

What kind of “intelligence” is Artificial Intelligence?

¿Qué tipo de “inteligencia” es la Inteligencia Artificial?

PANIEL REYES CÁRDENAS¹

UPAEP Universidad, México / The University of Sheffield, UK
panielosberto.reyes@upaep.mx

ABSTRACT

In this article I carry a study of some comparative terms that aim to help narrowing down the meaning of “intelligence”. I proceed by looking back to the tradition that actually produced the term “intelligence” itself, acknowledging that it is a term that has been suitable for integrating the acts of a kind of agent, a mind that interprets and knows itself while knowing, and as the operations and functions of intelligent beings. With such a battery of meanings, a more integrated view of intelligence emerges, and hence sheds light about what kind of intelligence is the artificial one. This view is a cornerstone to understanding the problems of purely functional and psychological conceptions of intelligence, and hence we have been able to interpret to what extent we can properly use “artificial intelligence” in a proper sense.

Keywords: Philosophy of Artificial Intelligence, Emotions, Philosophy of Technology, Philosophy of Mind.

RESUMEN

En este artículo realizo un estudio de algunos términos comparativos que buscan acotar el significado del término “inteligencia”. Procedo revisando la tradición que produjo el término “inteligencia” mismo, reconociendo que es un término que ha sido utilizado propiamente para la integración de los actos de un cierto tipo de agente, una mente que interpreta y se conoce a sí misma mientras conoce las cosas. La inteligencia también emerge como el conjunto de operaciones y funciones que tienen los seres inteligentes. Con tales aspectos de la inteligencia se construye una visión más integrada y apropiada del término y así se adquiere una perspectiva para evaluar el tipo de inteligencia que esté en lo artificial. Esta perspectiva es un punto de partida necesario para entender los problemas de las visiones reduccionistas puramente psicologistas o funcionales de la inteligencia, y desde allí se juzga hasta dónde podemos interpretar a la “inteligencia artificial” como inteligencia en sentido propio.

Palabras clave: Filosofía de la inteligencia artificial, emociones, filosofía de la tecnología, filosofía de la mente.

¹ ORCID: <https://orcid.org/0000-0002-4430-0991>

A very important problem

The problem of defining what kind of “intelligence” is involved in the term “artificial intelligence” is very relevant. However, the problem itself is fruit of a very long tradition, attending that tradition can give us reflective clarity and help us to frame the philosophical problems of AI in a better way. This article will review different aspects of the concept of intelligence: we will start off by describing the senses in which the concept has been understood and its generalisations present in the inter-sectional and trans-disciplinary character of the uses of the term. It is my conviction that a better understanding will emerge from knowing the origins and theories of intelligence that were advanced by the great minds of the past, this will reveal both the limitations as well as the promising routes to undergo for better developments of the concept. Thus, we will be ready to account for the relevant complexities of the term “intelligence” in the present technological challenges and so better understand the key problems about what types of “artificial intelligence” there are and how we will pursue their understanding. In addition, it is also the case that our time has produced a renewed inquiry over the varieties of human intelligence and how AI can relate to these, it is also the case that there is a better consciousness in our time on the role that emotions play in our conscious and intellectual lives. In the last part of this paper we will recapitulate different aspects of intelligence to construct a more integrated view of AI: we aim for an enriched view that will include metaphysical aspects of intelligence, epistemological aspects of our knowledge of intelligent beings, as well as psychological traits. Following this path of inquiry, I propose to offer an account of intelligence that might be better suited for the necessary dialogue that our contemporary technological, theoretical and practical challenges demand.

The origin of the term “intelligence”

The English word “intelligence” comes from the Latin term “*intelligentia*”, which in turn it is a composite word of the roots “*intus*” and “*leggere*” and can be translated verbatim as “inner reading”. However, paying attention to the composite origin of the term we ought to follow St. Thomas Aquinas, who recognises not as much as “inner Reading” but as an inner act of “interpretation”. Every act of interpretation requires and interpreter, and such and interpreter is an intelligence that captures the sense or meaning of the interpretation. The doctrine of intelligence, that Aristotle introduced in his famous treatise “*De Anima*” (*On the Soul*) is examined by St. Thomas and incorporated to a metaphysical theory of intelligence and being. In his treatment he dis-

What kind “intelligence” is Artificial Intelligence?

tinguishes two senses of intelligence: a general aspect and a more specific one that concerns human intelligence. Broadly introduced, Thomas Aquinas tells us that the intellect in general can be conceived from the metaphysical theory of modes of being of act and potency, leading us to see that there are potential intellectual activities that belong to the “passive intellect” and intellectual activities that belong to the “intellect agent”. In his passive aspect the intellect contemplates the intellectual concepts by which we can think, while the intellect in act (intellect agency) carries mental operations while is conscious of them. In this way, the medieval theory of intelligence, particularly the Aristotelian stripe, is a theory that articulates a view of intelligence not only attending to its functional properties, but its contents too and, furthermore, includes also what is specific about a thinking being: consciousness. A thinking being thinks by being a conscious “agent” of thought.²

Going even further back in time, for thinkers as St. Augustine, who is one of the main sources for the understanding of the history of the concept of intelligence in the Middle Ages, Intelligence (*mens*) is a faculty that is above the simple ability to reason logically; it also includes the ability to remember, to think, to judge and to deliberate and decide meaningfully. This antecedent was important for St. Thomas and later Aristotelians like John Duns Scotus. St. Thomas is the first one that develops a theory of intelligence as the union of the two intellects explained above. For St. Thomas, as it has been stated, intelligence is an act of interpretation, he calls it “actu intelligibile”, which means, in the interpretation provided by Peter Geach³ that an act is effectively thought. Saint Thomas also not only understands intelligence as an activity (act) but as a *habitus*, he tells us that intelligence is a *habitus principiorum* (habit of principles), a continuous activity that happens according to principles. The relevance of these ideas is that they tell us what is the fundamental metaphysical nature of intelligence: intelligence is a mode of being, not only a set of operations and functions.

Early Medieval Islamic philosophers such as Ibn Sina (known in the Latin Medieval West as Avicenna) and Ibn Rushd (known in the Latin Medieval West as Averroes) used the concept of intelligence in order to make sense of the intelligible spheres of the natural world or “separated intelligences” as they found it in the Aristotelian tradition that they transmitted and interpreted. In doing their contribution to the topic of intelligence the Islamic philosophers explained intelligence as the ability to follow laws of thought that belong to a logic. The knowledge that an intelligence develops over a

² Cf. THEOL, S., I, q. 14, a 1, q. 55, a. 1, q. 79, a. 2; y Coment. In De An.; L. III, c. 2, lec. 2; c.4, lec. 7 y 9, c. 5, lec. 10; c. 7, lec. 12; c. 8, lec. 3, etc.

³ GEACH, P., *God and the Soul*, London: Routledge and Kegan Paul, 1969.

given topic depends on the logic in which that topic falls: an intelligence is a conscious activity that identifies the reasons of each thing in its own context and develops habits about such operations.

Having briefly reviewed the origins of the term we come to realise that some proper traits of intelligent lives are not only the functions and operations. In our own time and grounded on the above tradition, philosophers such as John Searle criticise contemporary theories of “Strong Artificial Intelligence” (henceforth S-AI) for being unduly reductive. Searle attacks some views of S-AI because they reduce intelligence to an ability of perform operations that forgets about the meaning and consciousness of them. In other words and following Searle’s reasoning, it is not only about describing the logical syntax of language, but to make sense of how can we have a consciousness of the semantics of language, i.e., an understanding of the “meaning” of the elements of language.

An American polymath and philosopher of the 19th Century, Charles Sanders Peirce, reminds us of thought as the using of signs; a sign can only be such if it is interpreted. Indeed, Peirce holds that “all thought is in signs” and, therefore, underlies the importance of understanding how signs work. Should we wanted to understand the nature of thought, we need to understand the nature of signs. Peirce formed a rigorous science of signs to give a proper account of thought, he called this science “semeiotics”. The act of interpretation is an essential aspect of the sign-relation: each sign needs interpretation in order to be a sign of something. There are three basic kinds of signs from the view point of the interpretant (the way of interpretation that we might have of it): icons, indices, and symbols. Each kind of sign involves three characteristics:

1. The sign as such (as for example, a traffic sign)
2. That of what the sign is about, i.e., the object of the sign (for example, a speed limit, road works ahead, etc.)
3. That by which the sign is interpreted as a sign of its object, the “interpretant” of the sign (e.g., the actions and conducts that you might expect for a driver after seeing the sign).

We could summarise the above by stating that “x is a sign of y for z”, or “z interprets x as a sign of y”, or “y is indicated with z by x”, and the like. According to Peirce, even the interpretant itself (z) is meant to become a further sign. As such, the interpretant must be interpreted by a further interpretant that is wider in scope, which in turn is interpreted for another... and so on. Therefore, there is no limit to the process of interpretation. Inasmuch as there is an intelligence, an intelligence will be so manifested as a capacity of interpreting signs, even the sign of its own reality. What Peirce calls “the

final interpretant” is an idea projected in a distant future, the direction in which the convergence of interpreting minds will approach in their converging interpretations.

Since Peirce holds that interpretants are themselves signs and, hence, they must be interpreted, he thinks that the interpretation of signs must be a common effort of collaboration that is extended and oriented to the future. To give sense to the manifold of our experience is, for Peirce, a collective effort that involves generations of communities of interpretation. Intelligence is, therefore, at least to a minimal sense, the ability to interpret signs in a self-conscious and self-controlled way, not only performing some functional operations.

Modern and contemporary conceptions of intelligence

The concept at play, namely “intelligence”, is usually compared with similar or related concepts such as understanding, intellect, reason. Ferrater Mora rightly points out that in our time there is a pervasive use of the term “intelligence” as a psychological function: “[...] as denoting certain ‘faculty’, certain ‘intellectual function’”.⁴ In such a case, intelligence can be interpreted as a function or faculty of the rational thought, as “[...] a capacity possessed by some organisms to adapt to new situations using for such a purpose the knowledge acquired in the use of former processes of adaptation”.⁵

Thus, there are many famous experiments in which it is shown how animals show and exhibit this kind of intelligence in the sense of being a psychological function of adaptation. In this sense of intelligence, W. Köhler’s experiments on chimps and their behaviour are presented to the reader as exploring “chimp’s intelligence”, but the use of the concept of intelligence is here totally reduced to processes of adaptation and reflective control exhibited by such primates.

There is certainly no doubt that the contributions to psychology are valuable to contribute to an enrichment on the understanding of what kind of behaviours intelligent beings exhibit, and these are important to qualify distinct intelligent behaviours. However, these ways of talking about intelligence might hamper rather than help an inquiry into what is the meaning of intelligence, since already take for granted that intelligence is adaptation, such kind of presupposition is a non-justified reductionism. A further problem

⁴ FERRATER MORA, J., *Diccionario de filosofía*, Madrid: Alianza, 1990, p. 1873.

⁵ FERRATER MORA, J., *Diccionario de filosofía*, p. 1873.

with these kind of viewpoints is the total absence of fundamental metaphysical and epistemological aspects of intelligence: a mind that interprets has to be a very particular kind of being, not only one that functionally responds to some operative expectations. This kind of reductionism about the concept of intelligence has been the main obstacle to overcome some views of intelligence that exaggerate the value of some technological achievements in the field of Artificial intelligence, as we will show below.

Intelligence and technology

The concept of “Artificial Intelligence” (henceforth AI) emerged as opposed to the concept of “natural intelligence”. Let us remember that here I have been talking about the usual functional sense that the term “intelligence” has as a connotation in the behavioural sciences. In this kind of understanding, living beings will exhibit certain functions that are acknowledged as intelligent. Ferrater Mora tells us in his famous dictionary of philosophy that this concept of AI is understood usually as “[...] a series of operations in the course of which calculations are carried, problems (mathematical and otherwise) are solved, games are played (e.g., a Chess game), learning is happening, distinctions between diverse patterns and forms are established, a language is translated, etc.”⁶

Hence AI is more frequently understood as the ability to programme an artificial system in order to perform some of all of the tasks mentioned above. However, even as of today, machines function with a kind of programming that operates under kinds of logics that can only reflect a very specific interpretation of information, what has been called “crisp judgements”. No doubt technology has made great advances at an accelerated rate, and even in the use of programming that does not necessarily depend on the logic mentioned above, but trying fuzzy logics, multi-valued logics, even quantum logic that has helped to relativise the problem of the “crisp judgements” of classic logics. But these advancements, though very impressive, still do not mean that the systems that perform these highly sophisticated tasks are conscious.

According to a number of thinkers that follow the ideas of Alan Turing, if a machine could develop in such a way that its operations and the linguistic expression of these were indistinguishable from the ones of a human being that is behaving rationally then this will entail the judgement that this AI is equivalent to the human intelligence. This is the origin of the so-called “Tur-

⁶ FERRATER MORA, J., *Diccionario de filosofía*, 1875.

ing test”. The test has produced a considerable confidence that conversations and operations that were thought only in the scope of intelligent human beings now are very satisfyingly performed by artificial systems.

However, American philosopher of ordinary language John Searle has identified a serious problem with the conception of AI that follows from Turing’s ideas. Indeed, Searle articulated a famous thought-experiment known as the “Chinese room” experiment that can be formulated as follows: The argument of the Chinese room pictures a digital computer that executes a programme without having anything like a “mind”, “understanding”, “consciousness” and more importantly “intelligence” in the strong sense explained above and still passing Turing’s test. No matter how close to the behaviour of an intelligent human being a computer is, still does not mean that the computer “knows” what it is doing. Searle’s argument appeared for the first time in the journal *Behavioural and Brain Sciences* in 1980 in the article “Minds, Brains, and Programs”.

The argument is clearly addressed against the philosophical views that understand intelligence through a computational or functional theory (i.e., they already presuppose a reductionist approach to intelligence). Most specifically the argument presupposes two interpretations of the concept of AI: a strong interpretation and a weak interpretation. The argument attacks the strong view of AI, in Searle’s words: “a computer appropriately programmed with correct inputs and outputs of information would have in virtue of that a mind in the exact sense that the human beings have minds”.⁷

Searle’s thought experiment begins with this hypothetical premise: suppose that some research on AI has been successful in construction a computer that behaves as if could understand the Chinese language (let us say, Cantonese). The programme assumes the Chinese characters and symbols in the input slot and, following the instructions of a computational programme that describes “if you get x then answer y”, gives back an output of information. Let us suppose that in the same way this computer carries its tasks so convincingly that without a problem outstandingly passes the Turing test, consequently convincing the Cantonese speakers that the program is a living speaker of Cantonese. Thus, to every relevant question that a person does, the output slot offers appropriate responses, in such a way that a proficient Chinese speaker of Cantonese is convinced that inside the room there is another human being that actually understands the language. Finally, then, the question that Searle wants to respond is this: “does the machine actually ‘understands’ Cantonese? Or, is it only simulating the ability to understand Cantonese that is assumed

⁷ SEARLE, J., “Minds, Brains and Program”, *Behavioral and Brain Sciences*, núm. 3, 198, p. 1.

when a person receives understandable responses? To the first view Searle calls "Strong AI" and to the second view he dubbed: "Weak AI".

Searle then, supposes that he himself is the one that is in the closed room and has a book with an English version of the interactive, he also has papers, pencils, erasers and a lot of information in relevant shelves. Searle could be receiving Cantonese characters through the slot in the door, he could process such characters with the book of instructions, and producing new pieces of paper as answers. If the computer has passed the test of Turing in this way it will necessarily follow that Searle did too, but they are only following the manual of instructions, neither Searle nor the computer understand what the whole thing means.

Searle argues that without "understanding" (or "intentionality") we cannot describe the activities of a machine as "intelligent" and, if there is not thought in the sense in which we think when we understand, we cannot then ascribe a mind to the machine in the usual sense of the term. In consequence, Searle concludes that the "Strong AI" is, so far, an invalid conclusion to draw from the Turing test. However, even if Searle's overall argument is perfectly sound (as it really shows that there is a missing key aspect of intelligence, our consciousness of experience), many computer scientists might claim that the colossal amount of computational networks has grown exponentially on the capacity of predicting human behaviour, concluding that computers know ourselves better than us. But prediction is only a statistical aspect of information, however massive the amount of information could be recombined, this still does not mean consciousness. This keeps being a qualitatively syntactic and functional aspect of the operations, and it does not become what David Chalmers has called the "hard" fact of consciousness: consciousness does not simply emerge from a great syntactical web, it is needed to navigate willingly in the network of meanings and experiences that the web might represent: it is the interpretant missing in the sheer representation.

Intelligence and emotions

A significant advantage in the understanding of intelligence is the ever more growing conviction that we need to overcome the non-justified opposition between a rational life and emotions, as though intelligence in order to function properly needs to nullify emotions. From the field of psychology,

authors such as Daniel Goleman⁸ have vindicated the Intelligent character of the integration of emotions and the psychological life of the mind. However, we can also find positive developments from philosophy on the integration of the value of emotions to the intellectual life: thus, just to invoke an appealing example, there are interesting texts such as Brady.⁹ Unfortunately, these important developments have been struggling with a prejudiced view that revokes emotions as fundamentally irrational. Contemporary philosophy of emotions, as opposed to the negative assessment of emotions, values a positive role in moral emotions in the formation of a mature and autonomous moral character. Philosophy of emotion has even discovered that the sheer desire of knowledge and the unrestricted search for truth is endorsed by an intellectual emotion: even the classical philosophy of Plato and Aristotle is in tune with this belief that intelligence starts with the desire for truth and knowledge. If we examine our intellectual lives we discover that our consciousness and awareness of all its elements is what makes us intelligent. The irreducible “being aware” includes the positive value of trusting our awareness of emotions. In other words: a fuller sense of awareness and consciousness emerges from the integration of our intelligence of our emotional states. In Hegel’s philosophy, for instance, particularly in his *Phenomenology of the Spirit*¹⁰ he introduces us to a gradual growth of consciousness that allows us the understanding of our experience and the intelligence of it not as a monolithic or uniform phenomenon, but as a mosaic of man levels of consciousness of different realities that are unified by intelligence: including the manyfold of our emotional experience. Thus, far from expelling emotions from our consciousness, we become more capable of evaluate our life as intelligent if we take on account the value of our emotions and a reflective equilibrium of our consciousness of these. This clearly goes far beyond a purely functional aspect of intelligence.

The need to recover a richer conception of intelligence

In this article I have introduced some comparative terms that aim to help narrowing down the meaning of “intelligence”. I have looked back to the tradition that actually produced the term itself, acknowledging that it is a term that has been suitable for integrating the acts of a kind of agent, a mind that interprets and knows itself while knowing, and as the operations and func-

⁸ GOLEMAN, D., *Emotional Intelligence: Why It Can Matter More Than IQ*, New York: Bantam Books, 1995; GOLEMAN, DANIEL, *Working with Emotional Intelligence*, New York: Bantam Books, 1998.

⁹ BRADY, MICHAEL, *Emotion: the basics*, London: Routledge, 2019.

¹⁰ HEGEL, G. W. F., *Phenomenology of Spirit*, translated by Terry Pinkard, Cambridge: Cambridge University Press, 2017.

tions of intelligent beings. With such a battery of meanings, a more integrated view of intelligence emerges. This view is a cornerstone to understanding the problems of purely functional and psychological conceptions of intelligence, and hence we have been able to interpret to what extent we can properly use “artificial intelligence” in a proper sense. In addition, we have also considered that there is a growing realisation of the emotional component of intelligence, such a component is definitely not a purely functional aspect of intelligence and therefore questions the functional reductionism. Considering the emotional aspect of intelligence has a liberating effect: it shows that a conscious life goes beyond consciousness of thought: it is also consciousness of feelings and the awareness of realities that are meaningful to us. An issue that still remains open here, however, is the explanation of how the intelligent life includes the consciousness of emotional experience as well as the moral experience and the aesthetic experience. This has to be emphasised in future explorations to avoid reductionists accounts of intelligence, but not only that, we also need to open the discussion to a greater deepening on the inner richness of human intelligence. Artificial intelligence, thus, appears in its correct hindsight: an artificial reconstruction of a very specific aspect of intelligence, but it will be wrong to mistake one for another.

References

- BRADY, MICHAEL, *Emotion: the basics*, London: Routledge, 2019
- FERRATER MORA, JOSÉ, “Inteligencia”, “Inteligencia Artificial”, “Intelligentia”, en FERRATER MORA, JOSÉ, *Diccionario de filosofía*, Madrid: Alianza, 1990.
- GEACH, PETER, *God and the Soul*, London: Routledge and Kegan Paul, 1969.
- GOLEMAN, DANIEL, *Emotional Intelligence: Why It Can Matter More Than IQ*, New York: Bantam Books, 1995.
- GOLEMAN, DANIEL, *Working with Emotional Intelligence*, New York: Bantam Books, 1998.
- HEGEL, G. W. F., *Phenomenology of Spirit*, translated by Terry Pinkard, Cambridge: Cambridge University Press, 2017.
- PEIRCE, CHARLES S.; HARTSHORNE, CHARLES; WEISS, PAUL and BURKS, ARTHUR W., *Collected papers of Charles Sanders Peirce*, Cambridge, Mass: Belknap Press of Harvard University Press, 1965.
- SEARLE, JOHN, “Minds, Brains and Program”, *Behavioral and Brain Sciences*, núm. 3, 198, pp. 417-57.