

SCHUBERT, Lenhart (schubert@cs.rochester.edu)
University of Rochester, US

PARTIAL ZOMBIES, PROSTHESES, AND A PRETTY HARD PROBLEM

The problem considered here is whether certain kinds of partial zombies are possible. The question is answered affirmatively, and this is argued to have consequences for any attempt to formulate a general theory of consciousness, one that allows prediction of consciousness (or otherwise) and of kinds of phenomenal experiences for any kind of sense-equipped, intelligent being, regardless of physical constitution and functional architecture. This is referred to as a "pretty hard problem", since predicting consciousness and qualia may be easier than explaining them, as required by David Chalmers' "hard problem" (Chalmers 1996). The immediate objection against seeking a general theory, that predictions of qualia in significantly different beings cannot be verified, is addressed as well.

The kinds of (partial) zombies considered here fall somewhere between (partial) behavioral and (partial) functional zombies. Behavioral zombies are sometimes treated in philosophy as obviously possible, since for instance a human could be mimicked by an elaborate puppet remotely controlled by a supercomputer or a person's nervous system (Polger 2000). However, in AI circles it is more often thought that human-like subjective experience would very likely be an inevitable consequence of duplicating human-like behavior in a self-operated (as opposed to remote-controlled) agent. Therefore, only the possibility of self-operated behavioral zombies is of interest here. This moves us closer to the notion of functional zombies, though we still want to allow for differences in information processing architecture. In addition, we set aside the issue of self-awareness here, since self-awareness is "merely" a kind of access consciousness (Block 1995). (Indeed, various researchers are working on the development of self-aware agents - see, e.g., AAAI-SS 2005, AAAI-SS 2007). Rather, the question of interest here is whether a self-operated agent could be behaviorally like a human but (partially or largely) lack phenomenal consciousness.

As examples of such behavioral zombies, we consider a deaf person equipped with glasses containing a micro-miniaturized speech and sound recognition system that displays the results to the wearer visually, in a way that is undetectable by onlookers; and (as a more extreme example), a blind person with a vision gadget, disguised as sunglasses, that whispers detailed descriptions of the scene confronting the wearer, allowing the wearer to act and communicate pretty much like a sighted person. Of course, these partial zombies still do have phenomenal experiences, but in "collapsed" modalities (perhaps only a linguistic modality). So, AI researchers should not assume that a human-like inner life is an inevitable byproduct of self-determined, apparently human-like behavior. We are then left with the "pretty hard" problem; a key question is at what level of functional description we should seek a basis for predicting phenomenal experience. Some suggestions will be offered having to do with the "geometry" of various qualia modalities (similar to the notion of the "logic" of qualia as discussed by Jeffrey Foss, 2000).

References

AAAI-SS 2005, 2005 AAAI Spring Symposium on Metacognition in Computation, AAAI Tech. Rep. SS-05-04, Stanford University.

AAAI-SS 2007, Logical Formalizations of Commonsense Reasoning, AAAI Spring Symposium, March 26-28, Stanford.

N. Block, 1995, "On a confusion about a function of consciousness." Behavioral and Brain Sciences 18(2):227-287.

D. Chalmers, 1996, The Conscious Mind: In Search of a Fundamental Theory, New York and Oxford: Oxford Univ. Press.

J. Foss, 2000, Science and the Riddle of Consciousness: A Solution. Kluwer Academic Publishers.

T. Polger, 2000, "Zombies Explained," in D. Ross, A. Brook, & D. Thompson (Eds.), Dennett's Philosophy: A Comprehensive Assessment, MIT Press, 2000. See also "Zombies", <http://host.uniroma3.it/progetti/kant/field/zombies.htm>

Keywords: phenomenal consciousness, zombies, artificial intelligence