Intentionality and Background: Searle and Dreyfus against Classical AI Theory

Intencionalidade e pano de fundo: Searle e Dreyfus contra a teoria clássica da inteligência artificial

Teodor Negru[1]
Al. I. Cuza University of Iasi

Abstract
According to the theory of artificial intelligence (AI), the human mind is a formal system made of symbols that operate following a set of instructions, which allow the manipulation of symbols according to their physical form. Against the conception that mental states and processes can be defined only from a syntactic perspective, John Searle used the Chinese Room thought experiment by which he intended to demonstrate that the human mind is more than a formal structure, having a semantic content as well. The semantic content of the human mind is given by intentionality, a feature that belongs exclusively to biological organisms. Starting from here, Searle shows that the logical structure of intentionality and the conditions for functioning of the intentional states cannot be explained by the computational approach of the human mind. Another perspective, which invalidated the AI theory, belongs to Hubert Dreyfus who considers that an adequate understanding of the human mind needs to start from the understanding of the phenomenological structures by means of which we relate to the world. Therefore, cognition and intentionality are explained from the perspective of an embodied being that, due to his body skills, is ontologically and dynamically coupled to the world. In this case, it is not the biological dimension of the human body that matters, but the phenomenological one that does not treat intentionality as knowing-that, whose role is to grasp the world's objective features, but as a way of constituting the world of the subject according to his concerns and interests.

Key words: AI theory, intentionality, logical analysis, phenomenology, background, context.

Resumo
De acordo com a teoria da inteligência artificial (IA), a mente humana é um sistema formal feita de símbolos que operam seguindo um conjunto de instruções, que permitem a manipulação de símbolos de acordo com sua forma física. Contra a concepção de que estados e processos mentais só podem ser definidos a partir de uma perspectiva sintática, John Searle usou a experiência do pensamento *Quarto Chinês* pela qual ele pretendia demonstrar que a mente humana é mais do que uma estrutura formal, ela também tem um conteúdo semântico. O conteúdo

---

[1] Al. I. Cuza University of Iasi. Carol I Boulevard, 11, 700506, Iasi, Romania. E-mail: theonegru@yahoo.com

semântico da mente humana é dado pela intencionalidade, um recurso que pertence exclusivamente a organismos biológicos. A partir daqui, Searle mostra que a estrutura lógica de intencionalidade e as condições de funcionamento dos estados intencionais não podem ser explicadas pela abordagem computacional da mente humana. Outra perspectiva, que invalidou a teoria IA, pertence a Hubert Dreyfus, que considera que uma adequada compreensão da mente humana precisa começar a partir da compreensão das estruturas fenomenológicas por meio da qual nos relacionamos com o mundo. Portanto, cognição e intencionalidade são explicadas a partir da perspectiva de um ser encarnado que, devido a suas habilidades corporais, é ontologicamente e dinamicamente acoplada ao mundo. Neste caso, não é a dimensão biológica do corpo humano o que importa, mas o fenomenológico que não trata intencionalidade como sabendo-que, cujo papel é o de compreender características objetivas do mundo, mas como uma forma de constituição do mundo do sujeito de acordo com seus interesses e preocupações.

Palavras-chave: Teoria IA, intencionalidade, análise lógica, fenomenologia, fundo, contexto.

The problems of artificial mind

The classical approach to AI developed by computational theory claimed that the human brain is similar to a data processing device and the human mind to the software run by this hardware. This conception – which, on the one hand, aimed at completing the behaviouristic description of the mind in observable input-output terms and, on the other hand, wanted to overcome the dualist mistake that transforms the mind into an entity that could not be approached from a scientific point of view – relied on the idea that the mind is a formal system made of symbols, following a set of instructions whose role is to guide the combination of symbols. The instructions were thought to be algorithm-type procedures, which allowed the manipulation of symbols according to their physical form. Processing symbols by means of syntactic rules used to be thought sufficient to guarantee both the transition from premises to conclusion and the semantic coherence of a sequence of symbols. The consequence of approaching the human mind from the perspective of computational relationships was its transformation into an invariant structure, independent both from the context it operates within and from the physical mechanism of producing it.

All these features belong to the conception John Searle called strong AI, which distinguishes from the weak AI version by the belief that the human mind can be entirely simulated by a computer. While weak AI claims that formal programs are just models of the mind, without implying that computers literary have a mind, strong AI version endorses that implementing a program designed properly it is enough to say that the computer has a mind. Starting from here the researchers imagined general mind simulation models, e.g. the Turing machine, and even tests that a machine should pass in order to be considered as functioning similarly to the human mind. According to the Turing test, if a machine succeeds in performing a kind of behaviour, which for an outside observer may not seem different from that of a human being, than it is enough to say that such machine has cognitive skills as the human one.

Against the conception that mental states and processes can only be defined from the syntactic perspective, John Searle imagined a thought experiment that demonstrates that the human mind is not only a formal structure, having content as well. The experiment Searle proposed is developed similarly to the

19

Turing behaviouristic test: hence, a person considered to have no knowledge of Chinese is locked in a room with a rulebook by means of which s/he can answer any question in Chinese. This rulebook is so accurate that it not only contains all the answers but the answers it offers "are indistinguishable from those of a native Chinese speaker." (Searle, 1984, p. 32) Therefore, the requirement of the Turing test is fully met, i.e., the person in the room behaves as if s/he were a highly competent user of Chinese. Similarly, the requirement of a formal system is met, the answers being given only following the syntactic rules of the symbols used, their meaning remaining totally unknown to the user locked in the room. However, this is the starting point of Searle's objection: nowhere in the process of syntactic use of Chinese ideograms occurs the possibility of understanding their meaning, i.e., understanding their semantic content. In other words, the formal character of computer programs unfolds only syntactic sequences, without the possibility of inferring the interpretation or assignment of meanings to symbols thereto.

> What goes for Chinese goes for other forms of cognition as well. Just manipulating the symbols is not by itself enough to guarantee cognition, perception, understanding, thinking and so forth. And since computers, qua computers, are symbol – manipulating devices, merely running the computer program is not enough to guarantee cognition (Searle, 1990, p. 26).

The conclusion to the Chinese Room argument, i.e., syntax is not enough to found and guarantee semantics, shall be completed by Searle with the observation that syntax, just like semantics, are observer-relative notions and not physical features intrinsic to the system. The inputs and outputs of a formal system are but mere physical items, which acquire their meaning or syntax from an external observer. To characterise a system as being computational means to assign a certain feature to that system, from the user's point of view, without indicating thus an essential feature of the system.

Starting from here, Searle highlights other weak points of the strong AI conception, e.g., if computational approach is nothing but a matter of interpretation, then it cannot become a scientific subject similar to natural sciences. The role of natural sciences is to uncover the essential features of the objects and phenomena they study, features intrinsic to them and not assigned by an observer. However, computational theory only provides descriptions of some features observer-dependent, which means that it cannot be considered a natural science.

Furthermore, starting from the idea that syntax is not a part of physics, but it is a feature assigned to physical system by a user, Searle takes again the critique with respect to the homunculus character of computationalism. According to the homunculus fallacy, a solution to assign intentionality to a formal program is to admit the existence of an internal agent to the system, which should possess a non-derived intentionality. A variant of this argument is offered by Daniel Dennett who imagines a hierarchy of homunculi, where the highest level rests on the lower one, made of less and less intelligent homunculi who are lacking more and more intentionality, working only following some 0-and-1 handling rules. Thus, the intention is to rely the system's intentionality on levels of computation made of more and more non-intentional elements. However, according to Searle, an explanation in these terms does not demonstrate that the lower homuncular levels are intrinsic features to the physical system. To put it differently, "without a homunculus that stands outside the recursive decomposition, we do not even have a syntax to operate with" (Searle, 1992, p. 213).

20

Following this line of thinking, Searle shows that a weak point of the comparison between a computer and the human mind is that computational processes lack causal power. Syntax, or the 1 and 0 sequence characterising computer software, does not have its own intentionality which would enable it to cause effects beyond its environment. If the human brain is built in such a way that it could consciously follow some rules, computers behave *as if* they followed some rules. Moreover, this behaviour is suggested only from the perspective of an external observer. Without the existence of the external observer or of the intrinsic intentionality, all that remains from the computer and the brain are some patterns, which cannot have causal power by themselves.

One last mistake of the computational approach is that it likens the way the brain processes information with the one of a computer. In the case of computers, someone from the exterior encodes information that is to be processed, from which results an output in a physical form. In the case of the human brain, the neurobiological processes are not observed-dependent, operating with information coming from modalities, which benefit from the intrinsic intentionality. Therefore, the biological level is endowed with a high level of concreteness that cannot be grasped by the abstract approach of computer programming in terms of processing some symbols. Consequently, it is inadequate to say that the brain processes information but rather that "it is a specific biological organ and its specific neurobiological processes cause specific forms of intentionality" (Searle, 1992, p. 226).

Searle's conclusion is that formal programs can be neither constitutive nor sufficient to produce mental phenomena. This means that the computational approach, which implies the existence of an intermediate symbolic level between the neuronal processes and the intentional states, is wrong. An accurate understanding of the human mind implies treating it as a biological phenomenon and understanding mental states as a result of brain processes. This goal cannot be achieved without understanding intentionality as an intrinsic feature of the human mind and of the logical conditions for functioning of the intentional states.

Another perspective contesting the possibility of simulating the human mind by AI is phenomenology. According to Hubert Dreyfus, the idea underlying the AI theory can be found ever since the ancient beginnings of metaphysics, when logic and geometry were invented, which inspired the idea that "all reasoning might be reduced to some kind of calculation" (Dreyfus, 1972, p. xv). If Plato's and Aristotle's conceptions represented mere attempts to put this idea to practice, however insisting on the importance of the content of thinking, starting with the Modern Age, the approach of thinking in terms of formal rules came to dominate western thought. Hence, Galileo, Hobbes or Leibniz started from the idea that the structure of the mind is syntactic, trying to find ways of formally approaching thinking which should not refer to its semantic content. Such idea was later embraced by mathematicians who, apart from trying to discover the laws governing the operations of the mind (e.g., George Boole), also began to imagine machines that would operate based on logical combinations (e.g., Charles Babbage). From the phenomenological perspective, the beginnings of building machines, which, by operating with syntactic rules were meant to imitate the human mind, were not only a consequence of the technical progress but they also represented a fulfilment of western metaphysics that aimed at operating with purely rational thinking, lacking any subjective element.

After the emergence of the theoretical fundamentals of AI, one tried to apply it to several problems, regarded as essential for the simulation of the human mind at the level of a machine: game playing, language translating, problem solving, and pattern recognition. On a general level, showed Dreyfus, the beginnings of AI research had two stages: the former was dominated by the interest in Cognitive

21

Simulation, "the use of heuristic programs to simulate human behavior by attempting to reproduce the steps by which human beings actually proceed" (Dreyfus, 1972, p. xxxxiii). This stage was marked by several significant successes, mainly with respect to carrying out simple tasks that did not involve a complex cognitive structure. At the same time, the difficulties of approaching the human mind in syntactic terms became obvious as such approach became less and less applicable to problem solving or game playing in ambiguous contexts or in the case of language translation dependent on contextual elements which could not be formalized. Hence, it was ascertained that natural language is characterized by ambiguity, which can be reduced by cues context-dependent which are not necessarily linguistic.

The latter stage was marked by disappointment as the attempts at approaching complex problems did not yield the results expected. This stage was dedicated mainly to semantic processing of information and to developing an artificial intelligence similar to the human mind. Just as in the former stage, the results reached led to the conclusion that "human beings do not deal with a mass of isolated facts as does a digital computer, and thus do not have to store and retrieve these facts by heuristics rules. Judging from their behavior, human beings avoid rather than resolve the difficulties confronting workers in Cognitive Simulation and Artificial Intelligence by avoiding the discrete information-processing techniques from which these difficulties arise" (Dreyfus, 1972, p. 59-60).

Starting from this historical analysis, Dreyfus concluded that the AI theory can be summarised as relying on four assumptions: the biological assumption claims the structural and functional similarity of the human brain with the computer, in the sense that at the neurophysiologic level, our brain operates with information in discrete operations which would correspond to the information units in the computer. The psychological assumption reduces reasoning to processing some information according to formal rules. This means that human behaviour must be explained in terms of an information-processing level, which differs from the neuronal one. The epistemological assumption considers that the entire knowledge can be formalised and "can be expressed in terms of logical relations, more exactly in terms of Boolean functions, the logical calculus which governs the way the bits are related according to rules" (Dreyfus, 1972, p. 68). Finally, according to the ontological assumption, all that is essential to the intelligent behaviour can be reduced to a set of elements determined independently from any context.

To each of the above assumptions, Dreyfus brings counter arguments, which generally rely on highlighting the oversight of man's phenomenological dimension. Hence, against the biological assumption, he shows that the digital model described by the classical AI theory is not the only way to process information. The brain's way of processing information is rather analogical, meaning that information is processed globally and not by assigning a symbol to every bit of information. Moreover, in the case of the brain we should also consider the way the cerebral mechanism operates which does not offer arguments for digital processing of information.

Against the psychological assumption, Dreyfus shows that the role of psychology is not to describe man as a device providing responses to outside inputs following certain rules. This mistake occurs whenever no rigorous distinction operates between the neuronal and the phenomenological level, resulting in an intermediate level that requires explanations in terms of a new vocabulary. Preserving the distinction between the two levels removes the intermediate level suggesting that, although man can be approached from a physical perspective as a data-processing device, his behaviour cannot be explained in these terms but only by considering his phenomenological dimension.

22

Dreyfus brings arguments against the epistemological assumption from physics and linguistics. Thus, according to the formal approach, if man as a data processing device, is part of the physical world and the physical world follows scientific laws that can be grasped in mathematical formulae, then man too is subject to such laws. However, this argument relies on the confusion between the physical laws and the data-processing rules, considering that the latter are sufficient to guarantee the existence of the former. Moreover, there is a difference between the information processed by a computer and the information processed by its natural analogous.

> It is not processing the information which is processed by the simulated analogue, but entirely different information concerning the physical or chemical properties of the analogue. Thus the strong claim that every form of information can be processed by a digital computer is misleading. One can only show that for any given type of information a digital computer can in principle be programmed to simulate a device which can process that information (Dreyfus, 1972, p. 107).

The second argument against the epistemological assumption refers to language whose rules cannot be entirely formalised. Apart from grammar rules, which can be formalised and simulated by a computer, there are linguistic performance rules, as Dreyfus calls them, which cannot be formalised due to two reasons. Firstly, because it would mean we are able to systematise our entire knowledge in a general theory, which is impossible, and secondly, because not all linguistic rules can bear an objective description. This is exemplified by the fact that we are also able to understand vague linguistic expressions that do not follow any rules.

In addition, an important argument against both the possibility of formalising language rules and the ontological assumption of independent primitives is the fact that the AI theory does not take into consideration the context a sentence is formulated or the context in which the cognitive agent carries out his activity. The idea that we can break down knowledge into simple elements, which combine themselves following some rules has been present in western philosophy ever since Plato. However, such a conception, which transforms knowledge into a sum of atomic facts, does not take into consideration the fact that everyday knowledge is a kind of situated context-dependent knowledge.

The holistic character of knowledge, highlighted by phenomenology, means that we do not know things in isolation but starting from the relationships the new objects of knowledge have with other already known objects or human beings. This means that human beings have a kind of implicit knowledge of the context they are in and this kind of knowledge is an essential condition to understand what is going on in the world. Nevertheless, a computer cannot simulate this contextual knowledge; for a computer to be able to recognize a context, that computer needs to recognize the relevant features of that context. Such a task exceeds the limited capacity of a computer as it assumes formalising the entire human knowledge. Thus, contextualisation remains an exclusive feature of the humans owing to their direct coping with the world.

Dreyfus concludes that an adequate understanding of the human mind needs to start from understanding the phenomenological structures by means of which we relate to the world. Thus, cognition needs to be approached not in terms of logical structures but starting from what opens access to the world: the human body. Starting from here, we will understand that intentionality is the product of dynamic interaction between the body and the environment and that the cognitive agent is a situated one, having a direct understanding of the world, which cannot be formalised. These features can be grasped only by means of a phenomenological approach of the world, to which the computational theory cannot have access.

23

Intentionality and Embodiment

The conclusion of the Chinese room thought experiment was that syntactic processing of information cannot explain the semantic content of our mind's mental states. According to Searle, the semantic content of the human mind is given by intentionality, a feature that belongs exclusively to biological organisms.

The starting point of Searle's conception of intentionality is the classical conception that considers intentionality "that feature of mental states by which they are directed at or about objects and states of affairs other than themselves" (Searle, 1999, p. 99). Notwithstanding, Searle means to offer an explanation of intentionality which should not transform it into a transcendental phenomenon, as in phenomenology, neither reduce it to the causality relationship between different aspects of the world tokening symbols in our mind, as it does in the computational theory. Intentionality, just like consciousness, is regarded as a biological phenomena, which are caused and developed in the structures of the brain. Hence, the only science that can offer an explanation to intentionality, and to consciousness, is biology and not the computational approaches of the mind or the phenomenological theories.

The mistake of assigning intentionality to formal systems comes from the lack of the distinction between intrinsic and derived intentionality. Such distinction is made based on the perception with respect to the existence of two types of features of the world: the observer-independent and the observer-dependent ones. The intrinsic intentionality belongs to observer-independent features. The biological processes in the animal and plant kingdoms, such as mitosis, meiosis or photosynthesis, belong to this category, and so do both the primitive forms of desire, such as hunger and thirst, and the higher manifestations of complex biological organisms, such as consciousness, intentionality, or mental states.

The derived intentionality is observer-dependent and is a consequence of the mind's power to offer intentionality to entities that lack their own intentionality by assigning conditions of satisfaction, which belong to a certain psychological state. Such is the case of language that acquires the capacity of representation from the intentional mental states, thus having derived intentionality.

> I couldn't make a statement without expressing a belief or make a promise without expressing an intention because the essential condition on the speech act has as conditions of satisfaction the same conditions of satisfaction as the expressed Intentional state. So I impose Intentionality on my utterances by intentionality conferring on them certain condition of satisfaction which are the conditions of satisfaction of certain psychological state (Searle, 1983, p. 28).

Just like language, computer programs have derived intentionality from the input used by the programmer. Therefore, the syntactic sequences of formal programs do nothing but simulate mental states, without operating on the semantic dimension, which makes the content of mental states and which is a consequence of the mind's intrinsic intentionality.

The content of intentional states is given by the object or the state of affairs in the world the mind is oriented onto. This content displays itself in a certain psychological mode, which represents the way the content relates to the world and can take the form of desire, belief, fear, hope, etc. Hence, intentional states are regarded as representations, defined not by formal structure but by their propositional content (e.g., it is warm outside) – which does not necessarily need to be a full sentence – presented in a psychological mode (i.e., I believe/hope/assume that it is warm outside). As the relation to the world is defined by means of a psychological

24

state, the intentional states have an aspectual shape, which cannot be grasped by the computational or functionalist approaches that analyse intentionality in terms of representational relationships between the mind and the world.

Furthermore, the intentional states relate the propositional content to the world in different ways. In Searle's terms, this means that they have different directions of fit: the truth value of the beliefs results from the adequacy between the propositional content and the state of affairs in the world. This is why we say that they have a mind-to-world direction of fit. Intentions and desires, whose satisfaction depends on what is going on in the world, have a world-to-mind direction of fit. And there is also another category of intentional acts, like shame and pride, which do not attempt to somehow fit to reality, nor do they attempt to fit reality to what they express, hence, they have a null direction of fit.

> Where the mental state is responsible for fitting an independently existing reality, we can say that the mental state has the "mind-to-world direction of fit", or alternatively, it has the "mind-to-world responsability of fit." [...] Some intentional states, though they have a propositional content, do not have a direction of fit because it is not their aim either to match them (the mind-to-world direction of fit) or to get reality to match them (the world-to-mind direction of fit). Rather, they take it for granted that the fit already exists (Searle, 1983, p. 168-169).

The direction of fit shows the conditions to be met in order to satisfy an intentional state. As we cannot say about all intentional states that they have truth values, such as desires, for instance, of which we cannot say whether they are true or false, but that they can be fulfilled or not, in the case of intentional states, we will speak about conditions of satisfaction. They are a consequence of the propositional content of intentional states and of how this content relates to the world by means of the direction of fit. Even the states that have a null direction of fit are regarded as having conditions of satisfaction by reducing them to states, such as beliefs or desires, which have directions of fit and, consequently, conditions of satisfaction. This is the reason why Searle considers that intentional states can be interpreted as representations of their conditions of satisfaction.

This means that between the content of intentional states and their conditions of satisfaction there is a logical connection, which Searle identifies as causality. But we cannot speak of causality as in Hume's interpretation, where the cause is exterior to the effect, but in the case of intentional states "the cause is a representation of the effect or the effect is a representation of the cause" (Searle, 1999, p. 105). Consequently, the structure of such causality is described as self-referential in the sense that the conditions of satisfaction refer to that very state in order to fulfil the intentional act. Even if this thing is not valid for all the cases of Intentional causation – as in the case of desires and beliefs, where intention may lead to fulfilling an action which is not part of the state, Searle concludes that any intentional state "either causes or is caused by its conditions of satisfaction" (Searle, 1983, p. 123). Therefore, causality is an essential feature of intentional states which cannot be met in the case of computer program. The syntax underlying software development is merely formal as there is no content which could determine its causal power by determining the conditions of satisfaction.

Causality is one of the features that make the difference with respect to the cognitive science interpretation of the unconscious mental states as well. According to the strong AI conception, computational operations conduct various mental processes in an unconscious way. One such example is Noam Chomsky's theory describing a universal grammar as some innate set of rules we follow when we learn a

25

language. Agreeing with strong AI, Searle admits that there are unconscious mental states, whose role is to constitute the conscious states, but, he adds, the theory of cognitive science does not explain how the conscious states cannot be differentiated from the unconscious ones. The unconscious states are actually unconscious neurobiological processes that can determine the content of a conscious state. Such states should be treated as neurobiological structures behaving as if they had intentionality, meaning that they have the power to determine conscious thoughts and behaviours, but whose motivational structure remains unknown.

This does not mean that we should reduce the intentional states to the biological process of the brain, as the latter, just like the computer programs, do not grasp, the aspectual shape by means of which the content of an intentional state is presented in a certain manner. Conversely, we should understand the neurobiological processes from the perspective of their ability to cause conscious states, a feature not to be met with formal programs, characterised merely by syntax. Therefore, understanding conscious human behaviour in terms of rule-following is not a satisfactory explanation, as is the case with how a computer operates.

> If talk about unconscious rule following is to be taken literally, then such rule following has to have these features: the rule functions causally with the world-to-rule direction of fit and at the rule-to-world direction of causation. The rules have to have an aspectual shape, be followed voluntarily, be followed in a way that is subject to different interpretations; and they have to be followed in real time. Some postulations of unconscious rule following, such as rule following in the performance of speech acts, meet these conditions. But many postulations of unconscious rule following, as in the cognitive science accounts of visual perception and language acquisition, do not meet these conditions (Searle, 2004, p. 256).

The conclusion Searle reaches against computationalism is that intentionality is a characteristic feature of a biological system that cannot be replicated by a formal program. Thus, Searle objects to the functionalist thesis of multiple realizability, which considers that the mind can be created in environments other than the human brain, as it is only a matter of software. Phenomena such as consciousness or intentionality can only appear within a biological organism endowed with a brain similar to the human brain. Furthermore, these phenomena are not the result of a programmer's input, but of evolution, which selected those biological traits significant for the organism to survive, such as consciousness, whose advantage is to coordinate simultaneously a large amount of (intentional) information.

Contrary to Searle's approach, which reduces intentionality to merely its logical conditions for functioning, Dreyfus considers that without a phenomenological explanation, man's relationship with the world continues to remain within the framework of a representationist approach, which also characterises the approach of mind in terms of the AI theory. Searle's conception, which refers to searching for the conditions of satisfaction of the intentional states, fails on a phenomenological level, as the existence of such logical conditions of intentionality also implies their representation in the subject's mind, "i.e., that they must be structures of a conscious subject separate from and standing over-against an object" (Dreyfus, 1993, p. 19).

According to Dreyfus, this idea is also supported by Searle's conception regarding the role of mental states in triggering and supporting our actions. For a bodily movement to pass as an action, it needs to be accompanied throughout its development by a representation of the goal of the action, which should play a causal role and which should be known as such by the subject. In other words, according to the self-referential causality of the intentional states, the conditions of

satisfaction include the requirement that the intention of fulfilling a goal to cause the action oriented to that very goal. Such an approach of intentionality is nothing but a mentalistic approach of the subject's relationship with the world, which, again, brings up the issue of the ontological difference between the subject and the object.

Against such a conception, Dreyfus resorts to Heidegger and Merleau-Ponty's contributions to the intentionality approach. Both conceptions start from the idea that context-independent knowledge, as the one Descartes theorized and which is still to be found in the AI approach is not valid for the human being's everyday situation. An authentic understanding of how the subject related to the world should consider the subject as an active being, which is ontologically situated in a world and whose intentionality is characterized by the lack of any intentional content. This means that the relationship with the world is not conducted by a representational content guiding our intentional states. Similarly, the idea of an extended mind, according to which, by means of the mind, man extends his cognitive capacities to the world thus achieving a comprehension of the latter, is not supported. Rather, the relationship between the cognitive agent and the world is direct implying a skillful coping, which is spontaneous and lacks any intrinsic goal.

> According to Heidegger intentional content isn't in the mind, nor in some 3rd realm (as it is for Husserl), nor in the world; it isn't anywhere. It's an embodied way of being towards. Thus for a Heideggerian, all forms of cognitivism externalism presuppose a more basic existential externalism where even to speak of "externalism" is misleading since such talk presupposes a contrast with the internal. Compared to this genuinely Heideggerian view, extended-mind externalism is contrived, trivial and irrelevant (Dreyfus, 2009, p. 53).

The conception of Merleau-Ponty, who understands the situated cognitive agent as an embodied agent, plays an important role in understanding how this everyday coping works. This means that the world of the subject is disclosed by his bodily sensory-motor skills. The human body is understood as a lived body by means of which we not only obtain the necessary information from the world but also enact the world-we-live-in. The world appears thus inseparable from the body and it represents an important requirement to constitute the subject's ipseity.

> In so far as, when I reflect on the essence of subjectivity, I find it bound up with that of the body and that of the world, this is because my existence as subjectivity is merely one with my existence as a body and with the existence of the world, and because the subject that I am, when taken concretely, is inseparable from this body and this world. The ontological world and body which we find at the core of the subject are not the world or body as idea, but on the one hand the world itself contracted into a comprehensive grasp, and on the other the body itself as knowing-body (Merleau-Ponty, 2005, p. 475).

Thus, the Cartesian idea that the mind can function independently from a body is invalidated by understanding our cognitive structures as the result of our body's sensory-motor patterns. The body has the role of Husserl's transcendental consciousness of constituting the meaning, but it does no longer act by anticipating meaning, which would imply the synthesis between form and content, thus also implying the existence of a formal rule. If machines can be at best built in such a way as to offer an answer by testing some hypotheses, which would have to fit the data stored, in the case of the human beings this operation is conducted owing to the spontaneous coping-to-the-world skill of the body, which by means of bodily skills provides an unlimited set of responses to the environmental changes. Bodily skills do

27

not offer a formal analysis of the response possibilities, but, owing to the body and world coupling, they select directly the relevant answer for a certain context, having at their disposal an indefinite number of solutions to the problems that may occur.

What makes possible such spontaneous coping is the intentional arc and tendency to achieve a maximal grip. Intentional arc means the direct connection between the body and the world, conducted at a pre-reflexive level, by means of which the world is given directly to consciousness. Intentional arc subtends the entire activity of consciousness, pre-constituting the field of experience toward which consciousness is oriented.

> Let us therefore say rather, borrowing a term from other works, that the life of consciousness—cognitive life, the life of desire or perceptual life — is subtended by an 'intentional arc' which projects round about us our past, our future, our human setting, our physical, ideological and moral situation, or rather which results in our being situated in all these respects. It is this intentional arc which brings about the unity of the senses, of intelligence, of sensibility and motility (Merleau-Ponty, 2005, p. 157).

The existence of the intentional arc, as a basis of immediate experience, is due to the body's acquisition of some skills, which are not representations used to act within reality, but rather dispositions by means of which we respond to the situations in the world. These embodied skills, which are but developments of some bodily predispositions to respond to the changes in the world, open up the world in three ways: as a biological environment, as an environment of figurative meanings, or by revealing its cultural dimension. In all of these situations, owing to the intentional arc, man's relationship with the world is neutral, but the world is perceived in terms of affordances, which help man reach its goals and solve its tasks.

The other component of the embodied relationship with the world, the maximum grip, refers to the fact that by means of this everyday skilful coping with the world, one looks for an "optimal body-environment relationship" (Dreyfus and Dreyfus, 1999, p. 111). Skilful coping does not imply the existence of a goal that should constitute the content of an intentional act or that should require its mental representation. This is also valid for any skilful action, which should be approached from the perspective of restoring the equilibrium between the body and the world, and not from the perspective of looking for the intention or the goal justifying its achievement. The existence of this maximum grip completes the image of intentionality without an intentional content Heidegger formulated: the objective of intentional states is not to carry out a goal, represented in their content, but to achieve the maximum grip on the world. Thus, the tendency toward maximum grip, which implies the profound involvement of the body in coping with a situation, implies the intentional arc, meaning the development of bodily skills with a view to immediately cope with the world changes.

Our body's spontaneous tendency to reach a state of equilibrium with the environment opens the possibility of approaching the mind and world interaction in the nonlinear terms of dynamic systems. This means that the mind is not understood based on an input-output scheme, as it was thought in the AI theory, which constructs a representation of the world state of affairs from many features perceived in the exterior. Such an approach is invalidated by the binding problem, i.e., by the difficulty to explain according to what these elements, otherwise separated in reality, are bound. The binding problem is solved, or better said, does not emerge, if we understand that the answer to what is going on in the environment is the result of an ensemble of brain cells that detect a certain affordance starting from previous experiences.

28

Dreyfus takes, as an example, the theory of the neurobiologist Walter Freeman, who describes the functioning of the brain at a given time to have the tendency to orient toward a minimum energetic state called attractor. Cerebral states, which tend toward the same attractor, form a basin of attraction that occurs with every input of the brain, creating thus a landscape attractor. Any new attractor occurring in the basin does not have the role of representing an external object, but it activates its past experiences with that object. Thus, the brain resonates to the new affordance directly without activating a certain representation. This means that it does not react to the experience of a physical stimulus (excitation of a sensory organ), but to the significance such stimulus has to itself, acquired from past experiences.

The consequence of experimenting new significance of the world is the occurrence of a new attractor, which rearranges the basins of attractions in the landscape. Hence, the connection with the world achieved by the intentional arc is not an invariant structure, based on the understanding of isolated experiences of the interaction between the organism and the world. Nevertheless, the intentional arc achieves a dynamic coupling with the world by means of which, "each time a new significance is encountered, the whole perceptual world of the animal changes so that the significance that is directly displayed in the world of the animal is continually enriched" (Dreyfus, 2009, p. 64).

The dynamic coupling with the world via bodily skills also solves the objection raised against the AI theory regarding the way relevant data are selected in a situation. According to the frame problem, the AI theory fails to explain how a computer, which possesses a determinate number of data (the programmer's input), following some rules, can select the requested data needed to solve a task in a certain context. As man is directly engaged in a situation, owing to his skilful coping with the world, he has an unmediated knowledge of the information relevant to the context, without assuming data processing as in the heuristics of computer software. From this perspective, the body is regarded as being characterised by three functions, which cannot be simulated or achieved by any computer:

> (i) the inner horizon, that is, the partially indeterminate, predelineated anticipation of partially indeterminate data (this does not mean the anticipation of some completely determinates alternatives, or the anticipation of completely unspecified alternatives, which would be the only possible digital implementantation); (ii) the global character of this anticipation which determines the meaning of the details it assimilates and is determined by them; (iii) the transferability of this anticipation from one sense modality and one organ of action to another (Dreyfus, 1972, p. 167).

Therefore, cognition and intentionality are no longer reduced to the capacity of handling some symbols, independently from any material substratum. However, both are thought from the perspective of an embodied being, which by means of the bodily skills it is ontologically and dynamically coupled with the world. In this case, it is not the biological dimension of the human being that matters, but the phenomenological one, the one that treats intentionality not as knowing-that, which has the role of grasping the objective features of the world, but as a way to constitute the world of the subject according to his concerns and interests.

Background and contextualization

To Searle, the consequence of approaching intentionality from the perspective of the logical structure of the way to relate to the world has as a result the debate on the conditions for functioning of intentional states. To this aim, intentional states

29

are not approached in isolation, functioning separately from the others. Conversely, Searle considers that they are conditioning one another creating an interactive Network by means of which content and the conditions of satisfaction of every intentional state are determined. In other words, the position within the Network is an essential requirement to the function of the intentional states as the conditions of satisfaction are determined in relation to the Network and not individually by each intentional state separately. Therefore, it is not an atomistic approach of the intentional states that can explain how the conditions of satisfaction are determined, but only a holistic approach.

Integrating intentional states in a Network does not explain how they relate to the world. Intentional states do not self-interpret, nor do they apply themselves to various situations in the world. Consequently, there is the need to postulate a pre-intentional level made of non-representational mental capacities, which do not belong to the Network, to carry out coping with the world. These capacities, skills, tendencies, dispositions (or more generally, this know-how) that cannot be the object of comprehensive research and that share the feature of not being subject to analysis in terms of intentional states, create the Background.

> Intentional phenomena such as meanings, understandings, interpretations, beliefs, desires, and and experiences only function within a set of Background capacities that are not themselves intentional. Another way to state this thesis is to say that all representation, whether in language, thought or experience, only suceeds in representing given a set of nonrepresentational capacities. In my technical jargon, intentional phenomena only determine conditions of satisfaction relative to a set of capacities that are not themselves intentional. (Searle, 1992, p. 175)

Approaching the Background in terms of the first order mental states (i.e., intentions, desires, etc.), is rather a language game lacking an appropriate vocabulary that would grasp its non-intentional and non-representational side, hence, Searle's preference for the use of some terms such as practices, capacities, and stances, rather than using a description in terms of assumptions and presuppositions. Such an approach would have a representational character by implicitly referring to the existence of propositional content, having the truth value and assuming the existence of some logical relationships with the conditions of satisfaction. Similarly, the Background cannot be understood in terms of rules because rules cannot self-interpret nor can they self-apply, but they need a Background against which they operate.

Therefore, the Background is not the consequence of the transcendental relationship between the subject and the world, as phenomenology considers, but it is the condition of possibility of man's representing the social and biological world. Its content is given, on the one hand, by the capacities all people share, as biological beings that belong to the same species (i.e., deep Background). On the other hand, its content is given by the local cultural practices influencing the individual as a member of a certain society (i.e., local Background). Overall, we can say that the Background is a set of enabling pre-intentional conditions by means of which our intentional contents can determine various conditions of satisfaction, thus describing the world in various ways.

The Background is also an important condition in understanding language. Besides syntax, which gives form, language is endowed with semantics as well, given by the meanings of the words. However, this is undetermined as words can have more meanings or may acquire new meanings depending on the context. Fixing meaning is not a feature of the word in itself, nor is it of the sentence, because, by adding words, according to syntactic rules, we only achieve the addition of new

30

indeterminate semantic contents. The consequence of this process would be an infinite regress, which perpetuates the issue of semantic indeterminacy.

> If representation presupposes a Background, then the Background cannot itself consist in representation without generating an infinite regress. We know that the infinit regress is empirically impossible because human intellectual capacities are finite. The sequence of cognitive steps in linguistic understanding comes to an end. One the conception presented here, it does not come to an end with the grasp of semantic content in isolation or even with semantic content in isolation or even with semantic content together with a set of presupposed beliefs, but rather the semantic content only functions against a Background that consists of cultural and biologic know-how, and it is this Background know-how which enables us to understand literal meanings (Searle, 1983, p. 148)

Therefore, the issue of determining meaning can only be clarified by the assumption of a Background of capacities and social practices, which has no relationship with language syntax or semantics. This has the role to determine the conditions of satisfaction of the sentence, based on which the meaning of the sentence is to be determined. Hence, meaning is not a matter that belongs to language syntax, as the AI theories used to think, but it is dependent on the context of the utterance, i.e., it is fixed against a pre-intentional Background of practices and skills.

Searle's conclusion is that the Background is a preintentional stance representing the occurring condition of the intentionality forms from the integrated flow of perception and action. Therefore, it cannot be reduced to a finite sequence of procedures that can be computer simulated. However, it represents those skills, dispositions, states, behaviour, savoire-faire, know-how, etc., which only occur when mental states exhibit an intentional content.

The idea that the groundings of our knowledge have a pre-intentional and non-representational character, which makes it irreducible to computational operations, is also found in Dreyfus. He starts from Heidegger's conception according to which the human being is a being-in-the-world, who relates to the objects around by means of various practices. With the help of his actions and practices, man structures the situation he is in according to his concerns and interests. This is why he always acts in a context whose meaning is already given to him. Therefore, the objects around us are not neutral entities whose properties and relationships we mirror by means of the mind, but they are contained in a semantic network which involves their holistic interpretation. Unlike Searle, to whom the Network is made of the system of propositional attitudes we have with respect to the world, Heidegger's understanding of the semantic network also contain the knowledgeable subject, who thus has a direct knowledge of the ways to solve the problems which occur in different situations.

> Human beings are somehow already situated in such a way that what they need in order to cope with things is distributed around them where they need it, not packed away like a trunk full of objects, or even carefully indexed in a filing cabinet. This system of relations which makes it possible to discover objects when they are needed in our home or our world (Dreyfus, 1972, p. 172).

Furthermore, to Heidegger, the Background is not made of "skills" or "capacities" that are activated when there is an intentional goal-oriented state, but it occurs as a feeling of familiarity with the *Dasein* to the world around him. This non-intentional relationship between the subject and his world is called ontological transcendence and it represents the ways in which the *Dasein* means of action in the world become available to him. In other words, the *Dasein* by its orientation toward various objects

31

does not discover their conditions of satisfaction but it exercises a general skilled grasp by means of which the world is disclosed and direct coping becomes possible. The conclusion is that *all* coping, including unready-to-hand coping, takes place on the background of this basic non-representational, holistic, absorbed, kind of intentionality, which calls being-in-the-world" (Dreyfus, 2009, p. 56).

This originary situation of the subject is also an important point to differentiate between how man and computer related to the world. The computational approach considers the context irrelevant to the data with which a computer operates. For a computer to be able to respond to a problem, it needs to operate with a set of determined data to which it should assign a set of determined values. Unlike the computer, man has the possibility of processing holistically the information in the world, operating even with undetermined data. This is possible owing to the intentional arc and embodied skills by means of which we have direct access both to what is going on in the world and to the available solutions to solve our tasks.

In order to highlight the difference between man and computer, Dreyfus distinguishes among four types of knowledge: associationistic, formally simple, formally complex, and non-formal. The associationistic knowledge includes the information acquired by simple associations, which are either innate or learned by repetition. The meaning and the context are irrelevant for this type of knowledge, which consists in simple procedures, such as decision-making trees, list searching or a template, applied to some elementary problems, e.g., memory games, trial and error, word-by-word translation, response to rigid patterns. The second type of knowledge, i.e., the formally simple one, suits artificial intelligence best, as learning is achieved by rules and the procedures applied are algorithm type. The problems are fully formalised and measurable and the meaning is explicit, though context-independent. The next type of knowledge, the formal knowledge, refers to those situations which cannot be exhausted by algorithms and which require heuristic programs. In this case, learning is achieved by rules and practice and meaning is explicit and dependent only on the inner context not on the outer one as well. The last type of knowledge, i.e., the non-formal one, includes "all those everyday activities in our human world which are regular but not rule governed" (Dreyfus, 1972, p. 206). This implies that the activity depends on a meaning and a context, not explicitly given, and learning is achieved by examples and relies on intuition, which cannot be reduced to a set of rules.

If the first two types of knowledge can be computer-simulated, the third can only be partially approached from the computational perspective whereas the fourth is specific to humans. This is due to the following three reasons: a computers lacks bodily organisation that could provide a global organising level of perception and integration of bodily skills; it is impossible for computers to select from a context the aspects relevant to the interpretation of the meaning of an utterance; it is impossible for software to grasp the context of an action. All these lead to consider issues such as game playing, language translation, problem solving, and pattern recognition human-specific as they pertain to man's direct relationship with the world. Dreyfus' conclusion is that man's feature to be a being-in-a-world, i.e., an embodied being that is originary in a situation, cannot be expressed in computational terms, nor can it be computer-simulated.

Conclusion

Against the computational approach to mind, Searle highlights the biological dimension of the human being whose complexity is a result of evolution. Therefore, we should analyse man's essential features as well as their intentionality or consciousness as coping abilities developed throughout the stages of evolution. This

32

means that the answers to the problems related to the nature of consciousness and of intentional states should be offered by evolutionist biology, atomic physics or neurobiology. There is no way phenomenology could solve such problems as it rejects any naturalistic approach of the world. Moreover, phenomenology is accused of belonging to the foundationalist tradition to the extent that, "Husserl is trying to find the conditions of knowledge and certainty, Heidegger is trying to find the conditions of intelligibility" (Searle, 2001, p. 89) Searle's conclusion is that phenomenology cannot offer but an incomplete description of how phenomena appear to us, and that only logical analysis can explain the structure of intentionality, as a biological phenomenon, and the conditions of possibility of our practices.

Starting from the positions of phenomenology, Dreyfus shows that it is not enough to consider man only as a biological being, but we need to understand that one of man's essential features is given by his special relationship with the world. This relationship cannot be grasped within the representational terms of the logical analysis of phenomena, which

> as an account of all human intentional behaviour and all functional stuff in the world, it is simply false because it ignores a more basic form of intentionality – ongoing coping – that makes this propositional form of intentionality possible, and a kind of comportment – background coping – the enables one to find one's way about in the world (Dreyfus, 2001, p. 336-337).

Similarly, approaching man and world relationship in the input - output procedural terms, does not offer a satisfying explanation of man coping with the world due to the direct experience of his bodily skills that achieve spontaneous knowledge of the environment he lives in. Consequently, it is only with the help of phenomenology that we can grasp the non-propositional dimension resulting from the dynamic interaction between the body and the environment, which cannot be formalised and which will never be able to be processed by a computer.

Acknowledgements

References

DREYFUS, H.L. 1972. *What Computers can't Do: A Critique of Artificial Reason.* New York, Harper & Row Publishers, 259 p.

DREYFUS, H.L. 1993. Heidegger's Critique of Husserl's (and Searle's) Account of Intentionality. *Social Research*, **60**(1):17-38.

DREYFUS, H.L. 2001. Reply to John Searle. *In:* M.A. WRATHALL; J. MALPAS (eds.), *Heidegger, Coping, and Cognitive Science: Essays in Honor of Hubert L. Dreyfus.* Cambridge, MIT Press, p. 323-337.

DREYFUS, H.L. 2009. How Representational Cognitivism Failed and is being replaced by Body/World Coupling. *In:* K. LEIDLMAIR (ed.), *After Cognitivism: A Reassessment of Cognitive Science and Philosophy.* Dordrecht, Springer, p. 39-75. http://dx.doi.org/10.1007/978-1-4020-9992-2_3

DREYFUS, H.L.; DREYFUS, S.E. 1999. The Challenge of Merleau-Ponty's Phenomenology of Embodiment for Cognitive Science. *In*: G. Weiss; G. HABER; H.F. HABER, *Perspectives on Embodiment: The Intersections of Nature and Culture.* New York, Routledge, p. 103-121.

33

MERLEAU-PONTY, M. 2005. *Phenomenology of Perception.* New York, Routledge, 544 p.

SEARLE, J. 1983. *Intentionality: An Essay in the Philosophy of Mind.* London, Cambridge University Press, 278 p. http://dx.doi.org/10.1017/CBO9781139173452

SEARLE, J. 1984. *Minds, Brains and Science: The 1984 Reith Lectures (Can Computers Think?).* London, British Broadcasting Corporation, 102 p.

SEARLE, J. 1990. Is the Brain's Mind a Computer Program? *Scientific American,* **262**(1):26-31. http://dx.doi.org/10.1038/scientificamerican0190-26

SEARLE, J. 1992. *The Rediscovery of the Mind.* Cambridge, MIT Press, 270 p.

SEARLE, J. 1999. *Mind, Language & Society: Doing Philosophy in the Real World.* London, Weidenfeld & Nicolson, 175 p.

SEARLE, J. 2001. The Limits of Phenomenology. *In:* M.A. WRATHALL; J. MALPAS (eds.), *Heidegger, Coping, and Cognitive Science: Essays in Honor of Hubert L. Dreyfus.* Cambridge, MIT Press, p. 71-93.

SEARLE, J. 2004. *Mind: A brief Introduction.* Oxford, Oxford University Press, 224 p.

34