At a time when computer technology is advancing at a rapid pace and when software developers are convincingly hawking their products as having artificial intelligence, the inevitable question has begun to take on a certain urgency: Can a computer think? Really think? In one form or another this is actually a very old question, dating back to such philosophers as Plato, Aristotle, and Descartes. And after nearly 3,000 years the most honest answer is still uncertain. After all, what does it mean to think? On the other hand, that is not a very satisfying answer. However, with his paper: Minds, brains and programs published in 1980, John Searle has had a huge impact on the artificial intelligence issue worldwide. This essay will focus on Searle's idea that computers are incapable of being conscious, and then analyse whether Searle is right in terms of his three main efforts: a critique of computationalism and strong Artificial Intelligence (AI); the development of a theory of intentionality; and the formulation of a naturalized theory of consciousness.

At the first place, the best-known example of Searle's critique of computationalism and strong AI is his Chinese Room Argument. The argument (1980) goes as follows: Searle supposes that, many years from now; 'we have constructed a computer, which behaves as if it understands Chinese.' In other words, the computer takes Chinese symbols as input, consults a large look-up table (as all computers can be described as doing), and then produces other Chinese symbols as output. 'Suppose that this computer performs this task so convincingly that it easily passes the Turing Test.' In other words, it convinces a human Chinese speaker that it is a Chinese speaker. All the questions the human asks are responded to appropriately, such that the Chinese speaker is convinced that he or she is talking to another Chinese speaker. The conclusion proponents of strong AI would like to draw is that the computer understands Chinese, just as the person does. Yet, Searle asks us to suppose that he is sitting inside the computer. In other words, he is in a small room in which he receives Chinese symbols, looks them up on look-up table, and returns the Chinese symbols that are indicated by the table. Searle notes, of course, that he does not understand a word of Chinese. Furthermore, his lack of understanding goes to show, he argues, that 'computers do not understand Chinese neither, because they are in the same situation as he is. They are mindless manipulators of symbols, just as he is - and they do not understand what they are saying, just as he doesn't.'

Searle's Chinese Room Argument seems to be logical initially. Yet, in a view of a system, it is clear that though Searle himself does not understand Chinese in the thought experiment, it is perfectly correct to say that Searle plus look-up table understand Chinese. In other words, the entire computer would understand Chinese, though perhaps the central processor or any other part might not. It is the entire system that matters for attributing understanding. Moreover, it can be noted that the reason we do not want to attribute understanding to the room, or a computer as described by Searle is that the system does not interact properly with the environment. This is also a reason to think the Turing Test is not adequate for attributing thinking or understanding. If, however, we fixed this problem - i.e. we put the computer in a robot

body that could interact with the environment, perceive things, move around, etc. - we would then be in a position to attribute understanding properly. Additionally, the person in the Chinese room would have to shuffle not just a few slips of paper but millions or billions of slips of paper. It would take him years to answer a question, if he could do it at all. In effect, Searle is postulating mental processes slowed down by a factor of millions, so no wonder it looks different. Searle's reply-that he could memorize the slips of paper and shuffle them in his head-sounds plausible enough. But it dangerously undermines his whole argument: once he memorizes everything, doesn't he now understand Chinese in the same way he understands English?

Secondly, the theory of intentionality plays an important role in Searle's philosophy. Searle (1983) analyses the intentional state as 'consisting of a representative content in a psychological mode.' Although many representative contents consist in an entire proposition, many do not, and it is not necessary that they do. Searle also analyses intentional states in terms of their directions of fit (which can be world-to-mind, mind-to-world, or null) and directions of causation (which can be mind-to-world or world-to-mind). An important feature of Searle's theory of intentionality is something he calls 'The Background'. The Background is theorised to be a set of skills, capacities, and presuppositions that, while being non-representational, makes all representation possible.

It is clear to everyone that when Searle uses the word 'intentionality', he is not just talking about an obscure technical matter. In this context intentionality is virtually synonymous with mind, soul, spirit, or awareness. However, the comparison is unfair. The programs that Searle demonstrated a very crude kind of understanding at best, and no one in AI seriously claims anything more for them. Even if they were correct in principle, genuine humanlike understanding would require much more powerful machines and much more sophisticated programs. As Dennett (1991) has pointed out, 'For Searle, intentionality is rather like a wonderful substance secreted by the brain the way the pancreas secretes insulin.' And make no mistake: Searle's concept of intentionality does require a biological brain. He explicitly denies that a robot could have intentionality, even if it were equipped with eyes, ears, arms, legs, and all the other accoutrements it needs to move around and perceive the world like a human being. Inside, he said, the robot would still just be manipulating formal symbols. So far as natural selection is concerned, Searle's robot is just as fit for survival as those of us with Searle style intentional brains. Evolution would make no distinction. Indeed, from a biological point of view, intentionality is irrelevant, as useless as the appendix.

Last but not least, Searle's theory of consciousness is a key factor in his whole argument. The basic principle grounding Searle's theory of consciousness is that consciousness is irreducible. For Searle, consciousness is essentially a first person, subjective phenomenon, and thus talk of conscious states cannot be reduced or eliminated in favor of third-person, objective talk about neural events. Searle (1995)

argues that ‘Any such attempt at reduction simply misses the essential features of conscious states -- that is, their subjective qualities.’ The problem with Searle’s this argument is it is an unfair reasoning. In fact, it rules out every possible way to inspect and verify it. If consciousness is irreducible because the subjectivity of consciousness is lost when we try to make this reduction, then consciousness is unknowable by reductionist methodology. But, notice the irreducibility principle is not a conclusion of a deductive argument, but an assertion of a common sense notion, according to Searle. It is put forth as the result of a thought experiment in which he makes the assumption that subjective consciousness is indivisible, because doing this would contradict the very nature of being subjective. Again, this is assumed not proved. Searle is basing his argument on assuming the conclusion. If Saerle’s this argument is right, I can only make sure that I am conscious. Therefore, nobody knows whether other human beings are conscious, not to speak of computers. So this theory cannot deny that computers are capable of being conscious.

After Searle failed to convince us of his definition about the consciousness of AI, the problem is how can we relate ‘consciousness’ with ‘artificial intelligence’. Daniel Dennett has given us a very reasonable way to thinking. He (1991) argues that ‘The best reason for believing that robots might some day become conscious is that we human beings are conscious, and we are a sort of robot ourselves.’ In his theory, human beings are just extraordinarily complex self-controlling, self-sustaining physical mechanisms, designed over the aeons by natural selection, and operating according to the same well-understood principles that govern all the other physical processes in living things: digestive and metabolic processes, self-repair and reproductive processes, for instance.

Undoubtedly, there must be lots of arguments against this theory. Firstly, Old-fashioned dualism can argue that robots are purely material things, and consciousness requires immaterial mind-stuff. Take Clone technology as example; before 23rd February 1997, nobody would believe that two ewes could reproduce a lamb without a ram. However, the birth of Dolly proved that impossible can be possible. So why cannot purely material things cannot be conscious? This doubt can be easily answered with the passage of time. The advanced technology can always prove that most people are lack of imagination. Furthermore, people also can dispute that robots will always just be much too simple to be conscious. After all, a normal human being is composed of trillions of parts (if we descend to the level of the macromolecules), and many of these rival in complexity and design cunning the fanciest artefacts that have ever been created. We consist of billions of cells, and a single human cell contains within itself complex "machinery" that is still well beyond the artifactual powers of engineers. We are composed of thousands of different kinds of cells, including thousands of different species of symbiotic visitors, some of whom might be as important to our consciousness as others are to our ability to digest our food. If all that complexity were needed for consciousness to exist, then the task of making a single conscious robot would dwarf the entire scientific and engineering

resources of the planet for millennia. Facing up to reality, it is impossible to make a conscious robot under the material and technical conditions now available. Yet it does not mean we cannot do it in the future.

Judging from the above, Searle's theory gives us an apparent idea that computers will never think like human beings. But his views about artificial intelligence are too restricted. By using completely humanity measurements, he cannot jump from the intrinsic ideas, which always point to analogies between human being and artificial intelligence. It will not benefit the advance of the human society. On the other hand, till now, nobody can prove that artificial intelligence will never be capable of being conscious in an unquestionable way. Therefore, just like Dennett's idea, we still can say that making conscious artificial intelligence is possible in the future.

Reference

Daniel, C. Dennett (1991), Consciousness Explained, Penguin Books, New York.

Searle, J. (1980) Minds, brains, and programs, Behavioral and Brain Sciences, 1: 417-24.

Searle, J. (1983) Intentionality: An Essay in the Philosophy of Mind, Cambridge University Press, New York.

Searle, J. (1997) The Mystery of Consciousness, New York Review Press, New York.

Bibliography

Armin Laux and Heinrich Wanshong (eds.) (1995) Knowledge and Belief in Hhilosophy and Artificial Intelligence, Aksd. Verl, Berlin.

Margaret, A. B. (1990) The Philosophy of Artificial Intelligence, Oxford University Press, UK.