Qeios

Peer Review

Review of: "The Hard Problems of AI"

Steven Frank[1]

1. Independent researcher

This paper treats the question of whether artificial intelligence (AI) can exhibit consciousness or sentience and convincingly argues that the concept is too definitionally elusive to provide a solid basis for evaluation. Equally convincingly, the author argues that the question can only be addressed from a pragmatic perspective, based on values that inform our decision whether to apply the label – and we should be honest and explicit about this. The article nicely considers whether "hard questions" are really meaningful questions at all and effectively challenges the assumption that particular consequences or rights ineluctably flow if an AI is deemed conscious.

The author is far more expert than I in the relevant intellectual foundation and analytical framework, so I offer the following thoughts. First, I would distinguish between consciousness, which is very hard to define, and concepts such as creativity and intelligence. Creativity is often defined as the ability to invent new ideas rather than merely synthesizing existing ones, and tests have been devised to measure it – e.g., the ability to devise alternative uses for everyday objects. Is creativity the same as intelligence? Maybe. The Turing test was conceived as a test of a machine's ability to exhibit intelligent behavior indistinguishable from that of a human. Is intelligence the same as consciousness? Who knows, but we can more readily define and measure intelligence, even if our definitions and measures are subject to debate. Consciousness, by contrast, is often defined or detected tautologically – e.g., in the cited Chalmers article, as (among other things) by "seeming conscious."

Second, if the Turing test was conceived (or at least operates) more as a measure of intelligence than consciousness, then really the consciousness goalposts haven't been moved; they just haven't been defined with enough specificity, for reasons the article ably articulates. The question remains definitional and therefore subjective. For example, suppose we decide that consciousness, as a concept and however defined, is only meaningful for creatures whose self-awareness equips them to survive and flourish in a physical world among other such creatures. In that case, a machine that clears whatever goalposts we define is still not conscious because the concept isn't meaningfully applied, and we are

guilty of, as the author puts it, "making a mistake in assuming that there are some objectively extant necessary and sufficient conditions for attributions of the predicate."

Declarations

Potential competing interests: No potential competing interests to declare.