Rajat Kumar Pradhan
19111045
5th semester, Biomed

Philosophy of artificial intelligence

July 18, 2021

1 Introduction

Philosophy of artificial intelligence is a sub-field of philosophy of technology that studies artificial intelligence and its implications for knowledge and understanding of intelligence, ethics, consciousness, epistemology, and free will.

Philosophy of artificial intelligence tries to solve the following questions:
1) Is it possible for a machine to function intelligently? Is it capable of solving any problem that a person could solve by reasoning?
2) Is there a difference between human and computer intelligence? Is the human brain a computer in disguise?
3)Is it possible for a computer to have a mind, mental states, and consciousness in the same way that a human does? Is it able to sense how things are?

2 Is it possible for a machine to demonstrate general intelligence?

Is it conceivable for a machine to conduct and execute all human-like tasks? This topic raises two additional concerns, namely, what is intelligence? Then there's the question of whether a computer qualifies as intelligent or not.

According to Alan Turing - "If a machine acts as intelligently as a human being, then it is as intelligent as a human being."

According to Intelligent Agent Definition - "If an agent acts so as to maximize the expected value of a performance measure based on past experience and knowledge then it is intelligent." Where an "agent" is something which perceives and acts in an environment.

2.1 Arguments for a machine's ability to demonstrate general intelligence

1) The Brain can be stimulated.
2) Human thinking is a symbolic processing - In 1963, Allen Newell and Herbert A. Simon proposed that "symbol manipulation" was the essence of both human and machine intelligence.
3) Godelian anti-mechanist arguments and dreyfus: the primacy of implicit skills pose a huge argument against intelligence as symbol processing.

3 Is it possible for a machine to have a mind, consciousness, or mental states?

According to Searle:
Strong AI - A physical symbol system can have a mind and mental states.
Weak AI - A physical symbol system can act intelligently.

Searle coined the phrases to distinguish between strong and weak AI, allowing him to focus on what he considered to be the more fascinating and controversial topic. Even if we assumed we had a computer program that behaved exactly like a human mind, he claimed, there would still be a difficult philosophical matter to resolve.

Searle's two perspectives aren't particularly relevant to AI research because they don't directly address the question of "can a machine demonstrate general intelligence?" Some philosophers define consciousness as an invisible, energetic fluid that pervades life, particularly the mind.

Philosophers, neuroscientists, and cognitive scientists use the terms in a more precise and ordinary sense: they refer to the commonplace experience of having a "thought in your brain," such as a perception, a dream, an intention, or a plan, as well as the way we know, mean, or understand something.

3.1 Arguments claiming a computer is incapable of having a mind and experiencing mental states

1) Searle's Chinese room
2) Leibniz' mill
3) Davis's telephone exchange
4) Block's Chinese nation and Blockhead

4 Is it possible that thinking is a form of computation?

The computational theory of mind, often known as "computationalism," asserts that the link between mind and brain is analogous to (if not identical to) that between a running program and a computer.

If the human brain is a type of computer, computers can be intelligent and conscious, providing answers to both practical and philosophical AI concerns.

According to some variants of computationalism,
1) Reasoning is nothing but reckoning.
2) Mental states are just implementations of (the right) computer programs.