

Saving physicalism/materialism: the Chalmers test

A. U. Thor*
The University of Uranus

November 5, 2018

1 Definitions

Let us refer to a "Turing test" in a generalized sense as a test of the capacity of any simulacrum-entity (including but not restricted to AI) to mimic or emulate a specific behavior or function displayed by a sapient, sentient person; if any other sapient, sentient person cannot distinguish it from a real sapient, sentient person based on an analysis of that specific behavior or function alone, the simulacrum is considered to have passed the test. From this definition, we see that any simulacrum of a conscious person that passes all possible Turing tests ascribable to that person is behaviorally indistinguishable from it - we will call such a simulacrum a *T-zombie*.

But plausibly behavior isn't everything - if a few lines of code suffice to fool someone into thinking it is a professional psychotherapist [1], we'd be hard-pressed to point this as proof (or even as evidence) that those lines of code possess some kind of consciousness or cognitive phenomenology. To do that, we'd need a means to define rigorously what the criteria for "consciousness" is, before we meaningfully talk of whether the lines of code obey those criteria or not - which at our current state of art we do not even know what they should be. Regardless, we may still, *in absentia*, define a (possibly metaphysical or noncomputational) oracular tester-entity *SEARLE* inside a black box that can apply the unknown criteria to any entity given as input and give as output an answer to the yes-or-no question posed by the Turing test above (the "yes" output meaning "pass", "no" meaning "fail") but with "behavior or function" replaced by "conscious or cognitive property"; such a scheme constitutes the *Chalmers test* - and similarly to the previous paragraph, we shall refer to any simulacrum of a conscious person that passes all possible Chalmers tests ascribable to that person a *C-zombie*. Furthermore, iff a *T-zombie* is also a *C-zombie*, we shall refer to it as a *P-zombie*¹.

*Email: glacatlan@hotmail.com

¹Notice the important detail that, according to this definition, *we* ourselves are trivially

We shall also have need to talk of world-simulacra, or as we will call them, *simulations*, which we will define in the following way:

Definition 1 *A X -simulation is any logically consistent simulation of the world which reproduces isomorphically all of its Y and is indistinguishable from the world if referred only to its internal parameters², for $(X, Y) \in \{(T, \text{physical properties}), (C, \text{phenomenology of consciousness}), (P, \text{properties and phenomenology})\}$.*

Of course, this implies that the world itself is a P -simulation, since it is trivially isomorphic to itself.

2 The standard argument³

1. All that exists in the world (including consciousness) is physically-based.
2. Any P -simulation must be a T -simulation.
3. Any P -simulation must not be a C -simulation.
4. Therefore, consciousness is not physically-based.

Analysis: $\neg 3 \therefore \neg 4$.

3 The counter-argument, or the "CPT proof"

1. All that exists in the world is physically-based, except maybe consciousness.
2. Any X -simulation must be peopled by X -zombies.
3. Any P -simulation must be both a T - and a C -simulation.
4. Therefore, consciousness is physically-based and physicality is consciously-based.

Ironically, an independent route to 4 was provided by Chalmers himself with his synthetic brain Gedankenexperiment [3].

¹ P -zombies: we will always have a 100% approval rate in any test that checks if we act and think like what we do.

²One can distinguish a high-resolution simulation of the ocean's surface from the real thing if, e.g., one refers to the fact one is looking at a computer screen.

³Adapted from [2], pp. 107-108.

References

- [1] Roger Penrose, *The Emperor's New Mind - Concerning Computers, Minds, and the Laws of Physics*, (Penguin Books, 1989)
- [2] David J. Chalmers, *The Character of Consciousness*, (Oxford University Press, 2010)
- [3] D. J. Chalmers, *Sci. Am.*, 1995, 273 (6), 80-86