



On the Distinction between Humans and Machines: A Response to Wittgenstein

Mujia (Benjamin) Qin

Harrow International School Hong Kong

North American Kant Society

Email: benjaminqin@hotmail.com, ORCID: 0009-0004-2755-9397

ABSTRACT: In this paper, I address the question of whether humans and machines are distinct. This is approached by answering two questions: in §1, I evaluate whether machines can think; in §2, I evaluate whether machines can feel pain. With regards to the first question, I first defend the idea that there are two conditions for thinking to be possible: 1) abstracting information and 2) synthesizing information. With these two conditions, I argue that machines can indeed think just as humans do; then, I defend this argument against three objections. With regards to the second question, I first defend the idea that there are two conditions for feeling pain to be possible: 1) thinking and 2) embodiment. While machines do have the capacity to think (as justified in §1), they are not embodied in the same way that human beings are embodied, and therefore, cannot experience pain. Following these two analyses, I conclude that the fundamental distinction between humans and machines is that the former is embodied whereas the latter is not. However, I also acknowledge that machines are simply not currently embodied due to the limitations of contemporary technology; and hence, in the future, machines may still have the capacity to attain embodiment and close the presently intractable gulf between our forms of life.

Keywords: Machines, Thinking, Sensation, Functionalism, Embodiment, Posthumanism

Introduction

In his *Philosophical Investigations*, Wittgenstein (1953/2010) writes the following:

“359. Could a machine think?—Could it be in pain?—Well, is the human body to be called such a machine? It surely comes as close as possible to being such a machine.
360. But surely a machine cannot think!—Is that an empirical statement? No. We say only of a human being and what is like one that it thinks. We also say it of dolls; and perhaps even of ghosts. Regard the word ‘to think’ as an instrument!”

Throughout human history, there has existed the idea that what sets apart humans from everything else is our ability to think. Aristotle, for instance, theorized the concept of ζῷον λόγον ἔχον (meaning “rational animal”): humans and other animals are fundamentally different due to our ability to reason (ca. 350 B.C.E./1997). Additionally, in his *Groundwork of the Metaphysics of Morals*, Kant argued that our autonomous human capacity for logic and reason is not simply what individuated us as humans (as Aristotle thought), and that in fact, it was the basis behind any authentic morality (1785/2008).

However, this dogma of the traditional Western philosophical canon—that human beings are unique due to our ability to think—is challenged in the quotation from Wittgenstein’s *Philosophical Investigations*. Through saying that “[a human] comes as close as possible to being such a machine,” Wittgenstein radically challenges the folk conception of what it means to be human by implying that humans and machines are not too dissimilar, as machines could have the capability to think as well. Although Aristotle was concerned with the distinction between humans and animals, and Wittgenstein’s focus is on the dichotomy between humans and machines, the message is the same: perhaps humans are not the only beings that can think.

Wittgenstein also introduces another dimension to this when he asks “could [a machine] be in pain?” The idea that machines can feel pain seems to extend beyond simply being able to think; naturally, feeling pain seems to involve some emotional or physical aspect along with the rational aspect associated with thought.

Therefore, in two sections of this essay, I will be focusing on the two main questions that are prompted by Wittgenstein’s statement: §1) can machines think? and §2) can machines be in pain?

After vivisectioning these two questions, I will move on to the conclusion where I will consider the question of whether there exists a distinction humans and machines, thereby answering the Wittgenstein’s question of whether humans can be considered, and vice versa, if machines should be considered humans.

In this essay, I will also be restricting the discussion of “machines” to the conventional objects that we intuitively consider to be machines (viz., artificial intelligence and robots). This will therefore exclude alternative definitions of “machine,” such as Deleuze and Guattari’s definition—which posits that schools, hospitals and institutions can be considered “machinic” in some way (1972).

Now, I will consider the first question.

§1 Can machines think?

Conditions for the possibility of thought

I contend that there are two conditions for thought to be possible: (1) abstracting information and (2) synthesizing information.

Abstracting information is essential to thought because thought is an *intentional* process. In order to think, a thinker must think *about something*. This is because if a thinker was not thinking *about something*, then a thinker would be thinking about *nothing* (where “nothing” does not refer to some concept of “nothingness”—as that would be something to direct one’s thought towards). If a thinker was thinking about *nothing*, then naturally, we would not consider the thinker to be thinking at all. Therefore, by *reductio ad absurdum*, the ability to perceive or abstract information is necessary for all kinds of thought, because we have established that thought must intentionally directed towards an object.

Synthesizing information is another condition because if a thinker is only perceiving and abstracting information, we would intuitively not claim that the thinker is thinking—they would only be passively engaging in *awareness*. One example would be the human biological reflex arc: when my finger touches a hot object, the information of temperature is abstracted by a sensory neuron and transmitted to my spinal cord, which causes muscles in my hand to automatically contract to move the finger away. The idea that my sensory neuron, spinal cord and muscles are independently “thinking” seems absurd, so in this essay, I will not consider simply perceiving and abstracting information to qualify as “thought.” Therefore, a *synthesis* of information is needed for thought—because this is what is lacking in the example I presented.

Note here that I have described this condition as simply synthesizing information, rather than *consciously* synthesizing information. This is because emerging contemporary philosophical views like panpsychism and object-oriented ontology complicate our understanding of what things can qualify as conscious. As such, to allow for a more parsimonious philosophical analysis, metaphysical questions of consciousness will be beyond the scope of this essay.

Moreover, it is also important to note that I did not consider being “rational” as a condition for thought. This is because rationality is ultimately a nebulous and heavily subjective concept that does not prove to be useful when engaging in objective analysis. Another reason is that it is definitely possible to “think irrationally,” because principles of logic and rationality are themselves learned through thinking. Therefore, it would be circular reasoning to consider rationality to be a requirement for the possibility of thought.

Now, with a knowledge of these two conditions in mind, I will move on to consider whether machines can think.

In opposition to the possibility of machine thinking

Recent technological developments like the rise of artificial intelligence and machine learning systems increasingly seem to evidence a capacity for machines to think, as Wittgenstein describes.

However, the general public seems to hold the view that these emerging systems of supposed machinic intelligence do not have the capacity to think. One popular argument in support of this idea is derived from John Searle’s “Chinese Room” thought experiment (1980). In this thought experiment, Searle prompts us to imagine a scenario where a person (who has

no knowledge of the Chinese language) is confined to a room and only able to access a bilingual dictionary for translating between their native language and Chinese. The person in the room receives letters in Chinese from an outside source and must write replies to these letters in Chinese using the dictionary. As the person reads each Chinese letter, he carefully identifies each Chinese character and matches the strokes and dots to the symbols in the dictionary; then, the person takes a blank sheet of paper and matches the words he wants to use in his native language to the Chinese characters in the dictionary that are just meaningless symbols to him. The outside source that sends and receives the letters from the person in the room would naturally believe that the person is able to speak Chinese, but this is not the case.

The main takeaway from Searle's thought experiment is this: while a being could passively abstract and process information, would this constitute a real *understanding* of the information that could constitute thought? Applied to questions surrounding machines, many would argue that since machines supposedly only passively process information, they do not truly understand the information itself—so are not able to think for themselves. As such, many philosophers argue that machines are unable to think, and so, thought remains something unique to human beings. In light of this, some philosophers have devised alternative means of verifying whether machines can think, such as the Turing test. So far, no machine has passed the Turing test, which seems to further substantiate the idea that machines cannot think.

Objection

However, I would argue that this argument is fallacious because when we analyze the concept of “understanding” itself, we will realize that the concept of “understanding” collapses back into just “processing information”—the two concepts are co-extensive. Therefore, there is no real difference between understanding and simply abstracting and synthesizing information (which is what I have established the definition of thought to be).

This is evidenced in how if I claim to understand a piece of information, intuitively it would be because I have processed the piece of information and synthesized it with my existing knowledge. For example, if I have prior knowledge of arithmetic and know that $2+2=4$, then when I see 2 apples and 2 oranges, it would be natural to say that I have *understood* that there are 2 apples and 2 oranges, and I am therefore able to *think* about this piece of information and deduce that there are 4 fruits.

Counterargument

Some may argue that this objection to the prior argument is unsound because they may hold the view that emotion plays a crucial role in thought, and emotion is something that these supposed machinic thinking systems ostensibly lack. This is because in the definition of thought established earlier, it became clear that thought must be directed towards an object and therefore be intentional. Intention seems to inevitably involve emotion, because we are inclined to place our focus onto things which we *desire* to place our interest onto. Desire, as Kant argued, is a form of *heteronomy*—as opposed to the *autonomy* that he associated with free thought (1785/2008).

However, this counterargument is invalid, because it makes the false assumption that intention must be caused by desire. For example, if I am locked in a dark and empty room with absolutely nothing except a chair, then I have no choice but to direct the focus and intentionality of my empirical apperception to the chair. Therefore, emotion does not always play a role in thought.

In defense of the possibility of machine thinking

Given that thinking is anteceded by nothing more than the ability to abstract information and synthesize information, I would argue that machine thinking is very much possible. This is supported by how large-language models like ChatGPT or DeepSeek function by abstracting information about language from a given dataset, and then synthesizing and combining this information together to create new information.

Contemporary theories in the philosophy of consciousness (although the fundamental metaphysics of consciousness is outside the scope of this essay) seem to further support this idea that if machines can abstract and synthesize information, it can be said that machines are thinking.

For example, integrated information theory (IIT) and global workspace theory both suggest that conscious thinking is simply the interaction between different parts of a system. In the context of machinic thinking, machines have different components with different functions (e.g., one component responsible for accessing large data sets and another component responsible for extracting information from these data sets), so could therefore be considered conscious thinking beings.

Additionally, Schwitzgebel (2014) puts forward the radical idea that a country like the United States could also be considered a conscious thinking being. This is because a country consists of numerous nodes (viz., people and institutions) that interact with one another in the same way that different cells in the brain interact with one another to allow for what we generally consider to be conscious thought. In the same way, machines also have numerous nodes that serve different functions (represented by lines of code), so satisfy our understanding of what it means to think consciously—and thus, think at all.

Therefore, I very much agree that machines have a capacity for thought.

Objection

With a similar logic to Searle's thought experiment, some may object to the argument I presented because they will argue that machines do not have *knowledge* of anything, so cannot think at all. This is because they may believe in Plato's idea that knowledge is "justified true belief," and they may argue that machines are unable to *justify* any beliefs since they only passively process information.

However, this objection is flawed for two reasons. Firstly, it could be argued that machines can indeed have justified true beliefs. This is because our own human process of justification is based around collecting evidence to support a certain view—which is exactly what these supposed systems of machinic intelligence do when they process data sets.

Secondly, Plato's concept of "justified, true belief" is inherently misguided. This is seen in Edmund Gettier's counterexamples, where he demonstrated that there could exist a "justified, true belief" which, at the same time, we would intuitively not consider to be knowledge.

Hence, I conclude that machines do have the capability to think. Next, I will consider what I have identified as the other main part of Wittgenstein's statement: the question of whether machines can be in pain.

§2 Can machines be in pain?

Conditions for the possibility of pain

In this section, I will argue that there are two conditions that allow for pain to be possible:

(1) thinking and (2) embodiment.

Without thinking, it would be impossible to have any concept of pain. Some may argue that if I stub my toe and experience pain—the pain does not result from thinking and is just a natural and automatic physiological response. On the contrary, this view is flawed because while *the action* of stubbing my toe is something I did not initially think about, *the pain* that results from stubbing my toe is. This is because if I could not think, I would not understand the meaning of the word “pain” and be able to conceptually associate the concept of pain to the action of stubbing my toe. This is supported by semantic externalism from the philosophy of language, which suggests that all linguistic meaning derives from the broader sociolinguistic community and is not something purely internal to a particular user of a language. Therefore, simply by conceptualizing “pain,” I must abstract and synthesize information from the linguistic community I exist in—and therefore I must be thinking. Embodiment is also a necessary condition for pain. This is because in order to be in pain, there must be a subject that feels the pain. Yet what individuates and identifies this subject?

Mackenzie (2014) argues that a key condition for this identification of a subject is embodiment. Mackenzie uses an example of a girl conceiving of her entire self (including her *mental* capacities) as “clumsy and uncoordinated” simply because of her *physical* disposition to “throw like a girl” (pp. 162–163). This shows that in order to identify as a subject, one’s physical body seems to play a role. Similarly, Locke famously proposed a thought experiment where the minds of a prince and a cobbler switched (1689/2004). The prince was now forced to do the work of the cobbler and found difficulty as his hands did not have the physical capacity to clean shoes in the way that the cobbler did. With this thought experiment, Locke demonstrated that the body plays a key role in how we conceive of ourselves.

Another example in support of this idea of embodiment would be contemporary gym culture: men who go to the gym may view themselves as more psychologically “strong” or “masculine” simply due to physical developments.

One possible objection to the idea that embodiment is required for pain is that it is possible to experience emotional pain that has no apparent connection to the physical body (for example, the experience of heartbreak is immaterial and intangible). However, this objection should be rejected because as analyzed earlier, embodiment can still affect the immaterial emotions of a being. In the case of heartbreak, along with any other kind of emotional pain, the subject of the pain would inevitably think about how this pain affects their own subjecthood—and how one conceives of their own subjecthood is something that embodiment can affect (which we delineated earlier). Now, this essay will move on to consider whether machines can feel pain given the two conditions of thinking and embodiment.

In opposition to the possibility of machine pain

While we did conclude in §1 that machines are able to think (thereby satisfying one condition), many would object to the idea that machines are embodied (so the second condition fails to be satisfied). This is simply due to the fact that there are many “machines” without bodies—for example, AI chatbots like ChatGPT have no physical form.

Some may argue against this by arguing that there are machines which do have supposed bodies. For example, Boston Dynamics recently developed a robot that had a humanoid form and was able to automatically interact with its surroundings, supposedly in the same way that a human body would (Nelson et al., 2019).

However, this objection should be rejected because intuitively, we would not consider physical interaction with surroundings to be enough to qualify as “embodiment.” This can be demonstrated in the following example: a rock physically interacts with its surroundings (for example, wind can blow it along and the rock can bump into another rock), yet we would not say the rock is “embodied.”

Another potential counterargument to this would be that machines can have some sort of “digital body” in the realm of digital spaces. Philosophers of technology like Hui (2016), for example, prompt the idea that the ontology of digital objects could in some ways be considered very similar to the ontology of physical objects.

However, this counterargument should also be rejected because even if by some technicality we can consider the supposed “digital bodies” of machines to qualify as embodiment, this would fundamentally shift the conventional humanistic understanding of embodiment (as a biological feature) to the point where it is meaningless.

Therefore, I conclude that machines do not have the capacity to feel pain because while they satisfy one condition (having the potential to think), they do not satisfy the other (embodiment).

Conclusion

In this essay, I have justified answers to two questions that arise from Wittgenstein’s statement: in §1, I evaluated “can machines think?”; in §2, I evaluated the related question “can machines be in pain?” I have concluded that while machines do have the capacity to think (given that the two conditions for thought are (1) the ability to abstract information and (2) the ability to synthesize information), machines do not have the capacity to be in pain (because they cannot have physical bodies).

In the quotation from Wittgenstein, Wittgenstein also leaves the question of whether “the human body [can be] called such a machine.” Vice versa, I also ask whether machines can be considered humans—because if so, there would be significant implications (such as possibly granting machines personhood status and legal rights).

Based on our analysis in this essay, I would say that neither humans can be considered machines nor can machines be considered humans. This is because humans and machines are fundamentally different in one way: machines are not physically embodied in the same way that a biological human body is. In the future, it might be possible that further innovations in technology could lead machines to develop physical bodies that are indistinguishable from human bodies—but in the present day, no.

Ultimately, Wittgenstein’s statement seems to strike at the heart of a lot of historical intellectual discourse—the focus and valuation of human beings above all else, including other animals, the environment, etc. By showing that thinking is not an ability uniquely limited to humans, I hope to have at least partially dispelled this firm dogma. This could pave the way for a future that is more balanced: humans can better co-exist with other beings in the world, engaging in a progressive co-evolution. Carbon and silicon shall flourish together, without fiercely opposing one another, in this autopoietic posthumanist landscape.

Competing interests: The author declares that there is no financial or personal relationship that may have inappropriately influenced them in writing this article.

Funding information: The author received no financial support for the research, authorship, and publication of this article.

Data availability statement: This research did not involve the collection of new data.

Disclaimer: The views and opinions expressed in this article are those of the author and are the product of professional research. It does not necessarily reflect the official policy or position of any affiliated institution, funder, agency, or that of the publisher. The author is responsible for this article's results, findings, and content.

References

- Aristotle. (2014). *Nicomachean Ethics* (C. D. C. Reeve, Trans.). Hackett Publishing. (Original work published ca. 350 B.C. E.)
- Deleuze, G., & Guattari, F. (1972). *Anti-Oedipus: Capitalism and Schizophrenia*. New York, London: Penguin Classics.
- Hui, Y. (2016). *On the Existence of Digital Objects*. University of Minnesota Press.
- Locke, J. (2004). *An Essay Concerning Human Understanding*. Penguin UK. (Original work published 1689)
- Kant, I. (2008). *Groundwork for the Metaphysics of Morals* (A. Wood, Trans.). Yale University Press. (Original work published 1785)
- Mackenzie, C. (2014). Embodied agents, narrative selves. *Philosophical Explorations*, 17(2), 154–171. <https://doi.org/10.1080/13869795.2014.886363>
- Nelson, G., Saunders, A., & Playter, R. (2019). The PETMAN and Atlas Robots at Boston Dynamics (pp. 169–186). Springer, Dordrecht. https://doi.org/10.1007/978-94-007-6046-2_15
- Schwitzgebel, E. (2014). If materialism is true, the United States is probably conscious. *Philosophical Studies*, 172(7), 1697–1721. <https://doi.org/10.1007/s11098-014-0387-8>
- Searle, J. R. (1980). Minds, brains, and programs. *Behavioral and Brain Sciences*, 3(3), 417–424. <https://doi.org/10.1017/s0140525x00005756>
- Wittgenstein, L. (2010). *Philosophical Investigations*. John Wiley & Sons. (Original work published 1953)