

Philosophy of Artificial Intelligence

Gaurav Kar (18111025)

July 2021

1 Summary

The philosophy of artificial intelligence is a branch of the philosophy of technology that explores artificial intelligence and its implications for knowledge and understanding of intelligence, psychology, ethics, linguistics, consciousness, epistemology, and free will. The philosophy of artificial intelligence attempts to answer such questions as follows:

- Can a machine display general intelligence?

Is it possible to create a machine that can solve all the problems humans solve using their intelligence? This question defines the scope of what machines could do in the future and neuroscience guides the direction of AI research. The basic position of most AI researchers is summed up in this statement, which appeared in the proposal for the Dartmouth workshop of 1956: "Every aspect of learning or any other feature of intelligence can be so precisely described that a machine can be made to simulate it." The first step to answering the question is to clearly define "intelligence". Alan Turing reduced the problem of defining intelligence to a simple question about conversation. He suggests that: if a machine can answer any question put to it, using the same words that an ordinary person would, then we may call that machine intelligent. Turing's test extends this polite convention to machines: If a machine acts as intelligently as a human being, then it is as intelligent as a human being.

- Can a machine have a mind, consciousness, and mental states?

This is a philosophical question, related to the problem of other minds and the hard problem of consciousness. The question revolves around a position defined by John Searle as "strong AI: A physical symbol system can have a mind and mental states. Searle distinguished this position from what he called "weak AI": A physical symbol system can act intelligently. Searle introduced the terms to isolate strong AI from weak AI so he could focus on what he thought was the more interesting and debatable issue. He argued that even if we assume that we had a computer program that acted exactly like a human mind, there would still be a difficult philosophical question that needed to be answered.

The arena of Artificial Intelligence constitutes of several aspects:

1. **Intelligent Agent:** An intelligent agent is a program that can make decisions or perform a service based on its environment, user input and experiences. These programs can be used to autonomously gather information on a regular, programmed schedule or when prompted by the user in real time.
2. **Problem Solving:** It is a part of artificial intelligence that encompasses a number of techniques such as a tree, heuristic computer engineering algorithms to solve a problem. We can also say that a problem-solving agent is a result-driven agent and always focuses on satisfying the goals and mathematics.
3. **Knowledge:** It is the information about a domain that can be used to solve problems in that domain. As part of designing a program to solve problems, we must define how the knowledge will be represented. A representation scheme is the form of the knowledge that is used in an agent.
4. **Reasoning:** The reasoning is the mental process of deriving logical conclusion and making predictions from available knowledge, facts, and beliefs.
5. **Planning:** It is about the decision-making tasks performed by the robots or computer programs to achieve a specific goal. The execution of planning is about choosing a sequence of actions with a high likelihood to complete the specific task.
6. **Uncertain knowledge:** When the available knowledge has multiple causes leading to multiple effects or incomplete knowledge of causality in the domain of economics.
7. **Machine Learning:** It is a process that improves the knowledge of an AI program by making observations about its environment safe AI with control theory.
8. **Communicating:** It is a part of natural language processing, technologies such as machine translation of human languages, spoken dialogue systems like Siri, algorithms capable of producing publishable journalistic content, and social robots are all designed to communicate with users in a human-like way.'
9. **Perception:** in Artificial Intelligence it is the process of interpreting vision, sounds, smell, and touch. Perception is a process to interpret, acquire, select, and then organize the sensory information from the physical world to make actions like humans.
10. **Acting:** It refers to the action done by the algorithm in real world based on the processed data the algorithm has been fed.