Strong Artificial Intelligence

Many people who work in cognitive science and in the philosophy of mind think that the most exciting idea of the past generation, indeed of the past two thousand years, is that the mind is a computer program. Specifically, the idea is that the mind is to the brain as the computer program is to the computer hardware. This view I have baptized "Strong AI," but it is sometimes called "Computer Functionalism.' In this lecture I explain the appeal of this view, but I also subject it to a decisive refutation—The Chinese Room Argument."

I.   The Theoretical Basis of Strong AI.
In order to explain the appear of Strong AI, I have to introduce five somewhat technical notions.

A. Turing Macines
1. The idea of a Turing macine is an abstract, mathematical notion. For practical purposes, ordinary computers—the kind that you buy in a store—are Turing machines.
2. The remarkable feature of a Turing machine is that it performs only four operations: Print "0;" erase "1;" print "1," erase "0;" move on squar left; move one square right. Modern machines perform these operations at the rate of millions per second.

B. Algorithm—an algorithm is a systematic procedure for solving a problem in a finite number of steps. Computer program s are algorithms.

C. Church's Thesis—this thesis states that any algorithm can be implemented on a Turing machine. For every computable function, there is a Turing machine that can compute that function.

D. Turing's Theorem—this theorem states that there is a

Universal machine with can simulate the behavior of any other Turing machine.

    E. The Turing Test—this test states that if an expert cannot distinguish the behavior from that of a human, then the machine has the same cognitive abilities as the human.

II.   Strong AI
If we put all these together, we get the idea that the brain is a Universal Turing machine and that human cognitive abilities are computer programs. We test this with the Turing test and come to the conclusion that artificial intelligence is, in principle, capable of creating minds.

III.   Refutation of Strong AI
    A. Strong AI and, with it, the Turing test, are subject to a decisive refutation, the Chinese Room Argument.

    B. The Chinese Room Argument claims that a monolingual English speaker who is locked in a room with a set of computer rules for answering questions in Chinese would in principle be able to pass the Turing test, but he would not thereby understand a word of Chinese. If the man doesn't understand Chinese, neither does any digital computer.

IV.   Attacks on the Chinese Room Argument. According to the systems reply, the man in the room does not understand Chinese, but he whole room does.