

Machine and Consciousness

Kyle C. Anulacion

San Francisco State University

Pt. 1: Can machines think? Does a computer have conscious thoughts?

In a 1946 lecture, popular French existentialist Jean-Paul Sartre explained one of his most popular intrigues — “Man is condemned to be free.” “Condemned,” Sartre says, “because he did not create himself, yet is nevertheless at liberty, and from the moment that he is thrown into this world he is responsible for everything he does.” Sartre was offering a cold analysis of the human condition by trying to illustrate a distinction between humans, and the rest of life that exists. It may be the case that plants and animals do in fact exist, but Sartre argues that they cannot reflect on their existence. Sartre makes an interesting, fundamental point about humanity. It could be said that the *being* of man is apart from the *being* of things, that conscious experience is unique for humans. What is it like to be a cat? Or a bird? Or a garden snail? Or a plant? Are all of them capable of thought?

How do we know whether a computer is conscious or not? Is it enough if a computer acts like a human to perfection, if its physical workings perfectly mimic human neurons, that it has *consciousness*, or *thinks*, or at least is *intelligent*? Alan Turing proposed that through an “imitation game” it can be said that a machine thinks, while John Searle argued that machines can be seen as intelligent but it does not necessarily give them thought.

English mathematician, scientist, and philosopher Alan Turing conceptualised the modern computer in his paper “On Computable Numbers, with an Application to the

Entscheidungsproblem” while he was studying for his Ph.D. at Princeton in 1936. His “Turing machines”, as they were called after, could solve any problem described on a strip of tape through a table of rules. The theory that one machine can do any number of computational tasks through instructions was revolutionary. It gave rise to the question of if machines can think or not.

After the war, Turing published his paper “Computing Machinery and Intelligence”, in which he pondered this question, but him being wary of the ambiguity of the terms “machine” and “think”, Turing devises a clearer question through a sort of test. He pictures an interrogator, who between two subjects, has to distinguish which one is the machine and which one is the man. The interrogator can only determine this through the questions he asks and the answers they send typewritten from the other side of a wall. Turing calls this the “imitation game”. He believed that a computer programmed cleverly enough could be indistinguishable from the man — and therefore the computer could be called intelligent.

Turing maintained that machines have consciousness if people were open to accept the imitation game, and that machines could replicate human error to perfect their mimicry of a human. He also said that machines aren’t any less capable of thought for not being kind, beautiful, moral, just, humourous, able to enjoy food, do something new, or any other example, because these examples are just as absurd and naive as “most English children deciding that everybody speaks English and that it is silly to learn French.” (Turing, 1950, 4)

Forty years after, a response paper to Turing's argument for AI was published by John Searle, a philosophy professor at the University of California, Berkeley, titled "Is the Brain's Mind a Computer Program?" Searle criticised researchers of artificial intelligence for their firm belief that "the mind is to the brain as the program is to hardware", that in essence machines *think* just by virtue of programming it, and that a viable way to see if they think is through Turing's imitation game. He believed that the human mind does not operate on the manipulation of formal symbols — that the mind is not programmed. Searle proves this through a thought experiment that affirms that human thought is apart from machine "thought", even more so affirming that machines do not think at all.

Searle asks us to imagine him sitting in a room with a rulebook that tells him what Chinese characters to write in response should he ever receive a message written in Chinese. Searle and the rulebook discriminates the symbols based on their shape and nothing else. Outside the room, those who understand Chinese hand in symbols that ask a question, to which he can flip through the rulebook, find the symbols that give the appropriate response, pick out the symbols he needs from baskets of symbols that surround him, and give symbols back that mean a well-written answer. Without a doubt Searle does not understand Chinese, but lo, he can respond to those outside the room in fluent Chinese. He passes the Turing test only because he *seems* to speak Chinese, he *seems* to be intelligent. "There is no way I could come to understand Chinese in the system as described, since there is no way that I can learn the meanings of any of the symbols," Searle writes:

Like a computer, I manipulate symbols, but I attach no meaning to the symbols...Digital computers merely manipulate formal symbols according to rules in the program...Just manipulating the symbols is not by itself enough to guarantee cognition, perception, understanding, thinking and so forth. And since computers, qua computers, are symbol-manipulating devices, merely running the computer program is not enough to guarantee cognition. This simple argument is decisive against the claims of strong AI.

In conclusion, Alan Turing believed that machines are capable of intelligence and consciousness, and that this can be proven through a sort of “imitation game” that he devised. He held unbridled optimism for machines, predicting that in half a century a machine will be made somewhat indistinguishable from man. John Searle believed that machines are not capable of thought because they only *imitate* intelligence. He illustrates that computers in essence manipulate symbols meaningless to them — something very different from human minds — through a thought experiment called the “Chinese room” argument. Who is right? The decision only leads to more questions about the machine and consciousness. What exactly is meant by “artificial intelligence”? Can a machine acknowledge its own existence? Can a machine make ethical choices? And will machines supercede, or evolve humanity?

Pt. 2: An opinion on machine and consciousness

The advent of artificial intelligence provoked much thought to what consciousness actually is. It is simple enough to question consciousness between humans and other animals because they fall under similar anatomical distinctions, that they behave and appear in a way that is innate and familiar to us, yet machines, with “minds” born out of silicon, gold, and copper, are an entirely different gulf. It is without a doubt that the machines of today do not think. “Artificial intelligence” is a quality of computers, but the term itself does not imply that they are capable of thought. Their intelligence is artificial because their proficiency in X is the result of their formal programming, that is to say a computer’s heuristic ability is dissimilar to a human’s heuristic ability, wherein human experience is not based on ones and zeroes, but rather it is semantic.

It is erroneous to say that a machine that has intelligence (because it passes Turing’s imitation game) has a consciousness. For a machine to have conscious, they must be aware of their own environment. This is done to an extent, where they may have cooling fans when their internal hardware gets too hot, or robots react to certain stimuli, but we wouldn’t say that they are conscious. It is not enough to demonstrate self-concept, a quality that some animals outside humans, like elephants, apes, and dolphins do. A machine cannot acknowledge its own existence. This is because a computer can exist, and it can refer to itself and refer to others, but it does not think about what this implies. John Searle’s “Chinese room” argument illustrates this premise. It is too essential that the “being” of computers are apart from the being of man.

Does a computer making ethical choices show that it thinks? Without a doubt, no. A machine can make ethical choices, but it can only do so because it is dependent on its programming and is therefore incapable of saying what is right or wrong. A computer by nature is only sufficient for the execution of tasks. In other words, it is not the *computer* that does the thinking, but rather it is the *program*. To illustrate this, a computer does not reject its programming because its code is unethical. If this is the case, this must only mean that there is an underlying program in the computer that defines ethics for the computer. A computer in essence is an electronic machine that carries out its task, and no less.

The computer is an indispensable tool to mankind today. The computers of today have the power to perform millions of complex mathematical calculations seemingly instantly, manage money, finance, global commerce, and multi-billion dollar industries, play a variety of games proficiently, simulate complex 3-dimensional structures and models, and other functions too numerous to mention. In its sense of utility, computers do not harbor the quality of “human error” (however one may code a program to replicate it). To draw a conclusion from if machines can make ethical choices or not, to be artificially intelligent is to be artifactually obedient. Unless a computer is directly programmed to, it will not take the “life” of another computer for any reason. Neither will it be benevolent to others. Computers cannot surpass humans in this way, for it does not “desire” nor has the innate capacity to be self-sustaining or self-creating.

In short, there are lots of things that computers can do better than humans because of clever

programming. But computers, incapable of reasoning, altogether evolve because of their dependency to man. Man does the thinking for the computers — it is simple enough that machines do man's bidding by virtue of its own ignorance. Humanity finds its limits in computers because the human brain is bumbling, inexact, social, and aware of its own morality. The computer isn't any of this. And it may be possible that computers can mimic evolution, but it can only do so by mortal impetus. It can thus be said that if it were the case that computers surpass humanity, this is only because man had created his own end. Man is condemned to be free. However, machine is condemned is to chains.

References Used

Sartre, Jean-Paul. "Existentialism is humanism." The Sorbonne, 1946, Paris, France. University lecture. Translated by Philip Mairet (1948),
<https://wmpeople.wm.edu/asset/index/cvance/sartre>

A.M. Turing (1950) Computing Machinery and Intelligence. In S.M. Cahn (Ed.), *Exploring Philosophy: An introductory anthology*. Oxford, New York: Oxford University Press

John Searle (1990) Is the Brain's Mind a Computer Program?. In S.M. Cahn (Ed.), *Exploring Philosophy: An introductory anthology*. Oxford, New York as "Can Computers Think?": Oxford University Press