

Representation, Analytic Pragmatism and AI

Raffaella Giovagnoli¹

Abstract: Our contribution aims at individuating a valid philosophical strategy for a fruitful confrontation between human and artificial representation. The ground for this theoretical option resides in the necessity to find a solution that overcomes, on the one side, strong AI (i.e. Haugeland) and, on the other side, the view that rules out AI as explanation of human capacities (i.e. Dreyfus). We try to argue for Analytic Pragmatism (AP) as a valid strategy to present arguments for a form of weak AI and to explain a notion of representation common to human and artificial agents.

1. Representation in AI

The notion of “representation” is at the basis of a lively debate that crosses philosophy and artificial intelligence. This is because the comparison starts from the analysis of “mental representations”. First, we move by adopting a fruitful distinction between the “symbolic” and the “connectionist” paradigms in AI [1]. This distinction is useful to highlight two different ways of explaining the notion of representation in AI.

An important challenge for AI is to simulate not only the “phonemic” and “syntactic” aspects of mental representation but also the “semantic” aspect. Traditionally, philosophers use the notion of “intentionality” to describe the representational nature of mental states namely intentional states are those that “represent” something, because mind is directed toward objects. The challenge for AI is therefore to approximate to human representations i.e. to the semantic content of human mental states. If we think that representation means to connect a symbol to the object of representation we focus on the discreteness of mental representations. On the contrary, it could be plausible to focus on the interrelation of mental representations. The first corresponds to the symbolic paradigm in AI, according to which mental representations are symbols. The second corresponds to connectionism in AI, according to which mental representations are distributed patterns [2].

The task to consider the similarity between human and artificial representation could involve the risk of skepticism about the possibility of “computing” this mental capacity. If we consider computationalism as defined in purely abstract syntactic terms then we are tempted to abandon it because human representation involves “real world constrains”. But, a new view of computationalism could be introduced that takes into consideration the limits of the classical notion and aims at providing a concrete, embodied, interactive and intentional foundation for a more realistic theory of mind [3].

We would like to highlight also an important and recent debate on “digital representation”[4] that focuses on the nature of representations in the computational theory of mind (or computationalism). The starting point is the nature of mental representations, and, particularly, if they are “material”. There are authors who maintain that mental representation are material [5] others thing that thought processes use conventional linguistic symbols [6]. The question of digital representation involves the “problem of physical computation [7] as well as the necessity of the notion of representation [8] so that we only have the problem of how to intend the very notion of representation [9]. But, there is also the possibility of understanding computation as a purely syntactic procedure or to include “every natural process” in a “computing universe” [10].

2. What is AP?

The core point of Brandom’s original book *Between Saying and Doing* [11] is to describe discursive practices and to introduce norms for deploying an autonomous vocabulary namely a vocabulary of a social practice (science, religion etc.). These norms are logical and are at the basis of an “inferential” notion of representation. But, inference in this sense, recalling Frege, is material [12]. Brandom refuses the explanation of representation in terms of syntactical operations as presented by “functionalism” in “strong” artificial intelligence (AI). He does not even accept weak AI (Searle), rather he aims to present a “logical functionalism” characterizing his analytic pragmatism (AP) [13]. Even though Brandom uses his account of representation to refuse computationalism, his

¹ Pontifical Lateran University

pragmatism is different from the Dreyfus's one, which rests on a non-linguistic know-how (logically and artificially not computable). According to Brandom, we are not only creatures who possess abilities such as to respond to environmental stimuli we share with thermostats and parrots but also "conceptual creatures" i.e. we are logical creatures in a peculiar way.

First, we introduce "practice-vocabulary sufficiency" or "PV-sufficiency" which obtains when exercising a specific set of abilities is sufficient for someone to count as deploying a specified vocabulary [14]. These are for instance "the ability to mean red by the word red" or "the capacity to refer to electrons by the word electrons" (Brandom includes even *intentions* to refer). Together with these basic abilities we must consider the relationship between these and the vocabulary in which we specify them. A second basic meaning-use relation is the "vocabulary-practice sufficiency" or just "VP-sufficiency" namely the relation that holds between a vocabulary and a set of practices-or-abilities when that vocabulary is sufficient to specify those practices-or-abilities.

In order to deploy any autonomous vocabulary we must consider the necessity of certain discursive practices defined as "asserting" and "inferring" that, according to Brandom, rule out computationalism [15]. According to the PV-necessity thesis, there are two abilities that must be had by any system that can deploy an autonomous vocabulary: the ability to respond differentially to some sentence-tokenings as expressing claims the system is disposed to *assert* and the ability to respond differentially to moves relating one set of such sentence-tokenings to another as *inferences* the system is disposed to *endorse*. By hypothesis, the system has the ability to respond differentially to the inference from p (premise) to q (conclusion) by accepting or rejecting it. It also must have the ability to produce tokenings of p and q in the form of asserting.

3. Why AP could be a fruitful strategy to simulate representation?

In this conclusive session I'll try to show that the notion of representation described in AP terms presents aspects that are common to human and artificial intelligence.

The PV- and VP-sufficiency thesis suggest that basic practices can be computationally implemented and this description corresponds to the Brandomian interpretation of the Turing test and, consequently, to the refusal of a classical symbolic interpretation

in AI (GOFAI) of the notion of human representation. Brandom introduces a pragmatic conception of artificial intelligence or "pragmatic AI" which means that any practice-or-ability P can be decomposed (pragmatically analyzed) into a set of primitive practices-or-abilities such that:

1. they are PP-sufficient for P, in the sense that P can be algorithmically elaborated from them (that is, that *all* you need in principle to be able to engage in or exercise P is to be able to engage in those abilities plus the algorithmic elaborative abilities, when these are all integrated as specified by some algorithm); and
2. one could have the capacity to engage or exercise *each* of those primitive practices-or-abilities without having the capacity to engage in or exercise the target practice-or-ability P [16].

For instance, the capacity to do long division is "substantively" algorithmically decomposable into the primitive capacities to do multiplication and subtraction. Namely, we can learn how to do multiplication and subtraction without yet having learning division.

On the contrary, the capacities to differentially respond to colors are not algorithmically decomposable into more basic capacities. This observation entails that there are human but also animal capacities that represent a challenge for strong AI (GOFAI), but nowadays not for new forms of computationalism. Starting from Sellars, we can call them *reliable differential capacities to respond to environmental stimuli* [17] but these capacities are common to humans, parrots and thermostats so that they do not need a notion of representation as symbol manipulation.

Along the line introduced by Sellars, Brandom intends the notion of representation in an "inferential" sense. It is grounded on the notion of "counterfactual robustness" that is bound to the so-called frame problem [18]. It is a cognitive skill namely the capacity to "ignore" factors that are not relevant for fruitful inferences. The problem for AI is not *how* to ignore but *what* to ignore. In Brandom's words: "Since non-linguistic creatures have no semantic, cognitive, or practical access at all to most of the complex relational properties they would have to distinguish to assess the goodness of many material inferences, there is no reason at all to expect that that sophisticated ability to distinguish ranges of counterfactual robustness involving them could be algorithmically elaborated from sorts of

abilities those creatures do have” [19]. Nevertheless, we could start by studying what “intelligence” really is by starting from the simplest cases.

Brandom introduces the notion of “counterfactual robustness” to overcome strong GOFAL, to avoid the primacy of prelinguistic background capacities and skills in weak AI (Searle) and phenomenology (Dreyfus). The notion of representation he introduces could work only if we embrace a peculiar form of inferentialism. Differently, we could read AP to analyze inferential capacities that are connected with logical laws common to human and artificial agents [20].

REFERENCES

- [1] M. Carter, *Minds and Computers*, Edimburgh University Press, Edimburgh, 2007.
- [2] Carter (2007), chap. 18.
- [3] M. Scheutz (ed.), *Computationalism. New Directions*, MIT, 2002.
- [4] V.C. Müller, *Representation in Digital Systems*, in *Current Issues in Computing and Philosophy*, A. Briggle et. Al. (Eds), IOS Press, 2008.
- [5] A. Clark, ‘Material Symbols’, *Philosophical Psychology* 19 (2006), 291-307; S. Sedivy, ‘Minds: Contents without vehicles’, *Philosophical Psychology* 17 (2004), 149-179.
- [6] J. Speaks, ‘Is mental content prior to linguistic meaning?’, *Nous* 40 (2006), 428-467.
- [7] O. Shagrir, ‘Why we view the brain as a computer’, *Synthese* 153 (2006), 393-416.
- [8] J.A. Fodor, ‘The mind-body problem’, *Scientific American* 244 (1981), 114-123.
- [9] G. O’Brien, ‘Connectionism, analogicity and mental content’, *Acta Analytica* 22 (1998), 111-131.
- [10] G. Dodig-Crnkovic, ‘Epistemology naturalized: The info-computationalist approach’, *APA Newsletter on Philosophy and Computers* 6 (2007), 9-14.
- [11] R. Brandom , *Between Saying and Doing*, Oxford University Press, Oxford.
- [12] R. Brandom, *Making It Explicit*, Cambridge University press, Cambridge, 1994, chap. 2; R. Giovagnoli, *Razionalità espressiva. Scorekeeping: inferenzialismo, pratiche sociali e autonomia*, Mimesis, Milano, 2004; R. Giovagnoli (ed.), ‘Prelinguistic Practices, Social Ontology and Semantics’, *Etica & Politica/Ethics & Politics*, vol. XI, n. 1, 2009.
- [13] Brandom (2008), chap. 2, chap. 3.
- [14] Brandom (2008), pp. 74-77.
- [15] Brandom (2008), pp. 77-83.
- [16] Brandom (2008), pp. 82-83.
- [17] W. Sellars, *Empiricism and the Philosophy of Mind*, Harvard University Press, Cambridge, 1957, 1997.
- [18] Brandom (2008), p. 79.
- [19], Brandom (2008), p. 83.
- [20] H. Boley, *Semantic ‘Web. Knowledge Representation and Inferencing’*, 2010, <http://www.cs.unb.ca/~boley/talks/DistriSemWeb.ppt>; Carter (2007); R. Evans, ‘The Logical Form of Status-Function Declarations’ in Giovagnoli (ed.) (2009); R. Evans, ‘Introducing Exclusion Logic as Deontic Logic’ in *Lecture Notes in Computer Science*, vol. 6181/2010; R. Giovagnoli, ‘Osservazioni sul concetto di “pratica autonoma discorsiva” in Robert Brandom’, in *Etica & Politica/Ethics and Politics*, IX, 1, 2008, pp. 223-235; R. Giovagnoli, ‘On Brandom’s “Logical Functionalism”’, *The Reasoner*, 4 (3), (2010), www.thereasoner.org.