of meaning, and the problem of consciousness. Flanagan considers the pet claim of John Searle, this year's Reith lecturer, that programs cannot exhibit mentality because they lack a semantics. Existing computers have a most impoverished mental life because they interact with the world in highly restricted ways. They manipulate symbols in a purely formal manner without any grasp of what the symbols denote. One area of current work in cognitive science is the development of programs that effectively take an input from a television camera and from it construct a three-dimensional representation of the objects in the world giving rise to the image. As the computer's representations of the world are enriched, so too it should be possible to enhance its grasp of meaning.

There remains the barrier of consciousness. Cognitive scientists are divided on the issue of whether or not it is a computational process. Flanagan writes, echoing many such claims: "It is hard to see how the right qualitative character could emerge from the wrong kind of organic stuff, let alone from a plastic and metal electrical device." This line of argument falls into the same trap as the "cogito" – the assumption that the nature of reality is illuminated by what one can doubt. But, so too does the contrary argument. The issue is only likely to be resolved by attempts to formulate, and to implement, the sort of parallel computational architecture that is called for by the division between conscious and unconscious mental processes. What is clear is that no existing computer begins to approximate this design.

The Science of the Mind touches upon many other challenges to the cognitive sciences – rationality, free will, and the extent to which genes determine mental phenomena. Its author deftly takes apart bad arguments, and puts the pieces together again to solve a series of intriguing puzzles. What the book lacks, however, is an overall thesis or organizing framework. This weakness shows up in the team it fields to represent the development of the cognitive sciences: Freud is in, but Helmholtz (the architect of the unconscious inference) is out. Skinner is in (for bypassing the mind), but Craik (who refused to take the detour) is not. E. O. Wilson is in but David Marr does not even reach the index. Owen Flanagan thinks of science as having a narrative structure; he tells us excellent stories about his team. His book would have been even better if it had also had a story of its own to tell.

The significance of squiggles

David Papineau

JOHN SEARLE
Minds, Brains and Science
102pp. BBC Publications. £8.95 (paperback £5.25).
0 563 20286 6

John Searle's Reith Lectures have been widely received as a timely exposé of those woolly-minded computer-lovers who believe that computers can think, and indeed that the human mind is just a biological computer. In print Professor Searle's lectures retain the same punchy and engaging style as they had on the air. They also give us the opportunity to decide whether there is anything behind Searle's antipathy to computers apart from simple prejudice.

Searle starts off well enough. In the first lecture he raises various questions about the mind's place in the material world. Firstly, there is the problem of causal interaction – how can the mind influence the brain, and vice versa? Then there are the twin problems of consciousness and subjectivity – how is it that mental states involve subjective feelings of conscious awareness? And then there is the problem of "intentionality" – how is it that mental states can be *about* things, can reach out and refer to things other than themselves?

Searle deals with the first problem, the problem of causal interaction, by endorsing the now standard modern version of materialism. Mental events stand to brain events as, say, the solidity of a table stands to its molecular structure. One doesn't, in the latter case, think of the solidity as something puzzlingly detached from the molecular structure. No more should one think of pains, or beliefs, or emotions, as inhabiting some realm cut off from the neurophysiology of the brain. In neither case are there two substances, two different kinds of stuff. There are simply two different scales of description, two perspectives with different levels of focus.

Searle's remarks on consciousness and subjectivity are less satisfactory, however. In effect, he simply says that feelings manifestly do exist, and so we'd better believe it. True enough. But how does this sit with Searle's materialism? There is a long and respectable tradition, stemming from Descartes, which takes the existence of consciousness to show there must be more to mind than matter. Searle obviously believes that this tradition is mistaken, and that there is no reason why certain purely material states shouldn't be consciously experienced by the beings that have them. But then he surely owes us at least some account of where the Cartesians go wrong, and of how we should resist their intuitions.

The problem of intentionality is the focus of the second lecture. My desire to visit Naples relates me to a place where I have never been. My belief that Bob Hawke won the Australian election reaches out and refers to a man on the other side of the world. How can one thing be *about* another in this way? This is where Searle's antipathy to computers comes in. For he thinks that however the trick is done, it is obvious that we can do it and computers can't.

His rationale for discriminating between humans and computers in this way is the "Chinese room argument". For those who haven't heard it, the Chinese room argument goes like this. A man is sitting in a room containing lots of bits of paper with squiggles on. Further bits of squiggle-scripted paper are passed into the room. The man has a book of rules specifying, for every combination of squiggles that comes in, which set of squiggles he should pass back out. Unbeknownst to the man, the incoming squiggles are all good Chinese questions. And the rule book is designed to ensure that the man always passes out the right Chinese answer. But obviously the man doesn't thereby understand Chinese. However, such formal manipulation of symbols is all computers ever do. So computers must lack understanding too.

This is not a good argument, but it does get under one's skin. One can see why Searle has dined out on it for so long. For it is extremely difficult to identify the exact point where he leads us astray. And so each philosopher who disagrees with Searle has his or her own theory of what is wrong with the argument. My own response to it is that understanding Chinese isn't a reasonable first task to set a computer. Understanding Chinese is a highly sophisticated activity. Or rather, since the special complexities of Chinese are not the point, understanding *any* language is a highly sophisticated activity. By no means all intentional states involve a grasp of some public language. A dog can remember where a bone is buried; a chimpanzee can believe there are a lot of bananas up a tree. Maybe dogs and chimps aren't fully self-conscious. But it is not feelings of conscious awareness that are presently at issue, just the possibility of having thoughts *about* things. So dogs and chimps seem able to have intentional states, even though they understand no language. Similarly it seems wrong to say that all human beliefs involve language. Even when I believe that there are bananas up that tree, there is no direct sense in which my belief depends on my linguistic abilities.

One can think of understanding a language as a matter of having beliefs about words. To understand a Chinese sentence is to have the belief that a certain string of symbols represents a certain state of affairs. But this then only emphasizes the point that understanding a language is a quite special and sophisticated intellectual ability. For not all beliefs are beliefs about the representational powers of words.

Why then does Searle take it that the appropriate test case is whether computers can understand Chinese? No doubt his thought is that, whatever goes on in dogs and chimps and humans, *computer* operations consist entirely of the formal manipulation of symbols. So if a computer can do any thinking at all, it surely ought to be able to understand such symbols.

There is a sense in which Searle is right to think that digital computers spend all their time formally manipulating symbols. But only someone who quite misunderstands the suggested comparison between computers and minds will infer that they must therefore understand those formal symbols. Suppose I wanted to build an artificial chimp. One of the first things I would do would be to give it some structured internal state which was triggered whenever its sense organs were presented with bananas, and which led it to return to the bananas when it needed nutrition. This structured state would consist of an array of internal switches, or, at a different level of description, of a set of numbers in certain registers in the robot-chimp's memory bank. To this extent one could think of the structured internal state as a string of formal symbols. And having such an internal state is just the kind of thing that would incline computer enthusiasts to say that the robot-chimp was capable of believing that there were bananas up that tree. But they wouldn't dream of saying that therefore the chimp *understood* the formal symbols, the structured array of switches inside its head. The chimp wouldn't have any beliefs about those symbols, let alone beliefs which told it what those symbols represented.

I might even try to design a chimp that understood Chinese. To do this I might give it further internal states, in virtue of which it could be said to associate certain Chinese symbols with such non-linguistic things as bananas. So, for instance, I might give it some further internal state, again consisting of some structured array of switches, which ensured that whenever it saw a certain set of squiggles it would get into the original believing-there's-a-banana-up-that-tree state. Given all this, one might argue that the chimp understood those Chinese symbols to refer to bananas. But it would just be a confusion to conflate this with the claim that the chimp understood the structured array of switches which gave it this understanding of Chinese.

Of course there is much more to understanding human languages than simple word-banana associations. But the point is that whatever kind of understanding anybody, or any computer, has of anything, it is not supposed to be an understanding of some structured array inside its head. Rather, the thinker is supposed to understand *other* things, *because* he has such a structured array of formal symbols inside his head. Nobody (except perhaps the programmer) needs to understand the internal formal symbols themselves. And so arguments about what would or would not be understood by a little man manipulating such symbols inside the head, or inside a Chinese room, are of no significance.

Having come this far, it should be admitted that we are left with a real philosophical problem. If having beliefs is just a matter of having sentence-like structures inside one's head, then we still need to explain where the "aboutness" of mental states comes from. One can't just rest, as many cognitive scientists seem to, with the terminology of "sentences", or "symbols", or "representations". For, without any internal homunculi to breathe significance into them, it is not at all clear how brute internal structures can stand for anything. And to this extent Searle is justified in urging that "syntax is not semantics". Where he is not justified is in his faith that an explanation of semantics will be possible for humans but not for computers. After all, there is no obvious reason why biological "wetware" should by any better at imbuing internal structures with semantic significance than silicon "hardware".

A natural suggestion at this point is that the "aboutness" of internal structures is something to do with their causal relations to external objects. Thus the array of switches inside the robot-chimp's head could be argued to be about bananas precisely because it was caused by the presence of bananas, and because it caused the chimp to go back to the bananas when it was hungry. As it happens, this suggestion is rather less straightforward than it seems. But if there is anything to this line of thought, then clearly it will work as well for silicon brains as our spongy ones. A corollary worth noting is that familiar office computers, and even big laboratory computers, won't in any serious sense have states with intentional powers. For unlike the robot-chimp's brain, and indeed unlike our brains, commercial digital computers aren't designed to interact with specific features of the external world, but simply to be easily adjustable labour-saving devices. It is probably also worth noting that nothing in all this is supposed to show that computers are conscious. One of the Radio 4 announcers advertised Searle as demolishing the theory that "thermostats have feelings". This is no doubt how many people have understood Searle, and indeed what gives the Chinese room argument, with its talk of *understanding* Chinese, and consequent connotations of some kind of inner illumination, much of its appeal. But that's not how the Chinese room argument is intended. Searle says very little about consciousness, and certainly his explicit statements present the Chinese room argument as about aboutness, not about feelings.

In the third lecture Searle continues his attack on cognitive science. Here the target is not just the strong claim that computers can think, but any attempt to use the computer analogy to illuminate the workings of our brains. Searle is sceptical about the existence of identifiable levels of cerebral organization, analogous to "software" programmes, mediating between our neuronal "wetware" and our mental life. But he admits that future research might yet prove him wrong, and indeed mentions examples of existing computer-influenced work, such as David Marr's book on *Vision*, which many would argue have already discredited his pessimism.

After this Searle elaborates his own picture of the sciences of man, in which, as one would expect, the intentionality of human thought plays a central role. There is a good argument in the fifth lecture. Many contemporary philosophers believe in the possibility of "special sciences" in the human realm, whose categories are indefinable in physical terms, but which nevertheless deliver some degree of predictive generalization. Searle points out that this won't do. For if there are general truths about what will happen, say, in given economic circumstances, then, as Searle puts it, "the molecules will have to be blowing in the right direction" on all such occasions. But this would be an absurd coincidence, given the supposed mismatch between the concepts of economics and the concepts of physics. Searle takes this argument to demonstrate a radical division between understanding in terms of intentionality and explanation in terms of predictive generalizations. But the argument could as well be run the other way. Those impressed by the extent to which economic life, and indeed interpersonal life, is predictable, might conclude instead that the concepts of the human sciences must after all be commensurate with those of physics.

The final lecture is about free will. Here Searle will disappoint those who expect him to use the supposed divide between mental and physical explanation to find a niche for human freedom. For Searle recognizes that being undetermined under a mental description doesn't give us free will, given that we are also physical beings and what happens is fixed by physics. He does allow that, as agents, we have an inescapable belief in our own free will. But, he nevertheless concludes, it ain't really so. One only wonders why someone who is so unmoved by our prejudice that we have special powers of free choice is so certain that we humans have special powers of intentional thought.

Understanding Identity Statements by Thomas V. Morris has recently been published (152pp. Aberdeen University Press. £12.50; paperback, £8.50. 0 08 030388 9). In Part One Morris analyses identity statements under such headings as "The Objectual Analysis" and "Identity, necessity and Information", while in Part Two he examines Leibniz's Law and the difficulties it raises in respect of cross-category identities."

Papineau, David. "The significance of sqiggles." The Times Literary Supplement, no. 4263, 14 Dec. 1984, p. 1442. The Times Literary Supplement Historical Archive, link.gale.com/apps/doc/EX1200446426/TLSH?u=tlsacc&sid=bookmark-TLSH. Accessed 23 Sept. 2023.