

Philosophy of Artificial Intelligence: Interpretations of Human Intelligence

To define the intelligence of the mind or human intelligence is difficult as it touches on complex concepts such as neurology, psychology, and philosophy. I looked at the definition from Wikipedia that paraphrases from Tirri's *Measuring Multiple Intelligences and Moral Sensitivities in Education*: "Human intelligence is the intellectual capability of humans, which is marked by complex cognitive feats and high levels of motivation and self-awareness." I found it interesting to contrast this definition to the class definition of Artificial Intelligence:

[Artificial Intelligence] is the intelligence exhibited by machines or software. It is an academic field of study which studies the goal of creating intelligence. Major AI researchers and textbooks define this field as "the study and design of intelligent agents", where an intelligent agent is a system that perceives its environment and takes actions that maximize its chances of success.

Computer scientists have expressed interest in programming human intelligence—to break down a human's "complex cognitive feats and high levels of motivation and self-awareness" to just "machines or software." The Chinese Room thought experiment presents the argument of strong and weak AI, whether a computer can understand or pretend to understand (Searle 417-457), and the China Brain thought experiment asks whether the consciousness can be broken down into neurons or units of two-way communications, and each neuron can be represented by an operator (Cole). But do these thought experiments capture the essence of human intelligence? To create a computer with human intelligence, I believe that computer scientists must interpret human intelligence from a more philosophical perspective beyond these thought experiments.

The Chinese Room and China Brain thought experiments offer valuable insights into adaptations of AI and human intelligence, yet I think these thought experiments threaten to misrepresent the human psyche. The Chinese Room experiment presents the high-level question of understanding versus simulation which I think can only be answered with more research and practical implementation as the experiment itself black boxes the implementation. Does the program use some sort of neural network to learn the Chinese language: to understand its semantics and grammatical structures? Or does it just have access to a language resource? Translation as an intelligence test also presents complexities as linguists argue the difficulty of a perfect translation as each language can have unique cultural nuances (Hafez 23, Heisig). This brings to question the validity of the thought experiment itself as a representation of artificial-human intelligence. The China Brain thought experiment threatens to be an oversimplification of the human psyche. The argument lacks attention to the subjectivity of an individual human mind (Tye 140) and limits the thought experiments. Scientists need to further contemplate the essential component of human intelligence to create a computer with human intelligence.

I believe that, with more investment and dedicated research, there can eventually exist a computer program that can simulate a human mind: a computer capable of learning and imitating the neurological processes of the human brain to build a specific human profile with their agendas, beliefs, personality, faults, and capabilities. I understand the effectiveness of creating different computers, each with its own perfected aspects of visual, language, or emotional processing. What, however, would be the pragmatism of combining these attributes into one human-computer? In science fiction and horror, I've noticed writers often focus on the anxiety surrounding notions of an artificial-human or androids. From the 19th century, Mary Shelly's *Frankenstein* warns of the unnatural as Dr. Frankenstein is ultimately punished for creating life on his terms. In the 20th century, Harlan Ellison's *I Have No Mouth and I Must Scream* presents the horrors of a supercomputer that kills most of humanity and tortures the few remaining humans for creating its self-aware and immortal existence. Other works with similar themes and anxieties include the anime *Psycho-Pass*, Frank Herbert's *Dune* series, and Philip K. Dick's *Blade Runner: Do Androids Dream of Electric Sheep?* I think humanity is capable of programming human intelligence, but, when considering how often science fiction explores the negative tropes of human-computers, I wonder about the practicality of programming a computer with human intelligence.

References

Tirri, Nokelainen (2011). Measuring Multiple Intelligences and Moral Sensitivities in Education. *Moral Development and Citizenship Education*. Springer. ISBN 978-94-6091-758-5.

[Searle, John](#) (1980), "[Minds, Brains and Programs](#)", *Behavioral and Brain Sciences*, **3** (3): 417–457, [doi:10.1017/S0140525X00005756](#).

David Cole (2009). "Section 2.3 The Chinese Nation". The Chinese Room Argument. *Stanford Encyclopedia of Philosophy*.

Tye, Michael, 2006. "Absent Qualia and the Mind-Body Problem" *Philosophical Review*, Vol. 115, No. 2, 2006, 140.cs

Hafez, and Sedarat Roger. "Perfect Translation." *World Literature Today*, vol. 84, no. 5, 1 Sep. 2010, 23.

Heisig, James W, and James W Heisig. "East Asian Philosophy and the Case Against Perfect Translations." *Comparative and Continental Philosophy*, vol. 2, no. 1, 1 Jan. 2010, pp. 81 - 80.

Ellison, Harlan. *The Fantasies of Harlan Ellison*. Gregg Press, 1979.

Schoene-Harwood, Berthold. Mary Shelley, Frankenstein. Columbia University Press, 2000.

[Not a source I actually used but an interesting read.]

Morgan, Jason M. "Artificial Intelligence and the Human Person." *The Human Life Review*, vol. 45, no. 3, 1 Jan. 2019, pp. 53 - 52.