and most powerful scientific theory in that world is one in which striking things in that way causes them to spin. For example, there may be many other cases of bumpy and heterogeneous spheres that are made to spin in certain ways in that world. In the case of the river, perhaps the whole river is under the influence of a gravitational field that causes all liquids to flow downhill at a certain rate of acceleration, whether they are homogeneous or not. Under such hypotheses, there could be a definite matter of fact as to whether a given sphere is spinning or a given river flowing.

PROBLEMS WITH R-L-S PERDURANTISM: SMALL POSSIBLE WORLDS Dean Zimmerman (1999) points out that R-L-S Perdurantists will have residual problems with small worlds, a point that Lewis himself conceded (Lewis 1999b). Suppose that the entire universe is filled with a single, uniform substance and that the laws of the world in question make it impossible for there to be any vacuums and impossible for there to exist any matter of any other kind. In such a world, any law of motion would be superfluous. The simplest system of laws would require the fluid to be essentially stationary. We would have to say that it is metaphysically impossible for any locomotion to exist in such a universe, which is surely an incorrect result. We could easily imagine that there are eternally recurring eddies and vortexes within the fluid, any of an infinite variety of kinds of motion.

Sider (2001: 233–234) considers a particular simple world that is very troubling for his theory, namely, a world consisting at all times of a single, spinning, homogeneous sphere, with a velocity vector field attached to the material points within the sphere. That is, the simplest laws of nature require us to assign vectors to each point in the sphere, a vector of the kind that we normally assign to moving things in our world. As Sider recognizes, there are at least two equally good candidates for the best system of laws in such a world:

> (1) Each material point location-event with a velocity vector is always genidentical to some later material point location-event that is located in the direction toward which the original velocity vector was pointing. (The sphere is rotating in the direction indicated by the velocity vectors.)
> (2) A material point location-event with a velocity vector is always genidentical to some later material point location-event that is located in exactly the same place as the first location-event. (No motion at all. The velocity vectors all point in some direction, but they are not nomologically connected with any locomotion.)

Sider admits that he cannot decide in a principled way between these two candidates. So he must give up the intuition that there could be any fact of the matter about whether such a sphere is rotating or not.

Sider recognizes that R-L-S Perdurantists face this difficulty, but, as he points out, it is a difficulty afflicting the Ramsey/Lewis Theory of the laws of nature, not one specific to the Replacementist component of R-L-S Perdurantism. Sider puts it this way, "The defender of best-system accounts is already accustomed to biting similar bullets" (Sider 2001: 234) To our mind, these small world counterexamples to the Ramsey/Lewis Theory of laws are about as devastating as things get in metaphysics.

### 24.2.2 Substratism vs. Classical Genidentity Theory

In light of the small worlds objections lodged in the last section, Replacementists might consider taking on board a metaphysically fundamental *genidentity* relation. On this account, although only instantaneous objects are metaphysically fundamental, there is a fundamental relation of *genidentity* that holds between certain slices and not others, in such a way that persisting things exist when and only when there is a genidentity-unified spacetime worm. This is just Classical Genidentity Theory.

Is there a real difference between Substratism with its fundamental persisting things and Classical Genidentity Theory with its fundamental *genidentity* relation? There does seem to be some difference between the two, since the latter takes *genidentity* to be a metaphysically primitive relation, while the former defines *genidentity* in terms of the more fundamental relation of *identity*. Suppose that the *genidentity* relation is necessarily reflexive, symmetric, and transitive (what mathematicians call an 'equivalence relation'):

**Def D24.9.1 Reflexivity.** A relation $R$ is *reflexive* if and only if for every $x$, $Rxx$.

**Def D24.9.2 Symmetry.** A relation $R$ is *symmetric* if and only if for every $x$ and $y$, if $Rxy$ then $Ryx$. [= Def D10.2]

**Def D24.9.3 Transitivity.** A relation $R$ is *transitive* if and only if for every $x$, $y$, and $z$, if $Rxy$ and $Ryz$, then $Rxz$.

**Genidentity an Equivalence Relation.** *Genidentity* is reflexive, symmetric, and transitive.

*Identity* is clearly reflexive, symmetric, and transitive, since everything is identical to itself, if $x$ is identical to $y$ then $y$ is identical to $x$, and if $x = y$ and $y = z$ then $x = z$. This is a theorem of logic because it follows from Leibniz's Law, which says that when $x$ and $y$ are identical, we can substitute one of $x$'s names for one of $y$'s names in a truth and still obtain a truth as a result. If $x = y$ and $y = z$, then we can replace '$y$' with '$x$' in the second sentence, obtaining the further truth that $x = z$.

Given the definition of genidentity and the transitivity of identity, the hypothesis that *genidentity* is an equivalence relation would follow logically from the hypothesis that each time-slice is a temporal part of only one persisting thing. This is No Temporal Coincidence:

**24.4T No Temporal Coincidence.** Necessarily, no instantaneous thing is a time-slice of more than one persisting thing.

**24.4A Temporal Coincidence.** It is possible for an instantaneous thing to be a time-slice of two distinct persisting things.

In a world with temporal coincidence, there could be two distinct spacetime worms, each corresponding to the career of a persisting thing, that perfectly coincide at an instant of time. The contrary thesis, No Temporal Coincidence, rules out such criss-crossing of spacetime worms that correspond to persisting things. Let's prove that No Temporal Coincidence entails the transitivity of *genidentity*, if we assume that *genidentity* is definable in terms of *identity*:

1 Assume No Temporal Coincidence.
2 Assume that $x$ and $y$ are genidentical and that $y$ and $z$ are genidentical.
3 By the definition of *genidentity*, $x$ and $y$ are time-slices of some persisting thing $A$, and $y$ and $z$ are time-slices of some persisting thing $B$.
4 No Temporal Coincidence entails that $A$ and $B$ are identical, since $y$ is a time-slice of each of them.

5 Consequently, $x$ and $z$ are both time-slices of $A$. (This step tacitly assumes that *identity* is eternal: that if $A$ and $B$ are identical at the time of $y$'s occurrence, then they were also identical at the time of $x$'s occurrence. We'll take up this issue in the next subsection.)
6 Thus, $x$ and $z$ are genidentical.

Thus, Substratists have an explanation for the reflexivity, symmetry, and transitivity of *genidentity*, something that Classical Genidentity Theorists have to posit as a brute metaphysical necessity. This provides the Substatists with an advantage in terms of Ockham's Razor (**PMeth 1.2**).

However, is it really the case that *genidentity* is transitive? If *genidentity* weren't transitive, then this failure of transitivity would give a decisive advantage to Classical Genidentity Theory, since Substratists are forced to take definable *genidentity* to be transitive.

There are a number of puzzling cases in which *genidentity* seems to be non-transitive, including possible cases of fusion and fission. Suppose an amoeba $A$ divides into two "daughters", $B$ and $C$. It seems natural to think that $A$ is the same amoeba as $B$, and also the same amoeba as $C$, even though obviously $B$ and $C$ are different amoeba.

*24.2.2.1 The eternity of identity and distinctness.* However, the very cases that raise questions about the transitivity of *genidentity* could be used to raise questions about the eternity of *identity*, which is needed to prove the transitivity of *genidentity*. Perhaps $A$ really is identical to both $B$ and $C$, even though $B$ and $C$ are not identical to each other. Or, to be more precise, perhaps $A$ *was* identical to $B$ and to $C$ (when it was one amoeba) but *is not now* identical to either of them, and so neither are $B$ and $C$ *now* identical to each other. Could $A$ have been identical to $B$ in the past but not in the present or future? Is *identity* the sort of relation that can change, like *love* or *proximity*? The transitivity of *identity* over time depends on the eternity of *identity*: if $A$ and $B$ are identical at one time, are they identical at all times? If so, *identity* over time must be transitive, since *identity* at a time is.

Saul Kripke (1980) gave a very influential argument for the necessity and eternity of identity.

1 Suppose $x = y$.
2 It is a truth of logic that $x = x$.
3 So, it is necessarily true that $x = x$.
4 So, $x$ has the property of *being necessarily identical to x*.
5 Since $x$ and $y$ are identical, $y$ must also have the property of *being necessarily identical to x*. (By an application of Leibniz's Law to 1 and 4)
6 So, it is necessarily true that $x = y$.

We can replace the occurrences of 'necessarily true' with 'permanently true' in Krikpe's argument in order to demonstrate that *temporary* identity is impossible. If $x = y$ at any time, then it is permanently true (true at all times) that $x = y$. Things can't be temporarily identical.

This argument can be extended to demonstrate the eternity of *distinctness* (see Williamson 1996):

1 Suppose that $b \neq c$.
2 Assume that at some time in the future $b = c$, for contradiction.
3 By the previous argument, we know that $b = c$ entails that it has always been true that $b = c$. So, $b = c$ at some time in the future entails that at some time in the future it will always have been true that $b = c$.
4 If at some time in the future it will always have been true that $b = c$, then $b = c$ now. What was or will be always the case is now the case.
5 Contradiction (between 1 and 4). So, it will never be the case that $b = c$.
6 By a similar argument, we can show that from 1 it follows that it never was the case that $b = c$.

Thus, if $b \neq c$ now, then the two have always been and always will be distinct.

*24.2.2.2 Is genidentity transitive?* To sum up: the eternity of *identity* and *distinctness* provides an argument for Substratism over Classical Genidentity Theory. In particular, Substratism can provide a simple explanation for the fact that the *genidentity* relation is reflexive, symmetric, and transitive. Classical Genidentity Theorists must posit this as a brute necessity. However, if *genidentity* is not transitive, then we have a strong argument for Replacementism of some kind. So, the issue of the transitivity of *genidentity* is a critical one.

Here are some apparent cases of non-transitive *genidentity*:

1 Fission and fusion. If $A$ is split in two and survives twice over, as both $B$ and $C$, then $A = B$ and $A = C$, but $B \neq C$, since the two are located at different places, with possibly different qualities and properties. Similarly, if $B$ and $C$ are originally distinct (so $B \neq C$) but fuse together into one simple thing, $A$, we might conclude that $B = A$ and $C = A$, even though $B \neq C$.
2 The Methuselah paradox (based on Thomas Reid's example of the brave officer—Reid 1785). If a person were to live for thousands of years, like Methuselah in the Bible, the younger Methuselah might be so different from the older Methuselah that the two are different people. However, each year's Methuselah is identical to the following year's Methuselah, so $M_1 = M_2$, $M_2 = M_3$, and so on until $M_{(n-1)} = M_n$. However, $M_1 \neq M_n$, another failure of transitivity. We could also apply Chisholm's paradox to a very long-lived particle, which changes over time from one kind of matter to an entirely different kind.

In the next section, we will consider the problem of the persistence of composite objects, including objects that can gain or lose parts. For the time being, let's suppose that we are only concerned with Hawthorne's *quality* or dynamically *first-class* objects. These are the sort of things whose *genidentity* matters when distinguishing spinning disks and flowing rivers from stationary ones. For such things, we have good reason to suppose that *genidentity* is necessarily transitive, thanks to the conservation laws (e.g., the conservation of mass-energy, charge, baryon number, and so on). If a dynamically first-class particle $A$ were to split into two, then the two fission products would each have to have less mass than $A$ had (unless one of the two were to have no mass at all). Thus, it is impossible for us to count both *B* and *C* as genidentical to *A*, since at least one, if not both, will differ from *A* in one or more essential quantities.

Similarly, conservation laws will guarantee that any quality object, no matter how long-lived, will always possess the very same essential quantities. If the particle were at some point to lose mass or charge, then we would have good grounds for saying that the particle had been destroyed at that point.

Thus, there are good grounds for thinking that *genidentity*, as applied to quality or first-class objects, is necessarily transitive. This provides some advantage to Substratists if they assume that No Temporal Coincidence is necessary. Under these assumptions, the transitivity of *identity*, which is a law of logic, is sufficient to explain the transitivity of *genidentity*.

And there is a reason Substratists might have for embracing No Temporal Coincidence. First, we have to assume that Substratists are willing to concede that there are in fact time-slices. That is, we must assume Time-Slice Substratism. According to Endurantists, there can be no question about No Temporal Coincidence because there simply are no time-slices at all. If Substratists do accept the existence of time-slices, they would be wise to consider time-slices to be logical constructions (like ordered pairs of persisting things and instants of time) or dependent entities (like spatial boundaries). If each time-slice is metaphysically dependent on the persisting thing of which it is a part, that is, if time-slices are a kind of internal temporal boundary of persisting things, then No Temporal Coincidence seems to follow, because the identity of the time-slice will be a function of the identity of the thing of which it is a temporal part.

As we have seen, if *genidentity* is transitive and No Temporal Coincidence is true, Substratists can explain the transitivity of *genidentity* by means of the logical truth that *identity* is transitive, while Classical Genidentity Theorists must treat the transitivity of *genidentity* as a brute metaphysical necessity. We've seen that the conservation laws give us good reason to believe that *genidentity* is transitive insofar as it connects stages of quality objects.

However, there is an objection to this argument to consider. Couldn't Classical Genidentity Theorists appeal directly to the conservation laws themselves as the ultimate explanation for the transitivity of *genidentity*? They could certainly appeal to something like the ordinary conservation laws, but not perhaps to the laws themselves. We would normally state a conservation law in something like the following way:

(1) The mass-energy of a particle or other quality object does not change over time, unless it gains or loses a quantity of mass-energy from some other quality objects.

However, Classical Genidentity Theorists can't appeal to such a conservation law, since they don't believe in the fundamental reality of quality objects. They must appeal to something like (2):

(2) The mass-energy associated with one instantaneous object must be equal to the mass-energy associated with any other object genidentical to the first, unless there has been some intervening interaction.

There is something more natural about the form of the conservation law in (1), as compared with (2). The law stated in (1) appeals to an a priori very plausible principle: if one and the same thing exists at two times, then it will be intrinsically the same at both times, unless there is some causal explanation of the change. The form of the law in (2) can make no such appeal, since the two time-slices are only genidentical to one another. They are not one and the same thing; they just bear some primitive, external relation to each other. So, it seems that Substatists have some advantage over Classical Genidentity Theorists in explaining the basis of the conservation laws, in respect of Ockham's Razor (**PMeth 1.4**).

## 24.3 The Metaphysics of Motion

In the last section, we examined the problem of intrinsic change. In this section, we turn to a kind of change that may (or may not) be extrinsic: change of position or *locomotion*. What is the nature of locomotion? If a thing has a certain velocity, what does this fact consist in? There are two main contenders: the At/At Theory of motion (developed by Bertrand Russell 1917) and Intrinsic Motion.

### 24.3.1 The At/At Theory vs. Intrinsic Motion

On Russell's At/At Theory, a body's motion consists in its being at different locations at different times, nothing more or less. This means that whether a thing is in motion at a given moment depends on where that thing was and will be at other moments. Its being in motion at a moment isn't intrinsic to its instantaneous condition at that moment. Intrinsic Motion, on the other hand, insists that facts about whether something is in motion at some time are intrinsic to its instantaneous condition at that time.

Intrinsic Motion has its roots in Zeno's paradox of the arrow (discussed by Aristotle in Chapter 9 of Book VI of the *Physics*). Zeno argued that no arrow is ever in motion, since, at every point of time, the arrow has a single location. Whatever occupies just one position in space is motionless. So, the arrow is motionless at every moment. Whatever is motionless at every moment is never in motion. Hence, the arrow is never in motion. The obvious answer to Zeno is to insist that the arrow is in motion at each instant during its flight. But, what does it mean for the arrow to be in motion at each instant? Is this an intrinsic fact about that arrow at and only at that moment, or does its being in motion at one instant consist in its being located at other places before and after that instant? This is just the dispute between Intrinsic Motion and the At/At Theory.

**Def D24.10 Intrinsicality at a Time.** A property $P$ is *intrinsic to a thing $x$ at time $t$* if and only if $x$'s being $P$ at $t$ is not partially grounded in its state at any other time, or on the state of anything separate from it. [Compare Def D2.3 Intrinsicality]

**24.5T Intrinsic Motion.** Motion is something intrinsic to the moving thing at each moment of its motion.

**24.5A Extrinsic Motion.** Motion is not intrinsic to each instantaneous time-slice of the moving thing.

**Def D24.11 Instantaneous Location Events.** An *instantaneous location-event* is the event of some things (or some kind of thing's) being at some location at some instant of time.

**24.5A.1T The At/At Theory of Motion.** The fundamental truths about locomotion are truths about instantaneous location events.

Are the At/At Theory and Intrinsic Motion incompatible because the At/At Theory entails Extrinsic Motion? On the At/At Theory, whether or not a thing is moving, and how it is moving, depends entirely on the set of location events involving the thing. If the At/At Theory is correct, one cannot tell, simply by looking at an instantaneous time-slice of a body, whether or how it is moving.

If we adopt the At/At Theory, it would make sense also to embrace a similar theory of velocity: something's velocity at a moment *t* is the first derivative (the slope) of the object's trajectory through spacetime at that point. This means that an object could exist at a moment without a single, well-defined velocity if its motion is not continuous or not differentiable at that moment. For example, if something were to abruptly change its velocity from 10 mph to 20 mph, it would have no well-defined velocity at the moment of the change. Similarly, if an object were to jump through space and time in a discontinuous way, moving through space up until moment *t*, then appearing for an instant at *t* somewhere in the distance, and then somewhere else in some succeeding instants, the object would have no velocity at all at instant *t*. Third, suppose an object were created at *t* and annihilated immediately afterward, having completed no trajectory in spacetime. Such an object could, on the At/At Theory, have no velocity during its sole moment of existence. In contrast, Intrinsic Motion could attribute some definite velocity to the objects in each of these three cases.

Kinetic energy and momentum do seem to be intrinsic properties of basic particles. These quantities are conserved in a variety of transactions, and the conservation of energy and momentum play a role in explaining observed motion. Consider, for example, Newton's cradle, in which three steel balls are suspended from a framework in such a way that the three balls are touching at rest. If one ball is raised and allowed to swing, hitting the second, the kinetic energy flows from the first through the second into the third, causing the third to move, even though the second does not move at all (assuming that the balls are perfectly rigid bodies).

At the same time, kinetic energy and momentum both seem to be functions of velocity (energy equal to mass times velocity squared, momentum to mass times velocity). If so, the middle ball in the cradle must have, for an instant, a positive velocity without actually changing position. At/At Theorists must either deny that kinetic energy and momentum are intrinsic properties of things or else they must deny that a thing's kinetic energy and momentum are determined by (and dependent on) its velocity. Either assumption seems implausible:

**Intrinsic Energy.** Kinetic energy and momentum are intrinsic properties of moving things.

**Energy a Function of Velocity.** Kinetic energy and momentum are determined by and dependent on velocity.

The best strategy for At/At Theorists is to challenge Intrinsic Energy, arguing that the explanatory role of kinetic energy and momentum comes entirely from the fact that the fundamental laws of motion make reference to them, not to the metaphysical status of the properties as intrinsic features of moving things. This turns on a question that we took up in Chapters 4–6: are powers or the laws of nature more fundamental? At/At Theorists should be attracted to either a Nomism or Neo-Humeism, while Intrinsic Motion fits best with Powerism.

The At/At Theory is compatible with relativity theory. The At/At Theory doesn't presuppose the existence of absolute rest or motion or of a substantial space. It can be formulated in terms of Minkowski spacetime, with four dimensions (three of space and one of time). Intrinsic Motion, in contrast, is incompatible with relativity theory, since it entails that there is a real, intrinsic difference between bodies that are moving and those that are at rest. This is a serious drawback to Intrinsic Motion.

### 24.3.2 The At/At Theory vs. Intervalism

We can distinguish between *moderately discontinuous* and *radically discontinuous* motion. In the case of moderately discontinuous motion, every location-event belongs to a process of movement, but two such processes may be discontinuous, with the object undergoing a *quantum jump* in location from one place to a distant one at an instant. In the case of radically discontinuous motion, there are spatially isolated location events, events that do not belong to an extended process of motion, neither to the interior nor to an extreme boundary of such a process. In radically discontinuous motion, an object might undergo continuous motion up to time *t*, jump discontinuously to some distant location at time *t*, and jump immediately afterward, undergoing a second period of continuous motion after *t* at another distant location. The object's motion at *t* in such a case is radically discontinuous, unconnected to any continuum of motion both before and after.

There are two reasons for thinking that moderately discontinuous motion might be impossible. First, there is the problem of securing the persisting identity of the moving body both before and after the discontinuous jump. Second, any such jump would involve a kind of action at a spatial distance, with the condition of the body after *t* dependent directly on its condition at *t*, with a finite distance separating the two.

**No Radically Discontinuous Motion.** No body occupies a spatial position *A* at time *t* while also occupying positions at some fixed finite distance *d* from *A* at all times both before and after *t*.

Intervalism (19.1A) is a third reason for rejecting the possibility of radically discontinuous motion. If intervals rather than instants are metaphysically fundamental, then it would be impossible for a moving body to occupy a location at an instant when that location-event is not part of a temporally extended process of locomotion. This suggests that there is a third alternative, one disagreeing with both the At/At Theory and Intrinsic Motion. This third theory would agree with the At/At Theory by rejecting the view that motion is intrinsic to a thing at a moment, but it would differ from the At/At view by insisting that the fundamental truths about location are not truths about instantaneous location events but rather truths about extended processes of motion. This third position is Motion Intervalism, since it builds on the foundation of Intervalism:

**24.5A.1A Motion Intervalism.** The fundamental truths about locomotion are truths not about location events but about extended processes of motion.

Motion Intervalism is the position adopted by Aristotle and many of his successors. It has a simple answer to Zeno's paradox: the history of the arrow does not consist in a series of momentary states, but in one or more continuous processes of motion. During each of these metaphysically fundamental motion-processes, the arrow does not have a unique location, but rather an extended trajectory through space and time. At the same time, it can explain why radically discontinuous motion is metaphysically impossible, in a way that the At/At Theory cannot.

## Notes

1 Note that Presentists don't have to be Endurantists. One can be a Perdurantist Presentist, though one would then be committed to denying Persistence. See the second argument for Endurantism, below.
2 Take our reference to vague and context-sensitive properties like *coldness* or *bentness* to be shorthand for picking out some specific, fundamental property (like *having a precise temperature* or *exact velocity*).
3 Quantum mechanics does not support the idea that motion can be discontinuous. We could say that a particle that is in a superposed state is simultaneously traveling along a very large number of different paths through space, but each of them is continuous.

# 25

# The Persistence of Composite Things

In Chapter 24, we saw that Substratism (24.1T.1T.1T) and Replacementism (24.1T.1T.1A) are the two major contending accounts of intrinsic change. In this chapter, we turn to the interaction between composition and persistence, as the persistence of composite objects provides a critical test case for evaluating these two accounts. In particular, we are concerned with a question about the interconnection between composition and persistence: under what conditions do composite things persist through time? This question divides in two. First, can a composite thing survive the loss or the addition of parts? And second, can two distinct things come to share exactly the same parts at the same time, while retaining their distinct identities? These are, respectively, questions about mereological *constancy* and mereological *coincidence*. If a thing is mereologically constant, then it is incapable of gaining or losing parts over time; same thing, same parts. Some complex things, including organisms, seem to be mereologically inconstant, that is, capable of gaining or losing parts. In Section 25.1, we try to sort out whether inconstancy is possible, and if so, how. Two things are mereologically coincident at a time if they have exactly the same proper parts at that time. In Section 25.2, we look at arguments both for and against the possibility of coincidence. If coincidence is impossible, we would have the metaphysical law: same parts (at a time), same thing.

These two issues pose some challenging problems for both Substratism and Replacementism. For example, if Substratism is true, then we should expect diachronic survival to be a transitive relation: if *A* is the survivor of *B* at some point of time, and *B* the survivor of *C* at another time, then Substratists should assume that numerically the same substrate underlies *A*, *B*, and *C* at their various locations in time. Since numerical identity is transitive, so must be survival through time. However, as we shall see, there are a number of paradoxical situations that challenge the assumption that survival is transitive. In addition, Substratists have to cope with the possibility of intermittent existence. For example, if substrate *x* exists at time $t_1$, ceases to exist at a later time $t_2$, and then exists

*The Atlas of Reality: A Comprehensive Guide to Metaphysics*, First Edition.
Robert C. Koons and Timothy H. Pickavance.

again at some still later time $t_3$, Substratists must provide an account of what it means for a substrate to exist at some times and not at others. They will also have to explain why some substrates can survive certain kinds of changes and not others, and why some kinds of substrates can survive changes that other kinds cannot.

Replacementists are not bound to any assumption of the transitivity of identity, as we have already seen. Nonetheless, the task of accounting for the various paradoxes of survival for inconstant objects is still not trivial for the Replacemenist.

## 25.1 Mereological Constancy and Inconstancy

Can persisting things gain and lose parts? It would seem that quality objects, that is, dynamically first-class things that obey the fundamental laws of physics, cannot gain or lose parts. Such dynamically first-class things must satisfy the law of conservation of mass-energy, among other conservation laws. Such things are mereologically constant.

**25.1T Mereological Constancy.** Necessarily, everything that persists has exactly the same parts at all times at which it exists.

**25.1A Mereological Inconstancy.** It is possible for a persisting thing to gain or lose parts over time.

Are there any mereologically inconstant things? Unstructured masses seem to be constant. So we have to look to structured things like natural formations (rivers, mountains), artifacts, and living organisms. All of these seem capable of surviving the gain and loss of parts. A river continues to exist, although it is constantly gaining new bits of water (as the rain falls on its watershed) and losing other bits (as they flow into the sea or evaporate from its surface). Artifacts can gain and lose parts, and it is obvious that living organisms do so all the time. But are any of these things fundamentally real? We've already established that they are not quality or dynamically first-class objects. That may provide some reason for doubting whether they are fundamentally real, but it is not all by itself a decisive reason. Why couldn't there exist real, fundamental things that do not themselves fall within the scope of the laws of physics? The laws of physics might apply only to a proper subset of the world's fundamental things.

Do we have any positive reasons for believing in mereologically inconstant things? Our common sense suggests that the world is filled with such things, as we've seen. Science might give us some reason for believing in the existence of unifying, continuous processes, like the flow of liquid water in a stream or of nuclear fusion in the heart of a star. In addition, we have strong Cartesian reasons for believing that each of us exists (see the discussion of the *Cogito* argument in Section 22.6.1), and each of us seem to be inconstant, since we are living organisms. In order to avoid that conclusion, we might suppose that we are immaterial souls that inhabit and animate our living bodies or even that we are each some small, special particle located somewhere in the brain.

There are two problems with such attempts to identify human beings with constant objects. First, many of our ethical and interpersonal beliefs and practices presuppose that we are composite, mereologically inconstant, (at least partly) material objects, and not mere souls. We treat rape and torture as violations of the person, and not as on a par with vandalizing a person's property. We believe that we see and touch other people and not merely material things animated by other people. Second, if there were no mereologically inconstant things, then we would have no bodies. All that would exist would be atoms and souls (and mereologically constant sums of these). If one's body does not exist, then there would be a very severe problem explaining how one (a pure immaterial soul) would be able to interact in a regular way with the physical universe. Mental/physical interaction is a hard enough problem for dualists. Denying the existence of a persisting, living body makes the problem insoluble. What would tie each soul to some congeries of particles if there is no enduring human body to serve as the locus of interaction?

### 25.1.1 Objections to inconstant things: Paradoxes of intransitive persistence

Even though the existence of inconstant things seems to be part of our common-sense view of the world, philosophers have long been aware of the fact that inconstant objects give rise to a number of difficult puzzles and paradoxes. Many have used these paradoxes as an argument for rejecting the real existence of inconstant objects, especially when Substratism is taken for granted.

If something really persists from time $t_1$ to time $t_2$, then there is something existing at time $t_1$ that is identical to something existing at time $t_2$. We have cases of diachronic (cross-time) identity. It is a basic principle of logic that identity is transitive:

**The Transitivity of Identity.** Necessarily, if $x = y$ and $y = z$, then $x = z$.

We have also seen in Section 24.2.2.1 that there is a strong argument (made by Kripke) for thinking that *identity* and *distinctness* are eternal. Once distinct, always distinct, and once identical, always identical. Consequently, if $x = y$ at any time and $y = z$ at any time, then $x = z$ at all times.

However, if there are inconstant things, then there are cases in which we want to say that $x$ (something existing at time $t_1$) is identical to $y$ (existing at time $t_2$), $y$ is identical to $z$ (existing at time $t_3$), and yet $x$ is not identical to anything existing at time $t_3$. In other cases, we want to say that $x$ (existing at time $t_1$) is identical to $y$ (at time $t_2$) and also identical to $z$ (at time $t_2$), even though $y$ and $z$ are not identical to each other.

PARADOX 1: THE SHIP OF THESEUS The ship of Theseus is an ancient puzzle about the persistence of material objects through time. We are to imagine a ship whose planks of wood are taken out, one by one, placed in a warehouse, and replaced by new planks. Eventually, all of the ship's wood has been replaced, and a second ship is constructed from the planks stored in the warehouse. Which ship is the *original*? It is tempting to say that both ships are identical to the original, but this leads to a conflict with the transitivity of identity, since it is obvious that there are two distinct ships at the end.

If we reject the existence of artifacts and other inconstant objects, we can avoid the problem by simply denying that there ever was or is a ship at all. All we can say is that there is *ship-ping* going on *here-ishly* and *there-ishly*. The question of which ship is identical to the original ship of Theseus cannot be properly posed.

PARADOX 2: TIB AND TIBBLES Suppose there was a rabbit that putatively survived the loss of its left foreleg. Let's call the rabbit before the removal of the leg 'Tibbles', and let's call the rabbit minus its leg 'Tib'. It seems that both Tib and Tibbles exist before the leg is removed. At that point in time, Tib is a proper part of the rabbit (all of the rabbit except for its left foreleg). It seems clear that Tib and Tibbles are not identical, since Tibbles has a left foreleg and Tib does not. However, after the leg is removed, the rabbit is identical to Tib, since the rabbit itself now lacks a left foreleg. So, we seem to be saying that, at $t_1$, Tib and Tibbles were not identical, and the rabbit and Tibbles were identical. At the later time $t_2$, the rabbit and Tib are identical. So, the rabbit was once not identical to Tib, and then later it is identical to Tib. (This problem is related to the problem of the many introduced in Section 22.4.1, but applied across time rather than merely at a single time.)

This problem can be dissolved by rejecting Mereological Inconstancy. There are two persisting things: Tib and Tibbles. Before the operation, Tibbles is a rabbit and Tib isn't (Tib is just part of a rabbit), and after the operation, Tib is a rabbit and Tibbles isn't (since Tibbles is now a scattered object). There simply is no such thing as the persisting rabbit. No rabbit ever persists: all that persist are mereologically constant objects like Tib and Tibbles. Defenders of Mereological Inconstancy can dissolve the paradox in another way: by denying Mereological Universalism (22.3T.1). This at least makes it possible to deny (though it does not entail that) there is no such thing as Tib prior to the loss of the leg, and no such thing that includes the leg after the loss.

PARADOX 3: THE KAFKA PARADOX The Kafka Paradox is similar to the ship of Theseus, except that we focus on the career of a single inconstant object. Suppose that a living organism, like a human being, can always survive the removal or addition or replacement of a single particle. We start with a human being, like Franz Kafka, and we gradually transform Kafka into a cockroach, one particle at a time. It seems clear that a human being cannot persist if all that remains is a cockroach. Human beings cannot become cockroaches, since cockroaches lack many of the essential features of humanity. However, each stage of the transformation seems to be genidentical to the next, and we have good reason to believe that genidentity is transitive.

PARADOX 4: FISSION AND FUSION Suppose that a human being could survive the loss of one-half of his body, including one-half of his brain. Now suppose that a human being is cut in half, and each half is reconstituted into a whole human being through the addition of transplanted organs. Call the original person 'Alpha', and the two fission-products 'Beta' and 'Gamma'. Since Alpha could have survived as Beta, and Alpha could have survived as Gamma, it seems that the successful production of both Beta and Gamma should not negatively affect Alpha's survival. In addition, let's suppose that the entire operation was perfectly symmetrical, so we have no reason for saying that it is Beta, rather than Gamma, that is the best candidate for being the continuation of Alpha, or vice versa.[1] Thus, we should say that Beta is Alpha and that Gamma is Alpha. However, it is clear that Beta and Gamma are not identical. They are located in different places, doing different things. Over time, Beta and Gamma could become different in many ways, including their memories, personalities, and legal status. Once again, since *identity* is eternal and transitive, we seem forced to say both that Beta and Gamma are identical and that they are not identical.

The possibility of fusion provides similar problems for inconstant objects. Suppose we take the organs and brain parts from two people Alpha and Beta and fuse them together, Frankenstein-style, into a single human being, Gamma. Have both Alpha and Beta survived? Was Gamma once two different people? To think so is to violate the eternity of *identity*. And yet it seems that Alpha and Beta have both in some sense survived, and that Gamma has had some sort of past existence.

### 25.1.2 Objections to inconstant things: Vagueness and conventionality

THE VAGUENESS OF IDENTITY Suppose that there once was a restaurant in Philadelphia called "Bookbinders". The restaurant is moved and subsequently changes owners and menus. Is it still the same restaurant or a different one? There will be borderline cases in which we feel that either answer is legitimate. It is hard to believe that there is a real, fundamental fact of the matter. We can make sense of this fact if inconstant objects like restaurants are just useful fictions.

THE CONVENTIONALITY OF IDENTITY: EXOTIC OBJECTS Eli Hirsch (1992, 1993) asks us to imagine a community with very exotic ideas about the persistence conditions of certain objects. For example, they do not believe in cars, but they do believe in *incars* and *outcars*. An incar consists in what we would call a car while and insofar as the car is inside a garage. An outcar is a car outside a garage. Let's call the members of this conceptually exotic community 'Hirschians'. When we back a car out of a garage, Hirschians would say that an incar is gradually shrinking until it vanishes, and an outcar comes into existence, first as a part of the rear bumper and gradually growing into a complete outcar. Hirschians don't believe in persisting cars. They think that the only really persisting things are incars and outcars. If you drive your outcar into a garage, you have destroyed it and replaced it with a new entity, an incar. There are similarly exotic objects that could be constructed out of natural formations. For example, we could consider *Sunday-rivers*. A Sunday-river is a river-like entity that comes into existence on midnight at the beginning of one Sunday and ceases to exist at the end of the next Sunday.

We find it hard to believe that incars or outcars are really there, but do we have a good reason for thinking so? Aren't the persistence conditions we assign to natural formations and artifacts just as arbitrary and conventional as the conditions used by Hirschians? A simple solution is to deny the existence of all inconstant objects, cars and incars alike.

A TRUTHMAKER OBJECTION As we argued in Section 23.4, Mereological Inconstancy raises truthmaker worries. If $A$ is mereologically inconstant, and $B$ is a part of $A$, then what is the truthmaker for the truth of $B$'s being a part of $A$? It can't be $A$ itself, since $A$ can exist without including $B$ as a part. Could it be $B$? Only if $B$ is a rigid part (Def D23.6) of $A$, something that couldn't exist without being a part of $A$. How could that happen? If neither $A$ nor $B$ is a truthmaker for the mereological truth, what is?

### 25.1.3 Responses to the objections

We'll break down the defense of inconstant objects into three parts: the defense of organisms, the defense of artifacts, and the defense of natural formations, like rivers and rocks. Finally, we will look at a second way of defending all inconstant objects through embracing ontological plenitude.

*25.1.3.1 In defense of organisms: emergent powers and the unity of lives.* There are at least two ways one might respond to these problems with respect to organisms. One view is that living organisms possess emergent and essentially unitary powers (cf. Section 22.6). The persistence of an organism must then be a metaphysically fundamental fact, since it is the basis for the persistence of the emergent and unitary powers. If an organism did not survive, its powers would not survive, and the world would be different at a fundamental level. The second view is that biological lives are fundamental entities and that the persistence of the life of a given organism grounds the persistence of that organism itself. These fundamental facts about powers or lives provide an answer to each of the paradoxes. (We ignore the possibility of intermittent existence, since it just seems impossible for organisms to exist intermittently.)

1. The ship of Theseus. If a human being were reconstructed from the sloughed off parts of a human being, then the reconstructed human being would not be the same as the continuously existing human. What enables a human to survive is the continuous survival of either the emergent powers of human life or the life itself, not the persistence of the same parts.

2. Tib and Tibbles. Since what provides unity to a composite object are its unitary powers or its biological life, one can deny that there is any such fundamental thing as Tib prior to Tibbles's losing its leg. Fundamentally speaking, there is no such thing as Tib before the amputation or Tibbles after the amputation, where these are thought of as mere collections of atoms, since neither collection of atoms corresponds to a single, unified process of rabbit-life. Tib corresponds to only part of the rabbit's life before the amputation, and Tibbles includes atoms that are separated from the rabbit's life after the amputation. Thus the puzzles about the identity of Tib and Tibbles do not arise.

3. The Kafka Paradox. There will be some point at which the removal of a particle fatally disrupts the persistence of the powers essential to the human life or the life itself. At that point, one organism is destroyed, with the possibility that a new organism, with a new set of powers, is generated. There may be no way to know at what point a life will or does end, but that does not preclude there being a metaphysical fact of the matter. It is that metaphysical fact that is important here.

4. Fission and fusion. Like with the Kafka Paradox, in cases of fission or fusion, the fundamental facts about the persistence of emergent powers or biological lives will settle the question of whether the original person has survived, and if so, whether he has survived as Beta or Gamma. We may not always be able to tell whether it is Beta or Gamma or neither that is identical to Alpha, but there will nonetheless always be an

answer. The answer will depend on how to explain the emergent powers or the lives of Beta and Gamma. In at most one of these cases can we explain the existence of the fission-product's emergent powers or life in terms of the persistence of Alpha's emergent powers or its life. For Alpha to persist is for it to persist with its essential powers or its life. Hence, persistence can be used to provide a metaphysical explanation for the presence of those same powers or that same life in the later entity. The kind of explanation involved is not explanation in terms of causes or the laws of nature but metaphysical explanation: giving the ultimate ground or truthmaker for the humanity of Beta or Gamma. Since it is metaphysically impossible for Alpha to persist as *both* Beta and Gamma, it is impossible for the persistence of Alpha to be the truthmaker or the metaphysical ground for the presence of the essentially human bundle of powers or essentially human life in *both* Beta and Gamma. The persistence of Alpha might ground the existence of one or the other or it might ground neither of them. In the first case, Alpha would be identical to one and not the other, and in the second case it would be identical with neither one.

Of course, this sort of solution to the problem of mereologically inconstant organisms depends upon a claim that could, in principle, be falsified by future biochemical and biophysical research, namely, the claim that there are genuinely emergent and essentially unitary powers (both active and passive) to be found at the biological level (see Section 22.6 for further discussion of this issue).

EXOTIC OBJECTS AND VAGUE IDENTITY Exotic objects and vague identity are no threat, since in neither of these cases do we have the survival of any emergent powers or of the same biological life.

*25.1.3.2 In defense of artifacts: the continuous history theory.* One account of the persistence and unity of an artifact exploits certain practices of use and maintenance. For example, if a watch persists, it is because there is a certain ongoing history of use of the watch as a watch and of maintenance of the watch as a watch. Let's suppose that these social processes or practices have a kind of unity through time, that the process as a whole is metaphysically fundamental, not the various instantaneous events that make up the process. The maintenance and use of a single, persisting watch is greater than the sum of its constituent events of watch-use and watch-repair. If that is so, then the metaphysical unity of the process over time can be used to ground the persistence of the artifact. This is the "Continuous History Theory" of artifactual persistence.

1. The ship of Theseus. Let's call the ship that has been in continuous operation 'Theseus-A', and the ship that is reconstructed from the abandoned planks 'Theseus-B'. The Continuous History Theory entails that it is Theseus-A and not Theseus-B that is identical to the ship as originally built. Theseus-A is associated with a continuous process of nautical use and maintenance, which is not the case with Theseus-B. When the abandoned planks are put together into a ship, a new ship is created because a new practice of use and maintenance is initiated.

2. Tib and Tibbles. There can be cases like Tib and Tibbles for artifacts. Consider, for example, a desk (the analogue of Tibbles) and a desk minus some small part of its top (the analogue of Tib). Plausibly, the desk minus some small part of its top is not what is

being *used* as a desk. The *whole* desk is used as a desk. Thus, one can deny that there is any such thing as the desk minus some small part of its top, prior to that chunk of the top being separated from the desk itself. Given this, the puzzles about the identity of desk and desk minus do not arise.

3. The Kafka Paradox. Each kind of artifact has its own sort of associated process. The essential characteristics of the artifact are mirrored by essential characteristics of the process of use. In a Kafka-like paradox for artifacts, if at some point in the transformation this process of use is fatally disrupted, the artifact ceases to be (with the possibility that a new artifact is created).

4. Fission and fusion. This is perhaps the most difficult case for the Continuous History Theory. We would have a simple answer if we could say that the relevant processes of use can themselves never be divided or fused together, but an opponent might object that we have no principled reason for saying so. The answer would have to be that it is essential to the processes involved that they remain focused on a single system, and it is also essential to them that they cannot be merged with any other process without being disrupted. Whenever that focus is lost or another process is brought into a complete intersection, the original process is brought to a halt, and the artifact is destroyed.

ARTIFACTS AND VAGUE IDENTITY We might here take the view that vagueness is merely a reflection of our ignorance. If we understood all there was to know about the processes of use and maintenance that are involved, we might always know the right thing to say about whether the artifact, like the restaurant Bookbinders, persists or not. Alternatively, we might suppose that the vagueness is ontological. Maybe restaurants are simply vague objects, sometimes indeterminate in existence or in identity.

ARTIFACTS AND CONVENTIONALITY The artifactual objects that exist do depend on our concepts and conventions, since those concepts and conventions shape our practices, and it is our practices that are the ground of persistence for artifacts. However, there may be natural limits to the kind of social practices that can exist. It is hard to imagine a set of social practices that would really die out or begin to exist simply by driving a car out of or into a garage. A car just isn't the sort of thing that can be built or maintained in that way.

There is an obvious objection that could be raised to the Continuous History Theory: doesn't it simply push the problem of persistence back a step? What is the principle that unifies the various spatial and temporal parts of a single practice of use and maintenance? Don't such practices simply correspond to myriads of overlapping microphysical processes, with no sharp boundaries in time or space?

These are deep questions, but the defenders of artifacts might well claim that social practices, including the practice of using and maintaining a particular artifact like a car or a watch, have emergent and strongly unitary powers, just as do living and sentient organisms. It certainly isn't obvious that all of the powers of such social practices are wholly grounded (without remainder) in a host of chemical and microphysical processes.

*25.1.3.3 In defense of natural formations: the anthropocentric theory and natural processes.* It is much harder to find a principled answer to the puzzles about the persistence of natural formations. This is one of the reasons that Heraclitus's paradoxical claim that one cannot step twice into the same river continues to resonate with us so many thousands of years later.

One possibility would be to give an anthropocentric account. Natural formations sometimes become part of our human practices in such a way that we can think of them as something like artifacts. Take the stars for example. The visible stars, organized into constellations, have played an important role in human life for thousands of years. We could ground the persistence conditions of stars in those practices of use. This would mean that a star that ceases to be visible, say by turning into a red giant or a neutron star, would really cease to exist as a persisting whole. Similarly, rivers would only come into existence once they have been discovered by us and incorporated into our practices, by way of becoming a means of transportation or a border marker. This approach clearly has its limitations. For one thing, it would preserve only a fraction of our ordinary beliefs about formations as persisting things. It would also have some odd consequences: for example, it would mean that stars and planets first come into existence (as fundamental entities) only when they are discovered and used in some way by human astronomers. An uninhabited planet might have large quantities of water on its surface, but it couldn't have oceans, lakes, or rivers without intelligent agents that make use of them.

A second account of natural formations would avoid these counterintuitive consequences of the anthropocentric theory. On the alternative account, the persistence of a natural formation depends on the metaphysical unity of some underlying natural process. For example, the existence of a normal star is constituted by a single, nearly continuous process of gravity-induced nuclear fusion at its core. Similarly, a river consists of a continuous process of the flow of liquid water. If these astronomological, hydrological, and geological processes have a real unity, then the diachronic unity of the formations can be parasitic on the unity of the underlying process, as organisms depend for their unity on the unity of their lives and artifacts on the unity of processes of maintenance and use.

The greatest challenge to this account comes from the plausibility of microphysicalism. Are there really any macrophysical processes, like astronomical or geological processes, over and above the constitutive microphysical interactions? If the macrophysical processes supervene entirely on the activities and interactions of microparticles, then Ockham's Razor (**PMeth 1**) provides grounds for denying the real or fundamental existence of macrophysical processes. And without real or fundamental macrophysical processes, we can't ground the persistence of macrophysical formations.

*25.1.3.4 Another sort of defense: Plenitudinous answers.* As we have seen, Physicalists or reductive materialists—those who believe that all fundamental powers are microphysical powers—have good reason to reject the theories we've offered on behalf of inconstant objects. Such physicalists face a choice. They must either deny that any inconstant objects are fundamental or offer a different sort of solution to the puzzles.

In search of a new sort of solution, we could return to Temporal Plenitude (24.3T.1). On this view, every spacetime worm corresponds to some kind of persisting thing. As we saw, most of these worms do not correspond to quality or dynamically first-class objects, but we already have learned that mereologically inconstant things, if they exist at all, are not quality objects. Temporal Plenitude populates the world with a very large number of complex, mereologically inconstant things. Each time-slice of any one of these things is shared by an innumerable multitude of temporally coincident objects, as the many different spacetime worms converge upon and diverge from that instantaneous object. Temporal Plenitude therefore implies Temporal Coincidence (24.4A).

On this view, RCK is not a single living organism with a single living body. Instead, he has trillions of mostly coincident human bodies, some older, some younger, some destined to live a long time, some ending at the present moment. If we allow discontinuous worms to correspond to living bodies, then there is an organism with RCK's past and present and THP's future, and another one with the temporal parts reversed.

It is pretty easy to see how Temporal Plenitude dissolves all of the puzzles. The ship of Theseus was really two ships to begin with, Theseus-A and Theseus-B. These two ships are spatially co-located at the beginning of the story and eventually become separated. The rabbit is in fact many rabbity-things, Tib and Tibbles included. When the leg is amputated, Tibbles is destroyed but Tib continues living. As Kafka is transformed into a cockroach, we see a large number of temporally overlapping organisms come and go. At first, Kafka is only human, and at some point we have both a human being and something sub-human that perfectly coincide. Eventually, the human being ceases to exist and only sub-human things are left, until eventually only cockroachy things remain.

In the case of fission, the original human being was really two temporally overlapping human beings all along. Both Beta and Gamma were in existence right from the start, and the two were never strictly identical to each other, since each one was destined to have a different future. The same story can be run in reverse in the case of fusion: both Alpha and Beta continue to exist after the fusion and they remain distinct from one another, even though they come to share the same time-slices.

In the case of vague identity, we can say that each artifact is actually a host of distinct artifacts, each with slightly different persistence conditions. We can say that *this* Bookbinders restaurant is the same as *that* one, or we can say that they are different. In both cases we will be right because there are more than three restaurants involved: one that is located in both places at different times, and two more that are each located in only one place. The vagueness affects only our use of the name 'Bookbinders', a name we can legitimately apply to any of the relevant restaurants. We can use the same name to refer to different restaurants on different occasions without any danger of confusion, so long as the many restaurants involved all perfectly coincide during the relevant period of time.

Finally, exotic objects are no problem: they all exist! All of them and more, ones we can imagine and ones that are so complex and weird that we could never imagine them.

## 25.2 Coincident Things

As we have seen, by adopting Temporal Plenitude and Temporal Coincidence, a number of paradoxes involving mereological inconstancy can be dissolved. At the same time, there is something deeply counterintuitive about such coincidence. It preserves the existence and persistence of persons but only at the cost of utterly shattering the unity of the person. Whenever there is a thinking person, Temporal Coincidence entails that there are really a huge number of coincident and nearly coincident thinkers, all associated with the same body (or nearly the same body) at the same time. It as if Descartes should have reasoned, *We* think, and therefore we are. The first-person singular would never be appropriate.

There are actually two somewhat different conceptions of coincidence. One involves the sharing of a time-slice or instantaneous temporal part. This is Temporal Coincidence. A second kind of coincidence involves having exactly the same parts (whether spatially or temporally extended or both) at the same time:

**Def D25.1 Temporary Mereological Coincidence.** $x$ and $y$ are *coincident* entities at time $t$ if and only if $x \neq y$ and $x$ and $y$ have exactly the same proper parts at $t$ (at some sufficiently small scale of decomposition).

**25.2T Possibility of Temporary Mereological Coincidence.** It is possible for there to be two coincident entities.

**25.2A Impossibility of Temporary Mereological Coincidence.** It is impossible for there to be two coincident entities.

On some views, Temporal Coincidence and Temporary Mereological Coincidence are exactly the same things. In particular, Replacementists will see little or no difference between the two, since for Replacementists all of the fundamental things are instantaneous. The time-slice of a thing includes all of the instantaneous parts of the thing that co-exist in the same instant, and so, for Replacementists, the time-slice includes everything that is both fundamental and a part of the thing at that time. However, Endurantists (24.1T.1T.1T.1T) see an important difference between the two. Endurantists reject the very idea of instantaneous parts, and so the idea of a *time-slice*. If there are no time-slices, then there is no such thing as Temporal Coincidence (for Time-Slice Substratists 24.1T.1T.1T.1A, matters complicated in ways that would not be helpful to inventory here). However, Temporary Mereological Coincidence makes sense for all Substratists and Replacementists. Consequently, we will focus on the issue of whether temporary mereological coincidence is possible.

Besides the cases of temporally coincident objects that we considered in the last section, the most compelling case of temporary mereological coincidence is that of composite things and their matter. We, therefore, turn to that case.

### 25.2.1 Composite things and their matter

Consider Statue, which is a statue. Statue is made of Lump, which is a mass of clay. Are Statue and Lump one and the same thing? It seems that they cannot be. For example, Lump may have existed for millions of years, while Statue was just made today. Similarly, Statue could be destroyed without destroying Lump. Even if both Statue and Lump have existed forever and will exist forever, they can still differ in many of their properties. Statue is the sort of thing that would not survive being smashed into a ball, while Lump is the sort of thing that would survive such smashing. Statue might have the property of *being classical in style*, while Lump has no style, at least not essentially. By Leibniz's Law, Statue and Lump cannot be identical because they have different properties. They must be distinct.

So, it seems that we have two things that perfectly coincide with one another during some stretch of time (perhaps only for an instant if Statue is very short-lived). Not only do they coincide in space; they have exactly the same atomic parts or if they consist of gunk, the very same gunky parts (at some sufficiently small scale). How can two things made of the same parts be different? How can they have different powers, whether active, passive, or immanent?

What is the problem with distinct, coincident entities? Not the overcrowding of the physical space involved. There is no general rule that distinct physical entities cannot occupy the same space. It is a particular law of physics, the Pauli exclusion principle, that gives ordinary matter (composed of fermions) its power of mutual exclusivity. The Pauli principle forbids two bosons (including protons, neutrons, and electrons, which make up the bulk of ordinary matter) from occupying the same quantum state at the same place. However, non-fermions, like photons and neutrinos, are under no such prohibition, despite the fact that they are no less physical.

Coincidence is somewhat more problematic when living organisms or conscious agents are involved. Suppose there are two distinct but coincident people before fission (or after fusion). Each has two hands and a heart and each weighs 150 pounds. Ordinarily, that would mean that we would have four hands, two hearts, and at least 300 pounds of weight altogether, and yet these inferences clearly fail in this case.

The defenders of coincidence could respond that these cases are just extreme versions of Siamese twins. Just as two Siamese twins can share a certain amount of skin and flesh, and even an entire limb, so two coincident human beings can share all of their material parts and all of their mass.

But what about mental acts and passions? Suppose the coincident people experience a red sensation, form a mathematical thought or freely choose between two options. Do we have two sensations, two thoughts, and two acts of choice or just one of each? Sensations, thoughts, and choices do not seem to be the sorts of thing that could be shared by distinct persons.

An even greater metaphysical problem with coincidence is that of shared composition. If two entities are composites and the two have exactly the same basic parts, in what could their distinct existences consist? If we answer that the distinct existences consist in their distinct persistence conditions or in their belonging to distinct sorts, then we merely push the problem back a step. If two entities are composites with exactly the same basic parts, how can the two have different persistence conditions or belong to different sorts?

### 25.2.2 The grounding objection to coincidence

The most important objection to the Possibility of Temporary Mereological Coincidence is the *grounding objection*: if coincident objects differ in their modal properties (including their persistence conditions), what grounds this difference? By definition, two coincident objects are composed of exactly the same proper parts during some period of time. Those parts are related to each other in a certain way during that period. Obviously, this arrangement of the parts is also shared by the two coincident objects—same parts, in the same arrangement. How can the two differ? If the two composite objects are wholly grounded in the parts and their arrangement, and if grounding involves entailment (as asserted by 3.8T), then it must follow that both composite objects have exactly the same properties, including their modal properties, the properties that determine their persistence.

We can put the objection another way. The standard view of coincidence is one developed by David Wiggins (originally 1980, updated in Wiggins 2001). In Wiggins's account, two coincident objects belong to different kinds called 'sortals'. For example, Statue belongs to the sortal *statue* or *art work*, while Lump belongs to the sortal *piece of clay* or *piece of matter*. There are persistence conditions associated with sortals. Typically, different sortals will have different persistence conditions. A piece of matter can survive being smashed or ground into dust; a statue cannot. These persistence conditions, in turn, correspond to modal properties of the objects that belong to them. Thus, statues and pieces of clay have different modal properties. The piece of clay has the modal property of *being possibly in a condition of being perfectly flattened*; the statue lacks such a property. Now we can ask what makes the clay have this property and prevents the statue from doing so. What facts about the piece of clay are these modal properties grounded in?

The defender of the Possibility of Temporary Mereological Coincidence must say that the material parts of the clay and the statue have certain properties, like the property of *possibly composing something with the modal properties of a piece of clay* and the property of *possibly composing something with the modal properties of a statue*. When the material pieces are in the right condition, both of these potentialities are realized, resulting in the existence of two new objects. For the moment, let's call the objects "Tweedledee" and "Tweedledum". Now, we know that one of these objects must be a statue and the other must be a piece of clay. What is responsible for making Tweedledee the statue and Tweedledum the piece of clay or vice versa? Both Tweedledee and Tweedledum stand in exactly the same causal relation to the same material bits. The defenders of Coincidence cannot explain how they end with such different properties.

Probably the best answer for the defender of Coincidence is to claim that the two resultant objects stand in different causal relations to the underlying parts, since the parts have two distinct joint powers: one power of producing a piece of clay and another power of producing a statue. These powers are instances of what Aristotle called 'material causation'. However, this Aristotelian solution is one that involves a significant cost. If defenders of the Possibility of Temporary Mereological Coincidence want to claim that there are billions and trillions of coincident objects at each time and place, they will have to attribute a myriad of primitive material powers to the atoms.

The other approach to defending the Possibility of Temporary Mereological Coincidence relies instead on a deflationary understanding of the coincident objects. Priority Atomists (22.7T), who believe that no composite object is metaphysically fundamental, can countenance a large number of composite and coincident objects, since they do not believe that such objects make up any part of the fundamental structure of reality. This is especially so if the Priority Atomists take composite things to be *conceptually* grounded in their parts (as per Chapter 3.4), since this implies that the composite objects do not really exist. If we believe we can give a complete and adequate description of reality without mentioning composite and coincident objects, then we can treat our beliefs in coincident objects as Ontological Free Lunches. The large number of such objects is no longer of much concern, and we wouldn't need to resort to any special kind of causal explanation to account for their existence.

However, this ontological deflation comes at a high cost, since we (or at least our bodies) are among the entities whose pretensions to reality are being deflated. Can one reasonably believe that a perfectly complete description of reality could utterly omit to mention oneself?

### 25.2.3 Avoiding coincident entities: Nihilism and Near-Nihilism

The arguments for coincident entities depend on some principle of mereological composition. Obviously, Mereological Nihilists will be under no pressure to posit temporally coincident entities. The world would consist only of simple things, metaphysical atoms. If there are no composite things, then there are no mereologically inconstant things and so no coincidence.

Peter van Inwagen (1990a) has argued that it is obvious that there are mereologically inconstant things, namely, living organisms. He uses the existence of mereologically inconstant organisms as part of an argument against Mereological Universalism and in favor of a position that could be described as 'Near-Nihilism'. Van Inwagen denies the existence of all composite entities except for living organisms. Van Inwagen argues that if we were to accept Mereological Universalism, it would be impossible for any mereologically inconstant thing to exist. Consider again the Tib and Tibbles paradox. If we are Mereological Universalists, we should deny that any persisting thing has lost any part in this scenario. All that has happened is that one persisting thing, Tib, has become isolated from another, Tibbles, and that Tibbles has gone from a state of *being connected* to *being scattered*.

Van Inwagen claims that in order to believe in an inconstant object, we must give up Mereological Universalism. Then we can say that, both before and after the removal of the leg, there are only two kinds of things in existence: simple atoms and the rabbit. Both persist, but the rabbit loses some of the atoms as parts. There is no such thing as Tib before the leg is removed, nor such a thing as Tibbles after the leg is removed. The three-legged rabbit and the separated foreleg do not together compose anything, and the set of parts that compose Tibbles after the loss of the leg do not compose anything before the leg is removed. The only parts the rabbit has are atoms: it has no composite parts at all.

Van Inwagen's solution eliminates coincident entities. There is only the rabbit and the atomic parts. The atomic parts do not compose any entity other than the rabbit: there are no incomplete rabbits and no wholes composed of scattered rabbit-parts.

### 25.2.4 Avoiding coincidence through dominant sortals

Michael Burke (1994) proposed a position that is somewhat more liberal than van Inwagen's. In effect, van Inwagen proposed that there is only one sortal that picks out composite entities, namely *organism*. Burke proposed instead that there might be many such sortals, so long as they can be linearly ordered. Take any two sortals: one will always *dominate* the other. The object belonging to the less dominant sortal fails to exist whenever the more dominant sortal is instantiated. Lump ceases to exist when Statue is created. Tib ceases to exist when the Tibbles's leg is detached. Thus, Burke's account is, like van Inwagen's, able to avoid temporary coincidence entirely.

However, Burke's account is no panacea. Like Wiggins's account, Burke's account doesn't apply to puzzles of coincidence involving things of the same sort. Thus, it won't help with cases of fission or fusion or vague identity. In addition, Burke has to make some implausible claims about things' ceasing to exist. Why doesn't Tib persist? Burke replies that before amputation, it was only a part of a cat (and so not a living thing); in order to exist before the amputation, it would have to have been a cat (and so a living thing). However, as Sider (2001: 163–164) points out, Tib was *close to* being a living thing before amputation. How can something as minor and extrinsic as removing the tail bring Tib into being?

Burke also owes us an account of dominance that will guarantee that in every case we can find a maximally dominant sort. Burke suggests, "The dominant sortal is the one whose satisfaction entails the possession of the widest range of properties" (Burke 1994: 610). The "range" of properties must be measured by the number and importance of the categories to which the properties belong: physical, chemical, biological, ethical, and so on. How can Burke guarantee that one of each pair of sortals will always be maximal in the range of properties it entails? Michael Rea (1996) has offered a number of counterexamples of temporarily coincident objects whose sortals cannot be ranked. For example, there might be a statue that is also a pillar, an axe that is also a hammer, or a human being that is also a chess piece.

### 25.2.5 An Aristotelian solution to apparent coincidence

Aristotelians reject the Possibility of Temporary Mereological Coincidence. Aristotle seems to occupy a position between van Inwagen and Burke. He affirmed, in effect, two sortals, namely, *homogeneous blobs of matter* and *organisms*. Consequently, there are only three problem cases for Aristotelians to consider: (i) cases in which there are two homogeneous blobs that are temporarily coincident, (ii) cases in which two organisms are temporarily coincident, and (iii) cases in which an organism is temporarily coincident with a homogeneous blob. Aristotelians can plausibly claim that all three are metaphysically impossible.

**Case (i): two homogeneous blobs.** If two blobs have all the same proper parts at some time, then they are simply identical.

**Case (ii): two organisms.** The proper parts of an organism are all metaphysically dependent on the organism. They have the powers they do and they persist through time by virtue of the persisting nature of the organism as a whole. Thus, it is impossible for any material body to be a proper part of two different organisms at a time. In other words, for Aristotelians, it is impossible for two organisms even to overlap, much less to coincide.

**Case (iii): an organism and a homogeneous blob.** First of all, it is far from obvious that it is possible for any organisms to consist of a single, homogeneous blob. Second, even if this is possible, the Aristotelians could argue a blob is destroyed whenever an organism is engendered from it and that a new blob is generated whenever the organism is destroyed. The material parts of an organism have their identities, powers, and natures tied essentially to the organism, while apparently similar material blobs that are not parts of any organism have autonomous identities, powers, and natures. This profound difference provides Aristotelians with grounds for denying the persistence of the underlying blobs.

### 25.2.6 Replacementist solutions: temporal counterparts

Since the van Inwagenian and Aristotelian solutions require some heavy metaphysical commitments to the fundamental existence of organisms and the non-existence of the material parts of organisms (as fundamental entities), those solutions are not available to metaphysicians who wish to maintain the universal fundamentality of merely physical things. Such physicalists must again look to Replacementism, in one of two forms: Worm Theory or Stage Theory. Both views are committed to Ramsey-Lewis-Sider Perdurantism (24.3T.3, which we will re-label as '25.3T') and agree that there is an important connection between persistence and spacetime worms (Def 24.6). The difference concerns the identity of what we think of as persisting objects. According to Worm Theory, persisting objects are identical to spacetime worms. According to Stage Theory, persisting objects are identical to temporal parts of spacetime worms.

**25.3T Ramsey-Lewis-Sider Perdurantism.** A spacetime worm *S* over interval *T* corresponds to the existence of a derived thing persisting through *T* if and only if the simplest and most powerful scientific theory of the actual world assigns a persisting entity to *S*. [=24.3T.3]

**25.3T.1 Worm Theory.** Ramsey-Lewis-Sider Perdurantism is true, and persisting objects are identical to spacetime worms.

**25.3A.2 Stage Theory.** Ramsey-Lewis-Sider Perdurantism is true, and objects that can rightly be said to 'persist' are identical to temporal parts of spacetime worms.

As Sider (2001: 191–192) puts it, the difference between Worm Theory and the Stage Theory lies in the realm of semantics rather than metaphysics. Both agree exactly about what sorts of things exist. The difference between the two lies simply in their varying accounts of the correct truth-conditions to give to ordinary language assertions about the existence and diachronic identity of persisting things. For Stage Theorists, names pick out temporal parts; for Worm Theorists, they pick out spacetime worms.

Worm Theory has an obvious solution to all of these problems. Distinct spacetime worms that wholly (or nearly wholly) share a common segment throughout some temporal interval. However, Worm Theory cannot explain how Statue and Lump could differ if they in fact coincide throughout their entire lifetimes. If, in other words, the clay and the statue were created and destroyed at exactly the same times, they would still be distinct worms because they would still differ in their modal properties: Lump could have survived crushing, but Statue could not.

Here, Worm Theorists (and Stage Theorists as well) must appeal to Counterpart Theory (16.1A.1). Statue and Lump have different counterparts in other possible worlds.

Specifically, in a world in which Statue/Lump are crushed, the counterpart of Lump in that world persists, while the counterpart of Statue does not. Worm Theorists could take the relata of the *counterpart* relation to be spacetime worms in the various worlds if they then insist that different *counterpart* relations be used for different sortals (like *statue* and *piece of clay*). Thus, Counterpart Theory offers Worm Theorists a simple response to the grounding objection: it is our concepts and interests that make different *counterpart* relations relevant for statues and for lumps of clay. There need be no deep, metaphysical facts about which ground the difference.

A great advantage to Stage Theory is that it can make use of *counterpart* relations both for *de re* modality and for temporal persistence. A given time-slice can, under a given sortal, have both temporal counterparts at other times in the same world and modal counterparts in other worlds.

Peter Unger's Problem of the Many (Unger 1980) illustrates some common ground between Worm Theory and Stage Theory. Suppose that we have a single, vaguely composed and vaguely located object. On both theories, there are multiple, equally good candidates for the job of being the referent of the relevant singular term (that is, all of the spacetime worms that include candidate precisifications of the object's location and composition or all of the contemporary stages of those worms). In giving a semantic account of assertions involving the term, Stage Theorists have two choices: (i) a supervaluationist semantics, in which assertions come out as *super-true* or *super-false* just in case they are true (or false) under each of the candidate referents or (ii) a semantics that makes use of a notion of *approximate truth* or *near truth*. On the second account, it is literally true that there are trillions of persons in the room, but it is almost true that there are 14, since there are 14 clusters of nearly-identical, highly-coincident spacetime worms.

The difference in the two accounts comes out in the case of persistence through time. Unlike Worm Theorists, Stage Theorists deny that people, organisms, artifacts, and other ordinary objects *really* persist through time at all. Instead, each of necessity exists for only an instant. For this reason, Stage Theorists must produce a semantic account of our ordinary assertions of persistence that enables most of them to come out as true. This is where a temporalized version of Counterpart Theory comes into play. To say that entity $x$ will have some property $P$ at future time $t_2$ is to say that there is a counterpart of $x$ that exists at $t_2$ and that has property $P$. In other words, ordinary objects "persist" by having temporal counterparts at other times. In addition, Stage Theorists can make use of a family of *counterpart* relations, each corresponding to a different sortal.

Sider argues (Sider 2001: 194) that Stage Theorists can be neutral on the question of whether there are brute facts involving such *counterpart* relations or whether all such *counterpart* relations must be reducible to (or at least supervenient on) other, purely intrinsic and qualitative or spatiotemporal relations. (This is an issue that divides Classical Perdurantists 24.1T.1T.1A.1T from Classical Genidentity Theorists 24.1T.1T.1A.1A.) That is true, but a primitive *counterpart* relation, or perhaps a family of primitive *counterpart* relations, each bearing some further primitive relation to some sortal, would seem to be a highly mysterious and obscure sort of thing. In fact, it's clear that Sider believes that *counterpart* relations are reducible (via some analytic meaning postulates) to some logical complex of other relations.

A difficult problem for Stage Theory is that of cross-time counting. For example: fewer than a trillion people have lived in North America. Strictly speaking, this is false if Stage

Theory is true, since there have been infinitely many instantaneous person-stages living in North America (or, more precisely, in past continent-counterparts of North America). At this point, Sider recommends a hybrid theory, according to which the word 'person' sometimes picks out person-stages and sometimes person-worms (Sider 2001: 197).

How do Stage Theorists deal with the various paradoxes? In the case of Statue and Lump, they simply make use of two different *counterpart* relations—one for statues (or artifacts in general) and another for lumps of clay (masses of stuff). While they coincide, Statue and Lump are fully identical. However, use of the two names 'Statue' and 'Lump' will trigger (as a pragmatic matter) the use of different *counterpart* relations, making it possible to say that Lump will, and Statue will not, survive the smashing. The paradox ofTib and Tibbles) can be dissolved in a similar manner. There is one *counterpart* relation for cats (organisms) and another for sums of atoms or other cat-parts.

In the case of the fission and fusion paradoxes, Sider argues (2001: 202–203) that Stage Theory is superior to Worm Theory. Worm Theory fails to satisfy Fission-Product Interest below, given the assumption that Interest Requires Genidentity (a thesis suggested by the work of Derek Parfit 1985):

**Fission-Product Interest.** For any person $x$ each of $x$'s future fission products matters to $x$.

**Interest Requires Genidentity.** Person $x$ matters to person $y$ if and only if $x$ and $y$ are genidentical.

Suppose Alpha will undergo fission, and that two fission-products, Beta and Gamma, will result. On Worm Theory, there are, before fission, two distinct, overlapping persons corresponding prospectively to Beta and Gamma. Beta matters only to the first, Gamma only to the second. On Stage Theory, on the other hand, at each time pre-fission, there is only one person counterpart-related to both Beta and Gamma.

In the case of fusion, we have a similar difference. For Worm Theory, there are two distinct, coinciding persons post-fusion, whereas on Stage Theory there is only person, counterpart-related to the two prior branches. Sider sees this as another advantage of Stage Theory, remarking that it would be just to punish the post-fusion person for crimes committed by either pre-fusion branch. Not everyone will share this intuition, reflecting the lack of grip that the fusion story has on our considered judgments.

Stage Theory has no difficulty with the longevity or Methuselah paradox (see Section 24.2.2.2), since nothing prevents our using a non-transitive *counterpart* relation. In fact, such non-transitive cases are tailor-made for the use of temporal counterparts.

The paradoxes of vague and conventional identity similarly pose no problem for Stage Theory, since there are plenty of eligible candidates for the interpretation of names and singular terms, corresponding either to non-functional *counterpart* relations or to alternative eligible *counterpart* relations. Temporal Counterpart Theory offers a unified account of temporal and modal phenomena, since it is readily combinable with cross-world, modal *counterpart* relations. The modal version of the Statue/Lump case, in which Statue and Lump coincide throughout their actual life-spans, is thus no problem.

In fact, the paradoxes and puzzles dissolve so readily under Stage Theory that one has to suspect that Stage Theory is simply too good to be true, as Dean Zimmerman has suggested (reported in Sider 2001: 206–207). The nature and number of relevant *counterpart* relations is so unconstrained that it is hard to imagine any combination of intuitions about individual cases that Stage Theory couldn't easily accommodate. A very similar phenomenon is evident in the case of modal Counterpart Theory as developed in Lewis (1986a). We would put Zimmerman's complaint this way: Counterpart Theory is so flexible that, not only does it permit us to treat our ontological and modal intuitions as reliable, it actually renders them infallible in principle. This kind of epistemological overkill is exactly analogous to the mistake of the logical positivists and Phenomenalists (13.2T), who so tailored their metaphysical theories to our experience as to render all assertions of protocol (observation) sentences true by definition. (In Chapter 5, we voiced similar objections to the Ramsey/Lewis Theory of laws of nature, which renders our non-empirical criteria for theory choice infallible.) A complaint like Zimmerman's relies on the following principle of meta-metaphysics:

**Reasonable Reliability.** Any metaphysical theory that entails that we are either more or less reliable than we have a right to suppose is to be rejected.

Thus, metaphysical theories that entail unreasonable forms of skepticism or of infallibilism are equally objectionable. Let's illustrate some of the unpalatable consequences of the flexibility of Counterpart Theory. Imagine some metaphysically sophisticated defenders of slavery in the antebellum South, who regularly insist that slavery is just because all slaves are natural slaves. They defend the latter thesis by claiming that all slaves are essentially slaves. Moreover, they claim that emancipation is tantamount to murder, since no slave could survive such a process. It is child's play to show that what the slavery apologists say is true, given Counterpart Theory. All we have to do is to make use of a Counterpart Theory that provides each actual slave with no free counterparts, and each present slave with no future, free counterparts. Similarly, it would be easy to defend a position that rejects any anti-poverty measures by simply asserting that poor people are essentially poor and that no poor person could survive a significant increase in income. Sider could certainly reply that the use of such *counterpart* relations is wicked and reprehensible, but he could not take seriously even the possibility that their users were guilty of any intellectual error.

Finally, we should bear in mind that both Worm Theory and Stage Theory inherit all of the problems with Replacementism discussed in Chapter 24, especially the problem of making sense of the intrinsicality of motion.

## 25.3 Conclusion

In this chapter, we considered two questions about the relation between composition and persistence, namely those of constancy and coincidence. We canvassed important puzzles and paradoxes for those who believe that mereological inconstancy and mereological coincidence are possible. Three views emerged as front-runners, though none are absolutely unscathed. First, there is van Inwagen's Near-Nihilism, which countenances only material simples and organisms. Second is an Aristotelian view nearby Near-Nihilism that countenances homogeneous blobs as well as simples and organisms. Third, there are Worm and Stage Theory, which are Temporally Plenitudinous Replacementist views most plausibly coupled to Concretism and Counterpart Theory. In the final three chapters, beginning with the next, we will turn to a final set of issues, those concerning the relations of *causation*. There we may find evidence that will tilt our views in one direction or another.

## Note

1 One might object that human beings are not organically symmetrical. The left hemisphere of the brain is significantly different from the right hemisphere, for example. However, let's suppose that Alpha is a member of some humanoid species on another planet that is perfectly symmetrical or that Alpha is an abnormal, perfectly symmetrical human being. Either of these seems to be metaphysically possible.

# Part VIII

# Causation

# 26

# The Existence and Scope of Causation

The nature of causation has been one of the central questions of metaphysics since ancient times. For example, Plato (in *The Laws*, Book X) uses an argument about the cause of motion to prove the existence of God and the soul. Even earlier, Pre-Socratic philosophers like Thales, Anaximander, and Empedocles speculated about the cause of the origin of the universe and of natural phenomena, like the eclipse. We start in Section 26.1 by asking why we should believe that there are any causes at all. What does causation do for us? How does it help us make sense of the world? Then, in Section 26.2, we consider the issue of the scope of causation: how much of the world is caused? Is there a principled reason why some things are caused and others are not? In the following chapter, Chapter 27, we will be ready to take on the nature of causation. Is causation fundamentally a logical relation between truths or propositions or is it a real relation between things? If it is a relation between things, what sort of things—facts or events? Can absences and other negative things be causes? Finally, in Chapter 28 we will look at the relationship between causation and time, distinguishing between *discrete causation* (as an instantaneous relation between two things) and *continuous causation* (involving an infinite number of events in a temporally extended process).

## 26.1 Are there Causes?

We begin with our intuitive, commonsensical ideas of causation and ask: why think that anything causes anything? Should we opt for Causal Realism or Anti-Realism:

**26.1T Causal Realism.** Some things are caused.

**26.1A Causal Anti-Realism.** Nothing is caused.

*The Atlas of Reality: A Comprehensive Guide to Metaphysics*, First Edition.
Robert C. Koons and Timothy H. Pickavance.

### 26.1.1 Arguments against Causal Realism

Let's start by looking at the arguments for Causal Anti-Realism. Historically, there have been two main arguments for this view.

1 Causation not needed in fundamental physics. causation is scientifically obsolete. In 1913, Bertrand Russell published an article (Russell 1913) in which he argued that causation was an obsolete concept. Russell pointed out that modern physics, including Newtonian mechanics and Maxwellian electromagnetic theory as well as Einstein's special and general theories of relativity, describe the natural world by means of mathematical functions and equations. The words 'cause' and 'effect' do not appear in any of the fundamental laws of nature, nor do any of the other verbs that express causal relations, like 'push', 'pull', 'repel', 'attract', 'fracture', 'merge', and so on. The laws simply describe how certain quantities, like mass, energy, position, and velocity, change and do not change over time. To make this an argument for Causal Anti-Realism, one must assume that the best metaphysical theories should only appeal to the fundamental types of things in our best current physical theories. This is a non-trivial assumption.

2 Causation requires necessary connections between separate existences. The second principal objection to Causal Realism builds on the work of David Hume (Hume 1748). Hume observed that our idea of causation seems to involve some notion of a *necessary* or *non-accidental* connection between cause and effect. We think that if the cause is fully present and operational, and there are no obstacles or interferences, then in some sense the effect *must* follow. In addition, we think that if the cause had not occurred, the absence of the cause would have made the occurrence of the effect *impossible* in some sense.

At the same time, Hume argued that our ordinary conception of causation involves the separateness of the cause and effect. We don't think of things' causing themselves, either in whole or in part. Instead, one thing is thought to cause some other, completely separate thing. We think of the occurrence of a spark as causing a subsequent explosion, because the existence of the spark is wholly separate from the existence of the explosion. In contrast, we don't talk of John's loving Mary as the cause of Mary's being loved by John, since the two facts are inseparable, perhaps even identical.

These two facts about causation pose a problem. How can the existence of one thing necessitate the existence of another, separate thing? And how can the non-existence of something necessitate the non-existence of the other? Hume argued that there cannot be necessary connections of this sort. When two separate things are in view, we can always imagine one of the things existing without the other, and vice versa. By Imagination as Guide to Possibility (**PEpist 1**), we should conclude that there are no necessary connections between separate existences.

As we have noted, though, causation seems to be a relation between separate existences. For example, Hume argued that we could always imagine the cause's occurring without the effect (the spark without the explosion) or the effect's occurring without the cause (the explosion without the spark). If causation is a relation between separate existences, then causation cannot be a relation of necessitation. The two components of our intuitive notion of causation—necessity and separateness—cannot both be maintained. Our notion of causation, then, is incoherent.

However, we have seen that imagination is a fallible guide to possibility. We can in certain cases imagine even logical impossibilities, like the non-identity of Mark Twain and Samuel Clemens. Thus, it isn't entirely clear how much stock we should put in our ability to imagine a cause's existing without its effect, and vice versa, and more generally in Hume's claim that there can be no necessary connections between distinct existences. The argument is interesting, no doubt, but may not be ultimately successful, especially given the centrality of causation to our thinking.

Hume had a further, closely related argument against the reality of causation. He argued that our experience provides us with no direct acquaintance with necessary connections like causal relations. When we reflect on our experience of apparent causation, Hume claimed, we find that all we experience is the succession of one event by another. We never see (hear, feel, etc.) the connection between the events. If we have no acquaintance with causation or the associated necessary connections between events, then we must lack even a concept of causation. Our idea of causation is merely a confusion of several distinct concepts, namely, the concepts of regular succession and contiguity among events, and the concept of a felt propensity on our own part to infer one thing from another. Hume argued, in effect, that our concept of causation cannot be defined in terms of other, more fundamental concepts, and we can have no empirical knowledge of instances of causation. By Conceptual Acquaintance (PEpist 5), we ought to deny that we can know the truth of any proposition involving that concept.

Hume's claims have not gone unchallenged. Elizabeth Anscombe (1974), for example, has argued that we do have direct empirical knowledge of causal connections. We see things push, pull, break, hit, and dent other things (to pick just a few examples), and each of these actions involves a causal connection. George Berkeley (1710/2009) argued that, at the very least, we have direct experience of causation in our own case, when we move our own bodies or generate ideas in our minds at will.

### 26.1.2 Arguments for Causal Realism

Causal Realists have offered at least five arguments for their view.

1 Appeal to common sense. We believe that many ordinary events and situations are caused. As Anscombe pointed out, our experience of the world includes the experience of events like boys' kicking balls, rocks' breaking windows, horses' pulling carts, and so on. In addition, we seem to have implicit knowledge of our own agency. We do things, and what we do makes a difference to what subsequently happens. That is, our actions cause things to happen. If this were not so, then most of our beliefs about our everyday lives would be radically mistaken.

The process of making decisions involves a tacit commitment to the causal efficacy of our choices. Here is an example. Suppose that scientists discover that people who say 'Good morning' at least six times every day are likely to live ten years longer, not because saying 'Good morning' has any propensity to make one healthier or safer, but just because people with healthy genes are more likely to say 'Good morning' than those with unhealthy genes. It would be irrational, upon learning about this study, to begin saying 'Good morning' more often, even though doing so might raise the probability of long life. In order to make the right distinctions here between rational and irrational choices, we have to bring into account the effects of one's possible actions. An action is rational when it is believed that it *would cause* favorable results, not merely when it is *correlated* with favorable future events (see Skyrms 1980, Lewis 1981).[1]

2 Scientific practice, especially in the special sciences like medicine, astronomy, biology, and geology. As we mentioned above, Russell pointed out that the laws of physics don't mention causation. Nancy Cartwright (1983, 1994) has argued that the laws of physics "lie" in a certain sense. According to Cartwright, when we abstract the laws from actual contexts and turn them into abstract, mathematical formulas, we enter a fantasy world that is cut off from the actual practice of science. The laws of physics describe the tendencies or propensities of things. When we apply them to any actual setting, we must add various phenomenological "fudge" factors, as well as hedges and exceptions. We can say that two bodies will tend to accelerate toward one another at a rate proportional to the product of their masses and inversely proportional to the square of their distance, except insofar as other factors and forces (like friction or viscosity) interfere. Thus, the application of the laws of physics always involves reference to a network of actual and potential causes.

Moreover, as Cartwright and others have pointed out, the identification of causal factors is a central and indispensable part of all of the so-called *special* sciences, that is, of all of the sciences except for fundamental particle physics. In medicine, we seek the causes of health and disease, in economics the causes of inflation or unemployment, in sociology the causes of social conflict or harmony, and so on.

Finally, the problem of the interpretation of quantum mechanics has brought causal questions to the fore, even in pure, theoretical physics (Healey 1989, Bub 1997). Quantum systems seem to evolve in two quite different ways, sometimes in accordance with the linear dynamics of the Schrödinger equation, and sometimes by way of the collapse of a quantum wave in accordance with the probabilities associated with its instantaneous state. We seem to be forced to ask what causes the collapse of a quantum wave. Is it interaction with a macroscopic measuring instrument, interaction with a conscious observer or is collapse a spontaneous event that occurs with a certain probability? In addition, the Bell inequalities reveal that causal interaction in quantum mechanics is non-local. That is, the occurrence of a measurement in one location can *bring about* a simultaneous collapse in an associated quantum wave at any distance, no matter how great. This non-locality seems to conflict with the constraints of Einstein's relativity theory, according to which no signal or influence can travel faster than the speed of light.

Whatever the ultimate solution may be to the problems of measurement and non-locality in quantum mechanics, the mere fact that these have become central problems for physicists and philosophers of physics demonstrates that questions about causation have not become obsolete (see Cushing and McMullin 1992 and Albert 1994).

3 Causation needed in an adequate account of empirical knowledge. If nothing is caused, then sensory knowledge and memory are impossible. This fact was not always recognized. For a very long time, going back as far as Plato's *Theaetetus*, it was thought that knowledge could be defined as justified true belief. None of these three elements, whether justification, truth, or belief, involve causation in an obvious way. However,

Edmund Gettier (1963) crafted a persuasive refutation of the justified true belief theory of knowledge. Here are some Gettier-inspired examples of justified true belief that are not cases of genuine knowledge:

1 Smith has good evidence that Jones owns a Ford. On the basis of this evidence, Smith infers that Jones owns either a Ford or a Chevrolet. In fact, Jones owns a Chevrolet. Smith's belief that Jones owns a Ford or a Chevrolet is true and justified, but Smith doesn't *know* that Jones owns a Ford or a Chevrolet.
2 Smith is seeing a very lifelike hologram, and she believes that she sees a bright red apple in front of her. In fact, there is a bright red apple in front of Smith, of exactly the right kind and in exactly the right place, but behind an opaque screen. Smith justifiably believe that a red apple is there, and her belief is true, but she doesn't *know* that the apple is there.
3 Smith has completely forgotten a childhood visit to Disneyland, but a memory-like image of the Disney Matterhorn is placed in her mind by hypnosis. Smith justifiably believes that she remembers how the Disney Matterhorn looked, and her pseudo-memory does match exactly the appearance the Disney Matterhorn had at the time, but she does not *know* that it looked like that.
4 Smith justifiably believes that the stock market went up 2% yesterday on the basis of what looks like a copy of the *Wall Street Journal* delivered to her house. In fact, the copy Smith has is a fake produced as a prank by her neighbor. Coincidentally, the market did go up 2% yesterday, but she doesn't *know* that it did.

In each case, it seems that what is missing is a causal connection of the right kind between Smith's belief and the corresponding fact. Smith's belief that Jones owns either a Ford or a Chevrolet is not caused in any way by the fact that Jones does own a Chevrolet. Smith's red-apple-ish visual impression is not caused by the red apple, her memory-like image of Disneyland was not caused by an actual past experience of hers, and the accurate report about the stock market in the fake newspaper was not caused by the actual rise in the market (see Goldman 1977).

4 Causation needed for an adequate account of semantics. In his classic work on the philosophy of language, *Naming and Necessity* (Kripke 1980), Saul Kripke put forward a strong case for the thesis that what we refer to when we use a name depends at least in part on the actual causal connections between our use of the name and the origination of the name, when the first user "baptized" some object as the name's bearer. This Kripkean or causal/historical theory of reference has been further developed and applied to certain uses of descriptions (Donnellan 1966) and demonstratives, like 'this' or 'that' (Kaplan 1989, Evans 1982).

Kripke argues, for example, that one can refer to the great German-American mathematician Kurt Gödel by using the name 'Kurt Gödel', even if one knows virtually nothing about him, and even if the descriptions that you associate with the name really apply to someone else (like the mathematician David Hilbert, for example). What ties one's use of a name with a particular name bearer is a connected history of using the name in a certain way. The end of the historical chain is connected with the bearer by the intention and knowledge of some initial user, and each of the links in the chain involves an intentional connection between the earlier and later use: the later user intends to use the name as part of the ongoing practice as instantiated by the earlier use. Causation seems to play an indispensable role in this theory, both at the beginning and at each intermediate link. The bearer of the name impresses itself somehow upon the senses of the original user of the name, so that he or she is in a position to intend for this name to refer to that particular object, and each use of the name impresses itself in a similar way upon the senses of the next user.

Kripke's picture can be generalized to apply to giving the *semantics* (in a sense) of our mental ideas or concepts. RCK's concept of Dan Bonevac is something like a mental symbol that has Dan as its reference, by virtue of causal connections between Dan's features at various times and the *mental picture* that RCK associates with his concept of Dan as an individual. The causal connection needn't be knowledge-conferring: RCK might think that Dan is taller than he really is, if his sense impressions of him all occurred under misleading circumstances (in the context of some optical illusion, for example). What matters is that RCK thinks of Dan in a certain way because Dan (that very person) affected RCK's senses in the corresponding way and in the appropriate contexts. Thus, a full account of mental reference or intentionality, of the *aboutness* of ideas, seems to require the reality of causation.

5 Causation needed for an adequate account of modality and modal knowledge. If there is no such thing as causation, then there can certainly be no such thing as causal powers. If there are no causal powers, then an Aristotelian account of potentiality or real possibility, like Aristotelian Modality (15.2T.7), cannot be correct. Without the knowledge of the causal powers of things, we would have no reliable basis for beliefs about counterfactual possibilities. We can have no knowledge of how the world as a whole could have been without knowledge of how particular things might have acted or responded in actual past circumstances.

There is a simple reason for this: the world is a very big, very complex place. If the world were much smaller and simpler, we might experience a world that simply cycles over and over through some set of finite states. In such a world, we might notice that world-state *A* is sometimes followed by world-state *B* and at other times it is followed instead by world-state *C*. Observations like this would give us the basis for constructing models of alternative, counterfactual possibilities, like the possibility that the world actually passed from state *A* to *B* to *D*, but it might have passed instead from *A* to *C* to *E*.

However, it is impossible for us to observe anything like that in a world as large and complex as ours. Instead, what we can actually observe is that the world is made of a large number of separate things that fall into a relatively small number of natural kinds. We can observe that the members of each natural kind have certain powers and dispositions in common, and on the basis of this knowledge we can reliably construct models of how things might have gone counterfactually. The reliability of this method requires real powers, or, equivalently, causal laws of nature, which, in turn, require real causation.

Why is it so important that we have knowledge of any counterfactual possibilities? Couldn't we be content with knowing everything there is to know about the actual world? Who needs alternative possible worlds?

First of all, we saw in Chapters 4–6 and 14–16 that our knowledge of possibility is of central importance to both our common sense and our scientific view of the world. In addition, we would argue that we could not understand the actual world without knowing about the existence of other possible worlds. To understand the actual world, we must understand the natures of actual things. Understanding the nature of a thing, in turn, involves knowing, not only the details of its actual history, but also what alternative possibilities are consistent with its essence.

Finally, our knowledge of the semantics of general words and our knowledge of the content of our concepts both involve knowledge of when the word or concept would be true of something and when it wouldn't. To grasp the meaning of a word is to grasp its application-conditions, that is, what something has to be like in order for the word to be true of it. To understand the word 'red' is to know something about a range of possible situations, namely, which of them include something that would fall under the extension of 'red'. Similarly, one has a grasp of the concept of *redness* only if one can recognize, across a wide range of possible situations, which do and which do not include something that falls under that concept. Merely knowing which things are red is not enough. Suppose there were only one red thing in the world. Knowing that that thing is red and nothing else is would not be sufficient for understanding the meaning of the word 'red' or for grasping the concept of *redness*. Some knowledge of possible situations is required.

## 26.2 The Scope of Causation

Given Causal Realism, we can move on to the second question, which is about the scope of causation. How many things are caused? Everything, or only some things?

**26.1T.1T Universal Causation.** Everything is caused.

**26.1T.1A Special Causation.** Some things are caused, and some things are not.

### 26.2.1 Arguments against Universal Causation

There are four objections to Universal Causation, one assuming Mereological Universalism (22.3T.1), a second employing a *pluralized* version of Universal Causation, a third involving a *chain principle*, and a fourth including an argument against infinite causal regresses.

OBJECTION 1: ASSUMING MEREOLOGICAL UNIVERSALISM If Mereological Universalism were true, then the sum of all reality, *R*, would have to have a cause, *C*. But causes of existing things must themselves exist:

**Existence of Causes.** All actual causes exist.

Consequently, *C*, the cause of *R*, would have to be part of *R*. This would make *C* the cause of itself. But nothing can be self-caused:

**Impossibility of Self-Causation.** It is impossible for something (wholly) to cause itself.

We've already seen some reason to deny the possibility of causal loops, namely, the grandfather paradox (see Section 20.5.6). As we saw, Neo-Humeists (4.4T) have a way out of the paradox that is compatible with self-causation, since they can deny the Intrinsicality of Powers (**PMeta 2**).

Another reason for denying self-causation has to do with causal asymmetry. If the direction of causation between two events, which determines which event is the cause and which is the effect, is always reducible to other facts about the two events, then the basis of causal direction must be an asymmetric relation. If we assume that causation is transitive, then any successful reduction of causal direction to an asymmetric relation will impose a "partial order" on the class of events, ruling out any cases of self-causation. However, if causal direction is an irreducible, metaphysically fundamental fact about cause-effect pairs, then it would be much harder to rule out cases of circular causation a priori.

OBJECTION 2: A PLURALIZED CAUSAL PRINCIPLE We can avoid the appeal to Mereological Universalism by simply stating Universal Causation using plural quantification (Boolos 1984):

**Pluralized Universal Causation.** If the *x*'s are some things, then the *x*'s have (collectively) some causes.

We can then apply Pluralized Universal Causation to the plurality of all things that exist. These things would then have to have (collectively) some causes. These causes would have to exist themselves, and so they would be among their own effects. Consequently, at least one thing would have to be (at least in part) a cause of itself, in contradiction to the Impossibility of Self-Causation.

We could make this point by stating Pluralized Universal Causation principle in a stronger form:

**Strong Pluralized Universal Causation:** If the *x*'s are some things, then there are some other things, the *y*'s, such that the *y*'s collectively cause the *x*'s, and each of the *y*'s is distinct from each of the *x*'s.

OBJECTION 3: MEYER'S CAUSAL CHAIN PRINCIPLE Robert Meyer (1987) provides a similar objection to Universal Causation. A *causal chain* is a chain of entities, each of which is a cause of its immediate successor in the chain. Meyer assumes that if there is a causal chain, then there is a single thing that is a cause of every member of that chain. If we also assume that there is a set of causes, and we employ the usual axioms of set theory (including the Axiom of Choice), then Universal Causation entails that something causes itself. This violates the Impossibility of Self-Causation.

OBJECTION 4: AGAINST INFINITE REGRESSES Even if Mereological Universalism were false, Universal Causation would entail the existence of infinite causal regresses. If everything were caused, then each cause would itself be caused by something else, assuming again that nothing can cause itself. If causation is transitive, then an endless chain of causation would have to be infinitely long. If the chain were to form a closed loop, then each thing in the chain would cause itself, via the other links in the loop. That situation would also violate the Impossibility of Self-Causation.

At any rate, we are left to ask whether infinite causal regresses are possible. Many metaphysicians have found the idea absurd. If causal regresses were possible, then very specific structures and bodies of information could be self-causing and self-explaining. For example, one could imagine a world in which Tolstoy's novel *War and Peace* exists, even though no one ever created it. The story of the novel could simply be passed from one reader to another in an infinite regress, without an original author.

**Impossibility of Causal Regresses.** It is impossible for there to exist an infinite causal regress, that is, a series $x_1$, $x_2$, and so on of such a kind that $x_1$ is caused by $x_2$, $x_2$ by $x_3$, $x_3$ by $x_4$, and so on ad infinitum.

In addition, there are a number of arguments against the possibility of infinite causal regresses. The Grim Reaper paradox discussed in Section 19.1.2 provides the basis for one such argument. The argument requires a patchwork principle (**PMeta 5.1/5.2**). The basic idea is that if a certain kind of causal structure *C*, like an infinite regress, is possible, and a certain situation *S* consisting in the possession of certain causal powers and dispositions is possible, then there is a possible world in which the possible structure *C* is realized by placing a situation like *S* at every node in the structure.

This seems plausible, so long as the powers and dispositions that define the relations among the *S*-like situations are truly intrinsic to those situations. That is, this argument requires the Intrinsicality of Powers (**PMeta 2**). As we saw in Chapter 5, Neo-Humeists will disagree with this assumption, taking causal powers to be extrinsic to the things that possess them. According to Neo-Humeism, whether a thing has a certain causal power depends on the whole history of the world. Thus, Neo-Humeists have no reason to reject the possibility of infinite causal regresses, while those who see causal powers and dispositions as intrinsic to the things that have them do have good reason to think such regresses impossible.

In addition, the critic of the Grim Reaper argument can object to the patchwork principles. One could appeal to Sydney Shoemaker's Branch Principle (see Section 6.2.1) as a basis for this objection. If every possible world must branch off from the actual world at some point in time, then we have good independent grounds for rejecting the possibility of the infinite regress of Grim Reapers, so long as the actual world contains no such causally perverse regress.

### 26.2.2 Uncaused causes

If both self-causation and causal regresses are impossible and there is causation, then Special Causation must be true. In addition, the impossibility of both self-causation and infinite regresses entails that there must exist uncaused or "first" causes:

**26.1T.1A.1T Existence of Uncaused Causes.** Some uncaused things cause other things.

If there are such uncaused causes, what are they like? The only way for us to tell would be for us to appeal to some causal principle, a principle that states that everything of type *K* has a cause. We could then infer that all uncaused things are not of type *K*.

Here are some causal principles that have been proposed in the history of metaphysics:

**Causal Principle: Contingency-Based.** Every contingent thing is caused, and no necessary thing is caused. Proposed by al-Farabi (Craig 1980: 76–85), Avicenna or ibn Sina (Craig 1980: 86–97), Thomas Aquinas (Craig 1980: 158–204; Kretzmann, 1997), Samuel Clarke (Clarke 1704/1998), and Richard Taylor (Taylor 1991: Chapter 11).

**Causal Principle: Scotistic.** Every causable thing is caused. Proposed by John Duns Scotus (Scotus 1307/1987: 34–81).

**Causal Principle: Kalaam.** Everything that begins to exist is caused, and no beginningless thing is caused. Proposed by the medieval Kalaam philosophers, including al Kindi and al Ghazzali (Craig 1980: 61–75, 98–104), and defended more recently by William Lane Craig (Craig 1979; see also Koons 2014a).

**Causal Principle: Quantitative.** Every finite thing (i.e., every thing with a finitely measurable attribute) is caused, and no infinite thing is caused. Proposed by Duns Scotus (Scotus 1307/1987: 34–81; see also Koons 1997).

The Contingency-Based Causal Principle entails that all uncaused things are necessary beings, that is, beings that exist necessarily. The Scotistic Principle entails that all actually uncaused things are uncausable, the Kalaam Principle entails that no uncaused things began to exist, and the Quantitative Principle entails that all uncaused things are infinite, in the sense of lacking any attribute with a finite measure. The Quantitative Principle relies on the intuition that whenever something has a finite measure, there must be some causal explanation of why it has exactly that measure, rather than one slightly greater or slightly less. Hence, an uncaused thing must lack any such measurable attribute. All of its attributes must be absolutely unmeasurable or must have attributes that are metrically isolated, like the values 0 or absolute infinity.

Historically, these principles have been employed in proving the existence of God, with God understood as necessary, uncausable, beginningless, and infinite, and the cause of all other things. To pursue this question would require more space than we can spare here, but we will note that the First Cause argument as we've developed it does not by itself establish the existence of a unique first cause.

Must there be some such true causal principle? Could the set of caused things be undefinable? The main argument for the definability of a causal principle involves the claim that unprincipled causation would be incompatible with our having empirical knowledge. Consider any body of empirical data, including sensory experience, memories, physical traces, and testimony. If it is metaphysically possible for any of the data to exist uncaused, then it would seem prima facie possible for all of it to exist uncaused. However, the possibility that our empirical data is uncaused would be what epistemologists describe as an *undercutting defeater* to any claim to know something on the basis of that data. Uncaused perception, memories and so on cannot produce knowledge. If we cannot rule out the possibility that our data is uncaused, the live possibility that we do not in fact have knowledge might be enough to deprive us of knowledge in actuality. This possibility, if it cannot be rationally dismissed, would *defeat* any of the usual evidence for our claims to knowledge, since we would have to take seriously the possibility that all that evidence began to exist without any cause whatsoever.

Since the whole body of our empirical information could be vulnerable to this kind of defeater, we must be able to rule out the possibility that the information is uncaused on a priori, pre-empirical grounds. This would seem to require our having a priori knowledge of some appropriate causal principle, that is, knowledge that is prior to and independent of our empirical, scientific knowledge.

Any of the four principles listed above, if they could be known a priori, would preserve the possibility of empirical knowledge, since it is plausible to think that we can know a priori that all of our empirical knowledge is contingent, causable, finite, and had a beginning in time (see Koons 2008 for additional details).

OBJECTIONS TO THE FIRST CAUSE ARGUMENT There are two sorts of objections that could be lodged to these traditional First Cause arguments. First of all, one could object to Special Causation, which would involve arguing either that self-causation is possible or that causal infinite regresses are possible (or both).

A second sort of objection would involve accepting Special Causation but rejecting all of the Causal Principles. Such a position concedes that there are first or uncaused causes but denies that we can know anything about them. This is the thesis of Unprincipled Special Causation:

**26.1T.1A.2A Unprincipled Special Causation.** Some things are caused and others are not, and the two categories of things are in every other respect indistinguishable.

Defenders of Unprincipled Special Causation will still have to cope somehow with the problem of the defeat of our empirical knowledge by the fact that any of our evidence could have sprung into existence without cause. Without a principle of causation, there is no way to keep the possibility of uncaused events from penetrating deeply into the natural world—into our brains and our immediate environment.

The most common way of arguing for Unprincipled Special Causation is to argue that the sorts of things required by the various causal principles (Contingency-Based, Scotistic, etc.) are impossible. Here are the corresponding four types of objections:

Type 1: it is impossible for anything to exist necessarily. (Against the Contingency-Based Causal Principle)
Type 2: it is impossible for anything to be uncausable. (Against the Scotistic Principle)
Type 3: necessarily, everything began to exist at some point in time. (Against the Kalaam Principle)
Type 4: it is impossible for anything to be infinite in measure. (Against the Quantitative Causal Principle)

We will focus here on objections of Type 1, which are historically the most significant.

**Hume's Objection:** Imagination is the guide to possibility, and we can imagine the non-existence of anything.

In *Dialogues concerning Natural Religion* (Hume 1779), David Hume argued that the postulation of a necessarily existing being is absurd, since anything we can conceive of, we can conceive of as not existing. Hence, either we cannot conceive of a necessary being at all, or, if we can conceive of it, we can conceive of it as not existing. Anything we can conceive of as not existing, we must conceive of as possibly not existing, and so as not necessary. Thus, we cannot conceive of a necessary being at all.

We could take Hume to be appealing to Imagination as Guide to Possibility (PEpist 1). The argument might go something like this:

1 To exist necessarily is to be something whose non-existence is inconceivable.
2 If we can imagine the existence of something, then we can imagine its non-existence.
3 Hence, we cannot imagine the existence of something that exists necessarily. (From 1 and 2)
4 If we cannot imagine the existence of something, then we have strong grounds for denying its possibility. (Imagination as Guide to Possibility, **PEpist 1**)
5 Hence, we have strong grounds for denying the possibility of something that exists necessarily. (From 3 and 4)

However, as we have seen, imagination seems to be at best a fallible guide to possibility. So, even if it were true that anything we have a concept of is something we can *imagine* not existing, it doesn't follow that everything we have a concept of is *possibly* not existent (i.e., not something existing necessarily).

In addition, it is not at all clear that we can imagine something's not existing (contrary to step 2 of the argument), except under very special circumstances. If one knows that the Mona Lisa couldn't have existed without da Vinci's taking up painting, and one can imagine da Vinci's not being a painter (by imagining his spending his life sculpting instead), then one can (indirectly) imagine the Mona Lisa's not existing. Similarly, if one knows that electrons could not have existed unless the Big Bang reached a certain point of symmetry breaking, and one can imagine the Big Bang's not reaching that point, then one can imagine the non-existence of electrons. However, the concept of a necessary being doesn't seem to involve that thing's being dependent in any way on prior conditions, so it is far from obvious what it would be like to imagine such a being's non-existence.

**Kant's Objection:** The existence of a necessary being entails the validity of Anselm's ontological argument, but that argument is invalid.

In the "Transcendental Dialectic" section of *The Critique of Pure Reason* (Kant 1781/1787), Immanuel Kant argued that it is impossible to conceive of the existence of a necessary being. Kant argued that a necessary being would have to be a being whose existence was a logical truth, in the sense that it is a logical consequence of the basic axioms of logic and a set of coherent definitions. A necessary being would have to be the kind of thing that, in effect, exists by definition alone. For example, Anselm of Canterbury proposed that our concept of God is the concept of an absolutely perfect being, and perfection includes existence by definition. Hence, to suppose that God (so defined) doesn't exist would be self-contradictory, just like supposing that there is a triangle that doesn't contain three angles. This is one version of Anselm's *ontological* argument from his *Proslogion* (Anselm 1059/1998: 82–90—the name 'ontological' was first given to this argument by Kant in *The Critique of Pure Reason).*

However, Kant points out that nothing can exist by definition. You can define any sort of thing you want in any way, but you cannot possibly guarantee that there really is something that fits your definition. Even if the definition of 'God' included existence, there would still be an open question as to whether there exists anything in reality that fits the definition.

Kant's argument could be summarized thus:

1 A necessary being is one that exists as a matter of logical truth, that is, as a consequence of a set of definitions.
2 There is nothing that exists as a matter of logical truth, since no set of definitions by itself entails the existence of anything.
3 So, there is and can be no necessary being.

The defender of necessary beings must challenge premise 1. Logical truth is just one kind of necessary truth. A necessary being is simply a being that couldn't not have existed. There is no requirement that the non-existence of the being somehow contradicts a definition or the laws of logic.

Additionally, Kant's position is an unstable one. Supposed we ask whether premise 1 of Kant's argument is necessary or contingent. If Kant says it is necessary, then he must claim that it is a logical truth. This would mean that premise 1 follows from the definitions and the laws of logic. But, how can it? We haven't *defined* necessary truth as logical truth. If, instead, Kant supposes that premise 1 is a contingent truth, then we can ask on what basis does he accept it? If it could have been the case that some being existed necessarily but not as a matter of logical truth, then why couldn't that be true in the actual world?

## Note

1 In the middle of the twentieth century, an attempt was made by philosophers like Richard Jeffrey (Jeffrey 1965) to give a completely non-causal theory of decision-making. This non-causal decision theory relied entirely on subjective conditional probabilities. On this view, it is rational to choose to do action *A* just in case the conditional probability of good results, given *A*, is higher than the probability of those same goods conditional on any other action.

However, this non-causal decision theory was found to break down when considering certain possible choices. One example is the so-called *Newcomb problem* (Nozick 1969). Suppose that you are given the choice between taking either both or only the first of two boxes containing money. Your choice will not affect the amount of money in either box. It is surely obvious in such a case that the rational choice is to take both boxes. However, it is possible that the probability that you will become rich, conditional on taking only the first box, is higher than the probability that you will become rich, conditional on taking both. Suppose that you know that the second box contains only one dollar. Now suppose that a team of brilliant psychologists have been studying your behavior for years, and they have put \$1 million in the first box if and only if they have predicted that you will take only one box. Let's suppose that their prediction (in either case) has a 90% chance of being accurate. In such a situation, the probability that you will gain $1 million conditional on taking only the first box is much higher than the probability that you will gain $1 million conditional on taking both boxes. Your conditional expected utility for taking only the first box is about $900,000, while the conditional expected utility for taking both boxes is at most $101,000. Nonetheless, it seems crazy to leave the second box behind, since you don't cause any money to appear in the first box by not taking the second.

# 27

# Causation: A Relation between Things or Truths?

We can now turn to questions about the nature of causation. We have talked about causes and effects as 'things', but what sort of things? Is *causation* a relation between things, like *being next to* or *being taller than*, or is it something else entirely? Let's look at a simple example:

(1) Mary's kicking the ball caused the ball to enter the goal.

Should we think of causing as a kind of relation between two entities? If so, what entities? Apparently, Mary's kicking of the ball is the cause, and the ball's entering of the goal is the effect. What sort of entities are these? Philosophers in recent years have used a number of categories here, among them events, situations, facts, conditions, eventualities. We will in general use the word 'event' as the most general expression for the sort of thing that can stand in the *causal* relation.

There is another way to think about causal statements. We could rewrite (1) as (2):

(2) The ball entered the goal because Mary kicked it.

(2) seems to express pretty much the same thing as (1). In (2), there is no verb or preposition corresponding to causation. Instead, we have the word 'because', which joins two complete sentences into one compound sentence. (2) might suggest that causation is really a logical relation of some kind between two statements, something like the relations of conjunction (corresponding to the word 'and') or conditionality (corresponding to the words 'if' and 'then'). Compare (2) with (3) and (4):

(3) The ball entered the goal if Mary kicked it.
(4) The ball entered the goal, and Mary kicked it.

*The Atlas of Reality: A Comprehensive Guide to Metaphysics*, First Edition.
Robert C. Koons and Timothy H. Pickavance.

We're not suggesting that (3) or (4) come close to saying the same thing as (2). We're simply asking you to pay attention to the logical form of (2), (3), and (4). None of us are tempted to suppose that the 'and' in (4) refers to some kind of relation between the ball's entering the goal and Mary's kicking it. Instead, we naturally believe that there is some relation between the truth or proposition expressed by (4) and the truths or propositions expressed by (5) and (6):

(5) Mary kicked the ball.
(6) The ball entered the goal.

Perhaps (2) has a similar relation to (5) and (6). Here is a plausible idea: (2) is true when it's the case that the truth of (6) can be correctly *explained* in a certain way in terms of the truth of (5).

Thus, we seem to have two ways of thinking about causation. We might think of it as a real relation, the relation of *causal connection*, between things or events, or we might think of it as a logical relation, the relation of *causal explanation*, among truths. For metaphysicians, the crucial question is whether *causal connection* or *causal explanation* is more fundamental. We thus have the difference between Causal Connectionism and Causal Explanationism:

**27.1T Causal Connectionism.** Causation is fundamentally a relation (*causal connection*) between things (and not between truths).

**27.1A Causal Explanationism.** Causation is fundamentally a relation (*causal explanation*) between truths.

To clarify, Causal Explanationists need not deny the existence of events, even as fundamental entities. For example, Causal Explanationists might take the fundamental relation to be an explanation relation between propositions of the form 'event *c* occurred' and 'event *e* occurred'. In contrast, Connectionists must suppose that there is a fundamental relation between the events themselves. When that relation holds between events *c* and *e*, then it will follow that the fact that *c* occurred provides a veridical causal explanation of the fact that *e* occurred.

## 27.1 Causal Explanationism

Causal Explanationists take causal explanation to be metaphysically fundamental. For Explanationists, statements that seem to attribute causation as a relation between things, like events, are merely a different way of expressing the existence of a causal explanation between two truths. Consider again (1) and (2):

(1) Mary's kicking the ball caused the ball to enter the goal.
(2) The ball entered the goal because Mary kicked the ball.

For Explanationists, verbal noun phrases like 'Mary's kicking the ball' or 'the ball's entering the goal' do not really refer to things. They are merely ways of expressing the truths that are most properly expressed by the complete sentences 'Mary kicked the ball' and 'the ball entered the goal'. Hence, Explanationists have a prima facie advantage in respect of Ockham's Razor (**PMeth 1.4**): Explanationists' ontology is leaner, both quantitatively and qualitatively, than Connectionists'. Connectionists believe in many instances of a fundamentally new kind of entity, the sort of entity that can stand in a primitive causal relation to other entities.

There are three popular versions of Causal Explanationism, distinguished by their respective ways of explicating the idea that one truth can causally explain another:

**27.1A.1 Nomological-Deductive Theory of Causal Explanation.** The fact that truth *p* causally explains truth *q* consists in the fact that *q* can be deduced from *p*, together with truths about the laws of nature and relevant background conditions.

**27.1A.2 Counterfactual Conditional Theory of Causal Explanation.** If *p* and *q* are truths of the appropriate kind (truths about the occurrence of changes or becomings of some kind), and the truth that *p* causally explains the truth that *q*, then that causal-explanatory fact consists in the fact that *q* would be not have been true if *p* had not been true (i.e., that (Not-*p* []-> Not-*q*), using the Stalnaker-Lewis subjunctive conditional, '[]->').

**27.1A.3 Probability-Raising Theory of Causal Explanation.** The fact that truth *p* causally explains truth *q* consists in the fact that there is an appropriate body of background truths *K(p)* such that the probability of *q* conditional on the conjunction of *p* and *K(p)* is greater than the conditional probability of *q* on *K(p)* alone.

If the fundamental laws of causation were probabilistic (at least in some cases) rather than deterministic, then the Nomological-Deductive Theory would be inadequate, since probabilistic laws would *never* entail the truth of any specific truth about the consequences of any initial set-up. They would only entail that certain consequences are more likely than others. Many philosophers and physicists (but not all)[1] believe that the fundamental laws of quantum mechanics are probabilistic: the laws make it likely that a certain kind of atom will decay within a certain period, but they never entail that it certainly will decay. Other philosophers, including Anscombe (1975) and Cartwright (1983) think that even if the fundamental laws of physics are not probabilistic, it is still never the case that we can use them to deduce with certainty what would happen in any actual or hypothetical case, because the application of the laws to concrete cases involves the introduction of exceptions, hedges, and qualifications.

Phil Dowe (2000) argued that a certain "decay case" is prima facie problematic for the Probability-Raising and Counterfactual Theories. The decay case involves five events (see Figure 27.1). Event *A* represents the existence of an atom of type 1, which must decay either into an atom of type 3 (event *C*) or an atom of type 2 (event *B*), with probabilities $^1/_4$ and $^3/_4$, respectively. An atom of type 2 decays immediately with probability one into an atom of type 5 (event *E*), while the atom of type 3 may decay either into an atom of type 5 (with probability $^3/_4$) or into one of type 4 (with probability $^1/_4$). The actual sequence of events is *A*-*C*-*E*. Intuitively, we want to say that *C* is a cause of *E*, since *C* consists of the production of an atom of type 3 whose actual decay into an atom of type 5 is (in

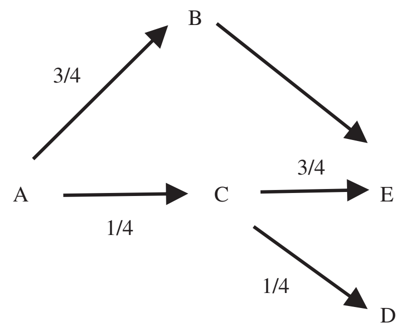


**Figure 27.1** The Dowe Decay Case

actuality) *E*. However, the occurrence of *C* actually lowers the probability of *E* from 1 to $^3/_4$. The Counterfactual Theory also seems to fail in this case, since it does not seem to be true that if *C* had not occurred, *E* would not have occurred. If *C* hadn't occurred, a type 2 atom would have been created instead, which would have decayed into an type 5 atom, and *E* is simply the production of an atom of type 5.

There is room, however, for Causal Explanationists to revise the two theories in such a way that they deliver the right answer in this case. In the case of the Counterfactual Theory, we could say (as David Lewis (1973a) recommends) that if *C* had *not* occurred, then *neither* would have *B*, and so *E* would not have resulted. The chain of events would simply have ended with *A*. Lewis recommends that when we evaluate a counterfactual conditional, we introduce a local "miracle" into the course of events, the smallest miracle needed to make the antecedent true. We should ignore what the laws of nature might say about what would have happened in place of the antecedent. The laws in Dowe's decay case are supposed to entail that either *C* or *B* succeeds event *A*. However, when evaluating the conditional 'If *C* had not occurred, *E* would not have occurred', we should ignore this law and consider what would have happened had neither *C* nor *B* occurred, a world in which the atom of type 1 decays and leaves (miraculously) no decay product at all.

Similarly, when applying the Probability-Raising Theory, we could ask what was the conditional probability of *E* given the absence of *C*, while holding the absence of *B* fixed. We should include the non-occurrence of *B* in the set of background truths relative to which we evaluate the posterior probability of *E* given the non-occurrence of *C*. Presumably, the probability of *E*, given the non-occurrence of both *B* and *C*, is zero or close to zero. Hence, the occurrence of *C* does raise the probability of *E*, after all.

Both the Counterfactual Theory and the Probability-Raising theory thus must provide some principled way of identifying those facts that should be held fixed when evaluating a counterfactual or the relevant conditional probabilities. One popular answer (Eells 1991, Kvart 2004) is to hold fixed all the propositions that are made true by something *earlier in time* than whatever makes true the proposed cause. However, such an answer ensures that it is metaphysically impossible for causes to be later in time than their effects, and it forces us to assume that causal directionality derives from temporal directionality, and not vice versa.

### 27.1.1 Objections to Causal Explanationism

There are two major objections to all three versions of Causal Explanationism: the problem of causal linkage and the problem of the direction or asymmetry of causation.

*27.1.1.1 Objection 1: The Problem of Causal Linkage.* The problem of causal linkage charges that Causal Explanationism provides an inadequate account of the linkage between causes and their effects. This lack of connection is most evident in cases of *redundant causation*. Redundant causation comes in two forms: symmetric *overdetermination* and *preemption*.

1a. The case of symmetrical causal overdetermination. Causation is redundant when several potential causes exist, each of which could be sufficient by itself and independently of the other causes to produce some specific effect. This kind of redundancy can occur in both deterministic and probabilistic settings.

Since we have already been discussing *deterministic* laws and causation, a definition of determinism is overdue.

**Def D27.1 Actually Deterministic Laws.** A set $S$ of causal laws is *actually deterministic* if and only if every member of $S$ is true, and every true proposition describing the occurrence of a localized, instantaneous event $E$ is logically entailed by the set of true propositions which includes $S$ as a subset and which includes each true proposition specifying a condition that is causally prior to $E$.

**Def D27.2 Essentially Deterministic Laws.** A set $S$ of causal laws is *essentially deterministic* if and only if $S$ is actually deterministic in every world $w$ at which all of the members of $S$ are true.

**Def D27.3 Deterministic Worlds.** A world $w$ is *deterministic* if and only if the set $S$ of causal laws true in $w$ is essentially deterministic.

Using these definitions, we can distinguish deterministic redundancy from probabilistic redundancy:

**Def D27.4 Deterministic Redundancy.** A truth $p$ has *deterministically redundant causes* if and only if there are two disjoint sets of truths $C_1$ and $C_2$; the truths in $C_1$ and $C_2$ are causally prior to $p$; no truth in $C_1$ is causally prior to any truth in $C_2$ or vice versa; there is a set of actual causal laws $S$ such that both the union of $S$ and $C_1$ entails $p$, and the union of $S$ and $C_2$ does so also.

A case of causal redundancy is *symmetric* if and only if there is no relevant difference between the two independent causes of the effect.

**Def D27.5 Probabilistic Symmetry.** A truth $p$ has *probabilistically symmetrical causes* if and only if (i) there are two disjoint sets of truths $C_1$ and $C_2$, (ii) the truths in $C_1$ and $C_2$ are causally prior to $E$, (iii) no truth in $C_1$ is causally prior to any event in $C_2$ or vice versa, (iv) there is a set of actual causal laws $S$ such that both the union of $S$ and $C_1$ raises the objective probability of the occurrence of $p$ (relative to background conditions), (v) the union of $S$ and $C_2$ raises the probability of $p$ to exactly the same degree, and (vi) the probability of $C_1$ and $C_2$ are mutually independent.

A classic example of symmetric deterministic redundancy is a symmetric firing squad, consisting of two infallible marksmen. Each firing is sufficient, by itself and independently of the other, to cause the victim's death. We are inclined to say either that both of the true propositions about the shootings are causal explanations of the death or that the conjunction of the two truths is a single, partially redundant explanation.

There are also cases of probabilistic symmetry. For example, take the case of two neutrons striking the same nucleus at the same time, each with a chancy propensity to initiate the nucleus's decay. If an event of decay actually occurs, the right answer seems to be that one or the other or both may be the true causal explanation. Suppose that each potential cause had a 50% chance of causing the decay. Intuitively, there is a 75% chance that the nucleus will decay when struck by both causes simultaneously: a 25% chance that it will decay as a result of the action of the first neutron alone, a 25% chance that it will decay as a result of the action of the second neutron alone, and a 25% chance that it will decay as a result of the simultaneous action of both neutrons (multiple, simultaneous causation of the effect). However, Causal Explanationists must say that there is a 75% chance that it will decay as a result of the action of *both* causes and a 0% chance that it will decay as a result of one but not the other. This is a counter-intuitive result.

1b. Preemption, especially *late* preemption and *trumping*. Preemption involves asymmetric redundancy. In preemption cases, there are two potential causes, but it is clear intuitively that one of the potential causes has preempted the other, making it the actual cause and the other a merely potential cause. To return to the firing squad case, we can turn a case of symmetric redundancy into preemption simply by having one marksman shoot a split second before the other, in such a way that the first marksman's bullet has already killed the victim before the second marksman's bullet can arrive. A simple example involves two bowling balls rolling toward the pins, with the first ball knocking down the pins before the arrival of the second. In such cases, it seems clear that one potential cause has preempted the other. Even though the preempted potential cause still entails (in the case of deterministic causation) or raises the probability of the occurrence of the effect, it in fact plays no causal role.

Preemption cases pose two dangers for any account. First, there is a danger of attributing causality to the merely potential, preempted cause. Second, there is a danger of denying causality to the actual, preempting cause. Any Explanationist theory is liable to one or both of these dangers since they provide no way of tying some particular effect to one rather than the other potential cause. Causal Connectionists, in contrast, can appeal to a richer account of the cases, one in which some kind of fundamental causal tie or connection breaks the logical or probabilistic symmetry.

Cases of preemption provide ready counterexamples to the Nomological-Deductive Theory. In the case of the staggered firing squad, the shooting of the delayed marksman satisfies the Nomological-Deductive Theory's conditions: the truth of the proposition that the second marksman fired is, together with the laws of nature and other background conditions, sufficient to entail the proposition that the victim is subsequently dead, despite the fact that this shooting has clearly been preempted from causing the death.

The same case poses a problem for the Counterfactual Conditional Theory as well. Here the problem is not that the second marksman's shooting counts as a cause, but that the first marksman's shooting does not. It is not the case that if the first marksman did not shoot, the victim would not have died. The victim would have died whether or not the first marksman had shot.

It is also relatively easy to create problem cases for the Probability-Raising Theory (like Eells's 1991 account). To do so, we must find a preempting cause that *lowers* the probability of the effect given the presence of the preempted cause. Douglas Ehring (1997) uses the example of a device that can deliver one of two drugs. It is very likely that *B* will be delivered, and the chances of survival conditional on *B*'s delivery are very high. It is possible, but unlikely, that *A* will be delivered instead, preempting the delivery of *B*. The chances of survival conditional on *A*'s delivery are good, but significantly worse than on *B*'s delivery. In the actual world, *A* is delivered and the patient survives. As we track the probability of survival, we find that it goes *down* when drug *A* is delivered, despite the fact that *A* is clearly the cause of survival.

DAVID LEWIS'S SOLUTION: CAUSATION AS A CHAIN OF CAUSAL DEPENDENCIES David Lewis (1973a) defines causation as the transitive closure, the *ancestral*, of the causal dependency relation. In the context of defending Causal Explanationism, Lewis's idea could be adapted in this way. First, we define *direct causal explanation* in terms of the Nomological-Deductive, Counterfactual Conditional or Probability-Raising Theory. Next, we define *causal explanation* as the ancestor of direct causal explanation. In other words:

**Def D27.6 Causal Explanation as the Ancestral of Direct Causal Explanation.** Truth *p causally explains* truth *q* if and only if either: (i) *p* directly explains *q*, or (ii) there are truths $r_1$, $r_2$, … $r_{(n)}$ such that: *p* explains $r_1$, each $r_{(i)}$ explains $r_{(i+1)}$, and $r_n$ explains *q*.

This Ludovician account has the advantage of providing the right answer in many cases of preemption, namely, in all those cases involving *early* preemption, in which the merely potential cause is preempted at a stage earlier than the effect. For example, suppose there are two potential causes, *c* and *c′*, of some event *e*. If not interrupted, *c′* would cause intermediate event *i′*, which would in turn cause *e*. Similarly, there is a potential *c-i-e* chain. Suppose both *c* and *c′* occur, but *c* prevents the occurrence of *i′*. In this case, *e* is causally dependent on *i* (since *i′* never occurred), and *i* is causally dependent on *c*, so *c* counts as a cause of *e*, even though if *c* hadn't occurred, *c′* would have caused *e*.

A similar story can be told in terms of probabilistic causation. Let's suppose that the definition of '*A* causes *B*' requires a chain of intermediate events $I_1, I_2, \ldots, I_n$, such that *A* raises the probability of $I_1$, each $I_{(j)}$ raises the probability of $I_{(j+1)}$, and $I_n$ raises the probability of *B*.

Obviously, this solution depends on the interposition of at least one pair of intermediate events (i.e., events *i* or *i′*). However, we can imagine cases of *late* preemption, cases where it is the effect *e* itself that preempts the alternate causal chain, without the failure of the realization of an intermediate event (like *i′* above). We could also imagine cases involving discrete time, in which the effect e occurs in the very next moment following the occurrence of the preempting cause, with no possibility of the occurrence of any intermediate event.

EARLY VS. LATE PREEMPTION Let's look at a pair of examples that illustrate the difference between *early* and *late* preemption:

**Early preemption:** Ball *A* knocks down the pins before ball *B* begins to be rolled. Because *A* has caused the effect, the other causal process is cut short (i.e., the second bowler never completes his bowling of ball *B*).

**Late preemption:** In this case, there are no intervening states. Ball *A* preempts the causal efficacy of ball *B* simply by knocking down the pins first, not by changing any of the preconditions of *B*'s action (Hall 2004).

When dealing with probabilistic causation, we should consider cases in which the preempting cause is simply the negation of the preempted cause. Suppose, in a variation on Dowe's decay case, atom 1 decays into atom 2 (event *B*), which in turn decays into atom 3 (event *C*). If the atom 1 had not decayed into atom 2 the probability that it would have decayed directly into atom 3 was 75%. Let's suppose that the probability that atom 2 would decay into atom 3 was only 50%. Thus, event *B* actually lowers the probability of event *C* in the circumstances, and yet *B* could be a cause of *C*. Let's suppose that the event *B* is intimately or intrinsically involved in the occurrence of *C*. The occurrence of *B* is not simply some partial disabling of the original *C*-producing mechanism but by itself constitutes a *C*-producing mechanism (let's say, of a significantly different kind from the direct *A*-to-*C* mechanism). *C* comes about in a different way, perhaps in an entirely different way, depending on whether it is due to the *A*-involving or *B*-involving mechanism. *B* had its own propensity to produce *C*, which was in fact exercised in the actual world. It seems clear that in such a case *B* is indeed a genuine cause of *C*, despite lowering *C*'s chances of occurring.

TRUMPING OR SIMULTANEOUS PREEMPTION Another kind of example that the Ludovician causal chain response cannot handle is preemption by trumping (first proposed by Jonathan Schaffer 2000). Suppose event *E* is produced simultaneously by two processes of two distinct kinds *K* and *K′*. In "Causation as Influence" (Lewis 2000), David Lewis uses orders by two officers of different and comparable ranks, such as simultaneous and conflicting orders from a captain and a sergeant to a private, to illustrate this type of situation. Whenever orders of these two kinds conflict, the order of kind *K* (from the superior officer) always wins. In the actual case, the two orders have exactly the same propensity: to produce *E*. Let's suppose, for example, that both the captain and the sergeant simultaneously order the private to fire. In such cases, the captain's order is the only cause of the private's actions, despite the fact that the two processes culminate at the same instant, with no preempted intermediaries. The sergeant's order is preempted by being trumped by the captain.[2]

1c. Lewis's Hybrid Counterfactual Theory. Before we move on, there is one historical complication to mention. David Lewis's own Counterfactual Theory of Causation (in Lewis 1973a) may not in fact have been a version of Causal Explanationism. Lewis posited the existence of *events* as the relata of causation. The causal connection between two events, *C* and *E*, is then analyzed by means of a counterfactual conditional:

**Simple Ludovician Counterfactual Theory:** *C* immediately causes *E* if and only if had *C* not occurred, *E* would not have occurred.

Since Lewis recognizes causation as involving a connection between things, namely, events, Ludovicians could respond to the problem of causal linkage by simply adding some further link between the two events, a link of *causal priority*:

**Hybrid Ludovician Theory:** *C* (immediately) causes *E* if and only if *C* is causally prior to *E* and, if *C* had not occurred, *E* would not have occurred.

The causal priority relation could then be thought of as some asymmetric connection between events, not to be analyzed in terms of counterfactuals.

Ludovicians could respond to the problem of linkage in another way, by way of event *fragility*. The Ludovician theory with fragility analyzes the causal connection between two events in terms of an asymmetric necessitation between the two events, considered as tokens rather than types. The idea is that *this* very event (the effect) couldn't have existed except in a world in which it results from *that* every event (the cause). Hence, if the cause had not existed, it would have been impossible for the effect to exist. This kind of counterfactual dependency would work even in cases of overdetermination and of preemption, including late preemption and trumping.

However, this fragility defense is not available to Causal Explanationists, since it requires some kind of fundamental tie between events to serve as the ground of the asymmetric necessitation of causes by effects. If Ludovicians fail to posit such a fundamental tie, then they must posit brute necessary truths linking the propositions that affirm the existence of the disjoint events. For example, in the case of trumping preemption, Ludovicians would have to suppose that the effect is fragile, that the very same effect could not have occurred in the absence of the trumping order. However, when we ask why the effect is tied in this way to the trumping order and not to the trumped order, the only plausible answer is that the trumping order is the cause and the trumped order is not. Hence, fragility cannot provide Ludovicians with a non-circular solution to the problem of linkage.

*27.1.1.2 Objection 2: The Problem of Causal Direction or Asymmetry.* The problem of causal direction or asymmetry charges that Causal Explanationism (in all of its forms) provides an inadequate account of the asymmetric direction of causation, from prior to posterior.

Deductive and probabilistic relations are independent of causal direction. For example, there are many cases in which we can deduce a cause from its effects (e.g., deducing the length of a flagpole from the length of its shadow). Similarly, an effect can raise the probability of one of its causes.

Counterfactuals may be more help. David Lewis (Lewis 1979b) argued that the Counterfactual Theory does provide an account of causal direction, since it is generally not true that if an event hadn't occurred, its cause would not have occurred. According to Lewis, all forward-tracking conditionals (from absence of cause to absence of effect) are true, while all backward-tracking conditionals (from absence of effect to absence of cause) are false.

Lewis argued that there are many cases in which conditionals of the form, 'If *A* had happened, *B* would have happened', are true when *A* occurs before *B*, but very few such conditionals are true when *B* happens before *A*. Thus, the truth-values of these conditionals can provide a basis for distinguishing time's direction, and thus for determining the direction of causal influence.

How does Lewis get this result, given that the fundamental laws of nature are all time-reversible, and given that his semantics for the conditional includes no explicit temporal bias? The result is supposed to follow from the way the arrangement of occurrent facts in our world interacts with Lewis's priorities for deciding *closeness* of possible worlds to our own. Lewis suggests four criteria for determining the closest possible worlds, for the purposes of evaluating counterfactual conditionals:

> (1) It is of the first importance to avoid big, widespread, diverse violations of causal law.
> (2) It is of the second importance to maximize the spatiotemporal region throughout which perfect match of particular fact prevails.
> (3) It is of the third importance to avoid even small, localized violations of law.
> (4) It is of little or no importance to secure approximate similarity of particular fact.

Recall the semantics for the counterfactual conditional discussed in Section 4.1. A conditional of the form '*p* []-> *q*' is true in the actual world *w* just in case *q* is true in all of the worlds in which *p* is true that are *closest* to the actual world. (If there are no maximally close worlds, then it is enough if *q* is true in all of the *p*-verifying worlds that are closer to the actual world than some reference point.) In a conditional of the form '*p* []-> *q*', *p* is the antecedent and *q* is the consequent. Thus, to evaluate such a conditional, we check if the consequent is true in all of the closest antecedent-verifying worlds.

Forward-tracking counterfactual conditionals with actually false consequents can often be true, since the closest world to the actual one in which the antecedent is true will be a world that matches the actual world throughout the past and right before the occurrence of the antecedent, the antecedent is realized as a result of a minor miracle, and the world thereafter evolves according to this world's natural laws. However, backward-tracking counterfactual conditionals with actually false consequents will rarely, if ever, be true, since meeting criterion 2 will force us to preserve an exact match with the actual world throughout the past, forcing the actually false consequent to remain false in the closest possible worlds.

However, this strategy for securing temporal direction and thereby causal asymmetry is problematic. Lewis has built into his application of criteria 1 and 2 a bias toward matching the past. We could instead consider a world that matches the actual world throughout the future, verify the antecedent by producing a small miracle immediately after its occurrence, and then evolve the history of the world backward, in accordance with the laws of nature. The past of such a world will be radically different from the actual past, just as the future of Lewis's preferred worlds are radically different from the actual future.

The crucial question, then, is which world has the smaller miracle: the world $w_1$, with a past like the actual past and a miracle immediately before the antecedent's occurrence, or world $w_2$, with a future like the actual future and a miracle immediately after the antecedent's occurrence? It is not at all easy to say. It certainly seems easier to figure out what kind of miracle is required in $w_1$ than in $w_2$, and that may give us some (fairly weak) reason for thinking that the $w_1$-miracles are smaller than $w_2$-miracles. For example, if our conditional is, 'Had Nixon pushed the button, there would have been thermonuclear war', we have a pretty good idea how to verify the antecedent: produce a miracle in Nixon's brain, causing certain relevant neurons to fire that don't fire in the actual world. It's much harder to figure out exactly what miracle would be required in the immediate future of Nixon's button-pressing that would be needed to verify the button-pushing when we evolve the resulting processes backward.

However, Alexander Pruss (2003) argues that if there is such an asymmetry, it is an anthropocentric asymmetry. Most counterfactuals are forward-tracking rather than backtracking simply because it is the forward-tracking ones that are of most interest to us. It may be that which propositions or facts our languages express impose this asymmetry on the counterfactuals of natural languages, but from a God's eye point of view, the entire set of possible propositions is innocent of any such bias. Hence, at the end of the day, the counterfactual theory of temporal direction is a subjective, anthropocentric one, very similar to accounts, like von Wright's (1971), that define temporal direction in terms of our contingent human abilities to do some things by doing other things.

In addition, Lewis's account doesn't work in cases of probabilistic causality. In the deterministic case, some violation of the laws of nature is required, so a late and small miracle is to be preferred over an early and large one. However, in the probabilistic case, we could verify the antecedent without any violation of law. Shouldn't the avoidance of a miracle be worth a small degree of backtracking? Douglas Ehring (1997) considers a case where we have the following links: *C*-(50%)-*B*-(100%)-*D*-(50%)-*E*. What would be the closest non-*D* world? Lewis must say a world where both *C* and *B* occur, and then a miracle prevents the occurrence of *D*. However, wouldn't a world in which *C* but not *B* occur be at least as close? Such a world avoids miracles altogether. This would mean that the backtracking conditional 'If *D* had not occurred, *B* would not have occurred' would come out as true, by Lewis's criteria (see also Elga 2001).

OTHER EXPLANATIONIST ACCOUNTS There are several other accounts of causal direction that Causal Explanationists might offer. There are four accounts that have been most popular: temporal asymmetry, the asymmetry of kaon decay, the asymmetry of entropy, and Reichenbach's account of *conjunctive forks*. Let's consider these one at a time.

2a. Temporal asymmetry. Here is a fairly simple account of the direction of causation (suggested by David Hume): causes always occur before their effects. This makes the direction of causation depend on a prior or more fundamental direction of time. There are several objections to this account:

**Objection 1.** The problem of the directionality of time is no easier to solve than the problem of the directionality of causation. Physics is (in its fundamental laws) perfectly time-reversible (except perhaps for kaon decay—see 2c below). In fact, if there is a direction to time, it seems plausible to think that the direction of time depends on the direction of causation, and not vice versa.

**Objection 2.** It seems plausible that in many cases the cause and effect are simultaneous. In fact, we shall see in the next section that it is plausible that all cases of *discrete* causation are instantaneous, with the action of the cause and the passion of the effect's being simultaneous.

**Objection 3.** It seems possible for there to be cases of temporally reversed causation, where the cause is later than the effect. This happens quite often in science fiction and fantasy stories involving time travel or the precognition of the future. Some philosophers have proposed that time-reversed causation actually occurs at the quantum level. A decision to observe a certain property of a particle here and now affects the past of the particle in such a way as to indirectly bring about a change far away at the same time as the observation.

2b. Entropy. Another very popular account of the direction of both time and causation involves an appeal to entropy. The Second Law of Thermodynamics states that the total entropy, or disorder, of an isolated system always increases. We see many examples of this. Whole eggs are succeeded by broken eggs, but never vice versa. Tornados tear houses apart; they never assemble houses.

However, modern physicists do not generally consider the Second Law of Thermodynamics to be a fundamental law of physics. Following the work of Ludwig Boltzmann (1844–1906), most scientists and philosophers consider the Second Law to be only statistical in nature, specifying what happens in most cases, not all. Since ordered situations (situations with low entropy) are so rare, they are always more likely to be followed by situations with higher entropy. However, in the whole history of the world, there are an equal number of cases of low-entropy states arising from high-entropy states as there are of low-entropy states decaying into the more common high-entropy states.

On this standard view, the increase in entropy we see around us is merely an artifact of the fact that our local universe began with the Big Bang in such an unusually low-entropy condition. If the world consists of many such universes, then there may be just as many that start in a high-entropy condition and evolve toward extremely low entropy as there are universes like ours that start out in the unusual low-entropy condition and evolve toward higher entropy. If that's not the case, we need some further explanation of why universes must start out, as ours did, in such an orderly condition (see Price 1996 for more details).

2c. Kaon decay. When we stated earlier that the fundamental laws of physics are time-reversible, we were over-simplifying somewhat. In fact, there is a difference between the actual and the time-reversed versions of many physical processes. If we know, for example, which particles are negatively charged and which are positively charged, we can tell which version of a process is actual and which is time-reversed. A time-reversed electron acts exactly as an ordinary positron (the positively-charged anti-particle of the electron) would. However, charge is itself perfectly symmetrical. A process in which both charge and time are simultaneously reversed would be indistinguishable from an actual process.

There is, though, a very rare and short-lived particle, the *kaon*, which decays in a way that violates the charge-time symmetry that is respected by all other physical processes. Phil Dowe (2000) has suggested that this charge-time asymmetry in kaon decay could be the fact that distinguishes the forward from the backward direction of time.

However, kaon decay, like all physical processes without exception, does satisfy a somewhat weaker symmetry principle: the charge-time-parity principle. If we simultaneously reverse time, charge, and left- vs. right-handedness, then we cannot distinguish kaon decay from kaon formation. Since parity (left- and right-handedness) are themselves intrinsically similar, the charge-time asymmetry of kaon decay does not pick out a unique direction to time. If kaon decay had occurred in the opposite direction, the only difference would have been that our universe would have been composed mostly of anti-matter instead of matter.

In any case, it seems implausible that the direction of causation depends on the occurrence of a single exotic and extremely rare kind of change. Causal asymmetry is both ubiquitous and local or intrinsic in character.

2d. Conjunctive forks. Hans Reichenbach (1958) suggested that the direction of time can be grounded in certain patterns of statistical correlation that exist in nature. Causal direction is to be defined in terms of the structure of *conjunctive forks*, making use of Reichenbach's principle of common cause.

**Principle of Common Cause.** If the joint probability of ($A$&$B$) is greater than the product of the probabilities of $A$ and $B$, and if $C$ is a proposition such that the joint probability of ($A$&$B$) given $C$ is equal to the product of the conditional probabilities $P(A/C)$ and $P(B/C)$, then $A$, $B$, and $C$ form a conjunctive fork (Reichenbach 1958).

**The Conjunctive Fork Theory of Causal Direction.** If propositions $A$, $B$, and $C$ form a conjunctive fork, then $C$ is causally prior to and causally connected to both $A$ and $B$.

The direction of time is determined by the predominant direction of causation. The temporal direction $D$ corresponds to the *earlier-than* relation if and only if most causes stand in the $D$ relation to their effects. The Conjunctive Fork Theory has been defended and further developed by Wesley Salmon (1984) and by Spirtes, Glymour, and Scheines (1993).

Frank Arntzenius (1993, 2010) has argued that the Conjunctive Fork Theory suffers from many scientific counterexamples, including quantum-mechanical correlations, correlations due to co-existence laws (like electromagnetism, gravitation or Pauli's exclusion principles), and correlations occurring in a deterministic universe. The theory only works when we can assume two things: an initial state of absolute, uncorrelated chaos, and spatial separation between the correlated properties. Thus, like the appeal to entropy, the appeal to Reichenbach's conjunctive forks depends on an apparently contingent and unexplained fact of our world's history, namely, its beginning in a condition in which certain properties are uncorrelated. The asymmetry of conjunctive forks is limited and derived, not universal and fundamental.

FURTHER PROBLEMS WITH EXPLANATIONIST ACCOUNTS OF CAUSAL DIRECTION There are a few additional Explanationist accounts of causal direction. For example, one might appeal to the fact that, at least in our local universe, the process of the formation of black holes seems to be irreversible. However, all Explanationist accounts must appeal to certain global features as the ground of causal direction, since none of the local features that ground causal explanation, whether logical deducibility, the truth of counterfactual conditionals, or the raising of probability, are intrinsically irreversible. There are two objections to all such global accounts of causal direction.

1. Inverted time thought experiments. Each of these accounts is vulnerable to inverted time thought experiments. Since the fundamental laws are (barring a couple minor exceptions) time-reversible, there is nothing inherently impossible about a universe in which entropy is decreasing or dispersed order converges (light waves or water ripples converging often to a single point) (see Huw Price 1996). It's true that the vast majority of states with low entropy move toward greater entropy, but, since the high-entropy states are more probable by an astronomically high factor, the high-entropy to low-entropy transitions will occur with the same probability as those in the opposite direction. Reichenbach is, in effect, guilty of the statistical fallacy of ignoring the base rate (ignoring the far greater probability of the high-entropy states).

2. Small world thought experiments and the intrinsicality of causal direction. In addition, all such accounts, and even accounts like Dowe's that appeal to exotic, irreversible events, have the consequence that the causal direction of a process is not intrinsic to that process. Whether, in a particular case, *E* caused *E′* or *E′* caused *E* will depend on what is going on in other, perhaps quite remote, parts of the universe. In fact, in some cases, the causal direction of the process will never be settled so long as the history of the universe continues, since events in the remote future might reverse the predominate direction of entropy increase or order dispersal or whatever else we fix on to define temoral order.

This problem can be illustrated by Tooley's small world thought experiments (Tooley 1987, 1990). It would seem to be possible for the history of the universe to be very simple, consisting only of a handful of causal interactions. In such a universe, there may be no change in entropy and no asymmetric dispersal of order whatsoever (and no kaon decay, either). Nonetheless it seems that there could still be a causal order to the events.

*27.1.1.3 Objection 3: Inclusion of Causally Irrelevant Factors.* Causal Explanation also seems to include causally irrelevant factors as causes. This objection challenges the sufficiency of Explanationist definitions of causation. That is, the objector attempts to show that pairs of propositions can satisfy Explanationist conditions without being genuine cases of causal explanation.

There are two cases to consider, namely, conjunctions (compound propositions formed by 'both' and 'and') and disjunctions (compound propositions formed by 'either' and 'or').

1. Addition of irrelevant conjuncts. Suppose that *A* is a genuine causal explanation of the truth of *B*. Let *C* be an irrelevant truth. Explanationist accounts wrongly count the conjunction *A*&*C* as a causal explanation of *B*.

- If *A*, together with the laws of nature, entails *B*, then so will *A*&*C*.
- If *A* raises the probability of *B*, relative to some background condition, and *C* is causally irrelevant, then *A*&*C* will also raise the probability of *B*.
- Suppose it is true that if *A* were not the case, then *B* would not be the case. Let's suppose that *C* is some causally irrelevant truth such that, if *A*&*C* were not the case, then *A* would not be the case. (In other words, to use Lewis's semantics, *A* is false in all of the closest worlds in which *A*&*C* is false.) On those suppositions, it will be true that if *A*&*C* were not the case, then *B* would not be the case, wrongly making *A*&*C* a cause of *B*.

2. Addition of irrelevant disjuncts. The version of the causal relevance objection that exploits irrelevant disjuncts applies only to the Probability-Raising Theory. Once again, suppose that *A* is a causal explanation of the truth of *B*. Let *C* be an irrelevant truth. If *A* raises the probability of *B*, relative to some background condition, and *C* is causally irrelevant, then (*A* or *C*) will also raise the probability of *B*.

The upshot of the foregoing is that Causal Explanationism faces serious challenges. We turn now to Causal Connectionism, to test whether it fares better.

## 27.2 Causal Connectionism

### 27.2.1 Truthmakers and the relata of the causal connection

Causal Connectionists posit a category of thing, events (or states or conditions), members of which stand in the fundamental relation of *causation*. Recall again (1) and (2):

(1) Mary's kicking the ball caused the ball's entering the goal.
(2) The ball entered the goal because Mary kicked the ball.

Connectionists propose that we are to understand (2) in terms of (1), not vice versa. The fundamental truth is (1). We can appropriately state something like (2) because the two events mentioned in (1) stand in the relation of *causation*. Causal Connectionists provide a causal explanation of a truth by accurately describing part of the actual causal structure of the world. But, what part of the world's causal structure? To explain the ball's entering the goal, we must describe a part of the causal structure of the world that is somehow relevant to the truth of the proposition that the ball entered the goal.

The most natural proposal would be to turn to the idea of truthmakers, as discussed in Chapters 2 and 3. We causally explain the truth of *p* by providing information about the causes of the truthmaker of *p*. On this view, truthmakers are the sorts of things that stand in the causal relation to each other. If that is right, then we could in principle provide a causal explanation of any truth that has a truthmaker, provided that the truthmaker itself has a cause.

Therefore, Causal Connectionism provides new grounds for Truthmaker Theory (2.1T/2.1A.1T/2.1A.1A.1T). In addition, it provides us with grounds for extending the truthmaker principle to various categories of true propositions. Any true proposition for which we can give a causal explanation must have a truthmaker. Ideally, each such truth will have exactly one truthmaker. Given this connection, any reason for extending the scope of causal explanation becomes a reason for extending the truthmaker principle.

Suppose, for example, that every contingent truth has a causal explanation. If so, then every contingent truth has a truthmaker. Similarly, if every truth attributing a finite attribute or every true proposition attributing a beginning to some condition has a cause, then every such truth also has a truthmaker.

*27.2.1.1 A Linguistic Argument for Causal Relata: Davidson's Theory of Adverbs.* Donald Davidson (1967) provides us with a further, linguistic argument for Causal Connectionism. We'll focus here on one piece of linguistic data that is best explained by positing the existence of events as truthmakers of certain propositions, namely, the use of adverbs. Consider (7) and (8):

(7) John walked.
(8) John walked quickly.

Clearly, the truth of (8) entails the truth of (7). John cannot walk quickly without walking. The best linguistic theory of adverbs would provide a simple and uniform explanation of this logical entailment. Davidson (1967) offers the following theory. Action verbs like 'walked' introduce an existential quantification over events. Consequently, the logical form of (7) and of (8) is best captured by (9) and (10):

(9) There is a walking, and it is by John.
(10) There is a walking, and it is by John, and it is quick.

Statement (9) now follows evidently from (10) by the rule of simplification (from *A*&*B*, infer *A*). The alternative theory would have to suppose that adverbs like 'quickly' represent functions from one set (the set of those who *V*) to another set (the set of those who *V* quickly). We would then have to introduce a new linguistic axiom to the effect that anything that *V*'s quickly also *V*'s. In fact, an infinite number of axioms of this kind would have to be added. For example, we would have to add an axiom to the effect that those who *V* quickly and gracefully also *V* quickly and *V* gracefully in order to capture the fact that (11) logically entails (7) and (8):

(11) John walked quickly and gracefully.

Davidson's theory captures all of these implications without the need to add additional axioms. The logical form of (11) can be seen in (12):

(12) There is a walking, and it is by John, and it is quick, and it is graceful.

In addition, Davidson's theory handles tense easily. The past tense of (7)–(12) can be captured simply by adding a clause to the effect that the walking is past.

Davidson's theory gives us good reason to extend the scope of Truthmaker Theory to include all atomic truths containing an action-verb, like 'walks', since any such sentence can meaningfully be modified by adding an adverb. The verb 'to be', in contrast, does not naturally accept adverbial modification. In English, we can say 'John is wise' but not 'John is wisely'. We can sometimes modify a verb phrase that contains 'to be', however. We can say 'John is effortlessly wise', for example. This suggests that the truthmaker principle can be applied to statements of condition or state as well as to action-statements.

Jaegwon Kim (1973, 1976) proposed a theory of *events* that can also be taken as a theory of truthmakers. Kim proposed that there is an event whenever there is a particular that bears a property at a time. This corresponds to the proposal that every statement of the following kind has a truthmaker that is (at least potentially) a cause or effect:

(13) Particular *A* has property *P* at time *t*.

Of course, Kim's proposal raises several metaphysical questions: what are particulars, properties, and times? Presumably, we should interpret the Kimian theory as applying only to fundamental particulars, properties, and times.

If events or causal relata are truthmakers, then each theory of truthmakers corresponds to a theory of the relata of causation. As we saw in earlier chapters, there are three principal accounts of the nature of truthmakers: tropes (both modules and modifiers), nexuses (between universals and particulars), and states of affairs (constituted by universals and particulars).

*27.2.1.2 Fine-Grained Causal Explanation and Truthmakers.* The phenomenon of *fine-grained* causal explanation provides further support for extending the truthmaker principle to all true statements involving the predication of fundamental properties. In explaining things causally, we often appeal, not only to what actions have been done, but also to how they are done. For example, L.A. Paul (2004) has argued that we can explain Susan's writing slowly by appealing to her broken arm, and her writing insightfully by appealing to her intelligence. The insightfulness of her writing can have effects that are not shared by the slowness of her writing. Thus, the causal structure involves more than just actions, where these actions are individuated coarsely (by, for example, the main verb occurring in the corresponding true statement).

(14), (15) and (16) can have different causes and effects:

(14) Susan wrote an article.
(15) Susan wrote an article insightfully.
(16) Susan wrote an article slowly.

If so, and if causal relata are truthmakers, then we must suppose that these three truths have different truthmakers. For example, the truthmaker for (15) must include some connection (by way of a trope, nexus or state of affairs) between Susan's writing and the property of intelligence. Neither (14) nor (16) require such a connection.

### 27.2.2 Objection to Causal Connectionism: Negative causation

The weightiest objection to Causal Connectionism concerns the phenomenon of negative causation. Causal Connectionism entails that causation is always a matter of some real relation between things (events or truthmakers). However, there are many cases in which we appeal to absences and other negations, like holes, shadows, privations, and inaction, to explain real phenomena. In those cases, there doesn't seem to be any real entity to bear the causal relation to the effect.

Causal Connectionists can respond in either of two ways. First, they can deny the reality of negative causation altogether. Second, they can attempt to provide a Connectionist account of negative causation. This accounts can take one of two forms. Either the account attempts to accommodate real negative causation without introducing absences or other negative entities or the account posits the real existence of negative things. We'll consider each of these responses in turn.

*27.2.2.1 Is Negative Causation Real?* Schaffer (2004) has provided a number of compelling examples of negative causation:

- Death by heart attack. People who die as a result of a heart attack die because their hearts are not pumping blood, which results in a shortage of oxygen in the bloodstream, which in turn leads to the failure of the cells to perform certain activities essential to the maintenance of life.
- Muscle contraction. The contraction of muscles is caused by the transmission of an electrical signal through the nervous system, which in turn involves the successive activation of synapses between serial nerve cells. The activation of a synapse occurs as the result of the removal of certain chemical obstacles to the movement of ions across the gap.
- Firing of a gun. Pushing the trigger removes an obstacle to the expanding of a spring, which triggers the firing of the bullet.
- Bombs can be designed in which pushing a button removes the shielding of one element from another within a bomb, resulting in an explosion.

Such examples can be multiplied. Droughts can cause famine by depriving plants of the water they need. As Schaffer argues, these are central and not peripheral examples of what we mean by 'causation'.

*27.2.2.2 Can Connectionists give an Account of Negative Causation?* It would seem, then, that Connectionists must provide some sort of account of negative causation. There are two avenues to explore: posit negative entities, such as real absences, to serve among the causal relata, or provide a disunified or disjunctive account. A disunified theory would have different implications for positive and negative causation. Positive causation would involve a connection between two entities, while negative causation would not.

## a. Disunified theories

A disunified theory of causation would have to be Connectionist about positive causation and Explanationist about negative causation. Phil Dowe defends such a view in *Physical Causation* (2000). All such accounts face an obvious challenge: if Explanationism is good enough for negative causation, why isn't it good enough for positive causation as well? The danger is that the arguments that Connectionists offer against the adequacy of Explanationism in the case of positive causation will have exact counterpart arguments against Explanationist theories of negative causation. It is relatively easy to produce cases of symmetric overdetermination or preemption involving negative causes. One's dying for lack of water can be preempted by one's dying for lack of oxygen, for example (see Koons 2003).

It would be hopeless to deny entirely that absences play any significant role in causation. However, Powerists (4.4A.3) can plausibly make a distinction between the role of positive things and their causal powers, on the one hand, and the role played by mere absences, on the other. Whenever an absence does play a role in causing a change or initiating a new process, it always does so in collaboration with some positive entities possessing appropriate active and passive causal powers. For example, when a gun is fired, the absence of the obstacle does play a role, but only in the presence of a spring or other mechanism with the active power of producing the appropriate motion in conjunction with the absence. Even when death is produced by an absence of water or oxygen, there is always a living organism with the capacity to produce, in those circumstances, an unliving corpse.

It would, therefore, be tempting for Connectionists to reduce the causal role of the absence to the fact that a certain result can be deduced from the proposition that something is absent, when that proposition is combined with propositions about the laws of nature and about the states of the other, positive participants in the causal interaction. Even if absences don't exist, there do exist propositions to the effect that certain things are absent, and we can always consider what follows logically from such propositions, either alone or in conjunction with others. Connectionists could distinguish between causes and enabling conditions, arguing that absences can never be causes in the strict sense, but only conditions that enable something else to act as a cause.

However, this maneuver comes at a significant cost, since it threatens to deprive Connectionists of the argument from causal linkage against Explanationism. This problem emerges when we focus on the other form of negative causation: prevention, or causation of absences. Consider a case of two, symmetrical potential causes $C_1$ and $C_2$ of some event $E$, and suppose that the effect $E$ is an absence, the absence of an event of type $K$. It still seems to make sense to ask, on each particular occasion, whether it was $C_1$ alone, $C_2$ alone, or $C_1$ and $C_2$ jointly that produced the absence? However, we will always be able to give a causal explanation of the proposition that no event of type $K$ occurred that refers to both $C_1$ and $C_2$. How then can we distinguish the three possibilities?

The answer to this problem might lie in hypothesizing that effects always consist in some change or absence of change in a persisting *patient*. When we prevent some event from happening, this always involves our preventing some thing, the *patient*, from changing in certain ways. We could now suppose that the continuing existence of any patient is a temporally extended *process* of a certain kind. (We explain more about what this means in Chapter 28.) The causal linkage between the *agent* and the absence of change in the patient consists in some powerful state of the agent's being an integral part of the process of the patient's continued, unchanging existence. If $C_1$ causes the absence and $C_2$ does not, then some state of $C_1$'s power will be literally a *part* of the patient's continued and unchanging existence, while no such state of $C_2$ will be contained in that process of continuing existence.

### b. Reified absences

The second option, reifying absences, also comes at a theoretical cost. Indeed, the cost might seem prohibitively high. If we reify all of the absences in the world, we will end up with Truthmaker Maximalism (2.1T.1). First, this involves a potentially very heavy cost in terms of ontological baggage. One must add one or more entirely new categories of things (absences, non-beings, privations, etc.), and one must populate the world with an infinity (at a very high level of cardinality) of such things. Absences of many kinds are potentially nearly ubiquitous. Think how many absences of hippopotamuses there are, to take just one example, not just in many stretches of river in Africa, but throughout the center of the earth, in every teaspoon of sugar, and throughout the vast stretches of outer space. In addition to the absence of hippopotamuses in general, we would have to consider the absences of this or that particular hippopotamus. We might have to recognize the existence of absences of merely possible hippopotamuses, to say nothing of all of the absences of impossible entities, like round squares.

Not only does Truthmaker Maximalism's addition of an infinite number of negative facts weigh against the theory by way of inflating its ontology, but also it requires (as we saw in Chapter 2) a huge number of mysterious necessary connections between positive and negative facts (**PMeth 1.2**). It must be impossible, for example, for the presence of water and the complete absence of water to coexist in the same place and time, and it must also be impossible for both the presence and the absence to be absent! There must be some necessary connection between the existence of a hippopotamus in some specific location and the non-existence of a hippo-absence in that same place, and a converse necessary connection between the existence of an absence of a hippopotamus somewhere and the non-existence of a hippopotamus there and then. This makes it impossible to change the world simply by deleting things. Each deletion necessarily involves a simultaneous creation. Deleting hippopotamuses necessarily increases the population of hippo-absences.

These objections can be at least partially met by introducing totality facts as the fundamental connections between universals and particulars. A negative atomic truth like '*b* is not *F*' is made true by some totality fact that includes the universal F-NESS but not the particular *b*. Such totality facts could serve as the relata of negative causation.

Alternatively, we could make use of the absences that we discussed in Section 17.4, especially the view according to which absences were identified with bundles of bare particulars and spatial-location qualities, with no associated corporeal qualities like mass or charge. A complete absence of matter from a spatial region, that is, a vacuum, could be identified with a plurality of such absence-bundles, one for each bare particular, and each bundling that particular with the spatial region but with no corporeal qualities. We might need to add to this plurality the totality fact for the corporeal universals, like mass and charge, to ensure that our vacuum-fact hasn't left out any possible bodies.

In addition, this ontological inflation can be contained to some extent if we can limit, perhaps severely, the class of absences that are needed in a full account of negative causation. In addition, if we can limit the class of negative entities to a small enough range, we might be able to find positive entities with which the absences can be identified, entirely eliminating the ontological cost of absences.

A promising strategy for limiting the range of absences to a manageable size is John Haldane's (2007) proposal of privative causality, which builds on a long Aristotelian tradition. On this view, absences are causally efficacious only by connection to some causal power or disposition of a positive substance or process. In some cases, a thing's nature makes use of certain absences, as in the case of the firing of a nerve synapse by means of the absence of certain blocking chemicals normally present in the synapse. In other cases, a thing's nature requires the presence of some feature, either within the thing itself or in its environment. The absence of this normal condition causes deformity or failure on the part of the positive thing whose nature is thereby thwarted. We can postulate that in these cases, the absence corresponds to the position of a real, negative property, either by the thing itself or by something in its environment. Thus, the absence of water is a real property of the immediate environment of the drought-stricken plant, and the absence of blocking chemicals a real feature of the nerve synapse during firing. We can deny, however, that there are any absences except in such a relation to power-bearing individuals. The absences come into being by virtue of the coming to be present of some appropriate bearer of causal powers.

On this view, a real absence of hippopotamuses can occur only in relation to something that requires the presence of hippopotamuses as part of its normal operation. There could be an absence of male hippopotamuses in the immediate environment of some female hippopotamus, but no such absence occurs in the remoteness of outer space.

Since absences always involve the presence of some appropriate entity, we could take the further step of supposing that the absence is actually a feature of the entity itself. The absence of water in the environment of the plant could be taken to be a property of the plant itself. This would involve understanding the plant to extend somewhat into its immediate environment.

We would still need to expand our ontology to the extent of adding fundamentally negative properties, and we will still need to postulate some necessary connections between negative properties and their positive counterparts. These additions could be limited, however, to just those cases in which the dynamic natures of the things involved licenses the attribution of causality to the negative properties.

It is interesting to note that all of Schaffer's examples fit the privative model. In each case of negative causation, we have either a living thing or an artifact whose normal condition involves the absence in question. In each case, the absence has an effect either by being incorporated into a special kind of operation, like firing a bullet or activating a nerve synapse, or by depriving some operation of one of its necessary conditions, like depriving cells of the oxygen they need for respiration.

## Notes

1 There are interpretations of quantum mechanics according to which all change is determined by prior conditions, including some versions of the Everett many-worlds interpretation and of the Bohm-de Broglie pilot-wave interpretation. However, all of these deterministic interpretations face serious objections. The many-worlds account has the problem of explaining the meaning of the probabilities associated with quantum predictions (since on that account all of the possible results occur with probability 1). The most promising solution, the Albert-Loewer many-minds interpretation (Albert and Loewer, 1988), re-introduces indeterministic causation into the picture, since individual minds are caused to travel along one branch of the evolving ensemble of branches in a probabilistic fashion. The Bohmian account has to introduce probability at the very beginning of the universe's history, suggesting a probabilistic cause of that initial state.

2 Parenthetically, trumping raises an interesting meta-philosophical or methodological question. In its original form (in Schaffer 2000), trumping involved two spells belonging to two different orders of magic, one higher and one lower. Such an example seems very remote from our view of how the actual world works. If the task of the theory of causation were simply conceptual analysis, then this remoteness from actuality wouldn't seem to matter, so long as our intuitions about the case are clear. If, in contrast, we are after the best theory about causation as a real feature of the world, then only examples that are clear examples of causation and non-causation in the actual world (or at least, in worlds we have reason to believe are genuinely, metaphysically possible) should count.

However, as Lewis (2000) points out, it's quite possible that something like trumping occurs in the actual world. For example, if Pauli's exclusion principle absolutely prevents an electron from going into a particular state because that state is already occupied by a different electron, we might want to say that the exclusion principle preempts by trumping any other process that might also, in the absence of the exclusion principle, have prevented the electron from entering that state.

# 28

# Discrete and Continuous Causation

Causal Connectionists need to provide an account of causal linkage and of causal direction. However, before we can turn to the details of such an account, we must distinguish between two kinds of causal connection, namely, *discrete* and *continuous*. A causal connection is discrete when there is either a direct causal link between the cause and the effect, or there is at most a finite number of such links between the two. Whenever we find an infinite number of intermediaries between cause and effect, we have a case of continuous causation.

**Def D28.1 Causal Betweenness.** Event *z* is *causally between* events *x* and *y* if and only if *x* causes *z* and *z* causes *y*, or *y* causes *z* and *z* causes *x*.[1]

**Def D28.2 Linear causal order.** A set of events *S* is a causal linear order just in case for any three distinct events *x*, *y*, and *z* in *S*, one of the events is causally between the other two.

**Def D28.3 Discrete Causation.** The causal connection between *x* and *y* is *discrete* if and only if *x* and *y* are causally connected, and the linear casual order consisting entirely of events between *x* and *y* is finite. (Compare Def D19.1 Immediate Causation.)

**Def D28.4 Continuous Causation.** The causal connection between *x* and *y* is *continuous* if and only if *x* and *y* are causally connected, and there is an infinte linear causal order consisting entirely of events between *x* and *y*.[2]

*The Atlas of Reality: A Comprehensive Guide to Metaphysics*, First Edition.
Robert C. Koons and Timothy H. Pickavance.

## 28.1 Is All Causation Discrete?

There seems to be good reason to think that there cannot be causation across a temporal gap. That is, if *C* causes *E*, then there cannot be a time gap between the end of *C*'s existence and the beginning of *E*'s existence. After all, to talk of causation as involving a causal connection between two entities seems to imply that the two entities both exist at the time at which they are connected.

**Partial Simultaneity of Discrete Causes.** If *x* is a cause of *y*, and *x* and *y* do not belong to any one process, then *x*'s time of occurrence includes the time at which *y* begins.

In fact, it seems that we can go further than this and insist that when a cause and effect are connected discretely, they must be fully simultaneous, beginning and ending at exactly the same time. If a cause pre-exists its discrete effect, then it seems that we can divide the supposed cause into two entities, one that exists entirely before the effect and the other that begins to exist at the same time as the effect. It is only the latter that is, strictly speaking, causally connected to the effect. Similarly, if an effect endures after the end of its cause, we can make a similar distinction between the effect's earlier and later parts.

**Full Simultaneity of Discrete Causes.** If *x* is a discrete cause of *y*, then *x* and *y* begin and end at the same time.

It also seems plausible that some causes are earlier than their effects, especially if causation is transitive. If this were not the case, then we could never explain anything by reference to things existing at earlier times, which is obviously not the case.

**Causation by Earlier Events.** Some events are caused by earlier events.

The principle of Full Simultaneity and Causation by Earlier Events jointly entail that at least some causation is continuous. If all causation were discrete, and all discrete causation were fully simultaneous, then it would be impossible for earlier events to cause later ones.

## 28.2 The Nature of Discrete Causation

Causal Connectionists have a number of options for explaining the linkage between causes and effects in the case of discrete causation. Here are four popular options:

1. *Causation* as a primitive relation (like *instantiation* or *distinctness*) holding between pairs of truthmakers.
2. A causal link as an entity in its own right, consisting of one or more relational tropes (or nexuses or states of affairs). This tie could either be a single relational modifier trope or a pair of modular tropes.

If there were something to the causal connection above and beyond the effect, then that additional entity would be a truthmaker of the causal connection. Is the addition of such a truthmaker a plus or a minus? If there is a correspondence between the class of truthmakers and the class of causal relata, then the introduction of causal truthmakers threatens to generate a new infinite regress. Suppose, for example, that event *A* causes event *B* and that *C* is the truthmaker linking *A* to *B*. If *B* is contingent, so will *C* be. Hence, we should expect a cause of *C*. This cause is either identical to *A* or it is a new entity, *D*. In either case, there will have to be a truthmaker for the *A*-to-*C* or *D*-to-*C* causal connection. This truthmaker, *E*, will require a further cause, and so on.

3 A causal link as the transfer of a conserved quantity (mass, energy, charge or information) or the transfer of a trope.

This is a theory that has been defended by Fair (2003), Dowe (1995), and Ehring (1997). We will consider a version of this account in more detail when considering continuous causation. The transfer theory provides a clear account of the linkage between causes and effects, since the linkage consists in the identity of some trope or quantity attached to both the cause and the effect. However, it is not at all clear how the transfer theory explains the grounding of the causal direction. What constitutes a quantity's being transferred *from* one thing *to* another, especially when the cause and effect are simultaneous? An answer to this question would seem to require a prior explanation of the nature of causal direction.

4 A causal link as the exercise of one or more causal powers.

If we adopt Powerism (4.4A.3), then it would be natural to identify instances of causation with exercises of causal powers. The problem of grounding the linkage between cause and effect would then become the problem of linking a causal power to the result of its exercise. Let's consider this picture in more detail. On the one hand, we have some substance or process with an appropriate causal power, such as a fire with the power to heat bodies in its vicinity. On the other hand, we have the result of the exercise of the power, such as the heating of some body of water. What about the exercise of the power? Is it some third thing, linking the power to its result? As we've seen above, introducing a third entity here runs the risk of starting an infinite causal regress. A traditional answer, going back to Aristotle's *Physics*, is to identify the exercise of the power with the result. The exercise of the fire's power to heat is simply the heating of this water.

How then is an exercise (or result) tied to its originating power? There seem to be two options. First, we could take this relation to be a primitive relational fact, like *instantiation*. Alternatively, we could identify it with a relation of *asymmetric token necessitation*. That is, what ties this effect *E* to that causal power *P* is the fact that this very effect *E* could not have existed were *P* not to have existed and were not to have been in circumstances appropriate to the exercise of *P*. That is, it is of the very essence of the kind of power *P* that it be capable of being exercised in circumstances of the kind of actual circumstances *C*, and it is of the very essence of this very truthmaker *E* that it could not possibly have existed except in worlds in which this very power *P* and circumstances *C* exist. The existence of *E* as a particular necessitates the existence of *P* and *C*, as particulars, and the nature of *P* is such as to permit the production of something like *E* in circumstances like *C*. (For more details about asymmetric token necessitation, see Koons 2000.)

This necessitation has to be asymmetric, in order to ground the directedness of causation. That is, the existence of *E* necessitates the existence of *P*, but not vice versa. It must be possible for *P* and *C* to exist without *E*'s existing, but impossible for *E* to exist without *C*'s and *P*'s existence. Such asymmetric necessitation fits very nicely in a world with indeterministic causal laws, since such laws never entail, in conjunction with the existence of the cause, the existence of the effect. However, it is possible to affirm both deterministic laws and asymmetric token necessitation, so long as it is metaphysically possible for each of the deterministic laws to be violated.

## 28.3 Is All Causation Continuous?

Let us now turn our attention to continuous causation. We have seen that there is some reason to think that some causation must be continuous, in order to explain the transmission of causal influence from earlier times to later times. Could all causation be continuous?

How are we to think about continuous causation? How can there be an infinite number of intermediaries between a cause and its effect? Wouldn't such an infinity of intermediaries involve the existence of infinite causal regresses? It seems, after all, that if $E_1$ causes $E_3$, and there is an intermediate cause $E_2$ such that $E_1$ causes $E_2$ and $E_2$ causes $E_3$, then the causal link between $E_1$ and $E_3$ should depend upon the links between $E_1$ and $E_2$ and $E_2$ and $E_3$. If there were further intermediate links between $E_1$ and $E_2$, and between $E_2$ and $E_3$, then those two links would depend on those further intermediate links, and so on ad infinitum. We seem to have an infinite regress of *causal dependency* relations.

The error in this way of thinking about continuous causation is that it tries to understand continuous causation in terms of discrete causation, as though continuous causation simply consists in an infinite number of discrete causal connections. We should instead take seriously the idea that continuous causation involves a continuum of events. In the basic case, two events are connected by continuous causation when they are both parts of a single process. A real process is a temporally extended whole that is more metaphysically fundamental than any of its unextended, instantaneous parts. Later parts of the process are dependent on earlier parts because both are parts of the same process, not because there is some discrete connection or chain of discrete connections between the two.

Doesn't this solution still involve a problematic infinite regress of dependency? We can find an infinite series of events, each earlier than its predecessor in the series. Consequently, each event in the series would *depend* in some sense on its successor, ad infinitum. Isn't this an objectionable kind of infinite regress? How does the metaphysical primacy of the whole process help?

What's needed here is a distinction between a grounded and an ungrounded infinite regress. We think that it's reasonable to believe in an infinite regress, in which event $E_1$ depends on event $E_2$, event $E_2$ on event $E_3$, and so on, so long as all of these dependency relations are themselves grounded in a common source, one that is independent of all of the members and that does not itself give rise to a further regress. The problem with

Bradley's Regress (Section 7.2.1.3), for example, was that there was no way of introducing such an ultimate ground of the *instantiation* relation without simply falling into another regress.

More specifically, when each new event emerges as the process unfolds, the new event $E_1$ is *immediately* dependent on the whole process up to that point and, therefore, on each prior event within the process. This dependency is modal or counterfactual: if the whole process had not unfolded as it did up to the occurrence of $E_1$, $E_1$ could not have occurred. Thus, if $E_1$, $E_2$, and $E_3$ all belong to the same process, with $E_3$ the earliest and $E_1$ the latest, then $E_1$ depends immediately on both $E_2$ and $E_3$, through its dependence on the whole process, even if it is also true (for similar reasons) that $E_2$ depends on $E_3$. We can find a fundamental ground for each *dependency* relation in the process itself, and no Bradley-like regress threatens.

With this clarification in mind, let's return to the question of whether all causation is continuous. If all causation were continuous, then all causally connected events would have to be parts of one and the same process. The whole world would consist of a single continuous causal process, without interruptions. Thus, the non-existence of discrete causation leads to a kind of Monism (11.2A). The world would consist of a single truthmaker.

However, as we have seen, there is some connection between the possibility of our knowledge of modality and the existence of multiple truthmakers. A Monistic world would be a world that provided us with no knowledge about the causal powers of things, which would make it impossible for us to know anything about alternative, contrary-to-fact possibilities.

Monist accounts of our knowledge of modality would seem to amount to a distinction between nomic and accidental regularities, probably in line with the Mill/Ramsey/Lewis (MRL) Theory of laws (see Section 5.2), according to which laws are the axioms of the best theory of the world, where being *best* consists in simultaneously maximizing simplicity, closeness of fit, and breadth of coverage. Sentences stating the boundary conditions of the universe are too messy and various to count as laws. Hence, such sentences count as being only contingently true, giving rise to alternative possibilities.

However, we've seen a number of powerful objections to the Ramsey/Lewis Theory of laws, especially the difficulties that *small worlds* pose for the theory. In addition, it is difficult to see what connection there could be between non-lawfulness as defined by the Ramsey/Lewis Theory) and contingency. Why assume that the laws are necessary? Why assume that non-laws are contingent?

On the alternative, Powerist account, we can have reliable knowledge of local necessities and contingencies, by way of gaining knowledge of the causal powers of particular things. Aristotelian Modality (15.2T.7), which fits nicely with Powerism, requires multiple, disjoint truthmakers, belonging to an evolving network of causes and effects. To gain knowledge of these truthmakers, we must encounter cases of discrete interaction between two or more processes. Thus, knowledge of modality seems to be connected to the occurrence of discrete causation. If we have no knowledge of alternative possibilities, we have no understanding of the actual world or of the contents of our thoughts or concepts.

## 28.4 The Nature of Continuous Processes

We've seen that continuous causation depends on cases in which two events are parts of the same process. When does a set of events constitute a single process? In addition to real processes (carriers of continuous causation), the world is filled with pseudo-processes (Reichenbach 1958). Moving spotlights and shadows are classic examples of such pseudo-processes. Suppose that a domed stadium has a rotating lamp at its center, projecting a moving spot of light that appears to revolve continuously around the stadium's inner wall. Each stage of the history of the moving spot is connected continuously with earlier and later stages. The intensity and position of the spot might change continuously. However, the various spot-events in this history are not really causally connected with one other. Each spot-event is dependent on the process by which light travels from the central lamp to the wall, but it isn't in any way dependent on the apparent process of the movement of the spot of light. You can't affect the future position or shape of the spot by trying to do something to the present condition of the spot.

What distinguishes real processes from pseudo-processes? There are two kinds of answers that we could give: reductive and non-reductive.

### 28.4.1 Reductive accounts of processes

1. Epistemic analyses. Bertrand Russell (1948) recognized this problem. Russell called real processes "causal lines", and he defined *causal lines* in terms of what we can correctly infer from facts about the parts of such lines:

> A 'causal line,' as I wish to define the term, is a temporal series of events so related that, given some of them, something can be inferred about the others whatever may be happening elsewhere. ]source[(Russell 1948: 459)

Wesley Salmon (1984) objected that we need an *ontic* and not an epistemic account. Surely facts about which lines are real processes and which are pseudo-processes do not depend on us or on our epistemic practices of inference. There were moving shadows and spots of light and real processes of movement, heating, and growth long before there were any humans around to infer anything. Moreover, Russell's analysis puts the epistemic cart before the metaphysical horse. The reason that it is correct to make inferences in some cases and not others is to be explained in terms of the metaphysical difference between processes and pseudo-processes, not the other way around.

2. Spatiotemporal continuity and quantitative constancy. Another simple but deeply flawed account defines a real process as any set of events that is spatiotemporally continuous and that is characterized throughout by a fixed amount of some fundamental quantity whose conservation is guaranteed by a law of nature, like mass-energy.

Phil Dowe's (1995, 2000) account is an example of such a theory. He proposes that continuous causation consists simply in the *transmission* of some quantity whose conservation is guaranteed by physical law. Dowe thinks that whether a quantity has been transferred along a certain continuum of events is wholly determined by the facts about how much quantity there is at each point of spacetime, together with the relevant laws of nature.

The central problem with this account is that there are pseudo-processes that satisfy this definition. Suppose that the spot of light moves in a continuous fashion across the wall, so that between any two times, the spot is located in a position spatially between the positions it held before and after, and the spot is located at the limit of the locations it occupies at each of the members of any infinite but bounded series of times. Suppose, in addition, that there is a fixed amount of mass and energy associated with the spot at each moment. Such a moving spot will satisfy this definition without being a real process.

In addition to the problem of failing to exclude certain pseudo-processes, Dowe's account also fails to resolve possible ambiguities in nature, like cases where it is unclear on empirical grounds which of two sets constitutes a process and which a mere pseudo-process. Consider a case of ambiguous exchange involving a pair of interacting particles. There are four events: $C$, $C'$, $E$, and $E'$. There are laws of nature linking the four pairs ($C$-$E$, $C'$-$E$, $C$- $E'$, $C'$- $E'$) in the following way: each of $C$ and $C'$ have a 50% probability of transferring their quantity of stuff (energy, momentum, or whatever) to each of $E$ and $E'$. In fact, the sum of the conserved quantity in $E$ and $E'$ is exactly identical to the sum of it in $C$ and $C'$. Where did the stuff in $C$ go? The laws of nature give no determinate answer. There are two equally good alternatives: a transfer from $C$ to $E$ (and a simultaneous transfer from $C'$ to $E'$), and a transfer from $C$ to $E'$ (and from $C'$ to $E$). Even though we cannot tell empirically which process is real and which pseudo, there should always be a fact of the matter as to which is which, a fact that Dowe's theory must deny.

Such ambiguity might occur even in deterministic cases. Suppose that two point-particles, each carrying an equal quantity of energy, are on intersecting paths. Two identical particles converge on the intersection point, and two identical particles emerge from it. Was there a collision or did the particles pass through one another without effect? The conservation laws can't answer that question, since, in either case conservation was preserved.

Finally, such constant-quantity theories wrongly exclude real processes that are mereologically inconstant, gaining or losing matter, energy, information, and other quantities. Many biological processes are like this. Photosynthesis, for example, takes in energy from sunlight at a variable rate. Even locomotion at the macroscopic level loses energy through friction, dissipating its energy into the surrounding environment.

3. Counterfactual dependency. One critical problem with the second theory was that it could not distinguish mere accidental continuity and quantitative constancy from the kind of continuity and quantitative constancy that are regularly and nomically associated with a real process. We might try to use counterfactual conditionals to distinguish between the two cases. Wesley Salmon's (1998: Chapter 16) Mark Transmission Theory is an example of such an approach. Salmon requires that real processes be capable of transmitting a mark, which comes to something like the following condition:

**28.1T Mark Transmission Theory.** A set of events $P$ constitutes a real process if and only if (i) the spatiotemporal locations of the events in $P$ form a continuum $C$, (ii) some conserved quantity $Q$ is associated with the events in $P$ at a constant value throughout $C$, and (iii) there is some action $M$ (the 'marking' action) which did not in fact occur during $C$, but which is such that, if $M$ had occurred at some spacetime location within $C$, then every subsequent event in $P$ (but none of the earlier events) would have been replaced by a different event associated with a different value of $Q$.

The marking event $M$ is something that would have altered the mass, energy, information or other conserved quantity associated with the process. It is clause (iii) of the theory that introduces the counterfactual conditional: if $M$ had occurred at time $t$, then the subsequent events of $P$ would have been replaced by a different set of events, a set of events appropriately *marked* by the action $M$.

Not surprisingly, this counterfactual conditional theory of continuous causation is subject to some of the same objections as is the Counterfactual Conditional Theory of Causal Explanation (27.1A.2). First, there are problems concerning linkage, or the unity of causal processes.

Second, the counterfactual conditionals can be *finked* (just as the counterfactual account of dispositions was: see Section 4.4). To return to the rotating spotlight example, let's suppose that the operator of the central lamp is disposed to move the spot of light in response to movements of a spectator located near the wall. If the spectator "bats" the spot as it moves past, the operator is disposed to alter the path of the spot, simulating a ricochet effect. In this case, the spectator could mark the spot with a new momentum in such way that if the spectator were to do so, the subsequent events in the pseudo-process would be affected. Thus, a pseudo-process can satisfy the Mark Transmission Theory.

Finally, there is the problem of accounting for the asymmetry of continuous causation. If we run the tape backward, the conserved quantity will apparently be transferred from the effect to the cause. As we have seen, counterfactual conditionals do not seem to be capable of grounding the asymmetry of causal direction.

4. Ehring's trope transmission theory. Douglas Ehring (1997) proposes that tropes can be used to provide a reductive, non-circular account of the difference between real processes and pseudo-processes. A set of events constitutes a real process just in case each event consists in the process of some trope at some spatiotemporal location, and the very same trope is present in every event in the set. The unity of a real process is grounded in the diachronic identity of its associated trope. One version of Ehring's trope theory would focus on tropes of a special kind, such as tropes of conserved quantities like energy.

Ehring's account does a good job of accounting for the unity of real processes, but it founders in its attempt to ground causal asymmetry. As Ehring recognizes, trope persistence is an entirely symmetrical relation. Hence, it cannot provide a basis for causal asymmetry.

In addition, Ehring's account doesn't seem to add anything to a simple, non-reductive account of processes (which we will discuss in the next section). Ehring assumes that tropes do not have fundamental temporal parts. Each trope is wholly present at each time during the lifespan of the process. In other words, tropes are simples extended in time. Why not simply take the processes themselves to be simple, extended, four-dimensional tropes?

### 28.4.2 A non-reductive account

We can draft a simple, non-reductive account of processes. Processes are mereologically simple, having no actual parts. A process is only *potentially* divisible into sub-processes. On this view, instantaneous events are dependent boundaries of processes.

What sort of things are processes? They could be properties (i.e., universals or tropes) of enduring substances, properties that are extended in four dimensions (time as well as space). In fact processes could be instances of temporally extended structural universals (see Section 10.3).

Let's take a simple example, namely, locomotion. Suppose that a baseball is struck by the batter and caught by the right fielder. The process of motion is a property of the ball that extends over time, from the time of the batting to the time of the catching, and across space, from the bat to the fielder's glove. The process has many potential parts, sub-processes that connect its starting point to various intermediate positions of the ball. Each sub-process is a distinct potentiality, namely, the ball's potentiality to be stopped at each of these intermediate positions.

On this view, the distinction between real processes and pseudo-processes is a simple one: processes exist and pseudo-processes do not. In the case of the moving spot of light, there is nothing that possesses any property corresponding to the pseudo-process. There is no spot of light, nor is the spot constituted by a plurality of real things. In contrast, a moving baseball consists of a large number of small physical objects, each of which has a real property corresponding to the locomotion.

This way of thinking about causal process fits well with the Intervalist and Proceduralist views discussed in Chapter 19:

**19.2A Procedural Intervalism.** Some extended processes are not composed of metaphysically fundamental time-slices.

If temporally extended processes are metaphysically fundamental entities, and not mere heaps of events, then the causal dependency of one event on another can be understood in terms of their common inclusion in a single, metaphysically fundamental process. The asymmetric causal dependency relations would then be relations between processes. A process $P'$ depends on process $P$ just in case $P$ is a proper part of $P'$, $P$ includes every part of $P'$ at or before time $t$, and $P'$ extends beyond time $t$ into the future. Although there will be infinitely many sub-processes that are causally intermediate between $P$ and $P'$, the dependency of $P'$ on $P$ does not itself depend any of those intermediaries.

## 28.5 Processes and the Direction of Continuous Causation

If some causation is discrete, and the exercise of causal powers provides a direction to discrete causation, then the causal direction of processes can be derived from the causal direction of discrete interactions. For example, there will be some joint exercise of causal powers at the beginning of each process, responsible for the existence of the process. In contrast, if a process is ended by the exercise of some causal powers, those powers will be responsible only for shortening of the process, not for its very existence.

In addition, it seems plausible to suppose that each causal power must pre-exist its first exercise. It doesn't seem possible for a thing to gain a new causal power and to exercise that power in the very same instant. We could justify this claim by appealing to the following principle:

**Temporal Separation of Power Acquisition and Exercise.** If event *A* is the acquisition of some power by a thing, and *B* is an exercise of that very power by that thing, then *A* and *B* cannot occur simultaneously.

The acquisition of a power and the exercise of that power are two distinct events, with the first a precondition of the second. If so, the necessary priority of the acquisition of powers to their exercises would provide a basis for distinguishing an intrinsic arrow of time. Processes may have natures that determine their evolution in a time-specific way. These seem to be true of many macroscopic phenomena, whether chemical, biological, social, and astronomical. Finally, we can appeal to asymmetric token necessitation. The later parts of a process token-necessitate the earlier parts, and not vice versa, since any token process could be ended by interruption at any time after its initiation.

## 28.6 Are Processes an Exception to Hume's Epistemic Principle?

David Hume famously made the following claim about our knowledge of causation in his *Enquiry Concerning Human Understanding* (Hume 1748):

**Hume's Epistemic Principle.** We cannot infer with certainty and on purely a priori grounds the existence of a causal connection between two events.

Hume offers, in defense of his principle, a set of thought-experiments that seem to show that we can always imagine any event's having a different result from the one we in fact observe. Hume assumes that if we can coherently imagine a claim *p* to be false, then we cannot know, on purely a priori grounds, that *p* is true.

Causal processes like locomotion (conceived of as continuous movement through space) may provide counterexamples to Hume's Epistemic Principle. Let *A* be the event of a material body's approaching closer and closer to location *l* as time draws closer and closer to time *t*. Let *B* be the event of that same body's being located at *l* at *t*. In this case, we can surely infer with certainty and on a priori grounds that *A* is the cause of *B*.

In response, Hume would have to ask us to imagine a world in which both *A* and *B* occur without being part of the same causal process of locomotion. We can imagine a world in which *A* occurs without *B*'s occurring. Maybe the body approaches position *l* and vanishes exactly at time *t*, never actually arriving at *l*. This could involve the body's ceasing to exist at *t*, or it could involve the body's being instantaneously teleported to some remote location at *t*. Similarly, we can imagine *B* occurring without *A* (or anything remotely like *A*). We could imagine the body being created at time t in location *l*, or we could imagine its being instantaneously teleported from some remote location to *l* at the instant *t*.

However, Hume asks us to imagine both *A* and *B* occurring, yet without a causal connection between the two. In order to accomplish this, we would have to imagine the body's being teleported away at *t*, so that it never arrives at *l* by the ordinary process of locomotion. In addition, we would have to imagine the body's also being teleported away from this remote location to *l* in such a way that it arrives in location *l* precisely at time *t*. In order to do all this, we would have to imagine the body's being in two, widely separated place at time *t*: both in some remote location (as a result of the "first" teleportation) and in location *l* (as a result of the "second" teleportation). This is bad enough, but in addition we must imagine two simultaneous events of teleportation, one moving the body to some remote location *r* and a second moving the body from that remote location *r* back to *l*. However, it seems absurd to suppose that there could be two simultaneous actions, the second of which requires as a precondition the completion of the first action.

## 28.7 Conclusion: The Consequences of Causation

We saw in Chapter 26 that causation is needed in accounting for a wide range of phenomena—knowledge, semantics, modality, science, decision theory, agency—as well as being deeply embedded in our common sense view of the world. These facts strongly favor theories that treat causation and causal powers as fundamental features of the world, since all anti-realist and reductionist theory face the very heavy burden of providing an account of causation that is adequate for all of its applications.

In particular, these last three chapters (26–28) tip the balance in favor of Strong Powerism and against its rivals, especially Neo-Humeism (4.4T) and Nomism (4.4A.2). The latter two theories are inextricably tied to Causal Explanationism (27.1A) and share in its struggle to explain the real linkage between particular causes and effects (especially in cases of late preemption and trumping—Section 27.1.1) and the asymmetric direction of causation and time (Section 27.1.2), phenomena that Powerism, especially when combined with Temporal and Procedural Intervalism (19.1A and 19.2T) and Causal Connectionism, can readily accommodate. The only serious drawback for Causal Connectionism is the difficulty that that theory faces in explaining negative causation. However, as we have seen (in Section 27.2.2.2), a number of plausible are available to the Connectionists, including the selective reifying of some absences and the use of Totality Facts.

As this chapter has shown, the package consisting of Powerism, Procedural Intervalism, and Connectionism (which we might call 'the Aristotelian package') offers the only way of accommodating both discrete and continuous causation, both of which seem to play an indispensable role in our world.

## Notes

1 We are assuming in this definition that causation is transitive. If it isn't, then we can stipulate instead that there be a finite chain of causal links from x to z and z to y, or vice versa.

2 In modern mathematical terminology, this ordering would be called 'dense' rather than 'continuous'. We use the word 'continuous' here to mark a connection with Aristotle and Brentano.

# 29

# Conclusion: The Four Packages

In this course of this book, a pattern has emerged. There are two quite coherent packages of answers to the issues that we've discussed: a neo-Humeist or Ludovician package (championed by David Lewis), and a neo-Aristotelian package. Ludovicians favor the following answers:

LUDOVICIAN PACKAGE

1 Truth supervenes on being.
2 Neo-Humeism about powers (the Humean mosaic as fundamental).
3 Nominalism—either Ostrich or Resemblance, with or without tropes.
4 Primitive *identity* and *similarity* relations.
5 Intentionalism, with a causal-representational theory of perception.
6 Mereological Universalism.
7 Modal Concretism, with Counterpart Theory.
8 Anti-Tensism (the B Theory).
9 Spatial and Temporal Pointillism and Infinitism.
10 Composition as Identity, and Compositional Equivalence.
11 Perdurantism, of either the worm or stage variety (or a hybrid of the two).
12 Causal Explanationism, including the Counterfactual Theory of causal explanation and the Mark Transmission theory of processes.

What unifies these 12 theses are two fundamental commitments: to Ockham's Razor and to the sufficiency of modern physics. Ludovicians put little weight on common sense beliefs, especially when they are embedded in ethical and legal practices, and they do not rely heavily on the "manifest image of the world" (to use Wilfred Sellars's phrase): that is, they do not read too much into how things appear to us. In addition, Ludovicians

*The Atlas of Reality: A Comprehensive Guide to Metaphysics*, First Edition.
Robert C. Koons and Timothy H. Pickavance.

share the proclivity of much of modern science to explain things in a *bottom-up* fashion, privileging the very small and very short-lived over the large and enduring.

ARISTOTELIAN PACKAGE Aristotelians favor positions on this issues that are typically inconsistent to those of the Ludovicians:

1. Atomic or Spectral Truthmaker Theory, with the One Truthmaker per Fundamental Property principle.
2. Strong Powerism.
3. Constituent Realism, with universals as parts of their instances.
4. Individuation by bare particulars or by a structure-relational (amphibian) form of Bundle Realism.
5. Direct Realism, with Perceptual Dualism or Meinongian direct realism.
6. Organicism (with emergent biological and social powers).
7. Modal Abstractionism, with an Aristotelian theory of possibility, actuality defined in terms of metaphysical fundamentality, and Transworld Identity.
8. Tensism (the A Theory), usually in a modal form (Changing Possibility or Changing Actuality, or both).
9. Moderate Anti-Pointillism and Temporal Intervalism (with dependent boundaries), along with Finitism.
10. Aristotelian Compositional Reductionism.
11. Classical Endurantism.
12. Causal Connectionism.

Like Ludovicians, Aristotelians are interested in discovering elegant and powerful theories in metaphysics—theories that can be the basis of real metaphysical explanations. Consequently, they share with Ludovicians a commitment to Ockham's Razor. However, Aristotelians seek to explain a broader and richer body of data, a body that includes insights into our selves and even into the nature of the world of matter that go beyond the formal conclusions of theoretical physics. In addition, Aristotelians rely more heavily on our semantic intuition about what could possibly be the truthmakers for familiar kinds of claims, including those involving tense and modality. They are, therefore, less willing to embrace radically revisionary semantic theories, even when these theories assign the right truth-values.

We leave it as an exercise for the reader to trace all of the interconnections and dependencies among the positions in each package. We are not claiming anything as strong as that these are the only two logically possible positions. In fact, there are at least two alternatives, each deviating from both the Ludovician and Aristotelian package in several ways.

First, there is a more micro-physicalist approach that deviates from the Aristotelian package by taking physical simples to be the only fundamental entities and which embraces Pointillism, Instantism, and Anti-Tensism, as represented by David Armstrong or John Heil. This view differs from Ludovicians by embracing either universals or trope theory, along with a Constituent Ontology, and by relying on Atomic Truthmaker Theory. We will call this group the 'Fortibracchians', translating Armstrongs's name into Latin ('fortis' for 'strong', and 'bracchium' for 'arm').

Second, there is a more Platonic approach that deviates from the Aristotelian package by rejecting the truthmaker principle altogether and eschews nearly all talk of *grounding* and *fundamentality*. This sort of philosopher makes little or no appeal to Ockham's Razor or ontological economy. They do not think of metaphysics on the model of a theoretical science, like physics, astronomy, or chemistry. Rather, the metaphysician's job is almost entirely *a priori*, to be performed from the philosopher's armchair. We don't associate this package with any one paradigmatic philosopher, as we do the other three (Aristotle, David Lewis, and David Armstrong). Instead, it is a kind of composite image, which a number of philosophers resemble to some degree, such as Roderick Chisholm, Trenton Merricks, Alvin Plantinga, George Bealer, or Peter van Inwagen. We will call the position, for want of a better term, 'Quietism', in the sense that its proponents do not aim to make revolutionary discoveries of a theoretical nature but rather to organize, synthesize, and harmonize the elements of reality that appear clearly to us in our most reflective moments.

In Table 29.1, we've listed the thesis most likely to be adopted by each package on each of the issues considered in this book. In most cases, the association is quite clear, but there are several that are more debatable. We've listed the Aristotelian package as including Meinongianism, ontological vagueness, and Tensism on the passage of time. It is certainly possible to be an Aristotelian and to adopt the opposite positions on these theses, but we think that Aristotelians should be somewhat more open to them than those committed to the other three packages, and so associating them with the Aristotelian package helps to further highlight the differences between the four.

Ludovicians are centrally committed to Neo-Humeism. Consequently, they reject universals, which would fit most naturally with Nomism. Ludovicians could be Ostrich Nominalists, but given their commitment to economy and given the resources of modal concretism, extreme resemblance nominalism seems the best fit. The Humean mosaic is a bottom-up picture, so Priority of Point-Parts follows, as does Instantism and Pointillism. Neo-Humeism is also consistent with gunk, so heaps should be as fundamental as atoms. The mosaic picture also leads most naturally to Anti-Tensism and to Ramsey-Lewis-Sider perdurantism.

Fortibracchians' core commitments are to Nomism and against Extreme Nominalism. These commitments fit well with a Constituent Ontology. On the question of properties, some Fortibracchians (like Armstrong himself) embrace universals only, but others adopt tropes instead of universals or a four-category ontology (with both universals and tropes). Given the importance of the immediate inclusion of universals and tropes in particulars for Constituent Ontologists, and given the Fortibracchian commitment to bottom-up grounding, gunk is a problem. The best metaphysical mereology for Fortibracchians is, therefore, Priority Atomism. Since the actual laws of nature, as formulated by scientists, refer to persisting particles, some form of endurantism is best. Why Anti-Tensism? Because science has no need of a metaphysically privileged present. Given the available resources of universals and laws of nature, Combinatorialism would be the best Fortibracchian account of modality.

Quietists (e.g., Chisholm, Plantinga, van Inwagen) believe in properties as abstract objects, but they do not believe that properties *ground* the character of particular things. Instead, they suppose that a particular instantiates a property like *whiteness* because it is white. For this reason, we classify Quietists as Ostrich Nominalists. Quietists tend to

**Table 29.1** Positions of the Four Packages

| | *Ludovician* | *Aristotelian* | *Fortibracchian* | *Quietist* |
|---|---|---|---|---|
| Truth | Truth Supervenes on Being (2.1A.1A.1T) | Classical (2.1T) or Spectral Truthmakers (2.1A.1T.1) | Atomic Truthmaker (2.1T.4) | No truthmaking (2.1A.1A.1A) |
| Undefinable Grounding? | Yes (3.1T.1T) | Yes (3.1T.1T) | Yes (3.1T.1T) | Definable? (3.1T.1A) |
| Powers | Neo-Humeism (4.4T) | Strong Powerism (4.4A.3) | Strong Nomism (4.4A.2) | Hypotheticalism (4.1T) |
| Universals | Extreme Resemblance Nominalism (8.1T.4.1T) | UP Realism (7.1T.1T) | UP Realism (7.1T.1T), or Moderate Nominalism (8.1T.4.1A) | Ostrich Nominalism (7.1A.1A) |
| Particulars | | Constituent Ontology (9.1T) | Constituent Ontology (9.1T) | |
| Relations | Reductive Separatism (10.1A.1A.1A) | Constituent Connectionism (10.1A.1T.1T.1T.1A) | Constituent Connectionism (10.1A.1T.1T.1T.1A) | Ostrich Separatism (10.1A.1A.1T) |
| Non-Existence | Possibilism (12.1A.1T) | Meinongianism (12.1A.1A), or Possibilism (12.1A.1T) | Possibilism (12.1A.1T) | Actualism (12.1T) |
| Vagueness | Multiple Meaning Theory (12.2A.1T) | Ontological Vagueness? (12.2A.1A) | | |
| Perception | Intentionalism (13.3A.1A.1A.2) | Perceptual Dualism (13.3A.1A.1T) or Meinongian Direct Realism (13.3A.1A.1A.1) | Intentionalism (13.3A.1A.1A.2) | Perceptual Dualism (13.3A.1A.1T) |
| Possibility | Concretism (14.1T.1T) | Aristotelian Modality (15.2T.7) | Combinatorialism (15.3T) | Magical Abstractionism (15.1T.1T) |
| Actuality | Modal Indexicalism (14.2A) | Modal Anti-Indexicalism (15.2T) | Modal Anti-Indexicalism (15.2T) | Modal Anti-Indexicalism (15.2T) |
| De Re Modality | Counterpart Theory (16.1A.1) | Transworld Identity (16.1T) | Transworld Identity (16.1T) | Transworld Identity (16.1T) |

(*continued*)

**Table 29.1** (*Continued*)

| | *Ludovician* | *Aristotelian* | *Fortibracchian* | *Quietist* |
|---|---|---|---|---|
| Space and Bodies | Spatial Monism (17.1T.1A.1A) | Aristotelian Relationism (17.1A.1A) | | Body-Space Dualism? (17.1T.1A.1T) |
| Regions and Boundaries | Spatial Pointillism (18.1T) | Volume-Boundary Dualism (18.1A.2A) | | |
| Bodies | Infinitism (18.4A) | Material Anti-Pointillism (18.3A), Finitism (18.4T) | | |
| Instants | Instantism (19.1T) | Moderate Intervalism (19.1A.1A), Temporal Finitism (19.3T) | | |
| Processes | Procedural Instantism (19.2T) | Procedural Intervalism (19.2A) | | |
| Time's Passage | Reductive Anti-Tensism (21.2A.1A) | Tensism (20.2T)? | Reductive Anti-Tensism (21.2A.1A) | Tensism (20.2T)? |
| Future/Past | Eternalism (20.2A.2T) | Indeterminacy of Future Contingents (20.3T) | Eternalism (20.2A.2T) | |
| Compositional Free Lunch? | Compositional Equivalence (22.1A.1T.1) | | Universal Bottom-Up Priority (22.1T.1T) | |
| Heaps | Fundamental Heapism (22.2T) | Anti-Heapism (22.2A) | Anti-Heapism (22.2A) | |
| Point-Parts | Priority of Spatial Point-Parts (22.3T) | Non-Priority of Spatial Point-Parts (22.3A) | | |
| Special Composition Question | Mereological Universalism (22.8T.1A.2) | Homogeneous Continua Plus Organisms (22.8T.1A.10), Organicism (22.6T), Anti-Artifactualism (22.5A) | Priority Atomism (22.7T), Universal Atomism (22.4T) | Organicism (22.6T) |

| | | | | |
|---|---|---|---|---|
| General Composition Question | Composition as Identity (23.1T.1.1) | Aristotelian Compositional Reductionism (23.1A.1T.1A.1) | Arbitrary Compositional Reductionism (23.1A.1T.1T) | |
| Persistence | Classical Perdurantism (24.1T.1T.1A.1T) | Classical Endurantism (24.1T.1T.1T.1T) | Classical Endurantism (24.1T.1T.1T.1T) | Classical Endurantism (24.1T.1T.1T.1T) |
| Constancy | Ramsey-Sider- Lewis Perdurantism (24.3T.3) | Mereological Inconstancy (25.1A) | Mereological Constancy (25.1T) | Mereological Inconstancy (25.1A) |
| Motion | At/At Theory (24.5A.1T) | Motion Intervalism (24.5A.1A) | | |
| Causation | Causal Explanationism (27.1A) | Existence of Uncaused Causes (26.1T.1A.1T) Principled Causation (26.1T.1A.2T), Causal Connectionism (27.1T) | Causal Explanationism (27.1A) | |

**Table 29.2** Principles Appealed to by the Four Packages

| | *Ludovician* | *Aristotelian* | *Fortibracchian* |
|---|---|---|---|
| PMeth 1 Ockham's Razor | 18.1T, 24.1T.1T.1A.1T | 17.1A.1A | |
| PMeth 1.1 | 2.1A.1A.1T, 4.4T | 2.1T, 4.4A.3 | 2.1T.4 |
| PMeth 1.2 | 18.1T, 22.2T | 7.1T.1T, 18.4T, 19.3T, 22.2A | 7.1T.1T |
| PMeth 1.3 | 4.4T | 4.4A.3 | |
| PMeth 1.4 | 8.1T.4.1T, 17.1T.1A.1A | 7.1T.1T | 7.1T.1T, 15.3T |
| PMeth 1.4.1 | 8.1T.4.1T | 7.1T.1T | 7.1T.1T |
| PMeth 2 Scientific Realism | 18.4A, 20.2A.2T | | 20.2A.2T |
| PMeth 2.1 | | 4.4A.3 | 4.4A.2 |
| PMeth 2.2 | | 4.4A.3 | 4.4A.2 |
| PMeth 3 Structuralism | | 18.4T | |
| PMeth 4 Redundancy | | 22.2A | |
| PEpist 1 Imagination Guide to Possibility | 4.4T, 19.1T, 20.2A.2T, 22.1T | 4.4A.3, 18.1A.2A, 19.1A.1A | 4.4A.2, 20.2A.2T |
| PEpist 2 Common Sense | | 20.2T | |
| PEpist 4.1 Sensory Error Minimization | | 13.3A.1T, 13.3A.1A.1T | |
| PTruth 1 One Truthmaker per Fundamental Property | | 7.1T.1T, 18.1A.2A | 7.1T.1T |
| PMeta 1 Test for Grounding | 4.4T | 4.4A.3 | 4.4A.2 |
| PMeta 2 Intrinsic Powers | | 4.4A.3. 19.3T, 20.2T | |
| PMeta 5.2 Patchwork Principle | | 18.4T, 19.1A.1A, 20.2T, 24.1T.1T.1T.1T | |
| PNatPhil 1 Possible Gunk | 22.2T, 22.6T.1A.2 | 22.8T.1A.10 | |
| PNatPhil 2 Motion and Substantial Change Independent | 22.2T | 18.1A.2A | 22.4T |
| PNatPhil 3 Continuous Motion | | 19.3T | 22.4T |
| Correspondence Theory of Truth (2.2) | 2.1A.1A.1T | 2.1T | 2.1T.4 |
| No Metaphysical Cheaters (2.2) | 2.1A.1A.1T | 2.1T | 2.1T.4 |
| Principle of Indifference (5.2.3) | | 4.4A.3 | 4.4A.2 |
| Priority of the Actual (15.1.2) | | 15.2T.6 | 15.2T.6 |
| Acquaintance Model (15.1.2) | 14.1T.1T | 14.1T.1A.1 | 15.3T |
| Physicalism (22.6.2) | 22.2T | | 22.7T |
| Connected Occupation (22.4.2) | 22.1T | | |

**Table 29.3** Special Principles Appealed to by Aristotelians, with Associated Theses

| | |
|---|---|
| Principle of Indifference (5.2.3) | 4.4A.3 |
| Phenomenal Exportation (13.3.1) | 13.3A.1T. 13.3A.1A.1T |
| Pure Physicality of Physical Things (13.3.2) | 13.3A.1T |
| Physics Carves Nature at Joints (13.3.2) | 13.3A.1T |
| Branch (15.3) | 15.2T.7 |
| Infinite Points in Space (18.3.1) | 18.4T |
| Symmetry (18.3.3) | 18.1A.2A |
| Unlimited Divisibility (18.3.3) | 18.1A.2A |
| Maximum Velocity (19.1.2) | 19.3T |
| Impotence of Identity (19.1.2) | 19.3T |
| Spatial Occupancy (22.4.1) | 22.2A |
| Intrinsic Composition (22.5.1) | 22.4A |
| Reasonable Reliability (25.2.6) | 24.1T.1T.1T.1T |
| Cogito ergo sum (22.6.1, 25.1) | 22.6T |
| No radically discontinuous motion (24.3.2) | 24.5A.1A |

believe in the objective truth of many subjunctive conditionals, but they typically do not give a reductive account of such truth in terms of possible worlds or causal powers. Plantinga (1974: 45–48), for example, treats the truth of counterfactual conditionals as an undefinable primitive, and uses such conditionals in his explication of notions like 'true-in-a-world' or 'exists-in-a-world'. Thus, we have classified Quietists as Hypotheticalists, but not as Strong Hypotheticalists, since Quietists don't aspire to offer accounts of powers or laws in terms of counterfactual conditionals. Most Quietists, including both Plantinga and van Inwagen, are ardent Actualists, and they favor Magical Abstractionism with Anti-Indexicalism and Transworld Identity as the best account of modality. We have attributed Organicism to the Quietists, treating van Inwagen's *Material Beings* (1990a) as a paradigm of Quietism. On other issues, we have assigned what we take to be the most common-sensical position, including Tensism, Endurantism, and Body-Space Dualism, but these would seem to be quite optional for the package as a whole.

In the following two tables we record appeal by each of the first three packages to each of our metaphysical principles (as tabulated in Appendix B). We list which specific theses associated with the package an argument for which included the principle as a premise. In this way, we can see clearly the methodological and foundational differences among the three approaches.

All four packages rely, to a considerable extent, on Imagination as Guide to Possibility. This is not surprising, given the central role of thought experiments within analytic philosophy generally. Aristotelians also rely very heavily on Patchwork Principles, as an adjunct to imagination in exploring modal space.

Ludovicians rely on just four sources of information: Ockham's Razor, scientific realism, Imagination as Guide to Possibility, and the principle that Truth Supervenes on Being. Fortibracchians are very similar in methodology to Ludovicains. In practice, Fortibracchians give more weight to a kind of naive realism with respect to scientific theory than do Ludovicians. That is, they tend to take the ontology of science at face value, leading to Classical Endurantism, in contrast to the Ramsey-Lewis-Sider Perdurantism of the Ludovicians.

Quietists are quite different. They rely almost exclusively on *a priori* intuitions and common sense.

Aristotelians represent a hybrid of the two methodologies. They rely strongly on Ockham's Razor and on the correspondence theory of truth, but they also strongly respect common sense and *a priori* intuition, especially intuition about the essences of material things.

Which package is most likely to be true? Obviously, our answer to this question will depend in large part in how much weight we give to the various sources of evidence, which is a matter of good judgment, rather than mere logical acumen.

Of course, the evidence we have surveyed in this book is not the final word. Our ability to ascertain the truth will improve as we understand more fully the implications of the theses in question, and as we gain more insight into the significance of what we already know. We hope that this book will be a stimulant to further metaphysical research. We have attempted to bring the reader as close as possible to the state of the art in metaphysics while presupposing very little prior knowledge of the field.

# Appendix A

# Metaphysical Theses and Antithesis

In this appendix, we have compiled the theses and anti-theses discussed throughout the book. Though the theses are arranged according to the chapters in which they appear, they are not always listed in the order of their appearance in the book. Rather, they are arranged here to display their relative structure: sub-theses (like 2.1T.1T and 2.1T.1A) are therefore listed immediately following their respective theses (2.1T). Each of the entries in this appendix has a corresponding entry in the index, allowing the reader to identify where in the text these theses are discussed

## 2. Truthmakers

2.1T Classical Truthmaker Theory. There are classical truthmakers for all truths, or for some very large sub-classes of truths.

  2.1T.1 Truthmaker Maximalism. Every truth has a strict truthmaker.

  2.1T.3 Totality Fact Maximalism. Every true proposition has a truthmaker, which includes one or more totality facts, possibly together with one or more ordinary existing things. Each universal is associated with at most one totality fact.

  2.1T.4 Atomic Truthmaker Theory. Every atomic (simple, positive) truth has a (classical) truthmaker.

2.1A No Classical Truthmaker. Few, if any, truths have classical truthmakers.

  2.1A.1T Non-Classical Truthmaker Theory. Metaphysical Fundamentalism is true, and propositions are made true by the way things are in the world, but there are no truthmakers.

    2.1A.1T.1 Spectral Truthmaker Theory. Every fundamental atomic truth is made true by something's existing *and being a certain way intrinsically*.

*The Atlas of Reality: A Comprehensive Guide to Metaphysics*, First Edition.
Robert C. Koons and Timothy H. Pickavance.

2.1A.1A No Truthmakers. Truths have no truthmakers, classical or otherwise.

2.1A.1A.1T Truth Supervenes on Being. The property of *being true* weakly supervenes on the property of being a truth about what things exist (and don't exist) and about exactly which natural properties they have and which natural relations they stand in.

2.1A.1A.1A Truth does not supervene on being.

## 3. Grounding, Ontological Dependence, and Fundamentality

3.1T Real Grounding. There is a relation of *metaphysical grounding*.

3.1T.1T Indefinable Grounding. Grounding is a unique and indefinable relation.

3.1T.1T.1 Conceptual Analysis and Indefinable Grounding. Grounding is not definable in terms of conceptual analysis.

3.1T.1T.2 Theoretical Reduction and Indefinable Grounding. Grounding cannot be defined in terms of theoretical reduction.

3.1T.1T.3 Ontological Reduction and Indefinable Grounding. Grounding cannot be defined in terms of ontological reduction.

3.1T.1T.4 Supervenience and Indefinable Grounding. Grounding is not definable in terms of metaphysical supervenience.

3.1T.1A Definable Grounding. The grounding relation is definable in terms of other relations.

3.3T Truthmaker to Grounding Link. The fact that $p$ (wholly) grounds the fact that $q$ if and only if: (i) $p$ & $q$, (ii) the truth of $p$ necessitates the truth of $q$, and (iii) necessarily, if $p$ & $q$, then every truthmaker for $p$ is a truthmaker for $q$.

3.3T.1 Generalized Truthmaker to Grounding Link. The facts corresponding to set $\Gamma$ of propositions (wholly) grounds the fact that $q$ if and only if (i) $q$ and every member of $\Gamma$ is true, (ii) the joint truth of the members of $\Gamma$ metaphysically necessitates the truth of $q$, and (iii) necessarily, if both $q$ and every member of $\Gamma$ are true, then, for every $p$ in $\Gamma$, every truthmaker for $p$ is contained in (is a part of) a truthmaker for $q$.

3.4T O-Fundamentality Entails G-Fundamentality. Necessarily, all O-fundamental entities are also G-fundamental.

3.4T.1 Formality of Grounding. Whenever $p$ grounds $q$, there exist forms $\Phi$ and $\Psi$ such that (i) for some $x$, $p = \Phi(x)$ and $q = \Psi(x)$, and (ii) for all true propositions $r$ and $s$ such that for some $y$, $r$ is $\Phi(y)$ and $s$ is $\Psi(y)$, the fact that $r$ grounds the fact that $s$.

3.4T.2 Mediation of Grounding by Essence. Every grounding fact of form (ii) (in Formality) is itself grounded in or derivable from some *essential truths* about the constituents of forms $\Phi$ and $\Psi$.

3.5T.1 Grounding-Essence Correspondence. The fact that $p$ grounds the fact that $q$ if and only if (i) $p$ & $q$, and (ii) some generalization of the form *For all x, if p(x), then q(x)* follows logically from the essences of $p$ and $q$, together with the essences of things that the facts $p$ and $q$ contain.

3.5T.2 Plural Grounding-Essence Correspondence. The facts that $p_1$, that $p_2$,..., and that $p_n$ jointly ground the fact that $q$ if and only if (i) $p_1$ & $p_2$ &... & $p_n$ & $q$; (ii) some generalization of the form *For all x, if* $p_1(x)$ & $p_2(x)$ &.... & $p_n(x)$, *then q(x)* follows logically from the essences of $p_1, p_2,..., p_n$, and $q$, together with the essences of things that the facts $p$ and $q$ contain; and (iii) if n = 1, i.e., there is only one grounding fact, then $p_1 \neq q$.

3.6T Conceptual-Grounding to Essence Link. The truth of the proposition $p$ grounds the truth of the proposition $q$ only if $p$ and $q$ are both true, and *if p, then q follows* logically from the essence of the conceptual and logical elements of the propositions $p$ and $q$.

3.7T Distinct Fundamental Properties. $F$-ness and $G$-ness are distinct fundamental properties if and only if it is necessarily the case that, for any $x$, a truthmaker for the proposition *x is F* is distinct from any truthmaker for the proposition *x is G*.

3.8T Grounding-Entailment Entailment. Necessarily, if the fact that $p$ grounds the fact that $q$, then $p$ metaphysically entails $q$ (i.e., necessarily if $p$ is true, then $q$ is true).

## 4. Conditionals

4.1T Hypotheticalism. There are fundamentally conditional or hypothetical truths.

4.1A Anti-Hypotheticalism. All conditional truths are grounded in non-conditional categorical facts.

4.2T Nomic Fundamentalism. Some truths about the laws of nature are fundamental.

4.2A Nomic Reductionism. No truths about the laws of nature are fundamental.

4.3T Power Fundamentalism. Some truths about the powers of particular things are fundamental.

4.3A Power Reductionism. No truths about powers are fundamental.

4.4T Neo-Humeism. None of the truths of counterfactual conditionals and none of the truths about laws of nature or about the powers of particulars are fundamental.

4.4A Anti-Humeism. Some truths about counterfactual conditionals, laws of nature, or powers of particulars are fundamental.

4.4A.1 Strong Hypotheticalism. Some of the truths of counterfactual conditionals are fundamental, but no truths about particular powers or the laws of nature are fundamental.

4.4A.2 Strong Nomism. Some of the truths about laws of nature are fundamental, but no truths about particular powers nor any of the truths of counterfactual conditionals are fundamental.

4.4A.3 Strong Powerism. Some of the truths about the powers of particular things are fundamental, but no truths about the laws of nature nor any of the truths of counterfactual conditionals are fundamental.

## 6. Powers

6.1T Causal Individuation of Properties (Weak Thesis). Some fundamental properties are individuated by the set of powers they confer.

6.1T.1T Strong Causal Individuation of Properties (Strong Thesis). All fundamental properties are individuated by the set of powers they confer.

6.1A Sicceity Theory. No fundamental properties are individuated by the set of powers they confer.

6.2T Essentiality of Powers (Weak Thesis). Some fundamental property has its causal profile of necessity.

6.2T.1T Strong Essentiality of Powers. All fundamental properties have all of their causal profile of necessity: that is, each power is either necessarily conferred or necessarily not conferred by each property.

6.2A Strong Accidentality of Powers. No fundamental property has any of its causal profile of necessity: that is, no property confers or fails to confer any power necessarily.

## 7. Universals

7.1T Realism. Universals exist and ground the character of ordinary objects.

7.1T.1T UP-Realism. There are fundamental universals and fundamental particulars, and the latter instantiate the former.

7.1A Nominalism. Universals do not ground the character of ordinary objects.

7.1A.1A Ostrich Nominalism. Nominalism is true, and there is no general explanation of the fact that some particulars have properties in common.

7.1A.1T Reductive Nominalism. Nominalism is true, and there is a general explanation of the fact that some particulars have properties in common. (=8.1T)

7.2T Universal-Particular Distinction. There is a clear and coherent distinction between universals and particulars.

7.2T.1 Aristotelian UP. Universals can be *predicated* of other things, while particulars cannot be predicated of anything.

7.2T.2 Russellian UP. Particulars are necessarily located in only *one place* at a time, while universals can be wholly present at *many places* at once.

7.2T.3 Wise UP. Any particular can coexist with another particular indistinguishable from it, but it is impossible for any universal to be indistinguishable from anything else.

7.3T Identity of Indiscernibles. If (thing *a* exemplifies *Fness* if and only if thing *b* exemplifies *Fness*), then *a* is identical to *b*.

7.3T.1 Restricted Identity of Indiscernibles. If (thing *a* exemplifies *Fness* if and only if thing *b* exemplifies *Fness*), where *Fness* is a natural, monadic property, then *a* is identical to *b*.

## 8. Nominalism

8.1T Reductive Nominalism. Nominalism is true, and there is a general explanation of the fact that some particulars have properties in common. (=7.1A.1T)

8.1T.1 Predicate Nominalism. Reductive Nominalism is true, and whenever two particulars have a property in common, this fact is grounded in the fundamental fact that the two fall under some one predicate of some language.

8.1T.2 Concept Nominalism. Reductive Nominalism is true, and whenever two particulars have a property in common, this fact is grounded in the fundamental fact that the two fall under some one concept.

8.1T.3 Class Nominalism. Reductive Nominalism is true, and whenever two particulars have a property in common, this fact is grounded in the fundamental fact that the two belong to some one set or class.

8.1T.4 Resemblance Nominalism. Reductive Nominalism is true, and whenever two particulars have a property in common, this fact is grounded in fundamental facts of resemblance between them.

8.1T.4.1T Extreme Resemblance Nominalism. There are only ordinary particulars, and whenever two ordinary particulars resemble each other, their resemblance is metaphysically fundamental.

8.1T.4.1A Moderate Resemblance Nominalism (Trope Nominalism). There are only particulars, and whenever two ordinary particulars resemble each other, their resemblance is grounded in the fundamental fact that the two are characterized by tropes whose resemblance is metaphysically fundamental.

8.2T Trope Theory. Tropes ground the character of ordinary objects.

8.2T.1T Modifier Trope Theory. Trope Theory is true, and tropes are modifying tropes.

8.2T.1A Modular Trope Theory. Trope Theory is true, and tropes are modular tropes.

8.2A Anti-Tropism. Tropes do not ground the character of ordinary objects.

8.2A.1 Extreme Nominalism. The character of ordinary objects is grounded neither in universals nor in tropes.

8.2A.2 Classical UP-Realism. The character of ordinary objects is grounded in universals and not in tropes.

## 9. Particulars and the Problem of Individuation

9.1T Constituent Ontology. When a substance instantiates a property, the instantiation relation between the two consists in the fact that the property is a part of the substance.

9.1T.1T Bundle Theory. The only constituents of each substance are its characterizing properties.

9.1T.1T.1T Trope Bundle Theory. Bundle Theory is true and tropes ground character.

9.1T.1T.1A Classical Bundle Theory. Bundle Theory is true and universals ground character.

9.1T.1T.1A.1 Constructive Bundle Theory. Substances correspond one-to-one to sets of universals.

9.1T.1T.1A.2 Simple Bundle Theory. Substances correspond one-to-one to co-instantiated sets of universals; that is, a substance S exists if and only some set of universals U is co-instantiated, and S exemplifies a universal F if and only if F is a member of U.

9.1T.1T.1A.3 Nuclear Bundle Theory. Substance correspond one-to-one to nuclei of universals. That is, a substance S exists if and only if some set of universals U has the property of being nuclearly co-instantiated, and S exemplifies a universal F if and only if F belongs to some peripherally co-instantiated set (a bundle) that contains U.

9.1T.1A Substrate Theory. Each substance has a constituent other than its characterizing properties, a substrate.

9.1T.1A.1T Trope Substrate Theory. Substrate Theory is true and tropes ground character.

9.1T.1A.1A Classical Substrate Theory. Substrate Theory is true and universals ground character.

9.1T.1A.2T Modular Substance Theory. Substrate Theory is true, and substrates have one property in themselves.

9.1T.1A.2A Bare Particular Theory. Substrate Theory is true, and substrates have no properties in themselves.

9.1A Relational Ontology. Instantiation is a fundamental relation between substances and properties, and instantiation is not a case of the *part-whole* relation.

9.1A.1T Classical Relational Realism. Relational Ontology and Extreme UP-Realism are both true.

9.1A.1A Trope Relational Ontology. Relational Ontology and Trope Theory are both true.

9.2T Primitive Identity. All pairs of substances are self-individuating.

9.2A Non-primitive Identity. Some pairs of substances are not self-individuating.

9.2A.1 Fundamental Relations of Distinctness. For any distinct substances *x* and *y*, the truthmaker for *x*'s distinctness from *y* is a fundamental distinctness nexus between *x* and *y*.

9.2A.2 Scotism. Substances have haecceities, and haecceities are natural properties.

## 10. Relations, Structures, and Quantities

10.1T Monadism. There are no fundamental relational truths.

10.1A Anti-Monadism. There are some fundamental relational truths.

10.1A.1T Connectionism. Some non-symmetric relational universals or non-symmetric relational tropes are fundamental entities.

10.1A.1T.1T Classical Connectionism. There are relational universals that are fundamental.

10.1A.1T.1T.1T Relational (Nexus) Connectionism. When some particulars (in a certain order) instantiate some relational universal, the truthmaker for the corresponding proposition is a relational *nexus* that ties those things (in that order) to the universal. The nexus does not contain either the relation or the relata as parts.

10.1A.1T.1T.1T.1A Constituent (State of Affairs) Connectionism. When some particulars (in a certain order) instantiate some relational universal, the truthmaker for the corresponding proposition is a state of affairs that contains those things and that universal as proper parts.

10.1A.1T.1A Trope Connectionism. There are relational tropes that are fundamental, but no fundamental relational universals.

10.1A.1A Separatism. There are no fundamental non-symmetric relational universals or relational tropes.

10.1A.1A.1T Ostrich Separatism. Separatism is true, and there is no general explanation of why some things taken in a certain order resemble other things taken in a certain order.

10.1A.1A.1A Reductive Separatism. Separatism is true, and there is some general explanation of why some things taken in a certain order resemble other things taken in a certain order.

10.2T Real numbers are universals.

10.2T.1 Determinate Universal Theory. Real numbers are relational universals, each instantiated by ordered pairs of determinate quantities.

10.2T.2 Simple Intensity Theory (Nexus version). Positive real numbers are relations between intensities, and intensities are universals that are jointly connected to particulars and determinable universals.

10.2T.3 Composite Intensity Theory. Positive real numbers are intensities. Each positive real number contains all smaller positive real numbers as proper parts, and the having of a determinate property consists in the joint instantiation of a determinable universal and an intensity.

10.2A Real numbers are not universals.

10.2A.1T Real numbers are tropes.

10.2A.1T.1 Determinate Trope Theory. Real numbers correspond to natural resemblance classes of relational modifying tropes, each of which modifies an ordered pair of quantities.

10.2A.1T.2 Simple Intensity Theory (Trope-nexus version). Positive real numbers are binary relations between intensity-tropes, and each intensity-trope is jointly attached to both a substrate and a trope of some determinable.

## 11. Nihilism and Monism

11.1T Aliquidism. Something exists.

11.1A Nihilism. Nothing exists.

11.2T Pluralism. More than one thing exists.

11.2T.1 Fundamental Atomism. Necessarily, every fundamental entity is atomic (lacking parts).

11.2A Monism. No more than one thing exists.

## 12. The Non-Existent and the Vaguely Existent

12.1T Actualism. Everything (actually) exists.

12.1A Anti-Actualism. Some things don't (actually) exist.

12.1A.1T Possibilism. Everything could exist, and some things could exist that don't actually exist.

12.1A.1A Meinongianism. Some things couldn't exist.

12.2T Epistemicism. Every boundary is determinate. Hence, any vagueness is merely epistemic.

12.2A Anti-Epistemicism. Some boundaries are indeterminate, resulting in failures of the law of bivalence.

12.2A.1T Ambiguity Theory of Vagueness. Some boundaries are indeterminate, but all such indeterminacy is merely a matter of ambiguous reference to fully determinate entities.

12.2A.1A Real Ontological Vagueness. Some entities lack determinate boundaries, independently of our knowledge or how we refer to them.

## 13. Solipsism, Idealism and the Problem of Perception

13.1T Idealism. Every fundamental particular is wholly mental.

13.1T.1 Solipsism. Every fundamental particular is wholly mental and part (or attribute or event intrinsic to the life of) of a single subject or mind.

13.1A Anti-idealism. Not every fundamental particular is wholly mental.

13.1A.1 Inferred Anti-Idealism. Non-mental things exist but cannot be perceived (either directly or indirectly).

13.2T Phenomenalism. Everything we are familiar with or that we have knowledge of, including *apparently* physical objects, is in reality wholly composed of and wholly grounded in wholly mental things.

13.2T.1 Phenomenalism (Solipsistic Version). For each subject of experience, the world as that subject knows it consists only in that very subject and its wholly mental parts and properties.

13.3T Veil of Perception. Wholly mental things are the only possible objects of sense perception.

13.3T.1 Veil of Perception (Solipsistic Version). The only possible objects of sense perception for each subject are the wholly mental and properties parts of that very subject.

13.3A Perceptual Realism. It is possible to perceive something other than wholly mental things.

13.3A.1T Indirect Realism. It is possible to perceive something other than wholly mental things, but only indirectly, *by virtue of* perceiving wholly mental things.

13.3A.1A Direct Realism. It is possible to directly perceive things other than wholly mental things (i.e., not by virtue of perceiving wholly mental things).

13.3A.1A.1T Perceptual Dualism. It is possible to directly perceive non-mental (physical) objects, as in veridical perception. It is also possible to be in states (hallucinatory states) in which one directly perceives a wholly mental object but perceives no physical object, and some hallucinatory states are introspectively indistinguishable from some veridical perceptions.

13.3A.1A.1A Unitary Direct Realism. Cases in which existing physical things are directly perceived are introspectively indistinguishable from cases in which no existing thing is directly perceived.

13.3A.1A.1A.1 Meinongian Direct Realism. It is possible to directly perceive both existing and non-existing physical objects, and some cases of the former are introspectively indistinguishable from the latter.

13.3A.1A.1A.2 Intentionalism. It is possible to directly perceive existing physical objects, and there are possible states introspectively indistinguishable from these in which one perceives nothing whatsoever (whether mental or physical, existing or non-existing).

## 14. Actuality, Possibility, and Necessity

14.1T Possible Worlds. There are possible worlds.

14.1T.1T Concretism. Possible worlds are maximal concrete objects.

14.1T.1A Abstractionism. Possible worlds are maximal possible abstract objects. (=15.1T)

14.1A No Possible Worlds. There are no possible worlds.

14.2T Modal Indexicalism. All attributions of actuality are indexical in character.

14.2A Modal Anti-Indexicalism. Some attributions of actuality are not indexical. (=15.2T)

## 15. Abstractionism: Worlds as Representations

15.1T Abstractionism. Possible worlds are maximal possible abstract objects. (=14.1T.1A)

15.1T.1T Magical Abstractionism. A world *w* represents that *p* if and only if it is necessarily true that if *w* were actual, then *p,* and there is no substantive account to give about why possible worlds represent what they do.

15.1T.1A Structural Abstractionism. Possible worlds represent what they do in virtue of their internal structure.

15.1T.1A.1 Linguistic Abstractionism. Structural Abstractionism is true, and worlds represent in the way linguistic things do, namely, by having components which represent things and which are arranged in according to a kind of "grammar".

15.1T.1A.2 Pictorial Abstractionism. Structural Abstractionism is true, and worlds represent in the way pictures do, namely, by having components that literally share features with what they represent.

15.2T Modal Anti-Indexicalism. Some attributions of actuality are not indexical. (=14.2A)

15.2T.1 Simple Anti-Indexicalism. *Actuality* is a simple, fundamental property of possible worlds.

15.2T.2 Actual-Truth-Defined Anti-Indexicalism. A possible world is actual if and only if it corresponds to the class of (actually) true propositions.

15.2T.3 Truth-Defined Anti-Indexicalism. A possible world is actual if and only if it corresponds to the class of true (simpliciter) propositions.

15.2T.4 Existence-Defined Anti-Indexicalism. The actual world is that unique possible world that is reality-bounded and not quantificationally deficient.

15.2T.5 Existence-Simpliciter-Defined Anti-Indexicalism. The actual world as that unique world *w* such that, for every *x*, *x* exists according to *w* iff *x* exists simpliciter.

15.2T.6 Fundamental-Truth-Defined Anti-Indexicalism. The actual world is that unique world $w$ such that every fundamental truth is true in $w$.

15.2T.7 Aristotelian Modality.

(i) It (tenselessly) is possible for $p$ to be the case if and only if there is, was or will be a time at which it is, was or will be possible for $p$ to be the case;
(ii) it is now *simply* possible for something to exist or to fail to exist iff something now has the power to cause or to prevent its coming into existence;
(iii) it is now *simply* possible for some things to stand in some natural relation iff something now has the power to make those things to stand in that relation;
(iv) it is now *possible* for $p$ to be the case iff either it is now simply possible for $p$ to be the case, or it is simply possible for it to be simply possible for $p$ to be the case, or it is simply possible that it is simply possible that it is simply possible that $p$, etc.

15.3T Combinatorialism. Every Tractarian world corresponds to a possible world.
15.3A Anti-Combinatorialism. Some Tractarian worlds correspond to no possible world.

## 16. De Re Modality and Modal Knowledge

16.1T Transworld Identity. Some things really exist (exist in a fundamental way) in more than one possible world.

16.1T.1 Overlapping Concrete Worlds. It is possible for something to fail to exist if and only if there is an concrete universe in which *it* (and not just a counterpart of it) does not exist.

16.1T.2 Strong Transworld Identity. There is an object $x$ and distinct worlds $w_1$ and $w_2$ such that both $w_1$ and $w_2$ represent $x$ in part because each is ontologically dependent on $x$.

16.1A Worldbound Individuals. Each possible individual exists, as a matter of necessity, in one and only one possible world.

16.1A.1 Ludovician Counterpart Theory. It is possible for something to fail to exist if and only if there is an isolated universe in which it has no counterpart, it is possible for something to have a property if and only if there is an isolated universe it which a counterpart of it has that property, and it is possible for some things to stand in some natural relation if and only if they have some counterparts that do stand in that relation in some isolated universe.

16.2T Conceivability Entails Possibility. Every conceivable scenario is true in some possible world.

16.2T.1 Lack of A Priori Falsity Entails Possibility. Every negative-epistemic conceivable scenario is true in some possible world.

16.3T Necessity Entails A Priori Knowability. If $p$ is necessarily true, then we can know a priori that $p$ is true in fact.

16.4T A Priori Knowability Entails Necessity. If it is knowable a priori that $p$ is false, then $p$ is true in no possible world.

## 17. Is Space Merely Relational?

17.1T Spatial Substantivalism. Places exist and are G-fundamental.

17.1T.1T Theory of Spatial Qualities. Places are fundamental properties or qualities, and location is predication: a place is predicated of the things located there.

17.1T.1A Spatial Particularism. Places are ordinary particulars (not properties or qualities), and location is an external relation between fundamental particulars (between locations and the occupiers of those locations).

17.1T.1A.1T Body-Space Dualism. Both bodies and places are fundamental particulars of different kinds, with location being an external relation between bodies and places.

17.1T.1A.1A Spatial Monism. The only concrete particulars are places (parts of space). A body is simply a special kind of place—one that is characterized by a quality of *being massive*, *body-ish*, or *en-mattered*.

17.1T.2T Spatial Externalism. Spatial distance and contiguity are external relations between places.

17.1T.2A Spatial Internalism. Spatial distance and contiguity are internal relations between places.

17.1A Spatial Relationism. Places are not G-fundamental.

17.1A.1T Modern Relationism. Spatial Relationism is true, and shapes and volumes are properties of material bodies that consist entirely in the holding of certain distances between point-sized parts of those bodies.

17.1A.1A Aristotelian Relationism. Spatial Relationism is true, and shapes, volumes and contiguity (or contact) are metaphysically fundamental properties of material bodies.

17.2T Fundamentality of Distance. The fundamental spatial property is distance between points.

17.2A Fundamentality of Shape and Contiguity. The fundamental spatial properties are shape, volume, and contiguity between extended things.

17.3T There are real absences.

17.3A There are no real absences.

## 18. The Structure of Space: Points vs. Regions

18.1T Spatial Pointillism (Extreme Indivisibilism). Indivisible, dimensionless parts of space (points) are more fundamental than extended regions; extended regions can be wholly grounded in points.

18.1A Spatial Anti-Pointillism. Finite regions or volumes are at least as G-fundamental as points.

18.1A.1T Fundamentally Gunky Space. All G-fundamental volumes of space have proper parts.

18.1A.1A Discrete Space. There are simple volumes.

18.1A.2T Voluminism. Entities of fewer than three dimensions (like boundaries that are points, curves, and surfaces) are only logical constructions from regions and extended entities.

18.1A.2A Volume-Boundary Dualism. Entities of fewer than three dimensions (like boundaries that are points, curves, and surfaces) exist and are not mere logical constructions, but they are not *more fundamental* than extended things.

18.1A.2A.1 Coincident Boundaries. There are spatially coincident points (curves, surfaces).

18.2T Fundamental Entity Space-Matter Correspondence. The fundamental spatial entities are points (as opposed to volumes) if and only if the fundamental occupiers of space are point-sized bodies (as opposed to voluminous bodies).

18.2A No Correspondence between Space-Matter Fundamentality.

18.2A.1 Extended Material Atoms in a Pointillist Space. There could be extended material atoms that occupy fundamental spatial points without point-sized parts occupying those points.

18.2A.2 Material Simples without Proper Location. There could be an extended body composed of infinitely many indivisible, zero-dimensional, volume-less material bodies, each without a unique and unshared fundamental location. That is, each indivisible body occupies infinitely many spatial regions, each of which is also occupied by infinitely many other bodies.

18.3T Material Pointillism. The only fundamental bodies are point-sized (dimensionless). Truths about extended bodies are wholly grounded conceptually in truths about point-sized bodies.

18.3A Material Anti-Pointillism. Necessarily, if there are any extended bodies (bodies with finite volume), then there are fundamental bodies with finite volume.

18.4T Finitism. There are (with the possible exception of sets and numbers) only finitely many actually existing things.

18.4A Infinitism. There are infinitely many actually existing things, other than sets and numbers.

18.5T Spatial Aristotelianism. Every region of space is Aristotelian.

18.5A Continuous Variation. Some quality or quantity varies continuously throughout some region of space.

## 19. The Structure of Time

19.1T Instantism. Temporal intervals are not G-fundamental entities: temporal intervals are wholly grounded in durationless instants.

19.1A Intervalism. There are extended temporal intervals that are G-fundamental.

19.1A.1T Strong Intervalism. Instants either don't exist at all or are derived entities—mere logical constructions from finite intervals.

19.1A.1A Interval-Instant Dualism (Moderate Intervalism). Both instants and intervals are fundamental entities.

19.1A.1A.1 Instants as Dependent Entities. Instants exist only when they are the actual boundaries of extended processes or events.

19.2T Procedural Instantism. No temporally extended process is fundamental: only their indivisible, dimensionless parts (time-slices) are fundamental.

19.2A Procedural Intervalism. Some extended processes are metaphysically fundamental.

19.3T Temporal Finitism. It is impossible for any temporal interval to have infinitely many actual temporal parts.
19.3A Temporal Infinitism. It is possible for temporal intervals to have infinitely many actual temporal parts.
19.4T Temporal Anti-Discretism (Unlimited Divisibility). All extended occupiers of time have proper parts that are temporally extended.
19.4A Temporal Discretism. There are extended occupiers of time without proper parts that are temporally extended.
19.5T Beginning of Time. Time necessarily has a beginning.
19.5T.1 Existence of a First Temporal Part. Time necessarily has a single part that is earliest.
19.5T.2 Finitude of the Past. The past is necessarily finite in extent.

## 20. Time's Passage

20.1T Temporalism. There are moments of time.
20.1T.1T Temporal Pluralism. There are, have been, and will be more moments of time than one.
20.1T.1A Solipsism of the Present Moment. There are, have been and will be only one moment of time (the present).
20.1A Atemporalism. There are no moments of time.
20.2T Tensism (A-Theory). Some tensed truths are metaphysically fundamental.
20.2T.1 Simple Tensism (Moving Spotlight Theory). There is a simple, primitive property of *presentness* that is uniquely and temporarily possessed by a moment of time.
20.2T.2 Falling Branches (Changing Possibility) Tensism. The present moment is the earliest moment such that all later times involve at least one contingent fact.
20.2T.3 Aristotelian (Changing Actuality) Tensism. The present moment is uniquely actual.
20.2T.4 Presentism (Existential Tensism). The present moment is the unique moment such that everything exists then.
20.2T.4.1 Possibilistic Presentism. The present moment is the unique moment at which everything that actually exists actually exists then.
20.2T.5 Serious Tensism. Only tensed truths are metaphysically fundamental.
20.2A Anti-Tensism (B-Theory). There are no metaphysically fundamental tensed truths.
21.2A.1T Eliminative Anti-Tensism. There are no tensed truths.
21.2A.1A Reductive Anti-Tensism. There are tensed truths, but all tensed truths are wholly grounded in tensed truths.
20.2A.2T Eternalism. Everything that exists in the past, present, or future exists simpliciter.
20.3T Indeterminacy of Future Contingents. Some future contingent B-propositions lack any truth-value.
20.3A Determinacy of Future Contingents. All future contingent B-propositions are either true or false.

## 22. Material Composition: The Special Question

22.1T Universal Compositional Priority. Whenever some parts compose a whole, either the parts are wholly grounded in the whole, or the whole is wholly grounded in the parts.

22.1T.1T Universal Bottom-Up Priority. Whenever some parts compose a whole, the whole is wholly grounded in the parts.

22.1A No Universal Compositional Priority. There are cases of composition in which neither the whole is wholly grounded in the parts, nor are the parts wholly grounded in the whole.

22.1A.1T No Compositional Priority. In cases of composition, the whole is never wholly grounded in its parts, nor are the parts wholly grounded in the whole.

22.1A.1T.1 Compositional Equivalence. In cases of composition, truths about the whole and truths about the parts are metaphysically equivalent.

22.1A.1T.1.1 Nihilistically Grounded Composition. All truths about wholes and their parts are wholly grounded in a class of nihilistic truths (truths that do not entail the existence of anything).

22.1A.1T.1.2 Atomistically Grounded Composition. All truths about wholes and their parts are wholly grounded in a class of truths about atomic things that never compose anything.

22.2T Fundamental Heapism. Some heaps exist fundamentally.

22.2A Anti-Heapism. No heaps exist fundamentally.

22.3T Priority of Spatial Point-Parts. If $x$ is a composite thing, then the location of $x$ is wholly grounded in the point-locations of point-sized parts of $x$.

22.3T.1 Priority of Spatial Point-Parts for Heaps. If $x$ is a heap, then the location of $x$ is wholly grounded in the locations of some proper parts of $x$.

22.3T.2 Priority of Spatial Point-Parts for Thorough Heaps. If $x$ is a thorough heap, then the location of $x$ is wholly grounded in the locations of its point-sized parts.

22.3A Non-Priority of Spatial Point-Parts. If $x$ is a composite thing, then the location of $x$ is *not* wholly grounded in the location of $x$'s point-sized parts.

22.4T Universal Atomism. Everything has an atomic part.

22.4A Existence of Mereological Gunk. Some things have no atomic part.

22.4A.1 Existence of Spatial Gunk. Something is spatially gunky.

22.5T Fundamental Artifactualism. Some material mere artifacts exist fundamentally.

22.5A Anti-Artifactualism. No material mere artifacts exist fundamentally.

22.6T Organicism. Some living organisms exist and are both fundamental things and composite material things.

22.6T.1 Composite Persons. Living organisms that are persons capable of free will exist and are fundamental composite, material things.

22.6A Anti-Organicism. Living organisms do not exist or are not fundamental composite, material things.

22.7T Priority Atomism. There are no fundamental composite material things. All fundamental material things are atoms.

22.8T Intelligible Composition. There is a finitely representable, non-trivial set of necessary and sufficient conditions $C$, such that for any set of simples $S$, the members of $S$ compose something if and only if they jointly satisfy $C$.

22.8T.1T Mereological Nihilism. No non-empty set of things ever composes anything (whether fundamental or not).

22.8T.1A Mereological Aliquidism. Some non-empty sets of things do compose something: there is some other principle of composition besides Mereological Nihilism.

22.8T.1A.1 Extreme Organicism. A set of entities composes something if and only if the members of that set participate in a single life.

22.8T.1A.2 Mereological Universalism. For every non-empty set *S*, the members of that set compose something. (Equivalent to (M6) Arbitrary Sums)

22.8T.1A.3 Fastenation. The *x*'s compose something if and only if they are fastened to each other.

22.8T.1A.4 Locomotive Unification. The *x*'s compose something if and only if they are locomotively unified.

22.8T.1A.5 Cohesion. The *x*'s compose something if and only if they cohere together.

22.8T.1A.6 Fusion. The *x*'s compose something if and only if they are fused together.

22.8T.1A.7 Democritean Fastenation. The *x*'s compose something if and only if they are all atoms and the *x*'s are maximally fastened.

22.8T.1A.8 Serial Bonding. The *x*'s compose something if and only if the *x*'s are atoms bonded in ways $R_1$ or $R_1$' (or $R_1$", etc.), or the *x*'s are things composed of atoms bonded in way $R_2$ or $R_2$' or…, or the *x*'s are things composed of things composed of atoms bonded in way $R_3$, or…

22.8T.1A.9 Homogeneous Continua. The *x*'s compose something if and only if they form a homogeneous material continuum (a material thing with no sharp internal boundaries).

22.8T.1A.10 Homogeneous Continua Plus Organisms. The *x's* compose something if and only if they form an organism or a homogeneous material continuum.

22.8T.1A.11 Composition of Artifacts (1). Some material things compose a mere material artifact if and only if they have been altered or arranged by some intelligent agent (or group of agents) for a single purpose or interdependent set of purposes, and they do not compose a living thing.

22.8T.1A.12 Composition of Artifacts (2). Some material things compose a mere material artifact if and only if they are used and maintained by some intelligent agent (or group of agents) for a single purpose or interdependent set of purposes, and they do not compose a living thing.

22.8A Brutal Composition. There is no such finite set of conditions.

22.9T Extreme Anti-Heapism. No heaps exist.

22.10T Extreme Anti-Artifactualism. No composite material artifacts exist.

## 23. Material Composition: The General Question

23.1T Compositional Realism. The *parthood* relation, or some closely related relation, like *joint-composition*, is either fundamental or corresponds to a relational universal.

23.1T.1 Fundamental Parthood. The *parthood* relation, or some closely related relation, like *joint-composition*, is fundamental.

23.1T.1.1 Composition as Identity (CAI). Every whole is identical to its parts: to be a proper part is to be one of the things that are (collectively) identical to the whole.

23.1T.2 Natural Parthood. The *parthood* relation, or some closely related relation, like *joint-composition*, corresponds to a relational universal or a natural resemblance class.

23.1A Compositional Anti-Realism. The *parthood* relation is neither fundamental nor natural. (=22.8T.1T Mereological Nihilism)

23.1A.1T Moderate (or Reductive) Compositional Anti-Realism. There are truths about proper parthood, but all truths about *proper parthood* are grounded in more fundamental, non-mereological truths.

23.1A.1T.1T Arbitrary Compositional Reductionism. All truths about *proper parthood* are grounded in more fundamental, non-mereological truths in such a way that Arbitrary Sums is true.

23.1A.1T.1A Non-Arbitrary Compositional Reductionism. All truths about *proper parthood* are grounded in more fundamental, non-mereological truths in such a way that Arbitrary Sums is false.

23.1A.1T.1A.1 Aristotelian Compositional Reductionism. Some things, the *x*'s, compose *y* at time *t* if and only if the *x*'s are more than one, and there is some process *P* involving exactly the *x*'s as participants such that (i) the existence of *y* at time *t* is causally explained by the continuation of *P* during some *I* that is terminated by *t*, (ii) all of the power-conferring properties of each of the *x*'s are at *t* wholly grounded in the power-conferring properties of *y* at *t*, and (iii) every property of *y* that confers an active or passive power on *y* at *t* ontologically depends on one or more of the *x*'s.

23.1A.1A Compositional Nihilism. Nothing is a proper part of anything else.

## 24. Change and Persistence

24.1T Something changes.

24.1T.T Strict Kineticism. Some things change intrinsically.

24.1T.1T.1T Enduring Substratism. Fundamental things can undergo qualitative change, existing before, during, and after the change.

24.1T.1T.1T.1T Classical Endurantism. Fundamental things can undergo qualitative change, existing before, during, and after the change, but they do not have any temporal parts.

24.1T.1T.1T.1A Time-Slice Substratism. Fundamental things can undergo qualitative change, existing before, during and after the change, and they have instantaneous temporal parts corresponding to each instant during which they exist.

24.1T.1T.1A Replacementism. It is possible for something fundamental to undergo substantial change but not qualitative change.

24.1T.1T.1A.1T Classical Perdurantism. The only fundamental things exist only for an instant, and the only fundamental relations between such time-slices are spatiotemporal relations.

24.1T.1T.1A.1A Classical Genidentity Theory. The only fundamental things exist only for an instant, and there is one fundamental relation (primitive *genidentity*) that is not spatiotemporal.

24.1T.1A Moderate Staticism. Some things undergo mere Cambridge change, but nothing changes intrinsically.

24.1A Nothing changes.

24.2T Persistence. Something persists (has existed or will exist at more than one instant).
24.3T Universal Instantaneous Temporal Parts (UITP). All persisting things have instantaneous temporal parts (time-slices), one corresponding to each instant during which they exist, and these time-slices are fundamental and not mere logical constructions.

24.3T.1 Temporal Plenitude. For every set *S* of instantaneous temporal parts of things, containing exactly instantaneous object for each instant in some finite interval *T*, there is a persisting thing that persists throughout *T* with exactly the members of *S* as its instantaneous temporal parts.

24.3T.2 Continuous Arbitrary Worms. For every spacetime worm *S* over interval *T*, if the spatial locations of the members of *S* form a continuous trajectory through spacetime, there is a persisting thing that persists throughout *T* with exactly the members of *S* as its instantaneous temporal parts.

24.3T.3 Ramsey-Lewis-Sider Perdurantism. A spacetime worm *S* over interval *T* corresponds to the existence of a derived thing persisting through *T* if and only if the simplest and most powerful scientific theory of the actual world assigns a persisting entity to *S*. (=25.3T)

24.3A Absence of Temporal Parts. Some persisting things *x* and some instants of time *t* are such that *x* exists at *t* but has no instantaneous part at *t* that is fundamental and not a mere logical construction.
24.4T No Temporal Coincidence. Necessarily, no instantaneous thing is a time-slice of more than one persisting thing.
24.4A Temporal Coincidence. It is possible for an instantaneous thing to be a time-slice of two distinct persisting things.
24.5T Intrinsic Motion. Motion is something intrinsic to the moving thing at each moment of its motion.
24.5A Extrinsic Motion. Motion is not intrinsic to each instantaneous time-slice of the moving thing.

24.5A.1T The At/At Theory of Motion. The fundamental truths about locomotion are truths about instantaneous location events.

24.5A.1A Motion Intervalism. The fundamental truths about locomotion are truths not about location events but about extended processes of motion.

## 25. The Persistence of Composite Things

25.1T Mereological Constancy. Necessarily, everything that persists has exactly the same parts at all times at which it exists.
25.1A Mereological Inconstancy. It is possible for a persisting thing to gain or lose parts over time.
25.2T Possibility of Temporary Mereological Coincidence. It is possible for there to be two coincident entities.
25.2A Impossibility of Temporary Mereological Coincidence. It is impossible for there to be two coincident entities.
25.3T Ramsey-Lewis-Sider Perdurantism. A spacetime worm *S* over interval *T* corresponds to the existence of a derived thing persisting through *T* if and only if the simplest and most powerful scientific theory of the actual world assigns a persisting entity to *S*. (=24.3T.3)

25.3T.1 Worm Theory. Ramsey-Lewis-Sider Perdurantism is true, and persisting objects are identical to spacetime worms.
25.3T.2 Stage Theory. Ramsey-Lewis-Sider Perdurantism is true, and objects that can rightly be said to 'persist' are identical to temporal parts of spacetime worms.

## 26. The Existence and Scope of Causation

26.1T Causal Realism. Something is caused.
26.1T.1T Universal Causation. Everything is caused.
26.1T.1A Special Causation. Some things are caused, and some are not.
26.1T.1A.1T Existence of Uncaused Causes. Some uncaused things cause other things.
26.1T.1A.2T Principled Causation. The set of things having causes is finitely definable.
26.1T.1A.2A Unprincipled Causation. The set of things having causes is not finitely definable.
26.1A Causal Anti-Realism. Nothing is caused.

## 27. Causation: A Relation between Things or Truths?

27.1T Causal Connectionism. Causation is fundamentally a relation (*causal connection*) between things (and not between truths).
27.1A Causal Explanationism. Causation is fundamentally a relation (*causal explanation*) between truths.
27.1A.1 Nomological-Deductive Theory of Causal Explanation. The fact that truth *p* causally explains truth *q* consists in the fact that *q* can be deduced from *p*, together with truths about the laws of nature and relevant background conditions.
27.1A.2 Counterfactual Theory of Causal Explanation. If *p* and *q* are truths of the appropriate kind, and *p* causally explains *q*, then this fact consists in the fact that *q* would be not have been true if *p* had not been true (i.e., that (~*p* []-> ~*q*)).
27.1A.3 Probability-Raising Theory of Causal Explanation. The fact that truth *p* causally explains truth *q* consists in the fact that there is an appropriate body of background truths *K(p)* such that the probability of *q* conditional on the conjunction of *p* and *K(p)* is greater than the conditional probability of *q* on *K(p)* alone.

## 28. Discrete and Continuous Causation

28.1T Mark Transmission Theory. A set of events *P* constitutes a process if and only if (i) the spatiotemporal locations of the events in *P* form a continuum *C*, (ii) some conserved quantity *Q* is associated with the events in *P* at a constant value throughout *C*, and (iii) there is some action *M* (the 'marking' action) which did not in fact occur during *C*, but which is such that, if *M* had occurred at some spacetime location within *C*, then every subsequent event in *P* (but none of the earlier events) would have been replaced by a different event associated with a different value of *Q*.

# Appendix B
# Table of Principles

This appendix collects together the principles of methodology, epistemology, truth, metaphysics, and mereology to which we have appealed throughout the book. We have listed, after each principle, the sections in which explicit appeal was made to that principle.

## I. Principles of Methodology

PMeth 1 Ockham's Razor. Other things being equal, adopt the simplest theory. (2.3, 3.9, 5.1, 8.1.3, 11.1, 11.2, 12.1.4, 13.3.2, 17.3, 18.4, 20.5.1, 22.4.1, 22.5.1, 22.7.2, 24.2, 25.1.3.3)

PMeth 1.0 The Zeroth Corollary of Ockham's Razor. Other things being equal, prefer the theory that posits the smallest number of fundamental entities. (3.0, 22.2)

PMeth 1.1 First Corollary of Ockham's Razor: Minimizing Rational Postulates. Other things being equal, prefer the theory that posits the fewest primitive, underivable postulates of reason. (2.3, 5.1, 5.2.3, 20.5.1)

PMeth 1.2 Second Corollary of Ockham's Razor. Other things being equal, adopt the theory with the fewest brute, inexplicable impossibilities and necessities. (2.4.2, 8.1.3, 17.2.2, 18.3.1, 18.4, 18.4.1, 19.1, 21.2.3, 21.6, 22.4.1, 24.2.2, 27.2.2.2)

PMeth 1.3 Third Corollary of Ockham's Razor. Other things being equal, adopt the theory that posits the fewest unexplained and uncaused non-internal relations between things. (5.1)

PMeth 1.4 Fourth Corollary of Ockham's Razor. Adopt the theory that jointly minimizes the following classes: (i) facts about the existence of fundamental things, (ii) facts about the holding of fundamental (natural) properties of and relations between fundamental things, and (iii) facts about brute metaphysical necessity. If more than one theory minimizes these classes (if some trade-off between them is inevitable),

*The Atlas of Reality: A Comprehensive Guide to Metaphysics*, First Edition.
Robert C. Koons and Timothy H. Pickavance.

adopt that theory which makes the best trade-offs. (7.2.4, 9.2, 11.1, 15.2.1.1, 17.2.2, 22.4.1, 24.2.2.2, 27.1)

PMeth 1.4.1 Addendum to Ockham's Razor: Prioritizing Qualitative Economy. When trading off quantitative economy (minimizing the number of fundamental things) and qualitative economy (minimizing the number of fundamental (natural) properties of and relations between fundamental things), always prefer the latter. (7.2.4, 11.1, 18.1.1)

PMeth 1.5 Fifth Corollary of Ockham's Razor. Other things being equal, adopt the theory that attributes the fewest ambiguities to natural language. (12.1.2, 17.4)

PMeth 1.6 Sixth Corollary of Ockham's Razor: Minimizing Possibilities. Other things being equal, prefer the theory that posits the smallest class of metaphysical possibilities. (15.1.1, 17.3)

PMeth 2 Scientific Realism. Other things being equal, adopt the theory that implies that our best scientific theories are straightforwardly true, as standardly represented. (5.2.1, 18.3.1, 20.5.6)

PMeth 2.1 Scientific Realism: Objectivity. Other things being equal, adopt the theory that implies that our best scientific theories are objectively true: true independently of our scientific preferences and practices. (5.2.1)

PMeth 2.2 Scientific Realism: Reliability. Other things being equal, adopt the theory that implies that we are reasonably reliable in finding scientific truth. (5.2.1)

PMeth 3 Structuralism. Other things being equal, adopt the theory that explains metaphysical impossibilities in terms of the essential structure of things—e.g., in the case of extended things, in terms of the essential structure of space or time. (15.1.1, 18.3.1)

PMeth 4 Redundancy for Composite Entities. Reject any theory that posits any kind of composite entity without both emergent active powers and essentially unitary passive or immanent powers. (22.3, 22.4.1, 22.4.2, 22.5.1)

## II. Principles of Epistemology

PEpist 1 Imagination as Guide to Possibility. If a scenario is imaginable in great detail without evident absurdity, then we have good reason to think that it represents a metaphysical possibility. (5.2.2, 6.3, 9.3.2.1, 16.2, 18.3, 19.1.1, 19.1.2, 20.5.6, 22.4.2, 26.1.1, 5.2.2, 9.3.2.1, 16.2, 19.1.1, 20.5.6, 26.1.1, 26.2.3)

PEpist 1.1 Limit of Possibles Itself Possible. If a series of scenarios is such that each represents a metaphysical possibility, and the series converges at the limit on a further scenario, then we have good reason to think that the latter scenario represents a metaphysical possibility. (9.3.2.1)

PEpist 2 Reidian Common Sense. It is prima facie plausible to presume the truth of every article of common sense. (13.6.1, 20.5.3, 21.2.1)

PEpist 2.1 Ethical Practices Presumption. It is prima facie plausible to suppose that all of the presuppositions of our fundamental ethical beliefs and practices are true. (11.1, 13.6.1, 22.6.1)

PEpist 2.2 Rational Practices Presumption. It is prima facie plausible to suppose that all of the presuppositions of rational practices (such as deliberation about the future or induction) are true. (13.6.1, 14.1)

PEpist 2.2.1 Presumption of Philosophical Discourse. It is prima facie plausible to presume the truth of all of the pragmatic presuppositions of philosophical discourse. (20)

PEpist 3 Appearance of Bodies and Minds. Perception and memory present us with what are apparently distinct physical things, including some embodying apparently distinct minds. (11.2, 13.3)

PEpist 4 Reliable Perception Presumption. It is prima facie plausible to suppose that human perception and memory are reliable. (11.2, 13.3, 13.4, 13.6.1, 17.4, 19.1.1, 20, 21.3)

PEpist 4.1 Sensory Error Minimization. Other things being equal, prefer a theory that posits the fewest and least severe sensory errors to human subjects. (13.3.1)

PEpist 5 Conceptual Acquaintance. It is possible to know the truth of a proposition that involves the use of a concept *C* only if either (i) one knows a definition of *C* in terms of other concepts, or (ii) one knows a non-empty class of instances of *C*. (13.4, 13.6.4)

## III. Principles of Truth

PTruth 1 One Truthmaker per Fundamental Property. If *p* is the true predication of a fundamental property *P* to $x_1$ through $x_n$, and *q* is the true predication of a different fundamental property *Q* to the same things $x_1$ through $x_n$, then *p* and *q* have distinct truthmakers. (2.5.2, 3.4.3, 8.1.3, 9.3.2.1, 18.4.1)

PTruth 2 Tarski's Schema. For any sentence *s*, if *S* is a name for *s*, then we should affirm the sentence of the form: *S* is true if and only if s. (2.3, 12.2.2)

## IV. Principles of Metaphysics

PMeta 1 The Necessary and Sufficient Condition Test for Necessary Grounding. If the truth of a proposition *p* is necessarily wholly grounded in the truth of proposition *q*, then the truth of *q* is a metaphysically necessary and sufficient condition for the truth of *p*. (4.4)

PMeta 1.1 The Necessary Condition Test for Necessary Grounding. If the truth of a proposition *p* is necessarily wholly grounded in the truth of proposition *q*, then the truth of *q* is a metaphysically necessary condition for the truth of *p*. (4.4)

PMeta 2 Intrinsicality of Powers. Having a power is an intrinsic property. (5.2.2, 19.1.2, 20.5.6, 26.2.1)

PMeta 3 Membership the Only Fundamental Set Relation. If *S* is a set, then the only fundamental relation involving *S* is the membership relation between *S*'s members and itself. (8.1.2, 10.2.1, 18.1.1, 18.4.1)

PMeta 4 General Principle of Constituent Identity. If *x* and *y* both have proper parts, and they have exactly the same proper parts, then they must be identical.
(23.3)

PMeta 4.1 PCI for Substances. If *A* and *B* are substances, and every constituent of *A* is a constituent of *B* and vice versa, then *A* is identical to *B*. (9.3.2)

PMeta 4.2 Weak PCI. If $x$ and $y$ are necessarily composite and necessarily have the same proper parts, then $x = y$. (9.3.2)

PMeta 4.3 PCI for Bundles. If $x$ and $y$ are bundles of properties, and $x$ and $y$ have the same proper parts, then $x = y$. (9.3.2)

PMeta 5.1 Finite spatiotemporal patchwork. If an event or process of (intrinsic) type $A$ is possible, as is an event or process of intrinsic type $B$, and if there is enough room in the history of the world to locate in it instances of both events (or processes) without overlap in time and space, then it is possible for events (or processes) of both types to occur together. (6.3, 16.2.3, 20.5.6, 24.1, 26.2.1)

PMeta 5.2 Infinite patchwork. If $T$ is a class of types of events or processes, and for each member of $T$, it is possible for an event or process of type $T$ to occur, and there is enough room in history of the world to locate within it instances of each of the types in $T$ without overlap in space and time between the instances, then it is possible for all the types in $T$ to be realized together. (16.2.3, 18.3.1, 19.1.2, 24.1, 26.2.1)

## V. Principles of Natural Philosophy

PNatPhil 1 Possibility of Ubiquitous Atomless Gunk. It is possible for nothing to exist but atomless gunk (things with parts but no atomic parts). (11.2.4, 22.4.2)

PNatPhil 2 Independence of Motion and Substantial Change. Neither the movement of two extended things nor the division of an extended thing into two parts can necessitate the creation or annihilation of material parts (not even point-masses). (18.3.3, 22.4.2)

PNatPhil 3 Continuity of Motion. It is impossible for any material thing to move discontinuously through spacetime. (19.1.2, 22.4.2, 24.2.1)

## VI. Principles of Mereology

Axioms of Classical Mereology (23.1, 23.3)

(MA1) Reflexivity. Everything is a part of itself.

(MA2) Anti-Symmetry. If $x$ is a part of $y$, and $y$ is a part of $x$, then $x = y$.

(MA3) Transitivity. If $x$ is a part of $y$, and $y$ is a part of $z$, then $x$ is a part of $z$.

(MA4) Weak Supplementation. If $x$ is a proper part of $y$, then $y$ has some part $z$ that does not overlap $x$.

(MA5) Strong Supplementation. If $x$ is not a part of $y$, then $x$ has some part $z$ that does not overlap $y$.

(MA6) Arbitrary Sums If $S$ is a non-empty set, then there exists at least one sum of $S$.

(MA6′) Arbitrary Parts. If there is a sum of the members of set $S$, and $S'$ is a subset of $S$, then there is a sum of the members of $S'$.

(MC1) Extensionality for Wholes. If $x$ and $y$ have the same proper parts, then $x = y$. (Follows from Anti-Symmetry and Strong Supplementation)

# References

Adams, R.M. (1974), Theories of Actuality, *Noûs* 8: 211–232.

Adams, R.M. (1979), Primitive Thisness and Primitive Identity, *Journal of Philosophy* 76: 5–26.

Adams, R.M. (1981), Actualism and Thisness, *Synthèse* 49: 3–41.

Albert, D.Z. (1994), *Quantum Mechanics and Experience* (Cambridge, MA: Harvard University Press).

Albert, D.Z. and Loewer, B. (1988), Interpreting the Many-Worlds Interpretation, *Synthèse* 77: 195–213.

Alexander, S. (1920), *Space, Time and Deity* (London: Macmillan and Company).

Alston, W.P. (1954), Particulars—Bare and Qualified, *Philosophy and Phenomenological Research* 15: 253–258.

Anscombe, G.E.M. (1964), Substance, *Proceedings of the Aristotelian Society*, supp. vol. 38: 69–78.

Anscombe, G.E.M. (1965), The Intentionality of Sensation: a Grammatical Feature, in R.J. Butler ed., *Analytical Philosophy: First Series* (Oxford: Blackwell), pp. 143–158.

Anscombe, G.E.M. (1974), 'Whatever has a Beginning of Existence Must have a Cause': Hume's Argument Exposed, *Analysis* 34: 145–151.

Anscombe, G.E.M. (1975) Causality and Determination, in E. Sosa, ed., *Causation and Conditionals* (Oxford: Oxford University Press), pp. 63–81.

Anselm of Canterbury (1059/1998), *The Major Works*, ed. B. Davies and G.R. Evans (Oxford: Oxford University Press).

Aristotle (1984), *Complete Works of Aristotle*, J. Barnes, ed. (Princeton, NJ: Princeton University Press).

Armstrong, D.M. (1963), Absolute and Relative Motion, *Mind* 72: 209–223.

Armstrong, D.M. (1968), *A Materialist Theory of Mind* (London: Routledge).

Armstrong, D.M. (1980), Identity through Time, in P. van Inwagen ed., *Time and Cause: Essays Presented to Richard Taylor* (Dordrecht: D. Reidel), pp. 67–78.

*The Atlas of Reality: A Comprehensive Guide to Metaphysics*, First Edition.
Robert C. Koons and Timothy H. Pickavance.

Armstrong, D.M. (1983), *What Is a Law of Nature?* (Cambridge: Cambridge University Press).
Armstrong, D.M. (1989a), *A Combinatorial Theory of Possibility* (Cambridge: Cambridge University Press).
Armstrong, D.M. (1989b), *Universals: An Opinionated Introduction* (Boulder, CO: Westview Press).
Armstrong, D.M. (1993), The Identification Problem and the Inference Problem, *Philosophy and Phenomenological Research*, 53: 421–422.
Armstrong, D.M. (1997), *A World of States of Affairs* (Cambridge: Cambridge University Press).
Armstrong, D.M. (2004), *Truth and Truthmakers* (Cambridge: Cambridge University Press).
Armstrong, D.M., Martin, C.B., & Place, U.T. (1996), *Dispositions: A Debate*, London: Routledge.
Arntzenius, F. (1993), The Common Cause Principle, *Philosophy of Science Association* 2: 227–237.
Arntzenius, F. (2010), Reichenbach's Common Cause Principle, *Stanford Encyclopedia of Philosophy*. Available at http://plato.stanford.edu/entries/physics-Rpcc/ (last accessed 5 June 2015).
Audi, P. (2012), Grounding: Toward a Theory of the In-Virtue-of Relation, *Journal of Philosophy* 109: 685–711.
Austin, J.L. (1962), *Sense and Sensibilia* (Oxford: Clarendon Press).
Azzouni, J. (2008), *Tracking Reason: Proof, Consequence, and Truth* (Oxford: Oxford University Press).
Barnes, E. (2010), Ontic Vagueness: A Guide for the Perplexed, *Noûs* 44 (4): 601–627.
Barnes, W.H.F. (1945), The Myth of Sense Data, *Proceedings of the Aristotelian Society, New Series* 45, pp. 89–117.
Barwise, J. and Etchemendy, J. (1987), *The Liar* (Oxford: Oxford University Press).
Barwise, J. and Perry, J. (1983), *Situations and Attitudes* (Cambridge, MA: MIT Press).
Baxter, D. (1988), Identity in the Loose and Popular Sense, *Mind* 97: 575–582.
Benardete, J.A. (1964), *Infinity: An Essay in Metaphysics* (Oxford: Oxford University Press).
Bennett, J. (1984), Counterfactuals and Temporal Direction, *Philosophical Review*, 93: 57–91.
Bennett, J. (2003), *A Philosophical Guide to Conditionals* (Oxford: Oxford University Press).
Berenstain, N. (forthcoming), What a Structuralist Theory of Properties Could Not Be, in A. Marmodoro and D. Yates eds., *The Metaphysics of Relations* (Oxford: Oxford University Press).
Berkeley, G. (1710/2009), *Principles of Human Knowledge and Three Dialogues* (New York: Oxford University Press).
Bigelow, J. (1988), *The Reality of Numbers: A Physicalist's Philosophy of Mathematics* (Oxford: Clarendon Press).
Bigelow, J. (1996), Presentism and Properties, in J. Tomberlin, James ed., *Philosophical Perspectives 10, Metaphysics* (Oxford: Blackwell), pp. 35–52.
Bird, A. (1998), Dispositions and Antidotes, *The Philosophical Quarterly* 48: 227–234.
Bird, A. (2010), *Nature's Metaphysics: Laws and Properties* (New York: Oxford University Press).
Bishop, R. (2005), Patching Physics and Chemistry Together, *Philosophy of Science* 72: 710–722.
Black, M. (1952), The Identity of Indiscernibles, *Mind* 61: 153–164.
Boghossian, P.A. (1989), The Rule-following Considerations, *Mind* 98: 507–549.

Bonevac, D., Dever, J., and Sosa, D. (2006), The Conditional Fallacy, *The Philosophical Review* 115: 273–316.

Boolos, G. (1984), To Be is to Be the Value of a Variable (or To Be Some Values of Some Variables), *Journal of Philosophy* 81: 430–449.

Boolos, G. (1998), *Logic, Logic and Logic* (Cambridge, MA: Harvard University Press).

Boscovich, R. (1763/1966), *A Theory of Natural Philosophy* (Cambridge, MA: MIT Press).

Bradley, F.H. (1893/1930), *Appearance and Reality* (Oxford: Clarendon Press).

Brentano, F. (1976/1988), *Philosophical Investigations on Space, Time and the Continuum*, ed. Stephen Körner and Roderick M. Chisholm, trans. Barry Smith (London: Croom Helm).

Broad, C.D. (1923/1959), *Scientific Thought* (London: Kegan Paul).

Broad, C.D. (1925), *Mind and its Place in Nature* (London: Kegan Paul).

Broad, C.D. (1938), *An Examination of McTaggart's Philosophy, Volume II* (Cambridge: Cambridge University Press).

Bub, J. (1997), *Interpreting the Quantum World* (Cambridge: Cambridge University Press).

Burge, T. (1979), Semantical Paradox, *Journal of Philosophy* 76: 169–198.

Burke, M. (1994), Preserving the Principle of One Object to a Place: A Novel Account of the Relations among Objects, Sorts, Sortals and Persistence Conditions, *Philosophy and Phenomenological Research* 54: 591–624.

Butler, J. (1839/1964), Of Personal Identity, in A. Flew ed., *Body, Mind, and Death* (New York: Macmillan), pp. 166–172.

Canfield, J.V. (1996), The Community View, *The Philosophical Review* 105: 469–488.

Carnap, R. (1928/1967), *The Logical Structure of the World*, trans. Rolf George (London: Routledge and Kegan Paul).

Carnap, R. (1934/37), *Logische Syntax der Sprache*, Vienna: Springer, 1934, revised edition; trans. by A. Smeaton, *The Logical Syntax of Language* (London: Kegan, Paul, Trench Teubner & Cie, 1937).

Carnap, R. (1947), *Meaning and Necessity* (Chicago: University of Chicago Press).

Cartwright, N. (1983), *How the Laws of Physics Lie* (Oxford: Oxford University Press).

Cartwright, N. (1994), *Nature's Capacities and Their Measurement* (Oxford: Oxford University Press).

Casati, R. and Varzi, A. (1994), *Holes and Other Superficialities* (Cambridge, MA: MIT Press).

Chalmers, D. (2002), Does Conceivability Entail Possibility?, in T. Gendler and J. Hawthorne, eds., *Conceivability and Possibility*, (Oxford: Oxford University Press, 2002), pp. 145–200.

Chisholm, R.M. (1957), *Perceiving: A Philosophical Study* (Ithaca, NY: Cornell University Press).

Chisholm, R.M. (1967), Identity through Possible Words: Some Questions, *Noûs* 1: 1–8.

Chisholm, R.M. (1973a), Homeless Objects, *Revue Internationale de Philosophie* 22: 207.

Chisholm, R.M. (1973b), Parts as Essential to their Wholes, *Review of Metaphysics*, 26: 581–603.

Chisholm, R.M. (1976), *Person and Object* (London: Allen and Unwin).

Chisholm, R.M. (1982), *Brentano and Meinong Studies* (Atlantic Highlands, NJ: Humanities Press).

Chisholm, R.M. (1989), *On Metaphysics* (Minneapolis: University of Minnesota Press).

Clarke, S. (1704/1998), *A Demonstrations of the Being and Attributes of God*, ed. Ezio Vailati (Cambridge: Cambridge University Press).

Collins, J., Hall, N. and Paul, L.A. (2004), *Causation and Counterfactuals* (Cambridge, MA: MIT Press).

Craig, W.L. (1979), *The Kalam Cosmological Argument* (London: Macmillan).
Craig, W.L. (1980), *The Cosmological Argument from Plato to Leibniz* (New York: Harper & Row).
Craig, W.L. (2008), The Metaphysics of Relativity: Three Views, in W.L. Craig and Q. Smith eds., *Einstein, Relativity, and Absolute Simultaneity* (Abindgon: Routledge), pp. 11–149.
Crane, T. and French, C. (2015), The Problem of Perception, *Stanford Encyclopedia of Philosophy*. Available at http://plato.stanford.edu/entries/perception–problem/.
Cross, R. (1999), *Duns Scotus* (Oxford: Oxford University Press).
Cushing, J.T. and McMullin, E. (eds.) (1992), *Philosophical Consequences of Quantum Theory: Reflections on Bell's Theorem* (South Bend, IN: University of Notre Dame Press).
Cutter, B. (2013), Panqualityism, unpublished manuscript.
Dancy, J. (2004), *Ethics without Principles*, Oxford: Clarendon Press.
Dasgupta, S. (2014), Metaphysical Rationalism, *Noûs* 49.
Davidson, D. (1967), The Logical Form of Action Sentences, in N. Rescher (ed.), *The Logic of Decision and Action* (Pittsburgh, PA: University of Pittsburgh Press).
Davis, R.L. (2011), Reconciling Realism and Common Sense about Color, unpublished manuscript.
De Valois, R. and De Valois, K. (1993), A Multi–Stage Model of Color Vision, *Vision Research* 33: 1053–1065.
Demos, R. (1917), A Discussion of Certain Types of Negative Propositions, *Mind* 26: 188–196.
Donnellan, K.S. (1966), Reference and Definite Descriptions, *The Philosophical Review* 75 (3): 281–304.
Dowe, P. (1995), Causality and Conserved Quantities: A Reply to Salmon, *Philosophy of Science* 62: 321–333.
Dowe, P. (2000), *Physical Causation* (Cambridge: Cambridge University Press).
Dretske, F. (1969), *Seeing and Knowing* (Chicago, IL: University of Chicago Press).
Dretske, F. (1977), Laws of Nature, *Philosophy of Science* 44: 248–268.
Dretske, F. (1988), *Explaining Behavior* (Cambridge, MA: MIT Press).
Ducasse, C.J. (1942), Moore's Refutation of Idealism, in P. Schilpp ed., *The Philosophy of G.E. Moore* (Chicago, IL: Northwestern University Press), pp. 223–252.
Dummett, M. (1969), The Reality of the Past, *Proceedings of the Aristotelian Society* 69: 239–258. (Reprinted in Dummett, M., *Truth and Other Enigmas* (London: Duckworth, 2nd edition).
Dummett, M. (1980), *Truth and Other Enigmas* (London: Duckworth, 2nd edition).
Eells, E. (1991), *Probabilistic Causality* (Cambridge: Cambridge University Press).
Ehring, D. (1997), *Causation and Persistence* (Oxford: Oxford University Press).
Ehring, D. (2004), Distinguishing Universals and Particulars, *Analysis* 64: 326–332.
Elga, A. (2001), Statistical Mechanics and the Asymmetry of Counterfactual Dependence, *Philosophy of Science* 68 (S1): S313–S324.
Ellis, B. (1999) Causal Powers and Laws of Nature, in H. Sankey, ed., *Causation and Laws of Nature* (Dordrecht: Kluwer Academic Publishers), pp. 19–34.
Ellis, B. (2001), *Scientific Essentialism* (New York: Cambridge University Press).
Evans, G. (1978), Can there be Vague Objects? *Analysis* 38: 208.
Fara, M. (2005), Dispositions and Habituals, *Noûs* 39: 43–82.
Fair, D. (2003), Physical Causation, *Mind* 112: 529–533.
Field, H. (1980), *Science without Numbers* (Oxford: Blackwell).
Field, H. (1984), Can We Dispense with Space-Time?, *PSA: Proceedings of the Philosophy of Science Association, Vol 2: Symposia and Invited Papers*, pp. 33–90.

Field, H. (2001). *Truth and Absence of Fact* (Oxford: Oxford University Press).
Fine, K. (1975), Vagueness, Truth and Logic, *Synthèse* 30, pp. 265–300.
Fine, K. (1985), Plantinga on Possibilistic Discourse, in J. Tomberlin and P. van Inwagen eds., *Profiles: Alvin Plantinga* (Dordrecht: D. Reidel), pp. 145–186.
Fine, K. (1994a), Essence and Modality: The Second Philosophical Perspectives Lecture, *Philosophical Perspectives* 8: 1–16.
Fine, K. (1994b), Ontological Dependence, *Proceedings of the Aristotelian Society* 95: 269–290.
Fine, K. (1994c), Senses of Essence, in N. Asher, D. Raffman, and W. Sinnott-Armstrong eds., *Modality, Morality, and Belief: Essays in Honor of Ruth Barcan Marcus* (Chicago: University of Chicago Press), pp. 53–73.
Fine, K. (2000), Neutral Relations, *Philosophical Review* 14: 1–33.
Fine, K. (2001), The Question of Realism, *Philosophers' Imprint* 1 (2): 1–30.
Fine, K. (2002). Varieties of Necessity, in T.S. Gendler and J. Hawthorne eds., *Conceivability and Possibility* (Oxford: Oxford University Press), pp. 253–281.
Fine, K. (2007), Response to Fraser McBride, *Dialectica* 61: 57–62.
Fine, K. (2012a), Guide to Ground, in F. Correia and Benjamin Schnieder eds., *Metaphysical Grounding: Understanding the Structure of Reality*, Cambridge: Cambridge University Press.
Fine, K. (2012b), The Pure Logic of Ground, *Review of Symbolic Logic* 5: 1–25.
Flew, A. (1964), (ed.), *Body, Mind, and Death* (New York: Macmillan).
Forbes, G. (1985), *The Metaphysics of Modality* (New York: Oxford).
Forrest, P. (1986), Ways Worlds Could Be, *Australasian Journal of Philosophy* 64: 15–24, 89–91.
Forrest, P. (1999), Supertasks and Material Objects, *Logique & Analyse* 167–168: 441–446.
Frege, G. (1892/1980), On Sense and Reference, in P. Geach and M. Black eds. and trans. *Translations from the Philosophical Writings of Gottlob Frege* (Oxford: Blackwell).
Gale, R. (1968), *The Language of Time* (London: Routledge & Kegan Paul).
Garcia, R.K. (2015), Two Ways to Particularize a Property, *Journal of the American Philosophical Association* 1: 635–652.
Garcia, R.K. (2016), Tropes as Character-Grounders, *Australasian Journal of Philosophy* 94: 499–515.
Geach, P. (1967), Intentional Identity, *Journal of Philosophy* 64 (20): 627–632.
Geach, P. (1972), *Logic Matters* (Berkeley, CA: University of California Press).
Georgi, H.M. (1989), Effective Field Theories, in P. C. W. Davies ed., *The New Physics* (New York: Cambridge University Press), pp. 446–457.
Gettier, E.L. (1963), Is Justified True Belief Knowledge?, *Analysis* 23: 121–123.
Glanzberg, M. (2001) The Liar in Context, *Philosophical Studies* 103: 217–251.
Glanzberg, M. (2004), A Contextual-hierarchical Approach to Truth and the Liar Paradox, *Journal of Philosophical Logic* 33: 27–88.
Goldman, A. (1977), Perceptual Objects, *Synthèse* 35: 257–284.
Goodman, N. (1951) *The Structure of Appearance* (Cambridge, MA: Harvard University Press).
Goodman, N. (1954), *Fact, Fiction, and Forecast* (Cambridge, MA: Harvard University Press).
Goss, M. (2006), Time and Human Nature: A Modest Defense of Eternalism, University of Texas at Austin PhD dissertation.
Grossmann, R. (1974), *Meinong* (London: Routledge).
Grover, D. (1992), *A Prosentential Theory of Truth* (Princeton, NJ: Princeton University Press).

Grünbaum, A. (1967), *Modern Science and Zeno's Paradoxes* (Middletown, CT: Wesleyan University Press), pp. 78–135.

Haldane, J. (2007), Privative Causality, *Analysis* 67: 180–186.

Hall, N. (2004), Two Concepts of Causation, in J. Collins, N. Hall, and L.A. Paul (2004), pp. 225–276.

Hardin, C.L. (1988/1993), *Color for Philosophers* (Indianapolis, IN: Hackett).

Hawthorne, J. (1996), The Bundle Theory of Substance and the Identity of Indiscernibles, *Analysis* 55: 191–196.

Hawthorne, J. (2000), Before-effect and Zeno Causality, *Noûs* 34: 622–633.

Hawthorne, J. (2001), Causal Structuralism, *Philosophical Perspectives* 15: 361–378.

Hawthorne, J. (2006), *Metaphysical Essays* (Oxford: Oxford University Press).

Hawthorne, J. and Kovakovich, K. (2006), Disjunctivism, *Proceedings of the Aristotelian Society Supplemental Volume* 80: 145–183.

Hawthorne, J. and Weatherson, B. (2004), Chopping up Gunk, *Monist* 87: 339–350.

Healey, R. (1989), *The Philosophy of Quantum Mechanics: An Interactive Interpretation* (Cambridge: Cambridge University Press).

Hendry, R.F. (2006), Is there Downward Causation in Chemistry? in D. Baird, L. McIntyre, and E.R. Scerri eds., *Philosophy of Chemistry: Synthesis of a New Discipline* (Dordrecht: Springer), pp. 173–189.

Hendry, R.F. (2010), Emergence vs. Reduction in Chemistry, in C. MacDonald and G. MacDonald eds., *Emergence in Mind* (Oxford: Oxford University Press), pp. 205–221.

Hirsch, E. (1992), *The Concept of Identity* (Oxford: Oxford University Press).

Hirsch, E. (1993), *Dividing Reality* (Oxford: Oxford University Press).

Hochberg, H. (1965), Universals, Particulars and Predication, *Review of Metaphysics* 19: 87–102.

Hochberg, H. (1999), *Complexes and Consciousness* (Stockholm: Thales).

Hochberg, H. (2004), Russell and Ramsey on Distinguishing between Universals and Particulars, *Grazer Philosophische Studien* 67: 195–207.

Hofweber, T. (2009), Ambitious, Yet Modest, Metaphysics, in D. Chalmers, D. Manley, and R. Wasserman eds., *Metametaphysics: New Essays on the Foundation of Ontology* (Oxford: Clarendon Press), pp. 260–289.

Horwich, P. (1998), *Truth* (Oxford: Blackwell, first edition 1990).

Hume, D. (1748/2007), *An Enquiry Concerning Human Understanding*, Peter Millican, ed. (Oxford: Oxford University Press).

Hume, D. (1779/1998), *Dialogues Concerning Natural Religion*, Richard H. Popkin, ed., 2nd edn. (Indianapolis: Hackett).

Hunt, R.G.W. (1982), A Model of Colour Vision for Predicting Colour Appearance, *Color Research and Application* 7: 95–112.

Jackson, F. (1977), *Perception: A Representative Theory* (Cambridge: Cambridge University Press).

Jackson, F. (1987), *Conditionals* (Oxford: Blackwell).

James, W. (1890), *Principles of Psychology* (New York: Dover).

Jeffrey, R.C. (1965), *The Logic of Decision* (New York: McGraw-Hill).

Johansson, I. (1989), *Ontological Investigations* (London: Routledge).

Johnson, W.E. (1921–24/1964), *Logic* (New York: Dover Publications Inc. 1964) (Cambridge University Press, vol. 1, 1921; vol. 2, 1922; vol. 3, 1924).

Johnston, M. (1983), Particulars and Persistence, Ph.D. dissertation, Princeton University.

Johnston, M. (1987), Is There a Problem About Persistence? *Proceedings of the Aristotelian Society (Supp)* 61: 107–135.

Johnston, M. (1992), How to Speak of the Colors, *Philosophical Studies* 68: 221–263.

Kamp, H. (1975), *T*wo Theories About Adjectives, in E.L. Keenan ed., *Formal Semantics of Natural Language* (Cambridge: Cambridge University Press).

Kant, I. (1965, orig. 1781/1787), *Critique of Pure Reason*, trans. Norman Kemp Smith (New York: St. Martin's Press). Originally *Kritik der reinen Vernunft.*

Kaplan, D. (1989), Demonstratives: An Essay on the Semantics, Logic, Metaphysics, and Epistemology of Demonstratives and Other Indexicals, in J. Almog, J. Perry, and H. Wettstein eds., *Themes from Kaplan* (Oxford: Oxford University Press), pp. 481–563.

Kenny, A. (1973), *Wittgenstein* (London: Allen Lane), Chapter 10.

Kim, J. (1973), Causation, Nomic Subsumption, and the Concept of an Event, *Journal of Philosophy* 70: 317–336.

Kim, J. (1976), Events as Property Exemplifications, in M. Brand and D. Walton (eds.), *Action Theory* (Dordrecht: D. Reidel Publishing), pp. 159–177.

Kim, J. (1992), Downward Causation, in A. Beckerman, H. Flohr, and J. Kim eds., *Emergentism and Reduction* (New York and Berlin: De Gruyter).

Kim, J. (1993), 'Strong' and 'Global' Supervenience Revisited, *Supervenience and Mind: Selected Philosophical Essays* (Cambridge: Cambridge University Press), pp. 79–91.

Kim, J. (1994), Explanatory Knowledge and Metaphysical Dependence, *Philosophical Issues* 5: 51–69.

Kim, J. (1998), *Mind in a Physical World* (Cambridge: Cambridge University Press).

Kim, J. (2001), Mental Causation and Consciousness: The Two Mind-Body Problems for the Physicalist, in C. Gillett and B. Loewer eds., *Physicalism and Its Discontents* (Cambridge: Cambridge University Press).

Kirk, G.S., Raven, J.E., and Schofield, M. (1983), *The Presocratic Philosophers* (Cambridge: Cambridge University Press, second edition).

Koons, R.C. (1992), *Paradoxes of Belief and Strategic Rationality* (Cambridge: Cambridge University Press).

Koons, R.C. (1997), A New Look at the Cosmological Argument, *American Philosophical Quarterly* 34: 171–192.

Koons, R.C. (2000), *Realism Regained: An Exact Theory of Causation, Teleology and the Mind* (Oxford: Oxford University Press).

Koons, R.C. (2003), Review of *Physical Causation*, *Philosophy and Phenomenological Research*, 67: 244–248.

Koons, R.C. (2008), Epistemological Foundations for the Cosmological Argument, *Oxford Studies in the Philosophy of Religion*, ed. J. Kvanvig (Oxford: Oxford University Press).

Koons, R.C. (2014a), A New Kalam Argument: Revenge of the Grim Reaper, *Noûs* 48: 256–267.

Koons, R.C. (2014b), Staunch vs. Faint-hearted Hylomorphism: Toward an Aristotelian Account of Composition, *Res Philosophica* 91: 1–27.

Koons, R.C. and Bealer, G. (2010), *The Waning of Materialism: New Essays in the Philosophy of Min*d (New York: Oxford University Press).

Kretzmann, N. (1997), *The Metaphysics of Theism: Aquinas's Natural Theology in Summa Contra Gentiles I* (Oxford: Oxford University Press).

Kripke, S.A. (1963), Semantical Analysis of Modal Logic I: Normal Modal Propositional Calculi, *Zeitschrift für mathematische Logik und Grundlagen der Mathematik* 9: 67–96.

Kripke, S.A. (1980), *Naming and Necessity* (Cambridge, MA: Harvard University Press).

Kuhn, T.A. (1970), *The Structure of Scientific Revolutions* (Chicago: University of Chicago Press)

Kvart, I. (2004), Causation: Probabilistic and Counterfactual Analyses, in J. Collins, N. Hall, and L.A. Paul, pp. 359–386.

Lange, M. (2004), A Note on Scientific Essentialism, Laws of Nature, and Counterfactual Conditionals, *Australasian Journal of Philosophy* 82: 227–241.

Leonard, H.S. and Goodman, N. (1940), The Calculus of Individuals and Its Uses, *Journal of Symbolic Logic* 5: 45–55.

Le Poidevin, R. (2005), *Travels in Four Dimensions: The Enigmas of Space and Time* (New York: Oxford University Press).

Leśniewski, S. (1916), *Podstawy ogólnej teoryi mnogosci. I* (Moscow: Prace Polskiego Kola Naukowego w Moskwie, Sekcya matematyczno-przyrodnicza), Eng. trans. by D.I. Barnett, Foundations of the General Theory of Sets. I', in S. Leśniewski, *Collected Works*, ed. S.J. Surma, J. Srzednicki, D.I. Barnett, and F.V. Rickey (Dordrecht: Kluwer, 1992), Vol. 1, pp. 129–173.

Lewis, C.I. (1932), Alternative Systems of Logic, *The Monist* 42: 481–507.

Lewis, D.K. (1970a), How to Define Theoretical Terms, *Journal of Philosophy* 67: 427–446.

Lewis, D.K. (1970b), Anselm and Actuality, *Noûs* 4: 175–188.

Lewis, D.K. (1973a), Causation, *Journal of Philosophy* 70: 556–567. Reprinted, with Postscript, in D.K. Lewis, Against Structural Universals, *Australasian Journal of Philosophy* 64: 25–46.

Lewis, D.K. (1973b), *Counterfactuals* (Oxford: Blackwell).

Lewis, D.K. (1976), The Paradoxes of Time Travel, *American Philosophical Quarterly* 13: 145–152.

Lewis, D.K. (1979a), Attitudes De Dicto and De Se, *Philosophical Review* 88: 513–543.

Lewis, D.K. (1979b), Counterfactual Dependence and Time's Arrow, *Noûs* 13: 455–476.

Lewis, D.K. (1980a), Survival and Identity, *Philosophical Papers—Volume 1* (New York: Oxford University Press), pp. 55–77.

Lewis, D.K. (1980b), A Subjectivist's Guide to Objective Chance, in R.C. Jeffrey, ed., *Studies in Inductive Logic and Probability*, Vol. II (Berkeley: University of California Press), pp. 263–293.

Lewis, D.K. (1981), Causal decision theory, *Australasian Journal of Philosophy* 59: 5–30.

Lewis, D.K. (1983), New Work for a Theory of Universals, *Australasian Journal of Philosophy* 61: 343–377.

Lewis, D.K. (1984), Putnam's Paradox, *Australasian Journal of Philosophy* 62: 221–236.

Lewis, D.K. (1986a), *On the Plurality of Worlds* (Oxford: Blackwell).

Lewis, D.K. (1986b), Against Structural Universals, *Australasian Journal of Philosophy* 64: 25–46.

Lewis, D.K. (1991), *Parts of Classes* (Oxford: Blackwell).

Lewis, D.K. (1993), Many, but Almost One, in J. Bacon ed., *Ontology, Causality and Mind: Essays in Honour of D M Armstrong*, (New York: Cambridge University Press).

Lewis, D.K. (1994), Humean Supervenience Debugged, *Mind* 103: 473–390.

Lewis, D.K. (1997), Finkish Dispositions, *Philosophical Quarterly* 47: 143–158.

Lewis, D.K. (1999a), *Papers in Metaphysics and Epistemology* (Cambridge: Cambridge University Press).

Lewis, D.K. (1999b), Zimmerman and the Spinning Sphere, *Australasian Journal of Philosophy* 77: 209–214.

Lewis, D.K. (2000), Causation as Influence, *Journal of Philosophy* 97: 182–197.
Lewis, D.K. (2001), Truthmaking and Difference-Making, *Noûs*, 35: 602–615.
Lewis, D.K. (2002), Tensing the Copula, *Mind*, 111: 1–14.
Lewis, D.K. (2004a), Void and Object, in J. Collins, N. Hall, and L.A. Paul, pp. 277–290.
Lewis, D.K. (2004b), Tensed Quantifiers, in D. Zimmerman ed., *Oxford Studies in Metaphysics Volume 1* (New York: Oxford University Press), pp. 3–14.
Lewis, D.K. and Lewis, S.R. (1970), Holes, *Australasian Journal of Philosophy* 48: 206–212.
Litland, J.E. (forthcoming), Grounding Ground, *Oxford Studies in Metaphysics*.
Locke, J. (1689/1979), *An Essay Concerning Human Understanding*, ed. Peter Nidditch (New York: Oxford University Press).
Loux, M. (2006), Aristotle's Constituent Ontology, in D. Zimmerman ed., *Oxford Studies in Metaphysics*, vol. 2 (Oxford: Clarendon Press), pp. 207–250.
Lowe, E.J. (1998), *The Possibility of Metaphysics* (Oxford: Oxford University Press).
Lowe, E.J. (2004), The Particular-Universal Distinction: A Reply to MacBride, *Dialectica* 58: 335–340.
Lowe, E.J. (2006), *The Four-Category Ontology: A Metaphysical Foundation for Natural Science* (Oxford: Clarendon Press).
Lowe, E.J. (2010), How Not to Think of Powers: A Deconstruction of the 'Dispositions and Conditionals' Debate, *Monist* 94: 19–33.
Ludlow, P. (1999), *Semantics, Tense, and Time: An Essay in the Metaphysics of Natural Language* (Cambridge, MA: MIT Press).
Łukasiewicz, J. (1967), On Determinism, in S. McCall ed., *Polish Logic* (Oxford: Oxford University Press), pp. 19–39.
MacBride, F. (1998), Where are Particulars and Universals? *Dialectica* 52: 203–227.
MacBride, F. (2004), Whence the Particular-Universal Distinction? *Grazer Philosophische Studien* 67: 181–194.
MacBride, F. (2005), The Particular-Universal Distinction: A Dogma of Metaphysics? *Mind* 114: 565–614.
Mach, E. (1960), *The Science of Mechanics: A Critical and Historical Account of its Development* (LaSalle, IL: Open Court).
Mackie, J.L. (1973), *Truth, Probability, and Paradox* (Oxford: Oxford University Press).
Mackie, J.L. (1977), *Ethics: Inventing Right and Wrong* (London: Penguin Books).
Mackie, J.L. (1983), Three Steps toward Absolutism, in R. Swinburne ed., *Space, Time and Causality* (Dordrecht: D. Reidel), pp. 3–22.
Mackie, P. and Jago, M. (2013), Transworld Identity, in Edward N. Zalta ed., *The Stanford Encyclopedia of Philosophy* (Fall 2013 Edition). Available at http://plato.stanford.edu/archives/fall2013/entries/identity-transworld/(accessed 5 June 2016).
Malcolm, N. (1989), Wittgenstein on Language and Rules, *Philosophy* 64: 5–28.
Manley, D. (2002), Properties and Resemblance Classes, *Noûs* 36: 75–96.
Mantegani, N. (2010), *Toward a Quinean Realism*, doctoral dissertation, University of Texas at Austin.
Markosian, N. (1992), On Language and the Passage of Time, *Philosophical Studies* 66: 1–26.
Markosian, N. (1993), How Fast does Time Pass? *Philosophy and Phenomenological Research* 53: 829–844.
Markosian, N. (1998), Brutal Composition, *Philosophical Studies* 92: 211–249.
Martin, C.B. (1994), Dispositions and Conditionals, *Philosophical Quarterly* 44: 1–8.

McCall, S. (1976), Objective Time Flow, *Philosophy of Science* 43: 337–362.

McCall, S. (1994), *A Model of the Universe* (Oxford: Clarendon Press).

McDaniel, K. (2004), Modal Realism with Overlap, *Australasian Journal of Philosophy* 82: 137–152.

McDaniel, K. (2010), Composition as Identity Does not Entail Universalism, *Erkenntnis* 73: 97–100.

McDowell, J. (1994), *Mind and World* (Cambridge, MA: Harvard University Press).

McTaggart, J.M.E. (1908), The Unreality of Time, *Mind* 18: 457–484.

Menzel, C. (1991), The True Modal Logic, *Journal of Philosophical Logic* 20: 331–374.

Merricks, T. (2001), Varieties of Vagueness, *Philosophy and Phenomenological Research* 62: 145–157.

Merricks, T. (2003), *Objects and Persons* (Oxford: Oxford University Press).

Merricks, T. (2007), *Truth and Ontology* (Oxford: Clarendon Press).

Meyer, R.K. (1987), God Exists! *Noûs* 21 (3): 345–361.

Mill, J.S. (1865/1963), An Examination of Sir William Hamilton's Philosophy, in J.M. Robson ed., *Collected Works*, Vol. 9 (Toronto: University of Toronto Press).

Molnar, G. (2003), *Powers: A Study in Metaphysics* (Oxford: Oxford University Press).

Moore, G.E. (1903a), *Principia Ethica* (New York: Cambridge University Press).

Moore, G.E. (1903b/1993), The Refutation of Idealism, in T. Baldwin ed., *Selected Writings* (London: Routledge), pp. 23–44.

Moore, G.E. (1939), Proof of an External World, *Proceedings of the British Academy* 25, 273–300. Reprinted in *G. E. Moore: Selected Writings*, ed. Thomas Baldwin (London: Routledge, 1993), pp. 147–170.

Moreland, J.P. (1998), Theories of Individuation: A Reconsideration of Bare Particulars, *Pacific Philosophical Quarterly* 79: 251–263.

Nagel, E. (1961), *The Structure of Science. Problems in the Logic of Explanation* (New York: Harcourt, Brace & World, Inc.).

Newton, I. (1687/1999), *The Principia: Mathematical Principles of Natural Philosophy*, I.B. Cohen and A. Whitman, trans. (Berkeley, CA: University of California Press).

Nozick, R. (1969), Newcomb's Problem and Two Principles of Choice, in N. Rescher ed., *Essays in Honor of Carl G. Hempel*, Synthèse Library (Dordrecht: D. Reidel), pp. 114–115.

O'Leary-Hawthorne, J. and Cortens, A. (1995), Towards Ontological Nihilism, *Philosophical Studies* 79: 143–165.

Oaklander, L.N. (1994), On the Experience of Tenseless Time, in L.N. Oaklander and Q. Smith eds., *The New Theory of Tim* (New Haven, CT: Yale University Press), pp. 344–350.

Oliver, A. (1996), The Metaphysics of Properties, *Mind* 105: 1–80.

Paśniczek, J. (1995), Are Contradictions Still Lurking in Meinongian Theories of Objects? *Grazer Philosophische Studien* 50: 293–303.

Parfit, D. (1985), *Reasons and Persons* (Oxford: Clarendon Press).

Parsons, J. (1999), There is no 'Truthmaker' Argument against Nominalism, *Australasian Journal of Philosophy* 77: 325–334.

Parsons, J. (2004), Distributional Properties, in F. Jackson and G. Priest eds., *Lewisian Themes* (Oxford: Oxford University Press), pp. 173–180.

Parsons, T. (1980), *Nonexistent Objects* (New Haven, CT: Yale University Press).

Paul, L.A. (2004), Aspect Causation, in J. Collins, N. Hall, and L.A. Paul (2004), pp. 205–224.

Pautz, A. (2006), Can the Physicalist Explain Colour Structure in Terms of Colour Experience? *Australasian Journal of Philosophy* 83: 535–564.

Pautz, A. (2007), Intentionalism and Perceptual Presence, *Philosophical Perspectives* 21 (1): 495–541.

Pautz, A. (2010), A Simple View of Consciousness, in R.C. Koons and G. Bealer eds., *The Waning of Materialism* (Oxford: Oxford University Press), pp. 26–66.

Pautz, A. (2014), Colour, Philosophical Aspects of, in T. Bayne, A. Cleeremans, and P. Wilken eds., *Oxford Companion to Consciousness* (Oxford: Oxford University Press), pp. 150–154.

Pearl, J. (2009), *Causality: Models, Reasoning and Inference* (Cambridge: Cambridge University Press).

Perry, J. (1979), The Problem of the Essential Indexical, *Noûs* 13: 3–21.

Perszyk, K.J. (1993), *Nonexistent Objects: Meinong and Contemporary Philosophy* (Dordrecht: Kluwer Academic).

Pickavance, T.H. (2014), Bare Particulars and Exemplification, *American Philosophical Quarterly* 51: 95–108.

Pickavance, T.H. (forthcoming), The Natural View of Properties of Identity, in R. Garcia ed., *Substance* (Munich: Philosophia Verlag).

Pickavance, T.H. (unpublished), The Way of Wisdom Toward the Universal-Particular Distinction.

Pickel, B. (2010), *Reduction, Ontology and the Limits of Convention*, doctoral dissertation, University of Texas at Austin.

Pickel, B. and Mantegani, N. (2012), A Quinean Critique of Ostrich Nominalism, *Philosophers' Imprint* 12 (6): 1–21.

Plantinga, A. (1974), *The Nature of Necessity* (Oxford: Clarendon Press).

Plato (1997), *Complete Works*, ed. John M. Cooper (Indianapolis: Hackett).

Price, H. (1996) *Time's Arrow and Archimedes' Point* (Oxford: Oxford University Press).

Priest, G. (2005), *Towards Non-Being. The Logic and Metaphysics of Intentionality* (Oxford: Clarendon Press).

Prior, A.N. (1949), Determinables, Determinates and Determinants, *Mind* 48: 178–194.

Prior, A.N. (1957), *Time and Modality* (Oxford: Clarendon Press).

Prior, A.N. (1959), Thank Goodness That's Over, *Philosophy* 34: 12–17.

Prior, A.N. (1967), *Past, Present, and Future* (Oxford: Clarendon Press).

Prior, A.N. (1968), *Papers on Time and Tense* (London: Oxford University Press).

Prior, E., Pargetter, R., and Jackson, F. (1982), Three Theses about Dispositions, *American Philosophical Quarterly* 19: 251–257.

Pruss, A.R. (2002), The Actual and the Possible, in R.M. Gale ed., *Blackwell Guide to Metaphysics* (Oxford: Blackwell).

Pruss, A.R. (2003), David Lewis's Counterfactual Arrow of Time, *Noûs* 37: 606–637.

Pruss, A.R. (2009), From the Grim Reaper Paradox to the Kalaam Argument, Available at http://alexanderpruss.blogspot.com/ (last accessed 5 June 2016).

Pruss, A.R. (2010), Presentism and Induction, Available at http://alexanderpruss.blogspot.com/2010/06/presentism-and-induction.html (last accessed 5 June 2016).

Pruss, A.R. (2011), *Actuality, Possibility, and Worlds* (New York: Continuum).

Putnam, H. (1975), *Mathematics, Matter and Method* (Cambridge: Cambridge University Press).

Quine, W.V.O. (1953/1980), *From a Logical Point of View* (Cambridge, MA: Harvard University Press).

Quine, W.V.O. (1957), The Scope and Language of Science, *British Journal for the Philosophy of Science* 8: 1–17.

Quine, W.V.O. (1960), *Word and Object* (Cambridge, MA: MIT Press).

Ramsey, F.P. (1925), Universals, *Mind* 34: 401–417.

Ramsey, F.P. (1927), Facts and Propositions, *Proceedings of the Aristotelian Society* 7 (Supplementary): 153–170.

Ramsey, F.P. (1928/1978), *Foundations* (London: Routledge and Kegan Paul).

Rea, M.C. (1996), *Material Constitution: A Reader* (Lanham, MD: Rowman & Littlefield).

Reed, L. (1924), *The Complete Limerick Book* (London: Jerrolds).

Reichenbach, H. (1958), *The Philosophy of Space and Time* (New York: Dover).

Reid, T. (1785), *Essay on the Intellectual Powers of Man* (Dublin: Bell and Robinson).

Ritchie, K. (2013), What are Groups? *Philosophical Studies* 166 (2): 257–272.

Ritchie, K. (2015), The Metaphysics of Social Groups, *Philosophy Compass* 10 (5): 310–321.

Rodriguez-Pereyra, G. (2002), *Resemblance Nominalism: A Solution to the Problem of Universals* (Oxford: Clarendon Press).

Rodriguez-Pereyra, G. (2004), The Bundle Theory is Compatible with Distinct but Indiscernible Substances, *Analysis* 64: 72–81.

Rodriguez-Pereyra, G. (2006), Truthmaking, Entailment, and the Conjunction Thesis, *Mind* 115: 957–981.

Rosen, G. (2010), Metaphysical Dependence: Grounding and Reduction, in B. Hale and A. Hoffman eds., *Modality: Metaphysics, Logic and Epistemology* (New York: Oxford University Press), pp. 109–136.

Routley, R. (1979), *Exploring Meinong's Jungle and Beyond: An Investigation of Noneism and the Theory of Items* (Canberra: Australian National University Press).

Routley, R. (1980), Exploring Meinong's Jungle, *Notre Dame Journal of Formal Logic* 21: 480.

Routley, R. (1983), Exploring Meinong's Jungle and Beyond, *Journal of Philosophy* 80 (3): 173–179.

Russell, B. (1903), *The Principles of Mathematics, I* (Cambridge: Cambridge University Press).

Russell, B. (1905), On Denoting, *Mind* 14 (56): 479–493.

Russell, B. (1911), On the Relations of Universals and Particulars, *Proceedings of the Aristotelian Society* (New Series) 12: 1–24.

Russell, B. (1912), The World of Universals, *The Problems of Philosophy* (Oxford: Oxford University Press).

Russell, B. (1913), On the Notion of Cause, *Proceedings of the Aristotelian Society* 13: 1–26.

Russell, B. (1917), *Mysticism and Logic* (London: George Allen and Unwin).

Russell, B. (1918–1919), The Philosophy of Logical Atomism, *The Monist*, 28: 495–527; 29: 190–222, 345–380.

Russell, B. (1927), *The Analysis of Matter* (London: Kegan Paul, Trench, Trubner).

Russell, B. (1948), *Human Knowledge* (New York: Simon and Schuster).

Russell, B. (1959), *My Philosophical Development*, London and New York: Routledge.

Russell, B. (1984), *Theory of Knowledge: The 1913 Manuscript* (London: George Allen & Unwin).

Russell, B. and Whitehead, A.N. (1962), *Principia Mathematica to *56* (New York: Cambridge University Press).

Ryle, G. (1949), *The Concept of Mind* (New York: Barnes and Noble).

Salmon, N. (2010), Vagaries about Vagueness, in R. Dietz and S. Moruzzi eds., *Cuts and Clouds: Vagueness, Its Nature and Its Logic* (Oxford: Oxford University Press).

Salmon, W. (1984), *Scientific Explanation and the Causal Structure of the World* (Princeton, NJ: Princeton University Press).
Salmon, W. (1998), *Causality and Explanation* (Oxford: Oxford University Press).
Santayana, G. (1955), Ultimate Skepticism, *Skepticism and Animal Faith: Introduction to a System of Philosophy* (New York: Dover), pp. 33–41.
Schaffer, J. (2000), Trumping Preemption, *Journal of Philosophy* 9: 165–181.
Schaffer, J. (2001), Causes as Probability-Raisers of Processes, *Journal of Philosophy* 98: 75–92.
Schaffer, J. (2004), Causes Need Not Be Physically Connected to their Effects: The Case for Negative Causation, in C. Hitchcock ed., *Contemporary Debates in Philosophy of Science* (Oxford: Blackwell), pp. 197–216.
Schaffer, J. (2005), Quidditistic Knowledge, *Philosophical Studies* 123: 1–32.
Schaffer, J. (2009), On What Grounds What, in D. Manley, D.J. Chalmers, and Ryan Wasserman eds., *Metametaphysics: New Essays on the Foundations of Ontology* (Oxford: Oxford University Press), pp. 347–384.
Schaffer, J. (2010), Monism: The Priority of the Whole, *Philosophical Review* 119: 31–76.
Schaffer, J. (2014), What Not to Multiply without Necessity, *Australasian Journal of Philosophy* 92(4): 644–664.
Schaffer, J. (2016), Grounding in the Image of Causation, *Philosophical Studies* 173: 49–100.
Schlesinger, G. (1980), *Aspects of Time* (Indianapolis: Hackett Publishing).
Scotus, J.D. (1307/1987), *Philosophical Writings: A Selection*, ed. Allan Wolter (Indianapolis: Hackett).
Searle, J. (1959), Determinables and the Notion of Resemblance, *Proceedings of the Aristotelian Society, Supplementary Volume* 33: 141–158.
Sellars, W. (1962), Philosophy and the Scientific Image of Man, in R. Colodny, ed. *Frontiers of Science and Philosophy* (Pittsburgh: University of Pittsburgh Press), pp. 35–78. Reprinted in *Science, Perception and Reality* (1963).
Sellars, W. (1963), *Science, Perception and Reality* (London: Routledge & Kegan Paul).
Shapiro, S. (1998), Proof and Truth: Through Thick and Thin, *Journal of Philosophy* 95: 493–521.
Shoemaker, S. (1980) Causality and Properties, in P. van Inwagen, ed., *Time and Cause: Essays Presented to Richard Taylor* (Dordrecht: D. Reidel Publishing), pp. 109–135.
Shoemaker, S. (1984), *Identity, Cause, and Mind* (Cambridge: Cambridge University Press).
Shoemaker, S. (1998) Causal and Metaphysical Necessity, *Pacific Philosophical Quarterly* 79: 59–77.
Sider, T. (1999), Presentism and Ontological Commitment, *Journal of Philosophy* 96: 325–347.
Sider, T. (2001), *Four Dimensionalism: An Ontology of Persistence and Time* (Oxford: Clarendon Press).
Sider, T. (2002), The Ersatz Pluriverse, *The Journal of Philosophy* 99: 279–315.
Sider, T. (2006), Bare Particulars, *Philosophical Perspectives* 20: 387–397.
Sider, T. (2007), Parthood, *Philosophical Review* 116 (1): 51–91.
Sider, T. (2011), *Writing the Book of the World* (Oxford: Oxford University Press).
Simmons, K. (1993), *Universality and the Liar* (Cambridge: Cambridge University Press).
Simons, P. (1994), Particulars in Particular Clothing: Three Trope Theories of Substance, *Philosophy and Phenomenological Research* 54: 553–575.
Sklar, L. (1974), *Space, Time and Spacetime* (Berkeley: University of California Press).
Skyrms, B. (1980), *Causal Necessity* (New Haven, CT: Yale University Press).
Skyrms, B. (1981), Tractarian Nominalism, *Philosophical Studies* 40: 199–206.

Skyrms, B. (1983), Zeno's Paradox of Measure in R.S. Cohen and L. Laudan eds., *Physics, Philosophy and Psychoanalysis* (Dordrecht: D. Reidel), pp. 223–254.

Smart, J.J.C. (1949), The River of Time, *Mind* 58: 483–494.

Smith, A.D. (1977), Dispositional Properties, *Mind* 86: 439–445.

Sorenson, R. (1984), We See in the Dark, *Nous* 38: 456–480.

Sorenson, R. (2009), Hearing Silence: the Perception and Introspection of Absences, in M. Nudds and C. O'Callaghan eds., *Sounds and Perception: New Philosophical Essays* (Oxford: Oxford University Press).

Spinoza, B. (1677a/1958a), *Ethics, Spinoza: Selections*, ed. J. Wild (New York: Scribner's and Sons), pp. 94–400.

Spinoza, B. (1677b/1958b), Short Treatise on God, Man and His WellBeing, *Spinoza: Selections*, ed. J. Wild (New York: Scribner's and Sons), pp. 45–93.

Spirtes, P., Glymour, C., and Scheines, R. (1993), *Causation, Prediction and Search* (Berlin: Springer Verlag).

Stalnaker, R. (1968), A Theory of Conditionals, in N. Rescher (ed.), *Studies in Logical Theory*, American Philosophical Quarterly Monograph Series, Vol. 2 (Oxford: Blackwell), pp. 98–112.

Stoljar, D. (2015), Physicalism, *Stanford Encyclopedia of Philosophy*. Available at http://plato.stanford.edu/entries/physicalism.

Strawson, P.F. (1959), *Individuals* (London: Methuen).

Tarski, A. (1956), Foundations of the Geometry of Solids, in J.H. Woodger, ed. *Logic, Semantics, Metamathematics*, (Oxford: Clarendon Press).

Tarski, A. (1956/1983), *Logic, Semantics, Metamathematics*, trans. J.H. Woodger, ed. J. Corcoran (Indianapolis: Hackett, second edition).

Taylor, R. (1991), *Metaphysics* (Englewood Cliffs, NJ: Prentice-Hall, fourth edition).

Thomson, J.F. (1954), Tasks and Super-Tasks, *Analysis* 15: 1–13.

Tooley, M. (1977), The Nature of Laws, *Canadian Journal of Philosophy* 7: 667–698.

Tooley, M. (1987), *Causation: A Realist Approach* (Oxford: Clarendon Press).

Tooley, M. (1988), In defense of the existence of states of motion, *Philosophical Topics* 16 (1): 225–254.

Tooley, M. (1990), Causation: Reduction vs. Realism, *Philosophy and Phenomenological Research* 50(Supplement): 215–236.

Tooley, M. (1997), *Time, Tense and Causation* (Oxford: Clarendon Press).

Trogdon, Kelly (2013), Grounding: Necessary or Contingent? *Pacific Philosophical Quarterly* 94: 465–485.

Tye, M. (2000), *Consciousness, Color, and Content* (Cambridge, MA: MIT/Bradford).

Unger, P. (1980), The Problem of the Many, *Midwest Studies in Philosophy* 5: 411–467.

van Fraassen, B. (1969), Presuppositions, Supervaluations and Free Logic, in K. Lambert, ed. *The Logical Way of Doing Things*, (New Haven, CT: Yale University Press), pp. 67–91.

van Fraassen, B. (1987), Armstrong on Laws and Probabilities, *Australasian Journal of Philosophy* 65: 243–259.

van Fraassen, B. (1989), *Laws and Symmetry* (Oxford: Clarendon Press).

van Inwagen, P. (1980), *Time and Cause: Essays Presented to Richard Taylor* (Dordrecht: D. Reidel).

van Inwagen, P. (1986a), Two Concepts of Possible Worlds, *Midwest Studies in Philosophy* 11: 185–213. Reprinted in P. van Inwagen, *Ontology, Identity, and Modality: Essays in Metaphysics* (Cambridge: Cambridge University Press).

van Inwagen, P. (1986b), *An Essay on Free Will* (Oxford: Oxford University Press).

van Inwagen, P. (1990a), *Material Beings* (Ithaca, NY: Cornell University Press).
van Inwagen, P. (1990b), Four–dimensional Objects, *Nous* 24: 245–255. Reprinted in P. van Inwagen, *Ontology, Identity, and Modality: Essays in Metaphysics* (Cambridge: Cambridge University Press).
van Inwagen, P. (1994), Composition as Identity, in J.E. Tomberlin ed., *Philosophical Perspectives, 8: Logic and Language* (Ridgeview: Atascadero), pp. 207–220.
van Inwagen, P. (2001), *Ontology, Identity, and Modality: Essays in Metaphysics* (Cambridge: Cambridge University Press).
von Wright, G. (1971), *Explanation and Understanding* (Ithaca, NY: Cornell University Press).
Whitehead, A.N. (1919), *An Enquiry into the Principles of Natural Knowledge* (Cambridge: Cambridge University Press).
Wiggins, D. (1980), *Sameness and Substance* (Cambridge, MA: Harvard University Press).
Wiggins, D. (2001), *Sameness and Substance Renewed* (Cambridge: Cambridge University Press).
Williams, B. (1973), *Problems of the Self* (Cambridge: Cambridge University Press).
Williams, D.C. (1986), Universals and Existents, *Australasian Journal of Philosophy* 64: 1–14.
Williamson, T. (1994), *Vagueness* (London: Routledge).
Williamson, T. (1996), The Necessity and Determinateness of Distinctness, in D. Wiggins, S. Lovibond, and S. G. Williams eds., *Essays for David Wiggins: Identity, Truth and Value* (Oxford: Oxford University Press).
Williamson, T. (1998), *Bare Possibilia, Erkenntnis* 48: 257–273.
Williamson, T. (2000), *Knowledge and its Limits* (New York: Oxford University Press).
Wilson, J. (2014), No Work for A Theory of Grounding, *Inquiry* 57 (5–6): 535–579.
Winnie, J. (1970), Special Relativity without One-Way Velocity Assumptions, *Philosophy of Science* 37: 222–238.
Wisdom, J. (1934), *Problems of Mind and Matter* (Cambridge: Cambridge University Press).
Wittgenstein, L. (1921/1974), *Tractatus Logico-Philosophicus*, ed. D. Pears and B. McGuinness (New York: Routledge).
Wittgenstein, L. (1953), *Philosophical Investigations*, ed. G.E.M. Anscombe and R. Rhees, trans. G.E.M. Anscombe (Oxford: Blackwell).
Wright, B. (2012), Darkness Visible? *Australasian Journal of Philosophy* 90 (1): 39–55.
Zahar, E. (1983), Absoluteness and Conspiracy, in Richard Swinburne ed., *Space, Time and Causality* (Dordrecht: D. Reidel), pp. 37–41.
Zalta, E. (1983), *Abstract Objects: An Introduction to Axiomatic Metaphysics* (Dordrecht: D. Reidel).
Zalta, E. (1988), *Intensional Logic and the Metaphysics of Intentionality* (Cambridge, MA: Bradford Books/MIT Press).
Zimmerman, D. (1996a), Could Extended Objects be Made Out of Simple Parts? An Argument for 'Atomless Gunk', *Philosophy and Phenomenological Research* 56: 1–29.
Zimmerman, D. (1996b), Indivisible Parts and Extended Objects: Some Philosophical Episodes from Topology's Prehistory, *The Monist* 79: 148–180.
Zimmerman, D. (1999), One Really Big Liquid Sphere: A Reply to Lewis, *Australasian Journal of Philosophy* 77: 213–215.

# Index



*The Atlas of Reality: A Comprehensive Guide to Metaphysics*, First Edition.
Robert C. Koons and Timothy H. Pickavance.