
Will machines ever become conscious?

Author: Barry Morisse

The nature of consciousness is something that has been debated for centuries by philosophers, neuroscientists, ethicists, biologists and the like, with very little progress having been made. We still know so little about what is undoubtedly the most important component of our human experience. The complexity still eludes our clumsy grasp.

The human brain, an incredible piece of matter that has been evolving over millions of years, continues to flummox us. Every new advancement in neuroscience tends to uncover a whole bunch of new questions we don't have the answers to. I would argue that we are only just starting to gain a true appreciation for what an amazing piece of evolution it is as we discover more and more of what we didn't know we didn't know—consciousness being one of those unknowns.

What is it? Where does it reside? Is it an illusion?

It's a debate that has become even more topical in recent times as the rapid advancement of computing capability, coupled with tangible signs of rudimentary artificial intelligence (AI) gives us a glimpse into what an artificial reality could look like. As Hollywood science fiction descends upon reality, it begs the question of whether we can recreate consciousness in an artificial being or not. Popular media is filled with stories of robots becoming conscious and then immediately deciding to destroy the human race – a trend that, while appealing for box office sales, doesn't quite do justice to the real deep questions which come up as a result of this kind of advancement. The huge leaps made within AI may be the best chance we have to discover more about consciousness itself.

Wading into this debate is challenging for two reasons. Firstly, because of how much we don't know – it is almost impossible to lay down a premise that can be defended in an intellectually honest way. Secondly, the limitations of language are dramatically exposed when trying to articulate the essence of what it means to be conscious; it is such a deeply personal experience. So, keeping these in mind, I'm going to take the approach of delving into the two polar extremes on both sides of the debate and leaving the wide spectrum in between open for you, as the reader, to slot yourself into.

The one side of the spectrum belongs to those who believe consciousness is a unique ethereal experience that has been given to us as living creatures as a result of evolution or from a higher power of some sort. It is often termed as a 'soul' in modern parlance – a term heavily influenced by religious tradition. This form of consciousness could almost be described as a separate entity entirely – being the driver of the body and brain, the thinker of thoughts. The CPU behind our eyes, if you wish. By its very nature, the premise here is that consciousness is something that has been endowed upon us and it would thus be impossible to create in a machine. Philosophically, this view is upheld by humanists and rationalists who argue that this self-awareness is unique and should be regarded as such.

The polar opposite side of the spectrum houses those who believe that consciousness is merely a very complex, interwoven mental illusion created by the firing off of millions of neurons and protons within our brains. They argue that what we see and experience as consciousness is nothing more than an evolutionarily-created response to the information we receive from the world around us through our senses. One way to try and articulate this is to imagine that everything we experience is data ingested by our nervous system and translated into a language which our eyes, ears and body can understand to better allow us to survive in the world around us. Advocates of this view suggest there is nothing magical about consciousness and that it is simply a computational problem we haven't solved yet. If consciousness is merely information processing, then the race is on to build a machine that is complex enough to simulate a human brain and, in so doing, we will create consciousness.

Now, those two extremes are both controversial and complicated in their own ways and with such limited information, I don't want to leap from the fence just yet. Consciousness is incredibly exciting because it is the ultimate human

connection that remains tantalisingly out of reach. And regardless of where you sit on the spectrum, what will become perhaps more important and intellectually constructive in the short term is to think about how we would know if a machine is conscious. How would we be sure?

There's a famous philosophical axiom from Descartes that suggests that the only thing I can know for sure is that I exist. This idea, now termed as solipsism, permeates this conversation in very interesting ways. When I look at another person, it sure seems like they are conscious because they exhibit similar body language, emotions and actions to me. For all intents and purposes, I have no reason to question it. Every instinct I have tells me that they are conscious. But philosophically, I can't really know for sure because I can't get inside of their head. It is possible that what I see is merely the illusion of consciousness – one that I take for granted.

This doubt is documented in various thought experiments (such as simulation theory), but also has come to the fore in technology as virtual reality continues to get better and better. As our technology improves, it becomes easier to fool the human brain, whether it be through editing images/audio/video, robotics, virtual reality or advanced artificial intelligence.

So, in the same way, it seems plausible that before we get to a stage where we have recreated a human brain in a machine (general artificial intelligence) we will be able to create machines that resemble humanity so well that they might gain the illusion of consciousness. If I can interact with a robot that looks like a human, speaks like a human, moves like a human and exhibits the same types of emotion and creativity as a human, I will be hard-pressed to deny it being conscious. Our brains will perform the anthropomorphising necessary to create the illusion of consciousness. The machine will pass the Turing test and we will be none the wiser.

So, perhaps the question we should be asking is not 'Will machines ever become conscious?' but rather 'How will we know if they are or not?'