

Review of: "On the subject part I: what is the subject?"

David Mumford¹

¹ Brown University

Potential competing interests: No potential competing interests to declare.

For me, the most intriguing part of this article is the author's introduction of the word "urge" as a central component of an interacting system of the world and an organism. He discusses *awareness/consciousness* as the key component of subjective experience of the organism and *agency* as the key component of the organism's interaction with an external world. But one needs to add some kind of *urge* to cause the *awareness* of the subjective entity to then seize its *agency* and so modify the world. Whether this be a eukaryotic cell swimming away from some toxin or a human voting for their preferred candidate in an election, the urge needs to always be there. On the other hand, there are things like genes and ideas that have much in common with organism/world, especially evolution, but lack urges.

The reason this struck home to me was my involvement in artificial intelligence. There is endless debate about the gap between the remarkably successful *large language models* (LLM) and a possible robot embodying full *general artificial intelligence* (GAI). Much of the current article can apply in illuminating ways to this debate. How, for example, can a LLM ever know what is true in the real world, can ever reject what its masters call its *hallucinations*, without senses giving it some fragment of awareness? Truly, it needs to have eyes and ears and to record its experiences. But the more controversial part of this debate are the potential uses of such computer systems. If and when they actually control things that are important to us, they will begin to have agency and, of course, to be flexible helpers in the real world, they cry out for mobility (even if tethered electronically to massive servers). And finally we will be forced to endow them with urges to make them truly helpful servants. Then we face the so-called *alignment* problem, first described in Goethe's poem *Der Zauberlehrling*.

I hope the author, in later parts of this piece, addresses also AI systems.