

## Peer Review

# Review of: "The Hard Problems of AI"

Zbigniew Orbik<sup>1</sup>

1. Silesian University of Technology, Gliwice, Poland

The author of the article presents, in the approach to the "hard problems" of consciousness, or more broadly consciousness as such, a philosophical point of view which I hold myself. However, a review of the scientific text should be critical, so I will try to formulate some questions rather than reservations. The article should be considered, in my opinion, as a valuable attempt to point out the fundamental issue: the importance for solving or at least clarifying the "hard problems" of consciousness of the conceptual (linguistic) nature of the issue being considered.

Olegas Algirdas Tiuninas believes that the so-called "hard problems" of consciousness, if I understand it well, are practically unsolvable or, more precisely, they are of a purely linguistic nature. By analogy to the famous concept of Chalmers, for whom hard problems are those that cannot be explained using cognitive science methods, he considers as examples of such problems "the questions of AI consciousness, sentience, agency, intelligence, capability to be genuinely creative, and so on". These are, in short, philosophical problems. The question arises: are these problems undecidable on the grounds of philosophy at all, or can we assume that further development of tools of philosophical analysis will allow these issues to be solved (explained)? In other words, can we definitively decide that these are only linguistic problems?

Problems with the definition of consciousness, which the author rightly draws attention to, are, however, not something exceptional. A cursory review of philosophical issues developed over the centuries allows us to notice that there has probably never been a definition of any fundamental concept that would satisfy everyone on philosophical grounds. The concept of consciousness is therefore no exception. The question arises about the nature of philosophical issues and progress in philosophy.

As we know, *qualia* play a key role in attempts to understand the phenomenon of consciousness. C.I. Lewis, considered the creator of the concept of *quale*, in his book entitled *Mind and the World Order*, writes: "In any presentation, this content is either a specific quale (such as the immediacy of redness or

loudness) or something analysable into a complex of such. The presentation as an event is, of course, unique, but the qualia which make it up are not. They are recognisable from one to another experience” (Lewis, 1929, 60). *Qualia* are therefore the properties of what is given. What is interesting, in my opinion, is the concept of “quasi space” (“the Private Arena Model”) introduced by the author of the text. He draws attention to the linguistic nature of the problem that appears. However, the question arises again: is the problem, in the case of consciousness, only linguistic? I have some doubts whether Wittgenstein’s analogy with a beetle in the box is entirely accurate. It explains the mechanism of using language in certain situations. But is it an adequate explanation of the problem of consciousness? In my opinion, probably not. It should be remembered that language is also one of the products of consciousness. The phenomenon of consciousness goes beyond the sphere of language. The essence of consciousness is intentionality and understanding, while the essence of *qualia* is experiencing. Even more mysterious is the phenomenon of self-awareness.

It seems to me that the basic difficulty in answering the fundamental question: whether computers can or will ever be able to think, whether they will become conscious, results from the fact that we do not really know what it means to think. The author of the article, referring to Plato, notes that the meanings of words such as consciousness or intelligence are fluid or fuzzy. He states: “I believe that the only reason we want there to be an essence to such entities is because the notion is so central to our world-modelling that we want it to have clear and distinct boundaries – making it more ‘real’ in our minds, more tangible (this is the same reason why the Humean account of causality as something continuous with correlation is unintuitive at first, as it challenges the perception of the central notion of causality as a separate entity). Thus, we try to artificially separate it from a melee of ambiguous terms with only approximate boundaries by insisting that the boundaries of this particular notion are sharp”. I think the author rightly draws attention to the inadequate nature of conceptual cognition. Among the great thinkers, Bergson, for example, wrote about this (Bergson & Andison, 2010). Olegas Tiuninas makes an interesting prediction regarding the future approach to the issue of AI system awareness: “I believe that as time passes and the capabilities of these systems become ever more advanced, it will just be too convenient to think of them as conscious for anyone to seriously ask the question”. Of course, the future will show how it will be. In my opinion, however, as the history of science and philosophy shows, there will always be researchers trying to capture the nature of consciousness. Another issue is whether this research can end in success. It seems that they are able to at least make us understand better the mechanisms of the mind.

Alan Turing noted decades ago: "The whole thinking process is still rather mysterious to us (...)" (Turing, 1951, 116). After many years, Chalmers repeated: "Conscious experience is at once the most familiar thing in the world and the most mysterious" (Chalmers, 2017, 3). Despite the passage of time and the undoubted progress in cognitive science, it seems that the phenomenon of consciousness has not ceased to be a mystery to us. The question arises: does it have to be this way? Stuart Hameroff and Roger Penrose point to three possibilities: "Consciousness is not an independent quality but arose, in terms of conventional physical processes (...) *Consciousness is a separate quality, distinct from physical actions and not controlled by physical laws, that has always been in the universe.* (...) Consciousness results from discrete physical events; such events have always existed in the universe as non-cognitive, proto-conscious events, these acting as part of precise physical laws not yet fully understood" (Hameroff, Penrose, 2014). Of course, I realize that the article was about a different problem. I wonder, however, if the author would be willing to accept either of these possibilities?

In my opinion, the considerations contained in the article are characterized by a depth of philosophical reflection. The author can be seen to be genuinely interested in the discussed issues. The article is undoubtedly very inspiring to reflect on the topic. It can be viewed as a part of the long tradition of attempts to capture the phenomenon of consciousness. It is undoubtedly a step forward on the path to solving the mystery of consciousness or at least clarifying the meaning of predicates such as "conscious" or "intelligent". This is also progress.

#### References

- Bergson, H., & Andison, M. L. (2010). *The creative mind: An introduction to metaphysics*. Courier Corporation.
- Chalmers, D. (2017). The hard problem of consciousness. *The Blackwell companion to consciousness*, 32-42.
- Hameroff, S., & Penrose, R. (2014). Consciousness in the universe: A review of the 'Orch OR' theory. *Physics of life reviews*, 11(1), 39-78.
- Lewis, C. I. 1929 (1956). *Mind and the World Order: An Outline of a Theory of Knowledge*. New York: Scribner's Sons. Reprint ed.
- Turing, A. M. (1951). Can digital computers think. *The Turing test: verbal behaviour as the hallmark of intelligence*, 111, 116.

## **Declarations**

**Potential competing interests:** No potential competing interests to declare.