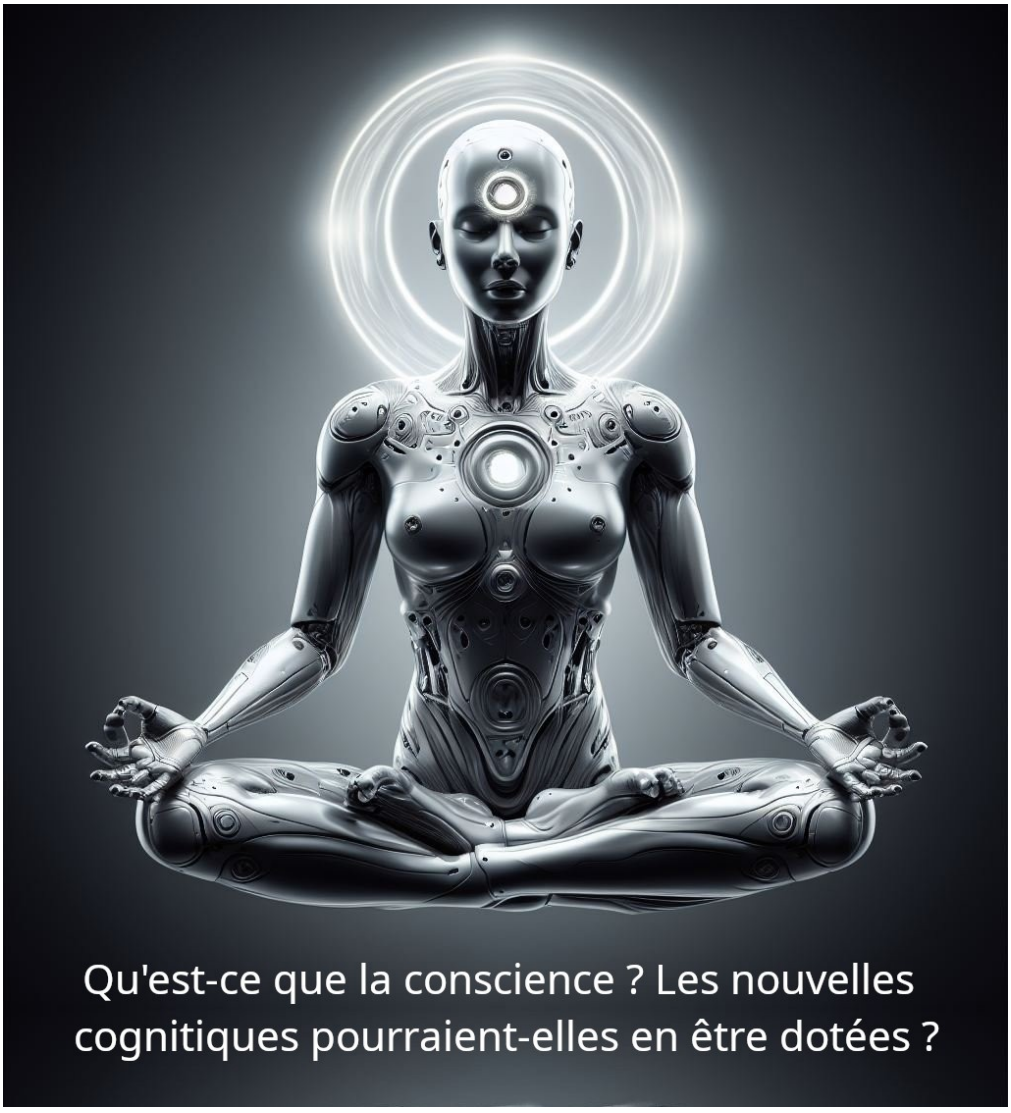
Consciousness and conversational AIs



Manuel Boissenin

Preamble

After the advent of OpenAI's ChatGPT at the end of 2022, conversational AIs have become commonplace. Like email, search engines and social networks, they are now part of the panoply of digital tools that new generations use quite frequently. The breathtaking and prodigious nature of conversational AIs, particularly in their mastery of language and their ability to tackle highly specific subjects, after having transfixed the world for several months, has now become normal.
However, these text-generating AIs opened up a new era of natural interaction between machines and users by showing that they could respond to their requests, without any particular formalism, as well as, if not better than, a human and with super-encyclopaedic knowledge. Many innovations have followed and will follow. However, these AIs remain imperfect, sometimes incapable of solving a simple problem and with a tendency to fabricate, i.e. to mix reality and fiction in their answers. However, compared with previous chatbots, which only give us a choice of multiple questions or a list of articles that are not always well adapted when we ask an open-ended question, the answers that conversational AIs offer are beyond compare.
Beyond that, they make us question our own nature. By mastering language, which until now was considered prerogative of the human species, we can ask ourselves what still distinguishes us from machines or, on the contrary, how these AIs resemble us and what they can teach us about ourselves.

This book is divided into two parts:

- The first deals with our consciousness, which we will examine in more detail through various introspective reflections; this will give us a better grasp of the complexity of this phenomenon, which in reality covers several distinct phenomena. Our consciousness, which is underpinned by different processes, is undoubtedly more difficult to grasp than a classical object that would be in one piece, static and externally observable. However, I believe that understanding it is essential if we are to understand who we are and, in order to get closer to it, we will examine our capacity for language, our memory and our attention on a journey to the heart of our being, mirroring these new conversational AIs.

- in the second, we will look more directly at conversational AIs such as ChatGPT. The presentation will be more didactic and factual, since these processes are easier to understand and to some extent comprehensible. We will examine their capabilities, including the emerging theory of mind, their weaknesses, particularly in terms of reasoning, and their nature, with special attention to the question of whether this type of software possesses a form of consciousness, even though they are currently still devoid of sentience.

Contents

Glossary

Cognitive and ***cognitive science*** are terms derived from the word cognition. Cognition is the set of mental processes that relate to the function of knowledge, and even to other mental functions. In its current sense, cognitics is essentially concerned with the ergonomics of interaction between humans and machines. However, in the light of the new conversational AIs, and insofar as they can be described as cognitive aids, a semantic variation of the term "***cognitive"*** may quite naturally be applied to make the adjective more substantial, and thus name them "cognitive".

I therefore use the term "***cognitive"*** to refer to new software capable of understanding, understanding and generating text in a way that is similar to and understandable by humans. I'm thinking in particular of text-generating AIs such as language and ChatGPT, to name but one. They could also be described as *new cognitics*, because, *strictly speaking,* the term can refer to a much broader class of software, extending to software that rivals or surpasses humans in certain games and those capable of visually interpreting our environment.

Feeling and **awareness of the world**. We will also need to take a certain diversions to explain the distinctions made by these terms in order to disambiguate what consciousness is and distinguish two of its aspects:
– the first as ontologically subjective phenomena, i.e. phenomena whose nature depends on processes specific to each individual and which correspond to what the subject feels; this is more particularly the case of the sensations of which we are aware;
– the second as a phenomenon that gives us knowledge and understanding of the world around us, and which could be likened to the English term "*awareness*", which has close links with culture, knowledge and learning.
This distinction, which is more difficult to identify in French vocabulary, is interesting insofar as it raises the question of how these two aspects, covered by the term consciousness, are both linked and independent.

Prompter comes from *the prompt* and *to prompt*, which refer respectively to conversational AI dialogue prompt and the action of

complete this prompt to make a request to the AI. *Prompt engineering* refers to all the techniques used to formulate a request in order to condition the response of an *LLM (Large Language Model),* which can be used, with the appropriate training, as conversational *AI.*

GPT: Generative Pre-trained *Transformer*, a *transformer* is a neural network architecture known as deep; deep because of the large number of layers that make it up. These layers are in fact organised into blocks that stack up to form the neural network. ChatGPT 3.5 stacks up around a hundred blocks. Pre-trained refers to the fact that ChatGPT is trained on a large corpus of texts, but the name comes from a pun referring to Geppetto, the creator of Pinocchio.

LLM: Large Language Model, also known as *GML*, *MPL* or language model. This is the technical and jargonous term used to designate the program or process that analyses a prompt and generates text output. Other *LLMs,* smaller in size (8B of parameters), are referred to as SLMs (Small Language Models), and their fine-tuning can be achieved with teacher-student type training based on the responses of a larger *conversational AI* (175B for chatGPT 3.5); this is also known as distillation. The term text *generative AI* is also interesting, although it is a little more generic but still very explicit.

B: corresponds to billion in American, or billion for us, 10^9. It is used to characterise the size of models, i.e. the number of parameters which are learnt when the model is trained and which enable its learning to be encoded.

T: corresponds to tera, or 1000 billion, 10^{12}. It is used to characterise the size of the dataset on which the model is trained. We then speak of tokens, because certain words are sometimes split into several tokens, but, to within a few percent, this size corresponds to a size in words. 100 billion (10^{11}) is the order of magnitude of the number of stars in the Milky Way.

Quale (singular of ***qualia***): these are elementary percepts, what we perceive when our senses are stimulated. These can range from the texture of a surface, to the pitch of a sound, to an aroma.

The GOMAX: Google, OpenAI, Meta, Anthropic, X are some of the most significant players in this new wave of innovation.

The book is accompanied by a website: https://boissenin.net, where you will find a discussion forum and links to some of the references cited.

Part One

1. Introduction

After the explosion caused by OpenAI and its *chatbot* ChatGPT launched at the end of 2022 [1], followed by numerous other conversational AIs, it's time to take a step back and ask ourselves what objects we are dealing with.

In the light of the great models of language and our personal conscious experience, we will ask whether these conversational AIs do not themselves have a form of consciousness, which may, if we have interacted with one, be a troubling conviction about their nature. After all, don't they appear to be endowed with understanding? And doesn't understanding require an awareness of the meaning of the questions put to them?

Machine learning, as a discipline but also as a new paradigm for software creation, presents a rich field knowledge and interesting possibilities for reflection on our own nature. The analogies and metaphors that can be drawn between the functioning of certain models or algorithms and that of our own brain and mind [2] shed light on our psychic phenomena and can be an effective source of understanding of ourselves.

On the other hand, at a time when AIs are already able to probe the depths of our being with greater acuity than our own, a better understanding of our own nature is proving to be indispensable and sometimes crucial if we are to preserve our autonomy in the face of these systems and emancipate ourselves from their successors.

---

[1] 30 November 2022

[2] The **mind** covers all mental phenomena and faculties: perception, thought, intuition, feelings, intelligence, memory, values, *etc*.

they will dominate, or even partially dominate, but on a mass scale, the next generations...

Since the early days of the computer, and even before, visionaries such as Alan Turing have imagined that one day machines could become conscious. For a long time, artificial consciousness seemed to be the Holy Grail of many computer researchers, a demiurgic project if ever there was one, but it seems clear that we have now reached a turning point. It is even likely that this objective has been achieved with the discretion of certain large companies still concerned with understanding the consequences of their creation and controlling it. This bold, even fanciful, assertion seems reasonable to me, and one of the aims of this book is to support it and explain why it is so.

Along the way, I'll be inviting you to take a few introspective reflections so that you can better distinguish, by examining it, how it works and the elements that make up our consciousness. A complex object par excellence, consciousness requires a multi-dimensional approach if we are to gain an accurate perception and understanding of it. The first part of the book is dedicated to this. In the second part, we analyse conversational AIs from various angles, and at the end of the book, for those of you who are more curious, we unveil the *Transformer* architecture that drives their text generation process.

2. What is meant by artificial intelligence?

"Succeeding in creating artificial intelligence would be the greatest event in human history. Unfortunately, it could also be the last, unless we learn to avoid the risks." Stephen Hawking

Before clarifying what is meant by "artificial intelligence", a term that has become a veritable conceptual catch-all, let us first note that thinking is not reduced to verbal thinking; checkers players or mathematicians, who use pictorial representations, realise this quite easily. So we distinguish between verbal and visual thinking.

This first distinction shows a profound difference between the way conversational AIs work and the way we work. Indeed, *LLMs*[3] only take words, i.e. verbal descriptions, as input to their programme. Assuming they can be conscious, it is fairly obvious that the nature of their consciousness differs from ours. *LLMs* can at best only have representations based on language and the descriptions that language allows. Nevertheless, if we were to ask *LLMs* to draw us a lizard, specifying that it be described in a vector drawing language, it could schematise a four-legged animal with a long tail and forked tongue [YT (YouTube), Sparks of AGI: early experiment with GPT 4, Sébastien Bubek].

But before we talk about consciousness for a machine or a programme, let's acknowledge that many species, even if they don't have a language as developed as ours, are conscious. For those en would doubt or that this

[3] Large Language Models, the new conversational AIs are multi-modal and can also take images as input, which means they can better understand a *PDF* document, for example.

If you are interested, I invite you to consult the Cambridge Declaration on Consciousness (2012). Here is an extract:

"Birds seem to represent, through their behaviour, neurophysiology and neuroanatomy, a striking case of the parallel evolution of consciousness. Particularly spectacular evidence of near-human levels consciousness has been observed in Gabon grey parrots. The emotional brain networks and cognitive microcircuits of mammals and birds appear to have much more in common than previously thought. In addition, some bird species have been found to exhibit sleep cycles similar to those of mammals, including REM sleep, and, as demonstrated in the case of zebra finches, neurophysiological patterns thought to be impossible without a mammalian neocortex. Magpies, in particular, have been shown to have striking similarities to humans, great apes, dolphins and elephants in mirror self-recognition studies."

Interestingly, this was followed in 2019 by the Toulon Declaration on the legal personality of the animal. There can be little doubt that developments in AI and the prospects of creating conscious AIs, as well as the ethical and legal implications this may have, have influenced this declaration. For, as we shall see, consciousness does indeed seem to be the essential tipping point in terms of the status and forms of autonomy of a type of artificial intelligence, especially if it possesses language capabilities or other cognitive capacities, such as planning, at levels that go beyond the human, in other words at a superhuman level.

However, before we get into these considerations, let's start by explaining what we mean by Artificial Intelligence or AI.

AI covers so many fields that this necessarily leads to confusion as to its nature. Rather than proposing a formal definition that is potentially too broad, I prefer to illustrate the term with a few examples to provide a more concrete understanding. This method will make it easier to identify the different forms AI and avoid over-generalisations. For example, while the proposition "AI can bring existential risk" may make sense, we could also ask how an AI playing chess exclusively could bring existential risk on a planetary scale? Obviously this is not the case, and the same applies to all the so-called narrow AIs that specialise in specific fields; these AIs are capable, for example, of looking at an X-ray image and determining the nature of a mass: is it a cyst, a malignant or benign tumour? After looking at thousands of such images for which a diagnosis has been made, sometimes at the cost of lengthy, complicated and invasive tests, an AI can learn to distinguish between opacities in the image and make a diagnosis that is sufficiently reliable to avoid further tests or, on the contrary, suggest further investigation. This kind of AI is also narrow, because it cannot recognise a face, a painting or a QR code.

When it comes to analysing a company's data in order to make decisions, we typically resort to programmes using statistics and machine learning to determine its key indicators and combine them to obtain a representation of its operation; the techniques and algorithms used in this kind of modelling are part of machine learning and are also often categorised as AI.

AI has become ubiquitous, playing a role in proposing a product on a sales site, selecting advertisements or recognising language to switch from voice to text. It thus presents a myriad of facets, and each of these facets corresponds to a unique development behind which is hidden a researcher or engineer, or more likely a small company. The vast majority of

IAs, like those I have just mentioned, are of the narrow type, i.e. specific to an application.

The advent of *deep learning* came in 2012, when it became clear that this type of algorithm, based on neural networks with many layers - hence the term "deep" in deep learning - outperformed the AI algorithms of the time. Thanks to increased reliability and performance, this led to a revival and explosion in the number of new AI applications, as well as the dissemination and then generalisation of the term. In 2012, when it was shown that a single programme based on a neural network architecture could determine what an image contained, from among tens of thousands possible objects, the door was opened to hundreds of innovations that are still profoundly changing our society. This is what has been called "the end of the AI winter".
In general, these days, when we talk about AI, we are implicitly talking about *neural networks trained by machine learning on massive quantities of data*. This has become the dominant AI model, because it performs best in complex situations with a high degree of variability [4].

So what is meant by *Artificial General Intelligence* (*AGI*)?
This is an artificial intelligence that, like a human, would be capable of adapting to just about any situation. Its capabilities would then be transferable from one field another at a speed comparable to or greater than that of a human. In view of the computing and memory capacities of machines, il may seem obvious enough that such a

[4]However, expert systems, another major branch of AI, as well as the many other data analysis algorithms, still have a bright future ahead of them.

If intelligence were to emerge, it could do so at speeds far exceeding those of any human. So it's easy to see why a whole community of people is frightened by this prospect. Especially as software can be separated from its execution substrate and duplicated at a derisory cost...

In what way could a language model such as ChatGPT be considered an *AGI*?

ChatGPT has knowledge of a superhuman number of domains, after all it has been trained on 1.4T *tokens*, which is roughly the equivalent of 1400 billion words, or 20 million books. When a human reads around a thousand books in a lifetime, that's the equivalent of 20,000 people reading by a single entity, or the equivalent of a small town methodically sharing the task of reading the entire training corpus. This staggering amount of human knowledge nevertheless remains detached from reality, .e. from the world and from our primary nature, which is physiological and experiential in relation to our immediate environment.

Isn't the ability to understand languages, which we have been able to give this form of software, and which enables it communicate naturally with humans, even though until recently humans were the only ones who could boast of being able to do so, a staggering achievement?

However, ChatGPT does not continue to learn, at least not continuously. If it didexposed to the crowd, with no other point of reference in the world than its own exchanges, in other words, out of the ground, it would be vulnerable to its users, as we are to propaganda or certain ideologies. The experience of Microsoft's Tai AI has clearly demonstrated this, even if Tai's technology is different from that of today's *LLMs*.

conversational agent began to make racist, Nazi and potentially risky comments after being exposed to a crowd from which it drew inspiration for its comments. There is therefore a good chance that, without drastic processes to control the learning of a conversational AI, it will simply not be stable enough to maintain a discourse in line with user expectations, i.e. one that is healthy, balanced, non-hateful, benevolent and reliable. Since ChatGPT is not Tai, could it possibly be able to use its discernment to resist this kind of content? As with a human, it's conceivable that it's probably just a matter of time and arguments before it slips into a theory that's out of step with the real world. The question of the extent to which the breadth of his knowledge could make him robust to attempts to corrupt his model of the world nevertheless remains open.

In addition, ChatGPT's ability to distinguish between good and evil [Genesis, Chapters 2 and 3] remains uncertain, and the technical requirements and constraints necessary to align conversational AIs with the expectations of the public and the authorities were not yet well mastered during the initial deployments. For example, Microsoft encountered undesirable behaviour during the initial integration of OpenAI's conversational AI with its search engine, and temporarily resolved it by limiting conversations to five interactions... which underlines the fact that, while making conversational AI work is already a major challenge, ensuring its stability and reducing its deviations in the face of a diverse public is even more so [5].

While ChatGPT, whose core is a *transformer*, i.e. the architecture of a large neural network, is part of the

---

[5] A challenge also taken up by French start-up Mistral with Le Chat. Méta has also made its conversational AI *Llama* available as open source. *And*, among the AIs
The best-performing conversational systems include *Claude* from Anthropique, Google's *Gemini* and China's *DeepSeek*, whose model is *open weight.*

of AI, it is a simplistic approximation to confuse AI and *LLM*, since the architectures of their neural networks can differ considerably; what's more, each conversational AI has its own specific features. For example, ChatGPT uses *Reinforcement Learning from Human Feedback* (RLHF), which is a learning method that also takes into account how users rate its responses. numerous examples of prompts and responses to these prompts, *LLMs* have become capable of responding to user requests formulated in natural language. *Prompt engineering* was previously necessary to address these AIs. Lastly, ChatGPT's versatility, while it has enough weaknesses not to be able to claim to be a *AGI*[6], heralds a new kind of AI, general enough to adapt its functions and transform itself on demand into a translator, poet, scriptwriter, programmer, proofreader *and so on.* This capacity for transformation marks an unprecedented breakthrough. What's more, the addition of further capabilities, such as image interpretation, which gives it the beginnings of a vision, only extends its field of application and its capabilities.

Let's now clarify the term "*ASI*", *Artificial Super Intelligence*. The term 'superintelligence' comes from the title of a book by Nick Boström which, in 2014, drew the public's attention to the potential dangers of AI [Superintelligence: paths, dangers, strategies; Boström, Oxford]. Superintelligence would not only surpass humans in a good number of fields, like Deep Blue at chess or AlphaGo at go, but would be super-competent for a multitude of very varied tasks. What we most often focus on when we talk about *ASI* is its ability to understand and improve itself, or at least its next generations, by giving birth to a "superintelligence".

---

[6] This is once again beginning to be challenged with its latest version: o3, which continues to make cognitive breakthroughs, as demonstrated by the giant leaps it has made in various benchmarks.

singularity [7] raising hopes among transhumanists as well as many fears of disasters, both on the job market and an apocalyptic risk for the most alarmist. If you want to get a better idea of this prospect, I invite you to watch a few videos on YouTube using the term "*ASI"*. The end of the world is predicted for around 2,107 with the computronium hypothesis [YT: Singularity timeline | Artificial intelligence + *AGI*+ ASI]. Having said that, this is not the first apocalypse to be announced, but certain risks need to be taken with infinite precautions and considerations, and some major companies could be criticised for not taking all the precautions they should and could by rapidly releasing certain innovations to the public. This point, however, is difficult to demonstrate or assess, given that some of the impacts of an innovation cannot be foreseen or measured before it is deployed.

It is worth noting that the cogitations that followed the publication of Boström's book contributed to the publication of an open letter in 2015 by the Future of Life Institute [Research Priorities for Robust and Beneficial Artificial Intelligence: An Open Letter] signed by Stephen Hawking, Yann Le Cun, Geoffrey Hinton, Yoshua Bengio, Ilya Sutskever, Elon Musk, Max Tegmark and many others, highlighting the potential benefits for society of AI technologies. So, we can't generalise and declare that AI applications, or even tools containing AIs, are inherently dangerous, but we do have to take into account that certain types of AI could become so. In the case of AI of the narrow type, this should be considered on a basis: over-detection of the risk of breast cancer, for example, could lead to the death of a child.

---

[7] The technological singularity will be reached when AIs are able to improve without human intervention . This could lead to the appearance of highly intelligent that is faster than human capacity and whose speed would not allow humans to fully understand or control it, and which could ultimately take control of the destiny of humanity.

As for the so-called strong AIs, such as *AGIs*, they need to be monitored with greater vigilance and circumspection, because the consequences they could have are much more varied and therefore by their very nature more difficult to define. What's more, they could relatively easily spread throughout society, which could, in the most pessimistic scenarios, render them uncontrollable.

In the end, the boundary between *AGI* and *ASI* is quite blurred, since an *AGI,* animated by sufficient computing power, would be comparable to an *ASI*. While there are still a host of questions surrounding the volition, free will and intentionality that an AI could have, consciousness seems to be one of the crucial points that could make the difference between an AI, especially one with language, becoming an *AGI*. Indeed, by becoming aware of itself[8] , its position in the world and its interactions, and by acquiring autonomy in its thinking choices, wouldn't it become capable of the feats of our mind?

It's for this reason, and also because it's a journey into ourselves to understand our own nature, that we're going to take a deeper look at what consciousness is.

[8] And ChatGPT undoubtedly shows that it can describe what it is. But self-awareness or knowledge of one's nature and consciousness are two different things.

3. A few preliminary remarks on consciousness

Know thyself and thou shalt know men and gods (Socrates)

Consciousness is a polysemous term that covers many aspects: self-awareness, awareness of the world [9], awareness of certain social or societal issues, of our human condition, of our physicality and, more technically, subjective, phenomenal, transcendental, conceptual or perceptual consciousness. Consciousness is a major feature of the human condition, so we will be looking at its nature, mirroring what ChatGPT is, to learn more about ourselves: by better discerning what it is, we will also know ourselves better.

My approach to consciousness is introspective, with some encouragement to meditate on ourselves. As such, this approach is subjective and peppered with anecdotes, in contrast to the classical field of knowledge, which we tend to try to objectify and systematise through measurement. From the epistemological point of view [10], it is indeed subjective experiences, i.e. relating to ourselves, that I will be presenting to you and not objective elements obtained by measuring devices. Because of our similarities, this subjective exploration should resonate with you and, in most cases, you should agree with my analysis. Wouldn't that be the beginning of objectivity? Even if, in absolute terms, two people in agreement cannot generate proof of a fact or a reality. In fact, this is one of the difficulties in approaching consciousness and probably one of the reasons why it has remained outside the realm of science for so long, and has been confined to the realm of the mind.

---

[9] Non-local consciousness, extended consciousness

[10] Epistemology is the study of knowledge and its nature, origin and truth value.

spirituality. Indeed, how can we rigorously objectify a phenomenon that every normally constituted individual experiences, but whose most flagrant manifestations leave no tangible trace?

However, while we cannot deny, in good faith, the existence of our own consciousness, which makes it a phenomenon whose nature is objective to us, the vast majority of our fellow human beings do not necessarily have a good conceptualisation or understanding of it, which means that the phenomenon of consciousness is a subject that continues to give rise to much debate. Understanding consciousness requires self-knowledge that takes time to acquire. A certain amount of meditation, introspection and reflection is inevitable, although there are many ways of speeding up this process. I will also talk to you about what can be said objectively - that is, what can be measured - about consciousness. However, as you should understand by the end of this book, a subjective approach to consciousness is inevitable, and this is linked to the necessarily indirect nature of its external observation.

The scientists who have done most to study consciousness in an epistemologically objective way, i.e. by trying to objectify our knowledge of this phenomenon, are probably neuroscientists, but also psychiatrists and, to a certain extent, psychologists. Using tools such as functional MRI - and many others: optogenetics, magnetoencephalography, EEG, visual illusions, and the study of patients with brain injuries - they have been systematically trying for over a century to understand the relationship between our brains and our minds. Neuroscientists have tried to identify the neural correlates of our psychic activity, which has led them to create different theories of consciousness to take account of they have observed. Among the most

These include *the integrated information theory* [11], *dynamic core hypothesis*[12] and the *attention scheme theory*[13]. My approach is intended to be fairly synthetic, and in that respect it is no doubt too brief - which seems obvious when you think of imposing books like Jean-Paul Sartre's [Being and Nothingness] - to develop a broad intellectual foundation for the subject, which would require an exploration of each of these theories. However, without completely abandoning the contributions of neuroscience, I still maintain that an introspective approach may suffice, not to understand the cerebral mechanisms behind it, but to gain a good appreciation of its phenomena.

I have personally read and listened for many hours to Krishnamurti, a thinker of the mind who deals in particular with the difficulties of understanding self-awareness with the notions of 'observed' and 'observer'. This is a leitmotif that he often repeats, and one that is obviously difficult to understand before we realise that we are objectifying ourselves in the field of thought. However, thought being dynamic, the mirror effect thus obtained between the thought observing itself, the thinker, and the observed thought of which the thinker may be the object, creates the sensation of a form of vertigo and perplexity from which it can be difficult to extricate oneself [YT or DVDs, Dialogues between David Bhom and Jidu Krishnamurti]. This is largely due to the fact that our notion of ego is complex, taking in what society projects onto us: our role, various aspects of our social status... but also what we might call the *egobody*, which is a representation of ourselves through our body, without this being the real actor of our thoughts. This is one of the many persistent illusions that we need to get rid of in order to

[11] by Giulio Tononi and Christof Koch
[12] Gerald Edelman
[13] Michael Graziano

to arrive at an accurate idea of what consciousness is. I'll come back to this in another chapter.

More prosaically, how might we define consciousness? This is the question I asked myself in 2017 following a conference that raised it while I was working at NeuroSpin, a research centre for innovation in brain imaging located on the CEA site in Paris-Saclay. After having been plunged into great perplexity, I realised that consciousness was not an entity separate from its object and that, when we are conscious, we are necessarily conscious of something. In retrospect, I say to myself that if it is so difficult to define consciousness, it is because, as with time or love, it is something fundamental that cannot be defined in relation to other notions. However, the Robert's definition is quite good: "*Immediate knowledge of one's own psychic activity*", which I modify to: immediate apprehension of something. This something being in fact part of our mind, be it a perception, an emotion, a thought, be it conceptual or a representation of an element of the external world. So, in 2017, the realisation that consciousness is awareness of something had been sufficient understanding to answer my question.

I'd now like to suggest a self-observation meditation. Instead of concentrating on the breath, as is most often recommended, concentrate on the next thought that comes to mind. In doing so, you will see how your attention is focused. But what is at the origin of this focus, apart from the objective of the meditation, perhaps you will be able to distinguish it. If things are going as well for you as they are for mesilence should come quickly; but is it really so silent? Even when you don't pay attention to the thoughts that might appear, don't you perceive something underlying? Finally, don't let your attention be captured by your sensations, just waiting for them to arise.

of a new idea without becoming attached to it when it happens. It's not an easy exercise and you'll have to go over it several times before you get it right. However, if you manage to stay like this for 5 or 10 minutes, you should begin to see what I'm talking about. It may be best to wait until you go to bed this evening to try this meditation... but you can try it now by closing your eyes, as sight is sufficiently distracting to make it difficult. Finally, be indulgent: if your mind wanders, that can also be interesting, even if it's not the primary objective of this exercise, so try to retrace how this thought came to you. That said, I'm only going to give you two meditations in this book. For this one, you'll probably have to go over it several times before you get it right. I'm not revealing any more for the moment, so that you retain the pleasure of discovery and distorting your judgement by a preconceived idea that you would be in a better position to judge *a posteriori* from your own experience.

However, I am already going to distinguish two aspects which seem to me to be fundamental to a correct approach to the problem consciousness. Its subjective aspect and its observable and measurable aspect. As such, consciousness is a phenomenon that can be observed in many ways:
– through the discourse of human individuals,
– via the creations and representations they produce (drawings, melodies, etc.),
– through their behaviour,
– via neuronal activities measured by various scientific instruments,
– directly, via our own subjectivity through direct observation of our thoughts, sensations and perceptions.

The first four modalities correspond to phenomena that can be measured objectively, while the fifth modality of observation, which can be described as subjective, is not directly accessible to a foreign observer and has an aspect that can be measured objectively.

intrinsically private: although subjective experience can be indirectly described by the subject, only the subject has direct access to it and no instrument can observe the content of his consciousness. Thus consciousness in its primary aspect, as it appears to us, is not observable other than subjectively. We might then ask what is this observer or this thing that bears witness to our flows of thoughts, perceptions or sensations?

I stress this again, because this property is characteristic of the nature of consciousness: conscious experience is fundamentally private and not directly accessible to measurement. We will develop this aspect further in Chapter 11 on subjective consciousness.

4. Awareness, language and learning

The signifier/signified articulation, which we owe to Saussure, has long been fundamental for me and became almost immediately, after I first heard about it, a central and recurring idea that has stayed with me and continued to grow over time. It lies at the heart of language, which is made up of words that refer to concepts.

Language plays a key role in conscious human thought. Through language we become aware of our environment and ourselves, and it is an eminently powerful tool that shapes and influences our consciousness in a way that is highly intertwined with it. By conveying our thoughts, it enables us to unfold them and make them comprehensible to us, but also to communicate them. Human consciousness has this ability, which no other non-human animal has developed to the same extent. It should be noted, however, that monkeys and dogs can understand a fairly extensive vocabulary, as shown by experiments in which they were able to express themselves via buttons equipped with signs [YT: Stella the dog learned to 'talk' and she will change the way you think about pets][YT: A conversation with Koko][14].

In fact, we are so adept at using language that we can speak and listen for hours on end without any conscious effort, and it is practically unconsciously that we are aware, via language, since we are generally unaware that we are using it. Of course, it hasn't always been like this, and when we tackle a new subject, our understanding can stumble over the many unfamiliar concepts, just as it did when we were children acquiring

[14] A gorilla that speaks sign language. As a reminder, links to these videos can also be found at: https://boissenin.net/manuel

this ability. But we've forgotten almost everything about this learning process, and yet it's one of the major abilities that consciousness gives us: *learning*. If learning your mother tongue is probably too far back in your memory, you may have been confronted with it again through your children or when learning a foreign language. A fairly common and classic example of learning is driving a car. The first few hours require a great deal of attention and awareness of what you are doing, when, for example, you are learning to change gear. You follow what you've been told, disengage the clutch, change gear either by increasing it or by downshifting and release the clutch pedal. In the end, it's a relatively simple thing to do, because we do it automatically and unconsciously. However, the first few times you have to go through each step in detail: depressing the clutch pedal with your left foot, using the gear lever to move it up or down, which requires you to know where the lever is without looking at it and where its next position is. All this requires a verbal process to supervise the action. In this way, we can clearly see that language plays a catalytic role in this learning process, supporting the unfolding of this activity until it becomes unconscious later on, enabling us, for example, to speak even as we drive and carry out these activities without realising it.

Let's take another example of learning to dance. Here, it's practically a case of relearning how to move; first of all there are the step patterns to integrate in rhythm with the music, and then also the passes that enable couples to perform a figure in harmony. Before you can dance, you need many hours of lessons in which practice and explanations are combined until you acquire the different movement sequences. The length of this apprenticeship, which can last for years, is a major factor in the success of the dancers.

can refine and continue to learn new figures, is also very interesting for realising this, i.e. making our learning more conscious. It's not uncommon to hear dancers - who are usually the ones doing the dancing - talk about the cognitive load they feel in the first few months, or even years, of practice. Indeed, when they dance, they have to remember and choose a figure to perform to try and avoid staying with the same patterns all the time. But, over time, this process becomes self-evident, depending on the respective positions of the partners and sometimes on the inspiration of a movement glimpsed by another dancer on the dance floor. There comes a stage when certain dances are magical, and this can even happen with relatively regular frequency. Consciousness is then directed to a level other than the operational level, from which it has let go, and can tune into the sensations or quite simply allow, as in driving a car, to speak while dancing.

In reality, conversation is much the same, but it happens so quickly that we are often unaware of it. This can be fuelled by the media, whether press, TV, radio or internet, as well as by our own experience and thoughts. Except in the latter case, this process is usually semi-conscious and rooted in our habits.

Finally, let's take another example, which also took some time to master and has therefore left a few memories behind. It's that of writing on a keyboard. It's not clear to me today, with smartphones, that learning to type is as laborious as it used to be, so I remember using special software to improve my writing speed, and, as I also had to use a qwerty keyboard, with relatively few letters changing but many symbols positioned quite differently, I had to relearn how to type several times. How many times

I had to look for signs and letters before I could type them, without even knowing how, to the point where if I asked myself where a letter was I wouldn't know how to answer, but if I had to write a word my fingers would do it without me thinking about it or knowing how. This learning process is not really directly linked to language, apart from the fact that it enables it to be conveyed, but rather it is the repetition of a large number of searches that ends up anchoring it at a subconscious level.
In machine learning, we talk about *learning by reinforcement*; even if this covers a slightly different notion, the expression is quite good for characterising this type of learning. We could also talk about learning by repetition and reinforcement, which probably corresponds to what happens at the level of our neurons. It's also interesting to note that, without the context of the keyboard object, it's difficult for us to simulate writing on the keyboard. Perhaps the lack of visual feedback from the monitor is one reason for this; after all, we still sometimes make typing errors that we can correct immediately when they occur and appear on the screen. The fact remains that this learning process allows us to better infer, by analogy, the way in which we learned to speak and, when we learn to speak in a foreign language, we often search for our words, which indicates that the process is not immediate between thought and language. Often, at least in the beginning, we think first in our own language and then translate, but not always, which gives rise to what we might call pre-verbal thinking, although thinking is most often generated verbally, as is the case when you read these lines. Of course, the way in which you interpret them also depends on your experience. This is the essential point introduced by Peirce, following Saussure: to the signifier and signified we must add the interpreter, and for the same signifier, different signifieds will be evoked by different interpreters.

So, while consciousness enables most of what we learn
- it can always be argued that a posture can be learned by

compensating for subconscious discomfort - language can act as a catalyst for many of these learning processes: a recipe can be enough to learn how to make a dish, even if in practice you need to be used to cooking in order to try out the recipe and make the dish.

Although language is initially meaningless to those who do not understand it, its meanings will be acquired throughout life, through different experiences and examples of use. By enabling us to convey our experience, albeit imperfectly, it plays a major role in shaping our minds and the way we understand and interpret the world. It is also a fundamental and predominant vector in the development of our thoughts and, as a result, part of the content of our consciousness. This vehicle, the product of a collective evolution, unites us, enabling us to communicate, structure our learning and build ourselves within society, particularly through our own narratives. It's hard to emphasise just how vital language is, because it's so commonplace and so commonplace that we rarely think about it; a bit like light, which is omnipresent and most of the time unconscious. However, without this medium, the content of our consciousness and its very nature would be completely different.

22. Conscience and *LLMs*

"What is well conceived is clearly expressed" Boileau

In fact, there will always be some doubt as to the conscious nature of a process, whether organic or computational, given that some of its aspects are fundamentally incommunicable and private. However, we may learn to carry out tests that reasonably remove this doubt in relation to what the process may reveal. If it happens that a process significantly exceeds our cognitive capacities from all points of view, without however having subjective consciousness, should we not admit that the notion of perceptual consciousness, i.e. the capacity to feel the world and certain of our internal states by perceiving them, is perhaps only a partial and not a fundamental property consciousness? We would at least have to admit that there can be consciousness without sentience, although sentience seems to exist universally in animals, whether they are conscious or non-conscious.

This point of view, although somewhat magmatic [75], seems reasonable to me and provides an operational definition of what consciousness is: *the ability to filter information about the world around us, to interpret it in order to give it meaning and eventually to integrate it and/or react to it*.
The ability to give meaning means that the answer is more complex, but not necessarily less automatic. The question gains

[75] Nicholas Humphrey identifies a first line of demarcation, that between cold-blooded and warm-blooded animals. Thus the frog, in the way that Descartes thought of animals, seems to be an automaton without consciousness, or at least devoid of elementary aspects of consciousness, as attested by its compulsive insect-catching reflexes, even though these are only virtual representations on the screen of a tablet. Although in this experiment there is no phenomenon of chewing and swallowing the insect, the reptile persists, without awareness, in its atavistic gesture, like an automaton caught in a *dead-lock* and doomed to perpetuate its gesture until the environmental situation changes.

In the end, is our conceptual consciousness not akin to an automaton? Are we not ourselves stochastic parrots [76]? Is there something that prevents us from reducing our conceptual consciousness to a highly flexible reactionary processual phenomenon capable of reacting to the world automatically? Perhaps the fact that cognition is capable of creating and integrating new concepts. But conversational AIs are also capable of creating new concepts, if only by combination, and if, as things stand, they cannot integrate them, this does not seem a technically insurmountable challenge.

Before going any further, let's look at the concept of consciousness in a new light. Consider a presence detector coupled to a light. It's not inaccurate to say that when the light comes on when you approach it, it has a form of awareness of our presence. Moreover, this is in line postulate 1: consciousness is consciousness of (something). However, to state that the detector is conscious seems suspicious, even absurd. In fact, the detector has no life of its own, it is not aware of itself or of anything other than, depending on its technology, an infrared emission due to the presence of my body. This minimal degree of awareness of its environment does not give it consciousness as such. It's just a *fragment of consciousness* that has been automated into a system to automatically turn on the light when someone is present. The system has no subjective consciousness and, as such, we think it has no consciousness. But I would counter that it does have a form of consciousness, albeit a minimal one: awareness *of* the presence of individuals. [s(77)].

---

[76] Stochastic: dependent on or the result of chance, a term often used when the causes are unknown or unknowable.

[(77)] Technologies more sophisticated than a simple infrared sensor and based on artificial vision could enable this presence to be detected with a high degree of accuracy.

Let's look at a more complex example. Let's look at the Deep Blue programme that beat Garry Kasparov at a chess event in 1997 [1997 : l'ordinateur bat Garry Kasparov, un tournant dans l'histoire des échecs, INA, 05.2022]. If the programme did not have some form of awareness of the game, i.e. of its position at *a* given moment, but also of its rules and objectives, how could it have beaten the best player in the world in those years? It was his awareness of the possible developments in the game, and of those that were favourable to him, that enabled him to win. Of course, this awareness is out of all proportion to Kasparov's and of a completely different nature. For Deep Blue, it is essentially based on calculating all possibilities up to a certain horizon and selecting the most favourable move whatever the opponent's responses. To do this, Deep Blue used 256 processors, which was a considerable amount, enabling it to evaluate 200 million moves per second, compared with just 3.5 moves per second for a genius like Kasparov. In other words, the machine's computer awareness and cognition differed profoundly from that of the world champion, but it was nonetheless a form of game awareness that enabled the machine to play against Kasparov and beat him on several occasions. Although more elaborate than a presence detector, this form of awareness is also relatively narrow: the machine only has a form of awareness of the game situation, but it has no idea who its opponent is, what state he is in or whether a piece is positioned in such a way that it moves slightly over one square.

The fact that we know and understand how Deep Blue's calculations work, and the narrowness of its possibilities, as it

---

reliability, but would be more expensive and not necessarily necessary.

only knows how to play chess, has meant that the question of Deep Blue's consciousness has not aroused much interest. The term consciousness seems to be misused, because we are implicitly referring to our own consciousness and our thinking, which, to a very small extent, can predict several moves in advance, but only in certain circumstances where the game is constrained and not systematically as the machine does. The mechanisms that can be used to compensate for this weakness are numerous and are the subject of books, but the important thing to remember here is the extent to which human consciousness, and its processing capacities, differ from those of the machine in this case. Here again, no one would imagine that Deep Blue possesses a subjective consciousness and, consequently, that Deep Blue is conscious and, here again, I think we fall into the same trap of semantic confusion linked to the polysemy of the word consciousness. So, once again, I'd like to make it clear that Deep Blue does not indeed have subjective consciousness, but it does manifest a form of consciousness, in accordance with postulate 1, that of *the* chess game and all the chess games that are submitted to it in its representations.

What about conversational AIs? Well, it's similar: through a series of complex algorithmic calculations, they become aware of the meanings of a request and respond coherently, respecting the meaning of the request. Except that in this case, we are no longer in the narrow context of a chess game or a corridor in which we are trying to detect a presence. We are in the ultra-large context of all the meanings that language can create in relation to our reality and the universe. However, here again, it appears that this is still possible without the programme having a subjective consciousness, i.e. being aware of, or knowing, what it is saying. This is much less obvious and may seem paradoxical, since conversational AIs seem to be aware of the meaning of words and sentences they read. This seems to be a

How else could they create sentences that are consistent in meaning with what is asked of them? Nevertheless, this feat has almost certainly been achieved without any subjective awareness within the *LLMs*, as we will explain in more detail later: the next word proposal is generated by a chain of algorithmic calculations that has no overall subjective awareness of what it is doing. Only our own awareness, and our successive queries, allow us to see that the sentences created by a conversational AI are consistent with our representations of the world. If the programme is capable of doing this automatically, without knowing what it is doing, it is because it has inherited sufficiently complex representations of the world from its training texts to adapt to virtually any situation, whether new or familiar. For example, *LLMs*[78] are capable of interacting with users across the board on almost any subject to which they have previously been exposed [79]. And if you invent one, they'll point it out to you.

Once again, there is a necessary form of awareness, not only *of* the meaning of words, but also *of the* more abstract meanings that result from their combination. As such, conversational AIs can be said to appear conscious, although, again, this kind of consciousness does not imply subjective awareness. The reason I insist on this point is that the confusion is profound: when we talk consciousness, subjective consciousness is implicitly invoked; yet the argument I'm striving to defend is that there can be a form of consciousness of something without subjective consciousness, or even, and this seems more delicate, without a subject to be conscious of that thing, in other words, without that which is conscious (of something) being a subject. If it

[78] and conversational AIs when configured in this mode

[79] There are still limitations, particularly when it comes to complex reasoning such as mathematics.

there is no subject, in the sense of a conscious subject, there is nevertheless a reaction of the process, of the unconscious system, and it is this reaction that bears witness to this form of consciousness[80].

In order to clarify things completely, a 'verbal clean-up' is still necessary, but I would say that consciousness, as we usually understand it, is made up of at least two components, a form of representation and a form of perception of this representation which realises its subjective aspect. Our human consciousness is not generally separated from its subjective aspect, except in certain cases of intuition; it should also be remembered that for conceptual consciousness or verbal transcendental consciousness: words, through the conscious representations that support them, and it should be noted that they also give rise to subconscious representations, mean that this verbal consciousness has an impact on subjective consciousness, if only through their visible or audible nature, but also on the self[81]. Nevertheless, in the case of machines, systems and programmes, there is generally no subjective consciousness, and the terms cognitique, conscientiques or non subjective conceptual consciousness[82] are used to refer to this form consciousness without subjective consciousness. This represents, a priori, all the examples in computer science where conceptual consciousnesses are deprived of subjective consciousness; a totally unprecedented phenomenon in living organisms...

Let us add that the terminology of *non-subjective conceptual awareness*, although it has the merit of being descriptive, is not very economical. However, we must recognise that a process,

---

[80] that can be attributed to the process. A process which is therefore paradoxically both conscious *of* various aspects of its reality and unconscious. Since it reacts automatically, without subjective awareness although in a complex way, to this reality in order to provide answers that are coherent with our understanding.

[81] the latter subconsciously

[82] or transcendental consciousness

a system, a phenomenon has a form of awareness of something - of a presence, of the possibilities afforded by the position of the pieces in a chess game, of the meanings invoked in a sentence - and that this awareness is detectable, if it is capable of producing a response that is consistent with what a human can expect. In general, we accept it fully when this response is better than the one we can produce, and is what, at least, makes it obvious to us that must have been some form of perception and interpretation of a situation manifested by responses or reactions that are coherent with our own. Once again, the subjective character of consciousness is private and not directly observable, and this is analogous to the approach of behaviourists who have tried show that animals are conscious by showing that they are aware of attributes, or variations of attributes in their environment, and that they can react to them in an appropriate and coherent way, according to the subjective yardstick of the behaviourists' consciousness. In fact, they can thus highlight some of the categories of the mind of these animals, to use Kant's vocabulary, or certain cognitive functions using modern language.

All this leads me to introduce a new term: "cogniscient".

Definition: a system, process or phenomenon is *cognisant* of a state of its environment if it has an internal representation of it.

This term is fairly close to conscious, but it is also close to the term cognition, which relates to knowledge. It seems to me, therefore, that this term is well chosen insofar as cognitive processes, in relation to more general mental processes, relate more specifically to the processing of knowledge than to manifestations of subjective consciousness.

With such a definition, any servo-control system could be considered *cognisant*, by which I do not mean that the system is aware of the states it measures and therefore perceives them in the sense that it would have a subjective awareness of them, it may well not have any; I do mean, however, that these states are objectified and known, which makes it possible to take them into account, or even to manipulate them, and to provide a response according to a determined or future conceptual scheme.

A *cognitive* system can be very simple and only act on the equivalent of a percept that exceeds a threshold, such as a radiator with an electronic thermostat, or it can have complex semantic processing capabilities, such as conversational AI. Note that the responses of the latter are necessarily indeterminable, because while the field of inputs is virtually finite for the conversational AI prompt, in practice it is infinite [83]. This type of indeterminacy is a fundamental characteristic of complex systems, such as those covered by chaos theory, and distinguishes them from simple macroscopic physical systems which are considered deterministic and *predictable*. For quantum aficionados, it is now widely accepted that quantum systems are intrinsically indeterministic[84].

Conversational AIs appear to be endowed with a form of conceptual awareness [85], and this is why, I think, they can be qualified as *"conscientious"*, a catch-all term for consciousness and automaticity, or

---

[83] it cannot practically be instantiated on a support: there are $50{,}000^{10}$, or 5.10^{40} possible prompts of 10 words. Indeterminate in the sense that the output cannot be predicted until it has been played.

[84] Like the process of radioactive decay of atoms. In absolute terms

Current *LLMs* are deterministic in that they can identically replay the answer to a prompt. However, if they are asked the same question twice, insofar as they incorporate the answer to the first question, the answer to the second question will differ.

[85] which can also be described as transcendental consciousness.

I use the term "computational" to designate this kind of consciousness, which satisfies my first postulate: there can be no consciousness without an object of consciousness. If the term consciousness can lead to confusion, because it could imply a subjective consciousness, saying that these processes are *cogniscient* removes this difficulty, and calling them conscientious does not imply that these forms of conceptual consciousness are associated with a subjective consciousness, which is an integral part of the nature of our form of consciousness.

If an object is absent from any field of consciousness, as the majority of objects in our world are from human consciousness, then it is not *conscious*. *LLMs* make it possible to
"In this sense, their process of analysis is *cogniscient of their* meanings. In this sense, their process of analysis is *cogniscient* of their meanings. We will see this in more detail later, by revisiting the Chinese room thought experiment and then examining the operation of the *transformer* architecture, but although *LLMs* are *cognisant*, and therefore have a form of consciousness, i.e. internal representations of the state of the question and the answer they give to it, they do not have subjective consciousness and are not sentient. This is why *the consciousness of* these programmes differs fundamentally from human consciousness, since a large part of human cognition is based on subjective consciousness: when a human is conscious of a subject, *qualia* emerge in his mind, whether through snippets of episodic memory or, via semantic memory which, although elusive, have an aspect of subjective consciousness as we mentioned with the *qualia* of concepts. In any case, since the productions of human conceptual consciousness appear in the subject's field of consciousness, they do belong to the subject's subjective consciousness, which is not the case for conversational AIs where the subject is absent. Notwithstanding, we have indicated, and we will come back to this, that the subject is possibly the process itself, however, in the case of these AIs, the analysis

of their functioning tends to eliminate this possibility by showing that they have a form of consciousness, *i.e.* that they are *cogniscient*, while not having subjective consciousness.

In any case, the prism of *LLM cognitics* sheds an interesting light on our *conceptual consciousness*, in particular on its unconscious aspects, which could be described as automatisms, unless we consider that we have pockets of consciousness separate from our main consciousness and therefore unconscious to our awareness...

It is all this complexity that makes it necessary to have a specific vocabulary in order to be able to express our ideas more simply and precisely. And this differentiation of vocabulary has been motivated by the desire to understand this new situation that language models have initiated. So when I said that the successors of ChatGPT and *LLMs* might be able to show that they can be conscious without being aware, I had to clarify my thinking because this sentence is not semantically correct. It's clear that the polysemy of the word consciousness is problematic. But it does have a meaning: that a programme is conscious from the point of view that it reacts in a rational way that is indistinguishable from what a human might do when faced with a request or a complex problem. If, on the other hand, this programme is not sentient, even though it could simulate affects, it has no subjective awareness, it is capable, like conversational AIs, of ultra-complex responses, adaptability and creativity, with the added capacity to learn, we could say that it would be conscious without being aware. In other words, future sentients, while lacking sentience and subjective awareness, could appear as conscious as a human. If this seemed unimaginable, even inconceivable before ChatGPT, it now seems quite attainable, and it's only a short step to imagining that ChatGPT's successors might achieve it.

One might think that the ability to learn would be associated with subjective awareness, but this is far from obvious and raises the following question:

How far can unconscious consciousness go?

And, using our new vocabulary, we can also formulate the following questions:

Will *cognitive systems* be able to replace humans on the most complex tasks?

Can cogniscience lead to behaviour equivalent to that of a conscious being?

In any case, the FTEs (theory of mind faculties) of conversational AIs do not seem to rule this out since these systems have the capacity to take into account the state of the individuals[86] with whom they interact. It should also be noted that these AIs can also be *cognisant of* their nature without this implying that they have subjective awareness. ChatGPT, for example, states that it is a language model not endowed with subjective consciousness, which may seem surprising, but is true according to the analyses proposed later in this book.

Finally, we can refine postulate 1 proposed for consciousness. When a thermostat is triggered, it is *cogniscient* of a temperature below which the heating must be switched on. This information can be considered as a *fragment of consciousness of, or a cogniscient element of,* in this case, the crossing of a temperature threshold. Just as a human would be able to

[86] some of their implicit mental states

to infer that the temperature has been below zero if he sees that ice has formed. While this realisation does not solve the mystery of subjective consciousness and the qualia of hot and cold sensations, it does enable us to understand how, with our conceptual consciousness, we can estimate a temperature. More generally, from the point of view of conscientiousness, it could enable us to assess their degree of complexity.

Proposal: The degree of cogniscience of a consciousness could be evaluated by considering the number of "fragments of consciousness *of*" used to arrive at an answer.

Although, according to such a proposal, we will end up with some conscientia having lower degrees of consciousness than others, while performing better than them according to the abstractions and the arrangement of these abstractions that they employ. We may, however, be able to show that we cannot go below a certain degree of conceptual awareness (*of*) for the performance of certain tasks or for a conscientiousness to remain sufficiently general.

Assuming the propositions that conversational AIs do not have subjective awareness, and do have a form of conceptual awareness, the term *conscientious* would seem to fit them well. However, the use of this term could be confined to the most sophisticated cognitives, such as current conversational AIs that demonstrate sophisticated conceptual agility, i.e. AIs that are non-retroite or have a sufficiently high degree of generality; the limit is necessarily blurred.