Article

Human Abductive Cognition Vindicated: Computational Locked Strategies, Dissipative Brains, and Eco-Cognitive Openness

Lorenzo Magnani iD

Department of Humanities, Philosophy Section and Computational Philosophy Laboratory, University of Pavia, 27100 Pavia, Italy; lmagnani@unipv.it

Abstract: *Locked* and *unlocked* strategies are illustrated in this article as concepts that deal with important cognitive aspects of deep learning systems. They indicate different inference routines that refer to poor (locked) to rich (unlocked) cases of creative production of creative cognition. I maintain that these differences lead to important consequences when we analyze computational deep learning programs, such as AlphaGo/AlphaZero, which are able to realize various types of abductive hypothetical reasoning. These programs embed what I call locked abductive strategies, so, even if they present spectacular performances for example in games, they are characterized by poor types of hypothetical creative cognition insofar as they are constrained in what I call *eco-cognitive openness*. This openness instead characterizes unlocked human cognition that pertains to higher kinds of abductive reasoning, in both the creative and diagnostic cases, in which cognitive strategies are instead unlocked. This special kind of "openness" is physically rooted in the fundamental character of the human brain as an open system constantly coupled with the environment (that is, an "open" or "dissipative" system): its activity is the uninterrupted attempt to achieve the equilibrium with the environment in which it is embedded, and this interplay can never be switched off without producing severe damage to the brain. The brain cannot be conceived as deprived of its physical quintessence that is its openness. In the brain, contrary to the computational case, ordering is not derived from the outside thanks to what I have called in a recent book "computational domestication of ignorant entities", but it is the direct product of an "internal" open dynamical process of the system.

Keywords: abduction; creativity; deep learning; eco-cognitive openness; locked and unlocked strategies; dissipative brain; go game; AlphaGo; AlphaZero

1. Introduction

In this article, the key terms refer to what I call *locked* and *unlocked* strategies, seen as concepts that deal with important cognitive features of deep learning systems. They denote various inference methods ranging from poor (locked) to rich (unlocked) situations of creative cognition generation. The first ones affect a computational system that refers to (or, importantly, selects) a rigid scenario of data. Another key term is *abduction*, a concept that indicates all kinds of hypothetical reasoning, both creative, for example in the case of scientific discovery, or diagnostic, for example in the case of medical reasoning. Indeed, when we look at computational deep learning programs such as AlphaGo/AlphaZero, which can perform various sorts of abductive hypothetical reasoning, I believe the distinction above is of crucial importance. These programs contain what I refer to as locked abductive strategies, and as a result, even if they produce spectacular results in games, they are characterized by poor types of hypothetical creative cognition in the sense that they are constrained in what I refer to as *eco-cognitive openness*, which is the third important concept introduced in this article. Instead, this openness characterizes unlocked human cognition in terms of higher types of abductive reasoning, in both creative and diagnostic contexts, in which cognitive

methods are unlocked. This unique kind of openness is physically rooted in the human brain's fundamental character as an open system constantly coupled with the environment (that is, a n"open" or "dissipative" system, the last key term introduced in this article): its activity is the continuous attempt to achieve equilibrium with the environment in which it is embedded, and this interplay can never be turned off without severe brain damage. The brain cannot be imagined without its physical essence, which is its openness. In contrast to the computational situation, ordering in the brain is not derived from the outside.

2. Deep Learning Cognitive Strategies Are Locked

In 2015 Google DeepMind's program AlphaGo (able to perform the famous Go game) beat Fan Hui, the European Go champion and a 2 dan (out of 9 dan) professional, five times out of five with no handicap on a full-size 19×19 board. In March 2016, Google also defeated Lee Sedol, a 9 dan player considered the best champion of the planet. The DeepMind program defeated Lee in four of the five games. The program was able to create a new—unprecedented—surprising move capable of creating new strategies and so phenomenologically simulating human beings' performances. This deep learning program "learned" a lot by attending thousand of games played by humans, such as the one played by Lee Sedol, thanks to so-called "reinforcement learning": the program in turn plays repeatedly against itself to further strengthen its own deep neural networks[1].

If we consider computational strategies as a compound of heuristic processes, that is, a process composed of good choices of the subsequent state of a cognitive routine—according to some opportunely chosen criteria—we can say that heuristics are used to arrive at a specific target thanks to their organization in strategies. In game theory, the concept of strategy is instead wider and includes the methods that agents adopt when dealing with other agents to the variegated related intertwined or collective cognitive actions. In turn, ecological thinking (or ecological rationality) [1–3] further sees strategies as processes that exploit a great quantity of information and knowledge, so involving high computational efforts; on the contrary, heuristics wonderfully perform simple and efficacious moves, even if less rigorous. In many areas of computer and cognitive science literature, cognitive heuristics are simply considered coincident with cognitive strategies. I will adopt here the considerably shared view within AI that sees strategies as a composition of successive selections of appropriate heuristics.

I think that it is from the perspective of the studies on abductive cognition [4] that we can suitably and profitably consider what I just called strategic cognition, stressing the contrast between what I called *locked* and *unlocked strategies* [5].[2] Some more words have to be added to illustrate some aspects of the concept of abduction. A hundred years ago, Charles Sanders Peirce [7] coined the concept of abduction in order to illustrate that the process of scientific discovery is not irrational and that a methodology of discovery is possible. Peirce interpreted abduction essentially as an "inferential" *creative process* of generating a new "explanatory" hypothesis. Abduction has a logical form (fallacious, if we model abduction using classical syllogistic logic)—the abductive inference rule corresponds to the well-known fallacy called affirming the consequent , which is distinct from deduction and induction. Many conclusions of reasoning that are not derived in a deductive manner are abductive. For instance, if we see a broken horizontal glass on the floor we might *explain* this fact by postulating the effect of wind blowing shortly before: this is not certainly a deductive consequence of the glass being broken (a cat may well have been responsible for it).

The concept was further studied in the field of epistemology by Hanson [8] and in the areas of artificial intelligence. Abduction is a popular term in many fields of AI, such as diagnosis, planning, natural language processing, motivation analysis, logic programming, and probability theory. Moreover, abduction is important in the interplay between AI and philosophy; cognitive science; historical, temporal, and narrative reasoning; decision-making; legal reasoning; and emotional cognition.[3] Six volumes (monographs and collections) are currently available [9–15] and three special issues of international journals (*Philosophica*,

1998 61(1); *Foundations of Science*, 2004, 9; 2008, 13(1); *Logic Journal of the IGPL*, 2006 14(2)). Of course, many articles from various disciplinary fields of research are continually published on this topic[4] .

I have to anticipate that deep learning machines (and so AlphaGo/AlphaZero programs) are characterized by locked strategies, a fact that strongly impacts the kind of creativity which is realized by them.

In my research on abduction, I have widely described the several types of human, animal, and computational hypothetical cognition that can be accounted for using this important concept. I introduced two kinds of abduction, *selective* [9]—for example in medical diagnosis (in which we have to "select" from a "repository" of already available hypotheses)—and *creative* (abduction that provides *new* hypotheses). In addition, I also always stressed that abduction is not only sentential, that is, executed thanks to the resources of human language (oral or written or artificially built using symbols, such as in the case of mathematics and logic), but also "model-based" and "manipulative". Model-based abduction concerns the exploitation of internal cognitive acts that utilize models such as simulations, visualizations, images, etc.; manipulative abduction exploits the so-called external character of human cognition, in which what I have called the "eco-cognitive" character of cognition is central because we have to consider all those cognitive processes (embodied, embedded, situated, and enacted) in which the function of external models (for example artifacts), is essential, and in which the structure of the cognitive performances is often concealed and not easily extractable. In this last case, the manipulative action can bring about new data—that were not available before—and new heuristics capable of enhancing the ways agents solve problems that require the innovative generation (or just the selection) of suitable hypotheses. I contended that manipulative abduction represents a kind of "thinking through doing" and not only, in a pragmatic sense, about doing (cf. [4] chapter one). In the case of deep learning machines (and in the case of the games we are considering in this article), it is apparent that we face cases of manipulative abduction: the cognitive processes are intrinsically related to the manipulation of the stones and several cognitive embodied moments are at stake, together with the needed visualization of the entire external context, the competitor, etc.

Abductive Cognition and AlphaGo/AlphaZero

It is well known that research on abduction has increased knowledge about creative cognition, such as when dealing with the simple case of a novel move in a Go game. The two concepts I have introduced during my studies on abduction of *knowledge-enhancing abduction* and *eco-cognitive openness* [20] are very appropriate to examine the *locked* and *unlocked abductive strategies* I have introduced above. Locked and unlocked strategies are fundamental as conceptual tools that favor the analysis of central cognitive aspects of deep learning machines. These strategies are both present in human cognition, but it is now impossible to find both in machines: indeed they generate creative outputs, which unfortunately are endowed with different levels of creativity, and in computational machines, the presence of locked strategies jeopardizes creativity. These differences can be clearly seen in the case of deep learning tools such as AlphaGo, which aim at mechanizing various types of abductive hypothetical cognition.

I indeed contend that these programs present what I call *locked* abductive strategies, which exhibit weak (even if astonishing) types of hypothetical creative cognition, because they are constrained in what I call *eco-cognitive openness*, which instead is typical of human beings' cognitive processes dealing with abductive creative reasoning: in this last human case, cognitive strategies are *unlocked*. The fact that these programs are not based on logic and that the main intellectual tradition related to a formalization of abduction was instead exactly related to logic does not create a problem for my arguments [5]. Indeed, abduction also occurs at a sub-symbolic level, in both humans and machines, so there is no surprise in seeing deep learning machines able to lead to abductive results. Furthermore, we have to remember that abduction is also characterized by multimodality (for example, it can

be performed exploiting diagrams), as I have described in [4,20]. In addition, we have to repeat that humans often guess abductive hypotheses thanks to the manipulation of the external environment, suitably enriched with cognitive representations and appropriate artifacts, but also thanks to embodied and unconscious capacities (which also characterize some aspects of cognition in higher mammals that surely do not take advantage of symbolic syntactic language). AI has always presented various methods, formalisms, and algorithms capable of leading to the construction of programs related to abductive performances.[5]

What characterizes the particular abductive performance occurring in deep learning AI program AlphaGo?

3. Natural, Artificial, and Computational Games
3.1. Locked and Unlocked Strategies in Natural and Artificial Environments

Go is a game played by human agents, and AlphaGo is an automatic computational deep learning program that also plays that game and is able to compete with humans. We can say that Go is already something "artificial" because it is created by human agents with established rules, a specific board, and other material entities, that is, stones. AlphaGo/AlphaZero is even more artificial, a fruit of a technological creativity, an engineered product of the cognitive skills typical of a small elite of human beings. There are also "natural cognitive games", such as in the case of the pre-linguistic cognitive "natural game" between humans and their surroundings, in which "unlocked" strategies (which I will soon describe) are at play, as wonderfully exemplified by the phenomenological tradition, a game encompassing embodied aspects and distributed cognition and also visual, kinesthetic, and motor sensations [22]. This "natural" game is characterized by the presence of *unlocked* strategies: no limits and local constraints are active. No predetermined backgrounds are fixed. In sum, humans in this case are not compelled to play constrained, for example, by a specific (unchangeable in its structure) Go board that indeed renders cognitive strategies *locked*. A fixed board and determined stones and rules render the scenario strongly constrained.

3.2. Reading Ahead as an Abductive Engine

Reading ahead, as Go players standardly say, is the generation of anticipation that aims to be very strong and reliable (both deeply minded and intuitive). It involves

1. Groups of potential being selected and their possible outcomes. A present scenario at time t_1 shown by the board "adumbrates"—a Husserlian concept!—a posterior possible and more productive scenario at time t_2, which is the fruit of an expected smart abduction that will be in turn followed by another abduction regarding the action that leads to a new move;
2. Possible countermoves to each move;
3. Further possibilities after each of those countermoves. It seems that some of the best competitors of the game Go can read up to 40 moves ahead even in enormous complex positions.

Other strategies that are productively followed by human players in the game Go are "global influence, interaction between distant stones, keeping the whole board in mind during local fights, and other issues that involve the overall game. It is therefore possible to allow a tactical loss when it confers a strategic advantage".

All these strategies, even multiple and variegated, are "locked": indeed the entities of each scenario remain same during the whole process of the game. The only modifications that are allowed are the ones concerning the number of possible stones in play and their positions on the board in a finite and fixed environment (no novel rules, no novel entities, no novel boards, etc.) These strategies cannot take advantage of information *different* from the ones presented in the constrained given scenario. Of course, the "human" player can exploit—to enhance and strengthen their strategies—internal sources not ineluctably connected to the previous time spent playing Go, but with other kinds of skills from other

variegated fields of cognition, but this does not cancel the fact that he is working in a locked situation: this kind of mental openness of the human just refers to the fact that, even if in the case of this human Go player, the strategies are locked strategies, they show a lesser degree of closure thank in the case of the computational program AlphaGo. In humans, strategies are locked with respect to the external fixed scenario but more open with respect to the mental references to other available extended strategic endowments; in AphaGo and in deep learning programs, the strategic storage cannot—at least presently—take advantage of that mental "openness" and versatility typical of human beings: the reservoir of strategies is merely generated, and the program learns to play the game by checking data of thousands of games (played by humans or by the program against itself) and not deriving from divergent founts.[6]

4. Locking Abductive Strategies Jeopardizes the Maximization of Eco-Cognitive Openness

I already said that a typical outcome of good abductive reasoning is the capacity to *enhance* knowledge in a more or less creative way (chapter seven [20]), ranging from the performance of medical diagnosis (in which at least we discover something new about a patient) to the new knowledge generated, for example, in the case of scientific discovery. We have already said that an automatic game player such as AlphaGo also performs knowledge-enhancing abductions thanks to a myriad of learned skillful cognitive strategies.

What I vigorously contend is that only thanks to an "openness" of the cognitive environment can reach we exceptional selective or creative optimal abductive result; that is, we need what I have called [20] *optimization of eco-cognitive situatedness*, in which eco-cognitive openness is fundamental [23]. An optimization of eco-cognitive situatedness is particularly necessary in the case of creative abduction: in a recent book [24], I have also emphasized that to favor *discoverability* and to protect human creativity, a kind of "ecology" of the cognitive environments available to human beings that are making abductions—and so occasionally guessing creative hypotheses—has to be implemented. In Section 7 below, I further describe that—first of all—to get good creative and selective abduction, cognitive strategies must not be "locked" in an external fixed eco-cognitive environment, that is, in a scenario marked by what Hintikka [25] would have called rigid and finite definitory rules, and by finite material entities, which would fixedly work as cognitive mediators capable of governing agents' reasoning.

In [20,23,26], I stressed the relevance in good abductive cognition of hypotheses that I called *optimization of situatedness*. For example, to obtain creative abductive outcomes in scientific reasoning, the "situatedness" of the related cognitive activities is surely related to what I call eco-cognitive features related to the situations in which knowledge is freely "traveling" and where the richness and the maximization of the information available is granted. The maximization regards the optimization of situatedness that can only be built by a *maximization of changeability* of the starting data that characterize the abductive cognitive processes: inputs have to be maximally expanded, reconstructed, or changed, and the same has to happen with respect to the knowledge enforced during the hypothetical reasoning process. We need an advantageous "cognitive environment" in which the accessible data have to be *optimally positioned*.[7] In summary, abductive processes leading to hypotheses—in an appreciable number of cases, for example, in science—are highly *information-sensitive* and are confronted with a stream of information and data that have to be continuous and fittingly promoted and upgraded when needed.

5. The Physics of Eco-Cognitive Openness

5.1. Eco-Cognitive Openness Characterizes Dissipative Brains

A further intellectual enrichment of the concept of eco-cognitive openness derives from research concerning the so-called dissipative brain [27–30], related to the extension of the quantum model of the brain to dissipative dynamics [31]. Indeed, in this context, two points are fundamental: (1) the brain is a system *permanently* coupled with the environment

(an *open* or *dissipative* system) in a continuous attempt to reach equilibrium with it and (2) a crucial property of quantum field theory (QFT) holds, i.e., the existence of infinitely many states of minimal energy, the so-called vacuum states or ground states. On each of these vacua, a full set (a space) of other states of nonzero energy can be built. The collection of all these spaces or representations constitutes a "memory space": "In the dissipative quantum model of the brain the vacuum code is taken to be the memory code. A given memory is represented by a given degree of ordering. A huge number of memory records can be thus stored, each one in a vacuum of given code. In the original model by Ricciardi and Umezawa only one vacuum is available for memory printing. In the dissipative model all the vacua are available for memory printing" ([28], p. 316). From this quantum perspective, a fundamental role is played by the process of

> [...] "the spontaneous break-down of symmetry" by which the invariance (the symmetry) of the field equations manifests itself into ordered patterns in the vacuum state. The symmetry is said to be broken since the vacuum state does not possess the full symmetry of the field equations (the dynamics). The order is indeed such a "symmetry". One can show that when symmetry is broken the invariance of the field equations implies the existence of quanta, the so-called Nambu-Goldstone (NG) quanta, which, propagating through the whole system volume, are the carrier of the ordering information, they are the long-range correlation modes: in the crystal, for example, the ordering information is the one specifying the lattice arrangement [28], pp. 318–319.

Consequently, order is *lack of symmetry*, and every breaking of symmetry becomes, at the macroscopic level, the recording of a new memory; it causes the formation of a new attractor. In this specific sense, memory is not a simple recording (and/or recollection) of information.[8]

As I have already said, the mechanism of spontaneous breakdown of symmetry is responsible for the order, as the product of a *self-organizing dynamics*: "The process of symmetry breaking is triggered by some external input; the 'choice' of the specific symmetry pattern which is actually realized is, on the contrary, 'internal' to the system. Therefore one speaks of self-organizing dynamics: ordering is an inner (spontaneous, indeed) dynamical process. [...] in the brain, contrary to the computer case, ordering is not imported from the outside, it is the outgrowth of an 'internal' dynamical process of the system." [28], pp. 316–317.

Given that in the quantum model of the brain a particular memory is associated with a specific degree of ordering (that is to a specific value of the vacuum code), and given that intertwining with the environment generates the irreversible evolution of the brain, we can understand how we can have a feeling of the past and the future: the arrow of time derives from brain dynamics, which is a dissipative dynamics. Furthermore, it is important to say that, differently from in the case of a computer, *the brain cannot avoid* getting information; it is, so to speak, always constitutively *open* to the world (and also open to our inner world) and so a dissipative system.[9] Furthermore, the brain state;s evolution is a "story" of its "irreversible" intertwining with the external world, not excluding the internal surroundings proper of a mammal body and external to the nervous central system. It is obvious that it is impossible, given the current state of computer science, to have something similar in the case of a computer.

We have to add that the mathematical formalism for quantum dissipation implies the doubling of the brain's degrees of freedom. The doubled degrees of freedom refer to the kind of relationships with the environment to which the brain is coupled, a doubling that governs the balance of the energy flux between the system and the environment. We see in this case that the physics of eco-cognitive openness explains the important fact that the environment is constantly and integrally represented by the doubled degrees of freedom that are characterized as the "time-reversed copy" (the Double) of the brain.[10] The environment is constantly reverberated in a "model" in the brain,[11] where the time-

reversed features indicate that the energy flux outgoing from the brain is incoming into the environment, and vice versa, in a physical framework naturally characterized by two components, noise and chaos: "Indeed, small differences in the codes associated to external inputs may lead to diverging differences in the corresponding memory paths" ([28], p. 326). Furthermore, weak perturbations can lead the system to macroscopic alterations of its configuration. The same stimulus in distinct contextual situations may drive distinct brain reactions, given the fact that the original stimulus is a weak one.

From the perspective of traditional representationalism, the doubled degrees of freedom of the dissipative brain, seen in a dynamical perspective, can also be fruitfully enriched by relating it to the problem of "conceptual blend" stressed by the research tradition in cognitive science of distributed cognition. For example, Hutchins usefully observes:

> First, there is the selectivity of perception that produces a filtered conceptual representation of the physical world. Second, there is selective projection in the process by which the prior conceptualization of the world (the "real space" representation) is blended with the other conceptual input. Is there any evidence that these are two separate processes? It seems preferable to assume that the selective attention to, and projection of, structure from the material world to the blended space is the perceptual process. That is, that selective perception is a conceptual process [33], p. 1561.

It is interesting to note that phenomena such as memory associations, memory confusion, even the possibility to forget some memories, or else difficulties in recovering memory, are described by the dissipative model. The coherent collective behavior also grants the system its stability. For example, resulting memories explain how memory remains relatively unchanging and well preserved even if the brain is a highly excited system, because of the enduring electrochemical processes and the uninterrupted response to external stimulation. We have to add that in the framework of the dissipative brain, memory mechanisms are modeled as distinct mechanisms from the electrochemical processes of neuro-synaptic dynamics:

> [...] the brain is then a "mixed" system involving two separate but interacting levels. The memory level is a quantum dynamical level, the electrochemical activity is at a classical level. The interaction between the two dynamical levels is possible because the memory state is a macroscopic quantum state due, indeed, to the coherence of the correlation modes. The coupling between the quantum dynamical level and the classical electrochemical level is then the coupling between two macroscopic entities. This is analogous to the coupling between classical acoustic waves and phonons in crystals (phonons are the crystal NG quanta). Such a coupling is possible since the macroscopic behavior of the crystal "resides" in the phonon modes, so that the coupling acoustic-waves/phonon is nothing but the coupling acoustic wave crystal [28], p. 326.

5.2. Abductive Errors Vindicated

A dissipative brain elaborates the relationship with its surroundings by cyclic processes of trial and error, that is, to procedures that are typically illustrated by the theories about human abductive cognition; uncertainty, doubt, surprise, and incomplete information characterize this process, in which new perspectives are continuously adopted, which are either opportunely—consciously—directed or random, also allowing "fuzziness in the initial conditions, the starting assumptions of our traveling in the memory space (our archive of certainties)" [28], pp. 335–336.[12] The dissipative brain is an *erratic* brain [34], in which the uniqueness of its identity and the unrepeatability of the emerged cognitive processes are granted. As Peirce observed, "It is a primary hypothesis underlying all abduction that the human mind is akin to the truth in the sense that in a finite number of guesses it will light upon the correct hypothesis" ([7], 7.220); this means that human abductive cognition is "akin to the truth" but not exempt from failures, making mistakes

a constitutive feature of human hypothetical cognition. We have illustrated above that the dissipative model of the brain emphasizes that the incoming flux of information that arrives at the brain through perceptions cannot be discontinued; we cannot switch off a mammal human brain as we can in the case of a laptop. It is this quality of the dissipative brain that can grant the emergence of entirely *new* cognitive perspectives and, at the same time, when the flux from information is not seriously inhibited (for example because of social and political disfunction), of new and highly creative ones.

Turing machines are not allowed making mistakes analogous to the errors performed by the dissipative brains: these errors are intrinsic to the system dynamics, so the errors are not in this case related to the ones done, for example, in measuring a quantity by an observer, as a "deviation from accuracy", that is, as a "system-observer relational feature". The quantum fluctuation processes of the dissipative brain are affected by unpredictability. Unpredictability is constitutively linked to mistakes, as ([34], p. 71) further observe:

> Sometimes mistakes are useful to introduce or observe unexpected behaviors or results. Testing a newly designed machine has in general the meaning of detecting erratic behaviors to be avoided in an improved design of the tested machine. When mistakes are not rejected, they constitute additions to or extensions of the observer knowledge. Examples are in production processes and discoveries made by chance. In these cases, the term by chance means indeed by mistake (with respect to what was expected). In some sense, the term dis-covery is equivalent to the term mistake (the discovery is always by chance, otherwise it is not a discovery). [...] Inside a given context, the unpredictable behavior is not a "negation", is not a "deviance" with respect to any possible behavior. It is a novelty.

6. Big Data: Huge but Locked

In Section 5.1 above, I quoted an interesting passage, written by Vitiello, regarding the so-called dissipative brain: "The process of symmetry breaking is triggered by some external input; the 'choice' of the specific symmetry pattern which is actually realized is, on the contrary, 'internal' to the system. Therefore one speaks of self-organizing dynamics: ordering is an inner (spontaneous, indeed) dynamical process. [...] in the brain, contrary to the computer case, ordering is not imported from the outside, it is the outgrowth of an 'internal' dynamical process of the system." [28], pp. 316–317.

If we accept that in the brain, contrary to in the computer case, ordering is not imported from the outside but it is the outgrowth of an "internal" dynamical process of the system, a question arises. What happens in the case of the relationship between machine learning (and deep learning) and big data, by now usually available in digital format? In light of the considerations I advanced in this paper, we have to note that in the case of the computational exploitation of big data, we still face a process of an "ordering" imported in the machine from the outside. This order is of course governed by the related software but is strongly imposed by the data that are offered to the system by human beings. These data, such as in the case of AlphaGo/AlphaZero, belong to a circumscribed field, where of course we can see the availability of a huge deluge of information, unfortunately always coming from limited and specific sources. Even in the case of AlphaZero, which exploits the data that it itself has previously generated, we face the computational manipulations of data that consist of a huge number of Go games. When we say that "in the brain, contrary to the computer case, ordering is not imported from the outside, it is the outgrowth of an 'internal' dynamical process of the system", we are just referring to the fact that the internal computational ordering processes are not able to "cleave" the locked—in the sense I have attributed to this adjective in Section 2 above—character of the informational scenario that feeds the system, making for more openness in the reservoir of data.

It is not the case that researchers that study the epistemological, cognitive, legal, and ethical problems of big data in general and of the relationship between big data and machine learning (and deep learning) have stressed the need for a *de-contextualisation* of

facts from their context of origin to the aim of improving the quality of their eventual computational manipulation. For example, in the case of biological data, "one of the main tasks of database curators is to decontextualise the data that are included in their resources, so that they can travel outside of their original production context and become available for integration with other datasets (thus forming a big data collection) [...] Despite constant advances, it is still impossible to automate the de-contextualisation of most types of biological data" [35], p. 4. In other words, adopting the lexicon I have introduced in this article, "unlocking" the data set is still a difficult—extra-computational—"human" task that involves both cognitive, institutional, and ethical problems related to the problem of "fitting standards" of epistemological decency in the subsequent computational treatment of big data: "researchers who wish to submit their data to a database need to make sure that the format that they use, and the metadata that they provide, fit existing standards—which in turn means acquiring updated knowledge on what the standards are and how they can be implemented, if at all; and taking time out of experiments and grantwriting" [35], p. 4.[13]

In summary, to pursue satisfactory epistemological virtues, big data have to be "unlocked" following different policies, of course from the well-known (and still to be deepened and clarified) legal and ethical points of view, but mainly because they have to be "curated" to be able to favor, for example, in biology, excellent abductive processes of scientific discovery across biological subareas. These policies cannot be performed by computations; only human beings can realize them. The various computational tools, and especially the ones based on machine learning (and deep learning), that are used to deal with big data are often opaque in the perspective of their functions and in their basic cognitive suppositions and often seem to reach results that are unreliable from the perspective of scientific rationality. As Leonelli concludes: "This increases the worry that big data science may be grounded upon, and ultimately supporting, the process of making human ingenuity hostage to an alien, artificial and ultimately unintelligible intelligence" [36].[14] A serious criticism of the abductive capacities of computational performances on large databases is contained in a recent article written by Calude and Longo [38]. The authors contend, taking advantage of deep classical results from ergodic theory, Ramsey theory, and algorithmic information theory, that the exploitation of large databases favors the generation of spurious correlations, which of course do not represent good creative abductions.[15]

7. Locked Strategies Limit Creativity

We have seen above that optimization of situatedness is related to unlocked strategies. Locked strategies, which characterize the Go game, AlphaGo/AlphaZero, and the deep learning systems, instead limit creativity. A poor scenario tends to minimize eco-cognitive openness, as I have illustrated in Section 3.2 above. I have said that in the game of Go, stones, board, and rules are fixed and therefore predetermined. We have to add that, in general, the presence of rigid and unchanging scenarios in the natural and artificial world, even if obviously related to the presence of stabilities, which in turn give origin to habits of behavior, surely tend to undermine human creativity. The creative brain is nourished by a free flux of information coming from an open and extended informational environment, as the knowledge deriving from physics teaches us, seeing human brains as constitutively open dissipative structures.

It is not possible, during a Go game, to play for a moment Chess or adopt a different rule or another weird cognitive process, affirming that that strange part of the game is still reliable and applicable to the game you agreed to play. On the contrary, in the case of scientific discovery, for example, the scientist (or the collective of scientists) often refers to variegated external models and freely modifies their reasoning strategies[16] for example to reach new analogies or to favor other unexpected cognitive fruitful cognitive processes (prediction, simplification, confirmation, falsification, etc.) to improve and enrich the abductive creative process.

In summary:

1. Contrarily to the case of high-level "human" creative abductive inferences the status of artificial games (and of their deep learning computational companions) is very poor from the point of view of the non-strategic knowledge that is exploited;

2. in Go (and similar games) and in deep learning systems such as AlphaGo/AlphaZero, in which strategies and heuristics are "locked", these are precisely the only part of the game that can be enhanced and made more fecund: strategies and related heuristics can be exploited in an innovative way, and new ones can be created. No other types of knowledge will be modified, and all the remaining aspects remain immutable.[17] Of course this preeminence of the strategies is the essence of Go, Chess, and other games, a fact that explains the impressiveness of the more smart moves of the human champions (and of course of AlphaGo/AphaZero). Unfortunately, this dominance of strategies is also the feature that renders the creativity at stake even weaker than the one that characterizes the most complicated cases of human selective abduction (medical diagnosis, for example). At the same time, this weakness also accounts for the easiness in building a deep learning computer simulation of games such as Chess or Go, with respect to the simulation of the strategies at play, for example, in scientific discovery.[18]

Many managers are exploiting and prospecting an exploitation of deep learning programs to aid science in resolving relevant real-world problems in healthcare and in other fields such as scientific reasoning. Many implementations in business, thanks to commercialization, of deep learning AI programs are currently at work. The reader interested in having a quick look at some pros and cons of deep learning programs can simply refer to the Wikipedia entry DeepMind (https://en.wikipedia.org/wiki/DeepMind, accessed on 28 November 2021), DeepMind is a British artificial intelligence company founded in September 2010 and taken by Google in 2014, the company also created the AlphaGo program that clearly reports the contested case concerning the so-called "NHS data-sharing controversy". However, it seems that beyond the epistemological and cognitive perplexities I have illustrated in this article and the limitations of the so-called locked strategies, deep learning systems offer chances for business and good integration in the market.

In my opinion, epistemologists, logicians, and cognitive scientists have to monitor the exploitation of these AI devices: good AI software, a big new opportunity in terms of data analytics, can be easily transformed into a tool that does not fulfill epistemological and/or ethical criteria of rigor. Even if I do not aim in this article to deal with this issue, I can just quote some recent observations about the current computational exploitation of big data, which can mistakenly lead to computer-discovered correlations affected by epistemological flecks. Calude and Longo ([38], p. 595) observe that, unfortunately, some "correlations appear only due to the size, not the nature, of data. In 'randomly' generated, large enough databases too much information tends to behave like very little information". Some spurious correlations as outcomes of deep learning supposed to be "creative abductions" can just be trivial generalizations, even if attained thanks to complicated engineered artifacts. I cannot further analyze the problems related to the effects of deep learning computational programs on ethics and society; in this article, I just considered some fundamental cognitive, logical, and epistemological aspects aiming at stressing the different qualities of locked and unlocked strategies in the case of human and computational intelligence.

8. Conclusions

In this article I have illustrated the concepts of locked and unlocked strategies, abduction, and optimization of eco-cognitive openness, as key concepts able to highlight the fundamental epistemological and cognitive differences between deep learning computational cognitive performances (I considered the example of the program Alphago/AlphaZero) and human ones. My research on abduction, which has stressed the importance of the concept of *eco-cognitive openness*, helped in showing how this openness is lacking in the case of deep learning programs but not in high-level human inferential creative and diagnostic inferences. Deep learning programs such as AlphaGo/AlphaZero are based on

locked abductive strategies that jeopardize eco-cognitive openness. On the contrary, unlocked abductive strategies, which respect what eco-cognitive openness demands, characterize the high-level types of abductive creative and diagnostic reasoning that are distinctive of human cognition. I also illustrated the important perspective, deriving from physics, that sees human brains as "open" and "dissipative systems": brains are permanently open and coupled with the environment in an uninterrupted attempt to achieve equilibrium with it. This interplay can never be switched off without producing severe damage to the brain. Consequently, physics teaches us that brains, contrary to computational machines, do not derive order from the outside thanks to what I have called in a recent book "computational domestication of ignorant entities", but determine it as direct product of the "internal" open dynamical process of the system.

Notes

1. AlphaGo Zero is a version of DeepMind's Go software AlphaGo; the recent AlphaZero further enriches AlphaGo Zero and learns by its own played games.
2. In a recent book [6], I have exploited these notions as components of the framework of an entirely new dynamic perspective on the nature of computation. In this theory I have highlighted the role of unconventional computation as an incessant and terrific process of cognitive domestication of ignorant entities.
3. A classical bibliography on abduction is given in [9].
4. General classical considerations on abduction in science and AI can also be found in [10,16–19].
5. A variety of tools already present at the time of traditional AI methods and formalisms, when I was cooperating with AI colleagues to build a Knowledge-Based System (KBS) able to develop medical abductive—diagnostic—reasoning [21].
6. I have to note that my notion of a locked strategy is not related to the standard nomenclature of the game theory.
7. I have furnished more cognitive and technical details to explain in [23] and in a recent book [24].
8. We face a kind of rearrangement of the "whole attractor landscape": this means that a new memory becomes situated in the context of the entire set of memories already acquired by the brain. This process of contextualization renders the newly arrived information—which is in itself without meaning in the Shannon sense—endowed with a specific meaning that in turn tends to change the meanings belonging to the whole set of memories.
9. It is well-known that isolation of an individual tends to generate various pathologies and not only at the psychological level.
10. The physical model also explains that the existence of the first cannot be independent of the existence of the second, and vice versa. The "brain/environment" system is treated as the closed system "brain and its Double". Regarding the relationship with the Double as a route to consciousness and on the role of objectiveness of the external world as the primary and necessary condition for consciousness to exist, see [28], p. 328–335; consciousness would be rooted in a restless dialog—entanglement—of the self with its Double.
11. Cognitive science has also stressed this fact when Brooks observed that, at the root of the more basic forms of cognition, it can be hypothesized that the "world serves as its own best model" [32], p. 145.
12. A fluctuating background jeopardizes the precise determination of the trajectories from initial conditions in a totally unpredictable way.
13. Further details can be found in [36].

[14] On the general problem of discoverability and its discontents, which encompasses the present one regarding the curation of big data, see my book [37].

[15] For a recent interesting discussion on the limits of the use of machine learning in prediction about climate change, dealing with scientists' replacement of physically based parameterizations with neural networks that do not represent physical processes, directly or indirectly, see [39]. A defense of deep learning and machine learning advances regarding their capacity to generate approximate causality thanks to the finding of correlations between indirect factors is described by [40]. Finally, the readers who are interested in a rich and extended survey of the present epistemological, social, and political problems related to big data, algorithms, machine learning, artificial intelligence, and social networks can refer to [41], which especially stresses the fact that more or less reliable predictions generated by computational systems become de facto "prescriptions" capable of surreptitiously modifying human behaviors.

[16] Many interesting examples are illustrated in the recent [42],

[17] Of course, different rules and new boards of different sizes can be advanced, but this will lead to novel kinds of games, a possibility that does not affect my arguments.

[18] Some data regarding the history of so-called automated scientific discovery in AI are illustrated ([4], chapter two, Section 2.7 "Automatic Abductive Scientists").

References

1. Gigerenzer, G.; Selten, R. *Bounded Rationality. The Adaptive Toolbox*; The MIT Press: Cambridge, MA, USA, 2002.
2. Raab, M.; Gigerenzer, G. Intelligence as smart heuristics. In *Cognition and Intelligence. Identifying the Mechanisms of the Mind*; Sternberg, R.J., Prets, J.E., Eds.; Cambridge University Press: Cambridge, MA, USA, 2005; pp. 188–207.
3. Gigerenzer, G.; Brighton, H. Homo heuristicus: Why biased minds make better inferences. *Top. Cogn. Sci.* **2009**, *1*, 107–143. [CrossRef] [PubMed]
4. Magnani, L. *Abductive Cognition. The Epistemological and Eco-Cognitive Dimensions of Hypothetical Reasoning*; Springer: Heidelberg/Berlin, Germany, 2009.
5. Magnani, L. AlphaGo, Locked Strategies, and Eco-Cognitive Openness. *Philosophies* **2019**, *4*, 8. [CrossRef]
6. Magnani, L. *Eco-Cognitive Computationalism. Cognitive Domestication of Ignorant Entities*; Springer: Cham, Switzerland, 2022.
7. Peirce, C.S. *Collected Papers of Charles Sanders Peirce [1866–1913]*; Hartshorne, C., Weiss, P., Burks, A., Eds.; Harvard University Press: Cambridge, MA, USA, 1931–1958; Volumes 1–8.
8. Hanson, N.R. *Patterns of Discovery. An Inquiry into the Conceptual Foundations of Science*; Cambridge University Press: London, UK, 1958.
9. Magnani, L. *Abduction, Reason, and Science. Processes of Discovery and Explanation*; Kluwer Academic/Plenum Publishers: New York, NY, USA, 2001.
10. Josephson, J.R.; Josephson, S.G. (Eds.) *Abductive Inference. Computation, Philosophy, Technology*; Cambridge University Press: Cambridge, UK, 1994.
11. Flach, P.; Kakas, A. (Eds.) *Abductive and Inductive Reasoning: Essays on Their Relation and Integration*; Kluwer Academic Publishers: Dordrecht, The Netherlands, 2000.
12. Kuipers, T. *From Instrumentalism to Constructive Realism. On Some Relations between Confirmation, Empirical Progress and Truth Approximation*; Kluwer Academic Publisher: Dordrecht, The Netherlands, 2000.
13. Gabbay, D.M.; Woods, J. *The Reach of Abduction*; North-Holland: Amsterdam, The Netherlands, 2005.
14. Aliseda, A. *Abductive Reasoning. Logical Investigations into Discovery and Explanation*; Springer: Heidelberg/Berlin, Germany, 2006.
15. Walton, D. *Abductive Reasoning*; The University of Alabama Press: Tuscaloosa, AL, USA, 2004.
16. Gooding, D. Creative rationality: Towards an abductive model of scientific change. *Philosophica* **1996**, *58*, 73–102.
17. Kuipers, T.A.F. Abduction aiming at empirical progress of even truth approximation leading to a challenge for computational modelling. *Found. Sci.* **1999**, *4*, 307–323. [CrossRef]
18. Thagard, P. *Computational Philosophy of Science*; The MIT Press: Cambridge, MA, USA, 1988.
19. Shrager, J.; Langley, P. (Eds.) *Computational Models of Scientific Discovery and Theory Formation*; Morgan Kaufmann: San Mateo, CA, USA, 1990.
20. Magnani, L. *The Abductive Structure of Scientific Creativity. An Essay on the Ecology of Cognition*; Springer: Cham, Switzerland, 2017.
21. Ramoni, M.; Stefanelli, M.; Magnani, L.; Barosi, G. An epistemological framework for medical knowledge-based systems. *IEEE Trans. Syst. Man Cybern.* **1992**, *22*, 1361–1375. [CrossRef]
22. Magnani, L. Playing with anticipations as abductions. Strategic reasoning in an eco-cognitive perspective. *J. Appl. Log. Ifcolog J. Logics Their Appl.* **2018**, *5*, 1061–1092.
23. Magnani, L. The eco-cognitive model of abduction. Irrelevance and implausibility exculpated. *J. Appl. Log.* **2016**, *15*, 94–129. [CrossRef]
24. Magnani, L. *Discoverability. The Urgent Need of an Ecology of Human Creativity*; Springer: Cham, Switzerland, 2022.
25. Hintikka, J. *Socratic Epistemology. Explorations of Knowledge-Seeking by Questioning*; Cambridge University Press: Cambridge, UK, 2007.

26.　Magnani, L. The eco-cognitive model of abduction. Ἀπαγωγή now: Naturalizing the logic of abduction. *J. Appl. Log.* **2015**, *13*, 285–315. [CrossRef]

27.　Vitiello, G. *My Double Unveiled. The Dissipative Quantum Model of Brain*; John Benjamins: Oxford, UK, 2001.

28.　Vitiello, G. The dissipative brain. In *Brain and Being. At the Boundary between Science, Philosophy, Language and Arts*; Globus, G., Pribram, K.H., Vitiello, G., Eds.; John Benjamins: Amsterdam, The Netherlands, 2012; pp. 317–338.

29.　Vitiello, G. Mind and matter. Two entangled parallel time-lines, one reconstructing the past in remembering, the other extrapolating into the future in predicting. In *Quanta and Mind. Essays on the Connection between Quantum Mechanics and Consciousness*; de Barros, J.A., Montemayor, C., Eds.; Springer: Cham, Switzerland, 2019; pp. 103–113.

30.　Vitiello, G. Brain, mind, and the ontological prejudice. *Act. Nerv. Super.* **2019**, *61*, 112–115. [CrossRef]

31.　Prigogine, I.; Stengers, I. *Order out of Chaos. Man's New Dialogue with Nature*; Bantam: London, UK, 1984.

32.　Brooks, R.A. Intelligence without representation. *Artif. Intell.* **1991**, *47*, 139–159. [CrossRef]

33.　Hutchins, E. Material anchors for conceptual blends. *J. Pragmat.* **2005**, *37*, 1555–1577. [CrossRef]

34.　Minati, G.; Vitiello, G. Mistake making machines. In *Systemics of Emergence: Research and Development*; Minari, G., Pessa, E., Abram, M., Eds.; Springer: Heidelberg, Germany, 2006; pp. 67–78.

35.　Leonelli, S. What difference does quantity make? On the epistemology of Big Data in biology. *Big Data Soc.* **2014**, *1*, 2053951714534395. [CrossRef] [PubMed]

36.　Leonelli, S. Scientific research and big data. In *The Stanford Encyclopedia of Philosophy*; Zalta, E.N., Ed.; Metaphysics Research Lab, Stanford University: Stanford, CA, USA, 2020.

37.　Magnani, L. Epistemic mediators and model-based discovery in science. In *Model-Based Reasoning: Science, Technology, Values*; Magnani, L., Nersessian, N.J., Eds.; Kluwer Academic/Plenum Publishers: New York, NY, USA, 2002; pp. 305–329.

38.　Calude, C.S.; Longo, G. The deluge of spurious correlations in big data. *Found. Sci.* **2017**, *22*, 595–612. [CrossRef]

39.　Kawamleh, S. Can machines learn how clouds work? The epistemic implications of machine learning methods in climate science. *Philos. Sci.* **2021**, *88*, 1008–1020. [CrossRef]

40.　Vallverdú, J. Approximate and situated causality in deep learning. *Philosophies* **2020**, *5*, 2. [CrossRef]

41.　Numerico, T. *Big Data e Algoritmi*; Carocci: Rome, Italy, 2021.

42.　Magnani, L.; Bertolotti, T. (Eds.) *Handbook of Model-Based Science*; Springer: Cham, Switzerland, 2017.