Artificial intelligenceFrom Wikipedia, the free encyclopedia

In computer science, artificial intelligence (AI), sometimes called machine intelligence, is intelligence demonstrated by machines, in contrast to the natural intelligence displayed by humans. Colloquially, the term "artificial intelligence" is often used to describe machines (or computers) that mimic "cognitive" functions that humans associate with the human mind, such as "learning" and "problem solving".[1]

As machines become increasingly capable, tasks considered to require "intelligence" are often removed from the definition of AI, a phenomenon known as the AI effect.[2] A quip in Tesler's Theorem says "AI is whatever hasn't been done yet."[3] For instance, optical character recognition is frequently excluded from things considered to be AI, having become a routine technology.[4] Modern machine capabilities generally classified as AI include successfully understanding human speech,[5] competing at the highest level in strategic game systems (such as chess and Go),[6] autonomously operating cars, intelligent routing in content delivery networks, and military simulations.

Artificial intelligence can be classified into three different types of systems: analytical, human-inspired, and humanized artificial intelligence.[7] Analytical AI has only characteristics consistent with cognitive intelligence; generating a cognitive representation of the world and using learning based on past experience to inform future decisions. Human-inspired AI has elements from cognitive and emotional intelligence; understanding human emotions, in addition to cognitive elements, and considering them in their decision making. Humanized AI shows characteristics of all types of competencies (i.e., cognitive, emotional, and social intelligence), is able to be self-conscious and is self-aware in interactions.

Artificial intelligence was founded as an academic discipline in 1956, and in the years since has experienced several waves of optimism,[8][9] followed by disappointment and the loss of funding (known as an "AI winter"),[10][11] followed by new approaches, success and renewed funding.[9][12] For most of its history, AI research has been divided into subfields that often fail to communicate with each other.[13] These sub-fields are based on technical considerations, such as particular goals (e.g. "robotics" or "machine learning"),[14] the use of particular tools ("logic" or artificial neural networks), or deep philosophical differences.[15][16][17] Subfields have also been based on social

factors (particular institutions or the work of particular researchers).[13]

The traditional problems (or goals) of AI research include reasoning, knowledge representation, planning, learning, natural language processing, perception and the ability to move and manipulate objects.[14] General intelligence is among the field's long-term goals.[18] Approaches include statistical methods, computational intelligence, and traditional symbolic AI. Many tools are used in AI, including versions of search and mathematical optimization, artificial neural networks, and methods based on statistics, probability and economics. The AI field draws upon computer science, information engineering, mathematics, psychology, linguistics, philosophy, and many other fields.

The field was founded on the claim that human intelligence "can be so precisely described that a machine can be made to simulate it".[19] This raises philosophical arguments about the nature of the mind and the ethics of creating artificial beings endowed with human-like intelligence which are issues that have been explored by myth, fiction and philosophy since antiquity.[20] Some people also consider AI to be a danger to humanity if it progresses unabated.[21] Others believe that AI, unlike previous technological revolutions, will create a risk of mass unemployment.[22]

In the twenty-first century, AI techniques have experienced a resurgence following concurrent advances in computer power, large amounts of data, and theoretical understanding; and AI techniques have become an essential part of the technology industry, helping to solve many challenging problems in computer science, software engineering and operations research.[23][12]

History.

Main articles: History of artificial intelligence and Timeline of artificial intelligence

Talos, an ancient mythical automaton with artificial intelligence

Thought-capable artificial beings appeared as storytelling devices in antiquity,[24] and have been common in fiction, as in Mary Shelley's Frankenstein or Karel ?apek's R.U.R. (Rossum's Universal Robots).[25] These characters and their fates raised many of the same issues now discussed in the ethics of artificial intelligence.[20]

The study of mechanical or "formal" reasoning began with philosophers and mathematicians in antiquity. The study of mathematical logic led directly to Alan Turing's theory of computation, which suggested that a machine, by shuffling symbols as simple as "0" and "1", could simulate any conceivable act of mathematical deduction. This insight, that digital computers can simulate any process of formal reasoning, is known as the Church?Turing thesis.[26] Along with concurrent discoveries in neurobiology, information theory and cybernetics, this led researchers to consider the possibility of building an electronic brain. Turing proposed that "if a human could not distinguish between responses from a machine and a human, the machine could be considered "intelligent".[27] The first work that is now generally recognized as AI was McCulloch and Pitts' 1943 formal design for Turing-complete "artificial neurons".[28]

The field of AI research was born at a workshop at Dartmouth College in 1956.[29] Attendees Allen Newell (CMU), Herbert Simon (CMU), John McCarthy (MIT), Marvin Minsky (MIT) and Arthur Samuel (IBM) became the founders and leaders of AI research.[30] They and their students produced programs that the press described as "astonishing":[31] computers were learning checkers strategies (c. 1954)[32] (and by 1959 were reportedly playing better than the average human),[33] solving word problems in algebra, proving logical theorems (Logic Theorist, first run c. 1956) and speaking English.[34] By the middle of the 1960s, research in the U.S. was heavily funded by the Department of Defense[35] and laboratories had been established around the world.[36] AI's founders were optimistic about the future: Herbert Simon predicted, "machines will be capable, within twenty years, of doing any work a man can do". Marvin Minsky agreed, writing, "within a generation ... the problem of creating 'artificial intelligence' will substantially be solved".[8]

They failed to recognize the difficulty of some of the remaining tasks. Progress slowed and in 1974, in response to the criticism of Sir James Lighthill[37] and ongoing pressure from the US Congress to fund more productive projects, both the U.S. and British governments cut off exploratory research in AI. The next few years would later be called an "AI winter",[10] a period when obtaining funding for AI projects was difficult.

In the early 1980s, AI research was revived by the commercial success of expert systems,[38] a form of AI program that simulated the knowledge and analytical skills of human experts. By 1985, the market for AI had reached over a billion dollars. At the same time, Japan's fifth generation computer project inspired the U.S and British governments to restore funding for academic research.[9] However, beginning with the collapse of the Lisp Machine market in 1987, AI once again fell into

disrepute, and a second, longer-lasting hiatus began.[11]

In the late 1990s and early 21st century, AI began to be used for logistics, data mining, medical diagnosis and other areas.[23] The success was due to increasing computational power (see Moore's law), greater emphasis on solving specific problems, new ties between AI and other fields (such as statistics, economics and mathematics), and a commitment by researchers to mathematical methods and scientific standards.[39] Deep Blue became the first computer chess-playing system to beat a reigning world chess champion, Garry Kasparov, on 11 May 1997.[40]

In 2011, 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.[41] Faster computers, algorithmic improvements, and access to large amounts of data enabled advances in machine learning and perception; data-hungry deep learning methods started to dominate accuracy benchmarks around 2012.[42] The Kinect, which provides a 3D body?motion interface for the Xbox 360 and the Xbox One, uses algorithms that emerged from lengthy AI research[43] as do intelligent personal assistants in smartphones.[44] In March 2016, AlphaGo won 4 out of 5 games of Go in a match with Go champion Lee Sedol, becoming the first computer Go-playing system to beat a professional Go player without handicaps.[6][45] In the 2017 Future of Go Summit, AlphaGo won a three-game match with Ke Jie,[46] who at the time continuously held the world No. 1 ranking for two years.[47][48] This marked the completion of a significant milestone in the development of Artificial Intelligence as Go is a relatively complex game, more so than Chess.

According to Bloomberg's Jack Clark, 2015 was a landmark year for artificial intelligence, with the number of software projects that use AI Google increased from a "sporadic usage" in 2012 to more than 2,700 projects. Clark also presents factual data indicating the improvements of AI since 2012 supported by lower error rates in image processing tasks.[49] He attributes this to an increase in affordable neural networks, due to a rise in cloud computing infrastructure and to an increase in research tools and datasets.[12] Other cited examples include Microsoft's development of a Skype system that can automatically translate from one language to another and Facebook's system that can describe images to blind people.[49] In a 2017 survey, one in five companies reported they had "incorporated AI in some offerings or processes".[50][51] Around 2016, China greatly accelerated its government funding; given its large supply of data and its rapidly increasing research output, some observers believe it may be on track to becoming an "AI superpower".[52][53] However, it has been acknowledged that reports regarding artificial intelligence have tended to be

exaggerated.[54][55][56]

Definitions

Computer science defines AI research as the study of "intelligent agents": any device that perceives its environment and takes actions that maximize its chance of successfully achieving its goals.[57] A more elaborate definition characterizes AI as "a system's ability to correctly interpret external data, to learn from such data, and to use those learnings to achieve specific goals and tasks through flexible adaptation."[58]

Basics

A typical AI analyzes its environment and takes actions that maximize its chance of success.[57] An AI's intended utility function (or goal) can be simple ("1 if the AI wins a game of Go, 0 otherwise") or complex ("Do mathematically similar actions to the ones succeeded in the past"). Goals can be explicitly defined, or induced. If the AI is programmed for "reinforcement learning", goals can be implicitly induced by rewarding some types of behavior or punishing others.[a] Alternatively, an evolutionary system can induce goals by using a "fitness function" to mutate and preferentially replicate high-scoring AI systems, similarly to how animals evolved to innately desire certain goals such as finding food.[59] Some AI systems, such as nearest-neighbor, instead of reason by analogy, these systems are not generally given goals, except to the degree that goals are implicit in their training data.[60] Such systems can still be benchmarked if the non-goal system is framed as a system whose "goal" is to successfully accomplish its narrow classification task.[61]

AI often revolves around the use of algorithms. An algorithm is a set of unambiguous instructions that a mechanical computer can execute.[b] A complex algorithm is often built on top of other, simpler, algorithms. A simple example of an algorithm is the following (optimal for first player) recipe for play at tic-tac-toe:[62]

If someone has a "threat" (that is, two in a row), take the remaining square. Otherwise,

if a move "forks" to create two threats at once, play that move. Otherwise,

take the center square if it is free. Otherwise,

if your opponent has played in a corner, take the opposite corner. Otherwise,

take an empty corner if one exists. Otherwise,

take any empty square.

Many AI algorithms are capable of learning from data; they can enhance themselves by learning new heuristics (strategies, or "rules of thumb", that have worked well in the past), or can themselves write other algorithms. Some of the "learners" described below, including Bayesian networks, decision trees, and nearest-neighbor, could theoretically, (given infinite data, time, and memory) learn to approximate any function, including which combination of mathematical functions would best describe the world[citation needed]. These learners could therefore, derive all possible knowledge, by considering every possible hypothesis and matching them against the data. In practice, it is almost never possible to consider every possibility, because of the phenomenon of "combinatorial explosion", where the amount of time needed to solve a problem grows exponentially. Much of AI research involves figuring out how to identify and avoid considering broad range of possibilities that are unlikely to be beneficial.[63][64] For example, when viewing a map and looking for the shortest driving route from Denver to New York in the East, one can in most cases skip looking at any path through San Francisco or other areas far to the West; thus, an AI wielding a pathfinding algorithm like A* can avoid the combinatorial explosion that would ensue if every possible route had to be ponderously considered in turn.[65]

The earliest (and easiest to understand) approach to AI was symbolism (such as formal logic): "If an otherwise healthy adult has a fever, then they may have influenza". A second, more general, approach is Bayesian inference: "If the current patient has a fever, adjust the probability they have influenza in such-and-such way". The third major approach, extremely popular in routine business AI applications, are analogizers such as SVM and nearest-neighbor: "After examining the records of known past patients whose temperature, symptoms, age, and other factors mostly match the current patient, X% of those patients turned out to have influenza". A fourth approach is harder to intuitively understand, but is inspired by how the brain's machinery works: the artificial neural network approach uses artificial "neurons" that can learn by comparing itself to the desired output and altering the strengths of the connections between its internal neurons to "reinforce" connections that seemed to be useful. These four main approaches can overlap with each other and with evolutionary systems; for example, neural nets can learn to make inferences, to generalize, and to make analogies. Some systems implicitly or explicitly use multiple of these approaches, alongside many other AI and non-AI algorithms;[66] the best approach is often different depending on the problem.[67][68]

Learning algorithms work on the basis that strategies, algorithms, and inferences that worked well in the past are likely to continue working well in the future. These inferences can be obvious, such as "since the sun rose every morning for the last 10,000 days, it will probably rise tomorrow morning as well". They can be nuanced, such as "X% of families have geographically separate species with color variants, so there is an Y% chance that undiscovered black swans exist". Learners also work on the basis of "Occam's razor": The simplest theory that explains the data is the likeliest. Therefore, according to Occam's razor principle, a learner must be designed such that it prefers simpler theories to complex theories, except in cases where the complex theory is proven substantially better.

The blue line could be an example of overfitting a linear function due to random noise.

Settling on a bad, overly complex theory gerrymandered to fit all the past training data is known as overfitting. Many systems attempt to reduce overfitting by rewarding a theory in accordance with how well it fits the data, but penalizing the theory in accordance with how complex the theory is.[69] Besides classic overfitting, learners can also disappoint by "learning the wrong lesson". A toy example is that an image classifier trained only on pictures of brown horses and black cats might conclude that all brown patches are likely to be horses.[70] A real-world example is that, unlike humans, current image classifiers don't determine the spatial relationship between components of the picture; instead, they learn abstract patterns of pixels that humans are oblivious to, but that linearly correlate with images of certain types of real objects. Faintly superimposing such a pattern on a legitimate image results in an "adversarial" image that the system misclassifies.[c][71][72][73]

A self-driving car system may use a neural network to determine which parts of the picture seem to match previous training images of pedestrians, and then model those areas as slow-moving but somewhat unpredictable rectangular prisms that must be avoided.[74][75]

Compared with humans, existing AI lacks several features of human "commonsense reasoning"; most notably, humans have powerful mechanisms for reasoning about "naive physics" such as space, time, and physical interactions. This enables even young children to easily make inferences like "If I roll this pen off a table, it will fall on the floor". Humans also have a powerful mechanism of "folk psychology" that helps them to interpret natural-language sentences such as "The city councilmen refused the demonstrators a permit because they advocated violence". (A generic AI has difficulty discerning whether the ones alleged to be advocating violence are the councilmen or the

demonstrators.)[76][77][78] This lack of "common knowledge" means that AI often makes different mistakes than humans make, in ways that can seem incomprehensible. For example, existing self-driving cars cannot reason about the location nor the intentions of pedestrians in the exact way that humans do, and instead must use non-human modes of reasoning to avoid accidents.[79][80][81]

Problems

The overall research goal of artificial intelligence is to create technology that allows computers and machines to function in an intelligent manner. The general problem of simulating (or creating) intelligence has been broken down into sub-problems. These consist of particular traits or capabilities that researchers expect an intelligent system to display. The traits described below have received the most attention.[14]

Reasoning, problem solving

Early researchers developed algorithms that imitated step-by-step reasoning that humans use when they solve puzzles or make logical deductions.[82] By the late 1980s and 1990s, AI research had developed methods for dealing with uncertain or incomplete information, employing concepts from probability and economics.[83]

These algorithms proved to be insufficient for solving large reasoning problems, because they experienced a "combinatorial explosion": they became exponentially slower as the problems grew larger.[63] In fact, even humans rarely use the step-by-step deduction that early AI research was able to model. They solve most of their problems using fast, intuitive judgements.[84]

Knowledge representation

An ontology represents knowledge as a set of concepts within a domain and the relationships between those concepts.

Main articles: Knowledge representation and Commonsense knowledge

Knowledge representation[85] and knowledge engineering[86] are central to classical AI research.

Some "expert systems" attempt to gather together explicit knowledge possessed by experts in some narrow domain. In addition, some projects attempt to gather the "commonsense knowledge" known to the average person into a database containing extensive knowledge about the world. Among the things a comprehensive commonsense knowledge base would contain are: objects, properties, categories and relations between objects;[87] situations, events, states and time;[88] causes and effects;[89] knowledge about knowledge (what we know about what other people know);[90] and many other, less well researched domains. A representation of "what exists" is an ontology: the set of objects, relations, concepts, and properties formally described so that software agents can interpret them. The semantics of these are captured as description logic concepts, roles, and individuals, and typically implemented as classes, properties, and individuals in the Web Ontology Language.[91] The most general ontologies are called upper ontologies, which attempt to provide a foundation for all other knowledge[92] by acting as mediators between domain ontologies that cover specific knowledge about a particular knowledge domain (field of interest or area of concern). Such formal knowledge representations can be used in content-based indexing and retrieval,[93] scene interpretation,[94] clinical decision support,[95] knowledge discovery (mining "interesting" and actionable inferences from large databases),[96] and other areas.[97]

Among the most difficult problems in knowledge representation are:

Default reasoning and the qualification problem

Many of the things people know take the form of "working assumptions". For example, if a bird comes up in conversation, people typically picture an animal that is fist-sized, sings, and flies. None of these things are true about all birds. John McCarthy identified this problem in 1969[98] as the qualification problem: for any commonsense rule that AI researchers care to represent, there tend to be a huge number of exceptions. Almost nothing is simply true or false in the way that abstract logic requires. AI research has explored a number of solutions to this problem.[99]

The breadth of commonsense knowledge

The number of atomic facts that the average person knows is very large. Research projects that attempt to build a complete knowledge base of commonsense knowledge (e.g., Cyc) require enormous amounts of laborious ontological engineering?they must be built, by hand, one complicated concept at a time.[100]

The subsymbolic form of some commonsense knowledge

Much of what people know is not represented as "facts" or "statements" that they could express verbally. For example, a chess master will avoid a particular chess position because it "feels too exposed"[101] or an art critic can take one look at a statue and realize that it is a fake.[102] These are non-conscious and sub-symbolic intuitions or tendencies in the human brain.[103] Knowledge like this informs, supports and provides a context for symbolic, conscious knowledge. As with the related problem of sub-symbolic reasoning, it is hoped that situated AI, computational intelligence, or statistical AI will provide ways to represent this kind of knowledge.[103]

Planning

A hierarchical control system is a form of control system in which a set of devices and governing software is arranged in a hierarchy.

Main article: Automated planning and scheduling

Intelligent agents must be able to set goals and achieve them.[104] They need a way to visualize the future?a representation of the state of the world and be able to make predictions about how their actions will change it?and be able to make choices that maximize the utility (or "value") of available choices.[105]

In classical planning problems, the agent can assume that it is the only system acting in the world, allowing the agent to be certain of the consequences of its actions.[106] However, if the agent is not the only actor, then it requires that the agent can reason under uncertainty. This calls for an agent that can not only assess its environment and make predictions, but also evaluate its predictions and adapt based on its assessment.[107]

Multi-agent planning uses the cooperation and competition of many agents to achieve a given goal. Emergent behavior such as this is used by evolutionary algorithms and swarm intelligence.[108]

Learning

Main article: Machine learning

Machine learning (ML), a fundamental concept of AI research since the field's inception,[109] is the study of computer algorithms that improve automatically through experience.[110][111]

Unsupervised learning is the ability to find patterns in a stream of input, without requiring a human to label the inputs first. Supervised learning includes both classification and numerical regression, which requires a human to label the input data first. Classification is used to determine what category something belongs in, and occurs after a program sees a number of examples of things from several categories. Regression is the attempt to produce a function that describes the relationship between inputs and outputs and predicts how the outputs should change as the inputs change.[111] Both classifiers and regression learners can be viewed as "function approximators" trying to learn an unknown (possibly implicit) function; for example, a spam classifier can be viewed as learning a function that maps from the text of an email to one of two categories, "spam" or "not spam". Computational learning theory can assess learners by computational complexity, by sample complexity (how much data is required), or by other notions of optimization.[112] In reinforcement learning[113] the agent is rewarded for good responses and punished for bad ones. The agent uses this sequence of rewards and punishments to form a strategy for operating in its problem space.

Natural language processing

A parse tree represents the syntactic structure of a sentence according to some formal grammar.

Main article: Natural language processing

Natural language processing[114] (NLP) gives machines the ability to read and understand human language. A sufficiently powerful natural language processing system would enable natural-language user interfaces and the acquisition of knowledge directly from human-written sources, such as newswire texts. Some straightforward applications of natural language processing include information retrieval, text mining, question answering[115] and machine translation.[116] Many current approaches use word co-occurrence frequencies to construct syntactic representations of text. "Keyword spotting" strategies for search are popular and scalable but dumb; a search query for "dog" might only match documents with the literal word "dog" and miss a document with the word "poodle". "Lexical affinity" strategies use the occurrence of words such as "accident" to assess the sentiment of a document. Modern statistical NLP approaches can combine all these strategies as well as others, and often achieve acceptable accuracy at the page or paragraph level, but continue to lack the semantic understanding required to classify isolated sentences well. Besides the usual difficulties with encoding semantic commonsense knowledge, existing semantic NLP sometimes scales too poorly to be viable in business applications. Beyond semantic NLP, the ultimate goal of

"narrative" NLP is to embody a full understanding of commonsense reasoning.[117]

Perception

Main articles: Machine perception, Computer vision, and Speech recognition

Feature detection (pictured: edge detection) helps AI compose informative abstract structures out of raw data.

Machine perception[118] is the ability to use input from sensors (such as cameras (visible spectrum or infrared), microphones, wireless signals, and active lidar, sonar, radar, and tactile sensors) to deduce aspects of the world. Applications include speech recognition,[119] facial recognition, and object recognition.[120] Computer vision is the ability to analyze visual input. Such input is usually ambiguous; a giant, fifty-meter-tall pedestrian far away may produce exactly the same pixels as a nearby normal-sized pedestrian, requiring the AI to judge the relative likelihood and reasonableness of different interpretations, for example by using its "object model" to assess that fifty-meter pedestrians do not exist.[121]

Motion and manipulation

Main article: Robotics

AI is heavily used in robotics.[122] Advanced robotic arms and other industrial robots, widely used in modern factories, can learn from experience how to move efficiently despite the presence of friction and gear slippage.[123] A modern mobile robot, when given a small, static, and visible environment, can easily determine its location and map its environment; however, dynamic environments, such as (in endoscopy) the interior of a patient's breathing body, pose a greater challenge. Motion planning is the process of breaking down a movement task into "primitives" such as individual joint movements. Such movement often involves compliant motion, a process where movement requires maintaining physical contact with an object.[124][125][126] Moravec's paradox generalizes that low-level sensorimotor skills that humans take for granted are, counterintuitively, difficult to program into a robot; the paradox is named after Hans Moravec, who stated in 1988 that "it is comparatively easy to make computers exhibit adult level performance on intelligence tests or playing checkers, and difficult or impossible to give them the skills of a one-year-old when it comes to perception and mobility".[127][128] This is attributed to the fact that, unlike checkers, physical dexterity has been a direct target of natural selection for millions of years.[129]

Social intelligence

Kismet, a robot with rudimentary social skills[130]

Moravec's paradox can be extended to many forms of social intelligence.[131][132] Distributed multi-agent coordination of autonomous vehicles remains a difficult problem.[133] Affective computing is an interdisciplinary umbrella that comprises systems which recognize, interpret, process, or simulate human affects.[134][135][136] Moderate successes related to affective computing include textual sentiment analysis and, more recently, multimodal affect analysis (see multimodal sentiment analysis), wherein AI classifies the affects displayed by a videotaped subject.[137]

In the long run, social skills and an understanding of human emotion and game theory would be valuable to a social agent. Being able to predict the actions of others by understanding their motives and emotional states would allow an agent to make better decisions. Some computer systems mimic human emotion and expressions to appear more sensitive to the emotional dynamics of human interaction, or to otherwise facilitate human?computer interaction.[138] Similarly, some virtual assistants are programmed to speak conversationally or even to banter humorously; this tends to give naive users an unrealistic conception of how intelligent existing computer agents actually are.[139]

General intelligence

Historically, projects such as the Cyc knowledge base (1984?) and the massive Japanese Fifth Generation Computer Systems initiative (1982?1992) attempted to cover the breadth of human cognition. These early projects failed to escape the limitations of non-quantitative symbolic logic models and, in retrospect, greatly underestimated the difficulty of cross-domain AI. Nowadays, the vast majority of current AI researchers work instead on tractable "narrow AI" applications (such as medical diagnosis or automobile navigation).[140] Many researchers predict that such "narrow AI" work in different individual domains will eventually be incorporated into a machine with artificial general intelligence (AGI), combining most of the narrow skills mentioned in this article and at some

point even exceeding human ability in most or all these areas.[18][141] Many advances have general, cross-domain significance. One high-profile example is that DeepMind in the 2010s developed a "generalized artificial intelligence" that could learn many diverse Atari games on its own, and later developed a variant of the system which succeeds at sequential learning.[142][143][144] Besides transfer learning,[145] hypothetical AGI breakthroughs could include the development of reflective architectures that can engage in decision-theoretic metareasoning, and figuring out how to "slurp up" a comprehensive knowledge base from the entire unstructured Web.[5] Some argue that some kind of (currently-undiscovered) conceptually straightforward, but mathematically difficult, "Master Algorithm" could lead to AGI.[146] Finally, a few "emergent" approaches look to simulating human intelligence extremely closely, and believe that anthropomorphic features like an artificial brain or simulated child development may someday reach a critical point where general intelligence emerges.[147][148]

Many of the problems in this article may also require general intelligence, if machines are to solve the problems as well as people do. For example, even specific straightforward tasks, like machine translation, require that a machine read and write in both languages (NLP), follow the author's argument (reason), know what is being talked about (knowledge), and faithfully reproduce the author's original intent (social intelligence). A problem like machine translation is considered "AI-complete", because all of these problems need to be solved simultaneously in order to reach human-level machine performance.

Approaches

There is no established unifying theory or paradigm that guides AI research. Researchers disagree about many issues.[149] A few of the most long standing questions that have remained unanswered are these: should artificial intelligence simulate natural intelligence by studying psychology or neurobiology? Or is human biology as irrelevant to AI research as bird biology is to aeronautical engineering?[15] Can intelligent behavior be described using simple, elegant principles (such as logic or optimization)? Or does it necessarily require solving a large number of completely unrelated problems?[16]

Cybernetics and brain simulation

Main articles: Cybernetics and Computational neuroscience

In the 1940s and 1950s, a number of researchers explored the connection between neurobiology, information theory, and cybernetics. Some of them built machines that used electronic networks to exhibit rudimentary intelligence, such as W. Grey Walter's turtles and the Johns Hopkins Beast. Many of these researchers gathered for meetings of the Teleological Society at Princeton University and the Ratio Club in England.[150] By 1960, this approach was largely abandoned, although elements of it would be revived in the 1980s.

Symbolic

Main article: Symbolic AI

When access to digital computers became possible in the middle 1950s, AI research began to explore the possibility that human intelligence could be reduced to symbol manipulation. The research was centered in three institutions: Carnegie Mellon University, Stanford and MIT, and as described below, each one developed its own style of research. John Haugeland named these symbolic approaches to AI "good old fashioned AI" or "GOFAI".[151] During the 1960s, symbolic approaches had achieved great success at simulating high-level thinking in small demonstration programs. Approaches based on cybernetics or artificial neural networks were abandoned or pushed into the background.[152] Researchers in the 1960s and the 1970s were convinced that symbolic approaches would eventually succeed in creating a machine with artificial general intelligence and considered this the goal of their field.

Cognitive simulation

Economist Herbert Simon and Allen Newell studied human problem-solving skills and attempted to formalize them, and their work laid the foundations of the field of artificial intelligence, as well as cognitive science, operations research and management science. Their research team used the results of psychological experiments to develop programs that simulated the techniques that people used to solve problems. This tradition, centered at Carnegie Mellon University would eventually culminate in the development of the Soar architecture in the middle 1980s.[153][154]

Logic-based

Unlike Simon and Newell, John McCarthy felt that machines did not need to simulate human thought, but should instead try to find the essence of abstract reasoning and problem-solving, regardless whether people used the same algorithms.[15] His laboratory at Stanford (SAIL) focused

on using formal logic to solve a wide variety of problems, including knowledge representation, planning and learning.[155] Logic was also the focus of the work at the University of Edinburgh and elsewhere in Europe which led to the development of the programming language Prolog and the science of logic programming.[156]

Anti-logic or scruffy

Researchers at MIT (such as Marvin Minsky and Seymour Papert)[157] found that solving difficult problems in vision and natural language processing required ad-hoc solutions?they argued that there was no simple and general principle (like logic) that would capture all the aspects of intelligent behavior. Roger Schank described their "anti-logic" approaches as "scruffy" (as opposed to the "neat" paradigms at CMU and Stanford).[16] Commonsense knowledge bases (such as Doug Lenat's Cyc) are an example of "scruffy" AI, since they must be built by hand, one complicated concept at a time.[158]

Knowledge-based

When computers with large memories became available around 1970, researchers from all three traditions began to build knowledge into AI applications.[159] This "knowledge revolution" led to the development and deployment of expert systems (introduced by Edward Feigenbaum), the first truly successful form of AI software.[38] A key component of the system architecture for all expert systems is the knowledge base, which stores facts and rules that illustrate AI.[160] The knowledge revolution was also driven by the realization that enormous amounts of knowledge would be required by many simple AI applications.

Sub-symbolic

By the 1980s, progress in symbolic AI seemed to stall and many believed that symbolic systems would never be able to imitate all the processes of human cognition, especially perception, robotics, learning and pattern recognition. A number of researchers began to look into "sub-symbolic" approaches to specific AI problems.[17] Sub-symbolic methods manage to approach intelligence without specific representations of knowledge.

Embodied intelligence

This includes embodied, situated, behavior-based, and nouvelle AI. Researchers from the related field of robotics, such as Rodney Brooks, rejected symbolic AI and focused on the basic engineering problems that would allow robots to move and survive.[161] Their work revived the non-symbolic point of view of the early cybernetics researchers of the 1950s and reintroduced the use of control theory in AI. This coincided with the development of the embodied mind thesis in the related field of cognitive science: the idea that aspects of the body (such as movement, perception and visualization) are required for higher intelligence.

Within developmental robotics, developmental learning approaches are elaborated upon to allow robots to accumulate repertoires of novel skills through autonomous self-exploration, social interaction with human teachers, and the use of guidance mechanisms (active learning, maturation, motor synergies, etc.).[162][163][164][165]

Computational intelligence and soft computing

Interest in neural networks and "connectionism" was revived by David Rumelhart and others in the middle of the 1980s.[166] Artificial neural networks are an example of soft computing?they are solutions to problems which cannot be solved with complete logical certainty, and where an approximate solution is often sufficient. Other soft computing approaches to AI include fuzzy systems, Grey system theory, evolutionary computation and many statistical tools. The application of soft computing to AI is studied collectively by the emerging discipline of computational intelligence.[167]

Statistical learning

Much of traditional GOFAI got bogged down on ad hoc patches to symbolic computation that worked on their own toy models but failed to generalize to real-world results. However, around the 1990s, AI researchers adopted sophisticated mathematical tools, such as hidden Markov models (HMM), information theory, and normative Bayesian decision theory to compare or to unify competing architectures. The shared mathematical language permitted a high level of collaboration with more established fields (like mathematics, economics or operations research).[d] Compared with GOFAI, new "statistical learning" techniques such as HMM and neural networks were gaining higher levels of accuracy in many practical domains such as data mining, without necessarily acquiring a semantic understanding of the datasets. The increased successes with real-world data led to increasing emphasis on comparing different approaches against shared test data to see which

approach performed best in a broader context than that provided by idiosyncratic toy models; AI research was becoming more scientific. Nowadays results of experiments are often rigorously measurable, and are sometimes (with difficulty) reproducible.[39][168] Different statistical learning techniques have different limitations; for example, basic HMM cannot model the infinite possible combinations of natural language.[169] Critics note that the shift from GOFAI to statistical learning is often also a shift away from explainable AI. In AGI research, some scholars caution against over-reliance on statistical learning, and argue that continuing research into GOFAI will still be necessary to attain general intelligence.[170][171]

Integrating the approaches

Intelligent agent paradigm

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. More complicated agents include human beings and organizations of human beings (such as firms). The paradigm allows researchers to directly compare or even combine different approaches to isolated problems, by asking which agent is best at maximizing a given "goal function". An agent that solves a specific problem can use any approach that works?some agents are symbolic and logical, some are sub-symbolic artificial neural networks and others may use new approaches. The paradigm also gives researchers a common language to communicate with other fields?such as decision theory and economics?that also use concepts of abstract agents. Building a complete agent requires researchers to address realistic problems of integration; for example, because sensory systems give uncertain information about the environment, planning systems must be able to function in the presence of uncertainty. The intelligent agent paradigm became widely accepted during the 1990s.[172]

Agent architectures and cognitive architectures

Researchers have designed systems to build intelligent systems out of interacting intelligent agents in a multi-agent system.[173] A hierarchical control system provides a bridge between sub-symbolic AI at its lowest, reactive levels and traditional symbolic AI at its highest levels, where relaxed time constraints permit planning and world modelling.[174] Some cognitive architectures are custom-built to solve a narrow problem; others, such as Soar, are designed to mimic human cognition and to provide insight into general intelligence. Modern extensions of Soar are hybrid intelligent systems that include both symbolic and sub-symbolic components.[175][176][177]

Tools

AI has developed a large number of tools to solve the most difficult problems in computer science. A few of the most general of these methods are discussed below.

Search and optimization

Main articles: Search algorithm, Mathematical optimization, and Evolutionary computation

Many problems in AI can be solved in theory by intelligently searching through many possible solutions:[178] Reasoning can be reduced to performing a search. For example, logical proof can be viewed as searching for a path that leads from premises to conclusions, where each step is the application of an inference rule.[179] Planning algorithms search through trees of goals and subgoals, attempting to find a path to a target goal, a process called means-ends analysis.[180] Robotics algorithms for moving limbs and grasping objects use local searches in configuration space.[123] Many learning algorithms use search algorithms based on optimization.

Simple exhaustive searches[181] are rarely sufficient for most real-world problems: the search space (the number of places to search) quickly grows to astronomical numbers. The result is a search that is too slow or never completes. The solution, for many problems, is to use "heuristics" or "rules of thumb" that prioritize choices in favor of those that are more likely to reach a goal and to do so in a shorter number of steps. In some search methodologies heuristics can also serve to entirely eliminate some choices that are unlikely to lead to a goal (called "pruning the search tree"). Heuristics supply the program with a "best guess" for the path on which the solution lies.[182] Heuristics limit the search for solutions into a smaller sample size.[124]

A very different kind of search came to prominence in the 1990s, based on the mathematical theory of optimization. For many problems, it is possible to begin the search with some form of a guess and then refine the guess incrementally until no more refinements can be made. These algorithms can be visualized as blind hill climbing: we begin the search at a random point on the landscape, and then, by jumps or steps, we keep moving our guess uphill, until we reach the top. Other optimization algorithms are simulated annealing, beam search and random optimization.[183]

A particle swarm seeking the global minimum

Evolutionary computation uses a form of optimization search. For example, they may begin with a population of organisms (the guesses) and then allow them to mutate and recombine, selecting only the fittest to survive each generation (refining the guesses). Classic evolutionary algorithms include genetic algorithms, gene expression programming, and genetic programming.[184] Alternatively, distributed search processes can coordinate via swarm intelligence algorithms. Two popular swarm algorithms used in search are particle swarm optimization (inspired by bird flocking) and ant colony optimization (inspired by ant trails).[185][186]

Logic

Main articles: Logic programming and Automated reasoning

Logic[187] is used for knowledge representation and problem solving, but it can be applied to other problems as well. For example, the satplan algorithm uses logic for planning[188] and inductive logic programming is a method for learning.[189]

Several different forms of logic are used in AI research. Propositional logic[190] involves truth functions such as "or" and "not". First-order logic[191] adds quantifiers and predicates, and can express facts about objects, their properties, and their relations with each other. Fuzzy set theory assigns a "degree of truth" (between 0 and 1) to vague statements such as "Alice is old" (or rich, or tall, or hungry) that are too linguistically imprecise to be completely true or false. Fuzzy logic is successfully used in control systems to allow experts to contribute vague rules such as "if you are close to the destination station and moving fast, increase the train's brake pressure"; these vague rules can then be numerically refined within the system. Fuzzy logic fails to scale well in knowledge bases; many AI researchers question the validity of chaining fuzzy-logic inferences.[e][193][194]

Default logics, non-monotonic logics and circumscription[99] are forms of logic designed to help with default reasoning and the qualification problem. Several extensions of logic have been designed to handle specific domains of knowledge, such as: description logics;[87] situation calculus, event calculus and fluent calculus (for representing events and time);[88] causal calculus;[89] belief calculus;[195] and modal logics.[90]

Overall, qualitative symbolic logic is brittle and scales poorly in the presence of noise or other uncertainty. Exceptions to rules are numerous, and it is difficult for logical systems to function in

the presence of contradictory rules.[196][197]

Probabilistic methods for uncertain reasoning

Main articles: Bayesian network, Hidden Markov model, Kalman filter, Particle filter, Decision theory, and Utility theory

Expectation-maximization clustering of Old Faithful eruption data starts from a random guess but then successfully converges on an accurate clustering of the two physically distinct modes of eruption.

Many problems in AI (in reasoning, planning, learning, perception, and robotics) require the agent to operate with incomplete or uncertain information. AI researchers have devised a number of powerful tools to solve these problems using methods from probability theory and economics.[198]

Bayesian networks[199] are a very general tool that can be used for a large number of problems: reasoning (using the Bayesian inference algorithm),[200] learning (using the expectation-maximization algorithm),[f][202] planning (using decision networks)[203] and perception (using dynamic Bayesian networks).[204] Probabilistic algorithms can also be used for filtering, prediction, smoothing and finding explanations for streams of data, helping perception systems to analyze processes that occur over time (e.g., hidden Markov models or Kalman filters).[204] Compared with symbolic logic, formal Bayesian inference is computationally expensive. For inference to be tractable, most observations must be conditionally independent of one another. Complicated graphs with diamonds or other "loops" (undirected cycles) can require a sophisticated method such as Markov chain Monte Carlo, which spreads an ensemble of random walkers throughout the Bayesian network and attempts to converge to an assessment of the conditional probabilities. Bayesian networks are used on Xbox Live to rate and match players; wins and losses are "evidence" of how good a player is[citation needed]. AdSense uses a Bayesian network with over 300 million edges to learn which ads to serve.[196]

A key concept from the science of economics is "utility": a measure of how valuable something is to an intelligent agent. Precise mathematical tools have been developed that analyze how an agent can make choices and plan, using decision theory, decision analysis,[205] and information value theory.[105] These tools include models such as Markov decision processes,[206] dynamic decision

networks,[204] game theory and mechanism design.[207]

Classifiers and statistical learning methods

Main articles: Classifier (mathematics), Statistical classification, and Machine learning

The simplest AI applications can be divided into two types: classifiers ("if shiny then diamond") and controllers ("if shiny then pick up"). Controllers do, however, also classify conditions before inferring actions, and therefore classification forms a central part of many AI systems. Classifiers are functions that use pattern matching to determine a closest match. They can be tuned according to examples, making them very attractive for use in AI. These examples are known as observations or patterns. In supervised learning, each pattern belongs to a certain predefined class. A class can be seen as a decision that has to be made. All the observations combined with their class labels are known as a data set. When a new observation is received, that observation is classified based on previous experience.[208]

A classifier can be trained in various ways; there are many statistical and machine learning approaches. The decision tree[209] is perhaps the most widely used machine learning algorithm.[210] Other widely used classifiers are the neural network,[211] k-nearest neighbor algorithm,[g][213] kernel methods such as the support vector machine (SVM),[h][215] Gaussian mixture model,[216] and the extremely popular naive Bayes classifier.[i][218] Classifier performance depends greatly on the characteristics of the data to be classified, such as the dataset size, distribution of samples across classes, the dimensionality, and the level of noise. Model-based classifiers perform well if the assumed model is an extremely good fit for the actual data. Otherwise, if no matching model is available, and if accuracy (rather than speed or scalability) is the sole concern, conventional wisdom is that discriminative classifiers (especially SVM) tend to be more accurate than model-based classifiers such as "naive Bayes" on most practical data sets.[219][220]

Artificial neural networks

Main articles: Artificial neural network and Connectionism

A neural network is an interconnected group of nodes, akin to the vast network of neurons in the human brain.

Neural networks were inspired by the architecture of neurons in the human brain. A simple "neuron" N accepts input from multiple other neurons, each of which, when activated (or "fired"), cast a weighted "vote" for or against whether neuron N should itself activate. Learning requires an algorithm to adjust these weights based on the training data; one simple algorithm (dubbed "fire together, wire together") is to increase the weight between two connected neurons when the activation of one triggers the successful activation of another. The neural network forms "concepts" that are distributed among a subnetwork of shared[j] neurons that tend to fire together; a concept meaning "leg" might be coupled with a subnetwork meaning "foot" that includes the sound for "foot". Neurons have a continuous spectrum of activation; in addition, neurons can process inputs in a nonlinear way rather than weighing straightforward votes. Modern neural networks can learn both continuous functions and, surprisingly, digital logical operations. Neural networks' early successes included predicting the stock market and (in 1995) a mostly self-driving car.[k][221] In the 2010s, advances in neural networks using deep learning thrust AI into widespread public consciousness and contributed to an enormous upshift in corporate AI spending; for example, AI-related M&A in 2017 was over 25 times as large as in 2015.[222][223]

The study of non-learning artificial neural networks[211] began in the decade before the field of AI research was founded, in the work of Walter Pitts and Warren McCullouch. Frank Rosenblatt invented the perceptron, a learning network with a single layer, similar to the old concept of linear regression. Early pioneers also include Alexey Grigorevich Ivakhnenko, Teuvo Kohonen, Stephen Grossberg, Kunihiko Fukushima, Christoph von der Malsburg, David Willshaw, Shun-Ichi Amari, Bernard Widrow, John Hopfield, Eduardo R. Caianiello, and others[citation needed].

The main categories of networks are acyclic or feedforward neural networks (where the signal passes in only one direction) and recurrent neural networks (which allow feedback and short-term memories of previous input events). Among the most popular feedforward networks are perceptrons, multi-layer perceptrons and radial basis networks.[224] Neural networks can be applied to the problem of intelligent control (for robotics) or learning, using such techniques as Hebbian learning ("fire together, wire together"), GMDH or competitive learning.[225]

Today, neural networks are often trained by the backpropagation algorithm, which had been around since 1970 as the reverse mode of automatic differentiation published by Seppo Linnainmaa,[226][227] and was introduced to neural networks by Paul Werbos.[228][229][230]

Hierarchical temporal memory is an approach that models some of the structural and algorithmic properties of the neocortex.[231]

To summarize, most neural networks use some form of gradient descent on a hand-created neural topology. However, some research groups, such as Uber, argue that simple neuroevolution to mutate new neural network topologies and weights may be competitive with sophisticated gradient descent approaches[citation needed]. One advantage of neuroevolution is that it may be less prone to get caught in "dead ends".[232]

Deep feedforward neural networks

Main article: Deep learning

Deep learning is any artificial neural network that can learn a long chain of causal links[dubious ? discuss]. For example, a feedforward network with six hidden layers can learn a seven-link causal chain (six hidden layers + output layer) and has a "credit assignment path" (CAP) depth of seven[citation needed]. Many deep learning systems need to be able to learn chains ten or more causal links in length.[233] Deep learning has transformed many important subfields of artificial intelligence[why?], including computer vision, speech recognition, natural language processing and others.[234][235][233]

According to one overview,[236] the expression "Deep Learning" was introduced to the machine learning community by Rina Dechter in 1986[237] and gained traction after Igor Aizenberg and colleagues introduced it to artificial neural networks in 2000.[238] The first functional Deep Learning networks were published by Alexey Grigorevich Ivakhnenko and V. G. Lapa in 1965.[239][page needed] These networks are trained one layer at a time. Ivakhnenko's 1971 paper[240] describes the learning of a deep feedforward multilayer perceptron with eight layers, already much deeper than many later networks. In 2006, a publication by Geoffrey Hinton and Ruslan Salakhutdinov introduced another way of pre-training many-layered feedforward neural networks (FNNs) one layer at a time, treating each layer in turn as an unsupervised restricted Boltzmann machine, then using supervised backpropagation for fine-tuning.[241] Similar to shallow artificial neural networks, deep neural networks can model complex non-linear relationships. Over the last few years, advances in both machine learning algorithms and computer hardware have led to more efficient methods for training

deep neural networks that contain many layers of non-linear hidden units and a very large output layer.[242]

Deep learning often uses convolutional neural networks (CNNs), whose origins can be traced back to the Neocognitron introduced by Kunihiko Fukushima in 1980.[243] In 1989, Yann LeCun and colleagues applied backpropagation to such an architecture. In the early 2000s, in an industrial application CNNs already processed an estimated 10% to 20% of all the checks written in the US.[244] Since 2011, fast implementations of CNNs on GPUs have won many visual pattern recognition competitions.[233]

CNNs with 12 convolutional layers were used in conjunction with reinforcement learning by Deepmind's "AlphaGo Lee", the program that beat a top Go champion in 2016.[245]

Deep recurrent neural networks

Main article: Recurrent neural networks

Early on, deep learning was also applied to sequence learning with recurrent neural networks (RNNs)[246] which are in theory Turing complete[247] and can run arbitrary programs to process arbitrary sequences of inputs. The depth of an RNN is unlimited and depends on the length of its input sequence; thus, an RNN is an example of deep learning.[233] RNNs can be trained by gradient descent[248][249][250] but suffer from the vanishing gradient problem.[234][251] In 1992, it was shown that unsupervised pre-training of a stack of recurrent neural networks can speed up subsequent supervised learning of deep sequential problems.[252]

Numerous researchers now use variants of a deep learning recurrent NN called the long short-term memory (LSTM) network published by Hochreiter & Schmidhuber in 1997.[253] LSTM is often trained by Connectionist Temporal Classification (CTC).[254] At Google, Microsoft and Baidu this approach has revolutionised speech recognition.[255][256][257] For example, in 2015, Google's speech recognition experienced a dramatic performance jump of 49% through CTC-trained LSTM, which is now available through Google Voice to billions of smartphone users.[258] Google also used LSTM to improve machine translation,[259] Language Modeling[260] and Multilingual Language Processing.[261] LSTM combined with CNNs also improved automatic image captioning[262] and a plethora of other applications.

Evaluating progress

Further information: Progress in artificial intelligence and Competitions and prizes in artificial intelligence

AI, like electricity or the steam engine, is a general purpose technology. There is no consensus on how to characterize which tasks AI tends to excel at.[263] While projects such as AlphaZero have succeeded in generating their own knowledge from scratch, many other machine learning projects require large training datasets.[264][265] Researcher Andrew Ng has suggested, as a "highly imperfect rule of thumb", that "almost anything a typical human can do with less than one second of mental thought, we can probably now or in the near future automate using AI."[266] Moravec's paradox suggests that AI lags humans at many tasks that the human brain has specifically evolved to perform well.[129]

Games provide a well-publicized benchmark for assessing rates of progress. AlphaGo around 2016 brought the era of classical board-game benchmarks to a close. Games of imperfect knowledge provide new challenges to AI in the area of game theory.[267][268] E-sports such as StarCraft continue to provide additional public benchmarks.[269][270] There are many competitions and prizes, such as the Imagenet Challenge, to promote research in artificial intelligence. The most common areas of competition include general machine intelligence, conversational behavior, data-mining, robotic cars, and robot soccer as well as conventional games.[271]

The "imitation game" (an interpretation of the 1950 Turing test that assesses whether a computer can imitate a human) is nowadays considered too exploitable to be a meaningful benchmark.[272] A derivative of the Turing test is the Completely Automated Public Turing test to tell Computers and Humans Apart (CAPTCHA). As the name implies, this helps to determine that a user is an actual person and not a computer posing as a human. In contrast to the standard Turing test, CAPTCHA is administered by a machine and targeted to a human as opposed to being administered by a human and targeted to a machine. A computer asks a user to complete a simple test then generates a grade for that test. Computers are unable to solve the problem, so correct solutions are deemed to be the result of a person taking the test. A common type of CAPTCHA is the test that requires the typing of distorted letters, numbers or symbols that appear in an image undecipherable by a computer.[273]

Proposed "universal intelligence" tests aim to compare how well machines, humans, and even non-human animals perform on problem sets that are generic as possible. At an extreme, the test suite can contain every possible problem, weighted by Kolmogorov complexity; unfortunately, these problem sets tend to be dominated by impoverished pattern-matching exercises where a tuned AI can easily exceed human performance levels.[274][275]

Applications

An automated online assistant providing customer service on a web page ? one of many very primitive applications of artificial intelligence

Main article: Applications of artificial intelligence

AI is relevant to any intellectual task.[276] Modern artificial intelligence techniques are pervasive and are too numerous to list here. Frequently, when a technique reaches mainstream use, it is no longer considered artificial intelligence; this phenomenon is described as the AI effect.[277]

High-profile examples of AI include autonomous vehicles (such as drones and self-driving cars), medical diagnosis, creating art (such as poetry), proving mathematical theorems, playing games (such as Chess or Go), search engines (such as Google search), online assistants (such as Siri), image recognition in photographs, spam filtering, predicting flight delays,[278] prediction of judicial decisions[279] and targeting online advertisements.[276][280][281]

With social media sites overtaking TV as a source for news for young people and news organizations increasingly reliant on social media platforms for generating distribution,[282] major publishers now use artificial intelligence (AI) technology to post stories more effectively and generate higher volumes of traffic.[283]

Healthcare

Main article: Artificial intelligence in healthcare

A patient-side surgical arm of Da Vinci Surgical System

AI is being applied to the high cost problem of dosage issues?where findings suggested that AI could save $16 billion. In 2016, a ground breaking study in California found that a mathematical formula developed with the help of AI correctly determined the accurate dose of immunosuppressant drugs to give to organ patients.[284]

X-ray of a hand, with automatic calculation of bone age by computer software

Artificial intelligence is breaking into the healthcare industry by assisting doctors. According to Bloomberg Technology, Microsoft has developed AI to help doctors find the right treatments for cancer.[285] There is a great amount of research and drugs developed relating to cancer. In detail, there are more than 800 medicines and vaccines to treat cancer. This negatively affects the doctors, because there are too many options to choose from, making it more difficult to choose the right drugs for the patients. Microsoft is working on a project to develop a machine called "Hanover"[citation needed]. Its goal is to memorize all the papers necessary to cancer and help predict which combinations of drugs will be most effective for each patient. One project that is being worked on at the moment is fighting myeloid leukemia, a fatal cancer where the treatment has not improved in decades. Another study was reported to have found that artificial intelligence was as good as trained doctors in identifying skin cancers.[286] Another study is using artificial intelligence to try and monitor multiple high-risk patients, and this is done by asking each patient numerous questions based on data acquired from live doctor to patient interactions.[287] One study was done with transfer learning, the machine performed a diagnosis similarly to a well-trained ophthalmologist, and could generate a decision within 30 seconds on whether or not the patient should be referred for treatment, with more than 95% accuracy.[288]

According to CNN, a recent study by surgeons at the Children's National Medical Center in Washington successfully demonstrated surgery with an autonomous robot. The team supervised the robot while it performed soft-tissue surgery, stitching together a pig's bowel during open surgery, and doing so better than a human surgeon, the team claimed.[289] IBM has created its own artificial intelligence computer, the IBM Watson, which has beaten human intelligence (at some levels). Watson not only won at the game show Jeopardy! against former champions,[290] but was declared a hero after successfully diagnosing a woman who was suffering from leukemia.[291]

Automotive

Main article: driverless cars

Advancements in AI have contributed to the growth of the automotive industry through the creation and evolution of self-driving vehicles. As of 2016, there are over 30 companies utilizing AI into the creation of driverless cars. A few companies involved with AI include Tesla, Google, and Apple.[292]

Many components contribute to the functioning of self-driving cars. These vehicles incorporate systems such as braking, lane changing, collision prevention, navigation and mapping. Together, these systems, as well as high performance computers, are integrated into one complex vehicle.[293]

Recent developments in autonomous automobiles have made the innovation of self-driving trucks possible, though they are still in the testing phase. The UK government has passed legislation to begin testing of self-driving truck platoons in 2018.[294] Self-driving truck platoons are a fleet of self-driving trucks following the lead of one non-self-driving truck, so the truck platoons aren't entirely autonomous yet. Meanwhile, the Daimler, a German automobile corporation, is testing the Freightliner Inspiration which is a semi-autonomous truck that will only be used on the highway.[295]

One main factor that influences the ability for a driver-less automobile to function is mapping. In general, the vehicle would be pre-programmed with a map of the area being driven. This map would include data on the approximations of street light and curb heights in order for the vehicle to be aware of its surroundings. However, Google has been working on an algorithm with the purpose of eliminating the need for pre-programmed maps and instead, creating a device that would be able to adjust to a variety of new surroundings.[296] Some self-driving cars are not equipped with steering wheels or brake pedals, so there has also been research focused on creating an algorithm that is capable of maintaining a safe environment for the passengers in the vehicle through awareness of speed and driving conditions.[297]

Another factor that is influencing the ability for a driver-less automobile is the safety of the passenger. To make a driver-less automobile, engineers must program it to handle high-risk situations. These situations could include a head-on collision with pedestrians. The car's main goal should be to make a decision that would avoid hitting the pedestrians and saving the passengers in the car. But there is a possibility the car would need to make a decision that would put someone in danger. In other words, the car would need to decide to save the pedestrians or the

passengers.[298] The programming of the car in these situations is crucial to a successful driver-less automobile.

Finance and economics

Financial institutions have long used artificial neural network systems to detect charges or claims outside of the norm, flagging these for human investigation. The use of AI in banking can be traced back to 1987 when Security Pacific National Bank in US set-up a Fraud Prevention Task force to counter the unauthorized use of debit cards[citation needed]. Programs like Kasisto and Moneystream are using AI in financial services.

Banks use artificial intelligence systems today to organize operations, maintain book-keeping, invest in stocks, and manage properties. AI can react to changes overnight or when business is not taking place.[299] In August 2001, robots beat humans in a simulated financial trading competition.[300] AI has also reduced fraud and financial crimes by monitoring behavioral patterns of users for any abnormal changes or anomalies.[301]

The use of AI machines in the market in applications such as online trading and decision making has changed major economic theories.[302] For example, AI based buying and selling platforms have changed the law of supply and demand in that it is now possible to easily estimate individualized demand and supply curves and thus individualized pricing. Furthermore, AI machines reduce information asymmetry in the market and thus making markets more efficient while reducing the volume of trades[citation needed]. Furthermore, AI in the markets limits the consequences of behavior in the markets again making markets more efficient[citation needed]. Other theories where AI has had impact include in rational choice, rational expectations, game theory, Lewis turning point, portfolio optimization and counterfactual thinking[citation needed].

Government

Main article: Artificial intelligence in government

See also: Mass surveillance in China

Artificial intelligence paired with facial recognition systems may be used for mass surveillance. This is already the case in some parts of China.[303][304]

Video games

Main article: Artificial intelligence (video games)

In video games, artificial intelligence is routinely used to generate dynamic purposeful behavior in non-player characters (NPCs). In addition, well-understood AI techniques are routinely used for pathfinding. Some researchers consider NPC AI in games to be a "solved problem" for most production tasks. Games with more atypical AI include the AI director of Left 4 Dead (2008) and the neuroevolutionary training of platoons in Supreme Commander 2 (2010).[305][306]

Military

Further information: Artificial intelligence arms race, Lethal autonomous weapon, and Unmanned combat aerial vehicle

The main military applications of Artificial Intelligence and Machine Learning are to enhance C2, Communications, Sensors, Integration and Interoperability[307]. Artifical Intelligence technologies enables coordination of sensors and effectors, threat detection and identification, marking of enemy positions, target acquisition, coordination and deconfliction of distributed Join Fires between networked combat vehicles and tanks also inside Manned and Unmanned Teams (MUM-T)[307].

Worldwide annual military spending on robotics rose from US$5.1 billion in 2010 to US$7.5 billion in 2015.[308][309] Military drones capable of autonomous action are widely considered a useful asset.[310] Many artificial intelligence researchers seek to distance themselves from military applications of AI.[311]

Audit

For financial statements audit, AI makes continuous audit possible. AI tools could analyze many sets of different information immediately. The potential benefit would be the overall audit risk will be reduced, the level of assurance will be increased and the time duration of audit will be reduced.[312]

Advertising

It is possible to use AI to predict or generalize the behavior of customers from their digital footprints

in order to target them with personalized promotions or build customer personas automatically.[313] A documented case reports that online gambling companies were using AI to improve customer targeting.[314]

Moreover, the application of Personality computing AI models can help reducing the cost of advertising campaigns by adding psychological targeting to more traditional sociodemographic or behavioral targeting.[315]

Art

Further information: Computer art

Artificial Intelligence has inspired numerous creative applications including its usage to produce visual art. The exhibition "Thinking Machines: Art and Design in the Computer Age, 1959?1989" at MoMA[316] provides a good overview of the historical applications of AI for art, architecture, and design. Recent exhibitions showcasing the usage of AI to produce art include the Google-sponsored benefit and auction at the Gray Area Foundation in San Francisco, where artists experimented with the deepdream algorithm[317] and the exhibition "Unhuman: Art in the Age of AI," which took place in Los Angeles and Frankfurt in the fall of 2017.[318][319] In the spring of 2018, the Association of Computing Machinery dedicated a special magazine issue to the subject of computers and art highlighting the role of machine learning in the arts.[320]

Philosophy and ethics

Main articles: Philosophy of artificial intelligence and Ethics of artificial intelligence

There are three philosophical questions related to AI:

Is artificial general intelligence possible? Can a machine solve any problem that a human being can solve using intelligence? Or are there hard limits to what a machine can accomplish?

Are intelligent machines dangerous? How can we ensure that machines behave ethically and that they are used ethically?

Can a machine have a mind, consciousness and mental states in exactly the same sense that human beings do? Can a machine be sentient, and thus deserve certain rights? Can a machine intentionally

cause harm?

The limits of artificial general intelligence

Main articles: Philosophy of AI, Turing test, Physical symbol systems hypothesis, Dreyfus' critique of AI, The Emperor's New Mind, and AI effect

Can a machine be intelligent? Can it "think"?

Alan Turing's "polite convention"

We need not decide if a machine can "think"; we need only decide if a machine can act as intelligently as a human being. This approach to the philosophical problems associated with artificial intelligence forms the basis of the Turing test.[321]

The Dartmouth proposal

"Every aspect of learning or any other feature of intelligence can be so precisely described that a machine can be made to simulate it." This conjecture was printed in the proposal for the Dartmouth Conference of 1956.[322]

Newell and Simon's physical symbol system hypothesis

"A physical symbol system has the necessary and sufficient means of general intelligent action." Newell and Simon argue that intelligence consists of formal operations on symbols.[323] Hubert Dreyfus argued that, on the contrary, human expertise depends on unconscious instinct rather than conscious symbol manipulation and on having a "feel" for the situation rather than explicit symbolic knowledge. (See Dreyfus' critique of AI.)[324][325]

Godelian arguments

Godel himself,[326] John Lucas (in 1961) and Roger Penrose (in a more detailed argument from 1989 onwards) made highly technical arguments that human mathematicians can consistently see the truth of their own "Godel statements" and therefore have computational abilities beyond that of mechanical Turing machines.[327] However, some people do not agree with the "Godelian arguments". [328][329][330]

The artificial brain argument

The brain can be simulated by machines and because brains are intelligent, simulated brains must also be intelligent; thus machines can be intelligent. Hans Moravec, Ray Kurzweil and others have argued that it is technologically feasible to copy the brain directly into hardware and software and

that such a simulation will be essentially identical to the original.[147]

The AI effect

Machines are already intelligent, but observers have failed to recognize it. When Deep Blue beat Garry Kasparov in chess, the machine was acting intelligently. However, onlookers commonly discount the behavior of an artificial intelligence program by arguing that it is not "real" intelligence after all; thus "real" intelligence is whatever intelligent behavior people can do that machines still cannot. This is known as the AI Effect: "AI is whatever hasn't been done yet."

Potential harm

Widespread use of artificial intelligence could have unintended consequences that are dangerous or undesirable. Scientists from the Future of Life Institute, among others, described some short-term research goals to see how AI influences the economy, the laws and ethics that are involved with AI and how to minimize AI security risks. In the long-term, the scientists have proposed to continue optimizing function while minimizing possible security risks that come along with new technologies.[331]

The potential negative effects of AI and automation are a major issue for Andrew Yang's presidential campaign.[332]

Existential risk

Main article: Existential risk from artificial general intelligence

Physicist Stephen Hawking, Microsoft founder Bill Gates, and SpaceX founder Elon Musk have expressed concerns about the possibility that AI could evolve to the point that humans could not control it, with Hawking theorizing that this could "spell the end of the human race". [333] [334] [335]

The development of full artificial intelligence could spell the end of the human race. Once humans develop artificial intelligence, it will take off on its own and redesign itself at an ever-increasing rate. Humans, who are limited by slow biological evolution, couldn't compete and would be superseded.

??Stephen Hawking[336]

In his book Superintelligence, Nick Bostrom provides an argument that artificial intelligence will pose a threat to humankind. He argues that sufficiently intelligent AI, if it chooses actions based on achieving some goal, will exhibit convergent behavior such as acquiring resources or protecting itself from being shut down. If this AI's goals do not reflect humanity's?one example is an AI told to compute as many digits of pi as possible?it might harm humanity in order to acquire more resources or prevent itself from being shut down, ultimately to better achieve its goal.

Concern over risk from artificial intelligence has led to some high-profile donations and investments. A group of prominent tech titans including Peter Thiel, Amazon Web Services and Musk have committed $1billion to OpenAI, a nonprofit company aimed at championing responsible AI development.[337] The opinion of experts within the field of artificial intelligence is mixed, with sizable fractions both concerned and unconcerned by risk from eventual superhumanly-capable AI.[338] Other technology industry leaders believe that artificial intelligence is helpful in its current form and will continue to assist humans. Oracle CEO Mark Hurd has stated that AI "will actually create more jobs, not less jobs" as humans will be needed to manage AI systems.[339] Facebook CEO Mark Zuckerberg believes AI will "unlock a huge amount of positive things," such as curing disease and increasing the safety of autonomous cars.[340]

In January 2015, Elon Musk donated ten million dollars to the Future of Life Institute to fund research on understanding AI decision making. The goal of the institute is to "grow wisdom with which we manage" the growing power of technology. Musk also funds companies developing artificial intelligence such as Google DeepMind and Vicarious to "just keep an eye on what's going on with artificial intelligence.[341] I think there is potentially a dangerous outcome there."[342][343]

For this danger to be realized, the hypothetical AI would have to overpower or out-think all of humanity, which a minority of experts argue is a possibility far enough in the future to not be worth researching.[344][345] Other counterarguments revolve around humans being either intrinsically or convergently valuable from the perspective of an artificial intelligence.[346]

Devaluation of humanity

Main article: Computer Power and Human Reason

Joseph Weizenbaum wrote that AI applications cannot, by definition, successfully simulate genuine

human empathy and that the use of AI technology in fields such as customer service or psychotherapy[347] was deeply misguided. Weizenbaum was also bothered that AI researchers (and some philosophers) were willing to view the human mind as nothing more than a computer program (a position is now known as computationalism). To Weizenbaum these points suggest that AI research devalues human life.[348]

Social justice

One concern is that AI programs may be programmed to be biased against certain groups, such as women and minorities, because most of the developers are wealthy Caucasian men.[349] Support for artificial intelligence is higher among men (with 47% approving) than women (35% approving).

Algorithms have a host of applications in today's legal system already, assisting officials ranging from judges to parole officers and public defenders in gauging the predicted likelihood of recidivism of defendants.[350] COMPAS (an acronym for Correctional Offender Management Profiling for Alternative Sanctions) counts among the most widely utilized commercially available solutions.[350] It has been suggested that COMPAS assigns an exceptionally elevated risk of recidivism to black defendants while, conversely, ascribing low risk estimate to white defendants significantly more often than statistically expected. [350]

Decrease in demand for human labor

Further information: Technological unemployment § 21st century

The relationship between automation and employment is complicated. While automation eliminates old jobs, it also creates new jobs through micro-economic and macro-economic effects.[351] Unlike previous waves of automation, many middle-class jobs may be eliminated by artificial intelligence; The Economist states that "the worry that AI could do to white-collar jobs what steam power did to blue-collar ones during the Industrial Revolution" is "worth taking seriously".[352] Subjective estimates of the risk vary widely; for example, Michael Osborne and Carl Benedikt Frey estimate 47% of U.S. jobs are at "high risk" of potential automation, while an OECD report classifies only 9% of U.S. jobs as "high risk".[353][354][355] Jobs at extreme risk range from paralegals to fast food cooks, while job demand is likely to increase for care-related professions ranging from personal healthcare to the clergy.[356] Author Martin Ford and others go further and argue that a large number of jobs are routine, repetitive and (to an AI) predictable; Ford warns that these jobs may be automated in

the next couple of decades, and that many of the new jobs may not be "accessible to people with average capability", even with retraining. Economists point out that in the past technology has tended to increase rather than reduce total employment, but acknowledge that "we're in uncharted territory" with AI.[22]

Autonomous weapons

See also: Lethal autonomous weapon

Currently, 50+ countries are researching battlefield robots, including the United States, China, Russia, and the United Kingdom. Many people concerned about risk from superintelligent AI also want to limit the use of artificial soldiers and drones.[357]

Ethical machines

Machines with intelligence have the potential to use their intelligence to prevent harm and minimize the risks; they may have the ability to use ethical reasoning to better choose their actions in the world. Research in this area includes machine ethics, artificial moral agents, and friendly AI.

Artificial moral agents

Wendell Wallach introduced the concept of artificial moral agents (AMA) in his book Moral Machines[358] For Wallach, AMAs have become a part of the research landscape of artificial intelligence as guided by its two central questions which he identifies as "Does Humanity Want Computers Making Moral Decisions"[359] and "Can (Ro)bots Really Be Moral".[360] For Wallach the question is not centered on the issue of whether machines can demonstrate the equivalent of moral behavior in contrast to the constraints which society may place on the development of AMAs.[361]

Machine ethics

Main article: Machine ethics

The field of machine ethics is concerned with giving machines ethical principles, or a procedure for discovering a way to resolve the ethical dilemmas they might encounter, enabling them to function in an ethically responsible manner through their own ethical decision making.[362] The field was delineated in the AAAI Fall 2005 Symposium on Machine Ethics: "Past research concerning the

relationship between technology and ethics has largely focused on responsible and irresponsible use of technology by human beings, with a few people being interested in how human beings ought to treat machines. In all cases, only human beings have engaged in ethical reasoning. The time has come for adding an ethical dimension to at least some machines. Recognition of the ethical ramifications of behavior involving machines, as well as recent and potential developments in machine autonomy, necessitate this. In contrast to computer hacking, software property issues, privacy issues and other topics normally ascribed to computer ethics, machine ethics is concerned with the behavior of machines towards human users and other machines. Research in machine ethics is key to alleviating concerns with autonomous systems?it could be argued that the notion of autonomous machines without such a dimension is at the root of all fear concerning machine intelligence. Further, investigation of machine ethics could enable the discovery of problems with current ethical theories, advancing our thinking about Ethics."[363] Machine ethics is sometimes referred to as machine morality, computational ethics or computational morality. A variety of perspectives of this nascent field can be found in the collected edition "Machine Ethics"[362] that stems from the AAAI Fall 2005 Symposium on Machine Ethics.[363]

Malevolent and friendly AI

Main article: Friendly AI

Political scientist Charles T. Rubin believes that AI can be neither designed nor guaranteed to be benevolent.[364] He argues that "any sufficiently advanced benevolence may be indistinguishable from malevolence." Humans should not assume machines or robots would treat us favorably because there is no a priori reason to believe that they would be sympathetic to our system of morality, which has evolved along with our particular biology (which AIs would not share). Hyper-intelligent software may not necessarily decide to support the continued existence of humanity and would be extremely difficult to stop. This topic has also recently begun to be discussed in academic publications as a real source of risks to civilization, humans, and planet Earth.

One proposal to deal with this is to ensure that the first generally intelligent AI is 'Friendly AI' and will be able to control subsequently developed AIs. Some question whether this kind of check could actually remain in place.

Leading AI researcher Rodney Brooks writes, "I think it is a mistake to be worrying about us developing malevolent AI anytime in the next few hundred years. I think the worry stems from a

fundamental error in not distinguishing the difference between the very real recent advances in a particular aspect of AI, and the enormity and complexity of building sentient volitional intelligence."[365]

Machine consciousness, sentience and mind

Main article: Artificial consciousness

If an AI system replicates all key aspects of human intelligence, will that system also be sentient?will it have a mind which has conscious experiences? This question is closely related to the philosophical problem as to the nature of human consciousness, generally referred to as the hard problem of consciousness.

Consciousness

Main articles: Hard problem of consciousness and Theory of mind

David Chalmers identified two problems in understanding the mind, which he named the "hard" and "easy" problems of consciousness.[366] The easy problem is understanding how the brain processes signals, makes plans and controls behavior. The hard problem is explaining how this feels or why it should feel like anything at all. Human information processing is easy to explain, however human subjective experience is difficult to explain.

For example, consider what happens when a person is shown a color swatch and identifies it, saying "it's red". The easy problem only requires understanding the machinery in the brain that makes it possible for a person to know that the color swatch is red. The hard problem is that people also know something else?they also know what red looks like. (Consider that a person born blind can know that something is red without knowing what red looks like.)[l] Everyone knows subjective experience exists, because they do it every day (e.g., all sighted people know what red looks like). The hard problem is explaining how the brain creates it, why it exists, and how it is different from knowledge and other aspects of the brain.

Computationalism and functionalism

Main articles: Computationalism and Functionalism (philosophy of mind)

Computationalism is the position in the philosophy of mind that the human mind or the human brain (or both) is an information processing system and that thinking is a form of computing.[367] Computationalism argues that the relationship between mind and body is similar or identical to the relationship between software and hardware and thus may be a solution to the mind-body problem. This philosophical position was inspired by the work of AI researchers and cognitive scientists in the 1960s and was originally proposed by philosophers Jerry Fodor and Hilary Putnam.

Strong AI hypothesis

Main article: Chinese room

The philosophical position that John Searle has named "strong AI" states: "The appropriately programmed computer with the right inputs and outputs would thereby have a mind in exactly the same sense human beings have minds."[368] Searle counters this assertion with his Chinese room argument, which asks us to look inside the computer and try to find where the "mind" might be.[369]

Robot rights

Main article: Robot rights

If a machine can be created that has intelligence, could it also feel? If it can feel, does it have the same rights as a human? This issue, now known as "robot rights", is currently being considered by, for example, California's Institute for the Future, although many critics believe that the discussion is premature.[370] Some critics of transhumanism argue that any hypothetical robot rights would lie on a spectrum with animal rights and human rights. [371] The subject is profoundly discussed in the 2010 documentary film Plug & Pray,[372] and many sci fi media such as Star Trek Next Generation, with the character of Commander Data, who fought being disassembled for research, and wanted to "become human", and the robotic holograms in Voyager.

Superintelligence

Main article: Superintelligence

Are there limits to how intelligent machines?or human-machine hybrids?can be? A superintelligence, hyperintelligence, or superhuman intelligence is a hypothetical agent that would possess intelligence far surpassing that of the brightest and most gifted human mind. Superintelligence may also refer to the form or degree of intelligence possessed by such an agent.[141]

Technological singularity

Main articles: Technological singularity and Moore's law

If research into Strong AI produced sufficiently intelligent software, it might be able to reprogram and improve itself. The improved software would be even better at improving itself, leading to recursive self-improvement.[373] The new intelligence could thus increase exponentially and dramatically surpass humans. Science fiction writer Vernor Vinge named this scenario "singularity".[374] Technological singularity is when accelerating progress in technologies will cause a runaway effect wherein artificial intelligence will exceed human intellectual capacity and control, thus radically changing or even ending civilization. Because the capabilities of such an intelligence may be impossible to comprehend, the technological singularity is an occurrence beyond which events are unpredictable or even unfathomable.[374][141]

Ray Kurzweil has used Moore's law (which describes the relentless exponential improvement in digital technology) to calculate that desktop computers will have the same processing power as human brains by the year 2029, and predicts that the singularity will occur in 2045.[374]

Transhumanism

Main article: Transhumanism

Robot designer Hans Moravec, cyberneticist Kevin Warwick and inventor Ray Kurzweil have predicted that humans and machines will merge in the future into cyborgs that are more capable and powerful than either.[375] This idea, called transhumanism, which has roots in Aldous Huxley and Robert Ettinger.

Edward Fredkin argues that "artificial intelligence is the next stage in evolution", an idea first proposed by Samuel Butler's "Darwin among the Machines" as far back as 1863, and expanded upon by George Dyson in his book of the same name in 1998.[376]

Economics

The long-term economic effects of AI are uncertain. A survey of economists showed disagreement

about whether the increasing use of robots and AI will cause a substantial increase in long-term unemployment, but they generally agree that it could be a net benefit, if productivity gains are redistributed.[377]

In fiction

Main article: Artificial intelligence in fiction

The word "robot" itself was coined by Karel ?apek in his 1921 play R.U.R., the title standing for "Rossum's Universal Robots"

Thought-capable artificial beings appeared as storytelling devices since antiquity,[24] and have been a persistent theme in science fiction.

A common trope in these works began with Mary Shelley's Frankenstein, where a human creation becomes a threat to its masters. This includes such works as Arthur C. Clarke's and Stanley Kubrick's 2001: A Space Odyssey (both 1968), with HAL 9000, the murderous computer in charge of the Discovery One spaceship, as well as The Terminator (1984) and The Matrix (1999). In contrast, the rare loyal robots such as Gort from The Day the Earth Stood Still (1951) and Bishop from Aliens (1986) are less prominent in popular culture.[378]

Isaac Asimov introduced the Three Laws of Robotics in many books and stories, most notably the "Multivac" series about a super-intelligent computer of the same name. Asimov's laws are often brought up during lay discussions of machine ethics;[379] while almost all artificial intelligence researchers are familiar with Asimov's laws through popular culture, they generally consider the laws useless for many reasons, one of which is their ambiguity.[380]

Transhumanism (the merging of humans and machines) is explored in the manga Ghost in the Shell and the science-fiction series Dune. In the 1980s, artist Hajime Sorayama's Sexy Robots series were painted and published in Japan depicting the actual organic human form with lifelike muscular metallic skins and later "the Gynoids" book followed that was used by or influenced movie makers including George Lucas and other creatives. Sorayama never considered these organic robots to be real part of nature but always unnatural product of the human mind, a fantasy existing in the mind

even when realized in actual form.

Several works use AI to force us to confront the fundamental of question of what makes us human, showing us artificial beings that have the ability to feel, and thus to suffer. This appears in Karel ?apek's "R.U.R.", the films "A.I. Artificial Intelligence" and "Ex Machina", as well as the novel Do Androids Dream of Electric Sheep?, by Philip K. Dick. Dick considers the idea that our understanding of human subjectivity is altered by technology created with artificial intelligence.[381]