Towards Artificial Intelligence. Possibilities and psycho-philosophical issues.

Ioannis D. Anyfantis
European Agency for Safety and Health at Work (EU-OSHA)

TOWARDS ARTIFICIAL INTELLIGENCE. POSSIBILITIES AND PSYCHO-PHILOSOPHICAL ISSUES.

Ioannis Anyfantis
MEng Electrical & Computer Engineering (Aristotle University of Thessaloniki, GR), MSc Communication Engineering (UMIST, UK), MSc Technology Management (Manchester School of Management, UK)

Abstract

During the last 30 years we all witnessed an incredible advance in digital technology. Modern technologies give us incredible computing power in our hands. Software development has provided intelligent ways of doing things. However, how far are we from the development of Artificial Intelligence machines? There are also of number of Psychological and Philosophical issues arisen by such attempts. Artificial Intelligence will be limited because we don't fully understand the brain. Thus it is argued that until we can dissect the human mind accurately describe the various elements that make up our very being and consciousness, we cannot expect to artificially replicate intelligence.

Keywords: Artificial Intelligence, artificially replicate intelligence, human mind, Philosophical Issues of AI, Psychological issues

Introduction

During the last 30 years we all witnessed an incredible advance in digital technology. Microprocessors become more powerful within just a few months, while the frequency of operation has reached the microwave band. More than one computing devices can be seen in the vast majority of the houses of every developed country, while modern smartphones give us incredible computing power in our hands “on the go”. So where is this going to stop and is it possible to reach a level of such a computing power that would be possible to simulate even a simple “version of brain” by the use of a capable software? In other words, is it only a matter of time before artificially intelligent machines are constructed?

Conceptually the creation of an artificial intelligence machine is by no means a recent idea. The first paper written on the subject of the electronic brain was published in 1943 by American scientists Warren

McCullock and Walter Pitts (McCullock and Pitts, 1943) on the subject of building electronic circuits to mimic neural networks, although this notion of artificial intelligence or artificially created humans can be traced back through time much earlier, to Homer's Iliad and Hellenic Egypt and later, to Von Kempelen's Turk and Babbage's Difference Engine (generally considered to be one of the first computers). However the increase in the power and ability of computers in recent years has brought the possibility of this attaining that concept further to reality.

What is Artificial Intelligence?

The term "Artificial Intelligence" was coined in 1956, by John McCarthy at the Massachusetts Institute of Technology (McCarthy, 1956). Actually, there are a lot of definitions for AI. One of those would say that AI is the branch of computer science concerned with making computers behave like humans. From a different point of view AI is the use of programs as tools in the study of intelligent processes, tools that help in the discovery of human abilities, like the thinking procedures and epistemological structures employed by intelligent creatures. More specifically, the ability to use language, the ability to perceive the world via sense data, or the ability to think, for instance, using a combination of traditional approaches to these topics and the use of computers.

Some of the areas in which AI is developed are: a) Systems: programming computers to make decisions in real-life situations (for example, some expert systems help doctors diagnose diseases based on symptoms). b) Natural language: programming computers to understand natural human languages. c) Neural networks: Systems that simulate intelligence by attempting to reproduce the types of physical connections that occur in animal brains d) Robotics: programming computers to see and hear and react to other sensory stimuli, e) Games playing: programming computers to play games such as chess and checkers expert. Thus, an AI machine could be a machine that could either play chess, or behaves like a human being. In other words there are many kinds of AI machines of several complexities that have of course different requirements in either software or hardware (Webopedia).

For example, Deep Blue, a chess computer created by Murray Campbell and colleagues at IBM, have beaten the best human chess player, Garry Kasparov (HSU, 2002). If playing chess requires intelligence then Deep Blue is actually an artificial intelligent machine. However chess could possibly described as a game where you could win if you were able to estimate and take into consideration any possible combination of a number of next moves. When a chess computer plays chess, it does not think about the move it makes but performs a series of calculations to formulate the appropriate move, based on the positions of the pieces on the board. Human chess players use judgement, previous experiences and also recognition of patterns to decide what moves to make. In that way, if it is considered as something like running an algorithm, then Deep Blue is not an AI machine. Actually, Alexander Kronrod, a Russian AI researcher, said ``Chess is the *Drosophila* of AI.'' That essay will concentrate on AI machines that will possibly have a way of "thinking" and reacting, similar to that of a human or generally of an animal.

The technological point of view

The core of an AI machine by today standards could be considered as an "effective procedure" – a fundamental concept to computer science – that constitutes of a computer program (software), together with an appropriate machine (hardware). People use to say, "programs tell computers what to do", and that "a computer can only do what you tell it to do". The program actually denotes a set of rules unambiguously specifying certain processes, which can be carried out by a machine processor built in such way as to accept these rules as instructions determining its operations. On the contrary, many programs use inferential procedures which (like human thought) work only *reasonably* well, *reasonably* often. Moreover, a programmer cannot always foresee every step the program will make. By taking that into consideration, some programs are so written, that certain decisions are left open, to be taken by the computer itself when the program is running, in light of the particular circumstances. And not only that, but also power programming can give the ability to the program, to learn from its mistakes when it will recognize that the result was wrong. However, all these decisions to be taken have to be specified *at some level* by the programmer, since only a program can tell a computer what to do even if he can just ignore the most basic detail of its thinking. By that way the slogan "a computer can only do what you tell it to do" can be misleading since there are situations in which the program will not do *all and only* what the programmer intended it to do.

One of the successes of artificial intelligence work in the 1980s was neural networks. Neural Networks actually try to simulate the brain of animals, since they consist of a number of artificial neurons – that simulate the basic unit of every brain, as that has been proposed by biologists – on several layers. The primary problem is that, by comparison with the brain, neural networks are small. Because most networks today are simulated on traditional computers, they are "limited" by the speed of such machines to a few hundreds of thousand neurons. (a cockroach's brain, by comparison, contains about 100,000 neurones while the human brain contains about 100 billion) (Economist, 1995). That is because neural networks take a long time to train. By that way, they can remember only a few things. Based on examples, together with some feedback from a “teacher”, we learn easily to recognise the letter A or distinguish a cat from a bird. Whereas you might have to teach a child the alphabet a hundred times, a neural network can require thousands of training sessions. This is fine when another computer can handle the drills at night; less fine when a human 'teacher' has to write his name over and over again. Different tasks, such as predicting purchasing trends or spotting good credit candidates, neural networks often require thousands of examples to learn from (Fausett, 1994).

However, neural networks are quite impressive. Let’s say that a neural network is trained by a set of data, in order to control a process that is traditionally controlled by a human since lets say the quality of the product that is the result of the process depends on its colour and flavour. After the training, the network will be able to produce an equal quality product for the same process, without have taken into consideration the colour and the flavour! That is because the network “learns” the way that each parameter of the process affects the final product. And even if in some cases the product is wrong, and a human corrects the parameters of the process, the neural network could be able to learn by its mistake, and don’t repeat the mistakes (Anyfantis, 1999).

If we estimate the computational capacity of the human brain, and allow ourselves to extrapolate available processor speed according to Moore's law (whether doing so is permissible will be discussed shortly), we can calculate how long it will take before computers have sufficient raw power to match a human intellect (Heap and Thomas *et al.*, 1995). That development of the hardware, compared with faster algorithms used for the training of neural networks, could possibly lead as a first thought at the implementation of an AI machine. Such a machine, could simulate a small “brain” that would be able to be trained fast in order to “think” or just know what to do on several different situations, understanding different objects, and learn from its mistake. And that would be just the beginning! However things are not that simple. There are several limitations for the development of technology up to that point. New processors architecture has to be designed, new materials for their construction have to be used and operational frequency has to be increased in order to get close to this goal. And even in that case there would be other limitations. Just imagine that today’s microprocessors that operate at the microwave band are still slow for the training of a neural network consisted of 100.000 neurons.

The relevance of artificial intelligence

But even in case that technology would be able to produce such a powerful processor there are a lot of other concepts to be taken into consideration, like psychological implications and philosophical issues.

Psychological issues

In general, the international perception of "representation" or internal modeling is vital to artificial intelligence. The question of how knowledge can be represented in a successful and flexible way was highlighted by M.L. Minsky (1961) some years ago and is now identified as one of the first priorities of computational research. In addition to the rather general psychological notions of justification, meaning, knowledge and representation, artificial intelligence make increasing use of more specific psychological terms such as purpose, plan, hypothesis, search, inference, assumptions, and the like. So, opposing to the popular opinion,

"…artificial intelligence researchers do not deduce their work as supporting the reductionist view that psychological explanations are in principle dispensable since everything mental is "really" just something happening in the brain. On the contrary, they choose to describe and explain their programs in mentalistic (many of which are borrowed from ordinary language, although others are newly coined), because they find it more natural and illuminating to do so than to refer merely to "behavioristic" input-output correlations or to "psychological" details of machine engineering." (Boden, 1990, p. 395)

Like humanist psychology, artificial intelligence avoids the prediction of quantitatively defined variables that characterize the natural sciences. But scientific understanding does not necessarily involve prediction. The question whether one or another possible explanation is sometimes more difficult. With respect to theories expressed as programs, someone could ask if is it possible a human performance that is simulated by actually carried out (in our minds) *in the same way*. Even that some critics of artificial intelligence treat this as an all-or-none question, it is not properly regarded as that. Two systems (people or programs) may think "in the same way" when their thought is represented at one level of detail, but "in different ways" when it is illustrated at another level. The aspect of thought concerned should be specified before one can ask the question. However, detailed comparison of programs with *human* alternatives is at present hardly possible, because of our theoretical lack of knowledge of human thought process.

By that way, the problem is that there is no generally accepted way to compare the "intelligence performance" of a machine with it's nearest human "equivalent". This lack is partly a function of the difficulties of evaluating scientific evidence in general, which is a philosophical issue in the sphere of confirmation theory that is not limited to programming contexts. Moreover, it is due to our unawareness of human thinking. This is why artificial intelligence is suggestive about, rather than definitive of, the information processing details of human thought. (Boden, 1990).

Philosophical Issues

There are many arguments within the philosophical literature to show that certain things done by people could not be done by computers. Michael Polanyi has stressed the role of tacit knowing in human thought, whether "mental" speculation or "motor" skill, and regards a complete computer simulation of human thinking as impossible (Polanyi, 1964). "Tacit knowing is the fundamental power of the mind which creates explicit knowing, lends meaning to it and controls its uses. Formalization of tacit knowing immensely expands the powers of the mind, by creating a machinery of precise thought, but it also opens up new paths to intuition"[1]

Moreover, many philosophers besides phenomenologists believe that a computer could not possibly simulate or have any understanding of emotion, since it makes no sense whatever to attribute "feelings" or "consciousness" to inorganic programmed systems. They agree that emotions are not mere feelings, or bodily sensations, but contain a strong cognitive component relating to the background circumstances in which the emotion is experienced. Thus a computer could have no real understanding of emotions – no matter how plausibly it used "emotional language", – on the ground that it supposedly cannot experience feelings, since a feeling is a complex emotional response drawing on a complicated conceptual base in the mind of a person (Bloomfield, 1987).

The power of language itself is a major problem. Explaining the process of human socialization in our natural language (English for example) let alone a computer or machine language is difficult. Notions such as "hope", "friendship", "trust" and their derivatives, are expressed as words but, "any understanding of them must be fundamentally metaphoric" (Weizenbaum, 1987). We understand the meaning of these words but through our socialization and learning as opposed to definitions. Intelligence gained by computers "must always be an intelligence alien to genuine human problems or concerns" (Weizenbaum, 1987) since it will not be subjected to exactly the same socializing processes. Searle tries to prove that with his Chinese Room, where he theoretically becomes a Chinese-speaking computer despite his lack of knowledge about Chinese (Searle, 1998). The Chinese test simplifies artificial intelligence into a game of symbol

[1] Polanyi, Michael, Logic of Tacit Inference, New York: Harper, 1964, p. 18

manipulation not demonstrative of thinking.[2] A quite impressing implementation of a 'natural language artificial intelligent system', is STRAT[3]. START is a software system designed to answer questions that are posed to it in natural language. START parses incoming questions, matches the queries created from the parse trees against its knowledge base and presents the appropriate information segments to the user. However as mentioned, of lacks the full understanding of the words.

Another example could be "The Turing test" (Turing, 1950). Alan Turing discussed conditions for considering a machine to be intelligent. He argued that if the machine could successfully pretend to be human to a knowledgeable observer then you certainly should consider it intelligent. This test would satisfy most people but not all philosophers. The observer could interact with the machine and a human by teletype (to avoid requiring that the machine imitate the appearance or voice of the person), and the human would try to persuade the observer that it was human and the machine would try to fool the observer.

By that way, machine that passes the test should certainly be considered intelligent, but a machine could still be considered intelligent without knowing enough about humans to imitate a human. Daniel Dennett (Dennett, 1998), makes an excellent discussion of the Turing test and the various partial Turing tests that have been implemented, i.e. with restrictions on the observer's knowledge of AI and the subject matter of questioning. It turns out that some people are easily led into believing that a rather dumb program is intelligent.

Conclusion

The amazing rapid development of digital devices could possibly lead at the production of a super-powerful processor, which by the use of a highly sophisticated program could possibly make feasible the development of a machine that would try to simulate the human brain, judge, take decisions and do many common "natural" things which a human can do, that could make people say: "it is intelligent." From that point of view, it is only a matter of time before artificially intelligent machines are constructed.

However, the human brain is the most complex part of the human body if not the most complex subject known to human kind. Although a computer may work on a complicated series of circuits and processors it is a relatively simple item to understand in comparison to the human brain. So many important functions of the brain are beyond our understanding and have un-quantifiable properties. Thus, Artificial Intelligence is limited

[2] http://www.msu.edu/user/vattervi/turing/reg/
[3] http://www.ai.mit.edu/projects/infolab/

because we don't fully understand the brain. We understand how the brain works at the cellular level; we understand that the brain has many specialized structures and that different parts of the cortex are important to different types of thought. But there is a gap in our knowledge of the brain since we are not sure how thought occurs at the cellular level.

Perhaps the whole problem is described best, by René Descartes. "I think, therefore I am" (Descartes, 1596-1650). Implementing an artificial intelligence may require the very notion of sentience to be considered. So, to conclude with, until we can dissect the human mind accurately describe the various elements that make up our very being and consciousness, we cannot expect to artificially replicate intelligence.

References:

Anyfantis I., (1999), *Implementation of a Neural Network control system, for the control of the sugar crystallization process*, Aristotle University of Thessaloniki.

Bloomfield P. Brian, (1987), *The question of artificial inteligence:Philosophical and Sociological Perspectives*, Croom helm.

Boden Margaret, (1990), *The Philosophy of Artificial Intelligence*, Oxford Readings in Philosophy, Oxford University Press,

Dennett D., (1998), *Brainchildren: Essays on Designing Minds*. MIT Press,

Fausett, L. (1994), *Fundamental of Neural Networks*, Prentice Hall, Upper Saddle River, New Jersey

Heap, N. and Thomas, R. et al., (1995), *"Information Technology and society"* The Open University , p93

Hsu, F. (2002). "Behind Deep Blue: Building the Computer that Defeated the World Chess Champion". Princeton University Press.

McCarthy J. (1956) *What is Artificial Inteligence?*, Computer Science Department

Stanford University, .

http://www.formal.stanford.edu/jmc/whatisai/whatisai.html

Minsky, M. (1961). "Steps toward Artificial Intelligence". *Proceedings of the IRE*49: 8–1. doi:10.1109/JRPROC.1961.287775

Polanyi M., (1964), *Personal knowledge: Towards a Post-Critical Philosophy*. New York: Harper

Searle J., *(1998,)Mind, Language and Society: Philosophy in the Real World*, New York: Basic Books

The Economist Newspaper – (15 April 1995) *"More in a cockroach's brain than your computer dream of."*

Turing A., (1950), *Computing machinery and intelligence*, *Mind*

Warren McCulloch and Walter Pitts, (1943), "A Logical Calculus of Ideas Immanent in Nervous Activity", Bulletin of Mathematical Biophysics 5:115-133.

Weizenbaum, J. (1987), *Artificial Intelligence*. Information Technology: Social Issues. OU.

http://www.ai.mit.edu/projects/infolab/ailab

http://www.msu.edu/user/vattervi/turing/reg/

Webopedia - Online dictionary for Computer and Internet technology http://webopedia.internet.com/TERM/a/artificial_intelligence.html