

Talking Heads ...

A Review of *Speaking Minds: Interviews with Twenty Eminent Cognitive Scientists*

Patrick J. Hayes and Kenneth M. Ford

The editors of this book were worried about cognitive science. They thought that the Chinese Room argument showed that computationalism could never fully account for the first-person perspective, that the "computer metaphor for the mind" might lead to some vital social questions being ignored, that passing the Turing Test¹ was nowhere in sight, and that symbol-processing AI based on Von Neumann architectures had foundered on the "commonsense problem." They conducted 20 interviews with a rather idiosyncratic collection of people, largely on the east and west coasts, to find out what the consensus was in the field. One of their happy discoveries was that connectionism (about which they initially knew little) was expected to overcome many of these obstacles. Each interview begins with a brief personal history of why the interviewee became involved with the subject and what they take it to be, and then moves into a discussion of contemporary issues which the editors find interesting. While the interviews do not conform to a set pattern, they return regularly to a few favorite themes: the Chinese Room, the importance of the Turing Test, why "symbolic AI" has failed (a claim that is made repeatedly throughout the book), and the significance of connectionism as a replacement for it.²

Their selection of interviewees betrays a certain bias. Of the 20 people interviewed, only Newell, Simon,

Wilensky, and Winograd could possibly be said to be active in mainstream AI; on the other hand there are seven or eight philosophers, of whom only Dennett has a sympathetic interest in AI; all the others have rejected its premises, and Dreyfus, Searle and Weizenbaum are notorious for their passionate and sustained attacks on the subject. This would be less important but for the fact that AI is the main subject matter of several of the interviews. Much of the interview with Dennett is focused tightly on

Speaking Minds: Interviews with Twenty Eminent Cognitive Scientists, Peter Baumgartner and Sabine Payr, editors, Princeton, N.J.: Princeton University Press, 1995. 376 pp. \$17.95 (paper). ISBN 0-691-02901-6. \$39.50 (cloth). ISBN 0-691-03678-0.

the Chinese room argument (the interviewers find this a vitally important topic, but seem unaware of the considerable amount Dennett had written on it). Weizenbaum, in contrast, was encouraged to expand at some length into such fantasies as what, in his opinion, motivates AI researchers (answer: male uterus envy); and Fodor and Putnam's contemptuous remarks on leading AI figures, which a sensitive editor might have removed or at least abbreviated, are reported in careful detail. Overall, one gets a very negative impression of AI from these pages. It is a pity that Marvin Minsky was not given an opportunity to express his point of

view, since he comes in for considerable vilification.

The interviews are also biased in significant ways. They regularly ask questions like "How do you see the history of AI in retrospect? What were the successes, the failures?" (to Pat Churchland), with the clear implication that whatever it was, it's all over now; or "I have heard many times that the early progress in AI has slowed down if not stopped all together" (to Herb Simon); or "What are the failures of AI?" (to John Searle); or finally "Will there be any important development in some fields of applications in the next years, or is AI just a complete failure?" (to Herb Dreyfus). This kind of question-begging is almost routine, and the interviewers often seem to fail to grasp replies which clearly explain their false assumptions.

It is conspicuous from the introduction and indeed throughout the book that the editors have only a sketchy and simplified grasp of the technical concepts which constitute the theory of cognitive science. For example, they talk about symbolists as being concerned only with software, while connectionists (the good guys) take the architecture into account as well. One can see what they mean, but this way of putting it is ludicrously oversimplified. In general, their grasp of computational science is about at the level of a children's encyclopedia.

Several of the people interviewed seem to share these rather naive views. The book is shot through with apparent quotes from well-known scientists which seem remarkably confused or simplistic in this way. Even if these words were dutifully transcribed from a recording, the reader probably needs to bear in mind the difficulty of expressing oneself clearly, while obeying the rules of polite conversation, on a subtle issue when talking to someone who displays such a profound level of ignorance and uninformed prejudice as these interviewers.

There are several places where the interviewee attempts to educate the interviewer, some of which are amusing to read. John Haugeland explodes

at the suggestion that when one is considering a trained skill, symbolic models are irrelevant because “the brain just does it,” and attempts valiantly to explain the concept of a virtual machine and show why the interviewer’s simplistic distinction between hardware and software is too elementary; and one can almost hear the care in James McLelland’s voice as he tries to dissect the interviewer’s simplistic notion of “symbolic.” Unfortunately, these valiant efforts do not seem to have resulted in the other interviews avoiding the same pitfalls. Sometimes the exchange consists of an incoherent, vague, or ridiculous question (“Couldn’t one argue that on the underlying level, the microprocessor and the circuits also form a pattern that is like the brain?” or, one of our favorites: “The symbol people invented LISP to support their point of view. Are connectionists working on a programming language...?”) followed by a puzzled, or politely noncommittal, reply; one has the impression that the interviewee just didn’t think it worth the effort to attempt the required amount of re-education. On other occasions, however, the interviewees seem to have decided to descend to the interviewer’s conversational level, allowing rules of polite discourse to override their usual intellectual care. The result is confusing if one is seeking for content, but can be insightful if one is interested in seeing how some famous thinkers’ minds run, since they here express themselves in a more casual way than their writings usually display.

This book is most useful not as a source of facts, but of opinions, attitudes, prejudices, hopes, and ambitions. The most interesting parts of the interviews are often the initial brief autobiographies in which people explain how they came to be working in cognitive science. Pat Churchland’s farming childhood accounts for her fascination with wetware; Newell’s account of the post-WW2 intellectual climate is fascinating, if brief, and complex forces helped form the young Weizenbaum’s views about binary thinking and the evils of Western sci-

ence. Dreyfus tells us that he has been interested in AI since 1957 as a kind of distillation of, and scapegoat for, a rationalist philosophical position which he rejects (although he doesn’t tell us why). Dreyfus knew from the start that AI was doomed, as it represents Husserl’s archaic rationalism, long refuted by Heidegger. He is not making a judgment based on evidence, but seeking arguments to bolster his faith; his relentlessly misinformed and shrilly emotional attacks on AI make more sense when seen against this background. Lakoff spends three pages telling us the history of his intellectual development and how it was at every stage somewhat ahead of Chomsky’s (who, regrettably, could not be interviewed.) His interview is also notable for the grandeur of its ambition, which at times approaches the poetic intensity, and indeed the poetic confusion, of William Blake. Lakoff is convinced that because humans think metaphorically, the human race will have to rebuild its intellectual foundations, civilization will rewrite its laws and constitutions, and our entire view of reality will be transformed. No doubt cognitive science can also expect a good brushing-over. Searle has been repeating the same arguments now for almost two decades and is still convinced that his Chinese room argument is correct. He shows a quite remarkable degree of self-confidence, bordering indeed on the delusional.

It is interesting that in some circles, most of them apparently orbiting Berkeley, the word “symbolic” has become a kind of generalized term of abuse, rather as “communist” was in much of American politics in the fifties and “liberal” is now. Its exact technical meaning seems adaptable from case to case, however. Lotfi Zadeh feels that the use of fuzzy logic will somehow escape the yoke of symbol-hood; Herb Dreyfus is confident that biological plausibility will enable connectionism to escape the Husserlian rigidity of symbolic rationality; Terry Winograd explains how conversations with Dreyfus led him to Heidegger, where he saw that “symbolic” models lacked the neces-

sary social connectivity, while Hauge-land (a former student of Dreyfus), feels that good old-fashioned symbolic AI lacks the sheer computational power needed to overcome the “commonsense problem.” Proponents of the symbolic perspective might almost be encouraged by the lack of coherence in the various positions arrayed against them here.

Many of the views expressed are quite predictable (Zadeh, for example, attributes all the problems of cognitive science to its attachment to binary truth-values), but some are more surprising. Hilary Putnam was one of the first philosophers to articulate some of the basic insights of computer modeling in psychology and made pioneering technical contributions to the subject (the Davis-Putnam inference method is still asymptotically as good as any), so the vehemence with which he criticizes the field and dismisses its achievements is striking. Like Jerry Fodor, who also comes across from this interview as a virulent critic of AI, Putnam dismisses the actual subject matter of AI as unimportant engineering details. Being unable to see anything else, he concludes that there is nothing there but empty boasting. On the other hand, the sociologist Aaron Cicourel carefully resists a barrage of confused and leading questions (“What do you think about the discipline of AI called ‘knowledge engineering’? Don’t you think their concept of ‘knowledge’ is totally formalized?”) and gives a thoughtful and well-informed analysis of the likely role of AI technology in society. Perhaps our experience has been unfortunate, but Cicourel seems to us to be unusual among sociologists in his clear grasp of these issues.

Several of those interviewed, especially the philosophers, identify consciousness as a crucial problem area for cognitive science. (Since these interviews were conducted, this topic has emerged quite suddenly into the general academic arena, with the publication of a number of books arguing extensively for or against the relevance of computational or neuroscientific modeling to consciousness studies.) It is important to appreciate

	Has AI Failed?	Is the Turing Test Useful?	Is the Chinese Room Convincing?	Common Sense Is AI's Waterloo?	Connectionism Is the Cure?	Degree of Symbolophobia
P. S. Churchland	✓	✗	-	-	✓	3
P. M. Churchland	✓	✗	✗	✓	✓	5
A. Cicourel	✗	✗	-	-	-	0
D. Dennett	✗	-	✗	-	✗	0
H. Dreyfus	✓	✓	-	✓	✓	5
J. Fodor	✓	✗	✗	-	✗	1
J. Haugeland	*	-	-	✓	✓	3
G. Lakoff	✓	-	-	-	✓	4
J. McClelland	✗	-	-	✗	✓	2
A. Newell	✗	✓	✗	✗	✗	0
S. Palmer	-	✓	-	✗	✓	0
H. Putnam	✓	-	✗	✗	✗	1
D. Rumelhart	-	✓	-	✗	✓	2
J. Searle	✓	✗	✓	✗	-	5
T. Sejnowski	✗	✗	-	✓	✓	5
H. Simon	✗	✓	✗	✗	✗	0
J. Weizenbaum	✓	-	✓	✗	✗	4
R. Wilensky	✗	✗	✗	✗	✗	0
T. Winograd	-	✗	✗	✓	✗	3
L. Zadeh	✓	✗	✓	✓	✗	3

✓ Yes ✗ No * Yes and no - No view was expressed

Table 1. A Handy Guide to the Condition of AI and Cognitive Science.

this interest, because it seems partly to motivate much of the opposition to “symbolic” AI displayed here, a trend that we have called symbolophobia³ (see the “On the Other Hand...” column in the Spring—volume 18 number 1—issue of *AI Magazine*). The unspoken argument seems to be this: conscious thinking is symbolic, and therefore symbolic AI only models conscious thought, or at any rate mental activity which is very similar to conscious thought. However, much of what goes on in the head doesn’t seem to be like conscious thought at all: it’s intuitive, spontaneous, inarticulate and apparently happens without conscious effort. Ergo, it can’t be modeled by symbolic theories or implemented on a serial computer. (We leave it as an exercise for the reader to find the errors in this argument, a task we often set for our undergraduate class.) Several of the connectionists, notably Sejnowski, articulate this position quite explicitly. Now, if the interview is read as an informal account of Sejnowski’s own motivation, this frankness is admirable, and indeed quite illuminating: but if it taken as an argument for a position, it is hardly convincing.

Many of those interviewed said things which would be textbook ex-

amples of fallacies if they were taken to be arguments, so it is probably more charitable to the authors to adopt the autobiographical reading throughout. With this interpretation, then, this book represents a kind of snapshot of fashionable opinions in a rather oddly selected cross-section of coastal academia from about five years ago. Indeed, its interest is already chiefly historical; there is already a wider and deeper understanding of the relationships between connectionist architectures and symbolic representations, and of the limitations and biological implausibility of many connectionist models, which make the uninformed enthusiasm of many of these interviews seem rather quaint. We summarize the opinions expressed on most of the central topics in a table above, which also gives our estimate of the various levels of symbolophobia present.

We have criticized this book rather harshly, but it contains some material of value. Perhaps we can say that for people in the field, who already have their views clear, it does make an interesting read; but that for students or laymen it would probably be a confusing, and potentially very misleading, way to find out anything about cognitive science.

Notes

1. As we have argued elsewhere (Hayes and Ford 1995), the Turing Test is a burden, damaging AI’s public reputation and its own intellectual coherence, but the editors’ apparent confusions about AI are more extensive than can be laid at its door.

2. Of course connectionism cannot be a replacement for AI — it *is* AI. This confusion arises from the common mistake of permitting AI, or any other field, to be defined by its methods rather than its goals. Connectionism is not a goal, but is one of AI’s many methods.

3. This disorder is described in the literature (Doctor and Kahn 1989) thusly, “Individuals may fear the symbols themselves with or without understanding their unconscious representation.” For more discussion of this distressing malady, see the “On the Other Hand...” column in the Spring (volume 18, number 1) issue of *AI Magazine*.

References

- Doctor, M., and Kahn P. K. 1989. *The Encyclopedia of Phobias, Fears, and Anxieties*. New York: Facts on File.
- Hayes, P. J., and Ford, K. M. 1995. “Turing Test Considered Harmful.” In Proceedings of the Fourteenth International Joint Conference on Artificial Intelligence (IJCAI-95). Menlo Park, Calif.: International Joint Conferences on Artificial Intelligence.