# AI viewed as a "Science of the Culture"

Jean-Gabriel Ganascia

LIP6 - University Pierre and Marie Curie, Paris, France

**Abstract.** Last twenty years, many people wanted to improve AI systems by making computer models more faithful to the reality. This paper shows that this tendency has no real justification, because it does not solve the observed limitations of AI. It proposes another view that is to extend the notion of "Sciences of the Artificial", which has been introduced by Herbert Simon, into to a "Science of the Culture"After an introduction and a description of some of the causes of the present AI limitations, the paper recalls what the notion of "Sciences of the Artificial" is and presents the "Sciences of the Culture". The last part explains the possible consequences of an extension of AI viewed as a "Science of the Culture".

## 1 Introduction

AI has sometimes been accused of failing to deliver on its promises. Obviously, it is very difficult to evaluate this point, since, compared to many other contemporaneous disciplines, AI has get many successes. This paper attempts to elaborate a tentative explanation of why AI has been seen as failing by questioning its philosophical foundations and by showing that there is a possible misunderstanding about its status. The starting point concerns the epistemological status on which AI groundings seem to be based. At first sight, it appears that many contemporaneous scientists tend to build AI on a model analogous to the one which the physical or the life sciences are based on. See for instance the recent project of Marcus Hutter [1] who pretends to scientifically ground *Universal Artificial Intelligence* on the Kolmogorov information theory. In the past, AI have often been understood as a "Science of the nature". For instance, the program of the very famous Dartmouth College Summer Research Project on Artificial Intelligence was 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.*". Such strong groundings would have been reassuring. Nevertheless, they are not totally satisfying and there were also some attempts to put in light other dimensions of AI than this reduction to a science of the nature. For instance, the notion of "Sciences of the Artificial" introduced by Herbert Simon [2] in his famous book opened many perspectives to AI. Do those "Sciences of the Artificial" differ from the traditional "Sciences of the Nature" or do they extend and renew them with new contemporaneous tools? Would it be possible today to extend the "Sciences of the Artificial"? On the one hand, most of the time, even for Simon, the aim assigned to AI is to naturalize – or

to computerize, but this is more or less equivalent from a philosophical point of view – social and psychological phenomenon, i.e. to reduce social and psychological phenomenon to mechanical processes that can be simulated on digital computers. On the other hand, the notion of knowledge that was introduced in AI seems not to be reducible to mechanical processes on physical symbols. The so-called *activity* of the "Knowledge Level" that was emphasized by Alan Newell in his famous paper [3] seems to express this irreducibility.

This paper attempts to elucidate the scientific status of AI. Is it only a "Science of the Nature"? Or, does it partially differ from it, as the notion of "Sciences of the Artificial", introduced by Herbert Simon, might suggest? To enlighten the epistemological nature of AI, understood as a "Science of the Artificial", we refer to the opposition between the "Sciences of the Nature" and the "Sciences of the Culture", i.e. the humanities, that was introduced in the first half of the $20^{th}$ century by some German Neo-Kantian philosophers among which he most famous were Heinrich Rickert (1863–1936) and Ernst Cassirer (1874–1945). The paper shows that the conceptual apparatus developed by the above mentioned Neo-Kantian philosophical school may be successfully applied to AI. More precisely, our working hypothesis is that AI can neither be fully reduced to a "Science of the Nature" nor to a "Science of the culture", but that it is what Heinrich Rickert calls an "intermediary domain". It means that both the objects of AI and its logic belong to fields covered simultaneously by the "Sciences of the nature" and by the "Sciences of the culture". The practical consequences of such philosophical considerations concern first the fields of application of AI, i.e. its objects, which cannot be reduced to the sole simulation of the nature, i.e. to a total and perfect reproduction of activity of intelligent beings. The AI influences also the culture, i.e. the medium of communication. Many of AI successes concern the way it changed – or it helped to change – the contemporaneous culture. Unfortunately, those successes have not been credited to AI. The second consequences are about the AI methods that cannot all be assimilated to logical generalizations of the diversity by general laws; for instance, the careful study of paradigmatic past cases, or more precisely of alleged past AI failures, can be valuable; it also belongs to the method of the "Sciences of the Culture", of which AI is a part of.

Apart this introduction, the paper is divided in four parts. The first one is a lesson drawn from the past AI failures. The second constitutes an attempts to explain why AI is supposed to have failed. The third briefly recalls the notion of "Sciences of the Artificial" as it was introduced by Herbert Simon and then describes the Neo-Kantian distinction between the "Sciences of the Nature" and the "Sciences of the Culture". The last one inventories some of the practical consequences of this distinction.

## 2 What Went Wrong?

In March 2006, at the AAAI Stanford Spring Symposium, in a workshop entitled "What Went Wrong and Why Workshop" numerous of cases of alleged AI failures

were presented. They were published last summer in a special issue of the AI Magazine [4]. The main lesson was that, most of the time, the difficulties were not due to technical impediments, but to the inadequacy of the AI systems to their social environment. This point is crucial. It originates the reflexion presented in this paper. For the sake of clarity, let us illustrate this point with two examples, one that was presented in [5] and one from my personal experience.

The first refers to user feedbacks with "elves", which are personal agents that act as efficient secretaries and help individuals to manage their diary, fix appointments, find rooms for meetings, organize travel, etc. The talk [5] reported technical successes but difficulties with inappropriate agent behaviors. For instance, one day, or rather one night, an elf rang his master at 3am to inform him that his 11 o'clock plane was going to be delayed. Another was unable to understand that his master was in his office for nobody, since he had to complete an important project... Many of these inappropriate actions make intelligent agents tiresome and a real nuisance. It causes their uselessness.

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 to aid customer were unable to provide relevant expert advices because they were only familiar with two or three products among the full range of available products. As a consequence, they advised systematically 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 and they asked to build a Knowledge Based-System able to act as an efficient adviser that helps customers to invest their money. My group succeed in building an efficient "investment adviser" by using the knowledge engineering techniques that were in use at this time. The resulting system asked relevant queries, diagnosed the situation of the customer 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 of the bank agency managers: they feared being reduced to a simple role of performers. The second came from customers who suspected the AI systems provided by the bank to serve the interests of the bank. Note that they astonishingly do not suspect so much the bank employees or even the bank softwares than the AI systems provided by the bank.

Those two examples show that social rejection is one of the main causes of AI system failures. In both previous cases, the AI programs were technically successful; they were not accepted because they did not answer to the requirements of the social environment. The causes of inappropriateness was not in the artificial system itself, but in the adequacy of the artificial system to its environment.

This conclusion is neither surprising 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 [6]). Moreover, it is in accordance with the pioneers of AI like Herbert Simon who insisted on the importance of the outer environment in his famous book "The Science of the Artificial" [2]: 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 the environment.

These conclusions are so obvious and conform with the predictions that the above mentioned AI failures would have had an incentive to persevere and 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: AI has been accused of oversimplifying the world. It has been said that the reproduction of high level cognitive abilities, for instance doing 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. The so-called "Moravec's Paradox" [7] summarized this point; it has been frequently invoked by specialists of robotics and AI last 20 years. As a consequence, the proposed solution would be to increase the complexity of the models, which will make us able to build powerful machines that effectively mimic physiological capacities [8]. 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, this does not exhaust the project of AI, which cannot be fully assimilated to a pure reproduction of the observable behavior, i.e. to a "naturalization" of the mind.

## 3 "Artificiality" vs. "Culturality"

### 3.1 The "Sciences of the Artificial"

Herbert Simon has introduced the distinction between the "Sciences of the Nature" and the "Sciences of the Artificial" in a famous essay published in 1962 in the "Proceedings of the American Philosophy Society". 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 [2]:

1. They are produced by human (or by intelligent beings) activity
2. They imitate the nature more or less the 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 computers, for instance airplanes and clocks, own all the above mentioned characteristics. However, computers greatly facilitate the building of artificial things.

Since the 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, Hebert Simon has introduced the concept of "artefact", 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 the 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 the nature" don't take into consideration the imperatives to which the 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.

### 3.2 Limits of the Artificiality

Two critics can be addressed to the 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 "artefacts", 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. the 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 [9] or the epistemology is identified to informational processes [10]. This tendency corresponds to the so-called "naturalization", which is very popular

nowadays among philosopher [11]. Nevertheless, despite the huge amount of researches done in this area for many years now, only a few results have been obtained.

The second critic is symmetric: the notion of "artefact" 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 the culture" is required. In other words, while the first critics 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 the culture" that will be presented in the next section.

## 4 The "Sciences of the Culture"

### 4.1 Origin of the "Sciences of the Culture"

The notion of "Sciences of the Culture" [12] was introduced in the beginning of the $20^{th}$ century by a German Neo-Kantian philosopher, Heinrich Rickert who has been 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. the disciplines like historic studies, sociology, laws, etc., on rigorous basis. More precisely, he wanted to scientifically characterize the sense of human activities, i.e. the culture understood as the result of goal oriented activities. In other words, he tried 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 psychology, which approached the mental phenomenon with the methods of the physical sciences. For him, spiritual phenomenon have a specificity that cannot be reduced to physical one, even if they can be submitted to a rational and empirical inquiry. The distinction between the "sciences of the nature" and the "sciences of the culture" had to precisely establish this specificity. As we shall see in the following, according to Rickert, the underlying logic of the "sciences of the culture" totally differs from the logic of the "sciences of the nature".

Before going more in the detailed characterization of those approaches, let us precise that the "sciences of the culture" have nothing to see with "cultural studies": the first attempt to characterize scientifically the results of human conscious activities – politics, art, religion, education, etc.– while the second 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 the culture" was introduced in the early $20^{th}$ century, while the

"cultural studies" exist only since the sixties. Lastly, the "sciences of the culture" have no more to do with any kind of cultural relativism that justifies unethical behaviors, e.g. the polygamy, as expressions of the identities.

As previously mentioned, the "sciences of the 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 the 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* participates to the "sciences of the culture" while some mathematical models of social phenomenon, e.g. game theory, participate to the "science of the nature". Moreover, the same discipline may simultaneously participate to the "sciences of the nature" and to the "sciences of the 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 first participate more to the logic of the "sciences of the nature" and the second to the logic of the "sciences of the culture". It even happens, in disciplines like medicine, that national traditions differs, some of them being more influenced by the "sciences of the nature", like the *evidence-based medicine*, while others participate more easily to the "sciences of the culture", like *clinical medicine* when it is based on the study of the patient history.

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

### 4.2 The Tree Logics

Ernst Cassirer clearly described the different logics of sciences in many of his essays [13, 14]. 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 the "sciences of the nature", which deal with physical perceptions, and the "sciences of the culture" that give sense to the world. According to him and to Heinrich Rickert, the "sciences of the 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 the "sciences of the 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 as 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 the "sciences of the nature", the logic of the "sciences of the 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 the "sciences of the culture" is to give sense to the world. The general methodology is to observe particular cases and to understand them. However, they have to choose, among the particulars, individuals that are paradigmatic, i.e. which can teach general lessons that may be reused in other circumstances. In other words, the "sciences of the culture" are not 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.

### 4.3 "Science of the Culture" vs "Science of the Artificial"

As previously said, the culture can be understood as the result of goal oriented human activities. For instance, the agriculture is the art and practice of working soils to produce crops and other vegetables. The "sciences of the 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 method of the "sciences of the 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 the nature", which, for instance, investigate physiological or physical mechanisms. Looking back to the "sciences of the artificial", it appears that they mainly belong to the "sciences of the nature", since they proceed by generalization of cases. Nevertheless, the characterization of artificial things by their functions, their goals and their adaptivity make them belong also to the "sciences of the culture".

The next section shows how methods of the "sciences of the culture" can play an important role in AI, even if AI cannot be reduced to a science of the culture. But the important point here concerns the distinction between the "sciences of the artificial" and the "sciences of the culture". As previously said, the artificiality, taken in the sense given by Herbert Simon, includes all the things that are produced by the activity of intelligent beings. The culture covers a broader area, since it also includes all the human activities, which, for many of them, are not reducible to the material things they produces. For instance, a statue is more than the bronze of which it is made of; a book is more than paper and ink, etc. Moreover, the logic of the "sciences of the culture" extends the logic used in the "sciences of the artificial" that remains similar to the logic of the "sciences of the nature".

## 5 Conclusion and perspectives

### 5.1 AI as an intermediary domain

The thesis developed here is that the causes of AI weaknesses do not relate to the oversimplification of AI models, like many people pretend nowadays, but to

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 the culture" introduced by the Neo-Kantian school in the beginning of the $20^{th}$ century.

From a philosophical point of view, it means that AI participates to the "sciences of the culture", i.e. that it cannot be entirely reducible to a "science of the nature" or to mathematics and theoretical sciences. But it is not more reducible to the "sciences of the culture". More precisely, it is what Heinrich Rickert identifies as an "intermediary domain" that belongs simultaneously to the theoretical sciences, i.e. to formal logic and mathematics, to the empirical sciences of the nature and to the empirical sciences of the 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.

### 5.2 Methods of AI

Since AI participates to the "sciences of the culture", it has to take advantage of the logic of the "sciences of the culture", which may enlarge the scope of its methods. Let us recall that the sciences of the 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 precise that it is not to observe singularity, but to study paradigmatic cases and to explain in what respect the individual cases under study can be universalized. An excellent example of such type of studies was done by a cognitive anthropologist, Edwin Hutchins, in the book titled "Cognition in the wild" [15] 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 participates to this dimension of AI. It is not to generalize all the individual failures by extracting their common properties, as it could be in any science of the nature, but to understand the logic of the failures, to see what lessons could be drawn from these bad experiences to generalize them and to learn from it. This is exactly what we are trying to do in the "What Went Wrong and Why?" workshops. In this respect, they participate to the logic of the "sciences of the culture".

### 5.3 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 that the present age is often called a "knowledge age". However, as Carl

Hewitt mentioned during the 2006 "What Went Wrong and Why?" workshop, in Stanford, it's a pity that AI did not participate more actively to cultural evolutions consecutive to the development of information technologies, for instance, to the Wikipedia free encyclopedia or to the social web.

More generally, the knowledge quest can be greatly accelerated by the use of AI technologies. For instance, my team is working in musicology [16], in textual criticism, in social sciences [17], in epistemology [18] 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 case of medicine, there already exist many attempts to model organs [19] and to simulate medical diagnosis; AI had part in these successful achievements; 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.

## References


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

[2] Simon, H.A.: The Sciences of the Artificial. 3rd edn. MIT Press, Cambridge, Massachusetts (1996)

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

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

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

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

[7] Moravec, H.: Mind Children. Harvard University Press, Cambridge, Massachusetts (1988)

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

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

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

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

[12] Rickert, H.: Kulturwissenschaft und Naturwissenschaft. 5th edn. J.C.B. Mohr (Paul Siebeck), Tubingen (1921)

[13] Cassirer, E.: Substance and Function. Open Court, Chicago (1923)

[14] Cassirer, E.: The Logic of the Humanities. Yale University Press, New Haven (1961)

[15] Hutchins, E.: Cognition in the Wild. MIT Press, Cambridge, Massachusetts (1995)

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

[17] Velcin, J., Ganascia, J.G.: Stereotype extraction with default clustering. In: proceedings of the 19th International Joint Conference on Artificial Intelligence (IJCAI). (2005) 883–888
[18] Ganascia, J.G., Debru, C.: Cybernard: A computational reconstruction of claude bernards scientific discoveries. In: Studies in Computational Intelligence (SCI), Springer-Verlag (2007) 497–510
[19] Nobel, D.: The Music of Life: Biology Beyond the Genome. Oxford University Press (2006)