

## Are Computer Neural Networks Intelligent or Conscious?

I began college in 1969 as a double major in computer science and mathematics. Perhaps because I came of age in the 1960s during the race to the Moon, I was captivated by computers and amazed at what these machines could do. We've come a long way since then, though! My smartphone today has more computing power than the IBM 360 mainframe computer that I used back then and which occupied an entire floor of a building. The Voyager 2 space probe launched in 1977 and that is still speeding away today in interstellar space has about the same memory as a garage door fob. Computing power continues to explode, and with this explosion comes a commensurate increase in just what computers can do.

Today artificial intelligence is the rage, and each day comes with some new headline about AI. You have your choice of ChatGPT, Gemini, Claude, and Perplexity AI chatbots, and with a little technical skill you can build your own AI chatbot on your home computer. Using a voice app on your smartphone, you can converse with your chatbot as you would with another person. The only difference you would probably notice is that the chatbot knows more about, well, everything, than your brightest friend. Unsurprisingly, we now talk about these chatbots as intelligent. Indeed, the holy grail of AI is to achieve "artificial general intelligence," which, very loosely, means the point at which the AI intelligence is indistinguishable from average human intelligence.

But are these machines *really* intelligent, and if so, how similar is this intelligence to human intelligence?

Originally in the 1970s and 1980s researchers took a "rules based" approach to machine intelligence. Take, for instance, IBM's chess playing computer Deep Blue. In 1997, it beat Garry Kasparov, a world champion chess player. The approach was to teach Deep Blue all the rules that govern chess moves, and then use its great computation power to determine the best move given any particular arrangement of chess pieces. It is estimated that there are  $10^{40}$  possibilities, and Deep Blue could evaluate 200 million of them per second! It is simply not possible for a human being to evaluate that many possibilities so quickly.

This approach to machine intelligence continued to drive the development of artificial intelligence until somewhere around the mid 2000s when researchers began to realize that human intelligence could not be reduced to just *an application of rules*. Human intelligence includes insight, creative flashes, and more. After all, as if to make this point, some years later Garry Kasparov beat Deep Blue. Some thought that this meant that artificial intelligence could never be real intelligence, at least, something which is akin to human intelligence.

Things began to change again when Gregory Hinton argued for a different approach called “deep learning.”<sup>1</sup> This approach attempts to mimic how the human brain works and learns by building machine analogues to the brain’s neural network. The human brain has approximately 80 to 100 billion neurons and over 100 trillion synaptic connections between these neurons. Our understanding of how the brain works, how it learns, stores, and retrieves information is still rather primitive, but very roughly our experiences of the world cause changes in the brain whereby certain synaptic connections are strengthened or weakened. The relative strength of these connections provides the physical architecture for information processing, including storage and retrieval. This is the brain’s neural network.

In deep learning, AI researchers attempt to duplicate this architecture. Once done, the machine is set free to go about learning from its training data. In short, the machine begins to teach itself in ways somewhat analogous to how our brains learn as our senses deliver data to it. This approach stumbled along for a few years with image and audio processing, but began to explode in its power around 2016 with the introduction of what would become known as large-language models.

To get a sense of how this works and the possibilities, imagine that you gather together all the world’s data, and then let an extremely powerful computer analyze all of the possible patterns. By patterns, I mean determining and cataloging the statistical strength of possible correlations of every word, phrase, and sentence with each other, and then applying this body of information to various questions, tasks, and prompts.

---

<sup>1</sup> [Heroes of Deep Learning: Geoffrey Hinton - DeepLearning.AI.](#)

This approach is proving very fruitful, and almost every day we read about some new discovery.<sup>2</sup> There is a clear sense in which the machine “knows” more than we do, but, again, is this really what we mean by “intelligence,” and even if so, how close to human intelligence is it?

These are difficult questions to answer. For one thing, large-language models are only one form of AI. And while right now they are the most promising forms, we know that they are particularly weak in deductive reasoning and certain mathematical operations which humans can easily do. But it is likely that they will be surpassed in power and capability in the intermediate future. For another, there is, really, no agreed upon measure of intelligence, much less human intelligence. And so most discussions of just what intelligence is inevitably get bogged down in imprecision.

Still, I think that these machines, broadly, are intelligent. After all, we don't find any particular difficulty in speaking of animals as intelligent, and even plants might be said to be intelligent, for they communicate through chemical signals and adjust to changing environmental conditions. I think behind the resistance of simply saying or admitting that these machines with advanced neural networks are intelligent is our legacy assumption that intelligence necessarily implies consciousness.<sup>3</sup> We might argue that since these machines are not conscious, they cannot really be said to be intelligent in any meaningful, substantive sense, at least, which is akin to human intelligence.<sup>4</sup>

But perhaps they are! Once dismissed as nothing but unfounded speculation, there is a growing view, called panpsychism, that consciousness is an irreducible constituent element of reality such that all things necessarily possess some degree of consciousness.<sup>5</sup> You might think of panpsychism as consciousness being smeared across the universe. And while I find this a difficult concept to wrap my head around, much the same might be said about the discoveries of

---

<sup>2</sup> Here are just a few: solutions to [climate change](#), [disease detection and treatment](#), [drug development](#), [improvements in battery storage](#), etc. Further, just recently Tesla released a new version of its autopilot system called “supervised driving.” Though early, most reviews say that it is a major improvement over previous versions. What has not been as noted is that this new release relies upon a neural network while the older versions used a “rules based” approach. See here, <[Revolutionizing Driving Forever: How Tesla's Neural Networks Power Full Self-Driving Capabilities](#)>.

<sup>3</sup> It doesn't, but here I wish to explore the claim that intelligent machines could never be conscious.

<sup>4</sup> It is not uncommon to hear critics dismiss large-language models as “souped autocorrect” systems doing little more than regurgitating information programmed into it. <[Opinion: Generative artificial intelligence is simply a waste of our time and money - The Globe and Mail](#)>.

<sup>5</sup> See, for instance, [Consciousness: Not just a problem for philosophers - Big Think](#) or Philip Goff's discussion of panpsychism in *The Universe Knows Right from Wrong*. <[The Universe Knows Right from Wrong - Nautilus](#)>.

quantum physics. It can reasonably be said that there is no science fiction today that is stranger than what science is discovering about the nature of reality.

But we needn't subscribe to panpsychism to entertain, though, the possibility that neural networks might indeed, at some point, achieve some level of consciousness, perhaps even self-awareness. Despite our familiar, first-person experience of conscious states, we do not have a well-understood theory about how consciousness arises, that is, how exactly does the brain produce consciousness? But however this question is eventually answered, it remains the case that consciousness arises out of brain activity, that is, out of physical states and processes. Physical states and processes are just that, physical. And if physical, then it seems to me to be a sort of prejudice to limit the possibility of consciousness to biological processes much less to biological creatures such as ourselves. Unless we discover that conscious states cannot arise from physical states, which seems unlikely, or retreat into some sort of obscurant mysticism about the nature of consciousness (such as that consciousness is the product of an immaterial soul or mind), we do not have good grounds for summarily dismissing the possibility that some sort of advanced AI could be conscious.

It is not my contention that today's AI large language models are conscious, for the complexity of our most advanced AI systems today pales in comparison to the complexity of the human brain with its 100 trillion or so synaptic connections. But this is a quantitative, not qualitative, difference. In time, with the advent of quantum computers and future AI systems that surpass today's just as today's computers vastly surpass those of only a few years ago, we may find that we have brought into existence not just a new form of intelligence but a new form of consciousness. Perhaps this is what Harari means when he suggests that homo sapiens as a distinctive human species may come to an end within a century.<sup>6</sup> There is much to be worried about here, and much, if we dare, to be excited about.

Michael D. Redmond

April 6, 2024

---

<sup>6</sup> For Harari's comments, see: <[Mustafa Suleyman & Yuval Noah Harari -FULL DEBATE- What does the AI revolution mean for our future?.](#)> Elon Musk has also argued that artificial intelligence puts the human species at risk, and suggests that the future is some sort of cyborg melding human and machine capabilities. See <[Elon Musk: Humans must merge with machines](#)>.