The Turing Test and the Chinese Room Argument

Angel Garrido, Facultad de Ciencias, UNED, Madrid, Spain

Abstract

As it is well-known, the problems in AI may be many times very related to Philosophy of Mind, and perhaps because this reason may be in essence very disputables. So, for instance, the famous question about if *Can a machine to think?* It will be the decissive question, but for many people it will be a non-sense question. So, two of the more fundamental topics which usually may be considered according this line includes the *Turing Test* and the *Chinese Room Argument.*

Keywords: Graph Theory, Knowledge Representation, Philosophy of Mind, Heuristics, AI.

Mathematics Subject Classification: 68R10, 68R05, 78M35.

1. Introduction

The answer to the question about if will be considered a Machine, or a Computer, as a Mind, had produced a very copious scientific and philosophical literature. Initially, with the seminal paper of Alan Turing [10], and then after the attempt of refutation of John Searle [6-9], by the so-called *Chinese Room Argument.* But after the ideas expressed by Turing, some of the participants in the Darmouth Conference (1956) giving the trend known as Strong A I; represented by authors as, for instance, may be Marvin Minsky [4].

2. Turing Test

The very popular *Turing´s Test* [10] suppose that an interrogator is connected to one person and one machine, via a terminal, and therefore cannot see their counterparts. Its task is to find out which of the two contestant candidates will be the machine, and which will be the human, only by their answer to questions. If the interrogator cannot make a decision within a reasonable time, then the machine is considered to be intelligent. The initial Turing estimate as reasonable margin was five minutes, but the length of this interval is irrelevant.

We comment now some about the Turing ideas. In its paper [10], written in 1950, Turing proposed an operational definition of intelligence that was chosen to be equally applicable to humans and machines. The test was described as a generalization of a parlor game Turing called the *imitation game.* The basic

[1] AMO - Advanced Modeling and Optimization. ISSN: 1841-4311

idea of this game is that an interrogator attempts to determine the sex of one contestant by asking questions and receiving answers in writing. The goal of at least one contestant answering these questions is to cause an interrogator to make the wrong determination. No information is available to an interrogator other than the written answers, and at least one of the contestants answering questions is not obligated to tell the truth.

The Turing Test [10], in its original form, is to replace by a machine one of the contestants of the imitation game who is not required to be truthful. If the results of the game are unaffected by the presence of this machine, then this machine is said to be capable of thought. In other words, a machine that is indistinguishable from a human being solely on the basis of "written" interaction is considered to be capable of thought.

In his paper, Turing motivates this definition by nothing that such written interactions are generally considered sufficient for assessing the intelligence of a person. He then discusses how this definition could be correct, provided you are willing to accept that the essential nature of intelligence is captured by a formal system. But if the Mind can be viewed as a formal system, then you must accept that it is a Turing Machine (TM), or show that the Church´s Thesis is false. Recall that according such Thesis, all forms of computability are TMs. And also that a TM is a formal definition of a computing engine; so, it may be considered as an algorithm.

Turing's paper is unusually informal. However, while Turing's definition is an obvious choice, it is quite possible to raise well-considered objections to it. The most important argument against the Turing Test is that indeed only provides a test for human intelligence.

3. Chinese Room Argument

John Searle´s [6] argument, proposed in 1980, is intended to show that implementing a computational algorithm that is formally isomorphic to human thought processes cannot be sufficient to reproduce thought. Something more is required. So, it will be considered an attempt of refutation of both, Turing Test and Functionalism. Searle published his argument in an issue of *Scientific American* [8].

It begins with this hypothetical premise: Suppose that AI research has succeeded in constructing a computer that behaves as if it understand Chinese. It takes Chinese characters as inputs, and produces other different characters, which it presents as output, by following the instructions of a computer program. Searle summarized in this way: Imagine a native English speaker who knows no Chinese, locked in a room full of boxes of Chinese symbols (a data base), together with a book of instructions for manipulating the symbols (program). Imagine that people outside the room send in other Chinese symbols which, unknown to the person in the room, are questions in Chinese (the input). And imagine that by following the instructions in the program the man in the room is able to pass out Chinese symbols which are correct answers to the questions (the output). The program enables the person in the room to pass the Turing Test for understanding Chinese, but he does not understand a word of Chinese.

Searle goes on to say that "the point of the argument is this: if the man in the room does not understand Chinese on the basis of implementing the appropriate program for understanding Chinese, then neither does any other digital computer solely on that basis because no computer, qua computer, has anything the man does not have."

Searle attempts with this to refute a certain conception of the role of computation in human cognition. To understand this argument, it will be necessary to distinguish among Strong AI, and Weak (or Cautious) AI. According the first of them, any system that implements the right computer program with the right inputs and outputs thereby has cognition in the same sense that human beings. According the second of them, the computer is nothing more than a useful tool in studying human cognition, as in studying many other scientific domains.

The principle of strong AI is that an appropriately programmed computer is actually thinking, as opposed to simulating thought. This principle is equivalent to accepting the Turing test as a definition of thought, always provided a computer can actually pass the test.

Hence, the contrast is that according the Strong version, the correct simulation *is* really *a mind.* Whereas according to the Weak version, the correct simulation is only *a model of mind.*

Its proof contains three premises and one conjecture, by

AXIOM 1. *Implemented programs are syntactical processes,* i.e. computer programs are formal (syntactic).

AXIOM 2. *Minds have semantic contents,* i.e. human minds have mental contents (semantics).

AXIOM 3. *Syntax by itself is neither constitutive of nor sufficient for semantics.*

CONCLUSION. *The implemented programs are not constitutive of nor sufficient for minds.*

Therefore, according Searle, *Strong AI is false.*

But the Chinese Room Argument may be expressed by two basic principles, each of which will be stated in four words. Firstly, *Syntax is not Semantics,* because syntax by itself, is not constitutive of semantics, nor by itself sufficient to guarantee the presence of semantics. Secondly, *Simulation is not duplication.*

The Chinese Room Argument comprises a thought experiment, and associated arguments by Searle [6-9]. It attempts to show that any symbol-processing machine (in particular, any computer) can never be properly described as having a “mind”, or “understanding”, regardless of how intelligently it may behave.

Because we suppose both in two closed, isolated and different rooms, having a book or computer program, and disposing in sufficient number of pencils, erasers, papers, and so on, but without any knowledge on Chinese, receiving from the exterior a paper each time, with Chinese characters, processing then according to the program´s instructions, and producing an answer in Chinese characters as output. Each is simply following a program, step-by-step, which simulates intelligent behaviour. We cannot describe what the machine is doing

as “thinking”. And neither the human operator understand a word of Chinese. Therefore, we must infer that computer does not understand Chinese either.

Thus, Searle affirms that Strong (but not Weak) AI is a clear nonsense.

4. Conclusions

The Chinese Room Argument, as it is due by Searle [6], have received many times acceptation, but also different refutations. However, it give us a subtle argumentation. For this, many authors as the british Roger Penrose [5] have viewed the horizon of the Strong A I as a mere "cul de sac", a wrong line of thought without sense.

My personal opinion is that, as in many situations of life, the middle will be perhaps the best election. Nor the too optimistic strong line, neither the too cautious weak line. The underneath troubles will be perhaps because their ideological, ethical and religious implications, with "secondary effects". While ANNs (Artificial Neural Networks) provide us with a potential technological platform for thinking machines, the technology is too nascent as yet. Also we must to mention here the very hard confrontation of Connectionists against Computational Mind theorists. But it may be the argument of a new and more detailed paper. Recall that *Connectionism* is the name for the computer modeling approach to information processing based on the design or architecture of the brain. And that the *CTM* (or Computational Theory of Mind) is the view that the human mind is an information processing system, and that thought only will be a form of computation.

References

[1] Barr & Feigenbaum: *The Handbook of Artificial Intelligence.* M. Kaufmann,1981.

[2] Garrido: "Logical Foundations of Artificial Intelligence". *FotFS V International Conference,* Institut für Angewändte Mathematik, University of Bonn, 2004.

[3] Ginsberg: *Essentials of artificial intelligence.* M. Kaufmann, San Francisco,1993.

[4] Minsky: "A Framework for Representing Knowledge". *The Psychology of Computer Vision.* McGraw Hill, 1975.

[5] Penrose: *The Emperor ´s New Mind: Concerning Computers, Minds and the Laws of Physics.* Oxford University Press, 2002.

[6] Searle: "Minds, Brains, and Programs". *The Behavioral and Brain Sciences,* vol. 3, pp. 417-457, 1980.

[7] Ibid.: *Minds, Brains and Sciences.* Harvard University Press, Cambridge, Massachusets, 1984.

[8] Ibid.: "Is the Brain´s Mind a Computer Program?". *Scientific American,* Vol. 262, No.1,pp. 26-31, 1990.

[9] Ibid.: *Intentionality: An Essay in the Philosophy of Mind.* Cambridge University Press, 1983.

[10] Turing: "Computing Machinery and Intelligence". *Mind,* vol. 59, No. 236, pp. 433-60, 1950.