Philosophy Of Artificial Intelligence

Sachin Malve
19111049
5th Semester
Biomedical Engineering

July 19, 2021

---

1 Introduction

The philosophy of AI dwells on philosophical discussion like whether AI should be experimented on Animals, Do machines and Humans have same intelligence, Can machines act intelligently and Can machines have a mind.
Some important propositions in philosophy of AI :-
Turing's "polite convention": If a machine behaves as intelligently as a human being, then it is as intelligent as a human being.
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.
Allen Newell and Herbert A. Simon's physical symbol system hypothes: "A physical symbol system has the necessary and sufficient means of general intelligent action.
John Searle's strong AI hypothesis: "The appropriately programmed computer with the right inputs and outputs would thereby have a mind in exactly the same sense human beings have minds."
Hobbes' mechanism: "For 'reason' ... is nothing but 'reckoning,' that is adding and subtracting, of the consequences of general names agreed upon for the 'marking' and 'signifying' of our thoughts.

2 Can a machine display general intelligence ?

Argument against the premise is that machines have a limit and they lack special part of human mind responsible for intelligence and it cannot be replicated. However the Turing Child Machine proposal sidesteps the need for precise design-time description all together.

2.1 Intelligence

Turing Test :- Turing reduced intelligence into simple questions about conversations, If the machine can answer any question into words used by ordinary person it is an intelligent as human. However there is criticism that just only if the machine uses human words doesn't mean it is as intelligent as human.

2.2 Intelligent agent

It is defined as if the machine can maximize the expected value by the help of past experiences and knowledge than it is intelligent.

2.3 Arguments that a machine can display general intelligence

There are arguments that machine would be able to simulate everything by year 2029 and we can simulate everything we will reach a breaking point. But critics argue that merely mimicking the functioning of a brain would in itself be an admission of ignorance regarding intelligence and the nature of the mind.
Some say that human thinking is a symbol processing and that A physical symbol system has the necessary and sufficient means of general intelligent action. However there are arguments that human thinking is not solely symbol processing and that even modern day AI uses mathematics and statistics for optimization and not only symbols.

2.4 Gödel's Incompleteness Theorem

Gödel surmised that that the human mind can correctly eventually determine the truth or falsity of any well-grounded mathematical statement (including any possible Gödel statement), and that therefore the human mind's power is not reducible to a mechanism. There are many critics of this theorem stating reason like human do not show any mathematical reasoning capabilities that are beyond machines and that human mind is very inconsistent and that the argument only applies to what can be theoretically proved .

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

Searle differentiated AI into Strong and Weak AI :-
Strong AI- A physical symbol that can have mind and mental states .
Weak AI – A physical system that can act intelligently.
However these statement do not answer our question of can machine have a mind ?

3.1 Consciousness, Mind and Mental States

Consciousness can be defined ranging from energetic fluid that permeates life to self awareness to something we know on a daily basis.

3.2 Arguments that Computer do not have mind

Searles Chinese Room - The Chinese room argument holds that a digital computer executing a program cannot have a "mind", "understanding" or "consciousness", regardless of how intelligently or human-like the program may make the computer behave.
Leibniz' mill, Davis's telephone exchange, Block's Chinese nation and Blockhead

4 Is thinking a kind of computation?

The computational theory of mind or "computationalism" claims that the relationship between mind and brain is similar.
our intelligence derives from a form of calculation, similar to arithmetic. This is the physical symbol system hypothesis discussed above, and it implies that artificial intelligence is possible Mental states are just implementations of (the right) computer programs.

4.1 There are other question which are debated upon on and they are necessary to be pondered.

Can a machine have emotions?
Can a machine be self-aware?
Can a machine be original or creative?
Can a machine imitate all human characteristics?
Can a machine have a soul?

4.2 Can a machine be benevolent or hostile?

There is the question of what will happen when the machines will achieve autonomy because it is argued machines will have millions times more intelligence in just a few years. Some question about the use of robots in military use because those robots are given a autonomy to some extent and not enough attention to there implications.

5 Views on the role of philosophy

Some scholars argue that the AI community's dismissal of philosophy is detrimental. Physicist David Deutsch argues that without an understanding of philosophy or its concepts, AI development would suffer from a lack of progress.

Important keywords :-

- Algorithm
- Artificial intelligence
- Cognitive computing:
- Consciousness
- Computation
- Deep learning
- Ethics
- Intelligent agents
- Machine Learning
- Unsupervised learning: