Artificial intelligence (**AI**) is the capability of computational systems to perform tasks typically associated with human intelligence, such as learning, reasoning, problem-solving, perception, and decision-making. It is a field of research in computer science that develops and studies methods and software that enable machines to perceive their environment and use learning and intelligence to take actions that maximize their chances of achieving defined goals.[1]

High-profile applications of AI include advanced web search engines (e.g., Google Search); recommendation systems (used by YouTube, Amazon, and Netflix); virtual assistants (e.g., Google Assistant, Siri, and Alexa); autonomous vehicles (e.g., Waymo); generative and creative tools (e.g., language models and AI art); and superhuman play and analysis in strategy games (e.g., chess and Go). However, many AI applications are not perceived as AI: "A lot of cutting edge AI has filtered into general applications, often without being called AI because once something becomes useful enough and common enough it's not labeled AI anymore."[2][3]

Various subfields of AI research are centered around particular goals and the use of particular tools. The traditional goals of AI research include learning, reasoning, knowledge representation, planning, natural language processing, perception, and support for robotics.[a] To reach these goals, AI researchers have adapted and integrated a wide range of techniques, including search and mathematical optimization, formal logic, artificial neural networks, and methods based on statistics, operations research, and economics.[b] AI also draws upon psychology, linguistics, philosophy, neuroscience, and other fields.[4] Some companies, such as OpenAI, Google DeepMind and Meta,[5] aim to create artificial general intelligence (AGI)—AI that can complete virtually any cognitive task at least as well as a human.

Artificial intelligence was founded as an academic discipline in 1956,[6] and the field went through multiple cycles of optimism throughout its history,[7][8] followed by periods of disappointment and loss of funding, known as AI winters.[9][10] Funding and interest vastly increased after 2012 when graphics processing units started being used to accelerate neural networks and deep learning outperformed previous AI techniques.[11] This growth accelerated further after 2017 with the transformer architecture.[12] In the 2020s, an ongoing period of rapid progress in advanced generative AI became known as the AI boom. Generative AI's ability to create and modify content has led to several unintended consequences and harms, while raising ethical concerns about AI's long-term effects and potential existential risks, prompting discussions about regulatory policies to ensure the safety and benefits of the technology.

Goals

The general problem of simulating (or creating) intelligence has been broken into subproblems. These consist of particular traits or capabilities that researchers expect an intelligent system to display. The traits described below have received the most attention and cover the scope of AI research.[a]

Reasoning and problem-solving

Early researchers developed algorithms that imitated step-by-step reasoning that humans use when they solve puzzles or make logical deductions.[13] By the late 1980s and 1990s, methods were developed for dealing with uncertain or incomplete information, employing concepts from probability and economics.[14]

Many of these algorithms are insufficient for solving large reasoning problems because they experience a "combinatorial explosion": They become exponentially slower as the problems grow.[15] Even humans rarely use the step-by-step deduction that early AI research could model. They solve most of their problems using fast, intuitive judgments.[16] Accurate and efficient reasoning is an unsolved problem.

Planning and decision-making

An "agent" is anything that perceives and takes actions in the world. A rational agent has goals or preferences and takes actions to make them happen.[d][32] In automated planning, the agent has a specific goal.[33] In automated decision-making, the agent has preferences—there are some situations it would prefer to be in, and some situations it is trying to avoid. The decision-making agent assigns a number to each situation (called the "utility") that measures how much the agent prefers it. For each possible action, it can calculate the "expected utility": the utility of all possible outcomes of the action, weighted by the probability that the outcome will occur. It can then choose the action with the maximum expected utility.[34]

In classical planning, the agent knows exactly what the effect of any action will be.[35] In most real-world problems, however, the agent may not be certain about the situation they are in (it is "unknown" or "unobservable") and it may not know for certain what will happen after each possible action (it is not "deterministic"). It must choose an action by making a probabilistic guess and then reassess the situation to see if the action worked.[36]

In some problems, the agent's preferences may be uncertain, especially if there are other agents or humans involved. These can be learned (e.g., with inverse reinforcement learning), or the agent can seek information to improve its preferences.[37] Information value theory can be used to weigh the value of exploratory or experimental actions.[38] The space of possible future actions and situations is typically intractably large, so the agents must take actions and evaluate situations while being uncertain of what the outcome will be.

A Markov decision process has a transition model that describes the probability that a particular action will change the state in a particular way and a reward function that supplies the utility of each state and the cost of each action. A policy associates a decision with each possible state. The policy could be calculated (e.g., by iteration), be heuristic, or it can be learned.[39]

Game theory describes the rational behavior of multiple interacting agents and is used in AI programs that make decisions that involve other agents.[40]

Natural language processing

Natural language processing (NLP) allows programs to read, write and communicate in human languages.[50] Specific problems include speech recognition, speech synthesis, machine translation, information extraction, information retrieval and question answering.[51]

Early work, based on Noam Chomsky's generative grammar and semantic networks, had difficulty with word-sense disambiguation[f] unless restricted to small domains called "micro-worlds" (due to the common sense knowledge problem[29]). Margaret Masterman believed that it was meaning and not grammar that was the key to understanding languages, and that thesauri and not dictionaries should be the basis of computational language structure.

Modern deep learning techniques for NLP include word embedding (representing words, typically as vectors encoding their meaning),[52] transformers (a deep learning architecture using an attention mechanism),[53] and others.[54] In 2019, generative pre-trained transformer (or "GPT") language models began to generate coherent text,[55][56] and by 2023, these models were able to get human-level scores on the bar exam, SAT test, GRE test, and many other real-world applications.[57]

Perception

Machine perception is the ability to use input from sensors (such as cameras, microphones, wireless signals, active lidar, sonar, radar, and tactile sensors) to deduce aspects of the world. Computer vision is the ability to analyze visual input.[58]

The field includes speech recognition,[59] image classification,[60] facial recognition, object recognition,[61] object tracking,[62] and robotic perception.[63]

Techniques

AI research uses a wide variety of techniques to accomplish the goals above.[b]

Search and optimization

AI can solve many problems by intelligently searching through many possible solutions.[69] There are two very different kinds of search used in AI: state space search and local search.

State space search

State space search searches through a tree of possible states to try to find a goal state.[70] For example, planning algorithms search through trees of goals and subgoals, attempting to find a path to a target goal, a process called means-ends analysis.[71]

Simple exhaustive searches[72] 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.[15] "Heuristics" or "rules of thumb" can help prioritize choices that are more likely to reach a goal.[73]

Adversarial search is used for game-playing programs, such as chess or Go. It searches through a tree of possible moves and countermoves, looking for a winning position.[74]