

Will artificial intelligence become conscious?

December 8 2017, by Subhash Kak



Credit: Mikhail Nilov from Pexels

Forget about today's modest incremental advances in artificial intelligence, such as the increasing abilities of cars to drive themselves. Waiting in the wings might be a groundbreaking development: a machine that is aware of itself and its surroundings, and that could take in and process massive amounts of data in real time. It could be sent on

dangerous missions, into space or combat. In addition to driving people around, it might be able to cook, clean, do laundry – and even keep humans company when other people aren't nearby.

A particularly advanced set of machines could replace humans at literally all jobs. That would save humanity from workaday drudgery, but it would also shake many societal foundations. A life of no work and only play may turn out to be a [dystopia](#).

Conscious machines would also raise troubling legal and ethical problems. Would a conscious machine be a "person" under law and be liable if its actions hurt someone, or if something goes wrong? To think of a more frightening scenario, might these machines rebel against humans and wish to eliminate us altogether? If yes, they represent the culmination of evolution.

As a professor of electrical engineering and computer science who works in machine learning and [quantum theory](#), I can say that researchers are divided on whether these sorts of hyperaware machines will ever exist. There's also debate about whether machines could or should be called "conscious" in the way we think of humans, and even some animals, as conscious. Some of the questions have to do with technology; others have to do with what consciousness actually is.

Is awareness enough?

Most computer scientists think that consciousness is a characteristic that will emerge as technology develops. Some believe that consciousness involves accepting new information, storing and retrieving old information and cognitive processing of it all into [perceptions and actions](#). If that's right, then one day machines will indeed be the ultimate [consciousness](#). They'll be able to gather more information than a human, store more than many libraries, access vast databases in milliseconds and

compute all of it into decisions more complex, and yet more logical, [than any person ever could](#).

On the other hand, there are physicists and philosophers who say there's something more about human behavior that cannot be computed by a machine. Creativity, for example, and the sense of freedom people possess don't appear to come from logic or [calculations](#).

Yet these are not the only views of what consciousness is, or whether machines could ever achieve it.

Quantum views

Another viewpoint on consciousness comes from quantum theory, which is the deepest theory of physics. According to the orthodox [Copenhagen Interpretation](#), consciousness and the physical world are complementary aspects of the same reality. When a person observes, or experiments on, some aspect of the physical world, that person's [conscious interaction causes discernible change](#). Since it takes consciousness as a given and no attempt is made to [derive it from physics](#), the Copenhagen Interpretation may be called the "big-C" view of consciousness, where it is a thing that exists by itself – although it requires brains to become real. This view was popular with the pioneers of quantum theory such as Niels Bohr, Werner Heisenberg and Erwin Schrödinger.

The interaction between consciousness and matter leads to paradoxes that remain unresolved after 80 years of debate. A well-known example of this is the paradox of [Schrödinger's cat](#), in which a cat is placed in a situation that results in it being equally likely to survive or die – and the act of observation itself [is what makes the outcome certain](#).

The opposing view is that consciousness emerges from biology, just as biology itself emerges from chemistry which, in turn, [emerges from](#)

[physics](#). We call this less expansive concept of consciousness "little-C." It agrees with the neuroscientists' view that the [processes of the mind are identical to states and processes of the brain](#). It also agrees with a more recent interpretation of quantum theory motivated by an attempt to rid it of paradoxes, the [Many Worlds Interpretation](#), in which observers are a part of the mathematics of physics.

Philosophers of science believe that these modern quantum physics views of consciousness have parallels in ancient philosophy. Big-C is like the theory of mind in Vedanta – in which consciousness is the [fundamental basis of reality](#), on [par with the physical universe](#).

Little-C, in contrast, is quite similar to Buddhism. Although the Buddha chose not to address the question of the nature of consciousness, his followers declared that mind and consciousness arise out of emptiness or [nothingness](#).

Big-C and scientific discovery

Scientists are also exploring whether consciousness is always a computational process. Some scholars have argued that the creative moment is not at the end of a deliberate computation. For instance, dreams or visions are supposed to have [inspired Elias Howe's](#) 1845 design of the modern sewing machine, and August Kekulé's discovery of the [structure of benzene in 1862](#).

A dramatic piece of evidence in favor of big-C consciousness existing all on its own is the life of self-taught Indian mathematician [Srinivasa Ramanujan](#), who died in 1920 at the age of 32. His notebook, which was lost and forgotten for about 50 years and [published only in 1988](#), contains several thousand formulas, without proof in different areas of mathematics, that were well ahead of their time. Furthermore, the methods by which he found the [formulas remain elusive](#). He himself

claimed that they were revealed to him by a goddess [while he was asleep](#).

The concept of big-C consciousness raises the questions of how it is related to matter, and how matter and mind mutually influence each other. Consciousness alone cannot make physical changes to the world, but perhaps it can change the probabilities in the evolution of quantum processes. The [act of observation can freeze](#) and even influence atoms' movements, as Cornell physicists [proved in 2015](#). This may very well be an explanation of how matter and mind interact.

Mind and self-organizing systems

It is possible that the phenomenon of [consciousness](#) requires a self-organizing system, like the brain's physical structure. If so, then current machines will come up short.

Scholars don't know if adaptive self-organizing machines can be designed to be as sophisticated as the human brain; we lack a mathematical theory of computation for systems like that. Perhaps it's true that only biological [machines](#) can be sufficiently creative and flexible. But then that suggests people should – or soon will – start working on engineering new biological structures that are, or could become, conscious.

This article was originally published on [The Conversation](#). Read the [original article](#).

Provided by The Conversation

Citation: Will artificial intelligence become conscious? (2017, December 8) retrieved 2 May 2024 from <https://techxplore.com/news/2017-12-artificial-intelligence-conscious.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.