

Philosophy of mind

The Reith Lectures for 1984 were delivered by Professor John Searle under the title "Minds, Brains and Science". Stuart Sutherland discusses some of the issues raised in the broadcasts.

FOR SOME time, one of the main occupations of philosophers has been to tease scientists. Richard Rorty and others have argued that science is valid only in terms of the conventions adopted at a given time, while Paul Feyerabend goes further and believes that within science (and presumably also within philosophy) "anything goes". In the recent series of Reith Lectures, broadcast on BBC Radio 4, Professor John Searle confined his teases to artificial intelligence: as befits a philosopher forged at Oxford and honed at Berkeley, he produced a carefully thought-out view of minds, brains and science. Although provocative, he was for the most part refreshingly clear, so much so that it is possible to refute him on certain points.

Searle believes that mental terms are simply another way of describing the activities of the brain, a view that has the merit of plausibility, though not of originality. In the lectures he likened the emergence of mental phenomena, such as consciousness, intention or desire, to the emergence of such properties as solidity or transparency which are produced by the arrangement of the molecules in a material. But he ignored an important difference: we understand rather fully how certain molecular arrangements can give rise to the higher level attributes of solidity or transparency, but we have no idea whatever of how any arrangement of neurones could give rise to consciousness. (This point was also made by Colin McGinn in an otherwise rather disappointing discussion of the talks during which Searle and his opponents rarely seemed to meet.) One may conclude that Searle has merely by-passed the mind-body problem; he has not made it go away.

Searle attacked three of the claims made by some of the AI fraternity. First, he maintained that no computer system could ever exhibit consciousness, since consciousness is a property only of brains. This claim is reasonable, although Searle defended it by assertion rather than argument.

His second counter-claim was that computers can never exhibit mental activity — they can never have knowledge, even unconscious knowledge. His main argument stemmed from his well-known analogy of the Chinese room. He asked the listener to imagine himself sitting in a room which contains a supply of Chinese symbols and a rule book in English for manipulating them. When the person in the room is handed a string of Chinese symbols, he assembles a sequence of his own symbols in a manner dictated by his

rule book. If the rules were comprehensive enough, and if the strings he received were genuine questions in Chinese, then the strings he assembled could be correct answers. The rules might correspond to those embedded in a computer program that answered questions in Chinese. But clearly the person manipulating the symbols does not understand Chinese nor *a fortiori* does the computer program. To



John Searle — a mind of his own

understand Chinese a person must not merely be able to manipulate the symbols according to rules, he must know the meaning of those symbols. Searle argued that the Chinese room analogy shows that so long as someone or something, whether it be a person or a computer, is automatically following a set of rules (a syntax) while unaware of their meaning (a semantics), he or it cannot be said to have a knowledge of whatever it is the symbols refer to.

The argument is highly ingenious, but it raises two problems. First, if the difference between computer programs and brains is that the latter have semantic knowledge while the former do not, it was surely incumbent on Searle to consider how it is that the brain acquires its semantics; but he did not even mention this question. Second, we would be more tempted to think of computers as having knowledge (a semantics), if they interacted with the world through sensors and mechanical limbs, if they could "learn" from these interactions and if their behaviour was governed by goals in the way that human behaviour is. Searle's reply to this point was too skimmed to carry conviction.

Searle's third and least plausible claim was that not only can computers never have mental activities, they cannot even simulate such activities. Now if the brain operated according to rules, it would clearly be possible to simulate its workings in computer programs. But since the brain and mental activity are one and the same according to Searle, he was driven to the desperate device of denying that the brain operates according to rules: indeed he suggested that no unconscious calculation is undertaken by the brain. But, to take one example, it is known that certain neurones in the striate cortex compute a function on parts of the retinal image. His comment on how the brain does mediate thoughts was distinctly unhelpful — "The brain just does them". The extent to which the brain carries out computations can surely be decided only by investigation not by philosophical *fiat*, and if it does not work like this, it is extremely hard, as Searle's own remarks made clear, to think of any other way in which it might operate.

Searle also attempted to show that there can be no social "science", at least in the meaning of the term arrogated by physical scientists. He claimed that there can be no "systematic correlation between phenomena identified in social and psychological terms and phenomena identified in physical terms". Anything — paper, metal, pigs — can serve as money: it depends only on the attitudes of the community. But this surely does not imply that there are not highly specific and, at least in theory, identifiable processes going on in the brain when someone uses the concept of money. It only means that these processes are likely to be complicated and to be interdependent on many other processes that underlie people's systems of belief. Moreover, there is an implicit contradiction in denying that certain concepts cannot have a systematic representation in the brain while claiming that mental activity *is* the working of the brain.

In his final lecture, which was on freedom of the will, Searle acknowledged defeat. If the mind is merely the brain under another guise, and if the physical world is determined, then there appears to be no room for freedom of the will. Searle said, with commendable honesty, "For reasons that I don't really understand, evolution has given us a form of experience of voluntary action where the experience of freedom is built into the very structure of ... voluntary ... behaviour". It is perhaps curious that he did not begin his lectures by saying the same thing about consciousness over which the same problem arises in an equally acute form. Despite Professor Searle's sleight of hand, the problem of consciousness continues to defeat the human mind. □

Stuart Sutherland is Director of the Centre for Research on Perception and Cognition at the University of Sussex.