# "The Extended Mind"—Extended

## Joseph S. Fulda

"Philosophy," it has been said, "is the finding of reasons for what one believes by instinct." In January, my Internet access was cut off for four days and I felt as if a part of me was gone, a part of my mind, with all the disorientation that one might expect. Later that month, I read a fascinating article by Clark and Chalmers [1] that gave me the reasoning for what I had believed by instinct. The extended mind is partly an artificial mind, composed of the brain and artifact alike and as such Clark and Chalmers' excursion into the philosophy of mind is equally a contribution to the philosophy of artificial intelligence.

We present their theory here and then extend it further so as to explain my impaired mental function during Internet blackouts. Clark and Chalmers introduce their subject thus:

> Where does the mind stop and the rest of the world begin? The question invites two standard replies. Some accept the boundaries of skin and skull, and say that what is outside the body is outside the mind. Others are impressed by arguments suggesting that the meaning of our words 'just ain't in our head', and hold that this externalism about meaning carries over into an externalism about mind.

Both views invite counterarguments of considerable weight and Clark and Chalmers have fashioned, with formidable ingenuity, a middle ground that does not appear to be susceptible to these standard objections. Whether other objections to their views will prevail, or whether their theory will withstand the test of time, is, of course, unknown: Philosophical arguments that have appeared cogent for centuries have been destroyed by fleeting flashes of inspiration or genius—witness Gettier's demolition of the idea that 'knowledge' is 'justified, true belief' or Kripke's demolition of the idea that a 'name' is shorthand for a 'bundle of properties' of the named. It is certain, in fact, that *my extension* of Clark and Chalmers, as powerful an explanation of certain mental phenomena as it may seem, will, not having undergone rigorous prior peer review, in time be seen as flawed. My hope is that the flaws will not be

fatal, that the line of analysis here put forth—based on a paradigm from AI—will be sustainable, even though the fine print of a philosophical position—extensive modifications, refinements, exceptions, and cautions will be required—as it always is.

Clark and Chalmers introduce their theory by considering a nontrivial mental task, rotation, which for the strength of the argument, we will take as reflection, that is rotation along the *z* axis. Consider, they say, three ways of performing this task: (1) the old-fashioned way, by performing the reflection "in one's head," (2) the modern way, by reflecting the object on a computer screen using a "retrieve image" and a "reflect image" command sequence, or (3) the way of the "cyberpunk future," in which a man with electrodes implanted under his skull, an implant that allows him to perform mental reflections effortlessly, sits in front of a screen with an image and can choose to use his implanted function or his natural function to rotate the image on the screen.[*] It is apparent to Clark and Chalmers and to this writer as well that (3) is philosophically no different from (1). On what grounds, then, can (2) be considered different? Not on the grounds that it is not all in the head, because that is the question at issue. Yet what else is there to distinguish (2) from the likes of (3)? If in both (1) and (3) the mind accomplishes the mental task, what is the "mind" in (2)? The answer, say Clark and Chalmers, is "the extended mind" which consists of the "coupled system" of man and machine. A simple and beautiful insight, but one never advanced before in the literature.

---

[*] Although the term "cyberpunk future" is from the *Analysis* paper, it is not hard to imagine. In [2], Robert Fox of the ACM informs us of this interesting Cal Tech experiment:

> A computer chip that interfaces directly with brain cells via standard integrated-circuit techniques has been developed and tested in rats.... Pitted with 16 depressions attached to a tiny electrode connected to a computer, researchers filled each depression with nerve-flourishing substances, then placed neurons from embryonic rat brains into each depression and allowed them to grow. Neuron extensions grew over the walls separating these wells, made connections with each other as they would in a developing brain, and detected individual firings between nerve cells,.... Eventually, say researchers, the chips could be used to enhance various brain functions.

Now we cannot here recount all the potential considerations against this view ably rebuffed by Clark and Chalmers, nor can we retrace in the space available to us their many *Gedanken* experiments in support of their position, although we do ask the skeptical reader to hold his fire until reading their intricate defense of their position. What we can and will do here, however, is relate their main conclusions from all the argumentation over the fine print as to what is and what is not necessary for an external component of a coupled mental system to be properly so-called.

First, the coupling must be *reliable*, so that *access* to the external component can be assumed, if not always, then just as often and as reliably and as reasonably as we have access to internal mental mechanisms. Second, what is not of concern is *how* the access is gained, e.g. whether by introspection or perception.

From this, Clark and Chalmers conclude that beliefs are also not in the head—but in the mind, and, in particular, in the extended mind. They set forth four conditions for beliefs: (1) the external component which has the information ascribed to the belief must be such that the person will rarely act on a matter relevant to the belief without first making use of the external component; (2) the information must be readily available; (3) upon retrieving the information, the person automatically endorses it as his belief; and (4) the information has been so endorsed by the person in the past and is there as a consequence of that endorsement.

Now these are not trivial conditions; on any view which does not reject outright the concept of the extended mind, they are quite restrictive. We have no wish to loosen the conditions, only to extend what qualifies as information sufficient to form a belief from declarative knowledge to procedural knowledge, in the AI tradition. [3] Thus, I am prepared to say that I have a belief about the time of day—all day, every day, about the sum of any two large 7-digit numbers, about what libraries held my latest book as of yesterday, about what indexing and abstracting services picked up last quarter's column, and about what rulings were handed down by the United States District Court for the Southern District of New York last week.

On the face of it, of course, these seem like preposterous claims of knowledge, but if one accepts the procedural view of knowledge—that having an algorithm to obtain the knowledge is a form of the knowledge, then all we need are external components that can be reliably accessed (a watch, a calculator or the Windows calculator, the

BIB database of the Research Libraries Group (RLG) and the WorldCat database of the Online Computer Library Center (OCLC), the indexing and abstracting services available online through OCLC, ISI, the AMS, and numerous other sources, and the Lexis database on the Web).

I will almost always consult my watch, calculator, or these databases if asked to act on my beliefs in these matters (for example, by answering a simple query as to what they are!): condition (1) met. I have—Internet blackouts aside—ready access to all these external components: condition (2) met. I would fall for a failing watch, a flawed Pentium chip, or an error in a database as reliable as the above—I do in fact endorse the information that comes *ex machina* pretty automatically, not noticeably less, at any rate, than information that comes from my "own" memory—and perhaps more: condition (3) met.

Condition (4) seems to be the problematic one, but if we bear in mind that it is the procedure that genuinely comprises the belief, the procedure in each case has been consciously endorsed before and the belief about the procedure is there because of its prior endorsement. Of course, as Clark and Chalmers briefly suggest might be true for the declarative case and is surely true for the procedural extension we propose, the fine print of the position involves getting to a coherent and correspondent (i.e., internally and externally consistent) account of meeting the fourth condition.

Although we do not provide such an account here, we *have* sketched out a theory that explains why someone who uses the Internet daily (and particularly the Web as a gateway to intranets) would feel that he was losing his mind, that his beliefs were slipping away, when his Internet access is cut off. How exactly do we feel? The same way we feel the day we forget our watch at home or (if we are under 35) cannot find a calculator when we have need of arithmetic. Perhaps the *Silicon Valley Sentinel-Observer* will report on a case that may yet come before the Supreme Court: Is twice being in jeopardy of loss of Internet access twice being in jeopardy of loss of limb?

## References

[1] Andy Clark and David Chalmers, "The Extended Mind," *Analysis* 58(January 1998): 7-19.

[2] Robert Fox, "Chip off the Old Gray Block," *News Track*, *Communications of the ACM* 41(January 1998): (1)9.

[3] An extensive dialogue between the declarative and procedural views of knowledge can be found in the first and best edition of Patrick Henry Winston's classic work, *Artificial Intelligence* (Addison-Wesley Publishing Company, 1977), pp. 390-392. (For a review of some of the defects of the third edition see my "PC in Computer Science?" *Measure* 113(February 1993): 5-6.)