Artificial Intelligence and Human Intellect

Víctor Velarde-Mayol

Seton Hall University

victor.velarde-mayol@shu.edu

Abstract

Many philosophers, computer scientists, and cyberneticists still consider it possible that a computer, described as a Turing machine, can effectively have intelligence. They are not talking about the present but the future of a possible computer to have the same properties as human intelligence. In this brief paper, I will provide some basic arguments that prove the impossibility of such a thinking machine. These arguments will be presented in four parts: the mental experiment of the Chinese Room, the argument from exactness, the argument from phenomenology, and the argument from abstraction. The first argument proves that mechanical manipulation of symbols is not understanding. The second argument proves that logic and mathematics are exact, which is a quality of understanding missing in the material world. The third argument proves that logical laws are unconditional, while physical laws are relative to material conditions. Finally, the last argument proves that abstraction is an operation of the intellect that is required for creative decision making. Neither animals nor any material system manifests abstraction.[*]

Keywords: *Abstraction, artificial intelligence, exactness, intelligence, logic, phenomenology, Turing machine, universal.*

1. Introduction

The idea of Artificial Intelligence (AI) has become part of our popular culture. It is commonly understood as a system that is intelligent or to which intelligence is attributed, although we also have experienced AI systems that at first appear intelligent, but with the passage of time seem less so, and even with greater perspective may be viewed as completely lacking intelligence. This means that frequently intelligence is attributed to a cybernetic system for the simple fact that it performs complicated operations in superhuman time, and

[*] The author wishes to acknowledge and thank Dr. Justin Anderson, Dr. Maria Jakubik, as many blind reviewers for their service in proofreading this manuscript.

what previously seemed to be something wonderful, is devalued over time. But, can intelligence be identified by complicated operations? These complicated operations can provoke in the imagination the idea that something is intelligent, but I think it can be considered obvious that there is no qualitative difference between a complicated system and a simple one. The complication itself does not add a new quality (intelligence or something else) by the mere fact of being complicated, but it is acceptable that a complicated system can be called "metaphorically" intelligent.

Other times, intelligence is identified with the ability to make decisions. In this way, there are cybernetic systems that seem to decide an "output" given certain "inputs." I think it is clear that this is just another metaphor for what a decision is, if by decision we mean a free, creative act that includes at least active indeterminacy that is not compatible with a merely mechanical system, and even with a system that includes the mere passive indeterminacy of matter. However, this is enough to lead us to "imagine" a cybernetic system as intelligent, but only in a metaphorical sense.

The name of cybernetics comes from the Greek χυβερνήτης (steersman), the man who governs the steering mechanism, which is indeed one of the earliest and well-known forms of feedback mechanisms. In the jargon of philosophy, these feedback mechanisms are a type of automation different from the self-movement idiosyncratic of life as it is different transient and immanent motions. Although AI has been methodologically separated from cybernetics, for AI is now considered a part of computer science, while cybernetics is concerned with control and automation (Wiener 4ff), I will treat them as a unified discipline (Bringsjord & Govindarajulu, 2020). AI and cybernetics have in common their interdisciplinary and transdisciplinary character, elements of learning, recognition, adaptation, and control, which are very useful in neuroscience and neurophilosophy research. Ultimately, as far as our subject is concerned, the common essence of both could be put into systems of feedback loops, although obviously intelligence is not only reduced to that (Elman, 1990).

Is AI, which is understood as a sophisticated feedback system, the cause of intelligence? This is going to be the subject of our brief study: proving its impossibility. However, there are many scientists and philosophers who reject this impossibility and maintain a disconcerting certainty—based more on faith than on scientific evidence—that a cybernetic system can be intelligent in a strong sense and not merely metaphorically. The central thesis would be that the brain is the hardware, while the mind is the software. For example, Jerry Fodor (1935–2017)—considered the father of the philosophy of mind in America—conceived the mind as a language that can be reproduced

computationally following the model of a Turing machine. Certainly, this has not been proven; therefore, it remains a mere hypothesis: the hypothesis of the Language of Thought (Fodor, 1975). Daniel Dennett (1942–) maintains that all mental activity, all consciousness is nothing more than pure linear or parallel brain computation according to a program, which would be brain software (Dennett, 1991). Ned Block asserts that the brain is a syntactic engine driving a semantic engine. This means that syntactics is a form of semantics in some way, and consequently, meaning is a computational result (Block, 1990, 1995).
In this paper I am going to provide a series of arguments—just a few among the very many that exist—that prove the impossibility of an essentially intelligent cybernetic system, especially the idea that intelligence can be reduced to it. The choice of these particular arguments is due to their strong relationship to cybernetics.

2. The Chinese Room Experiment

This argument proves that mere manipulation of words (symbols) can be done without any understanding of the meaning of words. Therefore, a computational system of symbols cannot qualify as intelligent.

John Searle (1932–) designed an argument by which he proved the impossibility of deriving the property of being intelligent from a computational physical system (Searle, 1980). Although the argument is valid for any cybernetic system, it is generally presented using the Turing machine model.
As background, the Turing machine is considered the basic model for AI, and it would coincide with what the mind does. Alan Turing (1912–1954) provided this theoretical model accepted by many as the essence of computation (Turing, 1937). This model, named in honor of this mathematical genius, is an extremely simple machine consisting of two subsystems: a reading head and a data tape. Simplifying greatly, its characteristics would be the following:

1) An "infinitely" long tape divided into sections. Obviously, this tape is physically impossible, but the reason the tape is infinitely long is so that the machine is without storage limitations, at least "ideally" (this is a really weak point). A Turing machine is a computer but with an important characteristic: it has an indeterminately large hard disk. In reality, Turing did not think that a Turing machine would ever have to deal with infinitely long strings of symbols, but he wanted it to be able to deal with "arbitrarily long," but still finite, strings of symbols. This infinity is just a convenient and common mathematical fiction for an indefinite amount of storage.

2) Each of the sections of the Turing tape can be either empty or filled with a single symbol.
3) The Turing tape runs through the machine reading head, with one important restriction that only one section is under the machine reading head at a specified time. Obviously, the purpose is to register the symbol the section contains in an orderly way.
4) The machine reading head can also perform mechanical operations on the section that is being registered. It can move the tape forward or backward, and it can delete and write a symbol.
5) Each Turing machine is controlled by a set of instructions placed in a table. These instructions determine what the Turing machine will do when it finds a certain symbol in a certain section. For the purpose of the Chinese Room argument, this is the most important characteristic of a Turing machine (Bermúdez, 2020).

In summary, a Turing machine's behavior is completely determined by the machine instruction-table, its current physical state, and the symbol in the section it is registering. It seems to me obvious that there is no room for the machine to exercise "judgment." It is, in fact, purely mechanical in exactly the way required for an algorithm.

It was the logician Alonzo Church (1903–1995) who established the important computational thesis—Church-Turing thesis—that anything done in mathematics by an algorithm can be done by a Turing machine (Smith, 2007). In other words, a Turing machine is a system that can do anything that can be algorithmically computed (Church, 1941). This suggested an important conclusion for the hypothetical thesis that the mind is the software: intelligence is a form of information processing, which can be understood as an algorithm.
The analogy between mind and digital computer became the analogy between intelligence and the Physical Symbol System Hypothesis (Simon, 1966). According to this hypothesis, all intelligent behavior essentially involves transforming physical symbols according to rules. The most sophisticated version of the Physical Symbol System Hypothesis was developed by the philosopher Jerry Fodor, who used the Turing machine as the model for any other computer system. Fodor developed a subtle and sophisticated argument for why symbolic information processing has to be linguistic (syntax and semantics). He argued that the architecture of the mind is built around a language of thought (Fodor & Pylyshyn, 1988).

Using this last idea, Searle developed his Chinese Room argument. His accurate intuition was that at the heart both of the Physical Symbol System Hypothesis and the very detailed Language of Thought Hypothesis is a sharp distinction between the syntax of information processing (the physical manipulation of symbol patterns) and the semantics of information processing. Searle proved in his Chinese Room argument that this distinction is fatally flawed.

The Physical Symbol System Hypothesis holds that we exhibit intelligent behavior when we have systems that manipulate symbols according to rules. The Language of Thought Hypothesis is a particular way of applying this model of intelligent behavior. The Language of Thought Hypothesis also tells us what the rules are going to be like and how they will end up producing intelligent behavior. These rules are fundamentally syntactic, transforming the physical symbols in ways that depend solely on their physical characteristics. These transformations will produce intelligent behavior because syntactic transformations of the physical symbols mirror semantic relations between the propositions that give meaning to the physical symbols. But no machine built according to the Physical Symbol System Hypothesis could possibly be capable of intelligent behavior: manipulating symbols is absolutely insufficient for intelligent behavior. Not only processing symbols (syntax) is not thinking semiotically, but even thinking syntactically is not thinking semiotically (Weed, 2003).

Searle asks us to imagine a person in what he calls a Chinese Room. The person receives pieces of paper through one window and passes out little paper notes through another window. The paper notes have symbols in Chinese written on them. The Chinese Room, in essence, is an input–output system, with symbols as inputs and outputs, as with a Turing machine. The way the input–ouput system works is determined by an instruction manual that tells the person in the room which paper notes to pass out depending on which paper notes he receives. The instruction manual is essentially just a way of synchronizing input symbols with output symbols. It is not written in Chinese and can be understood and followed by someone who knows no Chinese. All that the person inside the room needs to be able to do is to identify Chinese symbols in some sort of syntactic way—according to their shape, color, position, for example. This is enough for them to be able to find the right output for each input—where the right output is taken to be the output dictated by the instruction manual.

Imagine, first, that the manual has been written in such a way that the inputs are all questions in Chinese and the outputs are all correct answers, in English, to those questions. For all intents and purposes, therefore, the Chinese Room is answering questions in Chinese. Let us place in the room a person who does not know any Chinese. All he is doing is following the instructions in the manual (which is nicely written in English). What the Chinese Room shows is that it is perfectly possible for there to be syntactic symbol manipulation without any form of intelligence or understanding. As long as the person in the Chinese Room follows the instructions correctly, the semantic relations between input and output will be “preserved.” And yet, the Chinese Room (always with the person inside) does not understand Chinese: we have here a complete and complex computational system with no intelligence at all with respect to Chinese.

If the Chinese Room does not understand Chinese, then it is not behaving intelligently in any sense. To someone outside the room it might look as if there is intelligent behavior inside. The machine does, after all, respond to the Chinese questions it is asked with answers in English that make sense. But this is just an “imaginative” illusion of intelligence. The Chinese Room cannot be behaving intelligently if it does not understand Chinese. This situation is a powerful counter-example to the Physical Symbol System Hypothesis as intelligent.

What is more striking is that this mental experiment would pass the Turing test. This famous test proposed by Alan Turing is based on the imitation game, which is basically to affirm that if something behaves intelligently, it is intelligent (Davidson, 1990; Traiger, 2000; Turing, 1950). The Chinese Room looks to the outsider as though it is intelligent, but from the inside it is not.

3. Argument from Exactness

Logical and mathematical laws are exact, existing only in the intellect, but material reality is inexact. From here it is concluded that any system based on material reality cannot be intelligent.

This argument was devised by the great philosopher Franz Brentano (1838–1917), cited very frequently in the philosophy of mind for his discovery of intentionality as an exclusive property of mental phenomena versus the complete lack of intentionality of physical phenomena. The present argument is not directly related to the intentionality of mental activities, but to another property: *the exact character of logical laws* (and analogous laws, such as mathematics). Brentano (1995) made this argument for every physical system, especially brain neurological processes (Book I, pp. 49–58), but it has special relevance for computational systems as physical systems. What I am going to present here is a modification of the original argument adapted to cybernetic systems.
Any AI machine is based on the efficient causality of the moving parts that integrate the system. Every algorithm implemented in a physical system is a causal chain of material agents. Given the concept of a Turing machine as the

core of a computer, it is obvious that the reading head that scans the data tape and moves back and forth according to a program of instructions, which works by moving parts of the system (moving parts can be also at the quantum level), is a causal system. One part of the system causes the movement of another part, forming a causal chain in series or in parallel. It is important to note that the algorithm mathematically considered is not the algorithm exemplified in the AI machine. The algorithm is "absolutely" exact, and only exists in the understanding, while the machine that implements it is necessarily subject to some inaccuracy.

We know that in deductive logic the conclusion has to be necessary, exact; there is no room for compromises or approximations. It cannot be said that if A = B, and B = C, then, more or less, A = C. The conclusion is absolutely exact, without there being accidental cases where it does not occur. But in physical systems this accuracy is missing: there is a degree of inaccuracy that is part of the body's nature, and therefore, if it is "physically" computed that A = B, and B = C, it is possible that the output fails—that is, that the machine registers that A is not C. Logically this is impossible, but physically possible, for the contingent physical causes fail, and the output could be different due to physical indeterminacy.

A simple syllogism in Barbara (1-AAA: first figure, universal affirmative mode) provides an example where the conclusion is absolutely exact given the truth of the premises:

> Whatever is composed of integral parts is contingent.
> Whatever is material is composed of integral parts.
> Therefore whatever is material is contingent.

When it is translated into symbolic logic, we graphically see its inferential accuracy:

$$(x)\ (Cx \rightarrow Dx)$$
$$(x)\ (Mx \rightarrow Cx)$$
$$\therefore (x)\ (Mx \rightarrow Dx)$$

The inference from this deduction is completely exact and necessary. We know the exactness of that inference in our internal experience, and this is independent of our biological and neurological condition. One may be ill and yet experience the accuracy of a syllogistic inference; even when one is dreaming the most extravagant adventure, the experience of the necessity and exactness of an inference is immovable, and if for whatever reason the dreamed story infers a (logically) contradictory conclusion, we know it to be false. This fact was already observed by Rene Descartes (1596–1650) long before

Brentano. Logical and mathematical laws are not altered by a bad nightmare; they are not even affected by alcoholic intoxication of the brain. In dreams, as in drug intoxication, what is affected is the perception of reality (organs of perception), not the internal experience of logical (and mathematical) laws. This contrasts with what would happen if a computer were affected by some hidden malfunction. To give some color to this observation, the great mathematician Niels Abel is said to have developed some of his mathematical findings while he was suffering alcohol poisoning.

What do we get out of all this for our argument? Logic, which occurs only in our understanding, is not affected by material conditions. It is in the human intellect where these laws have an exact character and never in a causal system of material order. For everything that is material is contingent (that is, it can fail, which is the origin of inaccuracy), while logical laws never fail. This means that the so-called logical operations of a physical-computational system are not properly logical (exact) because they are necessarily contingent. A computer, being contingent, will not only sooner or later fail, but it also has an internal principle of inaccuracy. A logical law, which can only occur in the intellect, enjoys absolute exactness without exceptions. This discrepancy is not one of degree but qualitative: a contingent causal system (a computer) can never reach the exactitude that occurs in the intellect that forms logical laws. In other words, an AI system can never be the subject of logical laws, and therefore is not intelligent, except in a metaphorical sense.

Logical laws have the property of being exact (let's say they have the quality of exactness), while physical laws are qualitatively inaccurate to some degree. This indicates a radical division between what a computer can do, and ultimately is, and what logic is. Logic exists only in the intellect; in "material" reality we find patterns of behavior that can fail.

Albert Einstein also recognized this important difference when he indicated that there is a substantial difference between mathematical equations and physical reality: the former being exact, the latter inaccurate (Einstein, 1921). The former only occur in an intellect free from material processes; the latter is characteristic of material contingency.

It can be objected that the human intellect also fails; we commit shameful fallacies. Certainly, but they are failures due to the fact that the human intellect depends on the body (corporeal organs) that supplies the contents (not the logical form). This is obvious in tiredness, haste, incomplete experiences, and malfunctioning of the sense organs, but the intellect, in itself, when it forms logical laws, does so independently of the sensible contents: *the senses do not supply the logic*. That A = B, and B = C, then A = C is independent of the content of A, B, and C; they are relations of equality formed only in the intellect. The intellect accidentally fails due to what the sense organs supply us, but not essentially. An intelligence without a body, like the angelic intellect, would never fail. If the human intellect fails, it is not because it is an intellect but because it is human, that is, dependent on what the organic senses supply it (Aquinas, 2012, I.84.1). Logical errors come from "trusting" ill senses and bad will.

Neither the senses (organic faculties) nor any material system (a Turing machine, for example) can supply any logical law.

Another conclusion of this argument is that if, due to an impossible hypothesis, logic depends only on computational operations, logic would no longer be assured because at some point it would fail, and there would be no way to rectify that failure. There simply would be no logic. There is thus a radical division between the exactitude of the logos and physical inaccuracy.

Another conclusion of this argument is that if the entire material world is contingent, the intellect cannot be an effect of matter (Stephan & Klima, 2020). This is the famous thesis of Aristotle (384–322 BC) that the intellect is somehow separate from matter—that is, that it cannot be mixed with matter (Aristotle, 2016, 430a10–25).

4. Argument from Phenomenology

This argument starts from the phenomenological evidence of a logical law: universality without material restrictions. If one identifies logical laws with physical laws (of a system of AI), the universality is lost, and logic becomes relativistic to material conditions. Because this universality only is possible in the intellect, a physical computer cannot be logical, so it cannot be intelligent.

In continuity with the previous argument, but with greater scope, Edmund Husserl (1859–1938) dedicated several arguments against the psychological interpretation of logic or logical psychologism (Husserl, 2008). They are extremely valuable when applied to cybernetic systems, because they have similar problems from the perspective of the phenomenology of pure logic. It must be taken into account that Husserl, in addition to being one of the most important philosophers of the last century—founder of phenomenology—was originally a mathematician. His argumentation combines the natural mathematical instinct for proof with the experiential intuition of phenomenology. The argument is found in the first volume of his *Logical Investigations*, under the title of "Prolegomena to Pure Logic," especially in the fifth chapter. What I will develop here is a modification of the argument adapted for AI in the strong sense of the word "intelligence."

Take for example the law of non-contradiction. Its logical formulation is that it is impossible that "A" is "non-A" in the same sense: ~ (A • ~ A). This logical law has a universal value *without restrictions*, that is, without material conditions of any kind, and as a logical law it is formed "only" in the human intellect with the property of "universality." Now, if this logical law could also be formed in a computational system, it would be reduced to its physical individuality and to failures due to material contingency; that is, it would lose its universality, since it would be conditioned by the material restrictions of the computer. A logical law would be reduced to a physical law, the laws that rule the physics of a computer.

The logical laws are universal; they are valid for all situations. *Universal* here means that logical laws are fulfilled at all times, in all places, and in all cultures. On the contrary, the so-called physical laws are not universal in three cases: (a) they are approximate, there is indeterminacy; (b) they are not always fulfilled in all space and time, since, as is admitted in contemporary physics, many of the laws that are valid in the present time, would not be valid in another previous time, such as at the beginning of the Big Bang, and probably, many of the physical laws that apply to the visible world would not be valid for the world of dark matter and dark energy; and (c) physical laws are only valid for this current universe; there are infinite possible universes with different properties and different physical laws. This is in contrast with the unrestricted universality of logic valid for any possible world, which implies independence from material reality. The so-called physical laws, being restricted or conditioned by material indeterminacy, cannot be truly universal; therefore, it would be better to call them patterns of physical behavior.

Logic and mathematics are universal; physical reality is not, but individual and contingent. The logic of a theory is perfectly universal when it remains in the realm of the intellect, but the closer it gets to matter, we find it not only imprecise, but also lacking in universality. A physical AI system, which is a material causal system, could never formulate the exactness and universality of logic, it always has residual inaccuracy, uncertainty, and indeterminacy.

Everything that is made by or mixed with matter necessarily suffers from a degree of indeterminacy, which Werner Heisenberg (1901–1976) formulated at the subatomic level (Sen, 2014). Aristotle, long before, formulated it at all levels: the origin of the fallibility of physical causality is due to the essential indeterminacy of the fundamental matter (Aristotele, 1996, II.4). This would be the root of the contingency of physical processes incompatible with the universality of logical laws.

Husserl’s conclusion is that if logical laws are identified with a physical system (cerebral or computational), it would fall into a completely unacceptable *logical relativism*. Logical laws are not regular patterns of behavior (physical or neurological).

The operation of a computer depends on its physical condition, quantum level, external causes, and other factors. In contrast, no logical law has that dependence: it is valid without material conditions, which is an obvious phenomenological fact. If a logical law depended on a computer, that law would be conditioned to the physical state determined here and now of the computer. The same should be said of the brain as a cybernetic system: the logical laws would be conditioned by the physical (neurological) state of the moment. This poses two serious problems.

First, the logical law, which is universal without material conditions, would be transformed into a law with conditions of a particular singular moment. It would only be valid for those conditions, not for all possible ones. The logic would be relative to some physical conditions—which in themselves are virtually unrepeatable: this is a logical relativism, which destroys the essence of logic.

Second, how do we know that the system is working properly? To find out we would need to resort to the logical design (the idea) formed in our mind to contrast it. But if the mind is nothing more than a material cybernetic system based on the Turing machine, a causal physical system, then we find ourselves in an undesirable vicious circle. How to justify that something is a logical law if there is no universality? To say that logic is relative to material conditions is *eo ipso* the negation of the essence of a logical law, which is in contradiction with the phenomenological evidence that we have: a lived experience of the universal validity of any logical laws.

If a particular machine can form a logical law, a different machine will do it differently, with some variations—as few as you want, but variations. As a result, the unconditional universality of the logical laws is destroyed; that is, these laws would not be valid for every system. And if the mind is nothing more than a neurological system after the model of a Turing machine, no logical law can maintain its universality. The logical laws would depend on the neurological structure, its health, the environment, the chemistry at a given moment, the quantum state of its components, and ultimately on stochastic conditions proper to the contingency of the material world. In short, they would fall into an unwanted logical relativism. The law of non-contradiction would no longer be the same for all possible conditions. It would depend on the physical conditions of the system, and syllogisms as basic as the one mentioned above—the Barbara, or 1-AAA—would not be valid in all circumstances; it would depend on the circumstances of the computer.

It is phenomenologically evident that logic is universal and formed in our intellect. This universality means that it is free from material conditions. A law formed in a physical system (an AI based on a Turing machine or a neurological system, for example) is a law restricted to material conditions. Therefore, a logical law cannot be formed in a physical system, and no AI system can be properly intelligent.

A consequence from this conclusion is the following. Logical laws "occur" only in the intellect, and they cannot be conditioned by a material subject; the intellect itself must be free from matter. As shocking as this may sound, it is the natural conclusion required by the quality of unrestricted universality. A different problem would be to explain how the intellect, free from matter, can be in a human body. Max Planck captured it perfectly when he stated that this is one of man's oldest riddles: how to harmonize the mind with the physical world (Planck, 1932, pp. 107ff)—a philosophical problem of great interest but beyond the scope of this brief paper.

5. Argument from Abstraction

A creative decision to resolve an unexpected new problem implies the capability to form abstracts. But an abstract is not part of our reality. Abstracts are only formed in the intellect. So, a physical computational system cannot form abstracts (they are not part of the physical reality), and as a consequence, they are not intelligent.

Abstraction is the obtaining of a general concept from individuals. More specifically, it is the mental extraction of the nature of things that is common to various individuals (the abstract "man" from Peter, John, etc.). This abstracted nature is what is called a *universal*, which, as has been said, exists only in intelligence. There are no abstracted things in the real world. There is Peter, John; there is no the abstract "man."

Abstraction is very important to making decisions and solving problems. An argument can be made, based on experiments, that there are decisions that require abstractions. As the abstract does not exist in physical reality, but only in the understanding, then it can be concluded that there are intelligent decisions that cannot be made with (physical) cybernetic systems.

Intellectual abstractions are the basis of creative decision processes, which are in a way processes for solving new problems. Therefore, a decision theory that does not consider abstractive capacity does not seem to be a complete theory. This can be seen experimentally, and applied, without much change, to so-called cybernetic intelligent processes. Let's see this with two experiments: the first can be solved almost mechanically (although it doesn't appear feasible), but the second completely surpasses what AI can do.

In the experiments of Wolfgang Köhler (1887–1967), various species of animals, placed in their respective cages, see food through a grate. To get hold of it, they just have to turn around and go out the back gate, in the opposite direction from where the food is; that is, the animal has to move away from the food and then take it by going in the opposite direction. Depending on the animal, there are different responses. The hen goes back and forth by the back gate but cannot find the way out to take the food. Another animal with more cognitive ability, such as the dog, can solve the problem only if the food is far enough away from the grate, because if it is too close, the dog usually fails to capture the food. Not so with chimpanzees, better cognitively gifted: they end up solving the problem by the simple process of going out the back gate to take the food (Köhler, 1917/1976).

There is a problem of decision theory here, which is relatively easy to solve from the point of view of intellectual abstraction that these animals lack, but which becomes extremely complicated when one has to resort to random processes (the chicken), spatiality (the dog), or strategy (the chimpanzee). The chimpanzee strategy has traditionally been explained with the theory of the internal senses (imagination, memory, and estimative). The estimative faculty of the chimpanzee, much superior to that of other animals, allows what has been called *animal reason*, which is a comparison between singulars to solve a problem. But it is not an intellectual abstraction because the chimpanzee does not grasp the abstract—the universal—of space. In reality, the animal lacks creativity to resolve new problems. This can be easily seen in the following experiment.

The experiment of Ivan Pavlov (1849–1936) allows us to assess the power of abstraction or lack of it in animals, which will be very useful to build our argument about the impossibility that a material causal system can have any abstractive power (Windholz, 1987).

A chimpanzee is trained in two skills: (a) getting water from a tank with a bucket to put out a fire that prevents it from taking the desired food; and (b) getting water from the lake to cool off, pouring water on itself. The animal is placed on a sufficiently wide raft in a lake, and at one end of the raft the desired food is placed but obstructed by a flame that impedes its direct capture. On another raft, at some distance, the water tank is placed. This raft is accessible through a walkway. The chimpanzee, left alone on the food raft, crosses to the other raft to get water from the tank with the bucket, and goes back across the walkway to put out the fire and thus get the desired food. The chimpanzee does not think of taking water from the lake (easier to acquire) to put out the fire. But, even more interesting, if there is no water in the tank, it does not occur to the chimpanzee to take water from the lake to put out the fire and thus obtain food.

What has happened here? Simply, the animal has not been able to understand (abstract) the nature of water. If the chimpanzee had understood what water is, he would have realized that the water in the tank to put out the fire, and the water in the lake to cool off is the "same" water; they have the "same nature." But for this, the animal would have to have the ability to abstract the nature of water and verify that it is the same in both cases, something that does not occur to the chimpanzee. It must be taken into account that having the "same nature" is already an abstract (a universal) since the water in the bucket and the water in the lake are different individuals (not the "same" individual). The chimpanzee does not grasp the universal. It does not abstract. "Being of the same nature" is only formed in an intellect, which the animal obviously lacks.

If the mind were an instance of software exemplified in the hardware of the brain, then it could never abstract the common nature, the universal, as there are no abstracts in material reality: the brain as a neurological system is not the subject of an abstract universal, nor a physical AI system. The mere fact of abstracting something common is already forming a universal abstract, which has no place in the real world. This is easy to see if we try to describe a common human nature (an abstract): it is not tall, not short, not this color or the other, not this weight or the other, but it is valid for every concrete human with a size, color, and weight. Such an entity cannot exist, but it can be understood and applied to the corresponding individual, as the reader is now understanding it. If it cannot exist in the physical world, it cannot be realized in a Turing machine either.

In conclusion, intelligence allows the creation of new solutions for unexpected problems because of the ability to form abstract, common, universal concepts of reality. Abstraction as a source of creative decision making is not part of any material causal system. So, it is not part of any Turing machine system.

Abstraction is one of the operations of human intelligence, but not of a material system.

6. Some Philosophical Considerations

The four arguments briefly presented above imply a basic idea: that intelligence is ontologically immaterial and inorganic. Entering into this interesting philosophical subject is outside the modest limits of this short work and not absolutely necessary to the goal of this paper. I will limit myself to three points: the immaterial nature of knowledge, the anti-materialist paradox of the phenomenon of knowledge, and the ethical problem of reductionism of the intellect to the physical causality of a computer.

Intelligence is not explicable only through processes of material causality. This means that all so-called "natural" causality, efficient causality between bodies, implies an "exteriority" of the effect with respect to the cause. This is what in classical philosophy is called transient causality, whose model is that of the agent causing the movement. The motor agent causes the effect that is a movement, but both are different, mutually external. On the contrary, in knowledge the known is in the knower, not outside, but remains in the act of knowing itself. This is what in classical philosophy is called "immanent operation": the effect of the act of knowing is the act of knowing itself. This does not happen in any material agent, whose effect is always different from the activity of causing.

When fire is known, for example, it has to be in some way in the act of knowing, obviously it is not materially, because one does not get burned by knowing the fire. Knowledge of fire does not burn, but any material contact with fire does. This implies that the fire possessed by knowledge is in a different form, that if it is not material, it will have to be in a non-material form. This is why it is said in classical philosophy that the essence of knowledge is immateriality. For there to be knowledge, the object must be possessed immaterially. In philosophical terminology, the object and the act of knowing are said to be one and the same "formally", not materially.

An interesting consequence also follows from the above when the concept of intentionality is introduced (Searle 1983): the immaterial possession of fire intentionally refers to fire. Certainly, to know fire is to "possess" it immaterially (an immaterial identification with the known object) and to refer to that fire "intentionally." The phenomenology of intentionality, that is, the description of the phenomenon of intentionality manifests a "distance" between the act of knowing fire and known fire. This "distance" is not an effect of the act of seeing, since the act of seeing does not cause anything in the fire, or what is the same, the fire is not affected by the mere fact of being known. But this raises a problem that a materialistic conception of knowledge cannot solve: how is it possible that human knowledge is both a distance and an identity? Human knowledge is both an intentional distance and an immaterial identification with the object. The notion of "distance-identity" in which our knowledge consists is a materialist puzzle, an insurmountable stumbling block that manifests the internal contradiction of the materialist explanation of knowledge (Millan-Puelles, 1967).

Finally, I would like to suggest a moral problem in reductionism of intelligence to the mere material causality of a computer. What is the real reason for this reductionist interest? Is it merely a scientific or philosophical interest, or is there a moral interest in favor of reductionism? Is there an "interest" in the truth in this reductionism? There are more than enough voices that assure that it is a genuine interest in the truth, but I suspect that there is something else. The "interest" in something carries a certain moral value (a moral good) in human beings, so being interested in the truth is certainly a moral obligation. In the history of philosophy, there are countless valuable arguments against materialist reductionism, why then does there exist in our time this insistence on a materialist reductionism? One thing seems obvious to me, namely there is a degradation of the human person to a mere material component. For centuries, the dignity of the person was based on its rationality; however, with the reductionist project, the moral dignity of the person is dramatically "reduced" to the status of a mere "object", something merely "useful", which are characteristics of mere matter. As Kant stated forcefully, and before St. Thomas Aquinas, only a good will is absolutely good, which is one of the central elements (not the only one) of the dignity of the person: dignity is either absolute or it is not. Rocks, electrons, gases, circuits, etc. are at the level of means, mere objects, useful, that are used for an end, and when they do not serve they are discarded. Not so with the (absolute) dignity of the human person, for persons are not disposable, even in the case of useless; a person's dignity is not valued for what it produces. It has an absolute dignity, separated from the results, it is always an end in itself and never a mere means to achieve something. Materialistic reductionism eliminates that dignity by making human persons disposable, discarded if their lives are not useful, as it is the case with any material process. This is the true moral problem of the aforementioned reductionism: the degradation of the human person.

7. Conclusion

1. AI based on mere manipulation of symbols, as is the case of the paradigm of a Turing machine, cannot explain understanding. There is no syntactic replacement for semantic. A machine that translates from one language to another is not properly intelligent, for it can be done without “understanding” the language. An “intelligent” system without understanding is a perfect oxymoron.
2. Logical laws are exact and universal, while the so-called physical laws, by which a computer is governed, are never exact or universal. Absolute exactitude and universality exist only in intelligence. Therefore, if logical laws were the product of the cerebral physical system, which in turn behaves like a Turing machine, a causal physical system, the exactness and universality of logic would never be obtained, and it would fall into an undesirable logical (and mathematical) relativism. Logical laws are not physical laws, nor neurological laws, but laws of how we reason.
3. A cybernetic system to which decision properties are attributed has a fundamental limitation: the inability to abstract from the individual. This is essential for out-of-program decisions. Man can make completely new decisions (creative decisions) thanks to the fact that he captures the “common” nature of things, which is an abstract that only occurs in an intellect and never in the material world where computers belong.

References

Aquinas. (2012). *Summa theologiae* (The Aquinas Institute, Ed.). Emmaus Academic.

Aristotle, (1996). *Physics* (R. Waterfield, trans. & notes). Oxford University Press.

Aristotle, (2016). *De anima* (C. Shields, trans., introd., & commentary). Clarendon Press.

Bermúdez, J. L. (2020). *Cognitive Science. An Introduction to the Science of the Mind*. 3rd ed. Cambridge University Press.

Block, N. (1990). The computer model of the mind. In D. Osherson & E. Smith (Eds.), *An invitation to cognitive science* (Vol. 3). MIT Press.

Block, N. (1995). The mind as the software of the brain. *An invitation to cognitive science* (Vol. 3). MIT Press.

Brentano, F. (1995). *Psychology from an empirical standpoint* [Trans. from German 2nd ed., 1924]. New York: Routledge.

Bringsjord, S. & Govindarajulu, N. S. (2020). Artificial intelligence. *The Stanford encyclopedia of philosophy* (E. N. Zalta, Ed.). https://plato.stanford.edu/archives/sum2020/entries/artificial-intelligence/

Church, A. (1941). The calculi of lambda conversion. *Annals of Mathematics Studies* (Vol. 6). Princeton University Press.

Davidson, D. (1990). “Turing’s test.” In K. A. Mohyeldin Said, W. H. Newton-Smith, R. Viale, & K. V. Wilkes, Eds.), *Modelling the mind*. New York: Clarendon Press.

Dennet, D. (1991). *Consciousness explained*. New York: Little, Brown and Co.

Einstein, A. (1921, January 27). *Geometry and experience* [Address]. Prussian Academy of Sciences, Berlin. Methuen & Co. [1922]. https://mathshistory.st-andrews.ac.uk/Extras/Einstein_geometry/

Elman, J. (1990). Finding structure in time. *Cognitive Science*, *14*, 179–211.

Fodor, J. (1975). *The language of thought*. New York: Crowell Press.

Fodor, J., & Pylyshyn, Z. (1988). Connectionism and cognitive architecture: A critical analysis. *Cognition*, *28*(1–2), 3–71.

Husserl, E. (2008). *Logical investigations*: Vol. I. *Prolegomena to pure logic*. New York: Routledge.

Köhler, W. (1976). *The mentality of apes*. New York: W. W. Norton. (Original work published 1917).

Millán-Puelles, A. (1967), *La Estructura de la Subjetividad*, Madrid: Rialp.

Planck, M. (1932), *Where is science going?* (J. Murphy, trans.). New York: W.W. Norton & Company.

Searle, J. (1980). Minds, brains, and programs. *Behavioral and Brain Sciences*, *3*, 417–457.

Searle, J. (1983), *Intentionality*, New York: Cambridge University Press.

Sen, D. (2014). The uncertainty relations in quantum mechanics. *Current Science, 107,* 203–218 https://www.currentscience.ac.in/Volumes/107/02/0203.pdf

Simon, H. (1966). On reasoning about action [Tech. Rep. Complex Information Processing Paper #87]. Carnegie Institute of Technology.

Smith, P. (2007). *Church's thesis after 70 years*. http://www.logicmatters.net/resources/pdfs/CTT.pdf.

Stephan, K. D. & Klima, G., (2020). "Artificial intelligence and its natural limits," *AI and Society* 36:1-10.

Traiger, S. (2000). Making the right identification in the Turing test. *Minds and Machines*, *10*, 561.

Turing, A. (1937). "On Computable Numbers, with an Application to the *Entscheidungsproblem*," *Proceedings of the London Mathematical Society*, volume 2-42, issue 1. https://londmathsoc.onlinelibrary.wiley.com/doi/abs/10.1112/plms/s2-42.1.230.

Turing, A. (1950). Computing machinery and intelligence. *Mind*, *49*, 433–460. https://www.csee.umbc.edu/courses/471/papers/turing.pdf.

Weed, L. E. (2003). *The Structure of Thinking. A process-Oriented Account of Mind*. Imprint Academic.

Wiener, N. (1985). *Cyberetics or the Control and Communication in the Animal and the Machine*. Second Edition. Cambridge: The M.I.T. Press.

Windholz, G. (1987). Pavlov as a psychologist. A reprisal. *Pavlov J Biol Sci*., *22*,103–112.