

## Searle's Chinese Room

Searle's thesis regarding the claims put forth by supporters of Strong AI is: The implementation of a computer program, by itself, is not sufficient to guarantee intentional content, and thus it cannot duplicate a fully functioning mind.

In this paper, I will assess Searle's argument for this thesis, by way of his Chinese room analogy. In doing so I will also mention what needs to be added, in addition to the computer program, to give the machinery running the program actual intentional content\*. Before starting I will define some key philosophical terminology that I will use in discussion of Searle's thesis and arguments.

**Weak AI** is the idea that computers are a useful tool in cognitive science because they provide a means by which we can study the mind more closely by use of computer simulations.

**Strong AI** claims that computers can, not only, simulate the mind, but, with the right programs, act as fully functional minds all on their own.

**Syntax** has to do with rules governing formal symbol manipulation. Computer programs are purely syntactical because their only purpose is the manipulations of the formal symbols, 1 and 0.

**Semantics** has to do with meaning. Words are a representation for something other than the set of symbols that comprise them.

**Intentionality** is the feature of aboutness or directedness that certain mental states have. Like words, some mental states are a representation of something other than themselves, such as objects or events.

Searle has no objection to the claims of 'Weak AI', as it has no ambition to give computers fully functional minds. However, he strongly objects to Strong AI because it makes the claim that computers, with the right set of programs can have fully functional minds. He argues against Strong AI by way of his Chinese room analogy:

Suppose that Searle, a native speaker of English, is locked in a room with a pen, paper, and some sort of rulebook written in English. After some time, a person outside the room slips a paper under the door with a set of symbols on it. Searle looks at the paper, consults the rulebook, which tells him that this set of symbols correspond to a different set of symbols, which he is then required to write on a piece of paper and slip back under the door. Searle does so, and after some time another paper arrives, again with a set of symbols on it. Searle then repeats the same process as before: he consults the rulebook, writes the corresponding set of symbols down, and slips it back under the door. This event repeats itself many times, and is only occasionally interrupted by a piece of paper that has English writing on it. The papers containing English instruct Searle to answer some questions in English, which Searle does using his native understanding of the English language, and like the symbol papers, he is also to slip his answers under the door.

Unbeknownst to Searle, like the papers with English, the papers with the symbols that he received were questions in Chinese and the corresponding papers he sent out were the answers to those questions. To anyone outside the room, it would appear that the Searle has a fluent understanding of English and Chinese. However, while Searle is a native speaker of English, he has no comprehension of Chinese in any form, written or verbal. He couldn't even recognize Chinese symbols apart from Japanese or even meaningless squiggles. All Searle did was follow a very cleverly written rulebook that dealt with symbols, which were meaningless to him. Meanwhile, in the case with the English papers, he used semantics of the words to answer the questions along with the syntax of the English language.

A supporter of Strong AI would be forced to claim that there is absolutely no difference in the way Searle answers the English and the Chinese questions and that is proof that a program

(the rulebook), along with the computer (Searle), is a sufficient condition for a fully functional mind.

I believe Searle successfully argued against this specific claim. By interjecting himself as the computer, Searle showed that even something with proved capacity for semantics and intentionality cannot get a further understanding of anything simply by doing syntactical operations. Searle's running of the program might have produced the correct answers to the Chinese questions, but running the program in no way contributed to his understanding of Chinese. This is because formal manipulations of symbols have nothing to do with semantics. Moreover, his mental states concerning these symbols had no intentionality, besides referring to themselves, because to him, these symbols were meaningless. His argument can be framed in this way: (1) Computer programs are purely syntactical. (2) The mind has intentional contents and understands semantics. (3) Syntax is not sufficient for semantics or intentionality. (4) Therefore, a computer program cannot be considered a mind, on its own. (1) Is given by the Strong AI supporters, (2) is inferred from many examinations of the mind, (3) is demonstrated in the Chinese room analogy and (4) follows from (1) (2) and (3).

With this argument, I don't believe Searle is saying that computers and programs can never add up to a functional mind, they just can't add up to a functional mind all on their own. They are missing a mechanism to account for intentional content and the semantics of the symbols they are dealing with. Our minds have this mechanism, and it is the crux of our consciousness. How we artificially add this mechanism is still a mystery.

In this paper I have assessed Searle's argument for the idea that the implementation of a computer program, by itself, cannot duplicate a fully functioning mind. I hoped to have shed some light on why t line of reasoning supported by Strong AI should be abandoned.

## Works Cited

Searle, John R. "Minds, brains and programs." Behavioral & Brain Sciences 3 (1980):  
417-457.