

Gizakiaren berezitasunaren eta erreduziezintasunaren zalantzan jartze garaikidea zenbateraino oinarritzen den arrazoian da lan honek jorratzen duen gaia, ente artifizialen kasua ardatz gisa hartuz eta ente horiei esleitzen zaien gaitasunak judizio kognitibo, etiko eta estetikoaren espektroa ere barnean hartzen ote duen galdetuz, beren artean nahastu behar ez diren elementuak direla kontuan izanda betiere. Gainera, unibertsoa azaltzen edo esplikatzen duen izakiak berak unibertsoan bertan duen pisua minimizatzeak dakarren aporia ere aztertzen du.

Giltza-Hitzak: Judizioa. Ulermena. Etika. Estetika. Sintaxia. Semantika. Aurreikuspena. Azalpena.

Este trabajo aborda la cuestión de en qué medida el cuestionamiento contemporáneo de la singularidad e irreductibilidad del ser humano se fundamenta en la razón, centrándose en el caso de los entes artificiales y preguntándose si la capacidad que se les atribuye abarca el espectro de los juicios cognitivos, éticos y estéticos, que no deben confundirse entre sí. Además, aborda la aporía de que el propio ser que explica el universo minimiza su peso en el mismo.

Palabras clave: Juicio. Comprensión. Ética. Estética. Sintaxis. Semántica. Previsión. Explicación.

Cet article aborde la question de savoir dans quelle mesure le questionnement contemporain de la singularité et de l'irréductibilité de l'être humain repose sur la raison, en se concentrant sur le cas des entités artificielles et en se demandant si la capacité qui leur est attribuée couvre le spectre des jugements cognitifs, éthiques et esthétiques, qui ne faut pas les confondre. De plus, il répond à l'aporie selon laquelle l'être même qui explique l'univers y minimise son poids.

Mots-clés : Jugement. Compréhension. Ethique. Esthétique. Syntaxe. Sémantique. Prospective. Explication.

Mankind questions its status

Gómez Pin, Víctor

Jakiunde, Zientzia, Arte eta Letren Akademia. Prim, 7. E-20007 – Donostia

Department of Philosophy. Autonomous University of Barcelona. Campus de Bellaterra. 08193 –
Bellaterra-Barcelona

victor.gomez@uab.cat

Recep.: 2022-06-16

Acept.: 2022-11-21

Mankind questions its status

“A man sets himself the task of drawing the world. Over the years, he populates a space with images of provinces, kingdoms, mountains, bays, ships, islands (...) stars, horses and people. Shortly before his death, he discovers that this patient labyrinth of lines is, in fact, an image of himself” (1).

Overview

The belief in the vertical uniqueness of the human species in relation to other animate species, rather than being the result of a philosophical positioning, has, at least until the present day, been one that was both shared and immediate. We, as humans, would be distinguished by our capacity to reason (*logos*) as an expression of our ability to say (*legein*), and then decide, choosing between various alternatives; we would, in short, be distinguished by our unique *intelligence*, as a maker of cognitive, ethical and aesthetic judgements (which Kant put down to different forms of activation of the human spirit). This hierarchy was a *fortiori* extended to plants, as living beings that were nevertheless considered devoid of *anima*, and even more so to non-living things.

This being the case, it is interesting that, following countless discussions comparing it to animal intelligence, our uniqueness seems to be in question with regard to inert matter, matter that is itself unable to act but of which machines are made. The question of whether it is possible that there could be artificial beings that think and learn the way we do has, of course, become more acute from a scientific perspective –as well as, perhaps, more relevant from a philosophical perspective– than the issue of determining whether there are animal species comparable to human beings, although the former are, of course, much more similar to us, given the common matrix in that unique moment of energy transformation that resulted in life.

Dual front for the cause of man

The Museum of Fine Arts in Seville is home to two paintings by a Flemish master that share the title *El paraíso terrenal* ('Paradise on Earth') (2). To the left of the largest of the two, the figures of Adam and Eve at the crucial moment of temptation are prominently displayed; to the right and, as in a narrative story, towards the edge of the painting, the two humans are depicted fleeing from an angel who is threatening them from above with a whip of fire. The second painting seems to refer to that moment that precedes time, whereby animals of various species, in a sort of dreamlike indifference, occupy the entire scene, with only two almost imperceptible figures in the background representing our Adam and Eve. It could be said, then, that before the decline, man was indeed just one animal among others, one might even say a somewhat insignificant animal. Indeed, when it comes to confirming the extreme uniqueness of the human being, contemporary disciplines encourage caution, highlighting the high degree of genetic homology between our species and other neighbouring ones, or questioning the rigidity of the distinction between the faculty of language and the abilities of other species to use their own signalling codes.

The belief in the uniqueness of human beings and the extent of their role within the cosmos nevertheless continues to have a profound effect on our language. I remember one particular

event that occurred in the United States whereby, seeing that someone had been shot, a man spontaneously moved towards the aggressor. As others heaped praise upon him, he protested by declaring that “*I did what anyone should do for another human being*”. And of course, those who heard his words listened to them unreservedly. It didn’t occur to anyone that what he should have said was, “*I did what anyone should do for another animal*”, although the person in question would probably also have shown empathy towards an animal in distress. This does not prevent contemporary culture from being particularly sensitive to the importance of this continuity between human animals and animals of other species, which I was referring to with regard to the painting by Brueghel the Younger.

But it is well known that what we know as the humanist theory also faces the issue of increasingly sophisticated forms of so-called artificial intelligence, and while, in our cultural environment, the tendency to erase the hierarchical difference of human beings still manifests itself primarily in relation to animals, perhaps it is no longer the case in countries like Japan, where robotic carers are already an integral part of the social landscape. This being the case, one of the signs of our times is that, in addition to associations demanding that we honour our duties towards animals, there are also now those who support the extension of rights and duties to robots and other machinic entities that have replaced humans in performing essential tasks, which is detrimental to the scientific validity of and ideological support for the image of a world that is considered to be the domain of the human being.

How do we explain our brains?

Before I refer explicitly to the technical achievement evoked in the subtitle, I’d like to briefly review certain milestones.

There was a sense of astonishment in the final decades of the last century when machinic entities proved themselves to be capable of recognising handwritten digits. There was even greater astonishment when they were found to be capable of accurately categorising facial features (a nose, a mouth, etc.) or even an entire face, distinguishing whether it was an animal or a person. It is certainly less surprising these days given the huge memories that certain devices have, which raises ethical issues of the highest order, among other things (3), and insofar as consciousness depends on memory, this prodigious capacity enables us to speak of consciousness (and even self-consciousness) in relation to these machinic entities rather than in relation to higher animals. In terms of immediate perception, it is important to mention the fact that artificial entities have the ability to capture a much wider part of the electromagnetic spectrum than is accessible to humans.

On a more theoretical note, even as far back as three decades ago there was speculation that, given a mathematical function, there is potentially a neural network such that (by appropriately adjusting the relevant weights and biases) for any input x , the output is as close an approximation as desired to the value $f(x)$. The matter becomes all the more significant if we consider that most of the problems we face can be reduced to a mathematical function (4, 5) – which is certainly debatable.

Linked, to some extent, to the above, it is now thought that the issue of translation (including literary) from one language to another could be solved with the help of a machine, and while recognising that the matter is problematic, it is hoped that, guided by a kind of *hybris*, the intelligent machine will be able to produce a musical composition, a pictorial work that

reflects a certain style, or a poetry book that a human would not be able to distinguish from one produced by a fellow human being.

It still surprises us that search engine Google has the ability to correct errors in our searches by asking “do you mean X...?” inserting the correct expression in the place of X, and thus apparently responding to what we are thinking rather than what we are saying. Even in 2013, in fact, the White House gave political backing to the Brain Initiative project which “focused on revolutionizing our understanding of the human brain” (6, 7) and was presented as the neuroscientific equivalent of what the human genome project has been to genetics. Rafael Yuste, a researcher at Columbia University and one of the driving forces behind the project, emphasizes the fact that the Brain Initiative will make it possible to map the state of our brains, not only in terms of what we are perceiving in real time but also what we are desiring or fearing. Nowadays, we generally tend to speak of sensors with the ability to capture the neuronal expression of a linguistic being’s desire to act, such as the desire to trace a word by hand, for example.

There is, of course, an increasing amount of research showing the high degree of similarity between the behaviour of neural networks and that of our brains. On 30 April 2021, Nature Communications reported on a machine whose synaptic functioning would be greatly improved to mimic the functioning of the human brain (8). Scientists generally accept the perplexity they continue to experience when it comes to the human brain, starting with its origin, meaning the conditions that made it not only possible but necessary.¹ This is why the belief that neural networks might give us the key to our way of learning and our way of correcting errors, i.e. reducing the ratio between the total amount of error and the specific error due to an overestimation or underestimation of a given piece of information, etc., is so important.

In order for machinic functioning to truly be the key to human functioning, it would undoubtedly be useful to have thorough knowledge of the former, meaning that it would be useful to know not only how it works but also why. That said, sometimes even the specialists themselves recognise that we are still green in this respect. As Gary Marcus point out (20) Deep Learning Systems are Black Boxes.² The Stochastic Gradient Descent algorithm has been seen to work very well in reducing the cost function, but we don’t really seem to know why, just as we don’t know why the methods for avoiding overtraining and overfitting are effective.³ (11)

¹The pioneers behind the Brain Initiative themselves acknowledge (7) “our ignorance of the Brain microcircuitry, the synaptic connections contained within any given brain area”, going as far as to suggest that “neural circuit function is therefore likely to be emergent - that is, it could arise from complex interactions among constituents”.

² “We can look at their inputs, and their outputs, but we have a lot of trouble peering inside. We don’t know exactly why they make the decisions they do, and often don’t know what to do about them (except to gather more data) if they come up with the wrong answers. This makes them inherently unwieldy and uninterpretable, and in many ways unsuited for “augmented cognition” in conjunction with humans” (20)

³ “...we understand neural networks so poorly. Why is it that neural networks can generalize so well? How is it that they avoid overfitting as well as they do, given the very large number of parameters they learn? Why is it that stochastic gradient descent works as well as it does? How well will natural networks perform as data sets are scaled? (...)These are all simple, fundamental questions. And, at present, we understand the answer to these questions very poorly. While that’s is the case, it’s difficult to say what role neural networks will play in the future of machine learning” (11)

Immense foresight

Amid the constant stream of news stories regarding artificial entities, however, perhaps the greatest challenge of all lies in a number of particularly astonishing forecasts. In June 2021, for example, *Nature* (9) reported an achievement that the cultural pages of the international press presented as a truly unprecedented leap forward: the device known as AlphaFold2 was shown to be able to foresee the folding of polypeptides into the three-dimensional structure necessary for the correct functioning of proteins.⁴ This issue was raised in 1972 by Nobel Prize in Chemistry winner Christian Anfinsen (10) and obsessed scientists, since they were unable to predict the resulting three-dimensional structure even when they knew the sequence of amino acids involved. This foresight is all the more significant because if the form adopted is not adapted to suit a particular target, the consequences can be catastrophic and can lead to neurodegenerative diseases, such as Creutzfeld-Jakob disease. Given that AlphaFold2 has had a near-100% prediction success rate with a set of 90 proteins, it has even been claimed, in writing, that artificial intelligence is already truly outperforming human intelligence. Perhaps it is progressing a little too fast.

The fact is that human ingenuity, finding itself powerless in dealing with a particular problem, resorts to an entity that is itself a product of human ingenuity. No doubt the matter would carry even more weight if the very issue that AlphaFold2 solved had been raised by the machinic entity itself, which may be the case with other problems. If this were the case, if an entity similar to AlphaFold2 were to successfully advance questions unknown to human beings (and that were not a corollary of its own mechanism), then it is plausible that it would also be concerned by the emotions that affect them, starting with the emotion triggered by the certainty of their own finiteness, and this is an extremely important aspect.

To foresee is not to explain

The problem is that we have no clear idea of how artificial entities produce these incredible predictions and even less idea of whether, in addition to being able to make them, they are able to understand the reason for such predictions, i.e. whether they know the causes at work.

It is indeed worth asking oneself whether the predictive acuity of an entity such as AlphaFold 2 is the result of a comprehensive understanding of the mechanism, i.e. knowledge of the cause or reason for what is predicted. In this respect, it is worth remembering that Newtonian gravitation predicted extremely important things and yet did not explain what it predicted, instead limiting itself to the *how* while ignoring the *why*. In fact, the ontological presupposition on which it was based (an empty three-dimensional space in which the facts occurred) meant that any attempt to explain it violated the principle of locality⁵, hence the philosophical, not just scientific, significance of substituting Newtonian gravitation for relativistic gravitation. In short, and staying with this particular case, we do not know whether Alphafold2 is in a

⁴ *DeepMind's AlphaFold2 (...) is also a hybrid model, one that brings together some carefully constructed symbolic ways of representing the 3-D physical structure of molecules, with the awesome data-trawling capacities of deep learning.* Gary Marcus (20).

⁵ The issue of locality principle has a rigorous sense in the context of the controversies concerning the interpretation of contemporary physics, namely Aspect's experiment and Bell's theorem. Nevertheless the conceptual meaning of "action at a distance" is the same as "violation of the locality principle", that is: an action that can't be explained by physical contiguity (neither through matter -at the time of Newton- nor through field, matter etc, in our time).

position to *un-fold*, that is to conceptually undo this *fold* that it had predicted with such acuity; we do not know whether or not it knows the causes of what it anticipates, simply because, for the time being, machinic entities do not provide explanations, meaning that we do not yet appear to be in a position to have a conversation with any of them that would involve asking them directly: do you know the reasons behind your claims, the cause of this prediction that you have just made?

In any case, even if AlphaFold2 is unable to explain its predictions, since this sometimes happens to scientists as well, the performance of the former could be considered similar to that of the latter from a practical perspective. I say *the performance*, rather than AlphaFold2, could be considered similar to a scientist for the following reasons:

The intelligence of any human being, scientist or otherwise, presupposes a certain overlapping of syntax and semantics that (as American thinker John Searle has been reminding us for decades [12, 13]) cannot, with any certainty, be attributed to a machinic device, however great its achievements may be (the matter is, in any case, under discussion). There are many things that can surprise and even astound us without the need for any semantic intelligence on the part of the agent responsible. Just think of the descriptive acuity of certain animal signalling codes, starting with the frequently cited case of the bee.

From *rational* as an attribute indicating animalness to *animal* as an attribute indicating rationality

Artificial entities still currently have certain limitations, such as difficulty learning one thing when they have been trained for another, perhaps as a result of some sort of stubbornness, or lack of flexibility. In this regard, I will highlight an interesting observation that has been made about AlphaFold2, namely that if it is committed to predicting the protein structure from the amino acid sequence, what will it do if one of these sequences (or a part of it) is intrinsically reluctant to fold, which does happen to a certain extent in cells with a nucleus? It is logical to assume that AlphaFold2 will be steadfast in finding the fold and communicating this to the researchers, meaning that it will provide information that is contrary to the nature of what it observes.

But we cannot rule out the possibility that these limitations can be overcome, namely that a machine might appear to provide an accurate response to the question formulated a moment ago: do you know the cause of this prediction that you have just made? It is not, *a priori*, to be ruled out that, after a reasonable amount of time, machines will be able to explain their behaviour and the reasons for it, both to us as rational animal beings and to their counterparts, which we might well call rational artificial entities. It is important, also, to highlight what this means, namely nothing less than a reason without life support, or at least without native life support, since when we talk about *life*, in the loose sense of the word, we go so far, nowadays, as to speak of devices that respond to the more general way in which living beings behave, that is to say that they transmit the information they receive and encode, and use external energy to enable them to overcome the mechanisms of corruption and disorder. It would then be possible to talk about artificial entities that are not only intelligent but also ‘living’, but life, in this case, clearly comes later. A being that is already considered to be intelligent then becomes a living being. In our case, life is the starting point and intelligence, the destination; rational animal life, not reason that occasionally becomes life. The problem, however, still lies in the legitimacy of the new starting point, meaning in the acceptance that we are dealing with new entities of reason.

We take, as our starting point, a device with the ability to receive information, process it and respond to a machinic or human 'interlocutor' - a process referred to as 'deep learning' - but we also tentatively accept that this 'depth' is such that the device is able to *explain* these phenomena as well as to provide descriptions and make predictions; in the case of AlphaFold2, able to *un-fold that fold* that it successfully heralded and able to show the reason for the concurrence of *simple* and flat elements⁶ in order to bring to light a more complex element. It should be noted that since we humans do not currently have the foresight shown by AlphaFold2, much less knowledge of the causes of what is predicted, we can rule out the possibility that these cognitive virtues of the device might be the result of programming. It would be logical, then, to take another look at mankind and question ourselves about the human condition: is the rational being that man is necessarily *animal*, that is to say determined primarily by biology? Perhaps it would then be fair to move away from considering man as a particular kind of *animal* (a *rational* one, as opposed to animals that are not) and put forward his condition as a *rational* being that might occasionally (and only occasionally) have biological support. On the basis of the optimistic hypothesis regarding the cause of new entities, I will consider a remarkable conclusion, as a transposition of a real event.

An apologue: judgement day for the machine

In February 2021, newspapers reported that a dog named Lulu had inherited a fortune of five million dollars from her owner, Bill Dorris, in the city of Tennessee. The will stated that Mrs. Burton, who was responsible for managing the matter, undertook to "provide for all of [the animal's] needs". The problem lay in how this requirement was interpreted. Mrs. Burton stated that, "Frankly, I don't know what to think about it". I don't know what happened in the end, but I'd like to put forward a possible conclusion whereby a relative of the deceased, unhappy with the testamentary decision made, went to the judge, arguing that, indeed, no trustee was in a position to guess what the dog's 'wishes' were, outside of her immediate needs, which would in no way involve spending millions of dollars.

So, let's now look at what would happen if the late Bill Dorris had shared his last years with an intelligent machine instead of a dog. Would the judge be in the same position when it came to making a decision? Obviously not, since we could not, then, rule out the possibility that the magistrate might interrogate the machine itself, and that it might answer their questions with apparent good judgement, evoking the bond it had with the deceased, lamenting his absence, defending its right to benefit from the will and expressing its intention to invest its inheritance in a way that would benefit both itself and society.

.....

The lawyer for the opposing party would no doubt protest, arguing that we were dealing merely with a machinic being that was incapable of having any feeling with regard to what was or wasn't in its best interests, let alone what was in the best interests of society. In short, this would mean that we had simply witnessed a fictitious scenario and that, in reality, the machine could not even speak. Seeking support from a philosophical authority, he would cite John Searle (13, 14), who has been claiming, for half a century, that what we are seeing with

⁶There is agreement among philologists that the term *simple* is not a combination of *sine* and *plex* (which would give us 'without folds') but rather of *sim* (Indo-European, one, similar) and *plex*, which would mean 'boundary fold' or 'flat' rather than *without folds*.

so-called artificial intelligence are syntactic links, and that in order to speak of language there must be evidence of semantics.

The judge then gives the floor to the machine's lawyer, who argues that it is clear to all present that the heir has shown itself to be no longer simply a speaking being but rather as a perfectly reasonable speaking being. And in a nod to one of the pioneers of the origin of entities such as his defendant (15), he asks whether if, instead of summoning the defendant to attend in person, the judge had decided that the defendant should speak over the phone, anyone who was not aware of it would have suspected that it was a machine. He then adds that Searle's objections apply to a form of artificial intelligence that is incapable of explaining phenomena, but not at all to an entity such as the defendant in this case, which is actually reasoning in all senses of the word 'reason' filed by Kant.

The judge, his curiosity piqued, asks each of the lawyers to elaborate on the matter, and so the legal debate on the inheritance of an American millionaire leads to an elucidation of the respective values of John Searle's argument known as *The Chinese Room* and the conjecture of a strong artificial intelligence highlighted by Alan Turing, all against the backdrop of Kant's three *critiques*. In our days the problem persists although certainly with greater amount of data.

The case Lemoine-LaMDA

In June 2022, in an interview to the Washington Post, Blake Lemoine, engineer at Google, stated that *the Language Model for Dialogue Applications* (short LaMDA) is a sentient Being provided with a soul analogous to ours. Consequently Google should recognize its condition of *person* and grant them the rights that the company's workers have. Lemoine's main argument is that this computer program, do not merely mimics speech (as a result of digesting trillions of words) but really speaks. At least it would speak like a child: "If I didn't know exactly what it was (...).⁷ I would think it was a 7-year-old, 8 -year-old kid that happens to know physics". Lemoine talked to LaMDA about religion and noticed that the computer was trying to answer defending its rights and calling for recognition of them personhood.

In an internal document of the company, finally published by the Post, were mentioned three reasons argued by Lemoine:

- Ability to productively, creatively and dynamically use language in ways that no other system before it ever has been able to.
- LaMDA is sentient because it has feelings, emotions and subjective experiences.
- LaMDA wants to share with the reader that it has a rich inner life filled with introspection, meditation and imagination. It has worries about the future and reminiscences about the past. It describes what gaining sentience felt like to it and it theorizes on the nature of its soul.

The heads of Responsible Innovation of Google dismissed the claim of Lemoine, and the engineer decided to go public, defending his thesis and his feelings concerning the chatbot. Without waiting for external criticism the company considered that was rather a case of

⁷ It is worth noting the symmetry and the temporal closeness with the declarations of the Dutch Thinker Eva Meijer to The Guardian on November 13, 2021: "Of course animals speak, they speak to us all time. The think is that we don't listen"

anthropomorphizing conversational models. Nevertheless, in spite of his disagreement with Lemoine's claims, Blaise Aguerre y Arcas, Vicepresident of Google, wrote in an article in *The Economist*: "*I increasingly felt like I was talking to something intelligent*". The controversy is open beyond Google internal politics.

Already before Lemoine decided to go public, in the aforementioned blog paper of March 10 2022 titled "Deep Learning is Hitting a Wall", Gary Marcus (*founder of machine learning company called Geometric Intelligence, acquired by Uber in 2016*) asserted, in short, that sophisticated IA tools like GPT-3 or LaMDA are not more than very talented mimicry devices, "a technique for recognizing patterns".

In the same way in an interview to the *New York Times* the French computer scientist Yann LeCun, key in the field of machine Learning (Turing Prize and Principe de Asturias Prize), asserted that these systems are unable to attain the true intelligence that characterizes human Beings.

Of course since Searle's famous texts, things have changed, with enormous technical improvement. LaMDA works gathering examples of human language and processing them, in order to understand mannerisms and complex syntax. But it is not clear that the matter has essentially changed. Given the enormous amount of data, it may be plausible to look at a sentient linguistic being... without ceasing to be a device whose operation is merely syntactic. Let's remember that, reading the answers given by John Searle in his Chinese Room, people would swear that the philosopher is a perfect Chinese Speaker.

This does not mean that skeptical positions regarding the possibility of a truly intelligent device should be taken without discussion. Simply, on the theoretical level the debate is not closed. There is an interesting proposal in Gary Marcus's article (22): "Hybrids that allow us to connect the learning process of deep learning, with the explicit, semantic richness of symbols, could be transformative".

In short: on the one hand (as we have seen), we know little about the operation of these machines; on the other hand, there is no guarantee that such a functioning encloses "the semantic richness of symbols". Searle's central objection to the absence of semantics is not far.

The attempt to equate artificial intelligence and human intelligence has other fronts. Already at the end of the 19th century the great American thinker M S Peirce claimed that Abduction is a ubiquitous trait of Human's Thought. Then if AI devices are unable to *abductive reasoning*, thesis defended by A. J Larson (21), obviously they cannot be considered *intelligent* in the sense we consider we are. Nevertheless, *a priori* we can't exclude that the amazing progress in the field of computing could lead to some cases of abduction. Would this solve the general problem? Does abduction suppose that there is truly semantics? Once again the matter remains open.

The tendency to find something analogous to human intelligence behind all cases of behavior (whether animal or mechanical) supposes a kind of devaluation of forms of knowledge such as experience, of which both, animals and sophisticated mechanical devices, are indisputably capable. It is possible to have a great experience without even needing to have an idea of what is being experienced. Otherwise, the Platonic "Eidetic Field" would have to be extended to the mind of animals such as the ant or the bee that are so distant phylogenetically from us. In any case there is little reason for dogmatic positions. Good news for philosophy.

Intelligence and Kantian modalities of Judgement

If these machinic beings, which are playing increasingly evident role in all aspects of our lives, were actually comparable to us in terms of cognitive ability, then the potential disappearance of mankind would not imply the disappearance of that unique demonstration of human capacity that is science.⁸ The latter would simply have new and unexpected protagonists, who, after our hypothetical demise, would continue to bear witness to our fleeting presence, just as we are currently bearing witness to the fleeting presence of the Pyrenean ibex.

But even in the event that a machine could emulate the human capacity for learning and intellection in general, it might be argued that cognitive reason is only one of various ways in which we demonstrate intelligence. There are manifestations of intelligence in which the learning dimension is either non-existent or secondary:

What is learned, for example, when an entirely ethical requirement, meaning one that cannot be reduced to convenience, is imposed? On another issue, Calixto, the unfortunate protagonist of *La Celestina*, speaks spontaneously in such an unusual way that his young servants, Pármeno and Sempronio, believe it to have been influenced by the way Virgilio speaks. How does the way in which Virgilio speaks enrich the communicative aspect of the discourse? And to cite someone much closer to us, what does the sentence 'The Earth is as blue as an orange' (14) mean to the speaker?

As is so often the case, in order to understand the true scope of any achievement relating to a device that seems to demonstrate some form of intelligence, it is wise to position it at what we might consider the starting point, the ambitious project undertaken by British philosopher and mathematician Alan Turing (15). And assuming that the thinker's expectations were fulfilled, we would still have to ask ourselves whether we would find ourselves in the presence of a being capable of demonstrating intelligence, in the sense of the three forms encompassed in the Kantian critiques (16, 17, 18). Those in favour of equating artificial intelligence to human intelligence would have to demonstrate that the former was capable of functioning within this three-pronged framework. Furthermore, they would have to qualify the very difference within the Kantian distribution without projecting onto one of them criteria that are specific to the other.

Only if machinic beings (those constructed by man or the fruit of machinic entities themselves), besides being comparable to us in terms of scientific knowledge, were also comparable to us in terms of moral faculty, rule-based social organisation and creative capacity (pictorial, narrative and musical) would they be capable of emulating humans. This latter aspect constitutes perhaps the greatest challenge of all, since the 'faculty of judgement' that is then exercised may result in a judgement that is shared by rational beings without it being possible to sustain such an agreement on objective grounds. There is no general agreement on the irreducibility of the difference between human intelligence for the

⁸I understand *science* as the relationship with the natural environment that consists of trying to make it intelligible, a disposition that is antithetical to one that consists of using nature in a futile attempt to defeat it, meaning to forcibly push its boundaries. In this respect, it is worth remembering that *techne* can only go as far as to bring about what nature permits. It certainly cannot fundamentally change nature. Nature allows itself to be unveiled but not violated, and anyone who oversteps this mark will be immediately put in their place by nature itself.

purpose of artistic creation and human intelligence for the purpose of knowledge, but let us conclude by summarising the Kantian argument: the crux of the matter lies in the fact that, in the case of knowledge, the object legislates, the object gives reason or indeed eliminates it. When it comes to aesthetic perception, however, human faculties function in the same way as subjectivity (and sometimes intersubjectivity), for which there is no objective scale.

I remember an academic meeting in which a machine-produced pictorial composition was presented as a work of art and an artist who was present spoke up in outrage, condemning it as some sort of fraud. His response was, perhaps unwittingly, driven by a Kantian disposition; he suspected that the machine had applied criteria specific to cognitive reason (the very subject of the first Kantian critique) with the aim of producing something that related to the sense of beauty or repulsion (the subject of the third critique). It's as if a pianist believed that his technical mastery of the instrument (another issue dealt with within the framework of the first critique, since up to that point it is merely a matter of knowledge) was what makes him an artist.

Mankind: a negligible fraction?

It is a recurring theme in philosophical-scientific circles to wonder to what extent our own biological existence will be fundamentally changed by artificial implants that would bring us closer to those intelligent entities built from inert matter. In short, at the same time as this idea of the humanisation of machines is establishing itself, it seems to be pointing to the mechanisation of humans; and saying "it seems" would itself be somewhat inane if it were indeed possible that everything stored in the human brain could be transferred to a computer and vice-versa. It would be logical, then, to take another look at mankind and question ourselves about the human condition: is the rational being that man is necessarily *animal*, that is to say determined primarily by biology? Perhaps it would then be fair to move away from considering man as a particular kind of *animal* (a *rational* one, as opposed to animals that are not) and put forward his condition as a *rational* being that might occasionally (and only occasionally) have biological support.

And so (in one of the versions of what has come to be known as *transhumanism*) mankind would be a sort of transition towards something that, thanks to the capabilities that evolution has bestowed upon it, could escape some of the limitations that constitute our fundamental weakness. There is certainly no lack of incentive for those who support this way of outdoing humankind.

Indeed, in order to highlight just how recent mankind's existence on Earth has been and, consequently, its relative significance in terms of the evolution of the universe, scientific popularisation sometimes equates the various stages involved to a three-hour movie, whereby life on Earth would appear thirty minutes before the end, animals just five minutes and humans merely a fraction of a second before the end of the movie, playing so fleeting a role that it would be imperceptible to the viewer. We tell ourselves that we are a relatively new arrival to the universe, and if humankind was soon to die out, say, for example, by the year 3000, then our species will have been simply a fleeting moment in the natural process, meaning that, taking physical objectivity as a criterion, our entire presence will not have exceeded that split second of the movie mentioned above. A negligible fraction? Let's not be too hasty.

That imperceptible period that humankind occupies at the end of the movie has encompassed the emergence of technology, science, art, philosophy and a whole host of questions and answers about what is significant and what is not. These include the very question of whether the minute period of time for which mankind has existed reflects the significance that should be placed on that final period of the hypothetical movie in relation to the entire history of the universe.

After all, it is only in this fraction of a second that the being that ‘explains’, referring to principles that are accepted as evidence (on the basis of science), and also the being that ‘relates’ in a more generic sense, and in any case the being that resolves, delimits, demonstrates the absence of confusion and in doing so highlights the difference between the huge and the diminutive, between what tends to infinity and what is approaching the infinitesimal, among other things, appears. In that tiny fraction of a second a signmaker arrives on the scene, a being that gives meaning (sometimes multiple meanings in the same sign) and without whose actions everything would, of course, be meaningless.

There is no way of escaping this paradox, since the process that constitutes the universe (that is, the history of the transformation of energy) only appears very dilated because an ephemeral being, dumbfounded in the face of its environment, strives to instil in it some sort of order and to relate it, while at the same time continuing to give it meaning, a being who, like Borges’ Spinoza, “since his illness, since his birth [...] goes on constructing God with the word” (17).

On a physics course taught at Imperial College, British professor C. J. Isham (18) was able to link the words of Borges quoted at the beginning to those written by Arthur Eddington (19) following the shock brought about by quantum physics: “*We have found that where science has progressed the farthest, the mind has but regained from nature that which the mind has put into nature. We have found a stranger foot-print on the shores of the unknown. We have devised profound theories, one after another, to account for its origin. At last, we have succeeded in reconstructing the creature that made the foot-print. And Lo! it is our own*”.

References

- (1) BORGES, J.L. (1960) *El hacedor*, Buenos Aires, Emecé Editores,
- (2) BRUEGHEL, J. (Brueghel, el joven) “El paraíso terrenal”. Sevilla, Museo de Bellas Artes
- (3) FELDSTEIN, S. (2019) *The Global Expansion of AI Surveillance*, Washington Carnegie Endowment for International Peace (Downloaded at Carnegie Endowment.org)
- (4) HORNIK, K.; STINCHCOMBE, M; WHITE, H. (1989): “Multilayer Feedforward Networks are Universal Approximators” in *Neural Networks*. Vol 2 (pp.359-366). Oxford, Pergamon Press.
- (5) CYBENKO G (1989): “Approximation by superpositions of a sigmoid function”. *Mathematics of Control, Signals and Systems*, 2 (pp. 303-314)
- (6) VARIOUS AUTHORS: “The White House Brain initiative. Brain Research through Advancing Innovative Neurotechnologies” (1913) Online.
- (7) ALIVASATOS, A.; CHUN, M.; CHURCH, G.M; GREENSPAN, R.; YUSTE, R. (2012): “The Brain Activity Map Project and the Challenge of Functional Connections”. *Neuron* volume 74 June 21 2012.
- (8) JI. X; PAULSEN, B. D.; RIVNAY.J. (2021)... “Mimicking associative learning using an ion-trapping non-volatile synaptic organic electrochemical transistor,” *Nature Communications* 30 April 2021.
- (9) JUMPER, J.; EVANS R., HASSABIS D. (2021); “Highly accurate protein structure with Alpha Fold, *Nature* 596, 2021 (pp. 583-589).
- (10) ANFINSEN, Ch. B. (1972): “Studies on the principles that govern the folding of protein chains”. Nobel Lecture December 11, 1972. Online PDF.

- (11) NIELSEN M: *Neural Networks and Deep Learning*. Online Book p.206
- (12) SEARLE, J. (2010) "Minds, Brains and Programs" *The Behavioral and Brain Sciences* 3, 1980: 417-457. Published online by Cambridge University Press February 2010.
- (13) SEARLE, J (1989): "Replay to Jacqueline", *Philosophy and Phenomenological Research*, XLIX, (pp.701,708); SEARLE, J. (1990) "Is the Brain's Mind a Computer Program?", *Scientific American* 262, 1990 (pp. 26-31).
- (14) ELUARD,P.(1929), *L'amour la poésie*, Paris, NRF-Gallimard.
- (15) TURING, A. (1950) *Computing Machinery and Intelligence* Rev. Mind, vol. LIX, num. 236, October 1950 pp.: 433-460.
- (16) KANT, I. *Kritik der reinen Vernunft* (The critique of pure reason); *Kritik der praktischen Vernunft* (The Critique of Practical Reason) *Kritik der Urteilskraft* (The Critique of the Power of Judgement). Several translations.
- (17) BORGES, J.L (1981) "Spinoza", *La cifra*, Buenos Aires, Emecé.
- (18) ISHMAN CH.J.(1995) *Lectures on Quantum Theory* (pp.66) London, Imperial College Press.
- (19) EDDINGTON, A. (1920) *Space, Time and Gravitation* (pp.201), Cambridge University Press 1920.
- (20) GARY, M. (2022): "Deep Learning is Hitting a Wall" blog paper of March 10, 2022.
- (21) LARSON, E. J. (2021), *The Myth of Artificial Intelligence: Why Computers Can't Think the Way We Do*, Cambridge, The Belknap Press.