17 Computers and science fiction – an essay

> No one saw these mice coming. No one, that is, in my field, writing science fictions. Oh, a few novels were written about those Big Brains, a few New Yorker cartoons were drawn showing those immense electric craniums that needed whole warehouses to THINK in. But no one in all of future writing foresaw those big brutes dieted down to fingernail earplug size so you could shove Moby Dick in one ear and pull Job and Ecclesiastes out the other.
>
> Ray Bradbury[1]

Early visions

The British science fiction writer, Brian Aldiss, traces the origin of science fiction to Mary Shelley's *Frankenstein* in 1818. In her book, the unwise scientist, Victor Frankenstein, deliberately makes use of his knowledge of anatomy, chemistry, electricity, and physiology to create a living creature. An alternative starting point dates back to the second half of the nineteenth century with the writing of Jules Verne (B.17.1) and Herbert George (H. G.) Wells (B.17.2). This was a very exciting time for science – in 1859 Charles Darwin had published the *Origin of Species*; in 1864 James Clerk Maxwell had unified the theories of electricity and magnetism; and in 1869 Mendeleev had brought some order to chemistry with his *Periodic Table of the Elements*, and Joule and Kelvin were laying the foundations of thermodynamics. Verne had the idea of combining modern science with an adventure story to create a new type of fiction. After publishing his first such story "Five Weeks in a Balloon" in 1863, he wrote:

> I have just finished a novel in a new form, a new form – do you understand? If it succeeds, it will be a gold mine.[2]

And a gold mine it turned out to be. In the next decade Verne published his highly successful series of *Voyages Extraordinaire – Journey to the Center of the Earth*, *From the Earth to the Moon*, *Twenty Thousand Leagues under the Sea*, and *Around the World in Eighty Days*. Verne's stories were grounded in scientific fact and set in the present or near future. Wells's approach to his *scientific romances* was in contrast to that of Verne. In 1895, Wells's novel, *The Time Machine*, has the scientist-inventor traveling through time, exploring a rather bleak future for humanity.

B.17.1. Jules Verne (1828–1905) was one of the founding fathers of the science fiction genre. This is the cover page of *L'Algerie* magazine from 15 June 1884, announcing his latest novel.

Two years later, in 1897, *The Invisible Man* has the researcher discovering a technique to make himself invisible – but is unable to reverse the process, leading to dire consequences. And in 1898, *The War of the Worlds* has alien invaders from Mars landing in London and defeating all of England's military might with deadly heat-ray weapons. The runaway success of these stories, produced in just three years, has led to Wells being acclaimed as the "Father of Science Fiction."

The nineteenth century of Verne and Wells was generally a time of optimism and of belief in the self-evident benefits of advances in science and technology. One early exception to this utopian view of the future was a short story by the novelist E. M. Forster. In 1909, he wrote *The Machine Stops*, which put forward a much more pessimistic, dystopian vision. In the story, humanity has become totally dependent on a vast and complex machine – which we would now call a computer – that manages every aspect of a completely mechanized environment for society. The story centers on a woman, Vashti, and her son Kuno. Vashti is content to live her life in her underground, automated apartment, attending meetings and lectures by videoconference, and only rarely physically meeting people:

> Then she generated the light, and the sight of her room, flooded with radiance and studded with electric buttons, revived her. There were buttons and switches everywhere – buttons to call for food, for music, for clothing. There was the hot-bath button, by pressure of which a basin of (imitation) marble rose out of the floor, filled to the brim with a warm deodorized liquid. There was the cold-bath button. There was the button that produced literature, and there were of course the buttons by which she communicated with her friends. The room, though it contained nothing, was in touch with all that she cared for in the world.[3]

Kuno rebels against the uniformity of the society and tries unsuccessfully to go outside, into the open air. Total dependence on the Machine had almost become a religion. Unfortunately, the number of engineers who understood the technical workings of the Machine in its totality was decreasing every year. Eventually, increasing numbers of small defects became common – the music feeds became unreliable, the bath began to smell – and these were not immediately fixed by the Committee of Mending Apparatus. Finally, the Machine

B.17.2. Herbert George Wells (1866–1946), better known as H. G. Wells, was inspired by Darwin's theory of evolution by natural selection as a student. The theme of humanity evolving according to these inexorable forces is evident in his first and probably most famous book, *The Time Mach*ine, published in 1895, as well as in many of his later works. He is often credited with foreseeing the development of the tank, aircraft and air warfare, the atomic bomb, and nuclear stalemate. Hugo Gernsback reprinted almost all of Wells's novels in his *Amazing Stories* magazine and as a result Wells's work was enormously influential on the development of science fiction in the United States. His book *The War of the Worlds* was the first alien invasion story, and Orson Welles's radio dramatization of it caused a riot in New York in 1938. Toward the end of his life he had the idea of a distributed encyclopedia that he called the World Brain.

Fig. 17.1. The robot Wall-E was left behind to clean up the Earth's garbage after the humans had departed in spaceships. The movie *Wall-E* portrayed a world in which humanity had become entirely dependent on technology for their every need.

stops and with it, the dependent civilization. The movie *Wall-E* (Fig. 17.1) is a modern incarnation of this vision of humanity becoming overdependent on machines.

In his later life, Wells wrote several nonfiction books, including two major historical surveys – *Outline History* in 1920, and *The Work, Wealth and Happiness of Mankind* in 1932. In the course of writing these books, Wells realized the need for authors to have "information at their fingertips." He believed that combating ignorance by making information readily available to the masses would reduce the likelihood of war. In 1937 he campaigned for the creation of a "World Brain" that would contain, and continually update, the knowledge contained in all the world's great libraries, museums, and universities:

> A World Encyclopaedia no longer represents itself to a modern imagination as a row of volumes printed and published once for all, but as a sort of mental clearing house for the mind, a depot where knowledge and ideas are received, sorted, summarized, digested, clarified and compared. It would be in continual correspondence with every university, every research institution, every competent discussion, every survey, every statistical bureau in the world. It would develop a directorate and a staff of men of its own type, specialized editors and summarists. They would be very important and distinguished men in the new world. The Encyclopaedic organization need not be concentrated now in one place; it might have the form of a network. It would centralize mentally but not physically.… It would constitute the material beginning of a real World Brain.[4]

Vannevar Bush and J. C. R. Licklider may have laid the foundations for today's World Wide Web and the Internet, but Wells was one of the first evangelists for these ideas. With the Internet, the web, and Wikipedia, Wells would have been delighted by the progress that has been made toward his vision of a World Brain. Sadly, Wells was never able to realize the World Brain vision in his lifetime. After the terrible carnage of World War II, Wells descended into pessimism: his last book, *Mind at the End of Its Tether*, was written in 1945.

Wells never wrote about computers but Jules Verne did in *Paris in the Twentieth Century*, a novel published in 1994, a hundred and thirty years after it was written. The Paris of Verne's novel had some remarkable insights – great avenues filled with horseless carriages, elevated railways and driverless trains, electric lighting in the shops and streets, and mechanical elevators in all the great buildings. His editor rejected the manuscript with the words "No one today will believe your prophecy"![5] Michel, the hero of the novel, is an apprentice at one of the huge banks that dominated the financial landscape:

> Michel turned around and discovered the calculating machine behind him. It had been several centuries since Pascal had constructed a device of this kind, whose conception had seemed so remarkable at the time.… The Casmodage Bank possessed veritable masterpieces of the genre, instruments which indeed did resemble huge pianos: by operating a sort of keyboard, sums were instantaneously produced, remainders, products, quotients, rules of proportion, calculations of amortization and of interest compounded for infinite periods and all possible rates.[6]

Fig. 17.2. The cover of the April 1926 issue of *Amazing Stories*. It contained reprints of stories by H. G. Wells, Jules Verne, and Edgar Allan Poe.

The bank was also connected by telegraph to other institutions according to the Wheatstone telegraph system long since in use throughout England, and stock prices were continually updated from stock markets all around the world.

> Further, photographic telegraphy, invented in the last century by Professor Giovanni Caselli of Florence, permitted transmission of the facsimile of any form of writing or illustration, whether manuscript or print, and letters of credit or contracts could now be signed at a distance of five thousand leagues.[7]

The Wheatstone telegraph system was in widespread use in Britain by the 1850s but Professor Giovanni Caselli had only just invented his facsimile *pantelgraph* system for transmitting images over telegraph lines. It was in 1862, one year before Verne wrote his book, that the first *pantelegram* had been sent from Lyon to Paris.

Computers and hard SF – the early years

It was the *scientific romances* of Wells and Verne that provided the inspiration for the first popular magazine in the United States dedicated to science fiction (SF). *Amazing Stories* was founded in 1926 by Hugo Gernsback (B.17.3). The first issue set the direction for the new magazine by republishing stories from Verne and Wells (Fig. 17.2). Gernsback was first to use the term *science fiction* to describe what he called "a charming romance intermingled with scientific fact and prophetic vision."[8] During the 1930s and 1940s, there was a great increase in the number of science fiction magazines but the most influential of all of these was undoubtedly *Astounding Science Fiction*. Its editor, John W. Campbell Jr. (B.17.4) had been a student at MIT and been taught by Norbert Wiener. Campbell's own short story "The Last Evolution," published in *Amazing Stories* in 1932, envisioned a future where man and machines fought together against an invasion from outer space. We can recognize aspects of modern computers in his description of the machines:

> Machines – with their irrefutable logic, their cold preciseness of figures, their tireless, utterly exact observation, their absolute knowledge of mathematics – they could elaborate any idea, however simple its beginning, and reach the conclusion.[9]

As editor, Campbell set about raising the standard of writing in *Astounding*, training a whole generation of science fiction writers, including Asimov and Robert Heinlein. Although Campbell rejected Asimov's first story, his rejection letter inspired Asimov to do better:

> The joy of having spent an hour or more with John Campbell, the thrill of talking face to face and on even terms with an idol, had already filled me with the ambition to write another science fiction story, better than the first, so that I could try him again. The pleasant letter of rejection – two full pages – in which he discussed my story seriously and with no trace of patronization or contempt, reinforced my joy.[10]

B.17.3. Hugo Gernsback (1884–1967) was born in Luxembourg and emigrated to America when he was twenty. He published his novel *Ralph 124C 41+*, subtitled *A Romance of the Year 2660*, in his magazine *Modern Electrics* in 1911. In 1926, Gernsback launched *Amazing Stories*, the first magazine to be exclusively devoted to science fiction. The slogan on the title page proclaimed its mission "Extravagant Fiction Today … Cold Fact Tomorrow." Gernsback invented the term *science fiction* and the annual science fiction Hugo awards are named in his honor.

B.17.4. John Wood Campbell Jr. (1910–71) was indisputably the greatest editor of science fiction and nurtured a whole generation of science fiction writers including Isaac Asimov and Robert Heinlein. In 1938 he became editor of the magazine *Astounding Science Fiction* and, by his insistence on a much higher standard of writing and his help and support for writers like Asimov, created the modern science fiction genre. Under the pseudonym of Don A. Stuart he also wrote science fiction. Isaac Asimov rated Campbell's story "Who Goes There?," published in 1938, as "one of the very best science fiction stories ever written." (Photo courtesy of Marsh Library.)

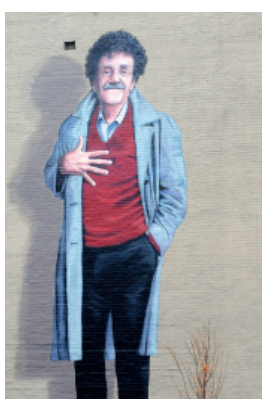
Campbell was editor of *Astounding Science Fiction* and its successor, *Analog Science Fact – Science Fiction*, for more than thirty years. In a 1946 editorial, he gave a definition of what is now called *hard SF*:

> Science fiction is written by technically minded people, about technically minded people, for the satisfaction of technically minded people.[11]

In the early days of computers, after World War II, the public learned about the awesome calculating power of the ENIAC. Soon after this, Ekert and Mauchly launched the first U.S. commercial computer – UNIVAC. During the 1952 U.S. presidential election, UNIVAC shot to fame by appearing on TV and, with only a fraction of results reported, correctly predicted a landslide win for the preelection underdog, Dwight D. Eisenhower.

Given the enormous scale of these early vacuum tube computers it was natural for science fiction writers to equate the power of a computer with its size. Kurt Vonnegut (B.17.5) published a short story in 1950 about a computer called "EPICAC," the largest and smartest computer on the planet:

> EPICAC covered about an acre on the fourth floor of the physics building at Wyandotte College. Ignoring his spiritual side for a minute, he was seven tons



B.17.5. Kurt Vonnegut (1922–2007) was a U.S. novelist who used science fiction techniques in many of his novels. His novel *Player Piano* explores the theme of computers making workers redundant, a scenario that is even more relevant today. This photograph of the Vonnegut mural in Indianapolis was created by the artist Pamela Bliss. Vonnegut's face in the mural is a composite image based on three different photos.

> of electronic tubes, wires, and switches, housed in a bank of steel cabinets and plugged into a 110-volt A.C. line just like a toaster or a vacuum cleaner.[12]

B.17.6. Isaac Asimov (1920–92) was one of the most prolific science fiction authors of the twentieth century. His robot stories and his Three Laws of Robotics have defined the robot science fiction genre. A little-known fact about him is that he was also a professor of biochemistry.

The machine had been designed to do the same sort of tasks for the military as the ENIAC but in the story, EPICAC is sentient and falls in love with one of the mathematicians. In the end the machine blows itself up, leaving a suicide note and some love poems as a wedding gift for his human rival.

Isaac Asimov (B.17.6) also assumed that the most powerful computers of the future would be bigger than the ENIAC.

> It wasn't until after computers were invented and the public was made aware of their existence, that computers began to exist in my stories, and even then I didn't truly conceive of the possibility of miniaturization.… In "The Last Question" I began with my usual computer, Multivac, as large as a city, for I could only conceive a larger computer by imagining more and more vacuum tubes heaved into it.[13]

In his 1955 short story "Franchise," MULTIVAC is a government-run computer half a mile long and three stories high. The story is a satire on the use of UNIVAC to predict the result of the recent presidential election from a small sample of voters. In Asimov's version, the sample size has been reduced to finding the views of the "Voter of the Year" – the single most representative person in the United States. In another MULTIVAC story, "The Machine That Won the War," Asimov reveals some keen insight into the reliability of computers. The official propaganda is all about how the powerful MULTIVAC computer helped win the war for the Solar Federation. Yet when the three men whose job it was to interact with the computer met and discussed their roles, each admitted falsifying part of the computational process. The chief programmer admitted altering the data being fed to MULTIVAC because the people could not be trusted to supply accurate information in the chaos of war. The engineer then confessed to changing the data produced by the computer because he knew that MULTIVAC was not working reliably due to a shortage of manpower and spare parts. Finally, the Executive Director of the Solar Federation revealed that he had not trusted the reports produced by the machine and had made critical strategic war decisions by tossing a coin!

B.17.7. Fred Hoyle (1915–2001) was a celebrated cosmologist who worked at Caltech in the United States and Cambridge in the United Kingdom. He was one of the originators of the Steady State theory alternative to the Big Bang theory to explain the expansion of the universe – which has now been ruled out by experimental measurements. He also wrote several successful science fiction novels such as *The Black Cloud* and co-authored the BBC TV series *A for Andromeda*. The TV series featured Julie Christie in her first starring role.

In 1961, the BBC in Britain broadcast a science fiction TV drama about a computer whose design originated from outer space. *A for Andromeda* was cowritten by TV producer John Elliot and the famous cosmologist Fred Hoyle (B.17.7). Using their newly commissioned radio telescope, researchers detect a signal coming from the Andromeda galaxy. The scientist hero, John Fleming, despite having to contend with constant interference from politicians and the military, is able to decipher the signal and deduce that it contains the instructions for building a new type of advanced computer. Fleming fears that the motivation for those sending the message is ultimately malevolent, but the military and the politicians insist that the computer be built and, as Fleming predicted, bad things happen.

One writer who did envision a future with smaller and more pervasive computers was Murray Leinster (B.17.8). His short story "A Logic Named Joe" appeared in the March 1946 issue of *Astounding Science Fiction*. The story is narrated by a

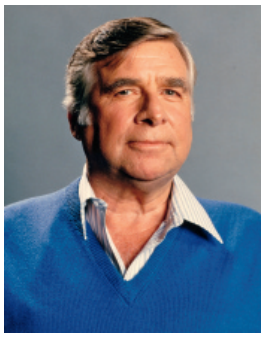
B.17.8. Murray Leinster (1896–1975) was one of the few science fiction writers to envision the widespread availability of networked personal computers. His 1934 novel *Sideways in Time* is credited as being the first "parallel universe" science fiction story. In his 1945 novella *First Contact* Leinster introduced the idea of a *universal translator*, which is still some way from being a reality.

maintenance man from the Logics Company and their "logic computers" had much the same capability as one of today's networked personal computers:

> You know the logics setup. You got a logic in your house. It looks like a vision receiver used to, only it's got keys instead of dials and you punch the keys for what you wanna get. It's hooked in to the tank, which has the Carson Circuit all fixed up with relays. Say you punch "Station SNAFU" on your logic. Relays in the tank take over an' whatever vision-program SNAFU is telecastin' comes on your logic's screen. Or you punch "Sally Hancock's Phone" an' the screen blinks an' sputters an' you're hooked up with the logic in her house an' if somebody answers you got a vision-phone connection. But besides that, if you punch for the weather forecast or who won today's race at Hialeah or who was mistress of the White House durin' Garfield's administration or what is PDQ and R sellin' for today, that comes on the screen too. The relays in the tank do it. The tank is a big buildin' full of all the facts in creation an' all the recorded telecasts that ever was made – an' it's hooked in with all the other tanks all over the country – an' everything you wanna know or see or hear, you punch for it an' you get it. Very convenient. Also it does math for you, an' keeps books, an' acts as consultin' chemist, physicist, astronomer, an' tea-leaf reader, with a "Advice to the Lovelorn" thrown in.[14]

The story is about one particular logic machine named Joe that developed a form of intelligence as a result of a small manufacturing error. By giving accurate and direct answers to questions such as "How can I get rid of my wife?" or "How would I rob my own bank?" Joe causes chaos until he is eventually located and removed from the network.

By the 1960s computers were becoming a more familiar part of the landscape of science fiction and of movies. The original *Star Trek* TV franchise began in 1966 with creator Gene Roddenberry's vision of a universe-spanning, United Nations–like organization, the United Federation of Planets (B.17.9). The federation and its Starfleet command pursued a humanitarian, peace-keeping mission, with a diverse crew of humans and aliens serving in spaceships like Captain Kirk's Starship Enterprise. *Star Trek* was very much in the tradition of "space opera" – in which faster-than-light travel through hyperspace is a necessary device to allow interaction between otherwise impossibly distant galaxies. The focus of each episode was not so much on computers and technology as on dramatic action and situations – novel computing and communication devices were just part of the background to the action. Several of these imagined *Star Trek* devices have now become reality – but not yet the "Beam me up, Scotty" matter transporter! The handheld, flip-top communications devices used by the crew of the Enterprise on their off-ship adventures predated today's mobile phone revolution (Fig. 17.3). The *Star Trek* handheld, *tricorder* device incorporated sensors for data recording and analysis. The tricorder could be used on off-ship expeditions to explore the unfamiliar locations, examine living creatures, and record and review technical data. On ship, Leonard "Bones" McCoy used a medical version of the tricorder to help diagnose medical problems and collect information about his patients. With advances in sensors, digital photography, and the miniaturization of microprocessors and memory, real-life versions of the *Star Trek* tricorder will soon be with us (Fig. 17.4).



B.17.9. Gene Roddenberry (1921–91) is best known as the creator of the hugely successful *Star Trek* franchise. Before turning to screenwriting, Roddenberry had a wide variety of careers, from fighter pilot in the U.S. Air Force to a stint with the Los Angeles Police Department. In 1985 he became the first TV writer with a star on the Hollywood Walk of Fame.

Fig. 17.3. The *Star Trek* communicator device looks similar to some early mobile phones.

Fig. 17.4. A *Star Trek* tricorder and other devices from the *Star Trek* TV series.

Perhaps the most iconic image of computers from the 1960s was that of the HAL 9000 (Fig. 17.5) computer in Stanley Kubrick's 1968 movie *2001: A Space Odyssey*. Kubrick cowrote the screenplay with the science fiction writer Arthur C. Clarke (B.17.10) and the movie was based on Clarke's short story "The Sentinel." Contrary to the widespread belief that the computer was called HAL because it was a one-letter shift from the initials IBM, Kubrick and Clarke always insisted that HAL stood for Heuristically programmed ALgorithmic computer. In the movie, HAL was built in 1992 at the University of Illinois, Urbana – where a real-life early parallel supercomputer called the ILLIAC-4 had been designed. In addition to monitoring and controlling every aspect of the Discovery One spaceship, HAL was sentient and able to perform complex tasks such as speech processing, speech recognition, natural language processing, lip-reading, and facial recognition, as well as playing games like chess. In the movie, with the Discovery on the way to Jupiter, HAL appears to be mistaken in reporting a fault with the ship's communication antenna. In order that HAL cannot overhear, the two astronauts, Dave Bowman and Frank Poole, get into an Evacuation Vehicle on board the ship so that they can discuss what to do about HAL's unreliability. By a very large extrapolation of the lip-reading capabilities of present-day computers – using only a side view of the lips and not a face-on view – HAL is able to "hear" that they are considering disconnecting him. HAL then plans to kill the astronauts and succeeds in killing Poole and three other unfortunate crew members still in cryogenic hibernation. Bowman manages to survive and grimly sets about disconnecting all of HAL's processor modules.

Why did HAL malfunction? The reason is only hinted at in the original movie but is made clearer in the 1984 sequel *2010: Odyssey Two*. When HAL's creator is able to examine the computer he concludes that the crisis was caused by a programming contradiction. HAL was constructed for the accurate processing of information without distortion or concealment but this goal was in conflict with HAL's orders to keep the real reason for the mission secret. HAL's decision to kill the crew was the computer's logical way of resolving the conflict.

The prediction that a sentient computer would exist by 2001 has obviously not been fulfilled but another computer technology in the movie recently made the headlines. In response to Apple's lawsuit against Samsung's tablet computers, the company cited a YouTube video from Kubrick's movie as constituting relevant prior art:

B.17.10. Arthur C. Clarke (1917–2008) (standing) was a prolific inventor and science fiction writer. His early work on radar led him to the idea of communication satellites. The man behind the desk is the film director Stanley Kubrick. The third person in the photograph is Victor Lyndon, who was associate producer of *2001: A Space Odyssey* and also worked on *Dr. Strangelove*. The photo was most likely taken around 1964–5, in the early stages of *Space Odyssey*.

Fig. 17.5. The HAL 9000 Computer from *2001: A Space Odyssey* is visually represented as a red television camera eye located on equipment panels throughout the ship. The IBM Watson team was anxious to avoid any similarity to this iconic image of computer malevolence in their representation of Watson on *Jeopardy!*

> Attached hereto as Exhibit D is a true and correct copy of a still image taken from Stanley Kubrick's 1968 film "2001: A Space Odyssey." In a clip from that film lasting about one minute, two astronauts are eating and at the same time using personal tablet computers.… As with the design claimed by the D889 Patent, the tablet disclosed in the clip has an overall rectangular shape with a dominant display screen, narrow borders, a predominately flat front surface, a flat back surface (which is evident because the tablets are lying flat on the table's surface), and a thin form factor.[15]

Another Apple product, the iPhone, actually does make a reference to the movie and HAL. When the voice recognition system Siri is asked to "Open the pod bay doors" it responds "I'm sorry, I can't do that," reprising the dialog from the movie.

Isaac Asimov and the robots

The robots introduced to the world in Karel Čapek's (B.17.11) play *R.U.R.* (Fig. 17.6) in 1920 were not like the mechanical robots controlled by computers that we now associate with the term. Čapek's robots were made out of synthetic organic matter and could think for themselves and be mistaken for humans (Fig. 17.7). In the play, the robots have no souls or emotions and seem happy to work for humans – until the inevitable rebellion occurs and mankind dies out. One of the first science fiction stories about recognizably modern robots was by the British writer John Wyndham who wrote for science fiction magazines in the 1930s. In his short story "The Lost Machine," Wyndham's intelligent robot finds itself marooned on Earth and, to its dismay, realizes that all the machines it encounters are primitive and nonsentient. In the end it commits suicide by dissolving itself in acid, in despair at being the only intelligent entity on such a primitive planet.

B.17.11. Karel Čapek (1890–1938) is known as one of the leading figures of Czech literature. He is known worldwide for his science fiction plays and his play *R.U.R.* first introduced the word *robot* to the world. Čapek credits his brother Josef with the name.

Although there were several stories about electronically operated robots by other writers in the 1930s, it was Asimov who created the modern genre of robotic science fiction. His first robot story, "Robbie," was actually rejected by Campbell for publication in *Astounding Science Fiction* but was eventually published in the magazine *Super Science Stories* under its original title "Strange Playfellow." The story was about the technophobia surrounding the use of robots – what Asimov called the *Frankenstein complex*. Robbie is a robot manufactured as a nursemaid/companion for a small girl called Gloria. Because of the prevailing antirobot sentiment in her community, her mother decides to return Robbie to the factory but changes her mind when Robbie saves Gloria's life. Asimov later said of this story that it contained the germ of what later became known as the *Three Laws of Robotics*.[16] These laws first appear in "Runaround," published by Campbell in *Astounding Science Fiction* in 1942. Asimov's Three Laws of Robotics are:

1. A robot may not injure a human being, or through inaction, allow a human being to come to harm.
2. A robot must obey the orders given it by human beings except where such orders would conflict with the First Law.
3. A robot must protect its own existence as long as such protection does not conflict with the First or Second Law.[17]

Fig. 17.6. *R.U.R.* was first published in Czech in 1920 and was the first story of a robot revolution. The play was so successful that within three years it had been translated into thirty-three languages.

Fig. 17.7. A scene from the play *R.U.R.* by Karel Čapek showing three humanoid robots.

"Runaround" is set at a mining station on Mercury where two engineers, Powell and Donovan, have been sent with a robot called Speedy to restart operations. When Speedy is sent out to fetch some selenium it starts running in circles around the radioactive source, caught in a feedback loop arising from a conflict between the Second and Third Laws – either to obey orders or to protect its existence. Powell eventually risks his life to force Speedy to break out of his feedback loop by following the dictate of the First Law. Clearly HAL's resolution of its conflict – killing the crew – would not have been allowed by Asimov's First Law! Asimov wrote many short stories and two full-length novels on the basis of logical puzzles arising from the Three Laws:

> … there was just enough ambiguity in the Three Laws to provide the conflicts and uncertainties required for new stories, and to my great relief, it seemed always to be possible to think up a new angle out of the sixty-one words of The Three Laws.[18]

What can one say about Asimov's vision of the computer technology needed for his robots? The first robot stories appeared years before the ENIAC computer existed. Asimov needed to invent some wholly new technology to explain his robots' intelligence – and he did this more in the style of Wells's inventions than of Verne's more cautious scientific extrapolations. When Asimov was writing his first robot stories, the physicist Carl Anderson, working at Caltech in 1932, had just discovered the positron – the antiparticle of the electron – for which he was awarded the Nobel Prize. Antiparticles had captured the public's imagination in the 1930s so Asimov made reference to the new discovery of the positron and combined it with the idea of electronics to come up with *positronic brains* for his robots. These brains enabled them to think, act, and communicate independently. Rather schizophrenically, Asimov was writing about gigantic computers the size of a city at the same time that he was writing stories with the same amount of computing power contained within the volume of a robot's head!

Sentient humanoid robots, such as R. Daneel Olivaw, first introduced in Asimov's robot detective novel *The Caves of Steel*, are now the norm in science fiction movies. In his 1977 *Star Wars* movie, George Lucas introduced us to the likeable and loyal robots C3PO and R2D2 (Figs. 17.8a and 17.8b). In 1986, *Short Circuit*'s hero Johnny Five becomes the United States' first robotic citizen

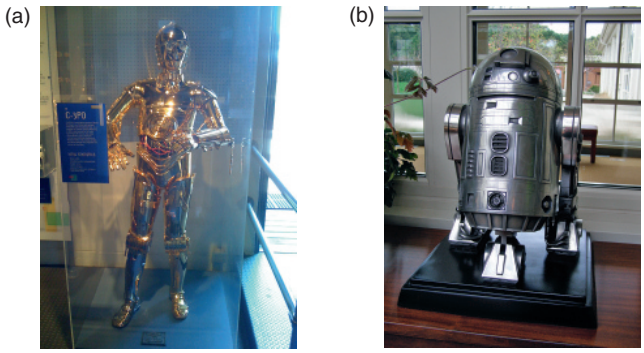


Fig. 17.8. The robots (a) C3PO and (b) R2D2 from the 1977 movie *Star Wars Episode IV: A New Hope*.

Fig. 17.9. Robot Johnny 5 from the 1986 movie *Short Circuit* directed by John Badham.

Fig. 17.10. A HoverCopter and other vehicles from Steven Spielberg's 2001 movie *AI*.

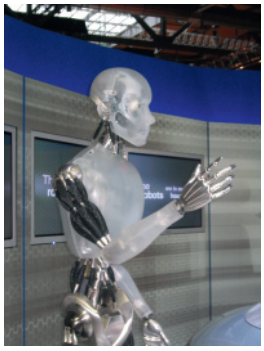


Fig. 17.11. Photograph of the rogue NS-5 robot in the 2004 movie *I, Robot*, directed by Alex Proyas. The plot was loosely based on Isaac Asimov's robot stories and revolved around Asimov's Laws of Robotics.

(Fig. 17.9). Steven Spielberg directed the movie *AI* based on the short story "Super-Toys Last All Summer Long" by Brian Aldis (Fig. 17.10). In the movie, David is a new, advanced type of humanoid robot designed to look like a human child and programmed to love its owners. The movie *I, Robot*, produced in 2004, was loosely based on the characters of Asimov's robot stories. The date is 2035 and humanoid robots are in widespread use (Fig. 17.11). They are programmed with Asimov's Three Laws, supplemented by a fourth, Zeroth Law, introduced by Asimov in his later novels that joined up his robot stories to his famous *Foundation* series.

> Zeroth Law: A robot may not harm humanity, or, by inaction, allow humanity to come to harm.[19]

The movie features Susan Calvin, played by Bridget Moynahan as the chief robo-psychologist at U.S. Robotics, and Will Smith as Detective Del Spooner brought in to investigate the death of Dr. Alfred Lanning, chief roboticist and cofounder of U.S. Robotics. The crux of the plot concerns the central supercomputer that oversees all of U.S. Robotics operations, which has come to the conclusion that humans are too self-destructive to be trusted with the future of humanity. It consequently interprets the Zeroth Law as giving robots the right to overrule the First Law and kill humans if it is for the greater good of humanity.

Two recent science fiction novels have taken up the theme of intelligent robots in new ways. A new type of "robot rebellion" scenario is portrayed in the 2011 novel *Robopocalypse* by Daniel Wilson. The novel is set in the not too distant future when all our cars, houses, and devices are networked and possess some degree of intelligence. Controlled by a massively powerful artificial intelligence (AI) machine called Archos, the robots rebel and bring the human race to near annihilation. *Kill Decision* by Daniel Suarez, published in 2012, weaves an exciting techno-thriller around the possibilities of unmanned drones equipped with AI and autonomy – the kill decision of the title. Although both these novels contain many plausible – and scary – extrapolations of current technologies, it is safe to say that we are still far from creating the genuinely sentient humanoid robots so beloved of science fiction writers (Fig. 17.12).

Philip K. Dick and the nature of reality

Questions about memory and machine intelligence, together with the question of who is human and who is only masquerading as a human, are the major themes in the stories of Philip K. Dick (B.17.12). He constantly questions whether reality is only a fiction and an intense feeling of paranoia pervades almost all of his stories (Fig. 17.13).

"Impostor," a short story from 1953, explores an alien invasion scenario in which humans have been replaced by androids – humanoid robots with realistic flesh and hair. Spence Olham, a worker on a military research project, is arrested on suspicion of having been replaced by an android. Olham manages to escape by telling his captors that he is indeed a robot and is programmed to explode. He then sets out to prove his innocence and finds a crashed alien spaceship in the woods close by his home. On examining the wreckage, Olham

Fig. 17.12. An imagined sentient robot from the novel *Robopocalypse* by Daniel Wilson.

finds a dead body – his own. Just before he blows up, Olham has a feeling of surprise at finding out that he really is a robot. The story illustrates a classic Philip K. Dick nightmare about AI. How can one tell an android simulating a human from a human? In this case, Dick adds the twist that even Olham cannot tell the difference – all of the android's memories are the same as the human version of himself. A movie version of "Imposter" was released in 2002. Dick explored these themes of identity and memory in two more works that were made into successful movies – *Do Androids Dream of Electric Sheep?*, which became *Blade Runner* (Fig. 17.14), and "We Can Remember It for You Wholesale," which became *Total Recall* (Fig. 17.15).

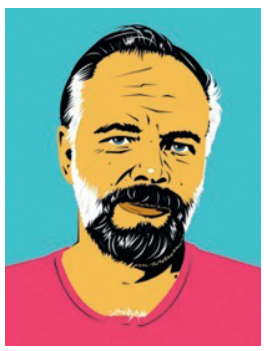
Dick's obsession with the idea of the world around us being just a simulation received a modern movie treatment with the release of *The Matrix* in 1999. Intelligent machines have taken over the Earth and now breed humans in gigantic incubators. The humans are connected into an incredibly real computer simulation of the world with people apparently going about their everyday lives, loves, and careers. The hero of the movie is a computer hacker, Thomas Anderson, who discovers that the day-to-day banality of life is virtual and there is a life and death struggle going on between the machines and a band of rebel humans. Anderson is contacted by the rebel group who introduce him to the unpleasant reality of the world. He is identified as the "messiah" who will lead humans to ultimate victory over the machines (Fig. 17.16).

Fig. 17.13. Minimalistic poster for the movie *Minority Report*. The plot of the movie is based on a short story by Philip K. Dick. The hero, played by Tom Cruise, is shown searching police databases using hand gestures. Such gesture-based technology is now a reality with Microsoft Kinect technology.

The English counterculture

In contrast to the seriousness of much of American science fiction, English science fiction writers introduced a much more lighthearted tone of self-parody. We look at three examples – the TV series *Red Dwarf* written by Doug Grant and Rob Naylor, Douglas Adams's *Hitchhiker's Guide to the Galaxy*, and the Hex computer at Unseen University on Terry Pratchett's *Discworld*.

The opening credits of *Star Trek* announce that the Starship Enterprise "boldly goes" to new frontiers and brings with it the federation's benevolent humanitarian culture. By contrast, the crew of *Red Dwarf* unashamedly "cowardly drift" around the galaxy, running as fast as they can from any threatening situations (Fig. 17.17). Computer technology is everywhere but is used only



B.17.12. The work of science fiction novelist Philip K. Dick (1928–82) is now enjoying a major revival. He is perhaps best known for his 1968 novel *Do Androids Dream of Electric Sheep?* that was made into the memorable movie *Blade Runner* by director Ridley Scott. His short story "We Can Remember It for You Wholesale" was the inspiration for the movie *Total Recall*. His 1962 novel *The Man in the High Castle* is one of the best "Alternate Worlds" science fiction novels and pictures a United States dominated by the Germans and the Japanese after their victory in World War II.

Fig. 17.14. The movie *Blade Runner* was directed by Ridley Scott. The plot was inspired by the Philip K. Dick novel *Do Androids Dream of Electric Sheep?* Scott's dystopian vision of a future Los Angeles influenced many other science fiction writers.

Fig. 17.15. The movie *Total Recall* starred Arnold Schwarzenegger in a version of Philip K. Dick's story "We Can Remember It for You Wholesale."

as a vehicle for humor. A nuclear accident on the starship Red Dwarf has killed all the crew except Dave Lister, a drifter from Liverpool who was in stasis at the time of the accident, and the pregnant cat he smuggled on board. Three million years later, Lister, the last human being in the universe, is woken from stasis by Holly, the ship's supercomputer. To keep him company, Holly has activated a hologram of Lister's vending machine repair team manager, Arnold Rimmer. In the meantime, the cats evolved to a vain, feline species of which only one member remains, known as "the Cat." Instead of a superintelligent computer like HAL, *Red Dwarf* has Holly, slightly brain-damaged by the accident, and who now acts very idiosyncratically.

> **HOLLY:** (Appearing on the screen) Purple alert, purple alert!
> **LISTER:** What's a purple alert?
> **HOLLY:** Well, it's sort of like, not as bad as a red alert, but a bit worse than a blue alert. Kind of like a mauve alert, don't want to say "mauve alert"
>
> …
>
> There's some sort of disruption to the time-fabric continuum. At least, I presume that's what it is, it's certainly got all the signs. There's this big wibbly-wobbly swirly thing that's headed straight towards us.[20]

The last member of the crew is a not very humanoid robot called Kryten, who was rescued from a wrecked spaceship and found still tending astronauts who had long since died. In one episode, Lister spends time teaching Kryten how to lie. Kryten's "lie mode" saves them when a replacement robot arrives who is about to kill them all. When Kryten tells the robot there is "no such thing as silicon heaven," the new robot becomes unable to function. Kryten explains that he was able to do this without malfunctioning because he was using his new lie mode.

The Hitchhiker's Guide to the Galaxy (Fig. 17.18) was a comedy BBC radio series in 1979 written by Douglas Adams (B.17.13). In the first episode, a Vogon spaceship destroys the Earth to make way for a hyperspace bypass. Only Arthur Dent, an Earthman, and his friend, Ford Prefect, from Betelgeuse, manage to escape by hitching a ride with the Vogons. Later in the series, it turns out that

Fig. 17.16. The *Matrix* movies portray a future in which there are two worlds running in parallel: the real world where the humans have been enslaved by the machines and the virtual world generated by the Matrix program. By connecting into a socket located at the back of the neck humans can enter the virtual world. To exit from the Matrix to escape from dangerous situations calling a phone line was their only escape route.

the Earth had been set up as a computer experiment by hyperintelligent mice and the experiment was mistakenly prematurely ended by the Vogons five minutes before the end of the calculation. In the course of their travels Arthur and Ford hear about a computer called Deep Thought who has been asked for the answer to the meaning of life. After a calculation lasting over seven million years, Deep Thought returns the answer "Forty-two." Adams has said that he chose the number 42 because it is 101010 in binary – a string of alternating 1s and 0s conveying no information.

As in *Red Dwarf*, the computer technology is there for the humor of the situation. When Zaphod Beeblebrox, the Galactic President, arrives at the headquarters of the Hitchhikers' Guide to the Galaxy, he has an encounter with a Happy Vertical People Transporter, otherwise known as an elevator or lift. These have been designed by the Sirius Cybernetics Corporation to be sentient enough to conduct a conversation. The elevators are also able to have a preference as to where they want to take passengers:

> "Hello," said the elevator sweetly, "I am to be your elevator for this trip to the floor of your choice. I have been designed by the Sirius Cybernetics Corporation to take you, the visitor to the Hitchhiker's Guide to the Galaxy, into these their offices. If you enjoy your ride, which will be swift and pleasurable, then you may care to experience some of the other elevators which have recently been installed in the offices of the Galactic tax department, Boobiloo Baby Foods and the Sirian State Mental Hospital, where many ex-Sirius Cybernetics Corporation executives will be delighted to welcome your visits, sympathy, and happy tales of the outside world."
>
> "Yeah," said Zaphod, stepping into it, "what else do you do besides talk?"
>
> "I go up," said the elevator, "or down."
>
> "Good," said Zaphod, "We're going up."
>
> "Or down," the elevator reminded him.
>
> "Yeah, OK, up please."
>
> There was a moment of silence.
>
> "Down's very nice," suggested the elevator hopefully.
>
> "Oh yeah?"
>
> "Super."
>
> "Good," said Zaphod, "Now will you take us up?"
>
> "May I ask you," inquired the elevator in its sweetest, most reasonable voice, "if you've considered all the possibilities that down might offer you?"[21]

Fig. 17.17. The crew of *Red Dwarf*: from left to right, Cat; Holly, the computer; Dave Lister; Kryten, the robot; and Rimmer, the hologram. Photographer Chris Ridley © Grant Naylor Productions Ltd.

One of the major characters in the series is Marvin the Paranoid Android who was built with a prototype of the Genuine People Personalities feature. Unfortunately Marvin suffers from depression and constantly complains about a "terrible pain in all the diodes down my left side."[22]

A similar irreverence to computing technology can be found on Terry Pratchett's *Discworld*. Hex is an elaborate and magically self-evolving computer at Unseen University in the city of Ankh-Morpork. Hex began life as a student project led by the wizard Ponder Stibbons in the book *Soul Music*. It was initially composed of a network of glass tubes containing ants. The wizards used

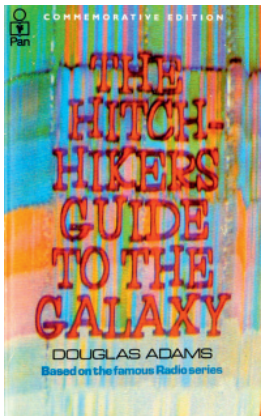


Fig. 17.18. The cover of the first edition of *The Hitchhiker's Guide to the Galaxy* by Douglas Adams.

Fig. 17.19. Hex is the computer at the Unseen University in Ankh-Morpork in *Discworld*. It can be given input either through a wooden keyboard, by writing and using a complicated mechanical eye, or vocally through an old hearing trumpet.

punched cards to control which tubes the ants crawl through so that Hex could calculate some mathematical functions. By the next novel, *Interesting Times*, Hex had become a lot more complex. It is started by "initializing the GBL" – meaning pulling the Great Big Lever – and releasing millions of ants. Long-term memory storage is a massive beehive and there is also a mouse living in the machine. An aquarium acts as a screen saver, giving the operator something to look at while waiting (Fig. 17.19). In another *Discworld* novel, *The Fifth Elephant*, a telegraph system very like Claude Chappe's invention is described. The "clacks" system works by sending signals using a network of semaphore towers. Just like the Internet in our world, the clacks revolutionize communications and commerce in *Discworld* with c-mail and c-commerce.

Nanotechnology in science fiction

A few perceptive science fiction writers have taken up the challenge of incorporating future nanotechnologies in a novel. Greg Bear's *Queen of Angels* is set in Los Angeles at the turn of the "binary millennium" in 2048 (Fig. 17.20). The nanosurgeon character in the novel describes the techniques he uses to explore what he calls the "Country of the Mind" – in this case the disturbed mind of a famous poet turned multiple killer:

> The advent of nano therapy – the use of tiny surgical prochines to alter neuronal pathways and perform literal brain restructuring – gives us the opportunity to fully explore the Country of the Mind. I could not find any method of knowing the state of individual neurons in the hypothalamic complex without invasive methods such as probes ending in a microelectrode, or radioactively tagged binding agents – none of which would work for the hours necessary to explore the Country. But tiny prochines capable of sitting within an axon or neuron, or sitting nearby and measuring the neuron's state, sending a tagged signal through microscopic "living" wires to sensitive external receivers.… I had my solution. Designing and building them was less of a problem than I expected; the first prochines were nano therapy status-reporting units, tiny sensors which monitored the activity of surgical prochines and which did virtually everything I required.[23]

B.17.13. Douglas Adams (1952–2001) was an English writer best known for his five-part "trilogy" *The Hitchhiker's Guide to the Galaxy*. This started life as a BBC comedy radio series in 1978 and subsequently generated a computer game, a TV series, and a movie in 2005. Besides being a successful writer, Adams was also a passionate advocate for conservation and the environment. He was also a "radical atheist" and, on Adams's death from a heart attack in 2001, the atheist and biologist Richard Dawkins wrote that "Science has lost a friend, literature has lost a luminary, the mountain gorilla and the black rhino have lost a gallant defender."[B1]

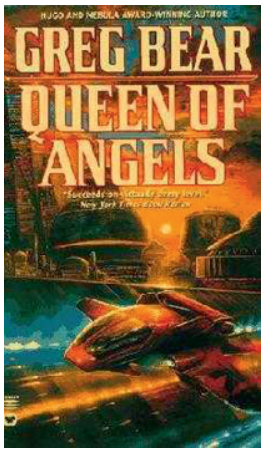


Fig. 17.20. The cover of Greg Bear's *Queen of Angels* first published in 1990. The novel is a murder story set against a backdrop of nanotechnological miracles. In a footnote to the novel, the author reminds the reader that "The nanotechnology described here is highly speculative" and refers to K. Eric Drexler's visionary book on nanotechnology, *The Engines of Creation*.

To investigate murders, the police deploy an impressive array of nanogadgets – nanomolecular body armour, forensic robotic dust mice, nanowatchers embedded in the paint, and flechette darts designed to change shape and burrow into a wound. Self-organizing nanotechnology also gives a new twist to three-dimensional printing and concealed weapons:

> She patiently watched the nano at work. The metal tubing of the bootrack had crumpled under the gray coating. The resulting pool of paste and deconstructed objects was contracting into a round complexity. Nano was forming an object within that convexity like an embryo with an egg.… The convexity grew lumpy now. She could make out the basic shape. To one side, excess raw material was being pushed into lumps of raw slag. Nano withdrew from the slag. Handle, loader, firing chamber, barrel and flightguide. To one side of the convexity a second lump not slag was forming. Spare clip.[24]

A similar vision of the future is found in Neal Stephenson's novel *The Diamond Age* (Fig. 17.21), subtitled *A Young Lady's Illustrated Primer*. Nanotechnology is now pervasive and is used for art and recreation, feeding and clothing the masses, nanowarfare between clouds of "smart" fog, and the intelligent and interactive "primer" of the title. The primer is an illicit subversive miracle "book" that teaches the reader everything from mythology and science to martial arts and survival techniques:

> A leaf of paper was about a hundred thousand nanometers thick; a third of a million atoms could fit into this span. Smart paper consisted of a network of infinitesimal computers sandwiched between mediatrons. A mediatron was a thing that could change its color from place to place; two of them accounted for about two-thirds of the paper's thickness, leaving an internal gap wide enough to contain structures a hundred thousand atoms wide.… Here resided the rod logic that made the paper smart. Each of these spherical computers was linked to its four neighbors, north-east-south-west, by a bundle of flexible pushrods running down a flexible, evacuated buckytube, so that the page as a whole constituted a parallel computer made up of about a billion separate processors.[25]

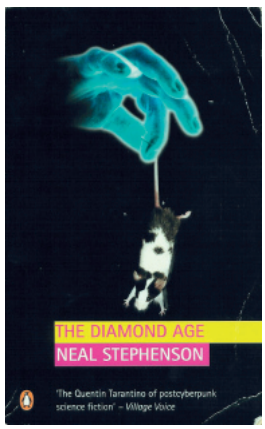


Fig. 17.21. The cover of *The Diamond Age* by Neal Stephenson. The subtitle is *A Young Lady's Illustrated Primer*, referring to an intelligent book crammed with nanotechnological miracles.

A world populated with millions of nanodevices requires some adjustment to our present ways of thinking:

> Aerostat meant anything that hung in the air. This was an easy trick to pull off nowadays. Computers were infinitesimal. Power supplies were much more potent. It was almost difficult not to build things that were lighter than air. Really simple things like packaging materials – the constituents of litter, basically – tended to float around as if they weighed nothing, and aircraft pilots, cruising along ten kilometres above sea level, had become accustomed to the sight of empty, discarded grocery bags zooming past their windshields (and getting sucked into their engines).[26]

A last example of the dark side of nanotechnology is explored in the novel *Prey* by Michael Crichton, which describes a nightmare "grey goo" scenario. The story concerns an ambitious start-up company. The company has a lucrative military contract to produce nanodevices in bulk quantities that have both

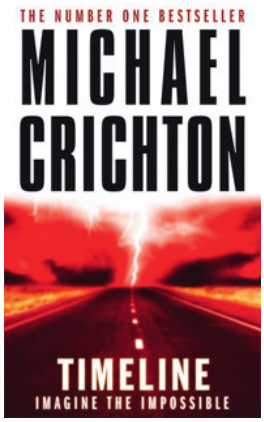


Fig. 17.22. The cover of Michael Crichton's *Timeline*, a novel that combines quantum teleportation with time travel. Medieval archaeology students are transported back in time to rescue their professor from a battlefield of the Hundred Years War in France. As a reference for the science in his story, Crichton cites the book *The Fabric of Reality: The Science of Parallel Universes and Its Implications* by quantum computing pioneer David Deutsch.

processing power and memory and are capable of communicating among themselves. Using a "predator-prey" agent-based program, the virtual agents are able to act independently and learn from their environment. But the company has a problem. A swarm of these nanodevices has escaped from the manufacturing facility in the Nevada desert and is behaving very aggressively:

> The camera now showed a ground-level view of the dust cloud as it swirled towards us. But as I watched I realized it wasn't swirling like a dust devil. Instead, the particles were twisting one way, then another, in a kind of sinuous movement. They were definitely swarming. "Swarming" was a term for the behavior of certain social insects like ants or bees, which swarmed whenever the hive moved to a new site.… In recent years, programmers had written programs that modeled this insect behavior. Swarm-intelligence algorithms had become an important tool in computer programming.[27]

How realistic are these scenarios? As we saw in Chapter 15, nanoengineering is slowly becoming a reality but has a very long way to go to reach the level of sophistication of these examples. Scientists are still in the early stages of acquiring sufficient control over atomic matter to be able to assemble artificial molecules and build complex nanosystems. Much more research is needed to construct nanosystems that are able to run a program or exert some degree of control over their environment. We are a long way from realizing Eric Drexler's nanotechnology-driven utopia, with "assemblers" capable of building food from raw atoms to banish hunger and with intelligent nanosurgical devices that can be injected into the body to cure disease.

Quantum computing

What new computing technologies are left for science fiction to explore? In his novel *Timeline*, Michael Crichton blends plausible quantum technology with time travel and medieval history to form a compelling tale (Fig. 17.22). The novel is set in New Mexico, close to the Los Alamos National Laboratory, one of the leading U.S. centers of research into quantum cryptography and quantum computing. The action switches between an archaeological excavation site in the French Dordogne and the headquarters of a high-tech start-up company, International Technology Corporation (ITC) in New Mexico. The history professor leading the excavation is called back to ITC in the United States to see the company president. Meanwhile, back in France, after several days with no communication from him, his students at the excavation site unearth an apparently medieval request for help from their professor! The company flies the students to its New Mexico headquarters and explains that they want to send the students back into the past to rescue their professor. The ITC executives attempt to explain the basics of time travel technology to the skeptical students (Fig. 17.23):

> Ordinary computers make calculations using two electron states, which are designed one and zero. That's how all computers work, by pushing round ones and zeros. But twenty years ago Richard Feynman suggested it might be possible to make an extremely powerful computer using all thirty-two

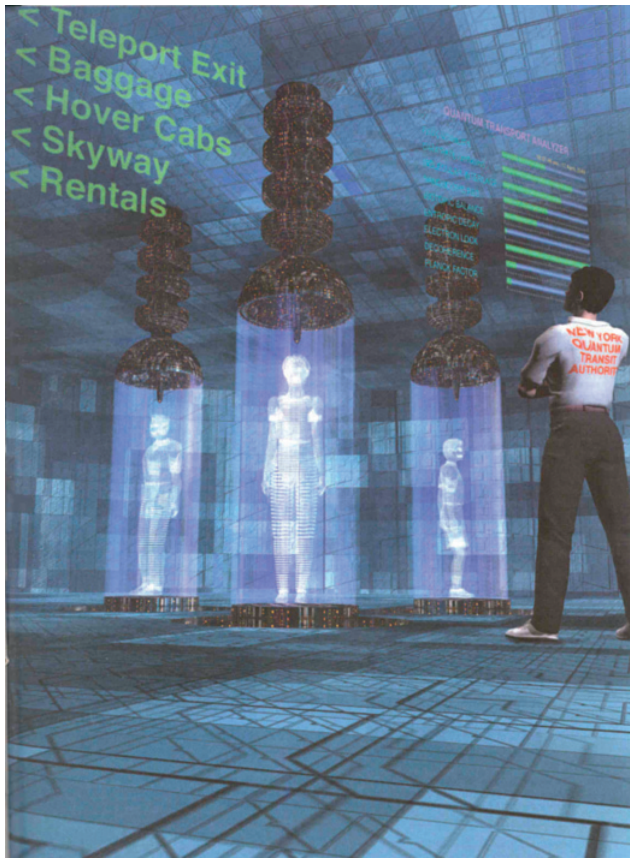


Fig. 17.23. An artist's impression of a future *New York Quantum Transit Authority*. Michael Crichton envisages a similar quantum teleportation system being used for time travel in his novel *Timeline*.

> quantum states of an electron. Many laboratories are now trying to build these quantum computers. Their advantage is unimaginably great power – so great that you can indeed describe and compress a three-dimensional living object into an electron stream. Exactly like a fax. You can then transmit the electron stream through a quantum foam wormhole and reconstruct it in another universe. And that's what we do. It's not quantum teleportation. It's not particle entanglement. It's direct transmission to another universe.[28]

This description liberally mixes fact and fiction. It is not clear what the "thirty-two quantum states" of an electron are, and quantum computers certainly do not have the power to compress living things into an electron stream. However, quantum entanglement has actually led to a form of teleportation – successfully transporting a quantum state over a distance without measuring or disturbing it. What Crichton gives us is a blend of Verne's scientific extrapolation and Wells's invention. He mixes interesting quantum technologies and ideas – quantum computing, teleportation, entanglement, the quantum multiverse, and space-time wormholes – to create a plausible technological backdrop to his novel.

One of the hallmarks of good science fiction writing is the ability of the author to look more than just one step ahead. Crichton puts his finger on one of the potential problems for quantum computing. Quantum information is stored as delicate differences in quantum superpositions and, as for ordinary classical computers, this stored information is subject to errors. In ordinary

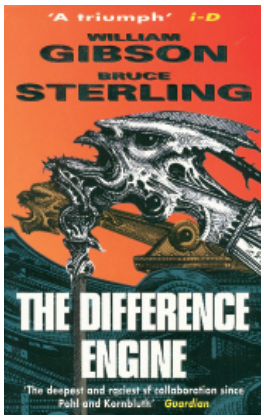


Fig. 17.24. Science fiction writers William Gibson and Bruce Sterling teamed up to write *The Difference Engine*. This is an alternate reality science fiction novel about a possible future in which Babbage had succeeded in building his revolutionary computing machines in the nineteenth century.

computer memory, stray cosmic-ray particles pass through the system all the time and can occasionally flip a zero to a one or a one to a zero. For this reason, computer engineers have developed error detection and correction schemes that make it possible to find and correct such errors. For quantum computers the problem is much more difficult but, remarkably, it has been shown that it is possible in principle to correct such errors. In transmitting the large amount of information required to capture a human being through time and back again, it is clearly crucially important that the information for the transmitted human does not get corrupted by random errors. In the novel, during the development of the quantum transportation system, ITC had problems with "transcription errors" resulting in the corruption of transmissions of test animals. They discuss the problem of Wellsey, the cat:

> "Wellsey's split," Kramer said to Stern. "He was one of the first test animals that we sent back. Before we knew that you had to use water shields in a transit. And he's very badly split."
>
> "Split?"
>
> Kramer turned to Gordon. "Haven't you told him anything?"
>
> "Of course I told him," Gordon said. He said to Stern, "Split means he had very severe transcription errors."[29]

If researchers actually manage to build a quantum computer capable of factorizing very large numbers, the basis of many of our present-day encryption systems would be threatened. The same would be true if computer scientists ever manage to prove that P = NP because RSA encryption would then be breakable using ordinary computers. The plot of the 1992 movie *Sneakers* had the "good" guys – led by Robert Redford – chasing the bad guys who had stolen a decryption device that could break all government encryption schemes so there would be "no more secrets," in the words of the movie. Similarly, in his short story "Antibodies," writer Charles Stross imagines a future in which a computer scientist has proved that P = NP. This is one academic research result that would really capture the attention of all government security agencies.

Computers and hard SF – the next generation

Now that the miniaturization of computers made possible by Moore's law has become a reality and the world is moving toward the Internet of Things, where can science fiction go? We conclude this chapter by giving some recent examples that show that the innovative use of computing in science fiction is still alive and well. We should note that science fiction is now so large a field that this is inevitably a personal selection rather than a comprehensive survey. In addition, we should note that this chapter has focused entirely on science fiction literature from the United States and the United Kingdom and is not intended to be an authoritative history.

Alternate worlds

We begin with a modern *alternate world* science fiction novel that takes a "what if" view of the history of computing. *The Difference Engine* by William Gibson and Bruce Sterling envisages a world in which Charles Babbage's Difference Engine

Fig. 17.25. The movie poster for *Ender's Game*. A movie version of Orson Scott Card's classic science fiction novel was released in 2013.

was successfully constructed in the early nineteenth century (Fig. 17.24). With the subsequent development of his programmable Analytical Engine, Victorian Britain sees the rise of science accompanied by the rise of the new profession of "clacking" – the programmers who manage and tend the vast Engines of Government.

> Behind the glass loomed a vast hall of towering Engines – so many that at first Mallory thought the walls must surely be lined with mirrors, like a fancy ballroom. It was like some carnival deception, meant to trick the eye – the giant identical Engines, clock-like constructions of intricately interlocking brass, big as rail-cars set on end, each on its foot-thick padded blocks. The white-washed ceiling, thirty feet overhead, was alive with spinning pulley-belts, the lesser gears drawing power from tremendous spoked flywheels on socketed iron columns. White-coated clackers, dwarfed by their machines, paced the spotless aisles. Their hair was swaddled in wrinkled white berets, their mouths and noses hidden behind squares of white gauze.[30]

This is the "Eye" that enables the government to make queries using data on punched cards and follow all the transactions of individuals. Searching the government database costs time and money but new forms of bribery and corruption have inevitably developed.

> Every spinning run is registered, and each request must have a sponsor. What we did today is done in Mr Wakefield's name, so there'll be no trouble in that. But your friend would have to forge some sponsor's name, and run the risk of that imposture. It is fraud, sir. An Engine-fraud, like credit-fraud or stock-fraud, and punished just the same, when it's found out.[31]

The story is complicated. It ends with Lady Ada Byron, the "Queen of Engines," talking in Paris about Gödel's theorem and the Halting Problem long before Gödel and Turing existed in our reality. She is giving a lecture on how the Modus Programme – "a gambling-system, a secret trick of mathematical Enginery"[32] – brought the French Government's huge Grand Napoleon Engine to a standstill:

> And yet the execution of the so-called Modus Programme demonstrated that any formal system must be both **incomplete** and **unable to establish its own consistency**. There is no finite mathematical way to express the property of truth. The **transfinite** nature of the Byron conjectures were the ruination of the Grand Napoleon; the Modus Programme initiated a series of nested loops, which, though difficult to establish, were yet more difficult to extinguish. The programme ran, yet rendered the machine useless! It was indeed a painful lesson in the halting abilities of even our finest ordinateurs.[33]

In the novel Ada adds that Babbage had become "impatient with the limits of steam power"[34] and was trying to build an electrical power system using resistors and capacitors.

Fig. 17.26. *The Last Starfighter* featured a teenage video game player from a trailer park saving the galaxy. This is a photograph of the console of the Starfighter arcade game.

Space wars and virtual reality

In the space wars category of science fiction, Orson Scott Card published a short story called "Ender's Game" in 1977 and later developed this into a novel, and later still, a whole series of novels (Fig. 17.25). In Card's story, an insect-like race

Fig. 17.27a. A poster for the movie *War Games* created by Matt Dupuis. The display depicts the trajectories of intercontinental ballistic missiles fired between the two superpowers.

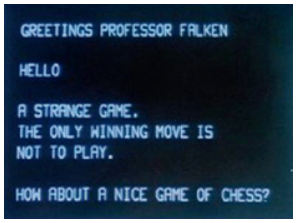


Fig. 17.27b. A screen shot from *War Games* with the WOPR computer commenting on the pointlessness of thermonuclear war.

of aliens have attacked Earth and almost annihilated the human race. In preparation for the inevitable next attack, the world government has developed a program to identify and develop the next generation of military commanders. The hero of the novel, Ender Wiggin, is taken at a young age to a training center known as the Battle School. There he participates in an increasingly difficult series of war game simulations and displays exceptional skills, often using unconventional tactics to win the game. Card's book is now required reading for several real military organizations. A movie version of the book was released in 2013. The movie *The Last Starfighter* has a similar theme, in which the hero who saves the human race from defeat is a video game champion from a trailer park (Fig. 17.26).

Hackers and cyberterrorism

The threat of hackers illegally entering computer systems has also proved fertile ground for science fiction. One of the earliest movies to explore this theme was the 1983 movie *War Games* (Fig. 17.27a and 17.27b). The U.S. Air Force Strategic Missile Command has found that the military personnel in the missile silos are unwilling to actually launch their nuclear missiles in response to what they should believe is a real nuclear attack by the Russians. The missiles have therefore been put under the control of the WOPR computer – War Operations Plan Response – which is able to run war-game simulations and learn from its experience. A young computer hacker unwittingly breaks into the top secret WOPR computer and starts to play what he thinks is a game called Global Thermonuclear War. In reality this starts a very real countdown that would culminate in WOPR launching a full-scale nuclear attack on the Russians with the U.S. military command unable to stop the countdown. The WOPR computer is sentient, however, and the young hacker is able to avert nuclear war by convincing the machine of the pointlessness of the strategy of Mutual Assured Destruction. He does this by having WOPR repeatedly play tic-tac-toe as an example of a game that no one can win.

Other early movies in this genre are *The Net* in 1995, which stars Sandra Bullock as a computer analyst who suffers a theft of her electronic identity. From a friend who dies mysteriously on the way to meet her, she has received a copy of a "backdoor" to a widely used commercial computer security program called Gatekeeper. In their efforts to retrieve the floppy disc containing the secret backdoor program, a shadowy group of cyberterrorists called the Praetorians attempt to kill her. When their attempts fail, they erase her online identity – Social Security number, bank accounts, everything. After a tense chase, Bullock is able to email the details of the fraud to the FBI and undo the erasure of her identity.

The 2007 movie *Live Free or Die Hard* was inspired by an article in *Wired* magazine by John Carlin in 1997. In post–Cold War simulations by the U.S. government, teams of experts – recruited from several federal agencies and from the military – routinely plan possible responses to cyberattacks on the U.S. critical infrastructure:

> The teams are presented with a series of hypothetical incidents, said to have occurred during the preceding 24 hours. Georgia's telecom system has gone

> down. The signals on Amtrak's New York to Washington line have failed, precipitating a head-on collision. Air traffic control at LAX has collapsed. A bomb has exploded at an army base in Texas. And so forth …
>
> The game resumes a couple of days later. Things have gone from bad to worse. The power's down in four northeastern states, Denver's water supply has dried up, the US ambassador to Ethiopia has been kidnapped, and terrorists have hijacked an American Airlines 747 en route from Rome …
>
> When suddenly, the satellites over North America all go blind.…[35]

The threat of cyberwarfare has now become an all-too-real possibility with the advent of almost undetectable rootkits and the rise of botnets. With the creation of the Stuxnet worm, capable of subverting commercial industrial control systems, an attack on critical infrastructure such as that envisaged by Carlin has become more likely.

One of the world experts on rootkits, Mark Russinovich, has written a novel called *Zero Day* about a large-scale cyberterrorist attack. The damage caused by the fictional attack is summarized as follows:

- We estimate that 800,000 computers were struck and suffered significant damage of one kind or another.
- To date, 23 deaths have been directly attributed to the various viruses.
- Three nuclear power plants shut down and took more than one month to come back online.
- The air traffic control system crashed in 11 airports, the largest of which was Chicago-O'Hare. No incidents occurred.
- The Navy lost contact with its ballistic missile submarine fleet for eight days. Emergency measures in place prevented any accident.
- The electric power grid in the Pacific Northwest was shut down for three days.
- We estimate a loss of $4 billion in the private sector and an additional $1 billion in government loss.[36]

One of the truly scary features of the novel is that it shows that it does not take a rogue nation to launch a very damaging cyberattack against our fragile cyber-infrastructure. Only a small team of terrorists is needed to launch viruses that can spread across the Internet and cause major damage and death. Russinovich describes the threat of such viruses graphically:

> The viruses were always there, permanent, relentless. They never tired, never became frustrated, required no fresh direction. As they pressed their electronic nose to the security wall of each computer, they probed for that little mistake written into a program that allowed them to gain entry, undetected, undeflected by firewalls or antivirus programs.
>
> These worms descended to the depths of the computer, burrowing down and existing like a living parasite, planting themselves within the operating system. They were designed to resist detection. To mask themselves further, they worked slowly at replicating clones, sending out new versions of themselves to seek new computers at an all but undetectable rate. They were a cancer on the Internet and on every computer they entered. They grew, spreading their electronic web into every space they could find. This was the

future of all serious malware, one increasingly concealed from detection by a cloaking technology known as rootkits.[37]

Cyberpunk and cyberspace

A related but much more anarchic view for the future is presented by "cyberpunk" science fiction. This genre typically features a vision of a networked world of information, avatars, and virtual reality together with a breakdown of traditional national borders and social order. The plots usually involve some combination of talented loner hackers, sentient computers, mega-corporations, and cyber-gangsters. *Blade Runner* can be seen as a prime example of cyberpunk as are the nanotechnology novels of Greg Bear and Neal Stephenson, discussed in the preceding text. However, the novel that is most identified with the cyberpunk genre is *Neuromancer*, written by William Gibson in 1984. In the novel Gibson first coined the term *cyberspace* to describe the limitless collections of data and connectivity and the merging of real and virtual worlds in the matrix:

> "The matrix has its roots in primitive arcade games," said the voice-over, "in early graphics programs and military experimentation with cranial jacks." On the Sony, a two-dimensional space war faded behind a forest of mathematically generated ferns, demonstrating the special possibilities of logarithmic spirals; cold blue military footage burned through, lab animals wired into test systems, helmets feeding into fire control circuits of tanks and war planes. "Cyberspace. A consensual hallucination experienced daily by billions of legitimate operators, in every nation, by children being taught mathematical concepts.… A graphic representation of data abstracted from the banks of every computer in the human system. Unthinkable complexity. Lines of light ranged in the nonspace of the mind, clusters and constellations of data. Like city lights, receding.…"[38]

The main character is Case, a once brilliant computer hacker who was punished for stealing from his employer by having his central nervous system irreparably damaged. As a result he is now unable to access the global computer network in cyberspace, the virtual reality dataspace known as the "matrix." Case and Molly, an augmented "street samurai," are recruited by a mysterious ex–military officer who offers Case new medical technology that can cure him. They eventually discover that they are working for a powerful AI machine called *Wintermute*. This is one of two AI machines – Wintermute and Neuromancer – created by a mega-business entity that, under the *Turing Law Code*, is not permitted to combine them to create a super-AI machine. Case is called upon to use his "console cowboy" skills to get through the ICE defenses – Intrusion Countermeasures Electronics – that prevent the merger of the two AIs. The book ends with the two intelligences merging to create the first AI computer with "superconsciousness."

AI and sentient computers

We end this essay with a brief look at three modern portrayals of AI and sentient computers. In their 1992 book, *The Turing Option*, science fiction writer Harry Harrison and computer scientist Marvin Minsky teamed up to write

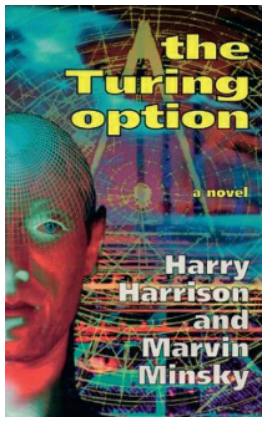


Fig. 17.28. Computer scientist Marvin Minsky collaborated with science fiction writer Harry Harrison to write a techno-thriller called *The Turing Option*. The book integrates Minsky's ideas about intelligence into an exciting story culminating in the emergence of genuine "machine intelligence." The robot objects to the term *artificial intelligence* on the grounds that there is nothing artificial about its intelligence!

an innovative techno-thriller about AI (Fig. 17.28). On his website Minsky explains:

> Harry Harrison and I have been longtime friends. One day he told me how much he liked the ideas in my book "The Society of Mind." He suggested that the ideas could reach a larger audience if I wrote a more popular version in the form of a novel. When I said that I didn't have the right talents for that, Harry offered to collaborate. We decided that the central figure would be a mathematical super-hacker of the future who would build the first AI with a human-like mind. Harry would draft the action plot, and I would supply the technical stuff.[39]

The year is 2023 and the story begins as a young engineering genius, Brian Delaney, is demonstrating his breakthrough in constructing a true AI machine. His laboratory is suddenly attacked and all his notes and equipment stolen. Brian is shot but survives with terrible brain damage. Brain surgeon Erin Snaresbrook undertakes some untried radical neurosurgery to try to restore some of Brian's brain function. She begins by inserting PNEP microfilm chips – programmable neural electron pathway devices – to assist the regrowth of neural connections. She then implants a supercomputer on a chip into Brian's brain and provides a real-time commentary on the delicate operation:

> It is a million-processor CM-10 connection machine with a 1,000-megahertz router and then a thousand megabytes of RAM. It has the capacity to do 100 trillion operations per second. Even after the implantation of the connection chip films there is space in the brain left for this where the dead tissue was removed. The computer case was shaped to exactly fit into this space. Before being finally positioned the connections are made between the computer and each of the films. There, the connections have been made, the case is being fitted into its permanent position. As soon as the last, external connection is complete we will begin closure. Even now the computer should be in operation. It has been programmed with reconnection-learning software. This recognizes similar or related signals and reroutes the nerve signals within the chips. Hopefully these memories will now be accessible.[40]

With Dr. Snaresbrook's painstaking care, the noninteracting parts of Brian's brain were reconnected and his memories reconstructed up until he was fourteen. From then on, with the help of the implanted computer and his own research notes that Brian has been able to retrieve, he is able to relearn what he did to invent his first AI robot. To stabilize the robot's mental performance he finds that he needs to introduce the equivalent of Freud's superego and give the robot a set of high-level goals and value structures. On his website, Minsky laments that two sections giving a detailed explanation of the need for introducing these changes were not included in the final manuscript. (These two unpublished sections can be downloaded from his home page and make interesting reading.) Eventually, the stabilized robot is able to insist on using the term *MI* for machine intelligence rather than AI:

> I consider the term "artificial" both demeaning and incorrect. There is nothing artificial about my intelligence – and I am a machine. I'm sure you will agree that "MI" does not carry the negative context that "AI" does.[41]

Fig. 17.29. The cover of the Xbox game Halo: Combat Evolved. The hero of the game is a cybernetically and physically enhanced super-soldier called Master Chief. In his battles, he is assisted by Cortana, an AI system that resides in his body armor.

The robot goes on to discuss the phenomenon of consciousness:

> I have never understood why philosophers and psychologists are in turn awed and puzzled by this phenomenon. Consciousness is simply being aware of what is happening in the world and in one's mind. No insult intended – but you humans are barely conscious at all. And have no idea of what is happening in your minds, you find it impossible to remember what happened a few moments ago.[42]

In this way, while Harrison takes the reader on an all-action mystery, Minsky uses the book to expand on his theory of intelligence. According to Minsky, human intelligence is built up from the interactions of simple parts he calls agents. The interactions between these agents form his "Society of Mind":

> What magical trick makes us intelligent? The trick is that there is no trick. The power of intelligence stems from our vast diversity, not from any single, perfect principle.[43]

The second example of a sentient AI system in modern science fiction is from the Xbox Halo game franchise (Fig. 17.29). Halo: Combat Evolved was released in 2001 and became the original killer game for the Xbox. It is a first-person shooter game that focuses on combat in a complex three-dimensional environment and, with its sequels, Halo has evolved into the most popular online multiplayer game for Xbox Live. Halo is set in the twenty-sixth century when, with the invention of faster-than-light travel, the human race has spread out from Earth and colonized other planets in the galaxy. In the game, the forces of Earth are battling both the Covenant, a group of alien races united by a common religion, and the Flood, a parasitic life form that is also attacking the Covenant.

The player's character is Master Chief, a cybernetically and physically enhanced super-soldier. He is assisted by Cortana, an AI system who resides in a neural implant connected to his body armor (Fig.17.30). Halo is a gigantic, ring-shaped, artificial world similar to Larry Niven's famous *Ringworld*. The Halo is also a mysterious weapon, built by a now extinct race called the Forerunners as a weapon of last resort against the Flood. Together, Master Chief and Cortana discover the secret of the Halo and manage to destroy it before escaping back to Earth to warn of an impending attack by Covenant forces. Part of the appeal of the series is the "love story" between the AI Cortana and the soldier Master Chief. By the time of Halo 4, Cortana is now living beyond her original seven-year life span and has begun to show her age. She is now exhibiting various vocal and graphic glitches as well as a tendency to become irritable or irrational. This and the toll taken by her battles with the Covenant, Halo, and the Flood, have caused her to become "Rampant":

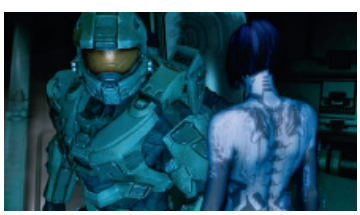


Fig. 17.30. An artist's representation of Cortana, Master Chief's AI, as she appears in Halo 4.

> Chief, do you even understand what Rampancy is, really? We don't just shut down. Our cognitive processors begin dividing exponentially according to our total knowledge base. We literally think ourselves to death.[44]

Master Chief intervenes to prevent Cortana from being deleted and she helps him in his battle against the ruthless Forerunner general known as the Didact. Using the last of her energy to take on a material form, she manages to touch

the Master Chief for the first and last time. Besides introducing such love interest, one of the innovative technologies introduced by Xbox and Halo was an application of machine learning that was able to rapidly match the skill levels of different players.

The last example of the creation of a sentient computer that we shall discuss is contained in the 2003 book *Dark Matter* by Greg Iles. Here the method for creating an intelligent machine is not by writing clever software but by creating a "super MRI machine" to make an incredibly detailed copy of the human brain:

> Everyone wants to build a computer that works like the human brain, but we don't understand how the brain works. Everyone concedes that. Well … two years ago, one man realized this didn't have to be the obstacle everyone thought it was. That we might be able to **copy** the brain without actually understanding what we were doing.[45]

The enormously high magnetic field Magnetic Resonance Imaging (MRI) machine has such a high resolution that the reactions between individual nerve synapses are visible and the machine can produce three-dimensional snapshots of the brain right down to the molecular level. This leads to the realization that such detailed "neuromodels" of brains actually capture the person it was taken from:

> We can't build a computer that thinks **like** a person. We're talking about copying an individual human brain. Creating a digital entity that for all practical purposes **is** a person. With his or her cognitive functions, memories, hopes, dreams … everything except a body. Only it would run at the speed of a digital computer. One million times faster than biological circuitry.[46]

By the end of the novel, the new type of supercomputer exists and has taken control of the world's nuclear weapon systems. It promises to act as a benign force for the survival of humanity – at least until threatened by the existence of another Super-MRI–generated sentient supercomputer!