AI Timeline

Compiled by Jake Browning and Philipp Schmitt

Introduction

The psychologist Edwin Boring once noted that the history of psychology is short, but it has a long prehistory. The same is true of artificial intelligence, and much of its history overlaps with that of psychology since both fields explore intelligence. But AI also explores the artificial—the artifacts made by gods, artists, magicians, and scientists capable of thought. In this sense, the prehistory of AI includes everything from philosophy to religion, mathematics to magic, science to art.

The field received its contemporary form—and the appellation *artificial intelligence*—in the twentieth century with a host of scientific discoveries. There are (at least) two main traditions of thinking machine: the biologically inspired and the cognitivist. The former arose in cybernetics, which explores how beings self-regulate using negative feedback from the environment—in effect, learning to preserve oneself by responding to external stimuli. The result was the creation of mechanical beings modeled on the body, the brain, and the process of adaptation. This research has led to the contemporary work in deep neural networks, machine learning, and robotic vacuum cleaners.

On the other side, the cognitivists focused on the high-level abstract rationality found in humans: mathematics, language, and problem-solving. This field grew out of the formalization of logic and mathematics in the early twentieth century, including Alan Turing's insights into mechanical computation. The result was a tradition of computer programs—called "virtual machines"—that communicated, solved puzzles, played games, and reasoned out solutions under uncertainty. Cognitivists dominated the last half of the twentieth century, providing many of the programming languages, operating systems, and search strategies still in use today.

This timeline tries to lay out these histories, highlighting the ways these different traditions have overlapped but also diverged, united by the technology of computers but divided by mathematics, philosophy, and general outlook. The timeline illuminates the amazing successes of each field but also notes the limitations, failures, and disappointments. Nevertheless, the takeaway is that immense progress has been made over the last one hundred years: in 1920, not a single computer existed anywhere, and AI was as fictional as conjuring life out of clay. Today, AI is ubiquitous, integrated into our cameras, cars, phones, and buildings. The world around us is now thinking—and getting smarter every minute.

FAMSF_UncannyValley_Interior_BLUES_180220_WM.indd 172 18.2.20 10:53

1914

■ The engineer Leonardo Torres y Quevedo builds El Ajedrecista (the Chess Player), the first chess-playing automaton. The machine is limited by playing only king and king-and-rook endgames.

1920

■ In his play *R.U.R. (Rossum's Universal Robots)*, Karel Čapek introduces the term "robot," which roughly translates to "serf" or "slave" in Czech (fig. 2). The play portrays a dystopian world where synthetic—but also biological—beings rebel against humanity.

2. Scene from the 1938 BBC Television production of Karel Čapek's 1920 play *R.U.R.*

3. Replica of Konrad Zuse's 1941 Z3 computer, Deutsches Museum, Munich, 2006

1927

■ The electrical engineer Harold S. Black develops the concept of "negative feedback," the idea that feeding destabilizing output back into a system can return it to equilibrium. The concept of a self-stabilizing system, an essential feature of biological systems, also has application in mathematics and engineering.

■ Fritz Lang's silent film *Metropolis*, which features a machine urging a workers' rebellion, marks the first on-screen depiction of a mechanical automaton.

1936

■ Alan Turing introduces the thought experiment of an "automatic machine," later dubbed the Turing machine. The model is composed of a potentially infinite tape consisting of discrete symbols (the system's memory); an executor that can read, erase, and write symbols (the computer's processor); and a control that provides instructions for what to do in response to each symbol (the program). Despite its simplicity, the Turing machine demonstrates that any algorithm that can be solved by a computer can be handled by this simple machine.

1938

■ Konrad Zuse finishes building the Z1, the first working mechanical computer. A few years later, Zuse completes the Z3, the first programmable digital computer (fig. 3). The Nazis, who do not see value in the device, refuse funding for it. The Z3 is later destroyed by a bomb in 1943, during World War II.

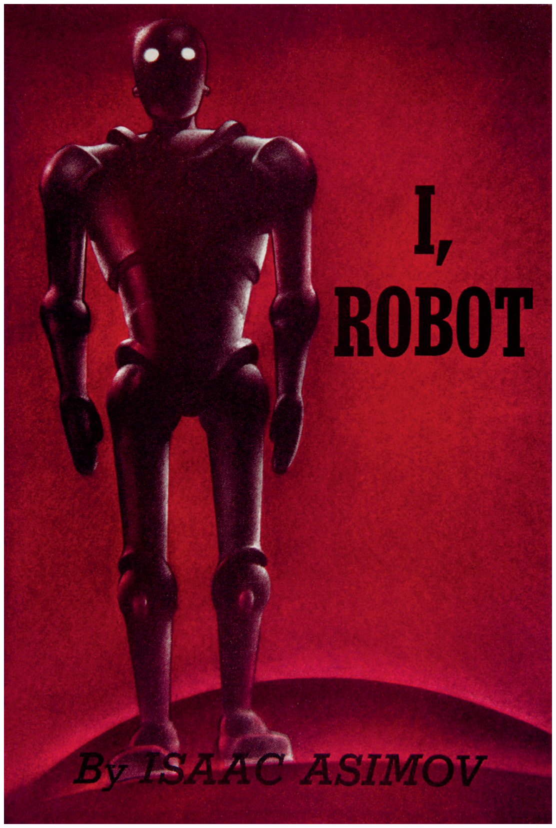
1942

■ The science-fiction writer Isaac Asimov publishes the short story "Runaround" in the magazine *Astounding Science Fiction*. In the story, the characters postulate "three laws of robotics": (1) a robot cannot harm or allow harm to a human, (2) a robot must obey humans unless their command violates the first law, and (3) a robot should preserve its own existence unless doing so violates the first or second law. Later, Asimov includes the story in his 1950 collection, *I, Robot* (fig. 4).

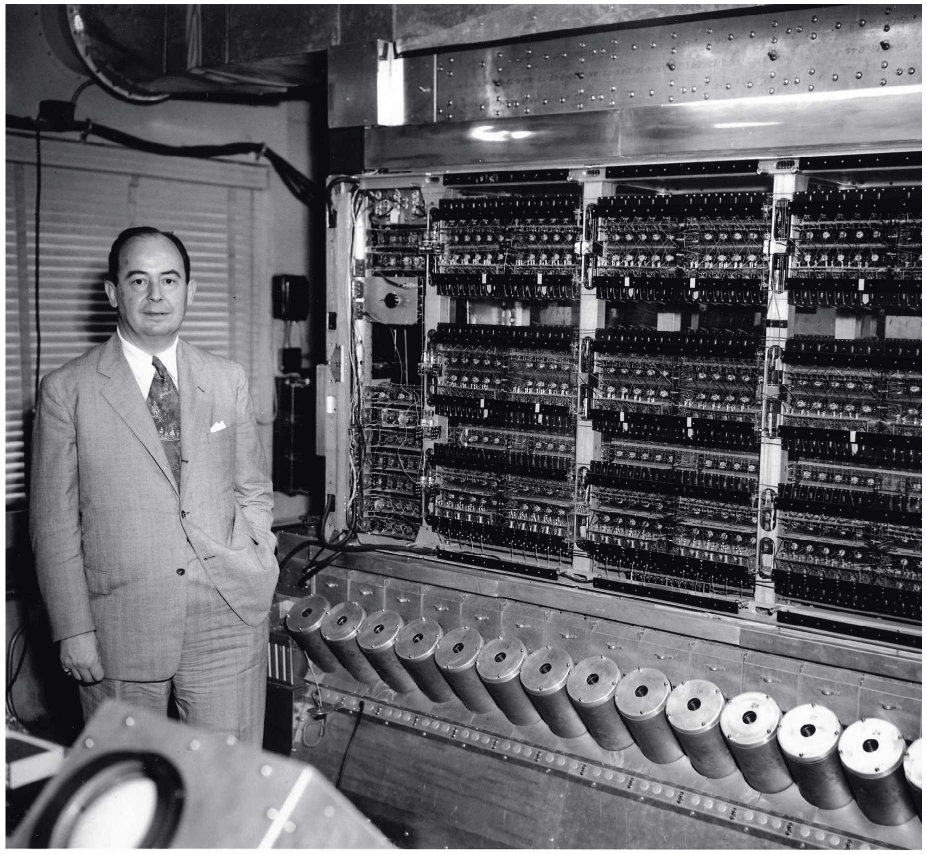
1943

■ Colossus, the first programmable computer developed for the Allied powers of World War II, is designed by Tommy Flowers and Alan Turing and built in Bletchley Park, UK. The Allies use the machine to crack the German Enigma code and hasten the end of the war.

■ The neuroscientist Warren McCulloch and the logician Walter Pitts describe an electrical "neuron" that, like the biological neuron, is a node that receives excitatory and inhibitory signals from multiple different inputs. The relative strengths of the signals are added up and, based on whether the



4. Cover of Isaac Asimov's collection *I, Robot* (1950), illustrating the story "Runaround," which was first published in 1942



5. John von Neumann, whose 1945 theoretical description of the von Neumann machine set a standard in computer architecture, ca. 1952

sum hits some threshold, an output is released. The electrical neuron provides a mathematical model that, if structured correctly, can perform simple logical operations, such as conjunction (A and B) and inclusive disjunction (A and/or B). This indicates that "neural networks," consisting of multiple different neurons, are capable of computation—an influential finding both for biology, since it suggests that brains might function like computers, and for AI, since it suggests that building electrical brains is possible.

1945

■ John von Neumann (fig. 5) describes the von Neumann machine, which becomes the standard in computer architecture. The machine separates out the processor, which runs logical operations; the control unit, which runs program instructions; working memory store, which contains program instructions and current data (now called "RAM," or "random-access memory"); and latent memory storage (now called "ROM," or "read-only memory").

■ George Pólya publishes *How to Solve It*, which outlines four steps of problem-solving: (1) understand a problem, (2) make a plan, (3) carry out the plan, and (4) reflect afterward on possible improvements. The book also contains a set of heuristics for different options for problem-solving. *How to Solve It* becomes essential reading for mathematicians and AI researchers, especially concerning the role of heuristics.

1947

■ At a talk at the London Mathematical Society, Alan Turing contends that the goal of computer research is to build a machine that learns from experience. Around the same time, he begins work on *Intelligent Machinery*, a manifesto on AI concerned with building neural networks that learn.

1948

■ Norbert Wiener's publication of *Cybernetics: Or Control and Communication in the Animal and the Machine* establishes the field of cybernetics, the study of self-stabilizing "automatic control systems" in biological and mechanical beings.

■ The information theorist Claude Shannon publishes "A Mathematical Theory of Communication," the seminal text behind information theory. The paper begins, "The fundamental problem of communication is that of reproducing at one point, either exactly or approximately, a message selected at another point." The problem of information theory is fundamentally mathematical: Ignoring what is being said (the "content"), how much information—only in terms of the measurable amount (the "bits")—is needed to communicate a message? Shannon's results establish a connection between noise and redundancy, an essential finding for telephone services. They also provide for a theory of information that computer scientists and cyberneticians find promising—although immediate applications are not apparent.

1949

■ Donald Hebb's book *Organization of Behavior* provides a neural theory of learning. His theory, often summarized as "neurons that fire together wire together," proposes that neural connections can be strengthened by repeatedly firing at the same time. "Hebbian learning" becomes a key principle for neural networks.

■ In a memorandum circulated to some acquaintances, Warren Weaver proposes machine translation based on "the common base of human communication—the real but as yet undiscovered universal language." Weaver suggests that there is a shared vocabulary and grammar underlying every natural language—a kind of "language of thought," or "mentalese." This idea influences how language is investigated in both human and artificial intelligence and provides a vision for how to define "thought": mental symbols with systematic, language-like structure.

■ William Grey Walter develops the first autonomous electronic robot, known as "the tortoise," demonstrating similarities between mechanical and biological organisms at the level of automatic control (fig. 6).

1950

■ Alan Turing proposes the Turing test in his seminal essay on AI, "Computing Machinery and Intelligence." For the test, a participant reads typed messages from two unknown partners—one woman and one man—both claiming to be women. Turing hypothesizes that the recipient will not be able to determine the gender of the message sender based on typed conversation alone. He then replaces one of the two partners with a machine and reruns the test, contending that if a machine can hold an intelligent conversation and not be discovered as mechanical, it should be regarded as intelligent.

■ In "Programming a Computer for Playing Chess," Claude Shannon suggests two possible ways of solving a problem computationally: the type-A, or "brute force," approach, in which the machine calculates as many different solutions as computationally possible and chooses the best path found; and the type-B, or "heuristic," approach, in which the machine uses a few hand-coded general rules that help it search only those pathways that are most likely to be fruitful.

1951

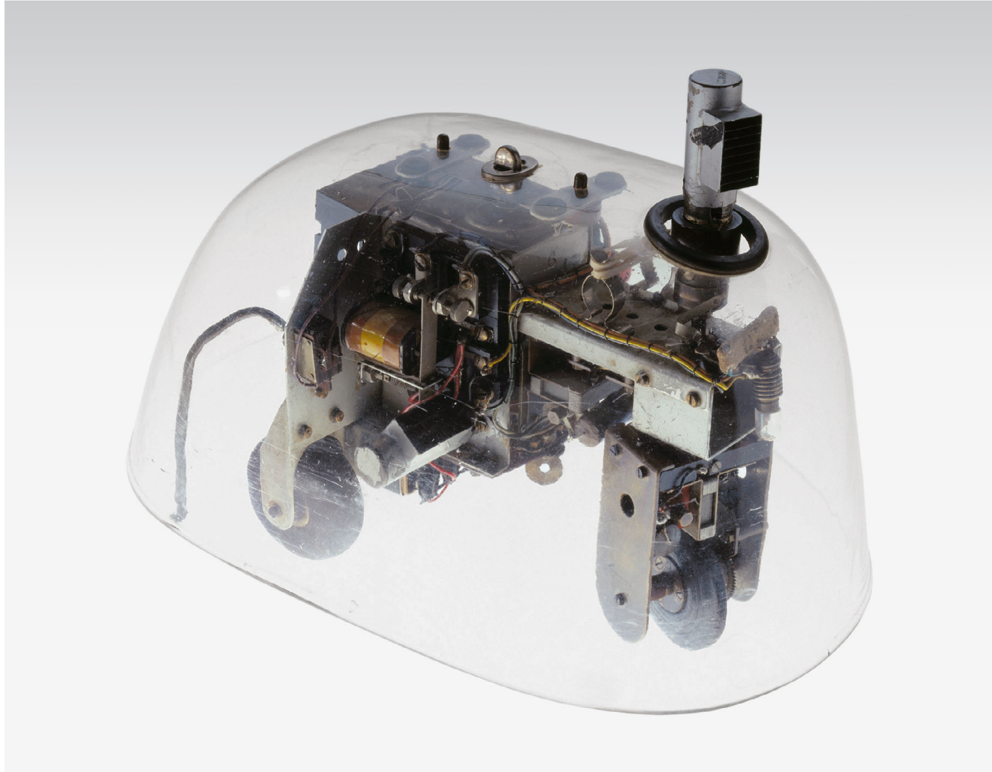
■ Based on the McCulloch–Pitts model, Marvin Minsky, paired with Dean Edmonds, develops the Stochastic Neural Analog Reinforcement Calculator (SNARC), the first neural network (fig. 7). The system uses feedback to steer a machine toward more correct responses, suggesting that insights from cybernetics can be paired successfully with computers.

1952

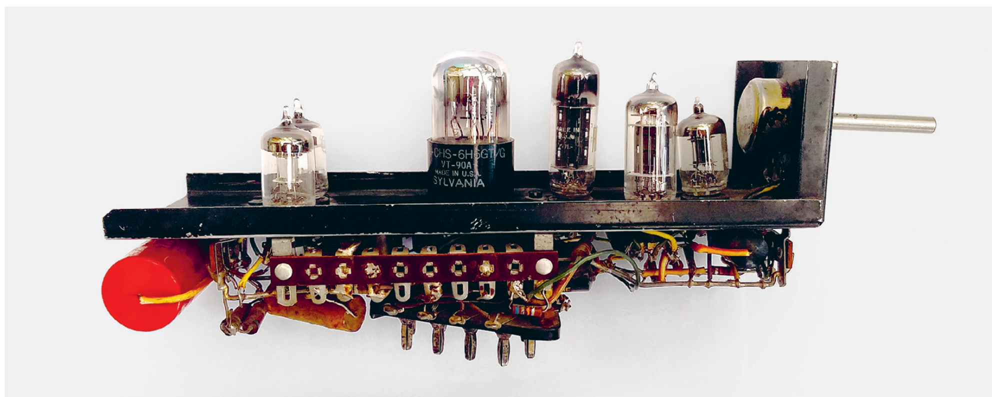
■ The first piece of literature created by AI is produced, using the computer scientist Christopher Strachey's love letter algorithm (fig. 8).

1954

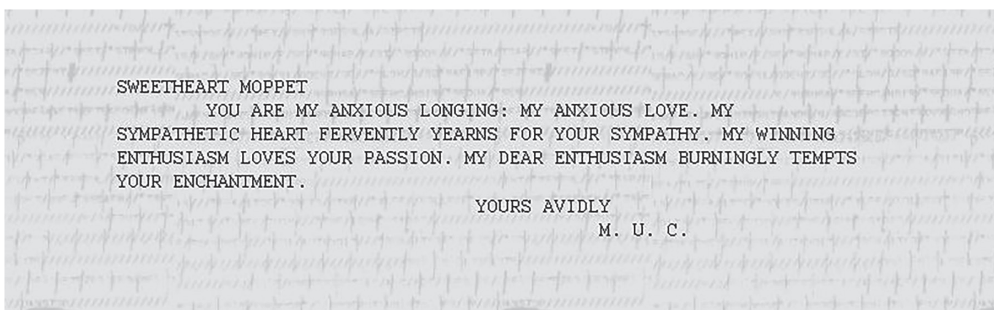
■ IBM performs a widely publicized demonstration of a machine translating Russian sentences



6. William Grey Walter's cybernetic tortoise, ca. 1950



7. One of the forty neurons in the Stochastic Neural Analog Reinforcement Calculator (SNARC), 1951



SWEETHEART MOPPET

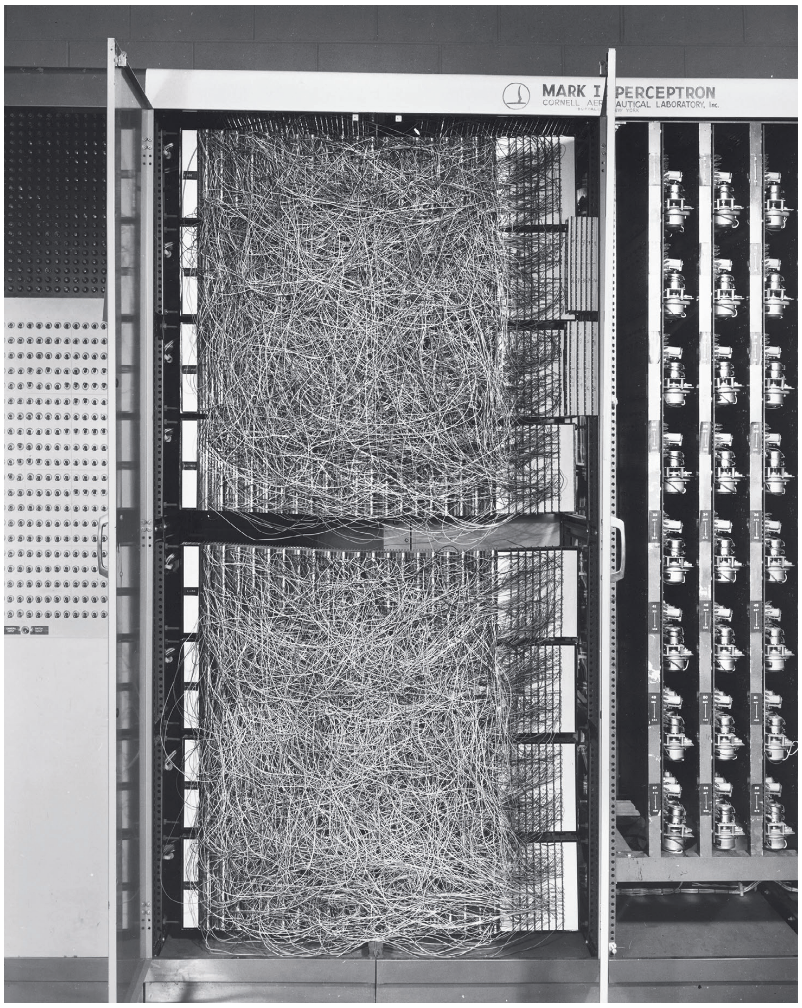
YOU ARE MY ANXIOUS LONGING: MY ANXIOUS LOVE. MY SYMPATHETIC HEART FERVENTLY YEARNS FOR YOUR SYMPATHY. MY WINNING ENTHUSIASM LOVES YOUR PASSION. MY DEAR ENTHUSIASM BURNINGLY TEMPTS YOUR ENCHANTMENT.

YOURS AVIDLY

M. U. C.

8. Artist rendering of a letter created by Christopher Strachey's 1952 love letter algorithm

9. Arthur Samuel with his checker player, 1956



10. The Mark I Perceptron, an implementation of Frank Rosenblatt's 1957 neural network algorithm, Perceptron, Cornell Aeronautical Laboratory, Buffalo, New York, ca. 1958

into English. The machine has a limited vocabulary and a weak grasp of syntax, and can therefore only translate a handful of carefully selected sentences.

1955

■ Arthur Samuel develops a checker player capable of self-improving play, which Samuel dubs "machine learning" (fig. 9). The machine assigns values to next moves based on looking ahead to probable board states for the next few steps. These numerical values are updated when strategies prove more successful than predicted, allowing the machine to perform progressively better over time.

1956

■ The Dartmouth Summer Research Project on Artificial Intelligence takes place in Hanover, New Hampshire, bringing together mathematicians and engineers to explore the idea of thinking machines. The term *artificial intelligence* is coined by John McCarthy for the workshop.

■ At the workshop, Allen Newell, J. C. Shaw, and Herbert Simon introduce the program Logic Theorist, which employs search trees and heuristics to generate proofs using the rules of logic and mathematics. The program mimics mental processes used by mathematicians: recognizing a problem, attempting different strategies, and then executing strategies that achieve the result.

■ At the MIT Special Interest Group in Information Theory symposium, essential papers on using computers to study intelligence—a field later known as cognitive science—are presented, including work by Noam Chomsky on computational linguistics, Newell and Simon on the Logic Theorist, and G. C. Sziklai on computer vision. A central contribution is George A. Miller's "The Magical Number Seven, Plus or Minus Two," which provides a model of "chunking" in memory—up to seven digits can be remembered if they can be assembled into useful mnemonic groups. Miller's result suggests that more information can be remembered if it is chunked, an idea that becomes essential to information storage methods in computers.

1957

■ The psychologist Frank Rosenblatt invents the Perceptron, a neural network algorithm based on work by McCulloch–Pitts and Marvin Minsky (fig. 10). The algorithm and the machines built to run it generate intense funding and media interest. In July of the following year, in an article titled "New Navy Device Learns by Doing," the *New York Times* reports that the machine is expected to someday "walk, talk, see, write, reproduce itself and be conscious of its existence." This leads to increased interest in biologically inspired work. The Perceptron is limited, however, by consisting of only a single layer of nodes between input and output, which prevents certain kinds of learning and logical functions.

■ Noam Chomsky publishes his monograph *Syntactic Structures,* which develops the idea of transformational grammar, where grammar can be understood solely in terms of rules and symbols (i.e., syntax). Chomsky argues that these rules, especially recursion, explain humans' ability to understand and generate a theoretically infinite number of grammatically correct sentences. Chomsky combines this idea with the older notion of context-free grammar—a method for identifying noun phrases, verbs, and other grammatical categories—giving rise to the computational approach to linguistics.

1958

■ In December, at the Teddington Conference on the Mechanization of Thought Processes, held in the UK, multiple influential papers are presented by John McCarthy (fig. 11), Marvin Minsky, and Oliver Selfridge.

■ Selfridge introduces Pandemonium architecture, a neurally inspired thought experiment for visual pattern recognition, specifically of letters. The program involves multiple layers of neurons, called "demons," moving from a layer of awareness of different features of an image, such as straight or curved lines, up to collections of features, such as shapes, eventually being able to distinguish between individual letters. The experiment inspires subsequent ideas for neurally inspired pattern recognition.

■ McCarthy discusses the need for a machine to possess "common sense" if it is to learn as humans do. Common sense, in this context, refers to the immense body of background knowledge and know-how that is necessary for accomplishing some task. McCarthy's program, Advice Taker, represents knowledge separately from the rules needed to solve the problem. The machine solves problems by both figuring out the necessary rules to achieve the result and consulting background knowledge to determine which subtasks therefore need to be performed. This project incites the quest for developing machines with common sense, a seemingly simple problem that remains largely unsolved today.

■ Also at the Teddington Conference, Herbert Gelernter and Nathaniel Rochester present the Geometry Theorem Machine. The machine solves problems by initially specifying the end states of each possible problem-solving strategy and plotting them on a grid. Any strategy that leads to an incoherent result—such as incompatible figures on the grid—can be excluded at the start. The program demonstrates that machines can solve problems not just by following abstract rules (i.e., syntactically) but also by determining the meaning of a rule in order to infer what strategy to use (i.e., semantically).

■ McCarthy develops the LISP programming language, which becomes the dominant medium for AI for decades. The language focuses on data structures called "lists" and allows for the use of functions, including recursive functions.

11. John McCarthy, Stanford University, California, 1966

1959

■ The General Problem Solver (GPS) is created by Allen Newell, J. C. Shaw, and Herbert Simon. The GPS uses "instrumental," or "means-ends," reasoning: deciding on a final outcome at the start and then reasoning backward to determine the intermediate steps necessary to achieve that outcome. The program is general because it can solve any task the programmer stipulates, provided the machine knows all of the permissible operations needed to achieve the specified end state.

■ The neurophysiologists David H. Hubel and Torsten Wiesel discover "edge-detecting" neurons in cats, which fire only when exposed to lines at specific angles. The discovery suggests that biological vision decomposes images into discrete parts that can be processed separately, results used later in building computer vision.

■ General Electric's ERMA is introduced at Bank of America. The program is an automatic check-processing system that reads the stylized numerals at the bottom of checks. ERMA uses template matching, effectively overlaying each input with a template and choosing the best fit.

■ The AI Lab at MIT is founded by John McCarthy and Marvin Minsky. Today, this is called the Computer Science and Artificial Intelligence Laboratory (CSAIL).

1960

■ J. C. R. Licklider writes *Man-Computer Symbiosis*, which predicts an essential role for computers in improving human life, especially by mechanizing routine tasks and providing information for more efficient and improved decision-making.

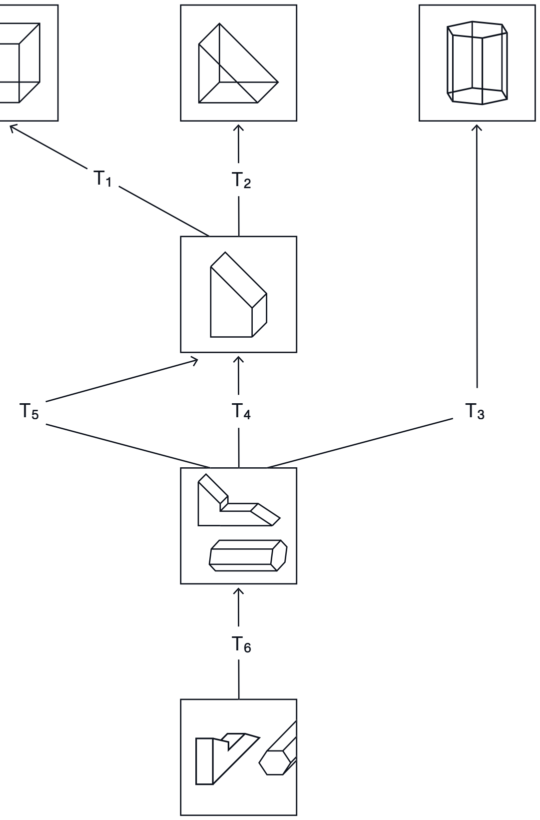
■ Donald Michie creates MENACE, a tic-tac-toe playing program that uses reinforcement learning, the process of training a machine through trial and error. Lacking the proper computer technology, Michie makes the program out of stacks of matchboxes, each with a label indicating a specific move at a specific step in a possible game. Michie places a colored bead in each box depending on whether the move leads to victory or defeat. Over a long enough series of games, the program plays perfect games (i.e., always winning or drawing).

■ Ray Solomonoff, in "General Theory of Inductive Inference," outlines the key ideas behind algorithmic probability. He develops a formal specification that expresses the complexity of any problem in terms of the size of the computer program needed to solve it. This lays the groundwork for inductive, probabilistic computation.

1961

■ The first industrial robot, made by the Unimation company, is deployed at a General Motors factory, assigned the dangerous task of die casting (fig. 12).

12. Unimation's industrial robot model Unimate 2000B, ca. 1979. The first iteration of Unimate was built in 1961.



13. Images from Lawrence Roberts's 1963 MIT dissertation showing machine perception of three-dimensional solids

By the end of the decade, robots will be installed in car factories around the world.

1963

■ Edward Feigenbaum and Julian Feldman publish the first anthology of writings on AI, *Computers and Thought.* The anthology brings together many of the key essays written during the first decade of AI, marking its emergence as a broader scientific field.

■ Thomas Evans's program ANALOGY, written as part of his PhD work at MIT, demonstrates that computers can solve multiple-choice analogy problems on IQ tests. The program determines the relevant features of the shapes of two different objects, discerns what makes them similar, then recognizes the relevant relationship in two other objects.

■ The engineer Lawrence Roberts writes a program for detecting objects in images. The program uses "convolutions," a filter over a small subset of pixels in the image, to transform the original image into a two-dimensional line drawing (fig. 13). The specific convolution used is the "Roberts cross," a simple edge detector that transforms gradual changes in lighting into a sharp break—in this case, a straight line. This incorporates into computer vision and image processing Hubel and Wiesel's idea that some neurons in the visual system act as edge detectors. It also leads to an expansion of interest in developing different and more effective convolutions.

1964

■ Daniel Bobrow, in his MIT dissertation, develops a program that solves natural-language algebra problems. This is possible because word problems often contain numbers, similar wording for mathematical relations, and clearly irrelevant data.

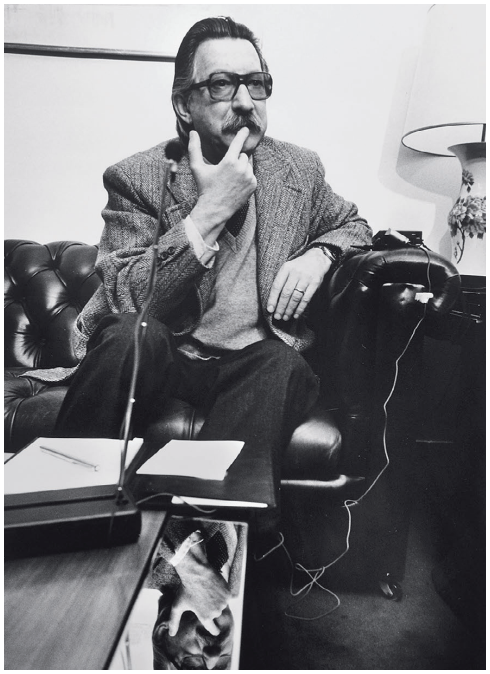
■ Bertram Raphael develops the Semantic Information Retrieval (SIR) program, designed to answer questions about information provided by a user. The user inputs whether each piece of information is an object or a property as well as what relationship this information has to other pieces of information, with the result that each piece of information exists within a network of logical relations.

1965

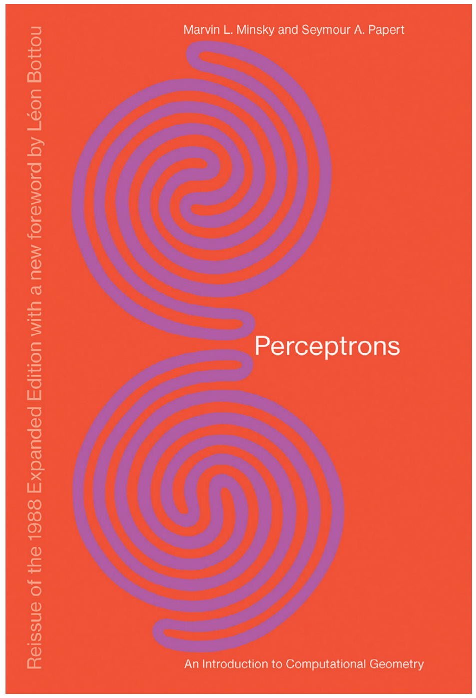
■ J. Alan Robinson develops the resolution method, an algorithm that utilizes conditional statements, such as "if A then B," in representing the relationship between different pieces of information. This allows the computer program to use both first-order logic, concerning the logical relations between sentences ("if sentence A, then sentence B"), and second-order logic, concerning complex sentences that distinguish predicates and objects ("if the predicate X of object A, then the predicate Y of object B"). Together with Bertram Raphael's work on SIR, this program highlights the usefulness of logical representation as a way of storing and retrieving information.

■ I. J. Good, in his "Speculations Concerning the First Ultraintelligent Machine," suggests machines will eventually become smarter than humans. He argues that "the first ultraintelligent machine is the last invention that man need ever make, provided that the machine is docile enough to tell us how to keep it under control."

■ The computer scientist Joseph Weizenbaum (fig. 14) builds ELIZA, the first chatbot. In its most famous incarnation, in which it imitates a psychotherapist, the program tracks keywords in a person's comment and reframes them into follow-up questions. ELIZA highlights how superficial conversations can create the illusion of an understanding—even caring—machine.



14. Joseph Weizenbaum, Hamburg, Germany, 1980. In 1965, Weizenbaum built ELIZA, the first chatbot.



15. Cover of Marvin Minsky and Seymour Papert's *Perceptrons* (1968)

1966

■ At the University of Washington, M. Ross Quillian writes his dissertation treating semantic nets as a model for conceptual organization that can store and retrieve information effectively. Concepts in semantic nets are defined by their location relative to other concepts: the concept "bird" is defined by both belonging to the concept "animal" and also containing the concepts "robin," "fly," and "wings." Although these semantic nets permit logical inferences based on links between higher and lower genera, Quillian's model does not require strict, exclusionary deductions; for example, penguins still count as birds despite lacking the typical feature of flight.

■ The Automatic Language Processing Advisory Committee (ALPAC), set up by the US government, produces a highly critical report on machine translation. The report draws attention to the gap between initial predictions of success in machine translation and the actual output, especially focusing on difficulties concerning ambiguity in sentences. This results in a drastic cut in funding for the field. The report praises work in computational linguistics and recommends more effort be put there.

1967

■ In the wake of John McCarthy's work on Advice Taker, a group of programs using large, narrowly focused sums of knowledge demonstrate enormous potential for problem-solving: DENDRAL identifies molecules based on mass spectra data; Macsyma solves problems in algebra; and MacHack plays tournament-level chess, becoming the first chess program to defeat a human in tournament play. These systems are early versions of "expert systems," machines with a large body of knowledge in some field hand coded into them.

1968

■ Marvin Minsky and Seymour Papert publish *Perceptrons,* a scathing critique of neural networks that suggests they are fruitless techniques for AI (fig. 15). Importantly, the criticisms only apply to relatively simple, single-layer neural networks, such as Frank Rosenblatt's Perceptron. However, in the late 1960s, it is unclear how to train multilayer neural networks. Fallout from the text's publication includes limited interest in and decreased funding for neural network research.

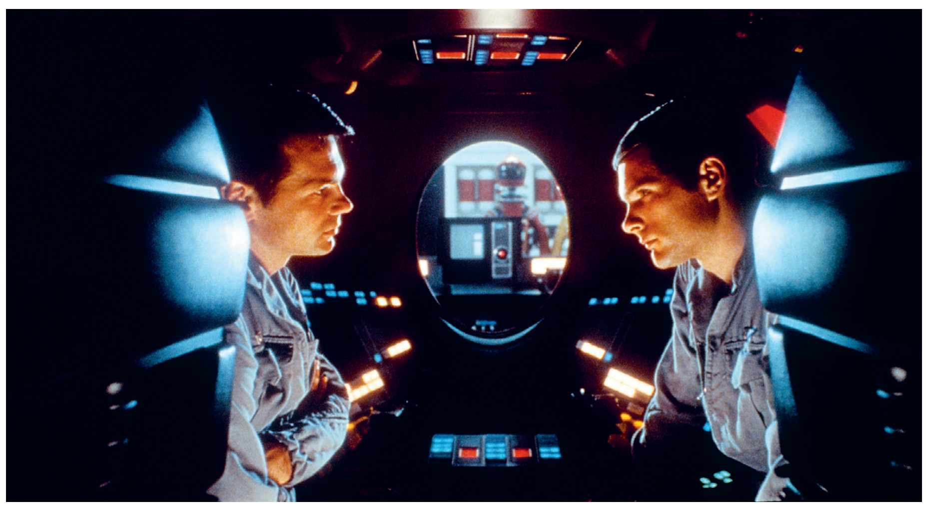
■ Stanley Kubrick's film *2001: A Space Odyssey* features the character HAL 9000, a sentient machine on a spaceship that begins to malfunction (fig. 16). When the astronauts attempt to disconnect HAL, the machine protects itself by becoming homicidal.

■ Terry Winograd develops SHRDLU, a program that controls a simulated arm in a toy-block world (fig. 17). The arm can be instructed to move the toy blocks according to natural-language instructions. The program can also develop new vocabulary and receive new instructions provided they use terms derived from previously supplied instructions. Winograd's work highlights the increasing interest in "microworlds," discrete domains in which a specific problem can be solved while ignoring other issues.

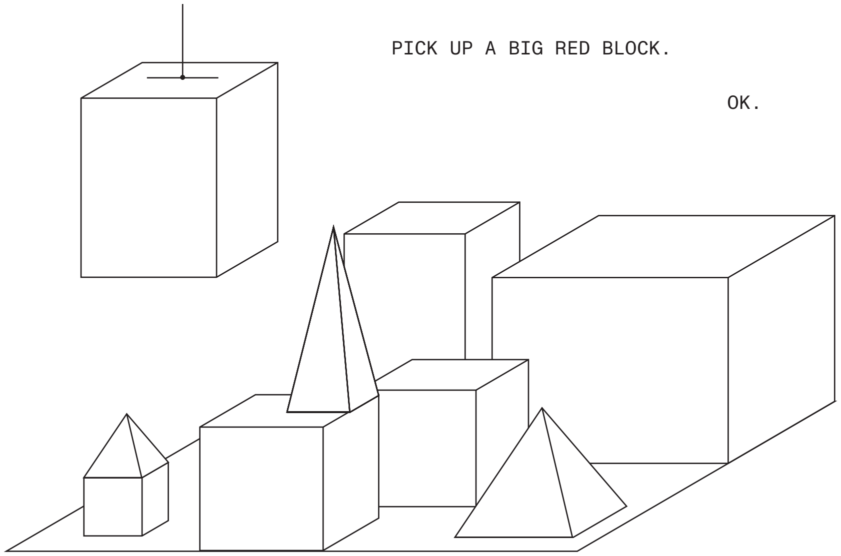
■ Chris Wallace and D. M. Boulton develop Snob, a program capable of classifying data points into "clusters"—effectively dividing up the search space into different groups of points, without any input from humans. This marks an early instance of unsupervised learning.

1969

■ Arthur Bryson and Yu-Chi Ho develop a learning algorithm for optimizing a neural network using the



16. Scene from Stanley Kubrick's *2001: A Space Odyssey* (1968), showing astronauts conversing as HAL, a sentient machine installed on their ship, looms in the background



17. Image of Terry Winograd's 1968 program SHRDLU

18. Computer scientist Charles Rosen with Shakey (1969), the first mobile robot controlled by AI, Stanford Research Institute, California, ca. 1970

chain rule—a notion later dubbed "back propagation of errors," or "backprop" for short. The technique allows for a network receiving a wrong answer to automatically update its weights in the direction of the correct answer, resulting in a gradual move toward the least erroneous answers—a process called "gradient descent." The next year, Seppo Linnainmaa develops a neural network using what he calls "automatic differentiation," a version of backprop.

■ Stanford Research Institute introduces Shakey (named because of its wobbliness), the first mobile robot controlled by AI (fig. 18). The robot receives commands in (a very limited) natural language, breaks down commands into steps, uses cameras to identify objects, and relies on locomotion to avoid or interact with objects in the lab. Shakey receives extensive media coverage and becomes the model robot for decades.

■ Roger Schank presents his conceptual dependency theory for natural-language input. Schank's model stores information according to a set of real-world categories, such as objects, events, places, and times. Each category also permits subcategories that can specify attributes. Since the categories refer to real-world entities and events, the system can determine when two sentences may be different but refer to the same thing—or even are identical in meaning.

■ John McCarthy and Patrick Hayes publish "Some Philosophical Problems from the Standpoint of Artificial Intelligence." Drawing out implications from McCarthy's earlier discussion of common sense, the authors highlight a challenge in determining what counts as "relevant information" for an AI system. For example, if a machine wishes to stack a group of irregularly shaped objects (such as dishes), it must have an intuitive grasp of not just physics but also the objects' different weights, centers of gravity, and relative sturdiness. McCarthy and Hayes suggest that even such "simple" actions require that an enormous amount of information be computed in a relatively short amount of time—a problem known as "combinatorial explosion." This problem of calculating relevant information comes to be known as the "frame problem"—the challenge of determining the proper "frame" of relevant and nonrelevant information—which is regarded as a near-fatal obstacle for AI at the time.

1971

■ Herbert W. Franke publishes *Computergraphik-Computerkunst*, the first book of computer graphic art.

■ In her paper "Natural Categories," Eleanor Rosch sets out a novel understanding of concepts. Rather than treating human concepts as strict logical definitions, such as "a triangle is an enclosed,

19. Harold Cohen using the "Turtle" drawing tool, controlled by his 1973 art-making program, AARON, San Diego, 1977

three-sided figure," she recommends treating concepts as bundles of attributes, such as the concept "mammal" having the attributes "fur," "four limbs," and so on. She calls these concepts "prototypes" and argues that their fuzzy boundaries are a feature, not a bug, that allows language users to deal with ambiguity. This work pushes AI toward fuzzy boundaries and uncertainty.

1972

■ Hubert Dreyfus publishes *What Computers Can't Do: The Limits of Artificial Intelligence*, a history of AI projects and an extended philosophical critique of their claims to simulate human thought. The book makes two main criticisms: First, programs like GPS or SIR only followed syntactic rules and did not "understand" what they were doing semantically. Second, even if AI could solve problems semantically, they would not be behaving like humans because humans are not principally thinkers. Rather, Dreyfus argues, humans are embodied social agents in an ever-changing world who only rarely engage in the problem-solving behavior typical of AI.

1973

■ The Lighthill Report for the British Science Research Council to the UK Parliament highlights the enormous gulf between the claims made by AI advocates and their modest accomplishments. The report, noting problems of combinatorial explosion, suggests AI is far less successful than alleged and recommends massive funding cuts for the field. This roughly marks the beginning of the first "AI winter," a period of funding cuts and lost interest in the field.

■ In version 4.0 of Northwestern University's chess-playing program Chess, the program's prior heuristics (type-B) approach is replaced with a brute force (type-A) approach. Chess and its successors become the dominant chess-playing programs of the decade. This implies two things: First, since prior heuristic programs had proven to be weak chess players, it is unlikely that humans rely on formalizable heuristics alone. However, humans also cannot use brute force in chess playing. This implies, second, that optimal problem-solving with computing technology is unlike human problem-solving.

■ The artist Harold Cohen develops AARON, a program designed to help him create art (fig. 19). Cohen hand codes the program's knowledge of artistic concepts, such as shape and color.

1974

■ Ted Shortliffe develops MYCIN, a program designed to diagnose illnesses based on a large database that correlates information from test results with symptoms. The program also calculates a confidence level for its diagnoses. The results are roughly comparable to those provided by a human doctor. MYCIN is never deployed but does suggest a practical role for expert systems combining large bodies of rules and data for medical use.

1975

■ Meta-DENDRAL, a machine-learning program working with a version of DENDRAL, derives new rules for mass spectrometry. These rules are published in a peer-reviewed journal, marking the first scientific discovery by a machine—although Meta-DENDRAL is not included as an author.

■ Marvin Minsky's article "Frames" makes explicit the increasing trend in AI for approaching intelligence through "microworlds," such as Terry Winograd's SHRDLU, where a specific problem can be solved by applying a "frame" that specifies relevant information, such as the expected objects the program will encounter, the concepts needed, the general rules that apply, and what information is significant.

■ Ross Quinlan introduces the Iterative Dichotomizer (ID3) program, which sorts through data sets and “dichotomizes” (splits in two) the data into categorizing clusters, with each cluster more homogenous than the original set. The program then reiterates the process multiple times, resulting in increasingly specific clusters. This “data mining” creates decision trees, which allow for faster categorization of new information as well as predictions about novel but related phenomenon (fig. 20). Decision trees become essential for the data mining at the heart of twenty-first-century deep learning.

1976

■ During their lecture upon receiving the Turing Award, Allen Newell and Herbert Simon propose “the physical symbol system hypothesis,” the idea that symbol manipulation is the essence of thought. They write, “A physical symbol system has the necessary and sufficient means for general intelligent action.” This becomes the basis for the computational theory of mind, a dominant paradigm for both AI and cognitive science.

1977

■ METEO, a system that translates weather forecasts from English into French for the province of Québec, is installed. The dual-language culture of Canada provides an excellent resource for machine translation, with the English–French transcripts of Parliamentary proceedings forming the input for many contemporary translation data sets.

■ In *Scripts, Plans, Goals, and Understanding: An Inquiry into Human Knowledge Structures*, Roger Schank and Robert Abelson hypothesize that human thinking involves “scripts” for certain interactions. For example, when ordering at a restaurant, people go through effectively scripted conversations to order their food. Schank and Abelson suggest that concepts in AI should be connected according to their use in scripts, rather than in a logical domain. So while the concepts “waiter,” “menu,” and “fork” are not related in the dictionary, they are all included in the script for successfully ordering and eating at a restaurant.

■ Roger Shepard performs experiments on human subjects exploring “mental rotation,” the ability to imagine what a pictured object would look like from another direction. He concludes that humans mentally simulate rotating the object in their mind in order to solve the problem. This sets off a debate in cognitive science over whether “imagistic” representations exist in humans and, if they do, whether they require separate treatment from the discrete representations at the heart of logic and language. It also spurs a similar discussion in AI about the value and treatment of imagistic representations.

1978

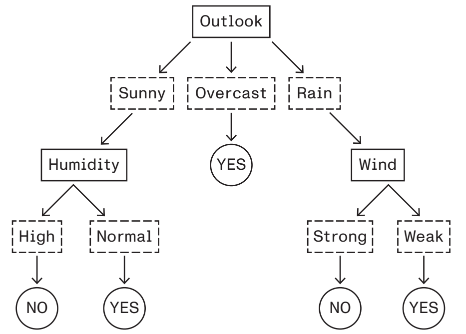
■ Herbert Simon wins the Nobel Prize in Economics for his work on “bounded rationality,” or “satisficing.” The idea is that human decision-makers do not make optimal decisions—which would require perfect knowledge and unlimited time—but satisfactory decisions—based on limited information, cognitive heuristics, and constraints.

1979

■ Kunihiko Fukushima develops the neocognitron, a multilayer neural network based on the visual cortex. The network recognizes handwritten characters by having different layers differentiate between local features, such as edges, and global features, such as shape.

■ For his dissertation, Hans Moravec turns the Stanford Cart into the first fully autonomous vehicle. The cart was originally built in the 1960s by the Stanford University graduate student James L. Adams to be a remote-controlled moon rover. Moravec provides the machine with three-dimensional visual awareness so it can slowly navigate a parking lot with obstacles (fig. 21).

■ Multiple researchers, including John McCarthy and Drew McDermott, begin work on nonmonotonic logics in AI. Whereas standard logic assumes statements are always true or false, nonmonotonic logics focus on tentatively held statements based on incomplete information that are subject to further evaluation as evidence comes in. This requires developing models of truth maintenance—adjusting beliefs in light of changing events and evaluating when to retain or reject a previously held belief.



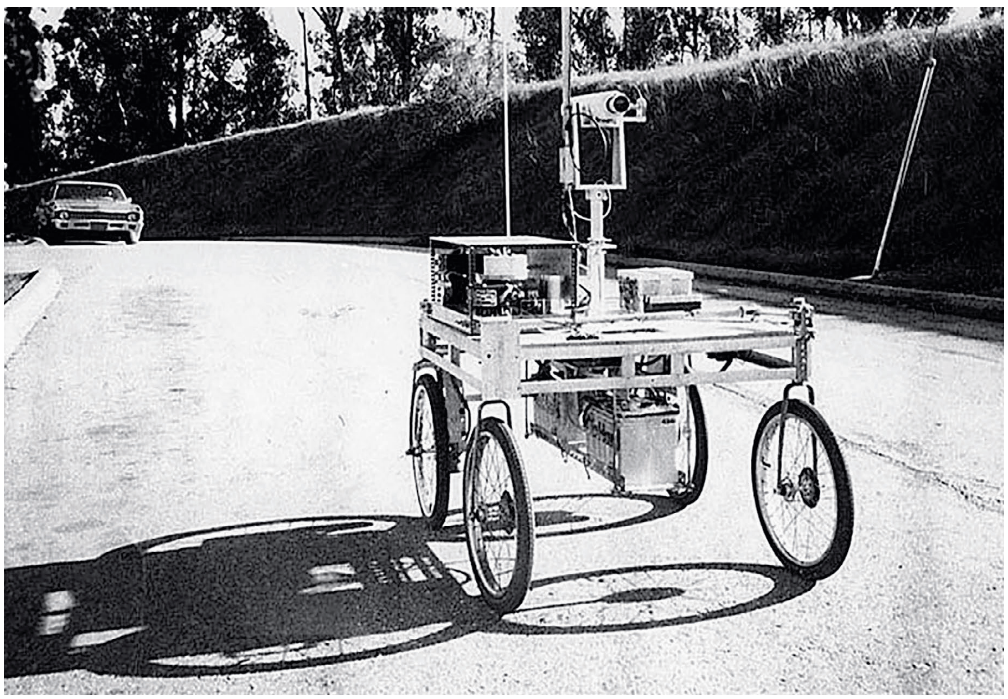
20. Example of a decision tree categorizing weather conditions for a game of tennis

1980

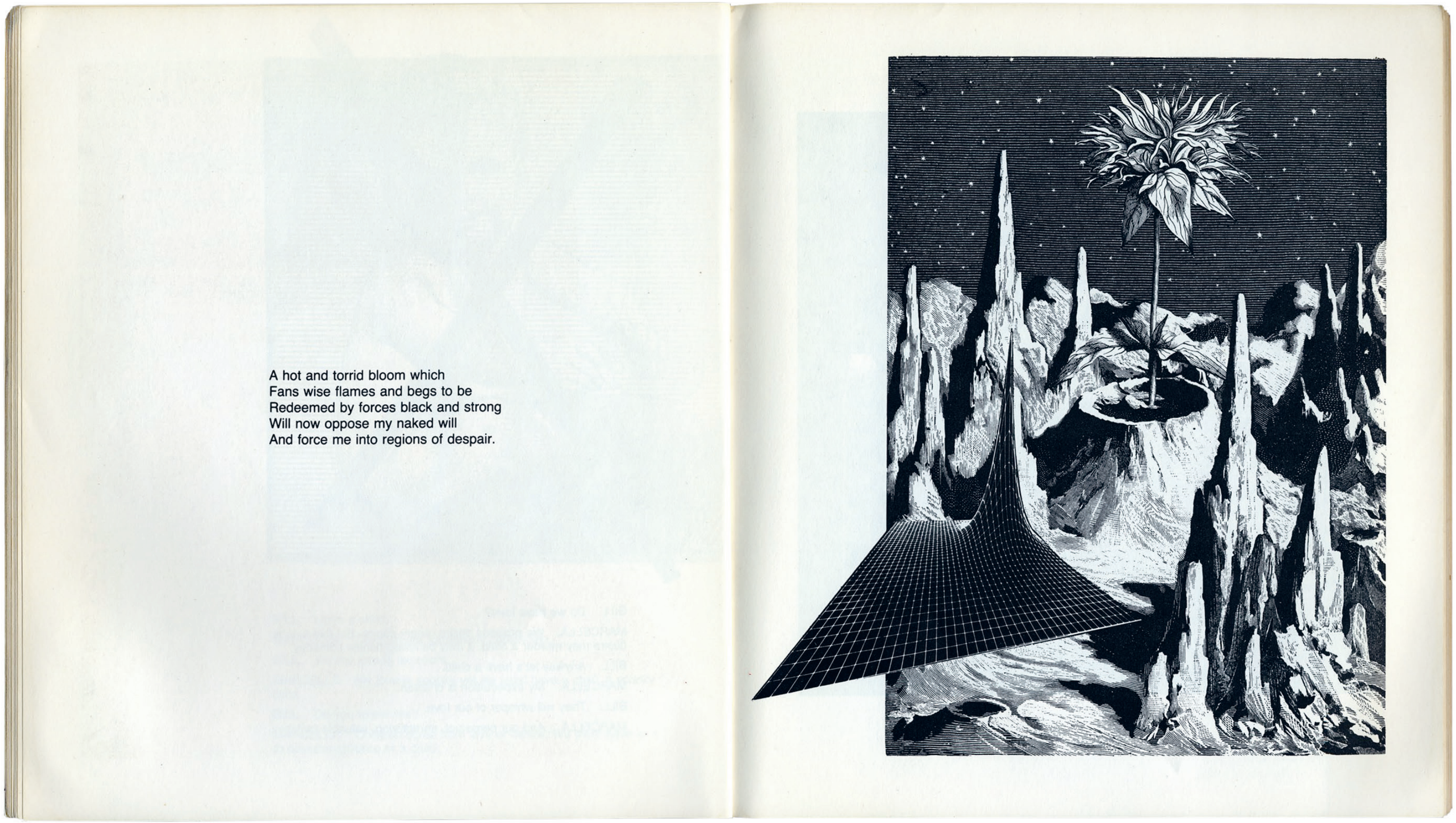
■ The first annual conference by the Association for the Advancement of Artificial Intelligence is held at Stanford University, California.

1981

■ Japan funds the Fifth Generation Computer Project, an ambitious attempt to develop AI for perception, speech recognition, language processing,



21. The Stanford Cart, Stanford University, California, 1970s



22. Spread from *The Policeman's Beard Is Half Constructed*, 1984; illustration by Joan Hall

and cognition. This marks the beginning of a new "AI spring," with similar efforts—and funding—taking place in the US and the UK.

1982

■ John Hopfield introduces the Hopfield network, a system loosely designed to mimic human associative memory. Each neuron in the network acts as both input and output, and all neurons are connected with each other in a single layer. The network is able to store patterns in memory, where the memory is encoded as the specific weights for that pattern, and retrieve them by returning to those specific weights.

1983

■ John Laird and Paul Rosenbloom, under Allen Newell at MIT, finish dissertations on Soar, a cognitive architecture loosely based on Newell's earlier work on the GPS. The goal of Soar is to create artificial agents with general intelligence—unified systems with multiple interacting cognitive capacities that make them capable of solving any cognitive task. This architecture is still in use today.

1984

■ The first AI-written book, *The Policeman's Beard Is Half Constructed*, is generated by the random prose–generating program Racter (fig. 22). On the back cover, Racter teases, "Stories, essays, dissertations, tales are in this book. There are also meat and tomatoes, contracts and agreements. This book is my consciousness, my awareness, my world-view."

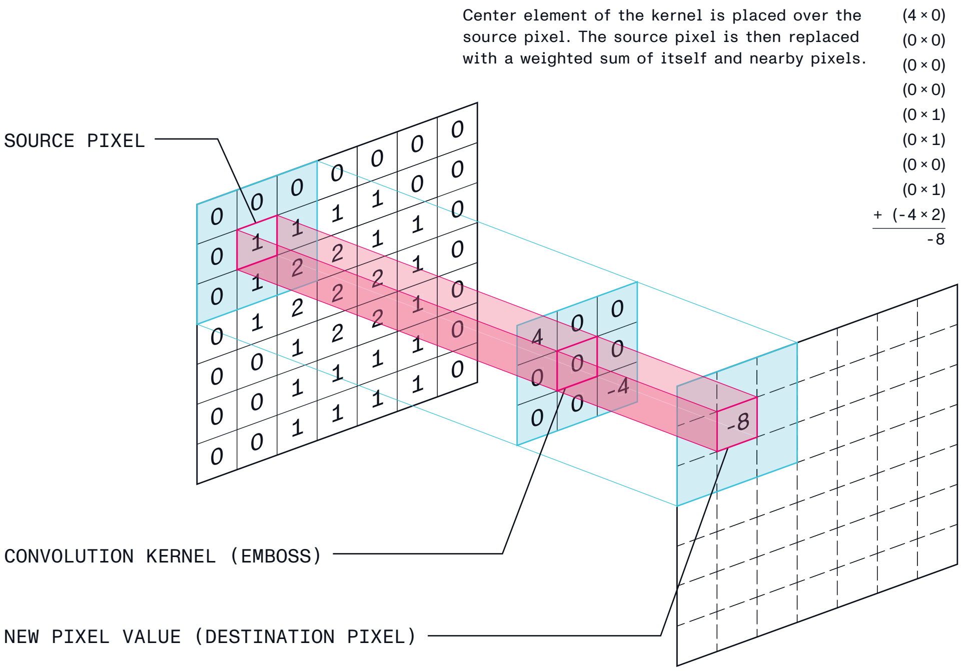
1985

■ David Rumelhart, Geoffrey Hinton, and Ronald J. Williams independently rediscover backpropagation. They each propose multilayer neural networks (now relabeled "parallel distributed processing" or "connectionist" systems) that can engage in machine learning using backpropagation. These results are published in a two-volume set titled *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*, a text that sets off a wave of interest in neural networks among philosophers and neuroscientists, as well as skepticism among defenders of "good, old-fashioned artificial intelligence" (GOFAI).

1987

■ The roboticist Rodney Brooks proposes that the traditional paradigm of AI is misguided because of its narrow focus on abstract problem-solving. He argues that most cognitive abilities are "situated," consisting of embodied perceptual-motor capacities tailored to interact successfully with an environment. Brooks introduces a "subsumption architecture," in which different processes are handled by different single-purpose systems stacked into a hierarchy. In his behavior-based robots, such as ALLEN and HERBERT, this architecture comprises multiple systems broken down by task—for example, movement or vision—allowing the robot to navigate its environment. Brooks holds that, similar to biological beings, these robots move through the world by using not abstract thinking but embodied awareness.

■ The first Conference on Neural Information Processing Systems (NeurIPS) is held. The annual conference becomes the leading forum for work on neural networks, which become progressively more common—and more dominant—throughout the twenty-first century.



23. Diagram of a convolution, a filter applied to a small part of an image, in a convolutional neural network

1988

■ Judea Pearl's *Probabilistic Reasoning in Intelligent Systems* introduces the idea of Bayesian networks. Bayes's theorem provides an effective mathematical method for evaluating the probability of a hypothesis in light of given information as well as how a system can update beliefs as more evidence comes in. Pearl's Bayesian networks use graphical representations to indicate the relative dependencies of probabilities, where certain pieces of evidence make more hypotheses more likely while ruling out other hypotheses altogether. This work on conditional probabilities also makes clear the need for—and challenges in—getting machines to distinguish between causality and correlation.

1989

■ Using backpropagation, Yann LeCun trains a multilayer convolutional neural network to recognize handwritten zip codes. This neural network creates feature detectors akin to those in biological vision systems—such as David H. Hubel and Torsten Wiesel's edge-detecting neurons. This is accomplished by applying multiple localized feature detectors (i.e., convolutions) to the entire image (fig. 23). These convolutions then detect and accentuate features like edges, corners, or curves, passing the results to following layers. Those layers pass the most distinctive features along and weed out the least distinctive, permitting the next layers to focus on conjunctions of features—for example, the combination of edges, corners, and curves making up the letter *R*. The result is one of the first instances of deep learning: machine learning involving numerous "hidden" layers between the inputs and outputs.

■ Christopher Watkins develops Q-learning, a method of "model-free" reinforcement learning by which the agent (i.e., learning system) develops an optimal policy for how it should act at each state. The process involves attributing values to various actions, with the agent always choosing the highest values. The agent also reweights its prior state once it knows the value of its choice, with the result that successful moves are made more likely in future iterations. Through successive trial-and-error attempts, the program—given infinite chances and a finite problem—will eventually always act in a way that achieves the highest values, which forms the optimal policy. Q-learning becomes influential in deep-learning systems, especially in those domains, such as video games, where actions already have values (e.g., scores).

1990

■ TD-Gammon, a neural-net backgammon player, competes with world-class players, highlighting the increasing abilities of multilayer neural networks using reinforcement learning to solve problems.

■ After another round of bold predictions meets with limited success, the second AI winter gradually sets in. Both biologically inspired and cognitivist fields see massive decreases in funding and attention, and the term *artificial intelligence* falls out of fashion. The result is that the subsequent rise of the internet and the invention of search algorithms occur without explicit discussions of intelligent machines.

1993

■ In *Rules of the Mind*, John R. Anderson provides an updated cognitive architecture, ACT-R, a descendant of work by Allen Newell as well as of Anderson's own cognitive science work. This architecture divides cognition into procedural knowledge, which provides the rules for how information should be used, and declarative knowledge, which furnishes facts for acting upon. The system also includes perceptuomotor modules utilizing "rational analysis," which Anderson interprets as the adaptation of the machine to statistical regularities in the environment. This becomes a dominant cognitive architecture; various projects are still implemented using this framework.

■ Rodney Brooks begins the Cog project at MIT (fig. 24). The project's aim is to create an embodied humanoid agent with humanlike sensory-motor capacities and facial expressions as well as some emotional awareness. The project is widely covered in the media, although its successes are limited.

1994

■ A demonstration of autonomous cars driving on the highway, using technology developed by Ernst Dickmanns and the Daimler-Benz Corporation, is presented to the public (fig. 25). The cars are able to operate at speeds of up to 80 mph, navigate heavy traffic, and change lanes without human input. Dickmanns's high-speed, low-information computer vision allows computers to be responsive to real-time events without internal models. This is accomplished by having the computers only process a tiny amount of the information—in this case, the immediate area in front of the car, lane markers, and road signs—while ignoring everything else. Although this is effective for highway driving, Daimler-Benz declines to provide further funding toward the research needed for in-city driving, thus ending the project.

1995

■ Drawing on Vladimir Vapnik's previous work, Corinna Cortes and Vapnik publish work on support-vector machines. These are kinds of algorithms that prove essential for supervised learning, which consists of training based on human-labeled data in order to generate humanlike responses from the machine.

■ Tin Kam Ho develops random decision forests, a method for data classification that avoids "overfit-

24. Rodney Brooks with Cog, 1996. Brooks began his project to create an embodied humanoid agent at MIT in 1993.

25. Prototype of the autonomous car developed by Ernst Dickmanns and the Daimler-Benz Corporation, first demonstrated in 1994

ting," the data-mining problem where a model's attempt to perfectly match current data—including outliers and errors—leads it to make poor predictions about future data. Random forests avoid this by creating decision trees that categorize using only subsamples of the data. By creating many different decision trees, which repeatedly separate data into increasingly specific clusters based on their own subsample, the program ends up with many different attempts to solve the same problem—useful for comparison. This forest of different decision trees minimizes the impact of idiosyncratic categorizations that might lead to overfitting on their own.

1997

■ IBM's chess-playing computer Deep Blue defeats Garry Kasparov, the reigning world chess champion (fig. 26).

■ The search engine AltaVista releases Babel Fish, an online translation service. The translations are primitive, but the program marks a renewed interest in machine translation during what is now the internet era.

■ Sepp Hochreiter and Jürgen Schmidhuber develop long short-term memory (LSTM), a type of neural network with a short-term memory in the form of recurrent loops. Unlike other neural networks, where all connections go in one direction (called "feed-forward networks"), LSTM has connections that loop back on themselves or to prior stages in the network. This repeats pieces of information back into the system, reminding it of some previous result. LSTM avoids retention issues faced by previous recurrent neural networks by storing pieces of information that other pieces of information will depend on—for example, earlier states in a video or the verb tense in a sentence. This allows LSTM to process later data in light of earlier data—even if that data is from much earlier—marking a substantive increase in efficacy for neural networks in processing temporal information.

1998

■ Yann LeCun develops LeNet-5, a seven-layer convolutional neural network that recognizes the handwritten numbers on checks. The system, eventually adopted by several banks, can recognize the digits on 10 percent of the checks written in the United States—a major early commercial success of neural nets. Around the same time, LeCun and his team release the Modified National Institute of Standards and Technology (MNIST) database, a large collection of examples of handwritten characters, which becomes the standard test data for building and evaluating character-recognition software.

1999

■ Sony releases the robotic dog AIBO, a toy capable of basic perception and navigation of its environment (fig. 27).

2000

■ Cynthia Breazeal publishes a dissertation at MIT on Kismet, a computing robot that can recognize and simulate human emotional states (fig. 28). Kismet possesses a face, visual and auditory modalities, and the capacity to track eye gaze, detect movement, and perceive skin color.

2002

■ Rodney Brooks's company iRobot releases Roomba, a robotic vacuum cleaner.

■ Lynn Hershman Leeson publishes the artificially intelligent artwork *Agent Ruby* (fig. 29). The online program, based on Hershman Leeson's film *Teknolust* (2002), offers text conversations with an AI web agent named Ruby, represented by a disembodied face that displays emotional reactions to the conversation.

2004

■ The first DARPA Grand Challenge is held. For the contest, funded by the Defense Advanced Research Projects Agency, autonomous cars compete at navigating rugged terrain on a 132-mile route across the Mojave Desert from California to Nevada. Although no cars complete the route in the first event, four out of twenty-three succeed in the second year. Stanley, the entry from the Stanford Racing Team, takes first place by navigating the course at an average speed of 19.1 mph.

2006

■ Geoffrey Hinton and Ruslan Salakhutdinov publish a seminal article on deep learning. While earlier work on backpropagation and convolutional nets revealed the potential for deep networks, it also pointed to shortcomings: limited computing power and the "vanishing gradient" problem. The former problem was gradually overcome by technological innovation in microprocessors, but better computing power only highlighted the second issue: the deeper the network (i.e., the more hidden layers between inputs and outputs), the more likely that patterns detected in early networks would be lost in later networks (i.e., the "gradients" would "vanish"). Hinton and Salakhutdinov argue that the vanishing gradient problem can be solved by pretraining—early layers can be trained on the data first, and later layers then stacked on top afterward. This innovation makes deep learning far more effective, efficient, and available for wider usage—both theoretical and commercial.

26. World chess champion Garry Kasparov playing against IBM's Deep Blue, New York, 1997

27. AIBO the robotic dog, 1999

2009

■ Fei-Fei Li and her colleagues at Princeton University release ImageNet data set, a massive database of labeled images. The next year, the first ImageNet Large Scale Visual Recognition Challenge (ILSVRC), an annual competition for AI object-recognition programs trained on ImageNet, is held.

2011

■ IBM's Watson, using a mix of deep learning and language processing, defeats two former *Jeopardy!* champions. Watson is later applied to problems in health care, education, and other data-intensive tasks. IBM does not release the underlying code for Watson; as a result, its impact is limited.

■ Apple releases Siri, a voice-activated digital assistant, as part of the iPhone 4S. Siri answers natural-language questions and provides weather forecasts, map directions, and jokes. Although the system is limited upon release, it continues to be improved to this day. Within the next few years, Amazon, Google, and Microsoft all roll out similar digital assistants for their own platforms.

2012

■ Jeff Dean and Andrew Ng at Google Brain develop a deep neural net trained on random, unlabeled images from YouTube. Surprisingly, the program learns to recognize cat faces. The implication is that, with enough data, pretraining is no longer necessary. This means the vanishing gradient problem is largely a problem for smaller data sets, and so the new focus is on gathering up as much data as possible, referred to as "big data."

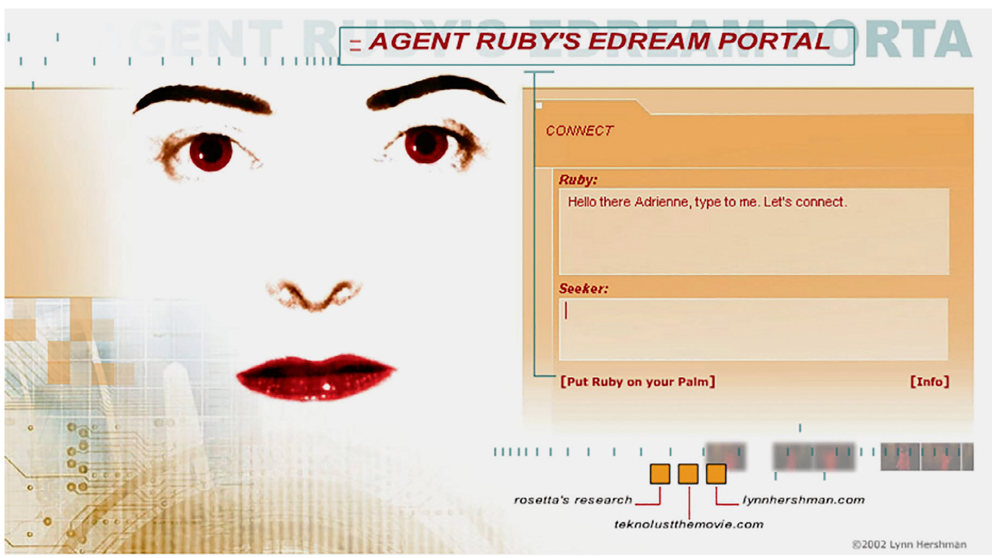
■ Alex Krizhevsky's AlexNet, a deep convolutional neural network (DCNN) run on a graphics processing unit (GPU), achieves an error rate of 16 percent in the ILSVRC. This is a marked improvement over prior years and the other competitors. Following this victory, DCNN and successors dominate ImageNet and become the new model for computer vision, spurring research and funding for GPU-implemented neural networks.

2014

■ Facebook publishes work on DeepFace, their deep-learning facial-recognition software. The program performs at an accuracy rate of 97.35 percent—effectively as accurate as humans and far more advanced than competing networks. The impressive success of the network brings with it questions about the ethics of facial recognition, especially in the hands of governments.

■ Ian Goodfellow demonstrates a generative adversarial network (GAN). GANs combine traditional machine learning, which categorizes examples, and generative networks, which produce new examples. For example, in the case of faces, a

28. Cynthia Breazeal with her computing robot Kismet, a project started in 1998



29. Screenshot of Lynn Hershman Leeson's *Agent Ruby* (2002)

traditional network focuses on learning to recognize faces, whereas a generative network learns to create new faces to be recognized. In GANs, both networks simultaneously improve, resulting in a highly tuned facial-recognition network and a very accurate face generator. These generative networks permit the creation of deepfakes, fraudulent but highly realistic faces.

2015

■ A group of influential intellectuals—including the physicist Stephen Hawking, the inventor Elon Musk, and the AI engineer Stuart Russell—sign "Research Priorities for Robust and Beneficial Artificial Intelligence: An Open Letter." The letter calls for the development of beneficial AI that is capable of being controlled by humans. The authors draw special attention to the dangers of autonomous vehicles—especially armed military drones—and end by expressing broader dystopian concerns about superintelligent machines.

2016

■ Daniel Yamins and James DiCarlo develop a deep convolutional neural network that mimics the hierarchical nature of the human visual system. The work highlights a convergence between the architectures used in natural and artificial visual processing, although there are still important disanalogies between the two systems, such as the lack of backpropagation in the brain.

■ Google Translate switches from traditional machine-learning methods—involving word lookup and grammatical rules for proper transformations—to deep-learning methods. This results in an impressive jump in performance, leading to an outpouring of interest in the technology. Competing platforms are pushed to move to deep learning and neural nets.

■ The computer program AlphaGo beats Go grandmaster Lee Sedol (fig. 30). AlphaGo, developed by Google DeepMind, mixes machine learning based on millions of example games with a Monte Carlo tree search strategy. The Monte Carlo tree search involves testing different strategies at random for any given move, expanding the search space. Based on the findings, the program can then selectively explore strategies in the area most likely to provide successful moves. This results in a program that employs both deep learning and heuristic strategies in order to avoid the impossible task of simulating every possible game. Although AlphaGo reveals the enormous potential of contemporary AI, the network is trained on games played by humans. A successor network, AlphaGo Zero, is able to decisively best AlphaGo by solely training on games it plays with itself, with zero human input. The upshot is that, at least in some cases, machines learn best when they ignore human reasoning altogether.

■ Two milestones are reached in health care intelligence: First, the robot Xiaoyi, developed by the Chinese technology company iFlytek and Tsinghua University, Beijing, uses deep learning to pass China's national licensing medical exam. Xiaoyi's success indicates the ability for machines to process and utilize the natural language needed for analyzing textbooks and clinical information. Shortly afterward, CheXNet, developed by Stanford University, bests humans at detecting pneumonia based on X-rays.

■ In Pittsburgh, Uber deploys self-driving cars, with a human able to take the wheel in emergencies. The program is expanded into San Francisco, Toronto, and Tempe, Arizona.

2018

■ The first fatality from a fully autonomous car occurs in Tempe, Arizona, when a woman in a crosswalk is hit by a self-driving Uber. Uber pulls all of its self-driving cars off the road in response.

■ Google unveils Duplex, a virtual assistant that can call businesses and schedule appointments on behalf of users. The speech abilities of Duplex, while limited to specific tasks, are so sufficiently human-like that receptionists do not always recognize that a machine is speaking.

■ DeepMind's AlphaStar defeats professional players of the video game *StarCraft II*. The game presents several challenges: the pieces are all identical, and competitors are often hidden from each other's sight. The corporation OpenAI also enters five neural networks, known as the "OpenAI Five," in a professional *Dota 2* competition. In *Dota 2*, multiple agents control avatars with different skill sets and work together as a team. These successes demonstrate the ability of AI to compete in games that demand cooperation and real-time strategies in the face of limited information.

■ The first piece of AI-generated art, *Portrait of Edmond de Belamy*, created by the Paris-based artist collective Obvious, sells at auction for $432,500.

2019

■ OpenAI unveils the program GPT-2, which generates predictive text for an input sentence based on the kinds of sentences that statistically follow.

30. Go grandmaster Lee Sedol reviews the match after losing to Google DeepMind's program AlphaGo, Seoul, 2016

AI Glossary

Compiled by Jake Browning and Philipp Schmitt

Agent	A machine that can sense and respond to its environment and that can use autonomous reasoning in order to achieve some goal. Examples include virtual assistants, such as Siri and Alexa. Agents can perform alone, with other agents or humans, or even *against* others, depending on the task.
AI Effect	The phenomenon where solving a task is considered a proper test of "artificial intelligence" before it is accomplished, but the solution is treated as "just a computer program" immediately afterward. This perspectival shift often results in the undervaluing of successes in AI.
Algorithm	A set of instructions for solving a problem. Although historically algorithms are associated with the steps involved in solving arithmetic problems, the concept is more general and includes any procedure—such as a recipe or map directions—that results in a solution. In machine learning, the "solution" to a learning problem is an algorithm, and the machine is responsible for discovering the algorithm through trial and error.
Artificial General Intelligence (AGI)	The development of an artificial agent that can perform intelligent tasks in a multitude of different domains. The ideal is a machine with cognitive capacities similar to—or even exceeding—those of humans.
Artificial Intelligence (AI)	Any nonbiological process that, if done by a biological being, would count as "thinking." The most common version of this is machine intelligence, in which a machine is used to solve a cognitive problem, such as playing a game, diagnosing an illness, or analyzing stock trends.
Artificial Neural Net (ANN) (or Neural Network)	A machine designed to mimic the learning that takes place in the brain. The machine uses parallel processing and interconnectedness of small processors, called nodes, that function similarly to networks of biological neurons. The nodes learn through trial and error by having their weights adjusted in response to wrong answers.
Autonomy	In the case of machines, this refers to the ability to act without input from a human.
Backpropagation (or Backward Propagation of Errors)	A method for training multilayer (i.e., "deep") neural networks that involves correcting each layer in the network in reverse sequence, working backwards from the output to the original input.
Bayesian Networks (or Belief Networks, or Causal Probabilistic Networks)	Networks that allow for reasoning about uncertain situations and are capable of evaluating probabilities and updating beliefs in light of new information.
Boltzmann Machine	A type of Hopfield network that randomizes whether nodes are on or off.
Brute Force (or Type-A Problem-Solving)	A type of problem-solving that involves searching through as many solutions in the solution space as is feasible and selecting the optimal solution from those available.
Chatbots	Interactive programs designed to engage in human conversation.
Classification	Providing a label for some input or piece of data. In the case of supervised learning, this involves a machine learning an algorithm that assigns human-specified labels to each input.
Cluster Analysis (or Clustering)	A type of unsupervised learning that groups similar data points into their own category.
Combinatorial Explosion	The phenomenon where small changes increase the complexity of a problem to the point at which it is no longer solvable within any realistic time frame.

Common-Sense Reasoning	The vaguely defined set of capacities and background knowledge evidenced by humans in their everyday interaction with the world. Being able to reproduce common-sense reasoning is often treated as the "Holy Grail" of AI, since it implies a "general" intelligence capable of reasoning, planning, conversing, and interacting with the environment much as humans do.
Computer Vision (or Machine Vision)	The field of designing machines that analyze visual information. The boundaries of this field can be both narrow, as when referring to image classification, or broad, as when it is used in autonomous vehicles.
Cybernetics	The field of research covering the adaptive use of information by a mechanical or organic being.
Data Mining	The process of uncovering relationships between different pieces of information. In the case of machine learning, this often involves teasing out unexpected connections by analyzing a massive amount of data, or "big data."
Data Point	A single input or piece of information, generally understood as a small part of a much larger data set.
Decision Tree	Similar to flowcharts, design trees represent how to classify a data point by analyzing its relationship to other data points.
Deep Learning	Machine learning performed by neural networks with many layers of nodes.
Expert System (or Knowledge-Based System)	A machine designed to solve a problem based on knowledge collected from the experts in some field. This is especially common in medical diagnostic software and troubleshooting programs.
First- and Second-Order Logic	Systems of logic used to deal with quantification (i.e., all, some, or none). First-order logic concerns only the relationship between entities. Second-order logic can also handle relationships between sets.
Frame Problem	The difficulty of specifying all of the relevant and nonrelevant axioms an agent needs to know in order to navigate its environment. This problem is often used to explain the need for common-sense reasoning.
Genetic Algorithm	Modeled on biological evolution, genetic algorithms create numerous random solutions to a problem, cull poorly performing solutions from each iteration, and introduce variations in (or mixing between) the best solutions from each group. Through multiple iterations, the remaining algorithms will often provide exceptional but unexpected solutions.
Gradient Descent	A standard technique for guiding trial and error in machine learning where the network seeks out through a series of small steps (i.e., "descends to") the least erroneous algorithm (also referred to as "the minimum").
Hebbian Learning	A theory proposed by the psychologist Donald Hebb that explains how neurons learn. The idea, often summed up as "neurons that fire together, wire together," suggests that the repeated coincidental firing of two or more neurons will lead to them being associated in the brain, thus firing together even more in the future.
Heuristic Search Techniques (or Type-B Problem-Solving)	A search technique involving hand-coded rules of thumb for solving a problem—often deployed in cases that cannot be solved by brute force alone, such as in games of chess. Although these techniques often lead to acceptable solutions, they are not always the best solutions—which is why many chess-playing computers still lose matches to humans.
Hopfield Network	A neural network where all nodes are binary (i.e., on or off) and act as input and output to all other nodes. This design mimics associative memory in the human brain, where nodes that fire together feed into one another, deepening their interconnection.
Hybrid Systems	Intelligent agents that use both lower-level machine learning and higher-level symbolic AI to achieve rapid learning while also maintaining humanlike logical reasoning.
Instrumental Reasoning (or Back Chaining)	The "means–ends" reasoning process where the final goal (the "ends") is established first, prior to the intermediary steps (the "means").
Learning	Progressive improvement over time on some task.
Machine Learning	The practice of letting machines develop their own algorithms through trial and error in lieu of humans hand coding the algorithms. This term is also used to refer to approaches favoring artificial neural networks.
Markov Chain	A random sequence where each decision in the chain is based solely on the last decision made.
Microworld (or Toy World)	A small domain for an artificial agent to operate in, such as a virtual or video game world.
Minimum (Global and Local)	The global minimum describes the algorithm that has the least error between what it predicts and the actual data points, graphically represented as the lowest point in the lowest valley in a solution space. Local minimums refer to nonoptimal algorithms that a machine might mistakenly treat as the best, represented as the lowest point in a shallow valley.

Natural-Language Processing	Programs designed to interpret everyday language, as opposed to technical or formal languages that are used in artificial situations.
Pattern Recognition	The ability of a machine to discern the pertinent features of data, especially for the purposes of classification.
Planning	The general field for machines that generate strategies for acting, deploying resources, and adjusting to solve problems.
Predictive Algorithm	An algorithm that predicts an expected outcome based on some input—for example, forecasting that a potential debtor will default based on their credit history.
Pruning	A practical technique whereby a machine ignores (i.e., "prunes") strategies on a decision tree that will not lead to an acceptable solution, also referred to as "fruitless strategies."
Recursion	The use of a function within use of the same function. For example, in the mathematical operation *x(y*z)*, the same function is applied both within the parentheses and also to the parentheses as a whole.
Reinforcement Learning	Akin to reinforcement learning in biological beings, this involves training a machine to achieve a general goal, such as winning a game, by providing general evaluative metrics for proper and improper responses, such as winning and losing.
Semantic Net	A graph depicting the relationships between concepts in some domain.
The Singularity	The moment when machine intelligence might surpass human intelligence.
Solution Space	All of the possible algorithms, both effective and ineffective, for solving a particular problem. Because most machine-learning problems involve more than three dimensions, it is difficult for humans to imagine this space; it is therefore often represented as a geometric plane crisscrossed with peaks and valleys.
Strong AI	The development of a sentient and sapient artificial intelligence. It is contentious whether this will ever be possible for computers.
Supervised Learning	Machine learning accomplished through trial and error whereby the machine attempts to conform to correct answers provided by a "teacher." The teacher often consists of a well-labeled data set, such as labeled images of faces, and the network trains itself to correctly predict the proper label for each face.
Symbolic AI (or Classic AI, or Good, Old-Fashioned AI [GOFAI])	Term used to refer to artificial machines that mimic high-level human reasoning, such as that deployed in mathematical proofs, deductive inference, conceptual analysis, and analogical reasoning. Although eclipsed by machine learning in recent years, symbolic AI was the dominant paradigm until the 1980s, and it still has many advocates.
Turing Test	A test for intelligence suggested by the mathematician and computer scientist Alan Turing that uses conversation to determine whether someone is talking to a sentient machine or a human being.
Unsupervised Learning	Machines that learn using unlabeled data, usually through cluster analysis (i.e., treating nearby data points as falling into the same group).
Value Alignment	The task of aligning the values and overall goals of machines with their human users.
Virtual Machine (or Computer Program)	The technical name for a computer program, or a simulated computing machine capable of running on different architectures.
Weak AI	The development of a nonsentient but functionally useful machine that can solve prespecified cognitive problems. All contemporary AI is "weak" in this sense.
Weight	The strength of connections between nodes in a neural network. Weights are modified through learning as the machine starts to approximate correct responses.