Philosophy of Artificial Intelligence  
Submitted by Vinit kumar  
28th July 2021

1 Introduction

Artificial intelligence is a branch of the philosophy of technology that explores  
artificial intelligence and its implications for knowledge and understanding of  
intelligence, ethics, consciousness, and free will. some scholars argue that the AI  
community’s dismissal of philosophy is detrimental. The philosophy of artificial  
intelligence attempts to answer such questions as follows: Can a machine act  
intelligently? CAN IT SOLVE ANY PROBLEM that a person would solve by  
thinking? Are human intelligence and machine intelligence the same? Is the  
human brain essentially a computer? Can machines have a mind, mental states,  
and consciousness in the same sense that a human being can? Can they feel  
how things are?

2 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 guides the directions of Ai research. To answer this ques-  
tion, it does not matter whether a machine is really thinking (a person thinks)  
or is just acting like it is thinking. Alan Turing reduced the problem of the  
defining intelligence to a simple question about conversation. He suggests that  
if a machine can answer any question using the same words that an ordinary  
person would, then we may call that machine intelligent. Modern version of his  
experimental design would use an online chart room, where one of the partic-  
ipants is a real person and one of them is a computer program. The program  
passes the test if no one can tell which of the two participants is human. The  
argument that a machine can display general intelligence was first introduced as  
early as 1943 and vividly described by Hans Moravec in 1988. Some critics of AI  
such as Hubert Dreyfus and John Searle argue that anything can be stimulated  
by a computer; thus, any process at all can technically be considered “computa-  
tion”. Modern AI, based on statistics and mathematical optimization, does not  
use the high level “symbol processing” that they discussed. G ̈odel’s theorems  
do not lead to any argument that humans have mathematical reasoning capa-  
bilities beyond what a machine could ever duplicate. The arguments had been

anticipated by Alan Turing in his 1950 paper computing machinery and intel-  
ligence. New models are being developed to capture our unconscious skills at  
perceptions and attention. These include neural nets, evolutionary algorithms  
and so on.

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

John Searle argued that even if we had a computer program that acted ex-  
actly like a human mind, there would still be a difficult philosophical question  
that needed to be answered. John Searle’s two positions do not directly an-  
swer the question “can a machine display general intelligence?” Science fiction  
writers use it to describe some essential property that makes us human: intelli-  
gence, desires, will, insight, pride and so on. John Searle asks us to consider a  
thought experiment: suppose we have written a computer program that passes  
the Turing test and demonstrates general intelligent action. He concludes that  
the Chinese room, or any other physical symbol system, cannot have a mind.  
Searle’s argument is just a version of the problem of other minds, applied to  
machines. The question is whether “consciousness” exist.

4 Is thinking a kind of computation?

Computationalism claims that the relationship between mind and brain is sim-  
ilar to that between a running program and a computer. If the human brain  
is a kind of computer then computers can be both intelligent and conscious,  
answering both the practical and philosophical questions of AI. The idea has  
philosophical roots in Hobbes and Leibniz, Hume and Kant. Other related ques-  
tions like Can a machine have emotions? Can a machine be self-aware? Can a  
machine be original or creative? Can a machine be benevolent or hostile? Can  
a machine imitate all human characteristics?

5 views on the role of philosophy

Some philosophers argue that the role of philosophy in AI is underappreciated.  
Physicist David Deutsch argues that without an understanding of philosophy  
or its concepts, AI development would suffer from a lack of progress. The  
Sandford Encyclopaedia of Philosophy argues that the AI community’s dismissal  
of philosophy is determined.