THE EMERGENCE OF ARTIFICIAL CONSCIOUSNESS AND ITS IMPORTANCE TO REACH THE TECHNOLOGICAL SINGULARITY

Alex Fernandes da Veiga Machado[1]
Luiz Raimundo Tadeu da Silva Silva[2]
Pablo de Lara Sanches[3]

Abstract: The men admired a way to swim the fish, but today they sail faster than anyone. They'd like flying like the birds, but have been a lot higher. They searched for wisdom, now they have all the knowledge accumulated in the story available in a few clicks. Human evolution is about to meet its peak through the Technological Singularity, which can be understood as the future milestone reached at the moment that a computer program can think like a human, yet with quick access to all information already registered by society. It will not be like a man, but more intelligent than all mankind in history. So we have a big question: will this new entity has consciousness? Through a study of the levels of intelligent agents autonomy and in a timeless dialogue with Alan Turing, René Descartes, Ludwic Wittgenstein, John Searle and Vernor Vinge, we show the possibility of an artificial consciousness and that the quest for intentionality, promoted by sophisticated algorithms of learning and machine discovery, is the key to reach of Technological Singularity.

Keywords: Artificial Intelligence. Machine discovery. Turing test. Chinese Room. Intentionality.

1. Introduction

The first computer application came in 1843 and with it the concept of intelligent computer program (FUEGI; FRANCIS, 2003). Since then, these digital systems have been gaining space in our daily life, from social networks, to apps where books are indicated, even to autonomous cars.

From a philosophical point of view, we may ask ourselves whether these systems really are intelligent or even if they may one day have consciousness.

In contrast, computing still finds barriers to reach the integration of current technologies with the goal of reaching the apex of computational evolution, the Technological Singularity.

[1]Instituto Federal de Educação, Ciência e Tecnologia do Sudeste de Minas Gerais. alex.machado@ifsudestemg.edu.br.
[2] Universidade Federal de Brasília. luiztutorsjdr@yahoo.com.br
[3]Instituto Federal de Educação, Ciência e Tecnologia do Sudeste de Minas Gerais. pablo_sanches@outlook.com

https://doi.org/10.36311/1984-8900.2018.v10.n25.08.p111

Today we realize that most of the greatest scientific developments in the area of computing have been achieved through bioinspired solutions. In this work we intend to explain (elements) that the next great technological evolution depends, essentially, on the understanding of the nature of human consciousness, returning to the interrelation with “life” for a better computational modeling of the solution.

For this purpose, we will base this possible great evolution, which we call here Technological Singularity (session 2). Through philosophy, we will not only justify the possibility of the existence of artificial intelligence and consciousness (session 3), but also counter-argue it (session 4). Subsequently, we will argue that machine autonomy will be a philosophically determinant factor (through the confirmation test of the Chinese Room through the intentionality argument) and scientifically (through machine discovery) for the design of this new system (session 5). Finally, we will associate the autonomy with the machine consciousness (session 6) to show the possibility of reaching the Technological Singularity (session 7).

2. Artificial Intelligence and Technological Singularity

Artificial Intelligence (AI) is a branch of Computer Science research that seeks, through computational symbols, to build algorithms that simulate the human capacity to think and solve problems, that is, to be rational. It uses techniques such as data mining, which proposes to gather information from databases and classify them as efficiently as possible (MONARD; BARANAUSKAS, 2003).

AI is a strand of cybernetics, an interdisciplinary science that studies and influences several fields of knowledge. The term "cybernetics" (MONARD; BARANAUSKAS, 2003), in Ancient Greek, has a political meaning, since it is related to the art of governing communities. Over the years, the term began to relate to computational systems, especially in the study of AI, language, mind and man-machine communication.

The idea of AI arose around 1943, in the study of Warren McCulloch and Walter Pitts (1943) which presented an early concept of Neural Networks, which only solved problems with data that were informed by a specialist of the area. This was the starting point for two approaches in neural network research: one in the field of biological processes in the brain and another one in the field of AI. From its creation, AI has evolved exponentially, with autonomous machines that can perform difficult tasks and that are possible to generate incredible reasoning without the assistance of humans (MONARD; BARANAUSKAS, 2003).

Without an artificial intelligence solution, several efforts on information (like data analysis) would take decades to be processed on a computer or it would need to be done manually by a specialist in the field. The first idea of AI appeared around 1943 ith the primitive Neural Networks which only solved the problem with the data that were informed by the specialist of the area. From its creation to the present day, AI has evolved exponentially, with autonomous machines that perform tasks and can generate incredible reasoning by themselves (MONARD; BARANAUSKAS, 2003).

Using AI, we can achieve something that would be one of the greatest milestones of humanity, Technological Singularity (VINGE, 1993). This is a hypothesis that relates the unbridled technological development of artificial superintelligence and irreversible transformations in human civilization (EDEN et al., 2012).

Although it seems a new idea, some philosophers already imagined that the Technological Singularity would somehow happen centuries ago. Perhaps the first idea that points to singularity appeared centuries ago, with Hegel (1807). In his studies he concluded that the rise of human culture becomes an ideal limit point of absolute knowledge. Although this statement does not deal directly with technology, it could be suggested that Hegel's position demonstrates that Technological Singularity represents only one more step of absolute spirit in its process of self-recognition.

In the 1960s, Good (1966) speculated that Artificial Intelligence could provoke a technological explosion by proposing a scenario where, as computers become more potent, there is an increased possibility of building machines with greater inventive capacity and solution problems than man himself. Increasingly capable machines would be developed from this, accelerating the recursive self-improvement, resulting in a huge qualitative change before any higher limits imposed by the laws of physics or theoretical computation.

It will be possible to solve complex problems due to the unimaginable processing power that this algorithm can achieve, such as the cure of cancer or travel between galaxies, for example. Machines will be able to learn our tastes, desires, and emotions so accurately as to be able to perfectly simulate our world and the entire universe, to facilitate market research, marketing, medicine, physics secrets, and more, beyond, of course, significant developments in the entertainment industry (LENARCIC; MOUSSET, 2004).

In this work, as in the studies of Vinge (1993), we assume that the Technological Singularity can be reached and that it will be reached in a few decades from now.

Therefore, what we have today are machines (hardware) and algorithms (software) in an accelerated evolution motivated by the globalized market. However, there are no significant studies projecting what is necessary for the conception of this "being" created by technology. In this context, the need of a philosophical investigation arises to propose new horizons that the traditional scientific methodology cannot clear.

In this article, we will discuss the possibility of the existence of an artificial consciousness and its importance for the reach of the Technological Singularity.

3. Robot’s Intelligence, Mind and Consciousness

In this session, we will present some philosophical aspects regarding Artificial Intelligence as well as authors and approaches that discuss the autonomy of the intelligent agents.

3.1 Turing Test: Proposal to Evaluate an “Intelligent Behavior”

In 1950, Alan Turing published an article entitled “Computing Machinery and Intelligence” (TURING, 1950) on the thinking capacity of machines. The article proposes the so-called Turing Test, which works as follows: two people are separated in a room by a kind of “barrier”, one interprets the interrogator and, the other, who’s interrogated. A computer with Artificial Intelligence will represent a "third person." The interrogator performs questions via keyboard to the human and the machine, not knowing who is responding. If the interrogator is not able to differentiate the human from the machine, that system can be classified as an intelligent system.

Turing (1950) Test is a method used to measure the degree of intelligence of a machine when confronted with human intelligence. However, it does not intend to create a machine that has the capacity to think identical to the capacity of a human. The purpose of the test is to establish reliable parameters to test theories concerning intelligence and processes of cognition of the tested AI.

This test was fundamental for the development of AI, since it launched the tripod of cognitive science (ZILIO, 2009): a) the test served to show the independence between hardware (physical part) and software (function), since it established bases to a mechanistic analysis beyond the physical matter (CLARK, 2001) that took the functionalist vision of the mind; b) changed the meaning of the word “think”. Thinking about Turing is the ability to process information in order to solve problems. Therefore, if the participating machine of the game is able to deceive the interrogator, it can be considered a thinking machine; c) laid the foundations for the approximation of pure and applied sciences, especially logic, mathematics and computation (CLARK, 2001). The results of this are the development of algorithms, mathematical formulas capable of receiving, processing and transmitting information.

In 2014 a computer managed to convince humans that it was not a machine, passing the Turing Test during an event held in the United Kingdom (YOU, 2015). Eugene Goostman's software passed as 13-year-old boy, where 10 of the 30 judges present said it was a boy and not a computer.

Although there is a controversy regarding the approval of an algorithm in the Turing Test, considering an increase of investigations and investments in the field of AI (KURZWEIL, 2015; YOU, 2015; EDEN et al., 2012), the fact is that, if it has not yet happened, we are at the threshold of breaking a great technological paradigm.

3.2 The Descartes Automaton

For Turing, a machine could accomplish everything that a human being does, but it would not have consciousness of what it does (TURING, 1950). This statement resumes what Descartes (2004) classified as a category of automata: first degree automata: machine and animal; second degree automata: man. The difference between the categories is the capacity for self-reflection, in this case what is lacking to the first degree automata is, at the very least, the spirit [4] (DESCARTES, 2004) Thus, he proposes that the human being is a combination of two distinct substances:

- The *res extensa* (from the Latin, “extensive thing”). The matter we know, and more specifically, the body, would be modifications of that substance. No extensive substance is capable of thinking. Thought is an operation that can never be attributed to the body. Where there is no soul, there is no thought.

[4] It is cited "at the very least" because for Descartes spirit is not only self-reflection, but also a substance.

- The *res cogitans* (from the Latin "thinking thing" or soul). The soul was characterized by Descartes as something immaterial. It was also defined as a thinking substance, but, contrary to what is often conceived today, thinking was not just reasoning. To think also is to be affected by joy, to suffer with a frustration or to wish to perform an activity to acquire pleasure.

Descartes (2004) understands the mind as a set of non-reflexive states of consciousness: composed of mechanical processes (first degree automaton) and reflexive processes (second degree automata).

Therefore, a well-built automaton could do everything a human being does, but never have a soul.

However, in the philosophy of the twentieth century, Teixeira (2015) contrasts, many philosophers of the mind came to a vigorous critique of the dualist position and, in the last decades, they would have leaned towards neuroscience, whose predominant tendency is the reduction of mind to the brain. In the philosophy of the contemporary mind the metaphysical problem of mind-body relations was reformulated and came to be called the explanatory hiatus.

3.3 Computational Functionalism

From Putnam's (1980) studies on philosophy of mind, the digital computer came to be seen as a metaphor of the mind-body relationship, so that mental states could be compared to functional states of a computer. Just as a computer program runs in different hardware configurations, similarly, a "psychological program" hypothetically could take place in varied biological as well as artificial systems.

In this context, the computational functionalism interprets the mental processes in terms of functional states. That is, it understands the mind as a formal system of manipulation of mutually related symbols in order to perform, under certain conditions, specific functions. Thus, a functional state is defined by the strict causal relation between input stimuli – these are understood as certain internal causal states – and output stimuli – their resulting effects or behaviors. In conceiving the mind in these terms, functionalism will recognize, in effect, the possibility that systems as distinct as human beings and calculating machines may somehow have mental states. For Block and Fodor (1980) the psychology of a system should not depend on the matter of which it is formed, but only on the way in which it is arranged. Since the functional state never depends on the physical structure of the system, it can be said, likewise, that the character of the mental state never refers to the physical structure.

From this, it is possible to understand the proposal of computational functionalism and the various theories of artificial intelligence. A computer machine exists exclusively for the purpose of performing functions. In this process, computers decode certain inputs, which are specific input commands, thus transforming them into an established set of formal rules, into responses or outputs that can be expressed in certain behaviors.

Reinforcing this idea, for Block:

> Computational-representational functionalism applies in an important case of functional explanation, namely, in the psychological explanation seen as analogous to a computer program for the mind. Any mystery about our mental life can be initially dissolved by the functional analysis of mental processes from a point where they can be seen as mechanical computations of a digital computer (...). Psychological states are seen as systematically representing the world via a language of thought, and psychological processes are seen as computations about such representations (BLOCK, 1980, p. 179).

However, in the case of humans, the inputs present themselves as certain sensitive stimuli, whether visual, auditory or otherwise, while the resulting outputs are expressed in certain types of behavior, such as when one feels a headache and presents a set of reactions behaviors among which is the attitude of taking a remedy. Therefore, if mental states are really comparable to functional states and functional roles always depend on input and output stimuli, nothing prevents the functionalist from proposing a strict analogy between the human mind and software, claiming that "the relational character of the mental" is always preserved in a computational machine.

4. Chinese Room: John Searle's Critique of Artificial Intelligence

John Searle (1980) criticized the cognitivism, which was characteristic of Turing's proposal. For the philosopher, the central hypothesis of the cognitive paradigm that a brain is a computer that processes information is incoherent because nothing is intrinsically a digital computer. In order to elaborate its critical arguments, he distinguished AI's into two approaches: weak AI - where artificial intelligence systems can simulate intelligent behavior; and strong AI - in which artificial intelligence systems are really intelligent entities, capable of thinking and having genuine minds as Turing intended. Currently the two approaches coexist in AI.

The Chinese Room was a philosophical argument built by John Searle John Searle (1980). He intended with this theory to oppose researchers in AI. This argument has some relation with the Turing Test that aims to verify if a Turing machine can or not be considered intelligent. Going further, Searle argued that it could not possess self-consciousness, not even think, for it would only be able to simulate reasoning. Searle's argument is based on the fact that computers can understand syntax only, but not semantics.

John Searle (1980) then proposes a hypothetical picture in which an individual is locked in a room with only one opening. Let us consider that this individual speaks only Portuguese, not understanding a word or symbol in Chinese. At your disposal there is only one box, with symbols in Chinese and a book with rules, where it is explicated that symbols should be sent in response when another symbol is received (for example, if a paper with a printed square is received by the opening, consulting the book is possible to come to the conclusion that as a suitable response a sheet should be sent with a circle). Thus is established a metaphor, in which the human being is the computer, the book of rules the program, and the sheets of paper the storage devices.

Now let us suppose that successive sheets of paper are sent with questions in Chinese; using the material that is available, the individual inside the room will be perfectly capable of sending the correct answers, also in Chinese, without, however, understanding what he is doing, without knowing the language. Anyone who is outside receiving the answers will think that whoever generated them is an “intelligent” being, who understands Chinese, which does not happen.

In this way, Searle demonstrates that a computer program, despite passing the Turing Test for the Chinese language, does not have any form of intelligence, because it has no cognitive abilities; only generates the appropriate answers, by means of previous programming, not necessarily understanding.

Searle concludes that saying that a machine understands something is a misconception, because the thoughts or “mental states” of a machine are over nothing. They have no relation to the external world. We cannot think unless we think of something and, to this capacity to think of something, Searle (1980) calls intentionality.

His conclusion is that intentional states are a representation of the world and cannot exist unconsciously. Thus consciousness would be a fundamental part of meaning.

In his biological naturalism, Searle even admits that if man could faithfully re-create a brain, there could be consciousness. But it does not accept that consciousness is reduced to input and output stimuli.

The argument of the Chinese Room is an icon of criticism to AI. Searle (1980) admits the idea of the weak AI approach, because artificial systems have become tools for studying mental processes, admitting physical tests.

5. Confirmation Test of John Searle's Absence of Intentionality through the Search of Machine Autonomy

Wittgenstein (1983) opposes Searle in his theory of meaning. In this conception, meaning is generated by the practices of language use and not by intentional mental states associated with words and beliefs. The conclusion thereafter is that the occupant of the Chinese Room could not generate symbols with meaning because of the private character of his language. Therefore, Searle's reductionism is wrong in treating the appearance of symbols on the screen as a form of behavior that could constitute a manifestation of thought. Typing symbol can only count as “thought” and “calculation” for a creature to which we can assign a wider range of psychological attributes. It is plausible to argue that neither experiments nor emotions can be attributed to a machine that merely reacts according to a program by pressing the keys. This in itself does not, however, prevent us from speaking here in thought or calculation. The most basic motivation for Wittgenstein's position is that following a rule presupposes that we do things for a reason, which is only possible for a creature endowed with autonomy, that is, that it can be interested in things and pursue goals.

From another point of view, the basis of intentionality and consciousness is associated with life. Without a living brain, they would not be possible. Soon this position of Searle can be questioned. After reflections on "life" Teixeira (2015) concludes, in agreement with Dennett (2017) that, due to the binary nature of DNA, life is a computable phenomenon, which would allow constructing a replica of the human brain. Reinforcing this argument, for Dennett (2017), consciousness is reducible to a related set of inputs and outputs. Therefore, it could have intentional states. This ideology also goes according to the computational functionalism thesis further elaborated by Block and Fodor (1980).

5.1 Autonomy in Karl Marx

Marx and Engels (1989) establish analytical parameters contrary to the absolute autonomy of the human being in The German Ideology. They consider the possibility of controlling the autonomy of man, given the social relations practiced in a historical context. This is the human which can be programmed to perform tasks of inputs and outputs without being aware of the meaning of their actions, as in the proposal of the Chinese Room.

In treating here intentionality directly related to autonomy in the sense of interest and freedom to make its decisions, we return to Karl Marx's (2005) considerations on human relations. For him, autonomy is the work of the existential human obligation to satiate his needs. It is defined as a moment prior to human enterprise when there is the act of deciding for a certain praxis, which, in itself, exists after a decision, and it goes through a metaphysical process of an individual that refers to the autonomy, the metaphysical essence of the human real movement. For everything in life, this action will be conditioned to the consciousness of the individual, that is, to their autonomy. In consciousness, autonomy is present as one of the determinants of praxis, and from this, the human being is realized, reproduces and produces society.

Consciousness, in the history of philosophical thought, is mostly what differs man from other things. In German philosophy, questions that pervade man's self-consciousness in the face of others and in the processes of understanding reality through are central questions. It happens that most of the philosophical tradition approaches the questions of the conscience from the dualism where the subjectivity superposes the objectivity: interior x exterior; subjective x objective. In short, consciousness has autonomy vis-a-vis the real world. Marx and Engels (1989) reterritorialized consciousness for the living world, experienced in the material social relations of each historical context. Awareness here is not a fact that explains socio-historical and cultural development. On the contrary, it derives from these objective relations of power between the ruling class and the dominated class. It is not the subjective autonomous consciousness that determines the practical life, but the life that determines consciousness.

From this relative autonomy of man derives the concept of alienation from Marx (2005) - which has a double meaning: estrangement and loss of consciousness. Ultimately, the loss of consciousness causes the thinking man to behave as a repeating machine of processes within a social system of production with inputs and outputs. That is, the alienation, produced by the labor regime in capitalist society, leads man to the loss of himself.

In his analysis of human alienation, Karl Marx (2005) assumed that the best of spiders will never be superior to the worst of the architects. Although the spider made magnificent webs, it was born knowing, was preprogrammed. And since it does not err, it does not create, it does not evolve for its own sake, because only error will dig into it the pleasure of innovation. What makes it decide things comes from outside it. For him, what the spider lacks is autonomy.

We have until then that the autonomy can be one of the tests that we needed to contrast to the theory of the Chinese Room. It is even important in the growth of the human individual, as Marx clarifies, therefore, analogously to the machine, we believe that it would be the key piece to reach the Technological Singularity. So the question is: can a computer program self-design yourself?

5.2 Autonomous Computing

Understanding the power that a machine can have in achieving Technological Singularity is not an easy task. However, over the years, the investments and (consequently) the studies related to Artificial Intelligence and the autonomy of the machines are progressing. Increasingly, there are theories and clues that can guide what the singularity would be.

With the exponential growth of technology and greater demand for more complex and extensive systems, the need for architectures that are independent of man for their management increases every day. Thereby, the automated systems paradigm arose, they are capable of an implementation that allows the self-management of their tasks, such as self-configuration, self-optimization, self-protection and self-repair (KEPHART & CHESS, 2003). Table 1 shows the four aspects of self-management, comparing conventional systems with the autonomous computing paradigm.

Thus, it is essential to study autonomous systems for the understanding of robotic evolution (KEPHART & CHESS, 2003), which, driven by the evolution of technology, enables us to create robots with active capacities, increasingly similar to humans.

Automation in	**Conventional Computing**	**Autonomous Computing**
Setting	Data centers have many suppliers and multiple platforms. System installation, configuration, and integration are time consuming and error prone.	Automatic configuration of components and systems from a high level monitoring. The systems adjust automatically.
Optimization	The systems have hundreds of manual adjustments, non-linear parameter adjustments and their numbers grow with each new version.	Components and systems seek opportunities to improve their own performance and efficiency.
Repair	Problem diagnostics on large, complex systems that can consume weeks of developer teams.	The system automatically detects, diagnoses and repairs software and hardware problems.
Protection	Detection and repair of attacks as well as failures is manual.	The system automatically defends attacks and failures. It is able to predict attacks and failures, anticipating the facts.

Table 1: Self-management: comparison between current systems and autonomous computing. (adapted from the work of Kephart and Chess (2003)).

5.3 Learning x Discovery: Demonstrations of machine autonomy

In a pioneering work for the time, Zytrokow (1993) clarified the difference between Learning (of Machine) and Discovery. The first depends on a knowledge base or a specialist to guide its learning, while the other will only be guaranteed without outside help, it will itself judge their decisions. Therefore, learning is easier than discovery.

In his studies he defined the scientific and mathematical foundations for the design of systems prepared for the discovery of machine as well as enumerated contemporary applications. He said that the fundamental difference between the two is that the discovery needs autonomy. For Zytrokow:

> Two complementary notions of autonomy are useful for us. First, the more an agent can do in the external world, the more autonomous it is. To increase this type of autonomy, we can give the agent more means, for instance, more sensors or better manipulators. Second, within the same means, an agent is more autonomous if it can make more choices, uphold more values, and explore a richer space of goals. An agent can have meager means to influence the external world, yet be able to set its cognitive goals in response to external situations, to internal values, and possessed knowledge, so that it can make its own decisions and understand as much of the world as its limited means permit. Philosophers call this an existential concept of autonomy (ZYTROKOW, 1993, p. 8)

From this point, whether through the concepts of Artificial Intelligence, Machine Learning or Machine Discovery, all interrelated, we realize that autonomy is a reality and represents an essential step for technological evolution.

6. The (Possible) Existence of Artificial Consciousness

By accepting autonomy as a characteristic of the system, we cannot forget that it refers to only part of the consciousness, according to Searle (1980). In this session we will discuss Giorgio Buttazzo's (2001) research in his work "Artificial consciousness: Utopia or real possibility?" he argues that an artificial consciousness can exist and bases these questions with Turing and Descartes. To this end, the discussion begins by raising the following philosophical questions:

- Is consciousness a human prerogative?
- Does consciousness depend on the substance that constitutes the human brain, or can computer hardware replicate consciousness?

Regarding the first question, from a purely philosophical perspective, we cannot verify the presence of consciousness in another brain, whether human or artificial, because only the possessor himself can verify this property, since we cannot enter the mind of another being, we cannot be sure of your consciousness. Thus the indetermination of the consciousness of the other infers the indeterminacy of the consciousness of the machine. Therefore, if it exists in man, it can exist outside.

Many scientists believe that consciousness cannot emerge from a silicon substrate because it is an intrinsic property of biological materials, such as neural cells. So we can reasonably ask: does consciousness depend on matter? To contrast this argument, we can simply observe around us: flying entities are made of different materials, such as birds, bats, insects, paper airplanes and spaceships. It soon makes sense that an information processing tool could be derived only from organic and chemical cells.

Generally, there is more than one way to build a particular type of machine. Several elements and combinations of elements may be able to do the same task of different shapes and efficiencies.

Thus, we arrive to a new question:

- Could a machine develop self- consciousness?

If consciousness is a product of a highly organized information processing system and this property does not depend on the hardware substratum that performs computation, soon can also arise on a machine, as the advocates of functionalism also believe Block and Fodor (1980) and Putnam (1980).

However, the most common objection to granting self-consciousness status of electronic-oriented computers is the realization that by working in a fully automated mode, they cannot display creativity, emotions, or free will. A computer, like a washing machine, is a slave operated by its components. Despite this, with the evolution of machine discovery, the problem of artificial consciousness becomes even more intriguing because these bioinspired algorithms replicate the basic electrical behavior of the brain and provide adequate support for the performance of a processing mechanism similar to that adopted by the brain.

7. Conclusions

The idea that machines will exceed the limit of human reasoning is nothing new. With the passage of time and the increasing expansion of technological capacity, this may cease to be just fiction. This can be the next step of evolution and the beginning of a new era. The Technological Singularity is not far, but it is not yet known what can happen after that.

We present in this article several authors who argue that artificial consciousness may be a reality, also listing important counterarguments.

For the advent of Technological Singularity to be possible, we reinforce an intrinsic characteristic of this new being: intentionality. This will be based on the levels of autonomy and discovery of the machine and will allow the (possible) emergence of consciousness in this entity.

Finally, an alert comes from Wittgenstein (1993) and Turing (1950): when computers have consciousness, we will not be able to understand their thoughts. Computers and brains are radically different (inorganic/organic) material instantiations, and therefore, there will always be a qualitative difference in the forms of movement of matter that occur in each one. And the interesting thing is that this situation is closer to happening than we can imagine. In an experiment by Google Brain Abadi and Andersen's (2016) with machine learning, two intelligent agents were able to communicate through a language created by themselves.

References

ABADI, M.; ANDERSEN, D. Learning to protect communications with adversarial neural cryptography. *ArXiv preprint arXiv*, 1612.06918, 2016

BAKER, G.; HACKER, P. Wittgenstein Meaning and Understanding. *Essays on the Philosophical Investigations,* Vol. 1. University of Chicago Press, 1983.

BLOCK, N.; FODOR, J. A. What psychological states are not. In: BLOCK, N. (Org.). *Readings in Philosophy of Psychology*. Cambridge: Harvard University Press, v. 1, 1980.

BLOCK, N. What is functionalism? *Readings in Philosophy of Psychology*. Cambridge: Harvard University Press, 1980.

BUTTAZZO, G. Artificial consciousness: Utopia or real possibility? *Computer*, v. 34, n. 7, p. 24-30, 2001.

CANAL, R. Quatro objeções de John Searle ao cognitivismo. *Kínesis*, Vol. I, n° 01, Março / 2009, p.171-185.

CLARK, A. *Mindware: an introduction to the philosophy of cognitive science.* New York. Oxford University Press. 2001.

CLIFF, D.; HUSBANDS, P. e HARVEY, I. Explorações na Robótica Evolutiva. *Adaptive Behavior*, 2 (1):73–110. 1993.

DENNETT, D. *Consciousness explained*. Little, Brown, 2017.

DESCARTES, R. Discurso del método. Ediciones Colihue SRL, 2004.

EDEN, A.; STEINHART, E.; PEARCE, D. & MOOR, J. H. Singularity hypotheses: an overview. *Singularity Hypotheses*. Springer Berlin Heidelberg. 2012.

FUEGI, J.; FRANCIS, J. *Lovelace & Babbage and the creation of the 1843 'notes'*. Annals of the History of Computing, 2003.

GOOD, I. Speculations concerning the first ultraintelligent machine. *Advances in computers*, v. 6, p. 31-88, 1966
HEGEL, G. *Wer denkt abstrakt*. Werke in zwanzig Bänden, 1807
KEPHART, J; CHESS, D. The vision of autonomic computing. *Computer***,** 36 (1):41–50. 2003
KURZWEIL, R. *The singularity is near: When humans transcend biology***.** Penguin, 2005.
LENARCIC, J.; MOUSSET, E. The open source singularity: A postmoderniist view. *Computers and Philosophy*, v. 37, p. 73-77, 2004
MARX, K. *Manuscritos econômico-filosóficos e outros textos escolhidos*. 1ª edição 1844. São Paulo, Nova Cultural, 1991.
______. *O Capital: crítica da economia política*. Livro I. Volume 1. Tradução de Reginaldo Sant'Anna. 2ª edição. Rio de Janeiro: Civilização Brasileira, 2005.
MARX, K.; ENGELS, F. *The German Ideology.* Christopher J. Arthur's Edits. New York: International Publishers, 1989.
MCCULLOCH. W.; PITTS, W. A logical calculus of the ideas immanent in nervous activity. *The bulletin of mathematical biophysics*, 5(4):115–133, 1943
MONARD, M.; BARANAUSKAS, J. Conceitos sobre aprendizado de máquina. *Sistemas Inteligentes-Fundamentos e Aplicações* 1.1. 2003.
PLATÃO. *Diálogos: seleção de textos de José Américo Motta Pessanha*. São Paulo. Nova Cultural, 5ª edição, 1991. (Os pensadores).
______. *Górgias.* Tradução Carlos Alberto Nunes. Bélem: Universidade Federal do Pará, 1986.
PYLYSHYN, Z. Computation and cognition. Cambridge: *The MIT Press*, 1986.
PUTNAM, H. The nature of mental states. In: BLOCK, N. (Org.). *Readings in Philosophy of Psychology*. Cambridge. Harvard University Press, 1980.
ROSENBLATT, F. *The perceptron, a perceiving and recognizing automaton***.** *Project Cornell* Aeronautical *Laboratory*, 1957.
SEARLE, J. Minds, brains, and programs. *Behavioral and brain sciences*, v. 3, n. 3, p. 417-424, 1980.
TEIXEIRA, J. *O cérebro e o robô: inteligência artificial, biotecnologia e a nova ética.* São Paulo: Paulus, 2015.
THE ORGANIZATION OF BEHAVIOR: *A neuropsychological theory.* D. O. Hebb. John Wiley And Sons, Inc., New York, 1949.
TURING, A. Computing machinery and intelligence. *Mind*, v. 59, n. 236, p. 433-460, 1950.
VINGE, V. *The Coming Technological Singularity. Whole Earth Review Winter issue.* Vita-More, N. Transhumanist Arts Statement, 1993.
YOU, J. Beyond the Turing Test. *Science***,** v. 347, n. 6218, p. 116-116, 2015.
ŻYTKOW, J. *Introduction: Cognitive autonomy in machine discovery*. Machine Learning, v. 12, n. 1, p. 7-16, 1993.
ZILIO, D. Inteligência artificial e pensamento: redefinindo os parâmetros da questão primordial de Turing. *Ciências & Cognição*, 2009; Vol 14 (1): 208-218.

Acknowledgements
The authors would like to thank IF Sudeste MG for all financial support.

Funding
This work was supported by IF Sudeste MG.

Recebido em: 21/09/2018
Aprovado em: 18/09/2018