

Review of: "Artificial Consciousness: Misconception(s) of a Self-Fulfilling Prophecy Nobody Wants"

Ashkan Farhadi¹

¹ University of California, Irvine

Potential competing interests: No potential competing interests to declare.

Review points

I had the pleasure of reviewing this well-written paper on the subject of AI and consciousness titled "Artificial Consciousness: Misconception(s) of a Self-Fulfilling Prophecy Nobody Wants" by Birgitta Dresch-Langley.

Strengths:

The article is well-written and benefits from a wide search of current literature. I personally enjoyed reading it and as a researcher in this field, I learned a lot from this review/opinion manuscript. The style starts from basic information about consciousness and gets into details of relevant theories and studies on the subject.

Major Comments:

- **Mixing the main concepts:** The author is using consciousness, awareness and attention loosely and interchangeably. At some points, author relates learning, memory, knowledge, novelty and experience in a close connection with awareness and consciousness. I do not blame the author for this indiscretion since the literature is filled with this interchangeable language, and even the hard problem of consciousness is in fact the hard problem of awareness. To clarify this very important point, I briefly describe these terms: Consciousness is a momentary state of being and does not have any subject while awareness is a subjective experience of raw or processed information in the form of informational or emotional intelligence. The subject of awareness could be sensation, memories, knowledge and so forth but nonetheless, awareness needs a subject. It is important to understand that the subject of awareness is not the awareness itself. Based on this definition awareness is needed for consciousness but is not enough for it, where the intentionality is key to complement awareness into a conscious being. I refer the author and readers for further reading to this newly published manuscript titled "Trilogy Theory of Mind".

https://mkscienceset.com/articles_file/600-_article1702457016.pdf

- **Plausibility of a conscious AI:** In a very elaborate manner, the author acknowledges that the path for AI to become conscious is very hard. In fact, author expresses several of these difficulties and even provides comments in some instances for those who may be interested to incorporate those modifications in the current AI configuration. In other words, the author portrays a difficult path to reach from point A, where we are now, to the point B where AI is conscious. This reminds me of an analogy. Currently we have no bridge across the ocean to travel the distance by car but anyone skilled in art or science can contemplate a solution to this problem that could be possible one day with advances in technology. However, if we limit the solution to a bridge with no support on either side or middle, it appears as if there's no solution. The solution to the latter problem may be hard to contemplate and one may think a solution to this problem is

beyond our current understanding of science of physics or structure.

Unfortunately, the author presents the problem of AI and consciousness like the first problem, where the problem is presented as hard to do but within the realm of possibility. In fact, the author provides a series of reasonable suggestions and instruction. In my opinion, due to lack of awareness, AI can never be conscious. I will refer to the author and reader to investigate this perspective in a published manuscript titled "There is no "I" in AI."

<https://link.springer.com/article/10.1007/s00146-020-01136-2>

- Ignoring the realistic but dark side of AI: I agree with the author that the idea that AI can be a human god is overblown but the idea that AI can do many mental functions that we brag about much better than humans is currently a reality. The article is completely devoid of any general perspective about that aspect of the issue and rests its case as long as "AI cannot be human God". We all know that AI is capable and in fact has already replaced many forms of human mental function, calculation, prediction and modeling that were once exclusive processes for humans. AI has also replaced so many physical and semi-physical task forces. Therefore, there is no question that AI is much more powerful in any single aspect of isolated mental functions such as memory, calculation, prediction, judgment, or making a proper decision. I liked to believe in the author's reassurance that since AI cannot possess a human mind, we can all have a sigh of relief since the human mind serves as a privilege for humans exclusively. In fact, the reality is far from the fuzzy and warm prediction provided by the author for several reasons:

- o Even though current AI may not be capable of doing all mental functions of one human brain in one place, or is not able to put all algorithms of human mental function in one unit, AI (or those who make AI) are capable of connecting all AI doing single tasks to performs in a functional network of AIs where multiple units at different locations and times can perform any mental function as well as or superior to a human. Does this large network fit into a skull? Of course, not and there is no need for it to. This network could be composed of a giant set of machines that communicate with each other or with their AI actuators to not only do mental functions of a human mind but also the physical work that is needed to be done.
- o The rate of advancement proposed in this article in AI is based on advancement in technology. However, the new generation of AI may take the lead in advancing their offspring, which would make many predictions about solving complex issues to perform mental functions like a human or even better than a human in much less time or in a larger magnitude. Therefore, as I explained, even though AI can never be conscious, their mental function could exceed humans' in the near future.
- o Another issue that keeps many scientists up at night could be the use of AI as a weapon. AI does not have to be human God but could be in the hands of some users that can turn them into God in the sense of an undefeatable world domination. There is no caution about this aspect of AI that if it is not real by now, it is not far from reality.

Minor comments:

The author states, " When we are born, all brain activity is more or less arbitrary, not necessarily random." I was wondering where we get this understanding. How can we know if any movement in an infant, or any natural intelligence for that matter, is arbitrary, on autopilot or random, if we can communicate with that entity.

