philosophy of AI

Submitted by Simran Meshram

July 2021

1 Introduction

Artificial intelligence philosophy is a subset of technology philosophy that investigates artificial intelligence and its implications for intelligence knowing and understanding of intelligence (described as the ability to perceive or infer information, and to retain it as knowledge) , ethics(seeks to resolve questions of human morality by defining concept such as good and evil, right and wrong , virtue and vice , justice and crime), consciousness(it is one's awareness of internal and external existence), epistemology(a branch of philosophy that deals with knowledge), and free will(an ability to choose between various possible outcomes). The philosophy of artificial intelligence attempts to answer such questions as follows: 1. Can a machine think? Is it capable of reasoning its way out of any problem? 2. Is there a distinction between human intellect and artificial intelligence? Is it possible that the human brain is a computer disguised as a human brain is a computer disguised as a human brain? 3. Can a computer have a mind, mental states, and consciousness in the same way as a human can? Is it capable of detecting changes in the environment? The scientific responses to these questions are based on the definitions of "intelligence" and "consciousness." Allen Newell and Herbert A. Simon's physical symbol system hypothesis: "A physical symbol system has the necessary and sufficient means of general intelligence action." John Scarle'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. All the above hypothesis are enough to prove that a machine can have a mind, consciousness and mental states (like understanding or perceiving) and ultimately, the experience of consciousness? It makes no difference whether a machine is genuinely thinking (as a human is) or simply acting like one. Most AI researchers core position can be summarized as follow: "Every facet of learning or any other feature of intelligence can be stated so accurately." Some of the arguments against AI's construction and operation claim that the system is impossible because the human mind possesses some unique property that is required for intelligent action but cannot be reproduced or stimulated by a computer. The term "intelligent agent" is then used. Intelligence is defined in AI research in terms of intelligent agents. A "performance measure" defines the performance of a "agent," which is a thing that observes and reacts to its surroundings. "An agent is intelligent

if it behaves to maximize the expected value of a performance measure based on past experience and knowledge." According to computationalism, the link between the mind and the brain is comparable (if not identical) to that between running program and a computer. This question relates to the previous ones: if the human brain is a type of computer, then computers can be intelligent and conscious, answering both the practical and philosophical questions of the previous question..In terms of the practical question of AI, can a machine display general intelligence? 1. To put it another way, intelligence is derived from a type of computation comparable to arithmetic. It also suggests that artificial intelligence is conceivable. Most variants of computationalism assert that a machine can have a mind, mental states, and consciousness? 2. Mental states are merely computer programmes in action That AI advancement would be hampered by a lack of knowledge of philosophy or its principles.