CORRESPONDENT'S REPORT

Recent Work in Philosophy of Interest to AI

Daniel Dennett

Philosophy Department, Tufts University, Medford, MA, U.S.A.

Some philosophers have written about AI directly, and some have written about topics that people in AI are, or ought to be, interested in. I will usually concentrate on the latter variety both because it is less likely to come to your attention on its own, and because it is generally better work. Philosophy's potential contribution to the field is not exhausted—or even very well represented—by philosophical proofs of the field's impossibility or futility, arguments about how to interpret the Turing test, or even friendly arguments about how AI solves philosophical problems. But there has been some informed discussion by philosophers about these topics, and since this is my first column I will take some backward glances, and not restrict myself to the most current work.

I expect that few will need to be introduced to John Searle's well-known paper, "Minds, Brains and Programs," published together with much furious commentary and a reply from Searle in *Behavioral and Brain Sciences* **3** (3) (1980) 417–458. A further critical discussion—indeed a fatal dissection—can be found in Douglas R. Hofstadter's "Reflections" on Searle's piece in *The Mind's I: Fantasies and Reflections on Self and Soul* (Basic Books, New York, 1981), composed and arranged by Hofstadter and myself. Searle has pointed out a minor (and inadvertent) misquotation in these 'Reflections' that will be corrected in any subsequent printings of the book: where Searle said "bits of paper" Hofstadter has him saying "a few slips of paper". Readers are hereby warned not to let the difference in connotation bias their appreciation of Hofstadter's rebuttal. I have been told that a review by Searle of The Mind's I is forth coming the New York Review of Books, which should make interesting reading since Searle (correctly) views our book as a defense of what he calls "strong AI"—which he takes himself to have refuted. While I am into self-advertisement, I might as well add that the chapter on "Further Reading"

in the Mind's I provides a wealth of references, with commentary, to the philosophical literature people in AI should know—up-to-date as of publication, 1981. One might in fact consider this column and its successors to be semi-annual supplements to that chapter.

Another swipe at Searle is taken by Richard Rorty in "Mind as Ineffable", forthcoming in *Synthese* this summer, along with a reply by me. Rorty's piece provides a useful overview of the recent history of 'philosophy of mind'—made no less useful by being drastically overstated (my reply to his paper tries to correct that).

A good, provocative piece of philosophy of mind is Paul Churchland's "Eliminative Materialism and the Propositional Attitudes" in *J. Philosophy* **78** (2) 1981. That journal is in general the place to look if you are going to keep track of just one philosophy journal. It is the more or less official journal of record of the whole field, and moreover it has been particularly strong in publishing in the areas that would tend to interest people in AI: philosophy of logic, of language, of mind, of science. For instance, if you want to learn about *situation semantics*—the hottest new topic in philosophical logic—you will find two of the three published papers in *J. Philosophy*.

What is situation semantics? It is a logic of 'propositional attitudes' and perception-talk, in some ways a successor or rival to Montague semantics. It is the brainchild of the mathematician Jon Barwise and the philosopher John Perry. Their analysis of perception-talk depends heavily on a distinction (drawn by Fred Dretske) between epistemic seeing and non-epistemic seeing: I can see, non-epistemically, a brown cow without realizing it is brown or a cow—I may 'get it all wrong' but still it is out there being seen by me. Epistemic seeing involves seeing things as this or that, and seeing that such and such is the case. One of Barwise's initial inspirations came from considering the AI use of the term *scene analysis* in computer vision projects. Barwise decided that scenes were promising items in an ontology that carves nature at its joints, and with Perry developed that concept of situations and scenes (which are, as one would expect, what one sees when one observes a situation). A co-authored book is 'forthappearing', according to the authors, and in the meantime one can read three papers. First, Barwise and Perry, "Semantic Innocence and Uncompromising Situations" in *Midwest Studies in Philosophy* **5** (University of Minnesota Press, 1981) 387–403. (Midwest Studies is not a journal, but a yearbook of invited articles by the leading philosophers on a particular topic. The first half dozen volumes have been excellent.) Then two papers by Barwise, "Scenes and Other Situations", *J. Philosophy* **78** (7) (1981) 69–97, and "Situations, and Attitudes", *J. Philosophy* **78** (11) (1981) 668–691.

J. Philosophy **78** (11) also contains a paper by Searle, "Intentionality and Method", which "attempts to draw some of the methodological conclusions implicit in a series of studies" he has published on intentionality. The last few pages of this paper offer some reflections on what Searle calls 'the back-

ground'—the mass of world lore (roughly speaking) that he thinks to be necessary equipment for any believer. This idea is of course a close descendant of the views of Searle's Berkeley colleague, Hubert Dreyfus. Searle's discussion of it sidles close to the Frame Problem of AI without explicitly mentioning it, or even (so far as I can see) illuminating it—but rather like a bulge in the orbit of Mercury, his remarks indirectly reveal the presence of a weighty problem in the vicinity.

I have recently embarked on something of campaign to introduce the Frame Problem to philosophers, in the expectation of stimulating some useful work from my field on this most philosophical of AI problems. I hope I will soon be able to recommend some results in successor columns.