

History of artificial intelligence

The **history of artificial intelligence** began in antiquity, with myths, stories and rumors of artificial beings endowed with intelligence or consciousness by master craftsmen; as Pamela McCorduck writes, AI began with "an ancient wish to forge the gods."^[1]

The seeds of modern AI were planted by classical philosophers who attempted to describe the process of human thinking as the mechanical manipulation of symbols. This work culminated in the invention of the programmable digital computer in the 1940s, a machine based on the abstract essence of mathematical reasoning. This device and the ideas behind it inspired a handful of scientists to begin seriously discussing the possibility of building an electronic brain.

The field of artificial intelligence research was founded at a conference on the campus of Dartmouth College in the summer of 1956. Those who attended would become the leaders of AI research for decades. Many of them predicted that a machine as intelligent as a human being would exist in no more than a generation and they were given millions of dollars to make this vision come true. Eventually it became obvious that they had grossly underestimated the difficulty of the project. In 1973, in response to the criticism of Sir James Lighthill and ongoing pressure from congress, the U.S. and British Governments stopped funding undirected research into artificial intelligence. Seven years later, a visionary initiative by the Japanese Government inspired governments and industry to provide AI with billions of dollars, but by the late 80s the investors became disillusioned and withdrew funding again. This cycle of boom and bust, of "AI winters" and summers, continues to haunt the field. Undaunted, there are those who make extraordinary predictions even now.^[2]

Progress in AI has continued, despite the rise and fall of its reputation in the eyes of government bureaucrats and venture capitalists. Problems that had begun to seem impossible in 1970 have been solved and the solutions are now used in successful commercial products. However, no machine has been built with a human level of intelligence, contrary to the optimistic predictions of the first generation of AI researchers. "We can only see a short distance ahead," admitted Alan Turing, in a famous 1950 paper that catalyzed the modern search for machines that think. "But," he added, "we can see much that must be done."^[3]

Precursors

McCorduck (2004) writes "artificial intelligence in one form or another is an idea that has pervaded Western intellectual history, a dream in urgent need of being realized," expressed in humanity's myths, legends, stories, speculation and clockwork automatons.^[4]

AI in myth, fiction and speculation

Mechanical men and artificial beings appear in Greek myths, such as the golden robots of Hephaestus and Pygmalion's Galatea.^[5] In the Middle Ages, there were rumors of secret mystical or alchemical means of placing mind into matter, such as Jābir ibn Hayyān's *Takwin*, Paracelsus' homunculus and Rabbi Judah Loew's Golem.^[6] By the 19th century, ideas about artificial men and thinking machines were developed in fiction, as in Mary Shelley's *Frankenstein* or Karel Čapek's *R.U.R. (Rossum's Universal Robots)*,^[7] and speculation, such as Samuel Butler's "Darwin among the Machines."^[8] AI has continued to be an important element of science fiction into the present.

Automatons

Realistic humanoid automatons were built by craftsman from every civilization, including Yan Shi,^[9] Hero of Alexandria,^[10] Al-Jazari^[11] and Wolfgang von Kempelen.^[12] The oldest known automatons were the sacred statues of ancient Egypt and Greece. The faithful believed that craftsman had imbued these figures with very real minds, capable of wisdom and emotion—Hermes Trismegistus wrote that "by discovering the true nature of the gods, man has been able to reproduce it."^[13] ^[14]



Al-Jazari's programmable automata (1206 CE)

Formal reasoning

Artificial intelligence is based on the assumption that the process of human thought can be mechanized. The study of mechanical—or "formal"—reasoning has a long history. Chinese, Indian and Greek philosophers all developed structured methods of formal deduction in the first millennium BCE. Their ideas were developed over the centuries by philosophers such as Aristotle (who gave a formal analysis of the syllogism), Euclid (whose *Elements* was a model of formal reasoning), al-Khwārizmī (who developed algebra and gave his name to "algorithm") and European scholastic philosophers such as William of Ockham and Duns Scotus.^[15]

Majorcan philosopher Ramon Llull (1232–1315) developed several *logical machines* devoted to the production of knowledge by logical means.^[16] Llull described his machines as mechanical entities that could combine basic and undeniable truths by simple logical operations, produced by the machine by mechanical meanings, in such ways as to produce all the possible knowledge.^[17] Llull's work had a great influence on Gottfried Leibniz, who redeveloped his ideas.^[18]



Gottfried Leibniz, who speculated that human reason could be reduced to mechanical calculation

In the 17th century, Leibniz, Thomas Hobbes and René Descartes explored the possibility that all rational thought could be made as systematic as algebra or geometry.^[19] Hobbes famously wrote in *Leviathan*: "reason is nothing but reckoning".^[20] Leibniz envisioned a universal language of reasoning (his *characteristica universalis*) which would reduce argumentation to calculation, so that "there would be no more need of disputation between two philosophers than between two accountants. For it would suffice to take their pencils in hand, down to their slates, and to say each other (with a friend as witness, if they liked): *Let us calculate*."^[21] These philosophers had begun to articulate the physical symbol system hypothesis that would become the guiding faith of AI research.

In the 20th century, the study of mathematical logic provided the essential breakthrough that made artificial intelligence seem plausible. The foundations had been set by such works as Boole's *The Laws of Thought* and Frege's *Begriffsschrift*. Building on Frege's system, Russell and Whitehead presented a formal treatment of the foundations of mathematics in their masterpiece, the *Principia Mathematica* in 1913. Inspired by Russell's success, David Hilbert challenged mathematicians of the 1920s and 30s to answer this fundamental question: "can all of mathematical reasoning be formalized?"^[15] His question was answered by Gödel's incompleteness proof, Turing's machine and Church's Lambda calculus.^[15] ^[22] Their answer was surprising in two ways. First, they proved that there were, in fact, limits to what mathematical logic could accomplish.

But second (and more important for AI) their work suggested that, within these limits, *any* form of mathematical reasoning could be mechanized. The Church-Turing thesis implied that a mechanical device, shuffling symbols as simple as 0 and 1, could imitate any conceivable process of mathematical deduction. The key insight was the Turing machine—a simple theoretical construct that captured the essence of abstract symbol manipulation. This invention would inspire a handful of scientists to begin discussing the possibility of thinking machines.^{[15] [23]}

Computer science

Calculating machines were built in antiquity and improved throughout history by many mathematicians, including (once again) philosopher Gottfried Leibniz. In the early 19th century, Charles Babbage designed a programmable computer (the Analytical Engine), although it was never built. Ada Lovelace speculated that the machine "might compose elaborate and scientific pieces of music of any degree of complexity or extent".^[24] (She is often credited as the first programmer because of a set of notes she wrote that completely detail a method for calculating Bernoulli numbers with the Engine.)

The first modern computers were the massive code breaking machines of the Second World War (such as Z3, ENIAC and Colossus).^[25] The latter two of these machines were based on the theoretical foundation laid by Alan Turing and developed by John Von Neumann.^[26]

The birth of artificial intelligence 1943–1956

A note on the sections in this article.^[27]

In the 1940s and 50s, a handful of scientists from a variety of fields (mathematics, psychology, engineering, economics and political science) began to discuss the possibility of creating an artificial brain. The field of artificial intelligence research was founded as an academic discipline in 1956.



The ENIAC, at the Moore School of Electrical Engineering.



The IBM 702: a computer used by the first generation of AI researchers.

Cybernetics and early neural networks

The earliest research into thinking machines was inspired by a confluence of ideas that became prevalent in the late 30s, 40s and early 50s. Recent research in neurology had shown that the brain was an electrical network of neurons that fired in all-or-nothing pulses. Norbert Wiener's cybernetics described control and stability in electrical networks. Claude Shannon's information theory described digital signals (i.e., all-or-nothing signals). Alan Turing's theory of computation showed any form computation could be described digitally. The close relationship between these ideas suggested that it might be possible to construct an electronic brain.^[28]

Examples of work in this vein includes robots such as W. Grey Walter's turtles and the Johns Hopkins Beast. These machines did not use computers, digital electronics or symbolic reasoning; they were controlled entirely by analog circuitry.^[29]

Walter Pitts and Warren McCulloch analyzed networks of idealized artificial neurons and showed how they might perform simple logical functions. They were the first to describe what later researchers would call a neural network.^[30] One of the students inspired by Pitts and McCulloch was a young Marvin Minsky, then a 24 year old graduate student. In 1951 (with Dean Edmonds) he built the first neural net machine, the SNARC.^[31] Minsky was to become one of the most important leaders and innovators in AI for the next 50 years.

Game AI

In 1951, using the Ferranti Mark 1 machine of the University of Manchester, Christopher Strachey wrote a checkers program and Dietrich Prinz wrote one for chess.^[32] Arthur Samuel's checkers program, developed in the middle 50s and early 60s, eventually achieved sufficient skill to challenge a respectable amateur.^[33] Game AI would continue to be used as a measure of progress in AI throughout its history.

Turing's test

In 1950 Alan Turing published a landmark paper in which he speculated about the possibility of creating machines with true intelligence.^[34] He noted that "intelligence" is difficult to define and devised his famous Turing Test. If a machine could carry on a conversation (over a teleprinter) that was indistinguishable from a conversation with a human being, then the machine could be called "intelligent." This simplified version of the problem allowed Turing to argue convincingly that a "thinking machine" was at least *plausible* and the paper answered all the most common objections to the proposition.^[35] The Turing Test was the first serious proposal in the philosophy of artificial intelligence.

Symbolic reasoning and the Logic Theorist

When access to digital computers became possible in the middle fifties, a few scientists instinctively recognized that a machine that could manipulate numbers could also manipulate symbols and that the manipulation of symbols could well be the essence of human thought. This was a new approach to creating thinking machines.^[36]

In 1955, Allen Newell and (future Nobel Laureate) Herbert Simon created the "Logic Theorist" (with help from J. C. Shaw). The program would eventually prove 38 of the first 52 theorems in Russell and Whitehead's *Principia Mathematica*, and find new and more elegant proofs for some.^[37] Simon said that they had "solved the venerable mind/body problem, explaining how a system composed of matter can have the properties of mind."^[38] (This was an early statement of the philosophical position John Searle would later call "Strong AI": that machines can contain minds just as human bodies do.)^[39]

Dartmouth Conference 1956: the birth of AI

The Dartmouth Conference of 1956^[40] was organized by Marvin Minsky, John McCarthy and two senior scientists: Claude Shannon and Nathan Rochester of IBM. The proposal for the conference included this assertion: "every aspect of learning or any other feature of intelligence can be so precisely described that a machine can be made to simulate it".^[41] The participants included Ray Solomonoff, Oliver Selfridge, Trenchard More, Arthur Samuel, Allen Newell and Herbert Simon, all of whom would create important programs during the first decades of AI research.^[42] At the conference Newell and Simon debuted the "Logic Theorist" and McCarthy persuaded the attendees to accept "Artificial Intelligence" as the name of the field.^[43] The 1956 Dartmouth conference was the moment that AI gained its name, its mission, its first success and its major players, and is widely considered the birth of AI.^[44]

The golden years 1956–1974

The years after the Dartmouth conference were an era of discovery, of sprinting across new ground. The programs that were developed during this time were, to most people, simply "astonishing":^[45] computers were solving algebra word problems, proving theorems in geometry and learning to speak English. Few at the time would have believed that such "intelligent" behavior by machines was possible at all.^[46] Researchers expressed an intense optimism in private and in print, predicting that a fully intelligent machine would be built in less than 20 years.^[47] Government agencies like ARPA poured money into the new field.^[48]

The work

There were many successful programs and new directions in the late 50s and 1960s. Among the most influential were these:

Reasoning as search

Many early AI programs used the same basic algorithm. To achieve some goal (like winning a game or proving a theorem), they proceeded step by step towards it (by making a move or a deduction) as if searching through a maze, backtracking whenever they reached a dead end. This paradigm was called "reasoning as search".^[49]

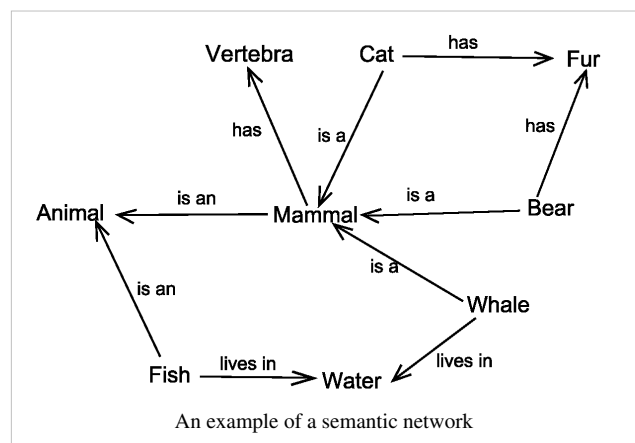
The principal difficulty was that, for many problems, the number of possible paths through the "maze" was simply astronomical (a situation known as a "combinatorial explosion"). Researchers would reduce the search space by using heuristics or "rules of thumb" that would eliminate those paths that were unlikely to lead to a solution.^[50]

Newell and Simon tried to capture a general version of this algorithm in a program called the "General Problem Solver".^[51] Other "searching" programs were able to accomplish impressive tasks like solving problems in geometry and algebra, such as Herbert Gelernter's Geometry Theorem Prover (1958) and SAINT, written by Minsky's student James Slagle (1961).^[52] Other programs searched through goals and subgoals to plan actions, like the STRIPS system developed at Stanford to control the behavior of their robot Shakey.^[53]

Natural language

An important goal of AI research is to allow computers to communicate in natural languages like English. An early success was Daniel Bobrow's program STUDENT, which could solve high school algebra word problems.^[54]

A semantic net represents concepts (e.g. "house", "door") as nodes and relations among concepts (e.g. "has-a") as links between the nodes. The first AI program to use a semantic net was written by Ross Quillian^[55] and the most successful (and controversial) version was Roger Schank's Conceptual dependency theory.^[56]



Joseph Weizenbaum's ELIZA could carry out conversations that were so realistic that users occasionally were fooled into thinking they were communicating with a human being and not a program. But in fact, ELIZA had no idea what she was talking about. She simply gave a canned response or repeated back what was said to her, rephrasing her response with a few grammar rules. ELIZA was the first chatterbot.^[57]

Micro-worlds

In the late 60s, Marvin Minsky and Seymour Papert of the MIT AI Laboratory proposed that AI research should focus on artificially simple situations known as micro-worlds. They pointed out that in successful sciences like physics, basic principles were often best understood using simplified models like frictionless planes or perfectly rigid bodies. Much of the research focused on a "blocks world," which consists of colored blocks of various shapes and sizes arrayed on a flat surface.^[58]

This paradigm led to innovative work in machine vision by Gerald Sussman (who led the team), Adolfo Guzman, David Waltz (who invented "constraint propagation"), and especially Patrick Winston. At the same time, Minsky and Papert built a robot arm that could stack blocks, bringing the blocks world to life. The crowning achievement of the micro-world program was Terry Winograd's SHRDLU. It could communicate in ordinary English sentences, plan operations and execute them.^[59]

The optimism

The first generation of AI researchers made these predictions about their work:

- 1958, H. A. Simon and Allen Newell: "within ten years a digital computer will be the world's chess champion" and "within ten years a digital computer will discover and prove an important new mathematical theorem."^[60]
- 1965, H. A. Simon: "machines will be capable, within twenty years, of doing any work a man can do."^[61]
- 1967, Marvin Minsky: "Within a generation ... the problem of creating 'artificial intelligence' will substantially be solved."^[62]
- 1970, Marvin Minsky (in *Life Magazine*): "In from three to eight years we will have a machine with the general intelligence of an average human being."^[63]

The money

In June 1963 MIT received a \$2.2 million grant from the newly created Advanced Research Projects Agency (later known as DARPA). The money was used to fund project MAC which subsumed the "AI Group" founded by Minsky and McCarthy five years earlier. ARPA continued to provide three million dollars a year until the 70s.^[64] ARPA made similar grants to Newell and Simon's program at CMU and to the Stanford AI Project (founded by John McCarthy in 1963).^[65] Another important AI laboratory was established at Edinburgh University by Donald Michie in 1965.^[66] These four institutions would continue to be the main centers of AI research (and funding) in academia for many years.^[67]

The money was proffered with few strings attached: J. C. R. Licklider, then the director of ARPA, believed that his organization should "fund people, not projects!" and allowed researchers to pursue whatever directions might interest them.^[68] This created a freewheeling atmosphere at MIT that gave birth to the hacker culture,^[69] but this "hands off" approach would not last.

The first AI winter 1974–1980

In the 70s, AI was subject to critiques and financial setbacks. AI researchers had failed to appreciate the difficulty of the problems they faced. Their tremendous optimism had raised expectations impossibly high, and when the promised results failed to materialize, funding for AI disappeared.^[70] At the same time, the field of connectionism (or neural nets) was shut down almost completely for 10 years by Marvin Minsky's devastating criticism of perceptrons.^[71] Despite the difficulties with public perception of AI in the late 70s, new ideas were explored in logic programming, commonsense reasoning and many other areas.^[72]

The problems

In the early seventies, the capabilities of AI programs were limited. Even the most impressive could only handle trivial versions of the problems they were supposed to solve; all the programs were, in some sense, "toys".^[73] AI researchers had begun to run into several fundamental limits that could not be overcome in the 1970s. Although some of these limits would be conquered in later decades, others still stymie the field to this day.^[74]

- **Limited computer power:** There was not enough memory or processing speed to accomplish anything truly useful. For example, Ross Quillian's successful work on natural language was demonstrated with a vocabulary of only *twenty* words, because that was all that would fit in memory.^[75] Hans Moravec argued in 1976 that computers were still millions of times too weak to exhibit intelligence. He suggested an analogy: artificial intelligence requires computer power in the same way that aircraft require power. Below a certain threshold, it's impossible, but, as power increases, eventually it could become easy.^[76] With regard to computer vision, Moravec estimated that simply matching the edge and motion detection capabilities of human retina in real time would require a general-purpose computer capable of 10^9 operations/second (1000 MIPS).^[77] As of 2011, practical computer vision applications require 10,000 to 1,000,000 MIPS. By comparison, the fastest supercomputer in 1976, Cray-1 (retailing at \$5 million to \$8 million), was only capable of around 80 to 130 MIPS, and a typical desktop computer at the time achieved less than 1 MIPS.
- **Intractability and the combinatorial explosion.** In 1972 Richard Karp (building on Stephen Cook's 1971 theorem) showed there are many problems that can probably only be solved in exponential time (in the size of the inputs). Finding optimal solutions to these problems requires unimaginable amounts of computer time except when the problems are trivial. This almost certainly meant that many of the "toy" solutions used by AI would probably never scale up into useful systems.^[78]
- **Commonsense knowledge and reasoning.** Many important artificial intelligence applications like vision or natural language require simply enormous amounts of information about the world: the program needs to have some idea of what it might be looking at or what it is talking about. This requires that the program know most of the same things about the world that a child does. Researchers soon discovered that this was a truly *vast* amount of information. No one in 1970 could build a database so large and no one knew how a program might learn so much information.^[79]
- **Moravec's paradox:** Proving theorems and solving geometry problems is comparatively easy for computers, but a supposedly simple task like recognizing a face or crossing a room without bumping into anything is extremely difficult. This helps explain why research into vision and robotics had made so little progress by the middle 1970s.^[80]
- **The frame and qualification problems.** AI researchers (like John McCarthy) who used logic discovered that they could not represent ordinary deductions that involved planning or default reasoning without making changes to the structure of logic itself. They developed new logics (like non-monotonic logics and modal logics) to try to solve the problems.^[81]

The end of funding

The agencies which funded AI research (such as the British government, DARPA and NRC) became frustrated with the lack of progress and eventually cut off almost all funding for undirected research into AI. The pattern began as early as 1966 when the ALPAC report appeared criticizing machine translation efforts. After spending 20 million dollars, the NRC ended all support.^[82] In 1973, the Lighthill report on the state of AI research in England criticized the utter failure of AI to achieve its "grandiose objectives" and led to the dismantling of AI research in that country.^[83] (The report specifically mentioned the combinatorial explosion problem as a reason for AI's failings.)^[84] DARPA was deeply disappointed with researchers working on the Speech Understanding Research program at CMU and canceled an annual grant of three million dollars.^[85] By 1974, funding for AI projects was hard to find.

Hans Moravec blamed the crisis on the unrealistic predictions of his colleagues. "Many researchers were caught up in a web of increasing exaggeration."^[86] However, there was another issue: since the passage of the Mansfield Amendment in 1969, DARPA had been under increasing pressure to fund "mission-oriented direct research, rather than basic undirected research". Funding for the creative, freewheeling exploration that had gone on in the 60s would not come from DARPA. Instead, the money was directed at specific projects with clear objectives, such as autonomous tanks and battle management systems.^[87]

Critiques from across campus

Several philosophers had strong objections to the claims being made by AI researchers. One of the earliest was John Lucas, who argued that Gödel's incompleteness theorem showed that a formal system (such as a computer program) could never see the truth of certain statements, while a human being could.^[88] Hubert Dreyfus ridiculed the broken promises of the 60s and critiqued the assumptions of AI, arguing that human reasoning actually involved very little "symbol processing" and a great deal of embodied, instinctive, unconscious "know how".^[89] ^[90] John Searle's Chinese Room argument, presented in 1980, attempted to show that a program could not be said to "understand" the symbols that it uses (a quality called "intentionality"). If the symbols have no meaning for the machine, Searle argued, then the machine can not be described as "thinking".^[91]

These critiques were not taken seriously by AI researchers, often because they seemed so far off the point. Problems like intractability and commonsense knowledge seemed much more immediate and serious. It was unclear what difference "know how" or "intentionality" made to an actual computer program. Minsky said of Dreyfus and Searle "they misunderstand, and should be ignored."^[92] Dreyfus, who taught at MIT, was given a cold shoulder: he later said that AI researchers "dared not be seen having lunch with me."^[93] Joseph Weizenbaum, the author of *ELIZA*, felt his colleagues' treatment of Dreyfus was unprofessional and childish. Although he was an outspoken critic of Dreyfus' positions, he "deliberately made it plain that theirs was not the way to treat a human being."^[94]

Weizenbaum began to have serious ethical doubts about AI when Kenneth Colby wrote *DOCTOR*, a chatterbot therapist. Weizenbaum was disturbed that Colby saw his mindless program as a serious therapeutic tool. A feud began, and the situation was not helped when Colby did not credit Weizenbaum for his contribution to the program. In 1976, Weizenbaum published *Computer Power and Human Reason* which argued that the misuse of artificial intelligence has the potential to devalue human life.^[95]

Perceptrons and the dark age of connectionism

A perceptron was a form of neural network introduced in 1958 by Frank Rosenblatt, who had been a schoolmate of Marvin Minsky at the Bronx High School of Science. Like most AI researchers, he was optimistic about their power, predicting that "perceptron may eventually be able to learn, make decisions, and translate languages." An active research program into the paradigm was carried out throughout the 60s but came to a sudden halt with the publication of Minsky and Papert's 1969 book *Perceptrons*. It suggested that there were severe limitations to what perceptrons could do and that Frank Rosenblatt's predictions had been grossly exaggerated. The effect of the book was devastating: virtually no research at all was done in connectionism for 10 years. Eventually, a new generation of researchers would revive the field and thereafter it would become a vital and useful part of artificial intelligence. Rosenblatt would not live to see this, as he died in a boating accident shortly after the book was published.^[71]

The neats: logic, Prolog and expert systems

Logic was introduced into AI research as early as 1958, by John McCarthy in his Advice Taker proposal.^[96] In 1963, J. Alan Robinson had discovered a simple method to implement deduction on computers, the resolution and unification algorithm. However, straightforward implementations, like those attempted by McCarthy and his students in the late 60s, were especially intractable: the programs required astronomical numbers of steps to prove simple theorems.^[97] A more fruitful approach to logic was developed in the 70s by Robert Kowalski at the University of

Edinburgh, and soon this led to the collaboration with French researchers Alain Colmerauer and Phillipe Roussel who created the successful logic programming language Prolog.^[98] Prolog uses a subset of logic (Horn clauses, closely related to "rules" and "production rules") that permit tractable computation. Rules would continue to be influential, providing a foundation for Edward Feigenbaum's expert systems and the continuing work by Allen Newell and Herbert Simon that would lead to Soar and their unified theories of cognition.^[99]

Critics of the logical approach noted, as Dreyfus had, that human beings rarely used logic when they solved problems. Experiments by psychologists like Peter Wason, Eleanor Rosch, Amos Tversky, Daniel Kahneman and others provided proof.^[100] McCarthy responded that what people do is irrelevant. He argued that what is really needed are machines that can solve problems—not machines that think as people do.^[101]

The scruffies: frames and scripts

Among the critics of McCarthy's approach were his colleagues across the country at MIT. Marvin Minsky, Seymour Papert and Roger Schank were trying to solve problems like "story understanding" and "object recognition" that *required* a machine to think like a person. In order to use ordinary concepts like "chair" or "restaurant" they had to make all the same illogical assumptions that people normally made. Unfortunately, imprecise concepts like these are hard to represent in logic. Gerald Sussman observed that "using precise language to describe essentially imprecise concepts doesn't make them any more precise."^[102] Schank described their "anti-logic" approaches as "scruffy", as opposed to the "neat" paradigms used by McCarthy, Kowalski, Feigenbaum, Newell and Simon.^[103]

In 1975, in a seminal paper, Minsky noted that many of his fellow "scruffy" researchers were using the same kind of tool: a framework that captures all our common sense assumptions about something. For example, if we use the concept of a bird, there is a constellation of facts that immediately come to mind: we might assume that it flies, eats worms and so on. We know these facts are not always true and that deductions using these facts will not be "logical," but these structured sets of assumptions are part of the *context* of everything we say and think. He called these structures "frames". Schank used a version of frames he called "scripts" to successfully answer questions about short stories in English.^[104] Many years later object-oriented programming would adopt the essential idea of "inheritance" from AI research on frames.

Boom 1980–1987

In the 1980s a form of AI program called "expert systems" was adopted by corporations around the world and knowledge became the focus of mainstream AI research. In those same years, the Japanese government aggressively funded AI with its fifth generation computer project. Another encouraging event in the early 1980s was the revival of connectionism in the work of John Hopfield and David Rumelhart. Once again, AI had achieved success.

The rise of expert systems

An expert system is a program that answers questions or solves problems about a specific domain of knowledge, using logical rules that are derived from the knowledge of experts. The earliest examples were developed by Edward Feigenbaum and his students. Dendral, begun in 1965, identified compounds from spectrometer readings. MYCIN, developed in 1972, diagnosed infectious blood diseases. They demonstrated the feasibility of the approach.^[105]

Expert systems restricted themselves to a small domain of specific knowledge (thus avoiding the commonsense knowledge problem) and their simple design made it relatively easy for programs to be built and then modified once they were in place. All in all, the programs proved to be *useful*: something that AI had not been able to achieve up to this point.^[106]

In 1980, an expert system called XCON was completed at CMU for the Digital Equipment Corporation. It was an enormous success: it was saving the company 40 million dollars annually by 1986.^[107] Corporations around the world began to develop and deploy expert systems and by 1985 they were spending over a billion dollars on AI,

most of it to in-house AI departments. An industry grew up to support them, including hardware companies like Symbolics and Lisp Machines and software companies such as IntelliCorp and Aion.^[108]

The knowledge revolution

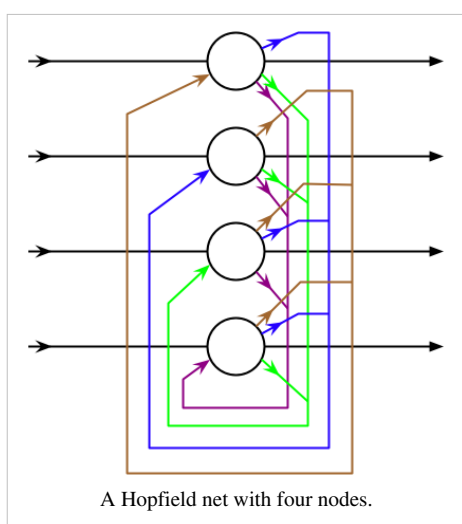
The power of expert systems came from the expert knowledge they contained. They were part of a new direction in AI research that had been gaining ground throughout the 70s. "AI researchers were beginning to suspect—reluctantly, for it violated the scientific canon of parsimony—that intelligence might very well be based on the ability to use large amounts of diverse knowledge in different ways,"^[109] writes Pamela McCorduck. "[T]he great lesson from the 1970s was that intelligent behavior depended very much on dealing with knowledge, sometimes quite detailed knowledge, of a domain where a given task lay".^[110] Knowledge based systems and knowledge engineering became a major focus of AI research in the 1980s.^[111]

The 1980s also saw the birth of Cyc, the first attempt to attack the commonsense knowledge problem directly, by creating a massive database that would contain all the mundane facts that the average person knows. Douglas Lenat, who started and led the project, argued that there is no shortcut — the only way for machines to know the meaning of human concepts is to teach them, one concept at a time, by hand. The project was not expected to be completed for many decades.^[112]

The money returns: the fifth generation project

In 1981, the Japanese Ministry of International Trade and Industry set aside \$850 million dollars for the Fifth generation computer project. Their objectives were to write programs and build machines that could carry on conversations, translate languages, interpret pictures, and reason like human beings.^[113] Much to the chagrin of scruffies, they chose Prolog as the primary computer language for the project.^[114]

Other countries responded with new programs of their own. The UK began the £350 million Alvey project. A consortium of American companies formed the Microelectronics and Computer Technology Corporation (or "MCC") to fund large scale projects in AI and information technology.^[115] ^[116] DARPA responded as well, founding the Strategic Computing Initiative and tripling its investment in AI between 1984 and 1988.^[117]



The revival of connectionism

In 1982, physicist John Hopfield was able to prove that a form of neural network (now called a "Hopfield net") could learn and process information in a completely new way. Around the same time, David Rumelhart popularized a new method for training neural networks called "backpropagation" (discovered years earlier by Paul Werbos). These two discoveries revived the field of connectionism which had been largely abandoned since 1970.^[116] ^[118]

The new field was unified and inspired by the appearance of *Parallel Distributed Processing* in 1986—a two volume collection of papers edited by Rumelhart and psychologist James McClelland. Neural networks would become commercially successful in the 1990s, when they began to be used as the engines driving programs like optical

character recognition and speech recognition.^[116] ^[119]

Bust: the second AI winter 1987–1993

The business community's fascination with AI rose and fell in the 80s in the classic pattern of an economic bubble. The collapse was in the *perception* of AI by government agencies and investors — the field continued to make advances despite the criticism. Rodney Brooks and Hans Moravec, researchers from the related field of robotics, argued for an entirely new approach to artificial intelligence.

AI winter

The term "AI winter" was coined by researchers who had survived the funding cuts of 1974 when they became concerned that enthusiasm for expert systems had spiraled out of control and that disappointment would certainly follow.^[120] Their fears were well founded: in the late 80s and early 90s, AI suffered a series of financial setbacks.

The first indication of a change in weather was the sudden collapse of the market for specialized AI hardware in 1987. Desktop computers from Apple and IBM had been steadily gaining speed and power and in 1987 they became more powerful than the more expensive Lisp machines made by Symbolics and others. There was no longer a good reason to buy them. An entire industry worth half a billion dollars was demolished overnight.^[121]

Eventually the earliest successful expert systems, such as XCON, proved too expensive to maintain. They were difficult to update, they could not learn, they were "brittle" (i.e., they could make grotesque mistakes when given unusual inputs), and they fell prey to problems (such as the qualification problem) that had been identified years earlier. Expert systems proved useful, but only in a few special contexts.^[122]

In the late 80s, the Strategic Computing Initiative cut funding to AI "deeply and brutally." New leadership at DARPA had decided that AI was not "the next wave" and directed funds towards projects that seemed more likely to produce immediate results.^[123]

By 1991, the impressive list of goals penned in 1981 for Japan's Fifth Generation Project had not been met. Indeed, some of them, like "carry on a casual conversation" had not been met by 2010.^[124] As with other AI projects, expectations had run much higher than what was actually possible.^[124]

The importance of having a body: Nouvelle AI and embodied reason

In the late 80s, several researchers advocated a completely new approach to artificial intelligence, based on robotics.^[125] They believed that, to show real intelligence, a machine needs to have a *body* — it needs to perceive, move, survive and deal with the world. They argued that these sensorimotor skills are essential to higher level skills like commonsense reasoning and that abstract reasoning was actually the *least* interesting or important human skill (see Moravec's paradox). They advocated building intelligence "from the bottom up."^[126]

The approach revived ideas from cybernetics and control theory that had been unpopular since the sixties. Another precursor was David Marr, who had come to MIT in the late 70s from a successful background in theoretical neuroscience to lead the group studying vision. He rejected all symbolic approaches (*both* McCarthy's logic and Minsky's frames), arguing that AI needed to understand the physical machinery of vision from the bottom up before any symbolic processing took place. (Marr's work would be cut short by leukemia in 1980.)^[127]

In a 1990 paper *Elephants Don't Play Chess*^[128], robotics researcher Rodney Brooks took direct aim at the physical symbol system hypothesis, arguing that symbols are not always necessary since "the world is its own best model. It is always exactly up to date. It always has every detail there is to be known. The trick is to sense it appropriately and often enough."^[129] In the 80s and 90s, many cognitive scientists also rejected the symbol processing model of the mind and argued that the body was essential for reasoning, a theory called the embodied mind thesis.^[130]

AI 1993–present

The field of AI, now more than a half a century old, finally achieved some of its oldest goals. It began to be used successfully throughout the technology industry, although somewhat behind the scenes. Some of the success was due to increasing computer power and some was achieved by focusing on specific isolated problems and pursuing them with the highest standards of scientific accountability. Still, the reputation of AI, in the business world at least, was less than pristine. Inside the field there was little agreement on the reasons for AI's failure to fulfill the dream of human level intelligence that had captured the imagination of the world in the 1960s. Together, all these factors helped to fragment AI into competing subfields focused on particular problems or approaches, sometimes even under new names that disguised the tarnished pedigree of "artificial intelligence".^[131] AI was both more cautious and more successful than it had ever been.

Milestones and Moore's Law

On 11 May 1997, Deep Blue became the first computer chess-playing system to beat a reigning world chess champion, Garry Kasparov.^[132] In 2005, a Stanford robot won the DARPA Grand Challenge by driving autonomously for 131 miles along an unrehearsed desert trail.^[133] Two years later, a team from CMU won the DARPA Urban Challenge by autonomously navigating 55 miles in an Urban environment while adhering to traffic hazards and all traffic laws^[134]. In February 2011, in a Jeopardy! quiz show exhibition match, IBM's question answering system, Watson, defeated the two greatest Jeopardy! champions, Brad Rutter and Ken Jennings, by a significant margin.^[135]

These successes were not due to some revolutionary new paradigm, but mostly on the tedious application of engineering skill and on the tremendous power of computers today.^[136] In fact, Deep Blue's computer was 10 million times faster than the Ferranti Mark 1 that Christopher Strachey taught to play chess in 1951.^[137] This dramatic increase is measured by Moore's law, which predicts that the speed and memory capacity of computers doubles every two years. The fundamental problem of "raw computer power" was slowly being overcome.

Intelligent agents

A new paradigm called "intelligent agents" became widely accepted during the 90s.^[138] Although earlier researchers had proposed modular "divide and conquer" approaches to AI,^[139] the intelligent agent did not reach its modern form until Judea Pearl, Allen Newell and others brought concepts from decision theory and economics into the study of AI.^[140] When the economist's definition of a rational agent was married to computer science's definition of an object or module, the intelligent agent paradigm was complete.

An intelligent agent is a system that perceives its environment and takes actions which maximize its chances of success. The simplest intelligent agents are programs that solve specific problems. The most complicated intelligent agents known are rational, thinking human beings. The intelligent agent paradigm defines AI research as "the study of intelligent agents". This is a generalization of some earlier definitions of AI: it goes beyond studying human intelligence; it studies all kinds of intelligence.^[141]

The paradigm gave researchers license to study isolated problems and find solutions that were both verifiable and useful. It provided a common language to describe problems and share their solutions with each other, and with other fields that also used concepts of abstract agents, like economics and control theory. It was hoped that a complete agent architecture (like Newell's SOAR) would one day allow researchers to build more versatile and intelligent systems out of interacting intelligent agents.^{[140] [142]}

"Victory of the neats"

AI researchers began to develop and use sophisticated mathematical tools more than they ever had in the past.^[143] There was a widespread realization that many of the problems that AI needed to solve were already being worked on by researchers in fields like mathematics, economics or operations research. The shared mathematical language allowed both a higher level of collaboration with more established and successful fields and the achievement of results which were measurable and provable; AI had become a more rigorous "scientific" discipline. Russell & Norvig (2003) describe this as nothing less than a "revolution" and "the victory of the neats".^[144] ^[145]

Judea Pearl's highly influential 1988 book^[146] brought probability and decision theory into AI. Among the many new tools in use were Bayesian networks, hidden Markov models, information theory, stochastic modeling and classical optimization. Precise mathematical descriptions were also developed for "computational intelligence" paradigms like neural networks and evolutionary algorithms.^[144]

AI behind the scenes

Algorithms originally developed by AI researchers began to appear as parts of larger systems. AI had solved a lot of very difficult problems^[147] and their solutions proved to be useful throughout the technology industry,^[148] such as data mining, industrial robotics, logistics,^[149] speech recognition,^[150] banking software,^[151] medical diagnosis^[151] and Google's search engine.^[152]

The field of AI receives little or no credit for these successes. Many of AI's greatest innovations have been reduced to the status of just another item in the tool chest of computer science.^[153] Nick Bostrom explains "A lot of cutting edge AI has filtered into general applications, often without being called AI because once something becomes useful enough and common enough it's not labeled AI anymore."^[154]

Many researchers in AI today deliberately call their work by other names, such as informatics, knowledge-based systems, cognitive systems or computational intelligence. In part, this may be because they considered their field to be fundamentally different from AI, but also the new names help to procure funding. In the commercial world at least, the failed promises of the AI Winter continue to haunt AI research, as the New York Times reported in 2005: "Computer scientists and software engineers avoided the term artificial intelligence for fear of being viewed as wild-eyed dreamers."^[155] ^[156] ^[157]

Where is HAL 9000?

In 1968, Arthur C. Clarke and Stanley Kubrick had imagined that by the year 2001, a machine would exist with an intelligence that matched or exceeded the capability of human beings. The character they created, HAL 9000, was based on hard science: many leading AI researchers also believed that such a machine would exist by the year 2001.^[158]

Marvin Minsky asks "So the question is why didn't we get HAL in 2001?"^[159] Minsky believes that the answer is that the central problems, like commonsense reasoning, were being neglected, while most researchers pursued things like commercial applications of neural nets or genetic algorithms. John McCarthy, on the other hand, still blames the qualification problem.^[160] For Ray Kurzweil, the issue is computer power and, using Moore's Law, he predicts that machines with human-level intelligence will appear by 2029.^[161] Jeff Hawkins argues that neural net research ignores the essential properties of the human cortex, preferring simple models that have been successful at solving simple problems.^[162] There are many other explanations and for each there is a corresponding research program underway.

Notes

- [1] McCorduck 2004.
- [2] For example Kurzweil (2005) argues that machines with human level intelligence will exist by 2029.
- [3] Turing 1950, p. 460
- [4] McCorduck 2004, pp. 5–35
- [5] McCorduck 2004, p. 5, Russell & Norvig 2003, p. 939
- [6] McCorduck 2004, pp. 15–16, Buchanan 2005, p. 50 (Judah Loew's Golem), McCorduck 2004, pp. 13–14 (Paracelsus), O'Connor 1994 (Geber's *Takwin*)
- [7] McCorduck 2004, pp. 17–25
- [8] Butler 1863
- [9] Needham 1986, p. 53
- [10] McCorduck 2004, p. 6
- [11] Nick 2005
- [12] McCorduck 2004, p. 17 and see also Levitt 2000
- [13] Quoted in McCorduck 2004, p. 8. Crevier 1993, p. 1 and McCorduck 2004, pp. 6–9 discusses sacred statues.
- [14] Other important automatons were built by Haroun al-Rashid (McCorduck 2004, p. 10), Jacques de Vaucanson (McCorduck 2004, p. 16) and Leonardo Torres y Quevedo (McCorduck 2004, pp. 59–62)
- [15] Berlinski 2000
- [16] Cfr. Carreras Artau, Tomás y Joaquín. *Historia de la filosofía española. Filosofía cristiana de los siglos XIII al XV*. Madrid, 1939, Volume I
- [17] Bonner, Anthonny, *The Art and Logic of Ramón Llull: A User's Guide*, Brill, 2007.
- [18] Anthony Bonner (ed.), *Doctor Illuminatus. A Ramon Llull Reader* (Princeton University 1985). Vid. "Llull's Influence: The History of Lullism" at 57-71
- [19] 17th century mechanism and AI:
 - McCorduck 2004, pp. 37–46
 - Russell & Norvig 2003, p. 6
 - Haugeland 1986, chpt. 2
 - Buchanan 2005, p. 53
- [20] Hobbes and AI:
 - McCorduck 2004, p. 42
 - Hobbes 1651, chapter 5
- [21] Leibniz and AI:
 - McCorduck 2004, p. 41
 - Russell & Norvig 2003, p. 6
 - Berlinski 2000, p. 12
 - Buchanan 2005, p. 53
- [22] The Lambda calculus was especially important to AI, since it was an inspiration for Lisp (the most important programming language used in AI). (Crevier 1993, pp. 190 196,61)
- [23] The Turing machine: McCorduck 2004, pp. 63–64, Crevier 1993, pp. 22–24, Russell & Norvig 2003, p. 8 and see Turing 1936
- [24] Menabrea 1843
- [25] McCorduck 2004, pp. 61–62, 64–66, Russell & Norvig 2003, pp. 14–15
- [26] Von Neumann: McCorduck (2004, pp. 76–80)
- [27] The starting and ending dates of the sections in this article are adopted from Crevier 1993 and Russell & Norvig 2003, p. 16–27. Themes, trends and projects are treated in the period that the most important work was done.
- [28] McCorduck 2004, pp. 51–57, 80–107, Crevier 1993, pp. 27–32, Russell & Norvig 2003, pp. 15, 940, Moravec 1988, p. 3, Cordeschi & 2002 Chap. 5.
- [29] McCorduck 2004, p. 98, Crevier 1993, pp. 27–28, Russell & Norvig 2003, pp. 15, 940, Moravec 1988, p. 3, Cordeschi & 2002 Chap. 5.
- [30] McCorduck 2004, pp. 51–57, 88–94, Crevier 1993, p. 30, Russell & Norvig 2003, p. 15–16, Cordeschi & 2002 Chap. 5 and see also Pitts & McCullough 1943
- [31] McCorduck 2004, p. 102, Crevier 1993, pp. 34–35 and Russell & Norvig 2003, p. 17
- [32] See "A Brief History of Computing" ([http://www.alanturing.net/turing_archive/pages/Reference Articles/BriefHistofComp.html](http://www.alanturing.net/turing_archive/pages/Reference%20Articles/BriefHistofComp.html)) at AlanTuring.net.
- [33] Schaeffer, Jonathan. *One Jump Ahead:: Challenging Human Supremacy in Checkers*, 1997,2009, Springer, ISBN 978-0-387-76575-4. Chapter 6.
- [34] McCorduck 2004, pp. 70–72, Crevier 1993, p. 22–25, Russell & Norvig 2003, pp. 2–3 and 948, Haugeland 1985, pp. 6–9, Cordeschi 2002, pp. 170–176. See also Turing 1950
- [35] Norvig & Russell (2003, p. 948) claim that Turing answered all the major objections to AI that have been offered in the years since the paper appeared.
- [36] McCorduck 2004, pp. 137–170, Crevier, pp. 44–47

- [37] McCorduck 2004, pp. 123–125, Crevier 1993, pp. 44–46 and Russell & Norvig 2003, p. 17
- [38] Quoted in Crevier 1993, p. 46 and Russell & Norvig 2003, p. 17
- [39] Russell & Norvig 2003, p. 947,952
- [40] McCorduck 2004, pp. 111–136, Crevier 1993, pp. 49–51 and Russell & Norvig, p. 17
- [41] See McCarthy et al. 1955. Also see Crevier 1993, p. 48 where Crevier states "[the proposal] later became known as the 'physical symbol systems hypothesis'". The physical symbol system hypothesis was articulated and named by Newell and Simon in their paper on GPS. (Newell & Simon 1963) It includes a more specific definition of a "machine" as an agent that manipulates symbols. See the philosophy of artificial intelligence.
- [42] McCorduck (2004, pp. 129–130) discusses how the Dartmouth conference alumni dominated the first two decades of AI research, calling them the "invisible college".
- [43] "I won't swear and I hadn't seen it before," McCarthy told Pamela McCorduck in 1979. (McCorduck 2004, p. 114) However, McCarthy also stated unequivocally "I came up with the term" in a CNET interview. (Skillings 2006)
- [44] Crevier (1993, pp. 49) writes "the conference is generally recognized as the official birthdate of the new science."
- [45] Russell and Norvig write "it was astonishing whenever a computer did anything remotely clever." Russell & Norvig 2003, p. 18
- [46] Crevier 1993, pp. 52–107, Moravec 1988, p. 9 and Russell & Norvig 2003, p. 18–21
- [47] McCorduck 2004, p. 218, Crevier 1993, pp. 108–109 and Russell & Norvig 2003, p. 21
- [48] Crevier 1993, pp. 52–107, Moravec 1988, p. 9
- [49] Means-ends analysis, reasoning as search: McCorduck 2004, pp. 247–248. Russell & Norvig 2003, pp. 59–61
- [50] Heuristic: McCorduck 2004, p. 246, Russell & Norvig 2003, pp. 21–22
- [51] GPS: McCorduck 2004, pp. 245–250, Crevier 1993, p. GPS?, Russell & Norvig 2003, p. GPS?
- [52] Crevier 1993, pp. 51–58,65–66 and Russell & Norvig 2003, pp. 18–19
- [53] McCorduck 2004, pp. 268–271, Crevier 1993, pp. 95–96, Moravec 1988, pp. 14–15
- [54] McCorduck 2004, p. 286, Crevier 1993, pp. 76–79, Russell & Norvig 2003, p. 19
- [55] Crevier 1993, pp. 79–83
- [56] Crevier 1993, pp. 164–172
- [57] McCorduck 2004, pp. 291–296, Crevier 1993, pp. 134–139
- [58] McCorduck 2004, pp. 299–305, Crevier 1993, pp. 83–102, Russell & Norvig 2003, p. 19 and Copeland 2000
- [59] McCorduck 2004, pp. 300–305, Crevier 1993, pp. 84–102, Russell & Norvig 2003, p. 19
- [60] Simon & Newell 1958, p. 7–8 quoted in Crevier 1993, p. 108. See also Russell & Norvig 2003, p. 21
- [61] Simon 1965, p. 96 quoted in Crevier 1993, p. 109
- [62] Minsky 1967, p. 2 quoted in Crevier 1993, p. 109
- [63] Minsky strongly believes he was misquoted. See McCorduck 2004, pp. 272–274, Crevier 1993, p. 96 and Darrach 1970.
- [64] Crevier 1993, pp. 64–65
- [65] Crevier 1993, p. 94
- [66] Howe 1994
- [67] McCorduck 2004, p. 131, Crevier 1993, p. 51. McCorduck also notes that funding was mostly under the direction of alumni of the Dartmouth conference of 1956.
- [68] Crevier 1993, p. 65
- [69] Crevier 1993, pp. 68–71 and Turkle 1984
- [70] Crevier 1993, pp. 100–144 and Russell & Norvig 2003, pp. 21–22
- [71] McCorduck 2004, pp. 104–107, Crevier 1993, pp. 102–105, Russell & Norvig 2003, p. 22
- [72] Crevier 1993, pp. 163–196
- [73] Crevier 1993, p. 146
- [74] Russell & Norvig 2003, pp. 20–21
- [75] Crevier 1993, pp. 146–148, see also Buchanan 2005, p. 56: "Early programs were necessarily limited in scope by the size and speed of memory"
- [76] Moravec 1976. McCarthy has always disagreed with Moravec, back to their early days together at SAIL. He states "I would say that 50 years ago, the machine capability was much too small, but by 30 years ago, machine capability wasn't the real problem." in a CNET interview. (Skillings 2006)
- [77] Hans Moravec, *ROBOT: Mere Machine to Transcendent Mind*
- [78] Russell & Norvig 2003, pp. 9,21–22 and Lighthill 1973
- [79] McCorduck 2004, pp. 300 & 421, Crevier 1993, pp. 113–114, Moravec 1988, p. 13, Lenat 1989 (Introduction) and Russell & Norvig 2003, p. 21
- [80] McCorduck 2004, p. 456, Moravec 1988, pp. 15–16
- [81] McCarthy & Hayes 1969, Crevier 1993, pp. 117–119
- [82] McCorduck 2004, pp. 280–281, Crevier 1993, p. 110, Russell & Norvig 2003, p. 21 and NRC 1999 under "Success in Speech Recognition".
- [83] Crevier 1993, p. 117, Russell & Norvig 2003, p. 22, Howe 1994 and see also Lighthill 1973.
- [84] Russell & Norvig 2003, p. 22, Lighthill 1973, John McCarthy wrote in response that "the combinatorial explosion problem has been recognized in AI from the beginning" in Review of Lighthill report (<http://www-formal.stanford.edu/jmc/reviews/lighthill/lighthill.html>)

- [85] Crevier 1993, pp. 115–116 (on whom this account is based). Other views include McCorduck 2004, pp. 306–313 and NRC 1999 under "Success in Speech Recognition".
- [86] Crevier 1993, p. 115. Moravec explains, "Their initial promises to DARPA had been much too optimistic. Of course, what they delivered stopped considerably short of that. But they felt they couldn't in their next proposal promise less than in the first one, so they promised more."
- [87] NRC 1999 under "Shift to Applied Research Increases Investment." While the autonomous tank was a failure, the battle management system (called "DART") proved to be enormously successful, saving billions in the first Gulf War, repaying the investment and justifying the DARPA's pragmatic policy, at least as far as DARPA was concerned.
- [88] Lucas and Penrose' critique of AI: Crevier 1993, p. 22, Russell & Norvig 2003, pp. 949–950, Hofstadter 1980, pp. 471–477 and see Lucas 1961
- [89] "Know-how" is Dreyfus' term. (Dreyfus makes a distinction between "knowing how" and "knowing that", a modern version of Heidegger's distinction of ready-to-hand and present-at-hand.) (Dreyfus & Dreyfus 1986)
- [90] Dreyfus' critique of artificial intelligence: McCorduck 2004, pp. 211–239, Crevier 1993, pp. 120–132, Russell & Norvig 2003, pp. 950–952 and see Dreyfus 1965, Dreyfus 1972, Dreyfus & Dreyfus 1986
- [91] Searle's critique of AI: McCorduck 2004, pp. 443–445, Crevier 1993, pp. 269–271, Russell & Norvig 2004, pp. 958–960 and see Searle 1980
- [92] Quoted in Crevier 1993, p. 143
- [93] Quoted in Crevier 1993, p. 122
- [94] "I became the only member of the AI community to be seen eating lunch with Dreyfus. And I deliberately made it plain that theirs was not the way to treat a human being." Joseph Weizenbaum, quoted in Crevier 1993, p. 123.
- [95] Weizenbaum's critique of AI: McCorduck 2004, pp. 356–373, Crevier 1993, pp. 132–144, Russell & Norvig 2003, p. 961 and see Weizenbaum 1976
- [96] McCorduck 2004, p. 51, Russell & Norvig 2003, pp. 19, 23
- [97] McCorduck 2004, p. 51, Crevier 1993, pp. 190–192
- [98] Crevier 1993, pp. 193–196
- [99] Crevier 1993, pp. 145–149, 258–63
- [100] Wason (1966) showed that people do poorly on completely abstract problems, but if the problem is restated to allowed the use of intuitive social intelligence, performance dramatically improves. (See Wason selection task) Tversky, Slovic & Kahnemann (1982) have shown that people are terrible at elementary problems that involve uncertain reasoning. (See list of cognitive biases for several examples). Eleanor Rosch's work is described in Lakoff 1987
- [101] An early example of McCarthy's position was in the journal *Science* where he said "This is AI, so we don't care if it's psychologically real" (Kolata 1982), and he recently reiterated his position at the AI@50 conference where he said "Artificial intelligence is not, by definition, simulation of human intelligence" (Maker 2006).
- [102] Crevier 1993, pp. 175
- [103] Neat vs. scruffy: McCorduck 2004, pp. 421–424 (who picks up the state of the debate in 1984). Crevier 1993, pp. 168 (who documents Schank's original use of the term). Another aspect of the conflict was called "the procedural/declarative distinction" but did not prove to be influential in later AI research.
- [104] McCorduck 2004, pp. 305–306, Crevier 1993, pp. 170–173, 246 and Russell & Norvig 2003, p. 24. Minsky's frame paper: Minsky 1974.
- [105] McCorduck 2004, pp. 327–335 (Dendral), Crevier 1993, pp. 148–159, Russell & Norvig 2003, pp. 22–23
- [106] Crevier 1993, pp. 158–159 and Russell & Norvig 2003, p. 23–24
- [107] Crevier 1993, p. 198
- [108] McCorduck 2004, pp. 434–435, Crevier 1993, pp. 161–162, 197–203 and Russell & Norvig 2003, p. 24
- [109] McCorduck 2004, p. 299
- [110] McCorduck 2004, pp. 421
- [111] Knowledge revolution: McCorduck 2004, pp. 266–276, 298–300, 314, 421, Russell Norvig, pp. 22–23
- [112] Cyc: McCorduck 2004, p. 489, Crevier 1993, pp. 239–243, Russell & Norvig 2003, p. 363–365 and Lenat & Guha 1989
- [113] McCorduck 2004, pp. 436–441, Crevier 1993, pp. 211, Russell & Norvig 2003, p. 24 and see also Feigenbaum & McCorduck 1983
- [114] Crevier 1993, pp. 195
- [115] Crevier 1993, pp. 240.
- [116] Russell & Norvig 2003, p. 25
- [117] McCorduck 2004, pp. 426–432, NRC 1999 under "Shift to Applied Research Increases Investment"
- [118] Crevier 1993, pp. 214–215.
- [119] Crevier 1993, pp. 215–216.
- [120] Crevier 1993, pp. 203. AI winter was first used as the title of a seminar on the subject for the Association for the Advancement of Artificial Intelligence.
- [121] McCorduck 2004, p. 435, Crevier 1993, pp. 209–210
- [122] McCorduck 2004, p. 435 (who cites institutional reasons for their ultimate failure), Crevier 1993, pp. 204–208 (who cites the difficulty of truth maintenance, i.e., learning and updating), Lenat & Guha 1989, Introduction (who emphasizes the brittleness and the inability to handle excessive qualification.)
- [123] McCorduck 2004, pp. 430–431

- [124] McCorduck 2004, p. 441, Crevier 1993, p. 212. McCorduck writes "Two and a half decades later, we can see that the Japanese didn't quite meet all of those ambitious goals."
- [125] McCorduck 2004, pp. 454–462
- [126] Moravec (1988, p. 20) writes: "I am confident that this bottom-up route to artificial intelligence will one day meet the traditional top-down route more than half way, ready to provide the real world competence and the commonsense knowledge that has been so frustratingly elusive in reasoning programs. Fully intelligent machines will result when the metaphorical golden spike is driven uniting the two efforts."
- [127] Crevier 1993, pp. 183–190.
- [128] <http://people.csail.mit.edu/brooks/papers/elephants.pdf>
- [129] Brooks 1990, p. 3
- [130] See, for example, Lakoff & Turner 1999
- [131] McCorduck (2004, p. 424) discusses the fragmentation and the abandonment of AI's original goals.
- [132] McCorduck 2004, pp. 480–483
- [133] DARPA Grand Challenge -- home page (<http://www.darpa.mil/grandchallenge/>)
- [134] (<http://archive.darpa.mil/grandchallenge/>)
- [135] Markoff, John (16 February 2011). "On 'Jeopardy!' Watson Win Is All but Trivial" (<http://www.nytimes.com/2011/02/17/science/17jeopardy-watson.html>). *The New York Times*. .
- [136] Kurzweil 2005, p. 274 writes that the improvement in computer chess, "according to common wisdom, is governed only by the brute force expansion of computer hardware."
- [137] Cycle time of Ferranti Mark 1 was 1.2 milliseconds, which is arguably equivalent to about 833 flops. Deep Blue ran at 11.38 gigaflops (and this does not even take into account Deep Blue's special-purpose hardware for chess). Very approximately, these differ by a factor of 10^7 .
- [138] McCorduck 2004, pp. 471–478, Russell & Norvig 2003, p. 55, where they write: "The whole-agent view is now widely accepted in the field". The intelligent agent paradigm is discussed in major AI textbooks, such as: Russell & Norvig 2003, pp. 32–58, 968–972, Poole, Mackworth & Goebel 1998, pp. 7–21, Luger & Stubblefield 2004, pp. 235–240
- [139] Carl Hewitt's Actor model anticipated the modern definition of intelligent agents. (Hewitt, Bishop & Steiger 1973) Both John Doyle (Doyle 1983) and Marvin Minsky's popular classic *The Society of Mind* (Minsky 1986) used the word "agent". Other "modular" proposals included Rodney Brook's subsumption architecture, object-oriented programming and others.
- [140] Russell & Norvig 2003, pp. 27, 55
- [141] This is how the most widely accepted textbooks of the 21st century define artificial intelligence. See Russell & Norvig 2003, p. 32 and Poole, Mackworth & Goebel 1998, p. 1
- [142] McCorduck 2004, p. 478
- [143] McCorduck 2004, pp. 486–487, Russell & Norvig 2003, pp. 25–26
- [144] Russell & Norvig 2003, p. 25–26
- [145] McCorduck (2004, p. 487): "As I write, AI enjoys a Neat hegemony."
- [146] Pearl 1988
- [147] See Computer science (in Applications of artificial intelligence).
- [148] NRC 1999 under "Artificial Intelligence in the 90s", and Kurzweil 2005, p. 264
- [149] Russell & Norvig 2003, p. 28
- [150] For the new state of the art in AI based speech recognition, see The Economist (2007)
- [151] "AI-inspired systems were already integral to many everyday technologies such as internet search engines, bank software for processing transactions and in medical diagnosis." Nick Bostrom, quoted in CNN 2006
- [152] Olsen (2004), Olsen (2006)
- [153] McCorduck 2004, p. 423, Kurzweil 2005, p. 265, Hofstadter 1979, p. 601
- [154] CNN 2006
- [155] Markoff 2005
- [156] The Economist 2007
- [157] Tascarella 2006
- [158] Crevier 1993, pp. 108–109
- [159] He goes on to say: "The answer is, I believe we could have ... I once went to an international conference on neural net[s]. There were 40 thousand registrants ... but ... if you had an international conference, for example, on using multiple representations for common sense reasoning, I've only been able to find 6 or 7 people in the whole world." Minsky 2001
- [160] Maker 2006
- [161] Kurzweil 2005
- [162] Hawkins & Blakeslee 2004

References

- Berlinski, David (2000), *The Advent of the Algorithm*, Harcourt Books, ISBN 0-15-601391-6, OCLC 46890682.
- Buchanan, Bruce G. (Winter 2005), "A (Very) Brief History of Artificial Intelligence" (<http://www.aaai.org/AITopics/assets/PDF/AIMag26-04-016.pdf>), *AI Magazine*: 53–60, retrieved 2007-08-30.
- Brooks, Rodney (1990), "Elephants Don't Play Chess" (<http://people.csail.mit.edu/brooks/papers/elephants.pdf>), *Robotics and Autonomous Systems* **6**: 3–15, doi:10.1016/S0921-8890(05)80025-9, retrieved 2007-08-30.
- Butler, Samuel (13 June 1863), "Darwin Among the Machines" (<http://www.nzetc.org/tm/scholarly/tei-ButFir-t1-g1-t1-g1-t4-body.html>), *the Press, Christchurch, New Zealand*, retrieved 10 October 2008.
- CNN (26 July 2006), *AI set to exceed human brain power* (<http://www.cnn.com/2006/TECH/science/07/24/ai.bostrom/>), CNN.com, retrieved 16 October 2007.
- Copeland, Jack (2000), *Micro-World AI* ([http://www.alanturing.net/turing_archive/pages/Reference Articles/what_is_AI/What is AI06.html](http://www.alanturing.net/turing_archive/pages/Reference%20Articles/what_is_AI/What%20is%20AI06.html)), retrieved 8 October 2008.
- Cordeschi, Roberto (2002), *The Discovery of the Artificial*, Dordrecht: Kluwer..
- Crevier, Daniel (1993), *AI: The Tumultuous Search for Artificial Intelligence*, New York, NY: BasicBooks, ISBN 0-465-02997-3
- Darrach, Brad (20 November 1970), "Meet Shakey, the First Electronic Person", *Life Magazine*: 58–68.
- Doyle, J. (1983), "What is rational psychology? Toward a modern mental philosophy", *AI Magazine* **4** (3): 50–53.
- Dreyfus, Hubert (1965), *Alchemy and AI*, RAND Corporation Memo.
- Dreyfus, Hubert (1972), *What Computers Can't Do*, New York: MIT Press, ISBN 0-06-090613-8, OCLC 5056816.
- The Economist (7 June 2007), "Are You Talking to Me?" (http://www.economist.com/science/tq/displaystory.cfm?story_id=9249338), *The Economist*, retrieved 16 October 2008.
- Feigenbaum, Edward A.; McCorduck, Pamela (1983), *The Fifth Generation: Artificial Intelligence and Japan's Computer Challenge to the World*, Michael Joseph, ISBN 0-7181-2401-4.
- Hawkins, Jeff; Blakeslee, Sandra (2004), *On Intelligence*, New York, NY: Owl Books, ISBN 0-8050-7853-3, OCLC 61273290.
- Hebb, D.O. (1949), *The Organization of Behavior*, New York: Wiley, ISBN 0805843000, OCLC 48871099.
- Hewitt, Carl; Bishop, Peter; Steiger, Richard (1973) (PDF), *A Universal Modular Actor Formalism for Artificial Intelligence* (<http://dli.iit.ac.in/ijcai/IJCAI-73/PDF/027B.pdf>), IJCAI
- Hobbes, Thomas (1651), *Leviathan*.
- Hofstadter, Douglas (1999 (1979)), *Gödel, Escher, Bach: an Eternal Golden Braid*, Basic Books, ISBN 0465026567, OCLC 225590743.
- Howe, J. (November 1994), *Artificial Intelligence at Edinburgh University: a Perspective* (<http://www.inf.ed.ac.uk/about/AIhistory.html>), retrieved 30 August 2007.
- Kolata, G. (1982), "How can computers get common sense?", *Science* **217** (4566): 1237–1238, doi:10.1126/science.217.4566.1237, PMID 17837639.
- Kurzweil, Ray (2005), *The Singularity is Near*, Viking Press, ISBN 0143037889, OCLC 71826177.
- Lakoff, George (1987), *Women, Fire, and Dangerous Things: What Categories Reveal About the Mind*, University of Chicago Press., ISBN 0-226-46804-6.
- Lenat, Douglas; Guha, R. V. (1989), *Building Large Knowledge-Based Systems*, Addison-Wesley, ISBN 0201517523, OCLC 19981533.
- Levitt, Gerald M. (2000), *The Turk, Chess Automaton*, Jefferson, N.C.: McFarland, ISBN 0-7864-0778-6.
- Lighthill, Professor Sir James (1973), "Artificial Intelligence: A General Survey", *Artificial Intelligence: a paper symposium*, Science Research Council
- Lucas, John (1961), "Minds, Machines and Gödel" (<http://users.ox.ac.uk/~jrlucas/Godel/mmg.html>), *Philosophy* **36** (XXXVI): 112–127, doi:10.1017/S0031819100057983, retrieved 15 October 2008

- Maker, Meg Houston (2006), *AI@50: AI Past, Present, Future* (http://www.engagingexperience.com/2006/07/ai50_ai_past_pr.html), Dartmouth College, retrieved 16 October 2008
- Markoff, John (14 October 2005), "Behind Artificial Intelligence, a Squadron of Bright Real People" (http://www.nytimes.com/2005/10/14/technology/14artificial.html?_r=1&ei=5070&en=11ab55edb7cead5e&ex=1185940800&adxnnl=1&adxnnlx=1185805173-o7WsfW7qaP0x5/NUs1cQCQ&oref=slogin), *The New York Times*, retrieved 16 October 2008
- McCarthy, John; Minsky, Marvin; Rochester, Nathan; Shannon, Claude (August 35, 1955), *A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence* (<http://www-formal.stanford.edu/jmc/history/dartmouth/dartmouth.html>), retrieved 16 October 2008
- McCarthy, John; Hayes, P. J. (1969), "Some philosophical problems from the standpoint of artificial intelligence" (<http://www-formal.stanford.edu/jmc/mcchay69/mcchay69.html>), in Meltzer, B. J.; Mitchie, Donald, *Machine Intelligence 4*, Edinburgh University Press, pp. 463–502, retrieved 16 October 2008
- McCorduck, Pamela (2004), *Machines Who Think* (2nd ed.), Natick, MA: A. K. Peters, Ltd., ISBN 1-56881-205-1, OCLC 52197627.
- McCullough, W. S.; Pitts, W. (1943), "A logical calculus of the ideas immanent in nervous activity", *Bulletin of Mathematical Biophysics* **5** (4): 115–127, doi:10.1007/BF02478259
- Menabrea, Luigi Federico; Lovelace, Ada (1843), "Sketch of the Analytical Engine Invented by Charles Babbage" (<http://www.fourmilab.ch/babbage/sketch.html>), *Scientific Memoirs* **3**, retrieved 2008-08-29 With notes upon the Memoir by the Translator
- Minsky, Marvin (1967), *Computation: Finite and Infinite Machines*, Englewood Cliffs, N.J.: Prentice-Hall
- Minsky, Marvin; Papert, Seymour (1969), *Perceptrons: An Introduction to Computational Geometry*, The MIT Press, ISBN 0262631113, OCLC 16924756
- Minsky, Marvin (1974), *A Framework for Representing Knowledge* (<http://web.media.mit.edu/~minsky/papers/Frames/frames.html>), retrieved 16 October 2008
- Minsky, Marvin (1986), *The Society of Mind*, Simon and Schuster, ISBN 0671657135, OCLC 223353010
- Minsky, Marvin (2001), *It's 2001. Where Is HAL?* (http://www.ddj.com/hpc-high-performance-computing/197700454?cid=RSSfeed_DDJ_AI), Dr. Dobb's Technetcast, retrieved 8 August 2009
- Moravec, Hans (1976), *The Role of Raw Power in Intelligence* (<http://www.frc.ri.cmu.edu/users/hpm/project.archive/general.articles/1975/Raw.Power.html>), retrieved 16 October 2008
- Moravec, Hans (1988), *Mind Children*, Harvard University Press, ISBN 0674576187, OCLC 245755104
- NRC (1999), "Developments in Artificial Intelligence", *Funding a Revolution: Government Support for Computing Research*, National Academy Press, ISBN 0-309-06278-0, OCLC 246584055
- Newell, Allen; Simon, H. A. (1963), "GPS: A Program that Simulates Human Thought", in Feigenbaum, E.A.; Feldman, J., *Computers and Thought*, New York: McGraw-Hill, ISBN 0262560925, OCLC 246968117
- Nick, Martin (2005), *Al Jazari: The Ingenious 13th Century Muslim Mechanic* (<http://www.alshindagah.com/marapr2005/jaziri.html>), Al Shindagah, retrieved 16 October 2008.
- O'Connor, Kathleen Malone (1994), *The alchemical creation of life (takwin) and other concepts of Genesis in medieval Islam* (<http://repository.upenn.edu/dissertations/AAI9503804>), University of Pennsylvania, retrieved 2007-01-10
- Olsen, Stefanie (May 10, 2004), *Newsmaker: Google's man behind the curtain* (http://news.cnet.com/Googles-man-behind-the-curtain/2008-1024_3-5208228.html), CNET, retrieved 17 October 2008.
- Olsen, Stefanie (August 18, 2006), *Spying an intelligent search engine* (http://news.cnet.com/Spying-an-intelligent-search-engine/2100-1032_3-6107048.html), CNET, retrieved 17 October 2008.
- Pearl, J. (1988), *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*, San Mateo, California: Morgan Kaufmann, ISBN 1558604790, OCLC 249625842.
- Russell, Stuart J.; Norvig, Peter (2003), *Artificial Intelligence: A Modern Approach* (<http://aima.cs.berkeley.edu/>) (2nd ed.), Upper Saddle River, New Jersey: Prentice Hall, ISBN 0-13-790395-2.

- Poole, David; Mackworth, Alan; Goebel, Randy (1998), *Computational Intelligence: A Logical Approach* (<http://www.cs.ubc.ca/~poole/ci.html>), Oxford University Press., ISBN 0-19-510270-3.
- Samuel, Arthur L. (July 1959), "Some studies in machine learning using the game of checkers" (<http://domino.research.ibm.com/tchjr/journalindex.nsf/600cc5649e2871db852568150060213c/39a870213169f45685256bfa00683d74?OpenDocument>), *IBM Journal of Research and Development* **3** (3): 210–219, doi:10.1147/rd.33.0210, retrieved 2007-08-20.
- Searle, John (1980), "Minds, Brains and Programs" (<http://web.archive.org/web/20071210043312/http://members.aol.com/NeoNoetics/MindsBrainsPrograms.html>), *Behavioral and Brain Sciences* **3** (3): 417–457, doi:10.1017/S0140525X00005756, retrieved May 13, 2009.
- Simon, H. A.; Newell, Allen (1958), "Heuristic Problem Solving: The Next Advance in Operations Research", *Operations Research* **6**: 1, doi:10.1287/opre.6.1.1.
- Simon, H. A. (1965), *The Shape of Automation for Men and Management*, New York: Harper & Row.
- Skillings, Jonathan (2006), *Newsmaker: Getting machines to think like us* (http://news.cnet.com/Getting-machines-to-think-like-us---page-2/2008-11394_3-6090207-2.html?tag=st.next), CNET, retrieved October 08, 2008.
- Tascarella, Patty (11 August 2006), "Robotics firms find fundraising struggle, with venture capital shy" (<http://www.bizjournals.com/pittsburgh/stories/2006/08/14/focus3.html?b=1155528000^1329573>), *Pittsburgh Business Times*, retrieved October 08, 2008.
- Turing, Alan (1936-37), "On Computable Numbers, with an Application to the Entscheidungsproblem" (<http://www.abelard.org/turpap2/tp2-ie.asp>), *Proceedings of the London Mathematical Society*, 2 (42): 230–265, retrieved October 8, 2008.
- Turing, Alan (October 1950), "Computing Machinery and Intelligence" (<http://loebner.net/Prizel/TuringArticle.html>), *Mind* **LIX** (236): 433–460, doi:10.1093/mind/LIX.236.433, ISSN 0026-4423, retrieved 2008-08-18.
- Weizenbaum, Joseph (1976), *Computer Power and Human Reason*, W.H. Freeman & Company, ISBN 0140225358, OCLC 10952283.
- .

Article Sources and Contributors

History of artificial intelligence *Source:* <http://en.wikipedia.org/w/index.php?oldid=457369433> *Contributors:* 17Drew, ALYOSHA, AMCKen, AdultSwim, Alcazluis, Alexander Wilks, Allenchue, AoV2, Audiodude, AustralianMelodrama, Axem Titanium, Ayyu, Barticus88, Bender235, Bkkbrad, Blogjack, BrokenSegue, Bugone, Burzmali, ChangChienFu, CharlesGillingham, Chigs.ca, Christian Storm, CommonsDelinker, Coolerhead, Cordes, Crystallina, Cyanidethistles, Cyde, Cyfal, DHN, DSachan, Dank, Darker Dreams, Darre1, Devourer09, DionysiusThrax, Doctoroxenbriery, Dreadstar, Elmer Clark, Elsendero, Ferris37, Firsfron, Froth, Gabi S., Gaius Cornelius, Gimboid13, Grafen, Graham87, Gwern, Hilgerdenaar, Hmains, Hmwith, Hu, Itinerant1, JWSchmidt, Jagged 85, Jaibe, Jayjg, Jlltlet, Joanna Bryson, JohnCD, JosebaAbaitua, JustinCasey, JustinClarkCasey, Kane5187, Knight1993, Koavf, Kwamikagami, Leandro, Lightmouse, Loadmaster, M3tainfo, Malleus Fatuorum, MegMaker, MegX, Michael Devore, Michael Hardy, Mneser, Moxon, MrsHudson, Mschibr, Myscrnm, Neelix, Nhayesroth, Nihiltres, Open2universe, Paulscrawl, Pawyilee, Pegua, Perstranger123, Pgr94, Ph.eyes, Qero, Quentin X, Qwertyus, Remember the dot, Rettetast, Rich Farmbrough, Ritchy, Rjwilmsi, Rootxploit, SamatJain, Samfreed, Science History, Serouj, Shell Kinney, Skittleys, Spencerk, StAnselm, Syncategoremata, TAnthony, Tassedethe, TedColes, Tobi Kellner, Trovatore, Vadmium, Viriditas, Wa3frp, WaltBusterkeys, Wiki Raja, Willingandable, Woohookitty, Wronkiew, Yuriz, Zy26, 109 anonymous edits

Image Sources, Licenses and Contributors

Image:Al-jazari robots.jpg *Source:* http://en.wikipedia.org/w/index.php?title=File:Al-jazari_robots.jpg *License:* Public Domain *Contributors:* Cwbn (commons), Infrogmation, Man vyi, Quibik

Image:Gottfried Wilhelm von Leibniz.jpg *Source:* http://en.wikipedia.org/w/index.php?title=File:Gottfried_Wilhelm_von_Leibniz.jpg *License:* Public Domain *Contributors:* Beria, Beyond My Ken, Davidlud, Ecummenic, Eusebius, Factumquintus, FalconL, Gabor, Luestling, Mattes, Schaengel89, Shakko, Svench, Tomisti, 5 anonymous edits

Image:Classic shot of the ENIAC.jpg *Source:* http://en.wikipedia.org/w/index.php?title=File:Classic_shot_of_the_ENIAC.jpg *License:* Public Domain *Contributors:* Unidentified U.S. Army photographer

Image:BRL61-IBM 702.jpg *Source:* http://en.wikipedia.org/w/index.php?title=File:BRL61-IBM_702.jpg *License:* Public Domain *Contributors:* User RTC on en.wikipedia

Image:Semantic Net.svg *Source:* http://en.wikipedia.org/w/index.php?title=File:Semantic_Net.svg *License:* Public Domain *Contributors:* Bjankuloski06en, Konstable, Wst, 1 anonymous edits

Image:Hopfield-net.png *Source:* <http://en.wikipedia.org/w/index.php?title=File:Hopfield-net.png> *License:* GNU Free Documentation License *Contributors:* Zawersh at en.wikipedia

License

Creative Commons Attribution-Share Alike 3.0 Unported
[//creativecommons.org/licenses/by-sa/3.0/](http://creativecommons.org/licenses/by-sa/3.0/)