

Does AI think?

RICARDO F. CRESPO

Universidad Austral

Consejo Nacional de Investigaciones Científicas y Técnicas

rcrespo@iae.edu.ar

ORCID: 0000-0002-7555-5479.

Abstract. This article raises a very relevant question today: “Does AI think?” To come up with an answer, it explores several conceptions of knowledge, from classical Aristotelian notions to modern philosophers’ takes on it. For Aristotle, knowledge starts with the senses, and, through epagoge and nous, it grows into an understanding of the nature of the thing known, which is something that does not happen with AI. Modern schools of thoughts are either rationalist or empiricist. For rationalists, knowledge is innate, while empiricists believe all knowledge comes from the senses. As Kant views it, knowledge is “built” on sensorial data with the use of theoretical reason categories. Based on these approaches, it is harder to claim that AI does not think by itself. The article also reviews current philosophers’ views on AI and AI thinking. Summing up, the answer to the initial question depends on the concept of knowledge adopted. As this article shares the classical notion of thinking, it concludes that AI does not think.

Keywords: Aristotle on Knowledge, Modern and contemporary visions of knowledge, Philosophers expert on AI notion of knowledge.

The title of this brief paper is inspired in an old article of Mario Bunge, entitled “Do Computers Think?” ([1956] 1959).¹ It maintains that concerning the AI there are no new reasons to reject Bunge’s arguments for his negative answer. The root of Bunge’s argument is that “Unlike nature, and unlike artifacts, man is able to build a mental level upon the physiological one” (Bunge [1956] 1959, 131). One can agree or not about where this metal level come from, but the center of the affirmation is that man has a mental level, not physical, something that artifacts lack. This mental level is the root of consciousness, awareness and the abstracting capacity (Bunge [1956] 1959, 137, 138, 141). Instead, “computing machines just perform certain operations without being aware of it: they do not know what they are doing nor even *that* they are doing anything –and this, simply because machines have no consciousness” (Bunge [1956] 1959, 136).

But what is to think or to know? Knowledge is a result of perceiving and thinking. It is something that is over the mere record of data, and the logical processes that can be made with them. What is the plus of thinking? It is something that has been discerned from the old times of classical Greek philosophy. For Aristotle, the essence of knowing is capturing the essence of things and causes of events that are impressed in our souls thanks to a determined process. However, this conception of knowledge has not been universally agreed, even in the Ancient and Medieval ages, and has changed in the Modern and Contemporary times.

Shane Legg and Marcus Hutter (2007) contains “A Collection of [70] Definitions of Intelligence”. They classified in “collective” (proposed by groups, and organizations including dictionaries) definitions, psychologist definitions and AI researcher definitions. In the 18 definitions of the first group the words knowledge, abstract knowledge, understanding, comprehend, thinking appears in almost all. The 35 psychologist definitions refer to the function of intelligence, though the expression “ab-

¹ It is also the initial question of Alan Turing 1950 article: “Can machines think?” (Turing 1950, 433).

stract thinking” also appears. In the 18 definitions of AI researchers the words thinking, knowledge, understanding completely disappear.

The first section of this paper will explain the Aristotelian conception of knowledge. The second section will pick up the ideas of contemporary thinkers about AI on the presence or not of this conception of thinking and knowing in AI. The third section will describe the Modern and contemporary vision of knowledge, showing that it is closer than the former to the possibility of AI to think.

1. Aristotle on Knowledge

Aristotle describes what he calls “dianoethic” virtues—that is, intellectual abilities—in Book VI of his *Nicomachean Ethics*. These virtues include technique (*téchne*), science (*epistéme*), practical wisdom (*phrónesis*), intuitive reason (*nous*), and philosophic wisdom (*sophia*), which embodies the union of intuitive reason and science (*Nicomachean Ethics* VI, 3). Technical and practical wisdom are dispositions towards action. All theoretical knowledge, Aristotle notes referring to science, “proceeds sometimes through induction (*epagogés*) and sometimes by syllogism” (*Nicomachean Ethics* VI, 3, 1139b 38–39).

Science, a deductive procedure, obeys the logical laws and performs its task rather automatically given the principles of science. The discovering of the latter involucrate *epagogé* and *nous*, which are the most “human” steps of thinking. There are relevant passages about the role of both in *Posterior Analytics* II, 19. In this section, Aristotle describes the cognitive steps leading to the knowledge of the first principles. He states that we do not possess these principles, but a previous capacity called perception, shared by animals, which is the first step towards them (99b 33–35). Then,

from perception there comes memory, as we call it, and from memory (when it occurs often in connection with the same item) experience; for memories which are many in number form a single experience. And from experience, or from all the universal which has come to rest in the soul [...], there comes

a principle of skill or understanding [episteme] –of skill if it deals with how things come about, or understanding if it deals with how things are (100a 3–9).

This is *epagogé*, the process.² However, what is the state that grasp the principle of skill or understanding? There is a “jump” between experience and grasping the principles. It is performed by *nous*, translated as “intuition” by the commentators and by “comprehension” by Jonathan Barnes (Barnes 1993, 267). The following passage explains:

Of the intellectual states by which we grasp truth some are always true and some admit falsehood (e.g. opinion and reasoning – whereas understanding [episteme] and comprehension [nous] are always true), and no kind other than comprehension is more certain than understanding, and the principles of demonstrations are more familiar, and all understanding involves an account – there will not be understanding of the principles; and since it is not possible for anything to be truer than understanding, except comprehension, there will be comprehension of the principles – both if we inquire from these facts and because demonstration is not a principle of demonstration so that understanding is not of understanding either – so if we have no other true kind apart from understanding, comprehension will be the principle of understanding. And the principle will be of the principle, and as a whole will be similarly related to the whole object (100b 6–17).

Comprehension (*nous*), thanks to induction (*epagogé*) grasps the principles from which understanding (*episteme*) demonstrates.

Thus, theoretical reason relies on empirical data concerning physical events, but it goes beyond them. When Aristotle begins his *Metaphysics* by saying that “all men by nature desire to know” (I, 1, 980a 21), he uses the term *eidenai*—the same term that he uses to mean the knowledge of principles (*Nicomachean Ethics* VI, 7 1141a 17) and that stems from the verb *eido*, to see: it means seeing. As already explained, this knowledge is not innate; it starts with the senses, the memory, and the experience, which ultimately facilitate a noetic or intuitive—not deductive—grasping

² See C. D. C. Reeve 2006 and 1995, 56ff. for a complete description of this process.

or comprehension of those first principles, essences, and causes, by *nous*. Induction is part of this process, but Aristotle's abstractive induction (*epagogé*), should not be confused with the modern concept of enumerative induction. Hintikka (1992, 34) explains the difference, "For Aristotle, the problem of induction was not first and foremost a problem of inference from particulars to a generalization. It was a problem of concept formation. Particular cases were stepping-stones to the concepts or forms 'induced' to be realized in the soul [...] Hence there is no such problem as the justification of induction for Aristotle." For Aristotle, induction is "a process of inducing in ourselves the right concepts" (Hintikka 1980, 429).

These concepts or forms are already general notions; then, we do not need to gather a complete number of instances to induce the general concept. It may be said that, for Aristotle, the theoretical truth features two levels. There is a pre-propositional level of (a richer or poorer) knowledge of essences (*Metaphysics* IX, 10, 1051b 17 – 1052a 4) and a propositional level of principles and judgments (*Metaphysics* IV, 7, 1011b 25–27 and VI, 4, 1027b 20–23). Aristotle views the truth not only as a logical category associated with judgments and propositions; the truth refers to reality not only as an actual correspondence to the facts of the composition or division of terms expressed by judgments, but also as the ability to grasp the actual beings that these terms designate, that become present in the souls. "Actual knowledge is identical with its objects" (*De Anima* III, 7, 431a 1): knowledge is an "intentional" possession of the form of what is known, not a representation of it (Hintikka 2004, 46). He states, "to think of x is to have the form of x in one's mind" (Hintikka 1980, 429). This form is an universal concept abstracted from the particulars thanks to the work of *epagogé* and *nous*.

2. Philosophers experts on AI knowledge

One can agree or not with this Aristotelian explanation of thinking and knowing, but what has been agreed by most people, including philosophers' experts on AI, is that the characteristic of human thinking is that it abstracts from particulars the essences and causes. By abstraction hu-

man beings captures or grasp the meaning of concepts, have a semantic power that surpasses the mere syntactic order, and they are aware of this power. This characteristic differentiates the human being from the AI.

Margaret Boden states: “today’s computers don’t grasp the meaning of what they “read” or “say”” (Boden 2016, 39). She puts an example, “the Google search engine, for instance, searches terms weighted by relevance –which is assessed statistically, not semantically (that is, *without* understanding)” (Boden 2016, 63). The explanation she offers is that life is necessary for mind, and the metabolism necessary for life cannot be instantiated by AI (Boden 2016, 144–5).

For Melanie Mitchell (2019) the computer “has no knowledge of the meaning of these [human interpretable concepts] symbols” (Mitchell 2019, 23) and does not need them. Human beings, she argues, have two fundamental capacities that originate knowledge, abstraction and analogy (Mitchell 2019, 242ff.), capacities that AI lacks (Mitchell 2019, 246). AI also lacks “metacognition”, “the ability to perceive and reflect on one’s own thinking” (Mitchell 2019, 260). A program may generate texts, essays, music, etc., but does not have real understanding of what has generated (Mitchell 2019, 274).

For Jobst Landgrebe and Barry Smith, the AI is a physical, not mental reality (2023, 11). The human mind, they argue, cannot be mathematically modelled (Landgrebe and Smith 2023, 13). Human beings have goals, desires, intentions that AI does not have (Landgrebe and Smith 2023, 39). Thinking is reasoning, planning, abstracting (Landgrebe and Smith 2023, 41), “the ability to conceive, and then deliberately plan and build” (Landgrebe and Smith 2023, 45). AI cannot have intentions (Landgrebe and Smith 2023, 227), experiences (2023, 233), and semantics and pragmatics (2023, 241). As Michael Polanyi explains, “the semantic operations attached to a formal system are functions of the mind which understands and correctly operates the system” (Polanyi 1952, 313).

For Stanley Jaki, only the human being is capable of abstracting and capturing universal concepts (Jaki 1969, 228). Instead, the machine, a formal system, cannot produce a truth (Jaki 1969, 216).

Even Alan Turing has recognized that “One may ‘play about’ with a machine and get the desired result, but not knowing the reason” (Blum

2010, 45). He feels “that my mind as I know it cannot be compared to a machine” (Blum 2010, 51).

Hubert and Stuart Dreyfus assert: “Computers are certainly more precise and more predictable than we, but precision and predictability are not what human intelligence is about” (Dreyfus 1986, xiv). For them, intelligence is the ability to recognize, to synthesize, to intuit. They state, “Each of us has, and uses every day, a power on intuitive intelligence that enables us to understand, to speak, and to cope skillfully with our everyday environment [...] Analysis and intuition work together in the human mind” (Dreyfus 1986, xiv). Intuition that leads to understanding is what AI lacks. Besides, it does not need it to develop its function.

In sum, AI enlarges our capacities and can make tasks that we human are not capable of making. However, this does not mean that AI thinks.

3. Modern and contemporary visions of knowledge

Modern and contemporary thought is very rich and cannot be described in a section of a paper. Therefore, the paper will only deal with some general characteristics of it.³

Two are the main themes of the Philosophy since the end of the medieval times: the anthropological condition of the human being and the nature and source of knowledge. Concerning the last theme there were initially two main currents, rationalism and empiricism. Then, Immanuel Kant proposed a theory taking elements from both currents. Aristotle is almost completely absent from philosophical thinking during all this period.

In the XVI Century Michel de Montaigne recovered old skeptical ideas of Sextus Empiricus. René Descartes answer him with his rationalist theory basing the possibility of knowledge in his own self-conscience and the idea of God. The source of knowledge for Descartes and posterior rationalists is reason and they sustain that we have innate ideas. For Thomas Hobbes, instead, the source of knowledge are senses. However,

³ I consulted Anthony Kenny 2006 and 2007, and Etienne Gilson and Thomas Langan 1963 for the first part of this section.

for him they are passive, they do not capture something external, but the occurrence of an image or fancy in the mind. In this he finishes likening Descartes. For John Locke there are not innate ideas, everything we know comes from experience. The mind is like a blank piece of paper, and it has ideas only by experience. There are two kinds of experience: sensation and reflection. Sensation is sense perception of the qualities of external objects: red, cold, hot, sweet, and other sensible qualities. Reflection is the perception of the internal operations of our minds. For him there are two kinds of ideas: simple and complex. All simple ideas come from experience, and complex ideas are a combination of simple ideas. In sum, ultimately all our ideas come from experience.

David Hume distinguishes between two kinds of perceptions, impressions and ideas. Impressions are the direct, vivid, products of immediate experience; ideas are merely frail copies of these original impressions. For Hume, we cannot rely on causal reasoning to convince us that there are external objects since such reasoning merely arises from our observation of a constant conjunction between causes and effects. In fact, Hume supposed, our belief in the reality of an external world is entirely non-rational. Although it is unjustifiable, belief in the external world is natural and unavoidable.

Kant manages to overcome the previous debate between rationalism and empiricism by stating that knowledge is part of experience, but not all knowledge comes from it. Thus, it speaks of two sources of knowledge, sensitivity, which provides the data of experience, and understanding, which gives shape to all the information coming from the senses, and which is independent of experience.

Kant thinks that Aristotle's metaphysics exceeds the possibilities of human reason. For Kant, theoretical reason cannot penetrate the nature of things, which he calls *noumenon*—that is, “a thing which must be cogitated not as an object of sense, but as a thing in itself [*Ding an sich*]” (Kant [1787] 1999, 362). As he puts it, “We have not insight into the possibility of such *noumena* and the domain outside of the sphere of appearances is empty (for us)” (Kant [1787] 1999, 362). We cannot grasp *noumena* with our theoretical reason.

Jens Kohne (2014) attributes a nominalist position to Immanuel Kant. More than nominalist Kant's position about knowledge has been labelled by some scholars as "conceptualist".⁴ In fact, Kohne speaks of conceptualization and concepts in the next quotations (my cursives). Kohne's characterization, regardless the correctness or not of labelling Kant as nominalist fits with the condition of AI. He states:

Instead of quitting the ontological enterprise now, Kant changed its purpose. In place of describing the nature of being, a characterization of the conceptual framework which constitutes reality is needed now. That means ontology is no longer interested in a direct identification and characterization of reality via categories but in a *conceptualization* of how reality is represented in my mind or rather through my mind, inasmuch as the difference between the Aristotelian ontology and the idealistic one is between two levels of considering reality (Kohne 2014, 87).

The influence of Kant has been impressive. Kant has broken the confidence in our capacity of knowing the world. From Kant to our days, philosophy has been mainly constructionist: to know is to build "reality" by the knower.

Where AI fits in the previous map? AI collects data and transforms them in signs that it combines, without understanding them. It has a nominalist foot in the sense that it stops in the sign, without intending a semantic understanding of it, that it does not need to accomplish its function. It has also an empiricist foot in the sense that all its "knowledge" comes from data and the combination and working on them.

Conclusion

Does AI think? As stated in the Introduction, it all depends on our notion of thinking and knowledge. If we believe that we are able to access reality and we can therefore—though imperfectly—know the essence and causes

⁴ See the complete article of Colin McLear (2020) about the conceptualist or non conceptualist position of Kant. Michael Oberst (2015) explains the kind of nominalism that can be attributed to Kant.

of events, we will conclude that AI cannot think, as noted by the AI experts quoted in the second section. If, instead, we believe that thinking is just turning data into signs and processing them, this is precisely what AI does. I uphold the former notion, along with Aristotle, Bunge, and the AI experts cited here.

Why? Because we all have the experience of thinking, hinting at the nature of things without the need to recollect a huge number of observations of those things. As Javier Sánchez Cañizares (2022) states, as a result of the freedom enabled by intellectual knowledge, human beings have a great capacity for growth. This is illustrated by the fact that, with our intellect, we can design artefacts to achieve specific ends. This is something that artefacts, including AI, cannot do. Only humans, like all natural beings, have intrinsic ends and can generate new ends.

References

- Barnes, Jonathan. 1993. *Aristotle Posterior Analytics*, translated with a commentary by Jonathan Barnes. Oxford: Clarendon.
- Blum, Paul Richard. 2010. "Michael Polanyi: Can the Mind be Represented by a Machine? Documents of the Discussion in 1949." *Polanyiana* 201(1–2): 35–60.
- Boden, Margaret. 2016. *AI, its nature and future*, Oxford: Oxford University Press.
- Bunge, Mario. [1956] 1959. "Do Computers Think?" *British Journal of the Philosophy of Science* 7(26): 139–148, reprinted in Mario Bunge (1959) *Metascientific Queries*, 124–152, Springfield, Ill.: Charles C. Thomas.
- Dreyfus, Hubert L. and Stuart E. 1986. *Mind over Machine. The Power of Human Intuition and Expertise in the Era of Computer*. New York: The Free Press.
- Gilson, E. and T. Langan. 1963. *A History of Philosophy, 3, Modern Philosophy. Descartes to Kant*. New York: Random House.
- Hintikka, Jaakko. 1980. "Aristotelian Induction." *Revue Internationale de Philosophie* 34(133–134): 422–439.
- Hintikka, Jaakko. 1992. "The concept of induction in the light of the interrogative approach to inquiry." In J. Earman (ed.) *Inference, explanation, and other frustrations: Essays in the philosophy of science*. Berkeley, CA: University of California Press.

- Hintikka, Jaakko. 2004. *Analyses of Aristotle*. Selected Papers 6. Hingham: Kluwer.
- Jaki, Stanley L. 1969. *Brain, Mind and Computers*. New York: Herder and Herder.
- Kant; Immanuel. [1787] 1999. *Critique of Pure Reason*, Translated and Edited by Paul Guyer and Allen Wood. Cambridge: Cambridge University Press.
- Kenny, Anthony. 2006. *The Rise of Modern Philosophy*. Oxford: Clarendon Press.
- Kenny, Anthony. 2007. *Philosophy in the Modern World*. Oxford: Clarendon Press.
- Kohne, Jens. 2014. "Ontology, its Origins and its Meaning in Information Science." In *Philosophy, Computing and Information science*, edited by Ruth Hagengruber, and Uwe Riss, 85–90. London: Pickering and Chatto. DOI: <https://doi.org/10.4324/9781315653938>
- Landgrebe, Jobst and Barry Smith. 2023. *Why Machines Will Never Rule the World*. New York and London: Routledge.
- Legg, Shane and Marcus Hutter. 2007. "A Collection of Definitions of Intelligence", IDSIA-07-07 Technical Report. <https://arxiv.org/abs/0706.3639>, Accessed 22 February 2024.
- McLear, C. 2020. "Kantian Conceptualism/Nonconceptualism." In *The Stanford Encyclopedia of Philosophy* (Fall 2021 Edition), edited by Edward N. Zalta. <https://plato.stanford.edu/archives/fall2021/entries/kant-conceptualism/>.
- Oberst, Michael. 2015. "Kant on Universals." *History of Philosophy Quarterly* 32(4): 335–352.
- Polanyi, Michael. 1952. "The Hypothesis of Cybernetics." *The British Journal for the Philosophy of Science* 2(8): 312–315. <https://www.journals.uchicago.edu/doi/10.1093/bjps/II.8.312>.
- Reeve, C.D.C. 2006. "Aristotle on the Virtues of Thought." In *The Blackwell Guide to Aristotle's Nicomachean Ethics*, edited by Richard Kraut, 199–217, Oxford: Oxford University Press.
- Reeve, C.D.C. 1995. *Practices of Reason. Aristotle's Nicomachean Ethics*. Oxford: Clarendon Press.
- Sanchez Cañizares, J. 2022. "Posibilidades y límites de la inteligencia artificial." *Toledana* 46(1): 233–250.
- Turing, Alan M. 1950. "Computing Machinery and Intelligence", *Mind* NS 59(236): 433–460. <http://dx.doi.org/10.1093/mind/LIX.236.433>.