Critical Analysis of the "No Relevant Difference" Argument in Defense of the Rights of Artificial Intelligences

Ali Reza Mazarian*

Abstract

There are many new philosophical queries about the moral status and rights of artificial intelligences; questions such as whether such entities can be considered as morally responsible entities and as having special rights. Recently, the contemporary philosophy of mind philosopher, Eric Schwitzgebel, has tried to defend the possibility of equal rights of AIs and human beings (in an imaginary future), by designing a new argument (2015). In this paper, after an introduction, the author reviews and analyzes the main argument and then deals with four criticisms witch Schwitzgebel himself has presented and rejected and ultimately, the author presents and explains two new critiques against the argument. According to the writer, though the argument regarding his new criticisms and some other probable critiques is not convincing, or at least needs to be corrected, however, it holds appeal and raises questions and may open a new horizon for future questions and researches.[1]

Ph.D. Graduate of the Institute for Cognitive Science Studies, Tehran | alirezamazarian@gmail.com

Keywords

Artificial Intelligence, rights, ethics, philosophy of mind, Schwitzgebel.

Introduction

Along with the remarkable progress of science, varied types of super computers, robots and artificial intelligences have emerged. The science of artificial intelligence is still at the initial stages of its awesome path. But at its very beginning, it has inspired new philosophical, ethical and legal questions.

A group of these philosophical questions are moral ones. In such questions, the main focus is on moral considerations about robots and other similar beings. In principle, what moral status may AIs have? Do we have moral responsibility to robots? What is our responsibility towards robots? Can robots have rights similar to animals and humans? Can robots be considered as moral or amoral agents as well? While we are designing and programming robots, should we assume that they have a moral responsibility for their operations and their consequences?

Recently, some philosophers of mind (such as Schwitzgebel: 2015) and Mara Garza have tried to provide an argument (with the title: "No Relevant Difference") in defense of AI rights. The author seeks to first describe and review the argument and then to criticize it. The "No Relevant Difference" argument has two versions and the main controversy is about the second one.

The second version: The similarity of the moral status of human and AIs

Premise 1. If Entity A deserves some particular degree of moral consideration and Entity B does not deserve that same degree of moral consideration, there must be some *relevant difference* between the two entities that creates the foundation of this difference in their moral status.

Premise 2. There are possible AIs who do not differ in any such relevant or involved respects in the discussion, from human beings and their rights are similar to those of human beings

Conclusion

Therefore, there are possible AIs who deserve a degree of moral consideration similar to that of human beings.

The most important critique against the argument is the supposed "possibility" of some particular beings in the second challenging premise of the argument. Many matters which can be considered epistemologically possible (in terms of their conceivability) are in fact impossible metaphysically and in the external world. There is an important distinction between genuine possibility and conceivability. If we really want to discover whether the property or state we are conceiving as possible to be realized in the world, is really a possible one or not,

we need to reasons which mostly depend on our scientific and empirical knowledge. Although conceivability can be a method for discovering metaphysical possibles and impossibles, nevertheless, it is fallible and can show possibles as impossible and vice versa.

I think that Schwitzgebel, and Garza have mixed up the two kinds of possibility regarding the possibility of the emergence of AIs who have the same psychological states as that of humans, without having a human body. Based on epistemic possibility, it can be conceived or imagined that it is possible that some robots, without having any human body, can have psychological states such as phenomenal consciousness, free will, feeling, introspection and so on. But considering possible metaphysical worlds which are more limited than conceivable possible worlds, we find that such beings are not metaphysically possible. My claim is that it is not metaphysically possible that a being can have human psychological properties but is not human.

References

- Bostrom, N. (2003). Are We Living in a Computer Simulation? *Philosophical Quarterly.* 53, 243-55. doi: 10.1111/1467-9213.00309.
- Chalmers, D. (2002). Does Conceivability Entail Possibility? in T. S. Gendler & J. Hawthorne (Eds.) *Conceivability and Possibility,* (pp. 145-200). Oxford: Oxford University Press.
- Egan, G. (1994). *Permutation City.* London: Millennium.
- Egan, G. (1997). *Diaspora.* London: Millennium.
- Fine, K. (2002). The varieties of necessity. T. S. Gendler & J. Hawthorne (Eds.). *Conceivability and possibility.* (pp. 253-281). Oxford: Oxford University Press.
- Hume, D. (1777/1975). *Enquiries Concerning Human Understanding and Concerning the Principle of Morals.* eds. L. A. Selby-Bigge & P. H. Nidditch, Oxford: Oxford University Press.
- Kripke, S. (1980). Naming and necessity. Cambridge: Harvard University Press.
- Putnam, H. (1960). Minds and Machines. *Journal of Symbolic Logic*, 57-80. doi: 10.2307/2271581.
- Schwitzgebel, E. & Garza, M. (2015). A defense of the Rights of Artificial Intelligences. *Midwest Studies in Philosophy*, 39, 98-119. doi: 10.1111/misp.12032.
- Soames, S. (2011). Kripke on epistemic and metaphysical possibility: Two routes to the necessary a posteriori. In A. Berger (Ed.), *Saul Kripke.* (pp.167-188). Cambridge: Cambridge University Press.