Memory and consciousness, an essay from artificial intelligence to cognitive neuroscience and philosophy

Yann Cherdo

I. Preamble

This essay constitutes a part of my doctoral thesis in the field of bio-inspired Artificial Intelligence (AI)[1], where I embarked, notably driven by curiosity about human nature.

II. Definitions

Here are the definitions used in this essay concerning certain key and recurring terms:

The self constitutes everything that allows an individual to define himself and distinguishing himself independently from the rest of the world and others.

Free-will constitutes an individual's ability to make a choice and control his destiny independently from the rest of the world and others.

Human nature refers to everything that can define and characterize the human being, from its physical structure to social, cognitive, intellectual behaviors, and mental objects, no matter how abstract they may be.

Reality, or truth, in this essay, refers to the objects of science or more broadly of thoughts, whether they are considered objective, universal, or subjective.

Reason is a faculty of the human mind to organize its relationships with reality[2]. It operates notably through modes of thinking such as determinism, causality, logic, or axiomatic system.

I will propose a definition of consciousness in section VII based on a set of considerations developed in section VI.

III. Motivation and Subject

AI, by its name, naturally invokes a personal questioning for each of us. In Western culture, intelligence is associated with human nature and seems to be part of the definition of the self, from the beginning of history through religion and the Enlightenment period[3]. But then, what does this field of AI has to say about human nature?

For example, human nature is revealed through the history of the evolution of techniques and machines. Today, AI is part of this process and also reflects human activities that a machine can reproduce, such as writing or driving thanks to autonomous vehicles. This process also extends beyond activities that humans can do by themselves and allows for augmentation, notably with aviation or big data processing.

As soon as humans invent a system that can perform a task instead of them, this task may appear to be less human or less constitutive of human nature. The more machines reproduce human activities, the less these activities seem to directly enlightening us about the nature of humans. This idea also applies to tasks performed by AI. Similarly, my readings in cognitive neuroscience concerning cortical mechanisms underlying learning, memory, or agency[4] led me to question the place of these cognitive properties in human nature.

The intimate relationship between AI and neuroscience has been a guiding thread throughout my doctoral thesis. For example, I have exploited and modeled spiking neurons and the sparse characteristics of activity and connectivity of the human cortex for the AI models I have developed. This was notably done in order to reduce their energy consumption[5].

Rather than discussing the concept of intelligence, I would like to focus on consciousness and the self as they are also prevalent and commonly used topics in discussions about AI. It could be proposed that the aforementioned personal questioning leads to a definition of consciousness independent of memory. This is the thesis that I propose to develop in this essay.

As will be discussed later in this essay, tackling the definition of consciousness head-on can be difficult. Consequently, a method used in physics seemed appropriate to me. It consists of defining a particle not by its structure or constituents but by its symmetries or invariant transformations. Indeed, a complementary understanding of consciousness could emerge from defining its conceptual boundaries. In this essay, I propose that cognitive neuroscience can help outline such a definition and provide an example with the concept of memory in section VIII.

IV. Understanding through Reason, Axiomatic System, and Causality

A classical, modern, and Western education in physics and mathematics can forge a powerful capacity for reasoning and induction allowing the resolution of complex problems. The

[1] https://www.college-de-france.fr/fr/agenda/seminaire/modeles-multi-echelles-et-reseaux-de-neurones-convolutifs/intelligence-artificielle-et-intelligence-naturelle-vers-ia-bio-inspiree

[2] https://www.cnrtl.fr/definition/raison

[3] Histoire de la philosophie occidentale, Russel, les belles lettres, 2011

[4] Agency theory, Shapiro, 2005

[5] Time series prediction and anomaly detection with recurrent spiking neural networks, Y Cherdo, B Miramond, A Pegatoquet, 2023 International Joint Conference on Neural Networks (IJCNN)

method of resolution is often based on a logical or causal understanding of a system, sometimes built by correlations. Axiomatic reasoning is key to this method: based on a set of assumptions, we deduce a theory or knowledge of a system and put it to the test through empirical tests and critical analysis by the scientific community.

Well-established models such as Kepler's or Newton's laws are valid only within their axiomatic system, much like mathematics. Therefore, a universal truth seems difficult to attain, and it is doubtful to claim to ever perfectly understand what consciousness is. It might therefore be difficult to implement it in an AI. In this regard, the philosopher Plato serves as a reference to illustrate the fragility of humans when they claim to possess universal knowledge through reasoning[6][7][8].

It follows that reasoning cannot lead to certainty, and the need for it responds to other motivations. One of the ends of reasoning might not be certainty but doubt. Doubt and curiosity are vectors of progress and seem to contribute to the joy and balance of an intellectual activity such as science.

Nevertheless, a finer understanding of consciousness can be obtained notably through the experience of it, humanities, and neuroscience. Thus, limited definitions of consciousness can be proposed and implemented in an AI.

V. Determinism and Indeterminism, Two Hypotheses to Approach the Self and Consciousness

The concepts of consciousness and the self can be perceived as singular. If they could be explained by axiomatic reasoning, using causal relationships just as neuroscience explains certain cognitive behaviors, what would become of these concepts and in particular the self? If a deterministic model could predict the actions, choices, or thoughts of an individual, what would remain of their self and free-will? And what would we be conscious of? As I read books and research papers, especially in cognitive neuroscience for the purpose of exploiting certain cortical mechanisms in AI, it seemed less trivial to me that memory and learning are constitutive components of the self. These concepts, although often presenting a statistical aspect, suddenly seemed to belong more to a deterministic mechanical system than to a being endowed with consciousness and a sense of the self.

For example, if the scientific community were to argue that such a deterministic model could exist, this would imply a loss of agency for individuals, with probably unique consequences for social organization and the concept of the self as it has been forged since antiquity in the western world. Justice, democracy, and free-will would then probably need to be redefined. It could also be considered that in the deterministic hypothesis, free-will, the self, and consciousness could be conceptual errors caused by a limited understanding of the world.

Although neuroscience explains certain cognitive behaviors deterministically, it is possible to choose the hypothesis where free-will, the self, and consciousness do exist. Many arguments can be found to support it, particularly in the functioning of current societies, arts, philosophy, and sciences. The very permanent sensation of the self and of an objective experience, which the reader can experience now, is part of these arguments. In what follows, I propose to use this hypothesis.

VI. Definitions of Consciousness

Having proposed to seek a better understanding of consciousness, it seems necessary to recall existing definitions. They are complex and vary from one domain or author to another.

A. Psychology

In psychology[9], for example, consciousness is associated with the duality that appears during the early years of a human's existence. By perceiving oneself, a person somehow splits himself, leading to the concept of a problem. Indeed, the idea of a problem seems nonexistent if an individual could not project himself into several possibilities. For example, the notion of a problem can arise when an individual regrets a past choice. This is possible because one can project himself into another choice that he believes could have led to a better outcome. He compares different versions of himself, thus seeming to split. Note that this mechanism of self-conceptualization, remembering past events, and projecting them into hypothetical events is made possible by memory. The notion of time is also a result of this, since without memory an individual could not remember who he is and therefore project into the past or future. He would then live only in the present moment. Note that this might not prevent the individual from experiencing consciousness or even the self, although it would no longer be possible to construct the latter around the individual's memory.

The idea that this duality is at the origin of the concept of a problem can be found in the word "diable" (devil) in French whose prefix "di" means two or double in its ancient Greek origin[10]. In Homer[11] and Plato[12], Ulysses is feared because he is capable of deceit and manipulation, he is double. In contrast, Achilles is adored, notably because he is consistent and true to himself, he is predictable and incapable of using cunning like Ulysses.

B. Etymology

The etymology of the word "consciousness"[13][14] comes from the Latin word "conscientia," a contraction of the terms

[6]Gorgias, Platon oeuvres complètes tome I, traduction par Léon Robin, Editions Gallimard, bibliothèque de la pléiade, 1950

[7]Phèdre, Platon oeuvres complètes tome II, traduction par Léon Robin, Editions Gallimard, bibliothèque de la pléiade, 1950

[8]Théétète, Platon oeuvres complètes tome II, traduction par Léon Robin, Editions Gallimard, bibliothèque de la pléiade, 1950

[9]The spiritual problem of modern man, Jung, 1928

[10]https://www.cnrtl.fr/definition/di-

[11]Iliade - Odyssée, Homère, traduction par Victor Bérard et Robert Flacelière, Editions Gallimard, bibliothèsue de la pléiade, 1955

[12]Le petit Hippias, Platon oeuvres complètes tome I, traduction par Léon Robin, Editions Gallimard, bibliothèque de la pléiade, 1950

[13]https://www.littre.org/definition/conscience

[14]https://gaffiot.fr/conscientia

”cum”[15][16] and ”scientia”[17] which can be translated as ”shared knowledge” or literally as ”with knowledge.” The difference from the word ”scientia” alone could stem from the interaction between a human’s self or consciousness and knowledge. It might seem too simplistic for a robot or AI to actually have knowledge by using a memory unit filled with information because there seems to be no fundamental difference between this memory unit, a robotic arm, the processor, or even the silicon it consists of: all of these are organized matter. As a result, the etymology of the word ”consciousness” seems to invoke a more singular approach to knowledge linked to its very creation and manifestation in the human mind.

It is interesting to note a certain perversion of language here, as in the use of words like memory, knowledge, or consciousness interchangeably between a human and a robot. By using the words ”memory” or ”knowledge” to describe a robot, for example in the sentence ”This robot is equipped with a memory in which knowledge has been downloaded”, this robot is implicitly attributed through these definitions with human qualities it does not possess. Indeed, the words memory and knowledge for a human go far beyond what they designate in a robot, which is only organized matter. Concepts, including complex ideas and those related to humans, reflections and products of societies and languages formed over thousands of years, are here projected onto a contemporary system such as the robot.

Finally, these considerations call for an attempt to understand the relationship between the self and consciousness on the one hand and the concept of organized matter on the other.

C. Philosophy

In Plato’s works, and unless I am mistaken, the word consciousness does not appear. Perhaps it was a term that did not exist in ancient Greek or a consequence of its translation. Another term is frequently used and constitutes a key concept in his works: the soul[18]. The description of the soul in Plato seems very similar to that of consciousness. The definition he gives of the soul is consistent with the previous definitions of consciousness due to the duality it proposes between the self or the soul and the physical world, including the body. The soul is singular because it is presented as immaterial and does not really follow the laws of physics, which seems to allow free-will to exist. This approach of the duality of body and soul, or sometimes called body and mind, is not only present in Plato’s works but is part of a tradition that can find some origins in various ancient authors such as Homer or Herodotus as well as in Christianity.

In this duality, a clear boundary is drawn between the soul or the self and the physical world where the body can be perceived as a blurry boundary between an individual and the rest of the world.

This unique approach to consciousness that defies reasoning has notably been the focus of works by the philosopher Michel Bitbol[19]. Consciousness is proposed there as an individual experience: the ”conscious experience.” Understanding consciousness appears as a paradox since the act of understanding itself is an activity of consciousness...

This circular difficulty of reasoning does not seem specific to the concept of consciousness since, for example, the void cannot be explicitly defined lest the object of that definition not be the void. Or the statement ”All humans are liars” leads to circularity of reasoning.

D. Neuroscience

In neuroscience[20][21], consciousness is associated with a global workspace that exchanges stimuli with neurons or assemblies of neurons called unconscious. This proposition is notably supported experimentally by the demonstration of neuronal activities located in this workspace, particularly in the prefrontal region of the cortex, when the studied individual performs a task consciously. This workspace would enable access to consciousness. A duality is observed between a certain set of neurons presenting global activity correlated with conscious acts and the rest of the neurons constituting the brain.

It should be noted that although significant progress has been made in neuroscience towards understanding consciousness, some uncertainty still remains regarding the link between the so-called conscious neuronal mechanism or access to consciousness and the phenomenon of conscious experience lived by each individual.

VII. A Proposed Definition of Consciousness

I do not have sufficient expertise to provide further references regarding the definition of consciousness. More information on this topic can be found in the fields of neuroscience, psychology, and phenomenology of the mind.

Based on the definitions previously presented, I propose to synthesize them as follows: by consciousness, I mean the system that allows a human to experience the self, free will, and his duality.

In this essay, the concepts of consciousness and the self are sometimes presented side by side. This reflects the intimate relationship that seems to exist between these two concepts, as the conscious experience may primarily be an experience of the self. It is indeed a subject that becomes conscious of an object.

[15]https://gaffiot.fr/con

[16]https://gaffiot.fr/cum

[17]https://gaffiot.fr/scientia

[18]Phédon, Platon oeuvres complètes tome I, traduction par Léon Robin, Editions Gallimard, bibliothèque de la pléiade, 1950

[19]Aux sources de la conscience, https://lejournal.cnrs.fr/articles/aux-sources-de-la-conscience, 2014

[20]What is Consciousness?, Wallace, 2004

[21]Global workspace theory of consciousness: toward a cognitive neuroscience of human experience, Baars, 2005

VIII. Results on Consciousness and Memory

In cognitive neuroscience[22][23][24], there are several models describing memory mechanisms. Neuronal plasticity partly explains the particular ability of mammals, especially humans, to learn throughout an individual's life and thus transmit culture. Reptilian species evolve and adapt rather through genetic mutations. In particular, the Spike-Time-Dependent Plasticity (STDP), derived from Hebb's law, describes this synaptic plasticity at the scale of the synapse that allows a neuron to transmit an electrical potential to another. The dynamics of this plasticity as well as the functioning of neurons explain, for example, why humans are capable of learning very long sequences such as texts or pieces of music, sometimes even unconsciously. The sparse activity of neurons explains the great robustness of memory as well as its great capacity for retaining information. The hierarchy and organization of the brain explain the abilities of abstraction and correlation between several senses like sight and touch, and thus of memorizing and representing the world in the manner of humans. These mechanisms make an individual capable of considering their constraints given their lived experience and what they have retained from it, and thus weighing their decisions. Other mechanisms such as competition (neuro-Darwinism), attention, interactions with the older parts of the brain could also be developed.

Finally, it is interesting to note that the human cortex, the privileged place of neuronal plasticity that concentrates about 80% of the brain neurons, is a memory but also the matrix that predominantly operates or constrains senses, motor control, language, imagination, etc.

These neurons and learning models are partly used today in AI models.

In sections V and VI-A, the relationships between the self and memory or determinism were questioned. It seems intuitive that memory participates in defining the self through memories but also cultural acquisitions such as language and social relationships. However, the memory mechanisms previously described are determined and do not account for consciousness. Moreover, the definition of consciousness is based on concepts partly contrary to determinism such as free-will or the self. Paradoxically, defining the self through memory and its mechanisms confines the self in a determinism that negates it, as already mentioned in section V. This reveals two possible forms of the self, one based on memory and an individual's environment, and the self that seems associated with conscious experience.

Consequently, it appears that memory is not a constitutive element of the self or consciousness.

It may be troubling to think that consciousness does not depend on memory. Indeed, how would we define ourselves or how would we act as individuals without memory? What would sensations, will, and intellectual activities be without memory? These questions all seem to provide some clues as to the relationship of these concepts with consciousness.

From these arguments, it is possible to question these relationships as well as certain conceptions, or sometimes certainties, regarding consciousness.

It should be noted that the concept of memory may not be reducible to the only models proposed in neuroscience and mentioned above, and that other relationships may exist between memory and consciousness. However, these different models account astonishingly well for the functioning of memory in humans.

IX. Difficulties of Indeterminism

Having repeatedly pushed determinism away from the definition of consciousness, what about indeterminism? Although it is used in language and in science, its existence is not certain.

Quantum physics has been used notably to develop random number generators. There are indeed generators of pseudo-random numbers and not purely random ones. In fact, in the book[25], the physicist Nicolas Gisin explains that the material realization of the Bell test experiment meant to prove the random nature of the generator is itself difficult to realise and could be imperfect. In such a case, a positive result would lose its meaning. The author argues that, although they obtained positive results, his team was unable to demonstrate that the experiment was perfect. He concludes that it is not certain that the numbers produced by a quantum generator are indeed indeterminate.

Moreover, Heisenberg's uncertainty principle states that it is impossible to simultaneously know the wave function and the position of a particle. But the fact that we cannot know them simultaneously does not mean that indeterminism exists.

The existence of choice as a consequence of free-will is one of these difficulties. A proposed experimental proof could be to confront the same individual several times with the same choice under perfectly identical conditions between each trial. This experiment seems unrealistic as it would require the implementation of simultaneity or reversibility of time.

X. Difficulties of Determinism and Reason

It would not be surprising if determinism fails to fully explain consciousness. In the history of science, convictions and certainties have dominated and then been replaced by others, contributing to the progress of knowledge. Moreover, science seems to discover more mysteries than certainties as it advances.

Determinism may only partially allow the conceptualization of knowledge. It could also be a consequence and a reflection of the physical neuronal system. In this case, the concept of determinism and its consequences could be reduced to a physical phenomenon, that of the human brain, by placing consciousness as a spectator or perhaps a user.

[22]Neuronal dynamics: From single neurons to networks and models of cognition, Gerstner, 2014

[23]On Intelligence, Hawkins, 2004

[24]The cerebral code: Thinking a thought in the mosaics of the mind, Calvin, 1998

[25]L'Impensable Hasard: Non-localité, téléportation et autres merveilles quantiques, Gisin, 2012

Erwin Schrödinger questioned scientific modes of reasoning, partly supposing them to be the cause of the singular conceptual difficulties encountered in quantum physics[26]. He proposes that these modes of reasoning are not universal but are the product of their history since antiquity in western Europe. Thus, the way of reasoning itself is a hypothesis that conditions and is part of the produced knowledge.

In physics, for example, the concept that matter possesses mass has become ingrained in the collective unconscious and western culture thanks to the scientific consensus it has enjoyed for centuries. However, another concept has emerged suggesting that matter does not possess mass[27]. For instance, in this new concept, particles without mass acquire mass through their interaction with their associated field.

Ettore Majorana questioned the deterministic nature of the physical world[28]. The concept of randomness can indeed be perceived as a mean for approximating a system insufficiently understood to be determined. This argument assumes that the system is indeed deterministic. For example, Brownian motion can be perceived either as an undetermined and indeterminable movement or as a determined but difficult to determine movement. He suggests that randomness could be a constituent of the physical world rather than just a conceptual approximation resulting from a lack of knowledge. In doing so, he offers a way to question determinism and how reason is used in the sciences.

Understanding reason itself is not trivial. Its functioning is complex and can lead to paradoxes, as shown by the work of the researcher Hugo Mercier: "Reason does not serve so much to guide our choices as to infer reasons to justify them."[29].

That said, the utility of determinism and the progress it has allowed are hardly questionable.

XI. Realism and Constructivism

Having confronted the idea that memory could be constitutive of the self or consciousness with the way neuroscience explains the functioning of memory, it seemed important to introduce a parallel here with the book of Jean Pierre Changeux and Alain Connes concerning realism and constructivism[30].

In this book, the mathematician Alain Connes describes his intuition that mathematics is a language allowing the discovery of a reality independent of humans, their culture, or their thought. This intuition arose from the significant robustness of the content of mathematics and the way this content evolves through cultures, history, and international scientific consensus. In this respect, this idea can be associated with realism according to Plato, who proposes the existence of a reality or truth that transcends humans and is independent of them.

On the other hand, the neurobiologist Jean Pierre Changeux suggests that this intuition could be false and that this convincing robustness of the content of mathematics, subject of an almost unique consensus in the history of humanity, would be the result and image of the structure of the human brain and its cognitive functions. This argument is nourished in particular by neuroscience, which today offers biological and physical explanations of human cognitive functions such as logic, imagination, or inference, which seem necessary for doing mathematics. The concepts of order, causality, and cognitive stability, or homeostasis, used to describe the human brain, seem to be the result of its biological structure. This argument can be associated with constructivism, which proposes that reality is a human construction.

The confrontation of these two modes of thoughts notably allows for a better understanding and definition of the boundaries of Alain Connes' intuition. Following the dialectical method, arguments after examples, the dialogue led by these two authors offers a clearer understanding of what mathematics are or are not, and by extension, of this intuition.

It should be noted that constructivism and realism do not seem incompatible. While assuming that there is a reality independent of humans, it is allowed to assume that knowledge is a product of human thought and reasoning whose object is this reality. Knowledge would then be dependent on humans while the reality it tends to approach would be independent of humans. Plato also argued that knowledge, or truth, is found within itself and is only accessible through itself, thanks to courage, wisdom, and dialectics, amongst other things. In this vision, knowledge is indeed the product of the human being who aims for truth, which exists according to Plato, but is unattainable[31][32][33].

In this essay, I attempt, with my modest means, to structure an argumentation similar to that of the book by Jean Pierre Changeux and Alain Connes, pushing the boundaries of complex and intuitive concepts like consciousness or the self by using neuroscience and confronting ideas rooted in culture such as the soul or free-will with materialism.

XII. Consciousness and Objectivity in Science

The work I have been able to produce throughout my scientific studies, my thesis, and within the engineering positions I have held seems notably enabled by my desire for objectivity and the mechanism of my consciousness.

In science, the act of solving a problem often implicitly relies on the independence between the subject, the scientist, and the object of the problem. Thanks to this independence, an objective solution or an understanding of the reality at stake can emerge. The solution can be considered objective,

[26]'Nature and the Greeks' and 'Science and Humanism', Schrödinger, 1996

[27]The Particle at the End of the Universe: How the Hunt for the Higgs Boson Leads us to the Edge of a New World, Carroll, 2013

[28]The value of statistical laws in physics and social sciences, Majorana and Mantegna, Springer, 2006

[29]Notre raison est-elle rationnelle?, https://lejournal.cnrs.fr/articles/notre-raison-est-elle-rationnelle, 2017

[30]Matière à pensée, Changeux and Connes, Odile Jacob, 1989

[31]Ménon, Platon oeuvres complètes tome I, traduction par Léon Robin, Editions Gallimard, bibliothèque de la pléiade, 1950

[32]Phèdre, Platon oeuvres complètes tome II, traduction par Léon Robin, Editions Gallimard, bibliothèque de la pléiade, 1950

[33]République, Platon oeuvres complètes tome I, traduction par Léon Robin, Editions Gallimard, bibliothèque de la pléiade, 1950

and the acquired knowledge is sometimes perceived as independent of humans, which seems hypothetical. Indeed, it is difficult to argue that the logical path of the scientist's thought is independent of his culture, neuronal activity, or consciousness, despite the significant and singular effort of science to make its content universal. This hypothesis of separation between subject and object has, amongst other things, been studied by the philosopher Edgar Morin [34]. Just as Erwin Schrödinger questioned the cultural substrate of modern science[35], this hypothesis could handicap science if it were not questioned and perhaps even integrated as a clear assumption of the scientific method.

It should be noted that the separation between subject and object can be found in the etymologies of the words conscience "cum scientia" and science "scientia." The word science, or "scientia," does not include the "cum" which seems to refer to the subject.

Science allows for the propagation of a rigorous and fascinating curiosity about knowledge and reality. However, it seems to be part of the acts enabled by consciousness. It is indeed difficult to imagine that this activity could spontaneously and naturally emanate from a non-conscious system such as a robot or an AI. If consciousness is indeed a necessary and constant mean for the activity of sciences, it may seem strange that science has revealed so much knowledge in various fields, sometimes even on distant objects like in astrophysics, and so little, it seems, concerning consciousness itself. This apparent imbalance may be due to the societal needs and challenges that science addresses. Since consciousness is part of the intellectual and sensory reality, it is an object of science. It constitutes a relatively thick mystery through which the scientist paradoxically seems to discover knowledge sometimes enveloped in a feeling of certainty. One might wonder where a certainty would come from when one of the means by which it is obtained, consciousness, still seems to elude understanding.

XIII. Conclusion

In this essay, it was first assumed that consciousness, the self, and free-will do indeed exist. This allowed addressing the conflict between these concepts and determinism, which seems to contradict their existence. It was thus proposed that determinism, despite its important role in science, might not suffice to understand what consciousness truly is. One way to circumvent this problem is to refine the boundaries of the definition of consciousness with the current knowledge. That is, to refine the understanding of consciousness by defining what it is not in its proximity. To do this, it was proposed to exclude from the definition of consciousness any phenomenon that appears deterministic. The example of memory was chosen since it is part of the cognitive functions that seem to have a connection with consciousness. Memory, as explained through deterministic models in neuroscience, was proposed as independent of consciousness.

In conclusion, this reasoning ends with the question of what consciousness could be without memory according to neuroscience, as well as a questioning of certain ideas associated with consciousness such as free-will and the self.

XIV. Discussion

This essay has led to uncertainties and questions. The concept of consciousness, studied through neuroscience, science, and philosophy, raises several difficulties, including the uncertainty of its definition and the singular and still incomprehensible mysteries it shelters.

A. AI

As a result, it is imprudent to claim that AI could experience consciousness or that the AI research community would approach it solely through increasing performances, according to the usual metrics in AI, on cognitive tasks, no matter how complicated they may be. Other ways of developing AI, emancipated from the sole objective of performance, might lead it toward conscious experience provided an implementable and determined definition of it exists. However, as we have seen in this essay, the different definitions of consciousness do not describe it deterministically and thus programmably. Indeed, the current rise of AI is largely facilitated by the precise definition and determined implementation of the desired performances.

However, it is important to note that AI can produce powerful illusions, or even generate addictions, thus having harmful effects on human health. These phenomena can be observed through social networks or applications such as the one aiming to "revive" a deceased person through an AI that has learned their personal language[36].

In this essay, it has been stated several times that a robot or an AI is only made of organized matter. This remains a supposition because it could also be argued that experiences of consciousness or the self could exclusively be consequences of the laws of physics.

B. Art and Science

It seems that consciousness manifests itself, amongst other things, through complex and singular activities. Art would be one of them. Poetry, painting, music could be means of expressing a reality, or a truth, without the intermediary of rational understanding[37]. This could explain the significant impact of art on humans because it would reveal a reality that can be recognized within oneself. Rational understanding alone may be perceived as incomplete in expressing reality or truth. In Plato's view, beauty is not an aesthetic concept but can be understood as the manifestation of truth in the human mind, or in its soul [38]. It is interesting

[34] Introduction à la pensée complexe, Edgar Morin, 2015

[35] 'Nature and the Greeks' and 'Science and Humanism', Schrödinger, 1996

[36] Reviving The Dead With AI: Is It Really Worth It?, https://www.forbes.com/sites/federicoguerrini/2023/04/15/reviving-the-dead-with-ai-is-it-really-worth-it/

[37] Ion, Platon oeuvres complètes tome I, traduction par Léon Robin, Editions Gallimard, bibliothèque de la pléiade, 1950

[38] République, Platon oeuvres complètes tome I, traduction par Léon Robin, Editions Gallimard, bibliothèque de la pléiade, 1950

to note that, from this perspective, mathematics or physics, for example, can be understood as arts, revealing a part of inner or outer truth to the human mind without truly understanding where it comes from or what it means.

In turn, AI can produce aesthetic poems or paintings with a certain degree of perfection, exploiting the learning of complex associations between human expressions from a very large database. Following what has been said previously, and despite a manifest performance, it would be risky to claim that such AI produces art. Indeed, it is necessary to distinguish which entity creates art from one that reproduces a certain understanding of the acts and productions that result from it.

For example, an AI can learn to reproduce certain painting styles and even combine several or transform a photo into these styles. If the logic and algorithm are given to it, it can also generate other styles from these styles. These acts are ultimately reproductions because they depend entirely on the human who implemented in an AI a logic allowing it to learn complex relationships in a database illustrating works of art. In both cases, the learning logic or more broadly the AI itself, as well as the works of art that constitute this database, are products of humans. Therefore, the actions of an AI, i.e., a system primarily constituted of organized matter, whether associated with art or any other task, are not independent of humans and are their production.

When a painter paints a picture, it is appropriate to say that the picture is the production of the painter and not of the brush, which is one of the means. When a photographer takes a photo, it is appropriate to say that this photo is the production of the photographer and not of the camera, which is one of the means. When a Formula 1 driver wins a race, it is appropriate to say that the driver won the race and not his vehicle, which plays an important role in that victory. Painting, photography, or racing are not the products of human products like AI; they are the creation of humans. One could argue that the difference with an AI is that once implemented and its learning completed, it operates independently of humans. However, this AI still needs humans during its inference phase since it needs energy, IT maintenance, etc.

On the other hand, and for example, the James Webb Space Telescope was created by humans and sent into space to take pictures, amongst other things. Once the telescope is launched, it might be said that it travels through the universe and takes pictures relatively independently of humans, although it follows its program, just like an AI. It could then be said that the James Webb Space Telescope took pictures of the universe. But it seems more accurate to say that humans took pictures of the universe through the James Webb Space Telescope.

Apart from the complexity and difficulty of explaining AI inferences, there is no fundamental difference between the algorithm controlling this probe and an AI algorithm. The actions of an AI remain a determined product of human thought, even though they often use the generation of pseudo-random numbers. So AI could in some cases be a mean through which humans could produce art, not an entity that produces art by itself.

Furthermore, the algorithmic logic that constitutes AI is, in the example of painting, a model of what humans have understood of the artistic act of painting, including how to be inspired by existing works through learning. On the one hand, it has been argued above that models derived from human reasoning are not universal and only incompletely express the object they aim to describe despite their qualities. Therefore, this AI would represent an incomplete idea of the artistic act and would be an approximation of it. On the other hand, it has also been said above that art is an act partly irrational and that this part that escapes reason is a condition of its existence. In other words, if art were determined and entirely described within a model, through AI for example, perhaps it would no longer be art. Finally, art has been described as another mean than reason to access reality or an inner or outer truth to humans. It is this bijective relationship between the work of art and the human that partly constitutes art and could explain the significant impact it has on humans. When a work of art is perceived by a human, it invokes truths that they recognize within themselves without fully understanding them. This act of recognition is personal and necessarily passes through the self and consciousness. One might wonder what the nature of the relationship between a work of art and a robot or AI would be.

Once again, we can observe one of the benefits of human production, AI, which, through the reflection of humanity it projects, allows for a questioning of human nature and notably the definitions of complex concepts such as art or knowledge, for example. This generates doubt, which is a driving force in abandoning certainties that could hinder scientific progress.

Since a finer and broader understanding of the concepts used in AI can emerge from the humanities or arts, philosophical considerations could contribute to the progress of AI development and the communication of the AI community's work. This broader approach in science is advocated by the researcher Edgar Morin through the "complex thinking"[39]. Ultimately, rigor demands questioning all certainties, and this ability can be nourished by an increased scientific curiosity.

Through this exercise of reflection and popularization, I hope to have offered the reader a moment of intellectual curiosity and joy. To attempt to conclude it harmoniously, allow me to quote Nietzsche:

"The magnificent and savage folly of poetry refutes you, followers of the useful. It is precisely the desire to free oneself from the useful that elevates man above himself." The Joyous Science, Nietzsche.

"La magnifique et sauvage déraison de la poésie vous réfute, sectateurs de l'utile. C'est justement la volonté de se délivrer de l'utile qui élève l'homme au-dessus de lui-même." Le Gai Savoir, Nietzsche.

XV. Acknowledgments

[39]Introduction à la pensée complexe, Morin, 2015