doi.org/10.29327/2194248.6.2-2

ON THE CONCEPT OF INTELLIGENCE IN THE CONTEXT OF THE TURING TEST

Victor Pereira Gomes
Universidade Estadual da Paraíba, Brasil
https://orcid.org/0000-0003-0054-019X

ABSTRACT: The Turing Test (TT), developed by Alan Turing as "the imitation game" in his paper *Computing Machinery and Intelligence* (1950), brought to light the discussion about the (im)possibility of thinking and intelligent digital machines. This paper revisits the Turing Test and analyzes the concept of intelligence under its light, exploring Turing's understanding of intelligence. The focus is on whether Turing views intelligence as human intelligence (referred to here as genuine intelligence) or as some other type of intelligence. The research results suggest: 1) The possibility to argue the Turing Test as not being developed to assess whether the digital computer involved possesses genuine (human) intelligence or not, but rather to evaluate if it can be intelligent in a different sense, termed hereby *Turing-intelligence*; 2) The Turing Test is achievable in practice, given some modifications, resulting in a new version of the test, named herewith the *Ideal Turing Test*. Therefore, the findings indicate that the TT can be interpreted to support the hypothesis that passing such test is a sufficient condition for an AI system to possess not genuine intelligence, but rather another type of intelligence, namely Turing-intelligence.
KEYWORDS: The Turing Test. Intelligence. Machine Intelligence. Turing-intelligence.

SOBRE O CONCEITO DE INTELIGÊNCIA NO CONTEXTO DO TESTE DE TURING

RESUMO: O teste de Turing (TT), desenvolvido por Alan Turing como "o jogo da imitação" em seu artigo *Computing Machinery and Intelligence* (1950), trouxe à baila a discussão acerca da (im)possibilidade de máquinas digitais pensantes e inteligentes existirem. O presente artigo objetiva revisitar o Teste de Turing e analisar o conceito de inteligência no contexto do referido teste, explorando o que Turing entende por inteligência. O foco está em saber se Turing vê a inteligência como inteligência humana (chamada aqui de inteligência genuína) ou como algum outro tipo de inteligência. Como resultados da pesquisa, argumenta-se que 1) é possível interpretar o Teste de Turing de forma a concluir que ele não foi desenvolvido objetivando avaliar se o computador digital envolvido nele possui inteligência genuína (humana) ou não, mas sim para avaliar se ele pode ser considerado inteligente, no sentido do que é aqui chamado de *Turing-inteligência*; 2) o Teste de Turing é possível, isto é, realizável na prática, contanto que passe por algumas modificações, resultando numa nova versão do teste, aqui denominada de *Teste de Turing Ideal*. De posse de tais resultados, tem-se que o TT pode ser interpretado de forma a sustentar a hipótese de que passar em tal teste é condição suficiente não para que um

sistema de IA possua inteligência genuína, mas sim um outro tipo de inteligência, a saber, a Turing-inteligência.

PALAVRAS-CHAVES: O Teste de Turing. Inteligência. Inteligência de Máquina. Turing-inteligência.

INTRODUCTION

In the aftermath of the Second World War, specifically from the late 1940s onward, significant advancements were made in the field of Artificial Intelligence (AI). The pivotal moment in its development occurred in 1950 when the English logician and mathematician Alan Turing published the seminal paper *Computing Machinery and Intelligence*. In this work, Turing posed the fundamental question "Can machines think?" (1950, p. 433). Given the complexity of this question and the necessity to define the concepts of machine and thought, Turing chose to reformulate it in light of an idea introduced in the same paper: the imitation game, which is now known as the Turing Test (TT).

Since the publication of Turing's *Computing Machinery and Intelligence* (1950), considered a milestone concerning the development of AI, cognitive scientists have been discussing the (im)possibility of thinking and intelligent digital machines[1] existing. Being a philosophical paper, it has been the target of many criticisms and objections from philosophers such as Ned Block (1981), Robert French (1990), and John Searle (1980), who interpret the test as a sufficient condition for a digital computer to possess genuine (human) intelligence. This interpretation is called here "the standard interpretation of the test".

The present paper[2] aims to analyze the Turing Test (TT), and to present a plausible solution to the following: "How does Turing understand the concept of intelligence within his test? Is it human intelligence (called here genuine intelligence),[3] or is it some other type of intelligence (whatever that intelligence may be)?". The relevance of this question is that the TT, as described by Turing, is subject to different interpretations and it would not be wise to focus on just one of them. Nonetheless, it should be noted that this question challenges the standard interpretation.

---

[1] Throughout the text, the terms "digital machine" and "digital computer" will be used as synonyms.
[2] This paper reflects part of the research I developed for my Ph.D. thesis.
[3] Genuine intelligence and human intelligence are here taken as synonyms. Although there are different definitions of "genuine intelligence", the present text does not intend to propose a definition for this kind of intelligence.

It becomes evident that the answer to this question is neither simple nor readily found in Philosophy or Cognitive Sciences texts. To address this, a study focused on Turing's understanding of the digital computer and the concept of intelligence is necessary. This study should align with the non-standard interpretation of the TT presented in this paper, which argues that Turing proposed a notion of intelligence, termed herewith Turing-intelligence. This is defined as the ability of a digital computer to perform tasks (exclusively through imitation[4] and simulation)[5] that require intelligence when performed by a human. Thus, in the TT context, a digital computer is considered Turing-intelligent if it can (exclusively through imitation and simulation) engage in dialogue with humans in the same manner humans converse with one another.

Regardless of claims that the TT is superficial or should not be taken seriously, it is still relevant. Since the publication of *Computing Machinery and Intelligence* (1950), just over 70 years ago, the TT has been the target of discussions. As an example of its ongoing relevance, two issues of relevant journals are presented, the 2020 special edition of the journal Minds and Machines, titled *Rethinking, Reworking, and Revolutionising the Turing Test*, as well as the 2023 special edition of the journal Philosophies, titled *Turing the Philosopher: Established Debates and New Developments*. As both titles imply, these special editions focused on the TT.

After a thorough study of papers and textbooks such as 1) *On Computable Numbers, with an Application to the Entscheidungsproblem* (1936), 2) *Computing Machinery and Intelligence* (1950), 3) *Can Digital Computers Think?* (1951), 4) *Can Automatic Calculating Machines be said to Think?* (1952), 5) *Parsing the Turing Test: Philosophical and Methodological Issues in the Quest for the Thinking Computer* (2009), 6) *A Clarification on Turing's Test and its Implications for Machine Intelligence* (2015), as well as other texts referenced throughout this paper, I argue, as a research result, that the TT is not a test designed to assess whether the digital computer involved possesses genuine intelligence, but rather to evaluate whether it can be considered Turing-intelligent.

In section 1, I will offer a concise introduction to Artificial Intelligence, emphasizing its definition, origins, and primary objectives. In section 2, I will introduce and elucidate the

---

[4] Imitation, in this context, refers to the performance of an act that is prompted by the perception of a similar act by something or someone.
[5] Simulation, in this context, refers to the use of algorithms in computers to replicate a real-world process or operation.

TT, explaining its nature and Turing's assertions regarding its implementation. In section 3, I will reinterpret the TT and further clarify the notion of Turing-intelligence. In section 4, I argue that The Turing Test is possible, i.e., achievable in practice, as long as it undergoes some modifications, resulting in a new version of the test denominated the Ideal Turing Test (ITT).

After a thorough study of the aforementioned texts, I propose that 1) Considering the non-standard interpretation of the TT, it was not conceived to evaluate whether the digital computer involved possesses genuine (human) intelligence, but rather to assess whether it is Turing-intelligent. In other words, it can be concluded that the success of a digital computer in the test implies that it possesses Turing-intelligence; 2) The Turing Test is possible, i.e., achievable in practice, as long as it undergoes some modifications, resulting in a new version of the test denominated the Ideal Turing Test (ITT). That being so, to affirm that an AI system[6] passes the ITT implies that such a system is Turing-intelligent, being classified in one of the levels of Turing-intelligence. Regarding Turing-intelligence levels, the reader will see in section 4 that they are necessary to classify AI systems, on a scale of 01-06, according to their performances in the ITT.

1 – ARTIFICIAL INTELLIGENCE

The term Artificial Intelligence (AI) was coined in 1956 by the North American computer scientist John McCarthy during a conference at Dartmouth College. Although he defined AI as "[...] the science and engineering of making intelligent machines" (McCarthy, 2007, p. 2), there is no consensus among researchers regarding this definition. Thus, it is not difficult to find other definitions for it. According to Jack Copeland, AI "[...] is usually defined as the science of making computers do things that require intelligence when done by humans" (Copeland, 2000b, online). Notwithstanding, following John Haugeland, it can also be defined as "[...] the exciting new efforts to make computers think [...] *machines with minds*, in the full and literal sense" (Haugeland, 1985, p. 2).

During the 1950s-1980s, the research field that gained prominence within AI was Symbolic AI, also known as Classic AI or GOFAI - Good Old Fashioned Artificial Intelligence (as highlighted by Haugeland (1985)). It was the first research field within AI based on the

[6] For an AI system, we mean a digital computer developed to imitate and simulate human intelligence.

symbolic representation of problems, logic, and search. Its basic idea was that aspects of human intelligence, such as thinking, problem solving, learning, and reasoning, could be achieved solely through symbol manipulation. Furthermore, its goals included developing digital computers capable of emulating human intelligence and decision-making. However, according to McCarthy (2007), Symbolic AI's primary aim was not to mimic human behavior, for "[...] most work in AI involves studying the problems the world presents to intelligence rather than studying people or animals" (2007, p. 3). Thus, AI researchers often employ methods independent of human behavior observation.

Another research field that gained prominence within AI since the 1980s was Connectionist AI. Although its concept emerged in 1943 through the work of McCulloch and Pitts, it was just in 1980 that it became popular. It differs from Symbolic AI in its basic premise, which is to simulate human brain components, such as neurons. It is worth noting that at the beginning, the contributions to this field were scarce since the hardware resources available were lesser. However, with the advancement of computing and computational tools, the current scenario is thoroughly different, and, as we know, the connectionist model of AI has largely evolved.

Since the 1980s, the belief that Symbolic AI has lost its usefulness has arisen due to the advancement of Machine Learning and Deep Learning[7] (especially with the advent of artificial neural networks).[8] However, this may not be the case, as Symbolic AI is still widely used in AI systems, such as natural language processing systems (human-machine verbal communication, such as Google Assistant, Alexa, and Siri) and logical inference systems (conclusion generation

[7] Machine Learning and Deep Learning are types of Connectionist AI which enable digital machines to interact more efficiently with humans. Machine Learning is a type of Connectionist AI that allows algorithmic applications to be more precise in the sense they can offer solutions and alternatives despite not being programmed to do so. That is possible because they can learn from experiences, i.e., from their mistakes and successes and from information acquired over time. Furthermore, they can identify patterns in massive data and make predictions. The central idea of Machine Learning is to develop algorithms that learn to read and understand new data and, through statistical analysis, determine answers within a finite number of possibilities. It is also important to highlight that there are two types of Machine Learning, namely, supervised and unsupervised ones. In the supervised one, the programmers must supply digital machines with information and offer feedback regarding their results. The unsupervised one performs tasks autonomously and the results are auto analyzed. Deep Learning, in turn, is a type of Connectionist AI that simulates the way humans process information and enables digital machines to do the same. Unsupervised Machine Learning uses Deep Learning to review the information received and reach conclusions.

[8] Computational model inspired by biological neuron networks (human brain neural networks), in which algorithms continuously analyze data and update their predictions just as human brains receive information, process them, and make decisions. Thus, each algorithm represents a neuron and calculates the output values from the given inputs.

based on previously given rules and evidence). Although Symbolic AI is still widely used, it is not as effective as Connectionist AI in terms of learning patterns.

Currently, AI constitutes one of the fields of study of sciences which investigate different aspects of human cognition, such as Cognitive Sciences, Computer Science, and Engineering. Cognitive Sciences explore the nature and functioning of the human mind, aiming to understand cognitive phenomena such as thinking, intelligence, learning, and understanding, Computer Science focuses on developing digital computers capable of performing tasks better than humans, and Engineering seeks to build instruments to support human intelligence, assisting in areas like medical diagnoses, and voice and writing recognition.

2 – THE TURING TEST

Alan Turing was an English logician and mathematician best known for his contributions during the Second World War (1939-1945). His work was fundamental in decoding messages encrypted by the "enigma," revealing vital orders and commands of the German military forces (*Wehrmacht*), which completely changed the war panorama. Furthermore, in the 1930s, Turing played a crucial role in developing what is now known as Computer Science, formalizing concepts of algorithm and computation through the theory of Turing machines, significantly influencing the development of modern computing, specially through the advent of digital computers in the 1950s.

With the digital computers development advancement in the 1950s, computing obtained the status of a science and began to be studied academically. AI is believed to have emerged between the late 1940s and early 1950s, inspired by Turing's reflections during a lecture on computational intelligence (computer intelligence) in London in 1947 and in his well-known and influential paper *Computing Machinery and Intelligence*, published in 1950.

Given the complexity of defining the concepts of "machine" (which Turing later defines as a deterministic working discrete state universal digital machine, in sections 4 and 5 of the paper) and "thought" (which is not defined in the text) before even seeking a direct answer to it, Turing concludes that the original question is "[...] too meaningless to deserve discussion." (Turing, 1950, p. 442). To avoid the ambiguity of the words "machine" and "thought", Turing reformulated the question to focus on what he introduced, in this same paper, as the imitation game, now known as the Turing Test (TT), described below.

A brief overview[9] of the test will suffice. The game aims to evaluate a digital machine's capability to simulate human intelligence, specifically, its ability to converse and respond to miscellaneous questions. In short, the game involves three players, A, B, and C, where C is an average human interrogator[10] (male or female), A is a digital machine, and B is a human being (male or female). The players are placed in separate rooms and are only able to interact through typewritten messages. Player C asks assorted questions to A and B and has to identify which of the two is the digital machine and which is the human based on their answers. Player A has to trick C into making the wrong identification, and the role of B is always to tell the truth, leading C to the correct identification. As an example, if C asks which player is the human being, B could answer: "I am the human being." On the other hand, A could respond: "do not listen to B; the human being is me," and vice versa. If, at the end of the game, C fails to correctly identify the human or does it incorrectly, then the digital machine has successfully performed the imitation game[11] and, consequently, passed the TT.

Two years after the publication of *Computing Machinery and Intelligence* (1950), Turing presented an updated version of the TT in a debate[12] with Newman, R. B. Braithwaite, and G. Jefferson, in 1952. In this version the single average interrogator from the previous version is replaced by a jury composed of average jurors. Players A and B are replaced by several other players, some of whom are digital machines, and some human beings. The judges communicate with the players via teletype messages, one participant at a time, and each judge must interact with the participants several times, considering that sometimes they will be interacting with a human being and not with a digital machine. We say a digital computer has passed this version of the test if it fools a sizable proportion of the jurors.

---

[9] See Turing (1950), and Turing; Newman; Braithwaite; Jefferson (1952) for a richer description of the imitation game.
[10] The interrogator must not possess specific knowledge about digital machines.
[11] The reader familiar with Turing's 1950 paper will realize that the first version of the imitation game is what is known today as the gender test. The gender test is to be played by three participants, A, B and C, where A is a man, B is a woman, and C can be either a man or a woman (gender is irrelevant here). At the end of the game, if the man manages to pretend to be a woman to the point of deceiving the interrogator, we say he was successful in the game. The second version of the game (discussed in this paper) is known as the species test - the most common interpretation of the test, where A is a digital computer (taking the place of the man in the gender test), and B and C are human beings (male or female). At the end of the game, if the digital computer manages to pretend to be a human to the point of deceiving the interrogator, we say that it has passed the test. In this case, the interrogator is aware that one of the two players is a digital computer, however, they do not know which one.
[12] The debate was recorded by BBC Radio on January 10, 1952, and broadcast by the BBC on January 14 and 23 of that year.

Notice that in the first version of the test (1950), the interrogator knows beforehand that one of the players, A or B, is a digital machine and the other is a human being, but they do not know which one is which. On the other hand, this does not occur in the second version of the test (1952) since each judge communicates with one player at a time, and that player can either be a digital machine or a human.

Regarding the questions to be addressed to the players in both versions of the TT, Turing (1950) states that the only restriction is that the interrogator cannot demand practical demonstrations from the players. In Turing's words,

> We do not wish to penalise the machine for its inability to shine in beauty competitions, nor to penalise a man for losing in a race against an aeroplane. The conditions of our game make these disabilities irrelevant. The 'witnesses' can brag, if they consider it advisable, as much as they please about their charms, strength or heroism, but the interrogator cannot demand practical demonstrations (1950, p. 435).

The test was so developed to prevent the interrogator from seeing, hearing the voices, or touching the players (otherwise, it would be easier to identify which is the digital machine and which is the human being). Furthermore, in the debate recorded and broadcast by the BBC in 1952, Turing, when talking about the questions that could be asked by the interrogator or the jurors in the TT, asserts that

> [...] the questions don't really have to be questions, any more than questions in a law court are really questions. [The question] "I put it to you that you are only pretending to be a man" would be quite in order. Likewise, the machine would be permitted all sorts of tricks so as to appear more man-like, such as waiting a bit before giving the answer [...] (Turing; Newman; Braithwaite; Jefferson, 1952, online).

After presenting the outline of the imitation game, Turing (1950) reformulated his initial question ("Can machines think?"), developing it in its final structure: is there a way to imagine a digital machine capable of satisfactorily playing the imitation game? To answer the question, Turing states the following:

> I believe that in about fifty years' time it will be possible to programme computers, with a storage capacity of about 10^9, to make them play the imitation game so well that an average interrogator will not have more than 70 per cent chance of making the right identification after five minutes of questioning (1950, p. 442).

Turing's prediction claims that by the end of the 20th century, computers would have no more than a 30 percent chance of fooling an average interrogator. This does not mean that if a digital computer fools the interrogator 30 percent of the time it passes the test. It merely states the probability of the computer fooling the interrogator as 30 percent, i.e., the interrogator has a 70 percent chance of making the correct identification.

Concerning the average interrogator, note that they (in Turing's point of view) must not be an expert in digital machines. Turing imposes this restriction because: 1) it is obvious that a specialist in digital machines would know how to identify their peculiarities, thus facilitating the correct identification of the players; 2) it would be possible for a specialist to predict the behavior of the digital machine, i.e., they would know how to identify the functioning rules of the digital machine and, as to Turing (1948), once they identify such rules, they would not consider the machine as intelligent. Regarding this, in the 1952 BBC debate Turing points out that

> […] as soon as one can see the cause and effect working themselves out in the brain, one regards it as not being thinking, but a sort of unimaginative donkey-work. From this point of view one might be tempted to define thinking as consisting of 'those mental processes that we don't understand'. If this is right then to make a thinking machine is to make one which does interesting things without our really understanding quite how it is done (Turing; Newman; Braithwaite; Jefferson, 1952, online).

That is, once we understand how our mental processes (perception, learning, language, thinking, intentionality, memory, emotion, etc.) happen, they may no longer be deemed intelligent. Considering this assertion, Turing seems to suggest that 1) a digital machine will only be intelligent if it manages to do "interesting" things (such as conversing and providing carefully considered responses to miscellaneous questions) without fully revealing how it does them, and 2) it is possible that human beings might not be inherently intelligent beings. In this sense, intelligence would be, as to Minsky (1988), nothing more than "[...] our name for whichever of those [solving-problem] processes we don't yet understand. [...] the very concept of intelligence is [...] like the concept of the *unexplored regions of Africa*, it disappears as soon as we discover it" (1988, p. 71).

Furthermore, Turing (1950) predicts that until 2000 it would be possible to discuss thinking machines with no risk of being contradicted. In the 1952 BBC debate, he states that a well-programmed digital machine could meet the requirements of his test in at least 100 years, i.e., around 2052. Copeland (2004) reformulated Turing's (aforementioned) prediction (1952) as follows:

> It will be at least 100 years (2052) before a computer is able to play the imitation game sufficiently well so that jurors will decide wrongly as often in man-imitates-woman

Revista Instante, v. 6, n. 2, p. 26 – 46, Jan./Jun., 2024
ISSN: 2674-8819   *Qualis A3*
Departamento de Filosofia, Universidade Estadual da Paraíba

imitation games as in computer-imitates-human imitation games, in each case no questions being barred (2004, p. 490).

In the same debate (1952), Turing affirms that the question of whether machines could really pass the test "[Is] not the same as 'Do machines think?'" (Turing; Newman; Braithwaite; Jefferson, 1952, online). With this statement, Turing presents compelling evidence that the digital machine passing the test and the digital machine engaging in human-like thinking are different things. From this statement, it is possible to affirm that passing the TT is not a sufficient condition for a digital computer to possess genuine intelligence. However, if the TT is so interpreted (which is possible), it follows that a digital computer that passes the TT is Turing-intelligent. In this sense, a digital machine is Turing-intelligent if, solely through imitation and simulation, it is capable of acting intelligently, performing tasks that typically demand human intelligence. In the context of TT, such tasks would include conversing coherently and providing skillful responses to all sorts of questions.

While discussing the TT, Turing (1950) was not extensive in details concerning the rules of the game, leaving aside some relevant information such as 1) the game duration (which I believe to be 5 minutes), 2) whether it is intended for a single interrogator or for a jury, 3) whether the interrogator or the jury are allowed to repeat questions that have already been asked during the test, and 4) whether the interrogator is allowed to play the game multiple times.

Despite extensive debate on the possibility of a digital machine satisfactorily executing the imitation game, the discussion persists[13] in the 21st century.

3 – REINTERPRETING THE TURING TEST

Some philosophers such as Ned Block (1981), Robert French (1990), and John Searle (1980) believe the TT proves that the digital computer can duplicate the human mind if it succeeds in the imitation game. In other words, if the digital computer passes the TT, it is

[13] The reader will find several papers referring to this question in Epstein; Roberts; Beber (2009), as well as in the special edition of the journal Minds and Machines, which was published in 2020, with the title *Rethinking, Reworking and Revolutionizing the Turing Test*, and in the special edition of the journal Philosophies, published in 2023, with the title *Turing the Philosopher: Established Debates and New Developments*. Moreover, Maguire; Moser; Maguire (2015) also address the issue.

Revista Instante, v. 6, n. 2, p. 26 – 46, Jan./Jun., 2024
ISSN: 2674-8819 *Qualis A3*
Departamento de Filosofia, Universidade Estadual da Paraíba

genuinely intelligent. That is the standard interpretation of the TT. Nevertheless, this is not accurate. In *Computing Machinery and Intelligence* (1950) Turing argues that

> It might be urged that when playing the 'imitation game' the best strategy for the machine may possibly be something other than imitation of the behaviour of a man. This may be, but I think it is unlikely that there is any great effect of this kind. In any case there is no intention to investigate here the theory of the game, and it will be assumed that the best strategy is to try to provide answers that would naturally be given by a man (1950, p. 435).

Thus, Turing is solely referring here to imitation and simulation. On the one hand, it is important to highlight that, in this context, imitation refers to the performance of an act stimulated by the perception of a similar act by something or someone. An example would be a child, who learns to talk by imitating the way adults do it. On the other hand, simulation involves employing formalizations (algorithms) in computers to imitate a real-world process or operation. In this context, when a physical system, namely, a computer, formally simulates a real physical system performing an action in the real world, it is not actually performing the same actions as the real physical system. An example would be the distinction between a real car, moving in the real world, and a driving simulator (a computational simulation of a car), which is not moving but is doing something formally equivalent in a fake (virtual) world.

That being so, Turing is putting these two definitions together. Therefore, when he affirms that the best strategy for the digital computer (to successfully pass the TT) is to try to provide human-like answers, such strategy would consist only of imitating and simulating the human ability to properly respond to questions. In this regard, in the debate recorded and broadcast by the BBC in 1952, Turing states: "the idea of the test is that the machine has to try and pretend to be a man, by answering questions put to it, and it will only pass if the pretence is reasonably convincing" (Turing; Newman; Braithwaite; Jefferson, 1952, online).

If the digital computer succeeds in the TT, such computer is only imitating and simulating the human ability to properly respond to questions. It is not capable of duplicating the human mind and possessing mental states. In other words, if such a computer passes the TT, to the human observers it will seem to possess traces of genuine intelligence (such as thinking, learning, reasoning, etc.), just because it was able to perform a task for which humans resort to intelligence to perform.

From that, we recall the definition of Turing-intelligence, namely, the ability that a digital computer has (solely through the resources of imitation and simulation) to perform tasks that require intelligence when performed by humans. In the TT context, we say a digital computer is Turing-intelligent if it is capable of (solely through the resources of imitation and simulation) conversing with human beings in the same way humans do with each other. However, it does not imply that the digital computer has mental states or even a conscious mind.

Jack Copeland corroborates this view when, in his paper *The Turing Test* (2000a), he affirms that the TT was (and still is) misunderstood by philosophers, who believe it aims to define intelligence and thinking. Throughout the paper, Copeland points out that it was not Turing's aim at all. About that, in the debate recorded and broadcast by the BBC in 1952, it is possible to perceive Turing's view on this issue. When asked by Jefferson if he had a mechanical definition of what thought would be, Turing states:

> I don't want to give a definition of thinking, but if I had to I should probably be unable to say anything more about it than that it was a sort of buzzing that went on inside my head. But I don't really see that we need to agree on a definition at all. The important thing is to try to draw a line between the properties of a brain, or of a man, that we want to discuss, and those that we don't. To take an extreme case, we are not interested in the fact that the brain has the consistency of cold porridge. We don't want to say "This machine's quite hard, so it isn't a brain, and so it can't think." I would like to suggest a particular kind of test that one might apply to a machine. You might call it a test to see whether the machine thinks, but it would be better to avoid begging the question, and say that the machines that pass are (let's say) Grade A machines (Turing; Newman; Braithwaite; Jefferson, 1952, online).

Thus, the TT is not a test meant to define intelligence and thinking. According to Harnad's thoughts on Turing's 1950 paper,[14] which was reprinted, with comments, in Epstein, Roberts, and Beber (2009), what Turing "[...] will go on to consider is not whether or not machines can think, but whether or not machines can do what thinkers like us can do [...]" (Harnad *apud* Epstein; Roberts; Beber, 2009, p. 23).

It should also be noted that the purpose of the TT is not to assess the intelligence level of the digital computer involved, but rather to identify intelligent acts (Turing-intelligence) in such digital computer, i.e., to answer questions addressed to it in like manner as humans would. That being so, in the paper entitled *A clarification on Turing's test and its implications for machine intelligence* (2015), the researchers Phil Maguire, Philippe Moser, and Rebecca Maguire state that Turing's idea was

[14] Turing's *Computing Machinery and Intelligence* (1950) was reprinted in Epstein; Roberts; Beber (2009). Throughout the paper there are comments by Kenneth Ford, Clark Glymour, Pat Hayes, Stevan Harnad, and Ayse Pinar Saygin.

Revista Instante, v. 6, n. 2, p. 26 – 46, Jan./Jun., 2024
ISSN: 2674-8819 *Qualis A3*
Departamento de Filosofia, Universidade Estadual da Paraíba

> [...] about the possibility of intelligence being evidenced *in practice*. His point was that, although intelligence is a Platonic ideal (i.e. cannot be decided in practice), it is somehow manifested in finite objects, meaning that finite tests can detect it with high confidence (Maguire; Moser; Maguire, 2015, p. 318).

That being said, regarding what has been pointed out so far, when Turing refers to intelligent digital computers, he is referring to digital computers that are capable, uniquely through imitation and simulation, of manifesting intelligent acts,[15] i.e., Turing-intelligent computers. Furthermore, it seems that Turing was foreseeing what we know now as machine intelligence in computer science, namely, “[...] advanced computing that enables a technology (a machine, device, or algorithm) to interact with its environment intelligently, meaning it can take actions to maximize its chance of successfully achieving its goals” (Enders; King; Murray; Koop, 2016, online).

4 – REBUILDING THE TURING TEST: THE IDEAL TURING TEST

Although I have presented an alternative interpretation for the TT, challenging the standard interpretation this argument does not provide evidence that the TT is feasible or achievable in practice. Therefore, this section aims to support the hypothesis that the TT is indeed possible in practice. To this end, it is necessary to revisit the two versions of the TT presented by Turing (1950, 1952) and attempt to fill the gaps he left, to present a new, more comprehensive, and practicable version of the test.

The present version of the test, which I refer to as the Ideal Turing Test (ITT), is based on the two original versions proposed by Turing (1950, 1952) (as previously presented), with some modifications. The ITT is conducted in rounds of five minutes[16] each and includes two fixed participants, namely, A and B, where A is an AI system[17] and B a human being (gender is irrelevant), as well as one variable participant, C, who is an average human interrogator (gender is irrelevant). This prevents human participants from being mistaken for AI systems.

[15] Observe that exhibiting intelligent act is not the same as possessing genuine intelligence.
[16] The five-minute time limit is adopted to maintain this version of the test as closely aligned as possible with the original version described by Turing (1950). However, in an implementation, the time limit may be adjusted according to the specific implementation requirements.
[17] By AI system, we mean any AI system capable of interacting with humans, such as Large Language Models (LLMs) and even virtual assistants (when programmed to try to pass the ITT), capable of carrying out this interaction via voice and/or text.

In this setup, A and B participate consistently in all rounds, while C varies at the end of each round, meaning the interrogator changes after each round. All three participants are placed in separate booths and communication occurs solely through typed messages.

The role of C in the game is to formulate a variety of questions to A and B and, based on the responses received, identify at the end of each round which of the two participants is the AI system and which is the human. If the interrogator cannot determine which participant is a human and which is an AI system, they should refrain from guessing and report their inability to make the identification. The role of A is to attempt to deceive C (so that C makes an incorrect identification), and the role of B is to always tell the truth (so that C makes the correct identification). It is important to note that C is permitted to ask any type of question[18] to A and B.

It is clear that A is the AI system and B is the human. However, although the interrogators will be informed of the test rules beforehand, during the test, the participants will not be referred to as A and B, since this would immediately reveal their identities. The interrogator will only know that they are interacting with two participants, one of which is an AI system and the other a human, and they will be referred to randomly (not respectively) as Participant 1 and Participant 2. This prevents the interrogator from identifying which is which.

At the end of each round, C will receive an identification sheet[19] and must indicate which of the two participants, 1 or 2, is the AI system (Participant A) and which is the human (Participant B). If, at the end of a round, C is unable to make an identification (marking "inconclusive" on the sheet) or makes an incorrect identification, then we say the AI system has passed the ITT. It is noteworthy to mention that if an AI system passes the ITT, it is considered Turing-intelligent at one of the levels of Turing-intelligence (ranging from 01 to 06), based on the percentage of interrogators who incorrectly identified it as human.

Note that this test does not aim to provide an operational definition for intelligence, nor does it seek to evaluate whether the AI system involved possesses human-level intelligence

[18] As Turing (1950, 1952) states.
[19] On the identification sheet, each interrogator will find two fields that must be filled out to indicate who Participants 1 and 2 are, i.e., who is the AI system (Participant A) and who is the human (Participant B). For example: Participant 1: _____________; Participant 2: _____________. If the interrogator cannot make the identification, they should select the following option: [ ] Inconclusive.

Revista Instante, v. 6, n. 2, p. 26 – 46, Jan./Jun., 2024
ISSN: 2674-8819 *Qualis A3*
Departamento de Filosofia, Universidade Estadual da Paraíba

(genuine intelligence). Furthermore, the test is not designed to evaluate humans,[20] as suggested by French (1990). The objective is to assess how well AI systems can mimic human-like interactions. In other words, the test aims to determine if AI systems can, solely through imitation and simulation, perform tasks associated with human intelligence, particularly engaging in meaningful dialogues and providing appropriate responses to questions.

Having defined the ITT, it is important to clarify certain points for better understanding. These will be addressed in the following subsections.

4.1 – What Questions Can Be Covered in the Test?

Considering the information presented in Section 1, C can ask any sort of question (with no content restrictions) to Participants 1 and 2. The only limitation is that C cannot request practical demonstrations from them due to reasons outlined by Turing (1950). Additionally, C is permitted to use any device (providing there is no violation to this limitation) that may assist in better identifying Participants 1 and 2. This includes repeating previously asked questions within the same round, if deemed useful, to observe how Participants 1 and 2 respond to this situation, which may (or may not) aid in their correct identification.

4.2 – On Turing-intelligence Levels

As previously mentioned, the AI system which successfully passes the ITT will be considered Turing-intelligent and classified into one of the 06 Turing-intelligence levels. This classification is based on the percentage of interrogators who failed to correctly identify the AI system at the end of the test. The table below illustrates the classification criteria.

TURING-INTELLIGENCE LEVEL	PERCENTAGE OF INTERROGATORS FOOLED BY THE AI SYSTEM
01	01-10%
02	11-20%

[20] French (1990) suggests that only humans can succeed in the TT. That being so, the TT would not be a tool to identify if the AI system is intelligent, but rather if it is human. As AI systems are not human, no AI system can succeed in the TT.

Revista Instante, v. 6, n. 2, p. 26 – 46, Jan./Jun., 2024
ISSN: 2674-8819 *Qualis A3*
Departamento de Filosofia, Universidade Estadual da Paraíba

03	21-30%
04	31-40%
05	41-50%
06	Over 50%

If the AI system involved deceives 01-10% of the interrogators, it is classified as Turing-intelligent level 01. If it deceives 11-20% of the interrogators, it is classified as Turing-intelligent level 02, and so on. If the AI system deceives over 50% of the interrogators, it is classified as Turing-intelligent level 06. It is important to emphasize that the classification of an AI system as Turing-intelligent (regardless of the level) is relative to the observer (the interrogator), since each interrogator's judgment is subjective. This is the very purpose of the Turing-intelligence levels. Without these, the amount of incorrect identifications needed for the AI system to be considered Turing-intelligent would be unclear. Would it be enough for 1 interrogator out of 100 to make an incorrect identification? Thus, the levels of Turing-intelligence were established to provide a clearer classification framework.

Regarding Turing-intelligence levels, AI systems classified at levels 01-03 are considered weak systems. These were unable to deceive more than 30% of the interrogators, meaning at least one of their responses has been deemed unsatisfactory[21] by most interrogators. Systems classified at levels 04-05 are considered intermediate systems. While performing significantly better than weak systems (with fewer responses deemed unsatisfactory), they still fail to deceive more than 50% of the interrogators.[22] Systems classified at level 06 are considered strong systems.[23] They manage to deceive more than 50% of the interrogators, which is an extraordinary performance, given that their responses would be considered nearly indistinguishable from those of humans confronted with the same questions.

[21] For an unsatisfactory answer, we mean an answer that would not supposedly be given by humans when confronted with the same questioning.
[22] The AI system that manages to fool 31-50% of interrogators will be considered a system that performed well in the test.
[23] Note that what we mean by weak, intermediate, and strong systems has no connection with Weak and Strong AI as defined by Searle (1980).

Nonetheless, it is important to note that the same AI system can be classified at different Turing-intelligence levels, as this classification largely depends on the observation of the interrogator. Hence, it is possible (even likely) for the same AI system to be attributed to different levels of Turing-intelligence, depending on the number of interacting interrogators. For instance, suppose only 1 interrogator participates in the ITT and, after conversing with Participants 1 and 2, incorrectly identifies Participant A as a human. In this case, Participant A (the AI system) would be classified as Turing-intelligent level 06, having deceived 100% of the interrogators. Now suppose the same test being applied with 100 interrogators and, after conversing with Participants 1 and 2 (the same ones as in the previous example), only 1 of them mistakes Participant A for a human. Here, Participant A would be classified as Turing-intelligent level 01, having deceived only 01% of the interrogators. This exemplifies the assertion of the classification of AI systems into a level of Turing-intelligence being largely dependent on the observer.

4.3 – On the AI System Capable of Passing the ITT

As seen in the previous section, AI systems that succeed in the ITT can deceive at least 01% of the interrogators. Consequently, it is plausible to believe that the AI systems with the best performance in the ITT (being classified at Turing-intelligence level 06) would be capable, through the aid of machine learning and deep learning (considering the current advancements in these AI models), of processing natural language and studying interactions between people. This system would analyze how individuals respond to one another and determine what makes such responses genuinely human to propose responses indistinguishable from those of a human.

It is important to clarify that the ITT does not aim to identify which AI model the system under test employs. Rather, given the advancements in machine learning and deep learning (in terms of their current computational capabilities), it is tempting to believe that an AI system running such algorithms has a high likelihood of succeeding in the following aspects: processing natural language, studying and analyzing human interactions from an extensive database, focusing on how humans respond to each other's questions, and understanding what makes these responses characteristically human. By successfully performing these operations, such AI systems would likely respond remarkably similarly to humans. This would enable them to succeed in the ITT and potentially be classified as Turing-intelligent level 06.

CONCLUSION

From the facts presented herein, the reader can realize that once we accept the interpretation of the TT I gave in this paper (a non-standard one), the purpose of the TT substantially changes. That being so, I argued, as research results, that 1) It is possible to interpret the TT to conclude that it was not developed to assess whether the digital computer involved possesses genuine (human) intelligence or not, but rather to evaluate if it possesses what I call Turing-intelligence. In other words, I argued that it is possible to interpret the referred test to conclude that a digital computer succeeding in it does not imply that such digital computer possesses genuine intelligence, but rather that it possesses Turing-intelligence; 2) The Turing Test is possible, i.e., achievable in practice, given some modifications, resulting in a new version of the test named the Ideal Turing Test (ITT).

Concerning result 1), the standard interpretation asserts that Turing's intention with the TT is to demonstrate that passing the test is a sufficient condition for a digital computer to possess genuine intelligence. Conversely, the non-standard interpretation posits that passing the TT is a sufficient condition for a digital computer to possess Turing-intelligence. Within the context of the TT, a digital computer is considered Turing-intelligent if it can, solely through imitation and simulation, engage in dialogue with humans in the same manner humans converse with one another. In other words, a digital computer is deemed Turing-intelligent if it succeeds in the TT. However, it is important to highlight that asserting a digital computer is Turing-intelligent does not imply that it possesses genuine intelligence. Furthermore, the concept of Turing-intelligence aligns with the definition of AI proposed by Copeland (2000b).

Regarding result 2), based on the non-standard interpretation of the TT, it was possible to lay bare its feasibility. Taking the two versions of TT proposed by Turing (1950, 1952) into consideration, a new version of TT, namely the ITT, was formulated and presented. The TT was thus reformulated (based on some adjustments made to its structure, resulting in the ITT), and the Turing-intelligence levels were defined. Additionally, I approached the question about which AI systems could pass the ITT.

It is important to clarify that the aim of this research is not to endorse or criticize the standard interpretation of the TT. Instead, it seeks to demonstrate that, through an analysis of the foundational texts *Computing Machinery and Intelligence* (1950), and *Can Automatic Calculating Machines be Said to Think?* (1952), it is possible to make a distinct interpretation

Revista Instante, v. 6, n. 2, p. 26 – 46, Jan./Jun., 2024
ISSN: 2674-8819 *Qualis A3*
Departamento de Filosofia, Universidade Estadual da Paraíba

of the TT from the standard one. Notably, interpreting the TT as presented herein (a non-standard interpretation) leads to the conclusion that an AI system passing this test is not a sufficient condition to establish that the system possesses genuine intelligence. Rather, it suggests that the system possesses Turing-intelligence.

REFERENCES

BLOCK, N. Psychologism and Behaviorism. *Philosophical Review*, no. 90, 1981.

COPELAND, J. *The Essential Turing*: Seminal Writings in Computing, Logic, Philosophy, Artificial Intelligence, and Artificial Life: Plus the Secrets of Enigma. New York: Oxford University Press, 2004.

___________. The Turing Test. *Minds and Machines*, 10, pp. 519-539, 2000a.

___________. What is Artificial Intelligence? *AlanTuring.net*: Reference Articles on Turing, 2000b. Available in: <http://www.alanturing.net/turing_archive/pages/Reference%20Articles/What%20is%20AI.html>.

ENDERS, S.; KING, C.; MURRAY, S; KOOP, A. "Machine Intelligence, Artificial Intelligence & Machine Learning". *Alberta Machine Intelligence Institute*, 2016. Available in: <https://www.amii.ca/latest-from-amii/machine-intelligence-artificial-intelligence-machine-learning/>.

EPSTEIN, R.; ROBERTS, G.; BEBER, G. *Parsing the Turing Test*: Philosophical and Methodological Issues in the Quest for the Thinking Computer. Springer, 2009.

FRENCH, R. Subcognition and the Limits of the Turing Test. *Mind*, vol. 99, no. 393, 1990, pp. 53-65.

HAUGELAND, J. *Artificial Intelligence*: The Very Idea. A Bradford Book. The MIT Press: Cambridge, Massachusetts, London, England, 1985.

MAGUIRE, P.; MOSER, P.; MAGUIRE, R. "A clarification on Turing's test and its implications for machine intelligence". *Proceedings of the 11th International Conference on Cognitive Science*, pp. 318-325, 2015.

MCCARTHY, J. "What is Artificial Intelligence?". *Artificial Intelligence*, 2007. Available in: <http://jmc.stanford.edu/artificial-intelligence/index.html>.

Revista Instante, v. 6, n. 2, p. 26 – 46, Jan./Jun., 2024
ISSN: 2674-8819 *Qualis A3*
Departamento de Filosofia, Universidade Estadual da Paraíba

MINSKY, M. *The Society of Mind*. New York: Touchstone, 1988.

SEARLE, J. R. Minds, Brains, and Programs. *Behavioral and Brain Sciences*, 3, 1980, pp. 417-457.

TURING, A. M. *Can Digital Computers Think?* Annotations of a Talk Broadcast on BBC Third Programme, 1951. Available in: <http://www.turingarchive.org/browse.php/B/5>.

__________. "Computing Machinery and Intelligence". *Mind*, New Series, vol. 59, no. 236, pp. 433-460, 1950.

__________. "Computing Machinery and Intelligence". In R. Epstein, G. Roberts and G. Beber (eds.) (2009), pp. 23-65.

__________. On Computable Numbers, with an Application to the Entscheidungsproblem. *Proceedings of the London Mathematical Society*, vol. 42, 1936, pp 230-265.

TURING, A. M.; NEWMAN, M. H. A.; JEFFERSON, G.; BRAITHWAITE, R. B. *Can Automatic Calculating Machines Be Said to Think?* Transcript of a broadcast discussion transmitted on BBC Third Programme, 1952. Available in: <http://www.turingarchive.org/browse.php/B/6>.

ACKNOWLEDGEMENTS
I would like to thank Vandikson Coelho Barbosa (vandikson.barbosa@ifpb.edu.br) for the proofreading and revision.

I – INFORMAÇÕES SOBRE O AUTOR

Possui Doutorado em Filosofia, pela Universidade Federal do Rio Grande do Norte, Brasil (UFRN, 2023). Possui Mestrado em Filosofia, pela Universidade Federal da Paraíba, Brasil (UFPB, 2016). Possui Graduação em Licenciatura Plena em Filosofia, pela Universidade Estadual da Paraíba, Brasil (UEPB, 2011). É integrante do Núcleo de Pesquisa em Filosofia, Interdisciplinaridade e IA – UEPB/CNPq. Foi integrante do Grupo de Pesquisa em Teoria da Computabilidade – UFPB/CNPq (2014-2020). Tem experiência na área de Filosofia (Lógica e Filosofia Analítica), com ênfase em Lógica Matemática, Teoria da Computabilidade, Filosofia da Inteligência Artificial, Filosofia da Mente e Filosofia da Linguagem, atuando principalmente nos seguintes temas: Lógica Clássica, Computabilidade Clássica, Inteligência Artificial (Teste de Turing) e Teorias Computacionais da Mente. Atualmente, é Professor do Departamento de Filosofia da Universidade Estadual da Paraíba, atuando nos cursos de Filosofia e afins. E-mail: victorpereiragomes@servidor.uepb.edu.br

II – INFORMAÇÕES SOBRE O ARTIGO
Recebido em: 20 de janeiro de 2024
Aprovado em: 06 de junho de 2024
Publicado em: 24 de junho de 2024