Noname manuscript No.
(will be inserted by the editor)

Epistemology of AI Revisited in the Light of the Philosophy of Information

Jean-Gabriel Ganascia

Received: date / Accepted: date

Abstract Artificial intelligence has often been seen as an attempt to reduce the natural mind to informational processes and, consequently, to naturalize philosophy. The many criticisms that were addressed to the so-called "old-fashioned AI" do not concern this attempt itself, but the methods it used, especially the reduction of the mind to a symbolic level of abstraction, which has often appeared to be inadequate to represent the richness of our mental activity. As a consequence, there were many efforts to evacuate the *semantical models* in favor of elementary physiological mechanisms simulated by information processes. However, these views, and the subsequent criticisms against artificial intelligence that they contain, miss the very nature of artificial intelligence, which is not reducible to a "science of the nature", but which directly impacts our culture. More precisely, they lead to evacuate the role of the semantic information. In other words, they tend to throw the baby out with the bath-water. This paper tries to revisit the epistemology of artificial intelligence in the light of the opposition between the "sciences of nature" and the "sciences of culture", which have been introduced by German neo-Kantian philosophers. It then shows how this epistemological view opens on the many contemporary applications of artificial intelligence that have already transformed – and will continue to transform – all our cultural activities and our world. Lastly, it places those perspectives in the context of the philosophy of information and more particularly it emphasizes the role played by the notions of *context* and *level of abstraction* in artificial intelligence.

Keywords Artificial Intelligence · Epistemology · Philosophy of Information · Humanities · Sciences of the Nature

Jean-Gabriel Ganascia
LIP6, University Pierre and Marie Curie (Paris VI)
4, place Jussieu, 75252, Paris, Cedex 05, France
Tel.: +33 (0) 1 44 27 37 27
Fax: +33 (0) 1 44 27 70 00
E-mail: Jean-Gabriel.Ganascia@lip6.fr

1 Introduction

Traditional Artificial Intelligence (AI) has often been accused of failing to deliver on its promises. Some people currently say that there is something wrong in what they condescendingly call the "Good Old-Fashioned" symbolic AI that has been accused of oversimplifying the world. More precisely, it has been said that the reproduction of high level cognitive abilities, for instance mathematics, reasoning or playing chess, were easier, from a computational point of view, but less valuable than the simulation of basic physiological mechanisms of perception and action. Frequently invoked by specialists of robotics and AI in the last 20 years, "Moravec's Paradox" (Moravec, 1988) summarizes this point: it claims that cognitive abilities, which require *ratiocination*, are easier to simulate on a computer with a few logical rules than low level cognitive abilities, like perception. For instance, nowadays, there are artificial intelligence programs playing chess, proving theorems or interpreting natural language queries. However, "low level" cognitive abilities, such as the ability to recognize faces or to clean the dishes, seem to be much more difficult to implement than those intellectual faculties. Similarly, basic animal behaviors, e.g. the capacity to perceive or awareness, seem very difficult to reproduce using logical and deterministic mechanisms. This is paradoxical, because the higher intelligence activities, which are proper to humans, seem to be easier to reproduce with classical AI techniques than the basic physiological mechanisms that almost all species possess.

As a consequence, many of those who mock and criticize traditional AI, which is restricted to the simulation of high level cognitive abilities, promote what they call a "Nouvelle AI" that would effectively mimic physiological processes. More generally, they propose to build efficient machines by increasing the complexity of the models and by designing powerful mechanisms that reproduce basic animal capacities (Brooks, 2002).

Obviously, it is very difficult to conclude the effective failure of AI, because, as we shall see further on in this article, compared to many other contemporary disciplines, AI has achieved many successes. Careful attention to the criticisms addressed to traditional AI shows that this disapproval is not really caused by so-called failures or by a non fruitfulness of AI, but by a philosophical divergence that implies a difference of attitude towards intelligence and science. Symmetrically, one can note that, since its initial promotion, more than twenty-five years ago, in the mid-eighties, the "Nouvelle AI" did not so much contribute to successful achievements, which would justify its claims unquestionably. Lastly, the conceptual basis of the "Nouvelle AI", which is rooted on *cybernetics* and *dynamic system theory*, are older than those of traditional Artificial Intelligence.

It therefore would seem that many of the accusations against AI are due to a misunderstanding of its project and are not caused by intrinsic weaknesses of AI techniques. Behind this misunderstanding, there is confusion on the epistemological status of Artificial Intelligence. This article constitutes an attempt to elucidate those points. It first shows, as we have just suggested, that most of the critics that were traditionally addressed against symbolic AI aren't justified. It then tries to elucidate the philosophical status of traditional AI by reference to both the pioneers of artificial intelligence and some philosophical works, in particular to Floridi's *Philosophy of Information* (PI) (Floridi, 2010) and to the traditional distinction between the "Sciences of Nature" and the "Sciences of Culture" that was introduced by the Neo-Kantian school of philosophy. Lastly, it demonstrates how it is different from both the "Nouvelle AI" philosophical status, which is no more than a promotion of an old materialist philosophy, and first cognitivism.

2 The Paradoxical Failure of AI

As previously noted, it is a commonplace nowadays to say that the “Good old-fashioned AI” has failed to deliver its promises. And, there is nothing more characteristic of a philosophical attitude than to analyze and to discuss commonplaces. Nevertheless, in the present case, it’s risky, since this commonplace is admitted by almost everybody without any discussion. Moreover, it could be wrongly interpreted and viewed as a defense of a corporation, while it is the philosophical scope of those criticisms that is of interest here. Our purpose is to show that there is some confusion behind the way AI is understood. More precisely, depending on the philosophical perspective that is adopted, AI may be envisaged differently. But, before investigating this point in more depth, let us recall the numerous successes of AI.

2.1 The Successes of AI

Let us first recall that AI is an academic discipline that was born in 1956 during the *Dartmouth Summer Research Conference on Artificial Intelligence*. This event was organized by a young logician who was less than thirty years old, John McCarthy. The proposal was signed by four persons: John McCarthy, Marvin Minsky, Nathaniel Rochester and Claude Shannon. It was explicitly based on “*the conjecture that every aspect of learning or any other feature of intelligence can in principle be so precisely described that a machine can be made to simulate it.*” (McCarthy et al., 1955). Accordingly, the original AI project was not to reify an intelligent being with a machine, but to decompose intelligence into numerous *features* and to simulate each of them with a computer. Note that, despite the fact that John McCarthy, the organizer of the Dartmouth Summer Research Conference on AI, was a logician, the proposed techniques to achieve this objective were not at all restricted to symbolic computation. For instance, in the McCarthy proposition, it was said (cf. (McCarthy et al., 1955)) that among the different aspects of the AI problem included:

1. Automatic Computers
2. How Can a Computer be Programmed to Use a Language
3. Neuron Nets
4. Theory of the Size of a Calculation
5. Self-Improvement
6. Abstractions
7. Randomness and Creativity

Now, let us examine some of the AI achievements more than fifty years after its birth. There were many successful attempts to reproduce the ratiocination. For instance *Automatic Theorem Proving* has been so well developed that now many mathematical activities have evolved, due to the introduction of computers that partially automate the proof or the proof checking. It even appears that *Mathematical Logic* has been strongly influenced by the development of artificial intelligence: new research areas have been elaborated to answer to the AI needs, for instance the *Non-Monotonic Logics*, the *Description Logics* or the numerous *Logics for Agents*. And, the reproduction of ratiocination is not restricted to mathematics: everyone knows that nowadays world chess champions are all defeated by machine. However, AI is not restricted to abstract logical and mathematical reasoning. *Natural Language Processing* has been intensively developed with many successful applications. It is not only to make the machine understand or translate natural language texts, but also to tag parts of speech, to extract meaningful patterns and to improve our knowledge of many syntactic and semantic phenomenon. Note that the extensive use of *Data Mining* and *Knowledge Extraction* has transformed the corpus linguistic, making possible the process of enormous quantities of texts. Perception has also been simulated with captors and pattern recognition techniques: it is now possible to understand speech, to identify visual patterns and to design robots able to perceive and to navigate using their own representation of the environment.

AI techniques also play a key role in the simulation of *memory*. Let us recall that the invention of the *hypertext* by Ted Nelson in 1965, which was designed as an external memory in reference to Vanevar Bush's MEMEX, was directly influenced by the development of AI techniques, in particular by list processing. It explains the title of Ted Nelson's seminal paper (Nelson, 1965), *A File Structure for The Complex, The Changing and Indeterminate*. Later on, in the seventies, the modeling of semantic memory, e.g. semantic networks (Quillian, 1968), took advantage of AI techniques. In return, semantics helped to design new knowledge representation techniques, e.g. *frames* (Minsky, 1975), which constitutes the semantic turn of AI. More recently, the web was designed by Tim Berners-Lee as a model of memory and then the *Semantic Web* (Berners-Lee et al., 2001) made an extensive use of AI Knowledge Representation techniques. For instance, the notion of *Resource Description Framework Schemas* is nothing more than semantic networks applied to describe meta-data of web resources.

From a practical point of view, AI greatly contributed to contemporary *information society* and, consequently to what Luciano Floridi calls the *Fourth Revolution* (Floridi, 2008, 2010). Let us take two examples. The first concerns *object oriented languages*, which are nowadays the most frequently used programming languages: they are directly inspired from AI semantic knowledge representation techniques. The second concerns the web. As previously said, it is a model of memory built on *hypertext* links, which are directly influenced by AI. And more recently the evolution of the web towards the *Semantic Web* using *ontologies*, *description logics* and other *Knowledge Representation* techniques shows how strong is the AI influence. In the same way, the notion of ambient intelligence refers also to AI.

As a conclusion, it is difficult to imagine a discipline whose concepts so strongly contributed to the evolution of society in less than half a century. Therefore, the mention of a so-called failure of AI and the condescending manner in which one refers to "Good Old-Fashioned AI" are a little bit paradoxical.

2.2 The Misunderstanding

We would like to understand the paradox according to which so many people argue against AI, while, as previously said, this academic discipline was both so successful and so influential in the contemporary information society. Undoubtedly, AI is a very fascinating project, which makes people simultaneously very excited and suspicious.

But, while this project is very ambitious, it is also ambiguous. Depending of the significance of the locution *Artificial Intelligence*, its aim can be interpreted differently: it is either to build an intelligent being using information technologies or to decompose intelligence and to describe precisely each of its components in a way that can be simulated on a computer.

In its first significance AI is understood as an attempt to reproduce consciousness or, at least, to reify the mind with a machine. This first significance is widespread and very popular. It made AI enjoyable. From a philosophical point of view, it means that our consciousness – or our mind – is nothing more than a sum of mechanisms, which satisfies most of our contemporaries. But, this significance is perforce very disappointing, because it proposes to produce a machine that owns a kind of consciousness, or at least of mind, which is too ambitious to be achieved in the short term. Even if the mind or the consciousness can, in principle, be simulated with a the machine, this simulation constitutes a horizon of possibility that is not attainable very soon. As a consequence, AI inevitably fails to fulfill the hopes of impatient persons. Regularly, people argue against AI saying that AI is not able to attain its objectives. Some say that the objectives of AI are not attainable; others say that the methods of AI are not appropriate. The so-called “Nouvelle AI” explains that the failure is caused by the restriction of the mind to symbolic manipulation, without paying attention to physiology and to the body part of intelligence.

In its second significance, AI is a scientific discipline that studies how intelligence can be so precisely decomposed in its different aspects that each of them can be reproduced on computers. This meaning corresponds to the definition given by John McCarthy in the above mentioned project for the *Dartmouth Summer Research Conference* in 1956 (cf. (McCarthy et al., 1955)). According to this second significance, AI investigates intelligence with resources of the artificial (cf. (Simon, 1996)), i.e. with *information processing* techniques, without having to produce an artificial mind. So doing, AI is what Herbert Simon named a *Science of the Artificial*; it is not restricted to what Allen Newell described as *Physical Symbol Systems* (Newell, 1980) and to psychological simulation; there may be many other levels of abstraction that can be simulated using AI techniques. In this respect, it is in total accordance with the principle of the Floridi’s *Philosophy of Information* (Floridi, 2010). Note, for instance, that *Neural Nets* were explicitly mentioned in the McCarthy project (McCarthy et al., 1955), which means that AI is not restricted to a naive Cartesian dualism, as some detractors want to let believe.

Clearly, AI was very successful and the results that were obtained greatly contributed to the fast development of information technologies. And, this discipline is always active and continues to produce new scientific results that will be useful in the future. However, understood as a scientific discipline that attempts to elucidate and simulate the intelligence with machines, AI is far less fascinating in its first significance. This explains why the debate about AI is mainly focused on the first significance, which is a pity, because many questions that are of interest from a philosophical point of view, and especially from the point of view of the *Philosophy of Information*, concern the possibility and the methods of AI understood as a scientific discipline. It does not mean that the discussion about the reduction and/or the reproduction of the mind and/or the consciousness with computers are not justified. However, even if these issues are more commonly discussed, they are far less current, because the practical results of AI understood as a *Science of the Artificial* strongly contributes to transform our contemporary world, while there is no evidence that it is already possible to very soon reproduce a mind or a consciousness with a computer, or with any other contemporary technology.

3 What Went Wrong in AI?

Two years ago, a special issue of AI Magazine (Shapiro and Goker, 2008) published numerous of cases of alleged faulty AI systems. The goal was to understand what made them wrong. The main lesson was that, most of the time, the difficulties were not due to technical impediments, but to the social inadequacy of the AI systems to their environment. This point is crucial. It has motivated the reflexion presented in this paper.

3.1 Elves Keep You

For the sake of clarity, let us take an example about the so-called electronic "elves", which are personal agents who act as efficient secretaries and help individuals to manage their diary, to fix appointments, to find rooms for meetings, to organize travel, etc. A paper (Knoblock et al., 2008) published in the above mentioned special issue of the *AI Magazine* (Shapiro and Goker, 2008) reported technical successes but difficulties with some inappropriate agent behaviors. For instance, one day, or rather one night, an elf rang his master at 3 am to inform him that his 11 o'clock plane was going to be delayed. Another was unable to understand that his master was not available for anybody in his office, since he had to complete an important project... Many of these inappropriate actions make those intelligent agents tiresome and a real nuisance, which causes their rejection by users.

3.2 An Embarrassing Investment Adviser

A few years ago, I was a consultant for a large French bank. The management wanted to introduce knowledge technologies in the company's culture. The reason was that the managers complained that in bank agencies, people in charge of helping customers were unable to provide relevant expert advices because they were only familiar with two or three products among the full range of available solutions. As a consequence, they systematically advised the products they knew, forgetting the others, even when they were more appropriate. The managers thought that a knowledge-based system could advantageously replace – or possibly train – those poor investment advisers. This is why they got in touch with my group who they asked to build a Knowledge Based-System able to act as an efficient adviser that helps customers to invest their money.

My group succeeded in building an efficient "investment adviser" by using the knowledge engineering techniques that were in used at that time. The resulting system asked relevant queries, diagnosed the situation of the customers and provided, for each of them, eligible, diversified and judicious investments that take advantage of all the products proposed by the bank. From a technical point of view, it seemed that it gave entire satisfaction.

However, the system has never been in use for two reasons. The first was the refusal from the bank agency managers: they feared being reduced to a simple role of performers.

The second came from customers who suspected the AI systems were provided by the bank to serve the interests of the bank. Note that surprisingly they were not so much suspicious of the bank employees nor the bank softwares, but rather the AI systems provided by the bank.

3.3 The Social Dimension of AI

Those two examples show that social inadequacy is the main cause of AI system rejection. In both cases, the AI programs were technically successful; they were not accepted because they did not answer to the requirements of the social environment where they had to be used. The causes of inappropriateness were not in the artificial system itself, but in the adequacy of the artificial system to the surroundings.

This conclusion is neither astonishing nor original. Many people have noticed that the failures of knowledge-based systems were mainly due to man-machine interfaces or to organizational impediments, which made them inefficient (cf. for instance (Hatchuel and Weil, 1995)). Moreover, it is in accordance with what the pioneers of AI had said, and in particular Herbert Simon who has insisted on the importance of the outer environment in his famous book "The Science of the Artificial" (Simon, 1996): according to him, "*Human beings, viewed as behaving systems, are quite simple. The apparent complexity of our behavior over time is largely a reflection of the complexity of the environment in which we find ourselves.*" In other words, the difficulty would not be in reproducing intelligent behaviors, but in adapting them to the complexity of their environment.

3.4 Why Would Something be Wrong?

These conclusions are so obvious and conform with the predictions that the above mentioned AI failures would have had an incentive to address both user-centered design and social studies. Nevertheless, surprisingly, since the eighties, the evolution of AI toward, for instance, the so-called "Nouvelle AI" has gone in a completely different direction: because it has been accused of oversimplifying the world and of ignoring the physical bodies, AI has been tempted to increase the complexity of its models and to build powerful machines able to effectively mimic physiological capacities.

This view tends to reduce AI to a simulation of the natural processes. It opens undoubtedly exciting prospects for scientists. However, as we shall see in the following section and as it was previously mentioned, this does not exhaust the project of AI, which cannot be fully assimilated to a pure reproduction of the cognitive abilities, i.e. to a "naturalization" of the mind. In other words, this project reduces AI to only one of its significances, i.e. to the production of an artificial mind, or of an artificial consciousness. However, it does not address the second significance of AI that is the discipline that investigates intelligence by simulating it with artificial devices.

In a way, the opposition between those two significances of AI reflect an old opposition, introduced in the beginning of the 20^{th} century, by neo-Kantian philosophers, between the "Sciences of nature" and the "Sciences of culture", i.e. the *humanities*.

Herbert Simon himself introduces the "Sciences of the artificial" to qualify the type of investigation that motivated AI, which suggests exploring more in depth this traditional opposition to better understand the epistemological status of AI understood as a scientific discipline and, subsequently the status of the Simon's "Sciences of the artificial".

4 "Artificiality" vs. "Culturality"

4.1 The "Sciences of the Artificial"

Herbert Simon has introduced the distinction between the "Sciences of Nature" and the "Sciences of the Artificial" in a famous essay published in 1962 in the "Proceedings of the American Philosophy Society". The question was of importance for him, since he worked for more 35 years on it, he has re-edited the same book three times in 1969, in 1980 and in 1996, and has considerably augmented the volume of the book: the first edition published in 1969 contained 123 pages, while the third edition, published in 1996, contained 231 pages.

The original point of Herbert Simon was to introduce the notion of artificiality to describe complex artificial systems in complex environments and to make them object of science. According to him, artificial systems have to be distinguished from natural systems, because they are produced by human beings – or, more generally, by intelligent beings – who have in mind some goals to achieve. More precisely, artificial things are characterized by the four following points (Simon, 1996):

1. They are produced by human (or by intelligent beings) activity.
2. They imitate more or less nature, while lacking the whole characteristics of natural things.
3. They can be characterized in terms of *functions*, *goals* and *adaptation*.
4. They can be discussed both in terms of imperatives or as descriptives.

Remark that the universe of artificial things is not reduced to the computerized world. Many artificial objects that were invented far before the existence and the development of electronic computers, for instance airplanes and clocks, own all the above mentioned characteristics. However, computers greatly facilitate the building of artificial things.

Since artificial things can be approached not only in descriptive terms of their structure, but with respect to their functions, their goals and their adaptive abilities, they cannot be reduced to natural things that have only to be objectively described from the outside, without any *a priori*. Their study can take into consideration the imperatives to which they are supposed to obey. As a consequence, the discipline that is in charge to study artificial things, i.e. the science of the artificial things, has to be distinguished from the sciences of the natural things. To characterize this discipline, Herbert Simon has introduced the concept of "artifact", which is defined as an interface between the "inner" environment, i.e. the internal environment of an agent, and the "outer" environment where it is plunged. As previously said, the "inner" environment is easy both to describe in terms of functions, goals and adaptation and to simulate with computers; its complexity results from the "outer" environment in which it operates.

It has to be recalled that artificial things can always be studied with the methods of the "sciences of nature", for instance a clock can be studied from a physical point of view, by analyzing the springs and the wheels it is composed of, but those "sciences of nature" don't take into consideration the imperatives to which artificial things are supposed to obey, their functions and their goals.

Symmetrically, natural things can be investigated by the "sciences of the artificial". More precisely, according to Herbert Simon, the "sciences of the artificial" can greatly help to improve our knowledge of the natural phenomenon. Any natural thing can be approached by building models, i.e. artificial things, that aim at simulating some of their functions. For instance, cognitive psychology has been very much improved by the use of computers that help to simulate many of our cognitive abilities.

4.2 Limits of the Artificiality

Two criticisms can be addressed to AI understood as a "science of the artificial".

The first is traditional and recurrent: for more than 20 years now, scientists and philosophers criticize the oversimplified models of the so-called "old-fashioned AI". In a word, they think that models have to be exact images of what they are intended to model. As a consequence, the "artifacts", taken in Herbert Simon terms, i.e. the interfaces between "inner" and "outer" environments, have no real value when the "inner" environments are too schematic. Therefore, the artificiality has to faithfully copy the reality, i.e. nature. As a consequence, many mental and social phenomenon are viewed as natural phenomenon. For instance, the mind is reduced to physical phenomenon that result from brain activity (Manzotti, 2007) or the epistemology is identified to informational processes (Chaitin, 2006). The AI itself has been mathematized by physicists as a unified and universal theory (Hutter, 2005), which gave birth to the *General Artificial Intelligence*. This tendency corresponds to the so-called "naturalization", which is very popular nowadays among philosophers (Dodig-Crnkovic, 2007). Nevertheless, despite the huge amount of researches done in this area for many years now, only a few results have been obtained.

The second criticism is symmetric: the notion of "artifact" does not allow to fully approach the semantical and cultural nature of all mental processes. For instance, Herbert Simon considers music as a science of the artificial, since everything that is said about the sciences of the artificial can be said about music: it requires formal structures and provokes emotions. It is partially true, however, music is not only a syntax; semantical and cultural dimensions of music exist and they are not taken into account in Simon models. Therefore, we pretend that an extension of the "science of the artificial" toward the "sciences of culture" is required.

In other words, while the first criticism opens on a naturalization, i.e. on a refinement of the models, the second pursues and extends the Herbert Simon "sciences of the artificial" by reference to the Neo-Kantian "sciences of culture" that will be presented in the next section.

5 The "Sciences of the Culture"

5.1 Origin of the "Sciences of the Culture"

The notion of "Sciences of the Culture" (Rickert, 1921) was introduced in the beginning of the 20^{th} century by a German Neo-Kantian philosopher, Heinrich Rickert who was very influential on many people among which were the sociologist Max Weber and the young Martin Heidegger. Its goal was to base the humanities, i.e. disciplines like historic studies, sociology, laws, etc., on rigorous basis. More precisely, he wanted to scientifically characterize the sense of human activities, i.e. culture understood as the result of goal oriented activities. In other words, he wanted to build an empirical science able to interpret human achievements as the results of mental processes. However, he thought that the scientific characterization of the mind had to be distinguished from the psychological science, i.e. from the psychology, which approached the mental phenomenon with the methods of physical sciences. For him, spiritual phenomenon have a specificity that cannot be reduced to a physical one, even if they can be submitted to a rational and empirical inquiry. The distinction between "sciences of nature" and "sciences of culture" had to precisely establish this specificity. As we shall see in the following, according to Rickert, the underlying logic of the "sciences of culture" totally differs from the logic of the "sciences of nature".

Before going more into the detailed characterization of those approaches, let us add that "sciences of culture" have nothing to do with "cultural studies": the former attempt to scientifically characterize the results of human conscious activities – politics, art, religion, education, etc.– while the latter try to identify and to differentiate cultural facts from various manifestations of human activities – dances, musics, writings, sculpture, etc.–. Very often cultural studies aim at exploring the cultural specificities and their conflict with official cultures and powers that tend to ignore them. As already said, the notion of "sciences of culture" was introduced in the early 20^{th} century, while "cultural studies" only exist since the sixties. Lastly, "sciences of culture" do not promote culture as the expression of identities, while "cultural studies" are often advocate of such expression.

As previously mentioned, the "sciences of culture" aim at understanding social phenomenon that result from human conscious activities. Obviously, physics and chemistry are out of the scope of the "sciences of culture" because they investigate the objective properties of the world, without any interference with human activities. On the contrary, the study of religion and discrimination may participate to the "sciences of the culture". But, the distinction is not so much a difference in the objects of study than in the methods of investigation. Therefore, the *history of physics* contributes to the "sciences of culture" while some mathematical models of social phenomenon, e.g. game theory, contribute to "sciences of nature". Moreover, the same discipline may simultaneously contributes to "sciences of nature" and to "sciences of culture"; it is what Rickert characterizes as an intermediary domain. For instance, medicine benefits simultaneously from large empirical studies and from individual case studies; the former enter more likely into the logic of "sciences of the nature" and the latter into the logic of "sciences of culture". It even happens, in disciplines like medicine, that national traditions differ, some of them being more influenced by the "sciences of nature", like *evidence-based medicine*, while others contribute more easily to the "sciences of culture", like *clinical medicine* when it is based on the study of the patient history.

In other words, the main distinction concerns different logics of sciences that are described in the next section.

5.2 The Tree Logics

Ernst Cassirer clearly described the different logics of sciences in many of his essays (Cassirer, 1923, 1961). Briefly speaking, he first distinguishes the theoretical sciences like mathematics, which deal with abstract and perfect entities as numbers, figures of functions, from empirical sciences that are confronted with the material reality of the world. Then, among the empirical sciences, Ernst Cassirer differentiates “sciences of nature”, which deal with physical perceptions, and “sciences of culture” that give sense to the world. According to him and to Heinrich Rickert, “sciences of nature” proceed by generalizing cases: they extract general properties of objects and they determine laws, i.e. constant relations between observations. As a consequence, the logic of “sciences of nature” is mainly inductive, even if the modalities of reasoning may be deductive or abductive. The important point is that the particular cases have to be forgotten; they have to be analyzed in general terms and composed of well defined objects that make no reference to the context of the situation. The validity of the scientific activity relies on the constance and the generality of the extracted laws.

By contrast to the logic of “sciences of nature”, the logic of the “sciences of culture” do not proceed by generalizing multiple cases. It does not extract laws, i.e. relations between observations; it does not even work with physical perceptions, but with meaningful objects that have to be understood. In brief, the main function of “sciences of culture” is to give sense to the world. Their way of investigation is to understand particulars. The general methodology is to observe individual cases and to understand them. However, they have to choose, among the particulars, individuals that are paradigmatic, i.e. who can teach general lessons that may be reused in other circumstances. In other words, “sciences of culture” are not properly interested in the singularity of cases, which has to be forgotten, but in the understandability of individuals under study. Their methods help to give sense to observations of complex individual cases.

5.3 “Science of Culture” vs “Science of Artificial”

As previously said, culture can be understood as the result of goal oriented human activities. For instance, agriculture is the art and practice of working soils to produce crops and other vegetables. The “sciences of culture” try to understand the human activities, i.e. the human goals and the ways humans take to reach them. Since AI tries to reproduce intelligent human activities, it can obviously benefit from the methods of the “sciences of culture”. However, it can also benefit from the theoretical sciences that work on abstract entities, i.e. from mathematics and logic, and from the “sciences of nature”, which, for instance, investigate physiological or physical mechanisms. Looking back to the “sciences of the artificial”, it appears that they belong both to “sciences of nature”, since they proceed by generalization of cases, and to “sciences of culture”, because they characterize artificial things by their functions, their goals and their adaptivity and not only by their structure.

The next section shows how methods of “sciences of culture” can play an important role in AI, even if AI cannot be reduced to a “science of culture”. Nevertheless, the important point here concerns the distinction between the “sciences of the artificial” and the “sciences of culture”. As previously said, the artificiality, taken in the sense given by Herbert Simon, includes not only the things that are produced by the activity of intelligent beings, but also the goals to which they are designed for. Human productions are not reducible to the material things they achieved. For instance, a statue is more than the bronze it contains; a clock is more than the metal it is made of; a book is more than paper and ink, etc. As a consequence, artificiality is also part of the culturality. The sciences that produce artifacts, i.e. the “sciences of the artificial” are undoubtedly part of "sciences of culture", while culture covers a broader area since it also includes pure interpretative activities like history. Moreover, the logic of "sciences of culture" extends the logic used in the "sciences of the artificial" that remains partially similar to the logic of "sciences of nature".

6 AI as an intermediary domain

The thesis developed here is that the alleged AI weaknesses are not caused by the oversimplification of AI models, as many people claim nowadays, but by their inadequacy to the "outer" environment. It has been shown that the notion of "science of the artificial", which was introduced by Herbert Simon, has to be extended by reference to the notion of "science of culture".

From a philosophical point of view, it means that AI participates to the "sciences of culture", i.e. that it cannot be entirely reducible to a "science of nature" or to mathematics and theoretical sciences. But it is not more reducible to "sciences of culture". More precisely, it is what Heinrich Rickert identifies as an "intermediary domain" that belongs simultaneously to theoretical sciences, i.e. to formal logic and mathematics, to empirical sciences of nature and to empirical sciences of culture. The practical consequences of such philosophical considerations are twofold: they have an impact on both the methods and the objects of application of AI.

6.1 Methods of AI

Since AI contributes to "sciences of culture", it has to take advantage of the logic of "sciences of culture", which may enlarge the scope of its methods. Let us recall that "sciences of culture" are empirical sciences, i.e. they build knowledge from the observation of particulars. However, they don't proceed by extracting properties common to observed cases; they do not abstract knowledge from particulars. They collect data about individual cases and they attempt to understand them, i.e. to find a common cause or to give a reason for them. Let us specify that it is not to extract singularities, but to investigate paradigmatic cases and to explain in what respect the individual cases under study can be universalized.

An excellent example of such studies was done by a cognitive anthropologist, Edwin Hutchins, in the book titled "Cognition in the wild" (Hutchins, 1995) where he attempted to identify the cognition in its natural habitat, in the circumstances a modern ship, and to model it. In practice, many preliminary studies should have recourse to such methods. It has to be the case with knowledge engineering and, more generally, when designing any AI concrete application.

Moreover, the attentive study of past failures contributes to this dimension of AI. It is not to generalize all the individual failures by extracting their common properties, as could be for any "science of nature", but to understand the logic of the failures, as did, for instance, Dietrich Dörner in his book *"The Logic of Failure"*(Dörner, 1997), to see what lessons could be drawn from these bad experiences and to learn from them. In this way, it could contribute to the logic of "sciences of culture".

6.2 Objects of AI

Lastly, the investigations of AI could focus more deliberately on cultural dimensions of the world, where there are many valuable applications. The information sciences and technologies greatly contribute to the advancement of knowledge to the point where the present age is often called the "knowledge age". However, it's a pity that AI did not participate more actively in cultural evolutions consecutive to the development of information technologies, for instance, to the Wikipedia free encyclopedia or to the social web.

6.3 Perspectives for AI

More generally, the knowledge quest can be greatly accelerated by the use of AI technologies. For instance, my team is working in musicology (Rolland and Ganascia, 2002, 1999; Ramalho et al., 1999), in textual criticism (Bourdaillet et al., 2007), in social sciences (Velcin and Ganascia, 2005), in epistemology (Ganascia and Debru, 2007; Ganascia, 2008), in ethics (Ganascia, 2007) etc. But there are many other fields of applications, not only in humanities. Let us insist that such applications of AI are directly connected with cultural dimensions. So, in the case of medicine, there already exist many attempts to model organs (Nobel, 2006) and to simulate medical diagnosis; AI played a part in these successful achievements, which are related to "sciences of nature"; but the new challenge now is to manage all the existing knowledge and to help researchers to find their way. This is undoubtedly the role of AI understood as a "science of culture" to help to achieve such tasks.

7 Conclusion and perspective

To conclude, let us first insist on our main point: AI can neither be reduced to a "Science of nature" nor to a "Science of culture"; it is what Rickert calls an "Intermediary domain". This has not only philosophical implications on the epistemological status of AI, but also practical consequences about both the objects and the methods of AI. Moreover, the reduction of AI to a "Science of nature" does not allow to understand the role it plays in the development of our *Information Society*. The concept of *knowledge* as it is commonly used today to qualify the present state of our societies does not only refer to the democratization of education or to the high qualifications that are required in a modern economy, but also to the formalization of interpretation processes, which render possible the storage, the access and the exchange of knowledge. For instance, the notion of *ontology* as it was developed in AI during the last few years takes its sense in the context of a "Science of culture", i.e. with respect to interpretation processes, but not with respect to a "Science of nature". To be convinced, let us quote Tom Gruber, one of the most influential persons in the field of ontology design in AI, who said in an interview that: "Every ontology is a treaty – a social agreement – among people with some common motive in sharing." (Gruber, 2004). More generally, the way AI attributes meaning to symbols, i.e. the semantic in AI, does not refer to a "Science of nature", but to a "Science of culture". To this respect, the notion of *Knowledge Level*, which was introduced by Alan Newell in 1982 (Newell, 1982) and which was so influential – and so controversial – in the field of *Knowledge Acquisition* in AI (Clancey,

1993) during the nineties, is illustrative: it does not reduce knowledge to symbols or to information, but it makes knowledge the result of an interpretative process. More generally, it refers knowledge to a specific *Level of Abstraction* which takes sense in *Context*.

Everything which has been said here concerning AI is also valid for most of automatic information processes. As an illustration, the way semantic information is extracted from data can neither be reduced to the sole induction, i.e. to a generalization from particulars, nor to a representation in a universal digital ontology. The knowledge, which is relevant semantic information, takes its sense within interpretative processes, at a *Level of Abstraction* and in a given *Context*, i.e. with respect to the key concepts of the Philosophy of Information (Floridi, 2010). More generally, most of the open problems of Philosophy of Information can be enlightened by being envisaged under the light of the opposition between the "Sciences of nature" and the "Sciences of culture".

References

Berners-Lee, T., Hendler, J., and Lassila, O. (2001). The semantic web. *Scientific American Magazine*.

Bourdaillet, J., Ganascia, J.-G., and Fénoglio, I. (2007). Machine assisted study of writers' rewriting processes. In *Proceedings of the 4th International Workshop on Natural Language Processing and Cognitive Science (NLPCS)*.

Brooks, R. (2002). *Flesh and Machines: How Robots Will Change Us*. Pantheon Books.

Cassirer, E. (1923). *Substance and Function*. Open Court, Chicago.

Cassirer, E. (1961). *The Logic of the Humanities*. Yale University Press, New Haven.

Chaitin, G. (2006). Epistemology as information theory. *COLLAPSE*, 1:27–51.

Clancey, W. (1993). The knowledge level reinterpreted. In Ford, K. M. and Bradshaw, J. M., editors, *Knowledge acquisition as modeling*. John Wiley & Sons.

Dodig-Crnkovic, G. (2007). Epistemology naturalized: The info-computationalist approach. *American Philosophy Association Newsletter on Philosophy and Computers*, 06(2).

Dörner, D. (1997). *The Logic Of Failure: Recognizing And Avoiding Error In Complex Situations*. A Merloyd Lawrence Book, Perseus Books, Cambridge, Massachusetts.

Floridi, L. (2008). Artificial intelligence - new frontier: Artificial companions and the fourth revolution. *Metaphilosophy*, 39(4/5):651–655.

Floridi, L. (2010). *The Philosophy of Information*. Oxford University Press, Oxford.

Ganascia, J.-G. (2007). Modeling ethical rules of lying with answer set programming. *Ethics and Information Technology*, 9:39–47.

Ganascia, J.-G. (2008). In silico' experiments : Towards a computerized epistemology. *Newsletter on Philosophy and Computers, American Philosophical Association Newsletters*, 7 (2):11–15.

Ganascia, J.-G. and Debru, C. (2007). Cybernard: A computational reconstruction of claude bernards scientific discoveries. In *Studies in Computational Intelligence (SCI)*, pages 497–510. Springer-Verlag.

Gruber, T. (2004). Every ontology is a treat. *AIS SIGSEMIS Bulletin*, 1(3). Tom Gruber interview.

Hatchuel, A. and Weil, B. (1995). *Experts in Organizations: A Knowledge-Based Perspective on Organizational Change*. Walter de Gruyter, Berlin-New-York.

Hutchins, E. (1995). *Cognition in the Wild.* MIT Press, Cambridge, Massachusetts.

Hutter, M. (2005). *Universal Artificial Intelligence: sequential Decisions Based on Algorithmic Probabilities.* Springer.

Knoblock, C., Ambite, J., Carman, M., Michelson, M., Szekely, P., and Tuchinda, R. (2008). Beyond the elves: Making intelligent agents intelligent. *AI Magazine*, 29(2):33–39.

Manzotti, R. (2007). Towards artificial consciousness. *American Philosophy Association Newsletter on Philosophy and Computers*, 07(1):12–15.

McCarthy, J., Minsky, M. L., Rochester, N., and Shannon, C. (1955). A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence. Technical report, Dartmouth College.

Minsky, M. (1975). A framework for representing knowledge. In Winston, P. H., editor, *The Psychology of Computer Vision.* McGraw-Hill, New York (U.S.A.).

Moravec, H. (1988). *Mind Children.* Harvard University Press, Cambridge, Massachusetts.

Nelson, T. (1965). A file structure for the complex, the changing and indeterminate. In *Proceedings of Association for Computing Machinery* 20^{th} *National Conference*, pages 84–100.

Newell, A. (1980). Physical symbol systems. *Cognitive Science*, 4:135–183.

Newell, A. (1982). The knowledge level. *Artificial Intelligence Journal*, 18:87–127.

Nobel, D. (2006). *The Music of Life: Biology Beyond the Genome.* Oxford University Press.

Quillian, M. R. (1968). Semantic memory. In Minsky, M., editor, *Semantic Information Processing*, pages 227–270. MIT Press, Cambridge, MA.

Ramalho, G., Rolland, P.-Y., and Ganascia, J.-G. (1999). An artificially intelligent jazz performer. *Journal of New Music Research.*

Rickert, H. (1921). *Kulturwissenschaft und Naturwissenschaft.* J.C.B. Mohr (Paul Siebeck), Tubingen, 5th edition.

Rolland, P.-Y. and Ganascia, J.-G. (1999). Musical pattern extraction and similarity assessment. In *Readings in Music and Artificial Intelligence. Contemporary Music Studies*, volume 20. Harwood Academic Publishers.

Rolland, P.-Y. and Ganascia, J.-G. (2002). Pattern detection and discovery: The case of music data mining. In *Pattern Detection and Discovery*, pages 190–198. Springer-Verlag.

Shapiro, D. and Goker, M. (2008). Advancing ai research and applications by learning from what went wrong and why. *AI Magazine*, 29(2):9–76.

Simon, H. A. (1996). *The Sciences of the Artificial.* MIT Press, Cambridge, Massachusetts, 3rd edition.

Velcin, J. and Ganascia, J.-G. (2005). Stereotype extraction with default clustering. In *proceedings of the 19th International Joint Conference on Artificial Intelligence (IJCAI)*, pages 883–888.