

Summary on Philosophy of Artificial Intelligence

Aditi Bhattacharya

17 July 2021

1 Introduction

This article aims to understand the whole journey of AI from a hypothesis to its present state, where does it aim to go discussing along with its pros and cons which affect the entire universe. The philosophy of AI finds its roots to answer questions as follows:

- Can a machine ever be referred to as intelligent? If so, would it solve a given problem using the same protocols as a human does?
- Can we ever develop an algorithm which works exactly as the human mind? If so, then would the human mind be also referred as a computer (with just a lot more computations and networks)?
- If we are able to mimic the human mind exactly, would the machine also have mental states, a mind of its own, and a sense of conscious experiencing as subjective as that of a human? If so, how do we test whether its conscious for real, and just not an algorithm which is designed to just answer and not feel anything.

Questions of these sort, allures the scientific community from all possible realms, from mathematicians, physicists, neuro-biologists to businessmen, who want to use this strong-yet-unpredictable technology to enhance the quality of living.

2 Machine displaying General Intelligence

To work on this realm of intelligence, it is very important to first define intelligence quantitatively without paying much focus on whether a machine is really thinking (as a person thinks) or is just acting like it is thinking.

For this, we need to define how an intelligent agent might work like and when would it be called intelligent. The Turing test helps us to pass through this test. Machine learning Algorithms and ANN nowadays, help us to solve many complicated problems without us being actually able to infer how it works. It is typically an example of black box.

Other arguments support that if we are able to mimic the working of the brain to the best possible resolution and then simulate it in a virtual environment, we might end up generating a general intelligence in a machine. This approach is hugely being researched as more and more technologies are helping us to study the brain with greater precision.

Symbol processing is an approach which takes into account that the human brain processes a stimulus through symbolic representations and to mimic this, a system needs to be designed which looks for patterns in the input rather than going for the object as a whole. Arguments against symbolic processing states that these arguments show that human thinking does not consist (solely) of high level symbol manipulation. They do not show that artificial intelligence is impossible, only that more than symbol processing is required.

Godel through his incompleteness theorem conjectured that the human mind is an exception to his theorem and is hence not reducible to a mechanistic system. A Turing machine would always fall prey too the incompleteness theorem and hence it becomes theoretically impossible to capture a human mind into a machine. However, later on many scientists have shown that human reasoning is also inconsistent and hence the above mentioned theory cannot be taken into account.

Hubert Dreyfus argued that human intelligence and expertise depended primarily on implicit skill rather than explicit symbolic manipulation, and argued that these skills would never be captured in formal rules.

3 Machine having mind, consciousness and displaying Mental States

This philosophical question brings along with the 'Hard Problem of Consciousness' and refers a future with 'Strong AI'. A strong AI apart from being intelligent would also possess the essence of experiencing qualia. It would not just know, but have the feeling of knowing as well. We are still on the very initial stage in this realm, and an immense research is required here to understand how might a biological system develop consciousness, and if so, how can it be proved theoretically that it's exhibiting such a phase. Various theories have been given by scientists which aim to take various approaches from quantum mechanics at play to quantifying consciousness into ϕ , none of them have yet crossed the hypothetical stage and given valuable proof.

Scientists like Searle and Lebiniz have hypothesized various thought experiments deducing that it would never be possible for machine to understand what's it processing, a power which only conscious beings possess. However, there are many people who beg to differ. It is a very interesting field to explore and brainstorm (because as of now, only beings with brains are asking this question).

4 Conclusion

To conclude, this branch of science helps a curious mind to quench its thirst of curiosity without ever running out of questions.