BRAIN. Broad Research in Artificial Intelligence and Neuroscience
e-ISSN: 2067-3957 | p-ISSN: 2068-0473
Covered in: Web of Science (ESCI); EBSCO; JERIH PLUS (hkdir.no); IndexCopernicus; Google Scholar; SHERPA/RoMEO; ArticleReach Direct; WorldCat; CrossRef; Peeref; Bridge of Knowledge (mostwiedzy.pl); abcdindex.com; Editage; Ingenta Connect Publication; OALib; scite.ai; Scholar9; Scientific and Technical Information Portal; FID Move; ADVANCED SCIENCES INDEX (European Science Evaluation Center, neredataltics.org); ivySCI; exaly.com; Journal Selector Tool (letpub.com); Citefactor.org; fatcat!; ZDB catalogue; Catalogue SUDOC (abes.fr); OpenAlex; Wikidata; The ISSN Portal; Socolar; KVK-Volltitel (kit.edu)
2024, Volume 15, Issue 4, pages: 354-374.
Submitted: September 2nd, 2024 | *Accepted for publication:* November 5th, 2024

Computational Locked and Unlocked Strategies and Dissipative Brains The Abductive Eco-Cognitive Perspective

Lorenzo Magnani
Department of Humanties, Philosophy Section,
University of Pavia, 27100 Pavia, Italy
e-mail: lorenzo.magnani@unipv.it
https://orcid.org/0000-0003-2344-4379

Abstract: *Eco-cognitive computationalism is a cognitive science perspective that views computing in context, focusing on embodied, situated, and distributed cognition. It emphasizes the role of Turing in the development of the Logical Universal Machine and the concept of machines as "domesticated ignorant entities". This perspective explains how machines can be dynamically active in distributed physical entities, allowing data to be encoded and decoded for appropriate results. In this perspective, we can clearly see that the concept of computation evolves over time due to historical and contextual factors, and it allows for the emergence of new types of computations that exploit new substrates. Taking advantage of this eco-cognitive framework I will also illustrate the concepts of "locked and unlocked strategies in deep learning systems, indicating different inference routines for creative results. Locked abductive strategies are characterized by poor hypothetical creative cognition due to the lack of what I call eco-cognitive openness, while unlocked human cognition involves higher kinds of creative abductive reasoning. 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. Brain cannot be conceived as deprived of its physical quintessence which 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: *eco-cognitive computationalism; cognitive science; logical universal machine; Alan Turing; philosophy; creative abductive reasoning; brain; abductive strategies; eco-cognitive openness.*

How to cite: Magnani, L. (2024). Computational locked and unlocked strategies and dissipative brains. The abductive eco-cognitive perspective. *BRAIN. Broad Research in Artificial Intelligence and Neuroscience, 15*(4), 354-374. https://doi.org/10.70594/brain/15.4/24

1. Introduction

In a previous book Magnani (2022b) I have proposed what I called “eco-cognitive computationalism”, which explores computing in context, adhering to some of the key ideas presented by modern cognitive science perspectives on embodied, situated, and distributed cognition. I think this intellectual framework is especially useful to depict various conceptualizations of machines that I consider basic and important because they offer 1) a naturalization of the concept of computation and 2) an antidote against dogmatic definitions or philosophical assumptions tending toward metaphysical outcomes or excessive ontological claims.

A first section

- *Alan Turing: conceptualising machines as unorganised brains that can be computationally organised mimicking human education*

is devoted to the fundamental conceptualization of machines furnished by Turing himself. Turing emphasizes the role of the “education” of the machine, paralleling it to the education of human brains, as the way for “educating” a machine to “think”. It is exactly this Turing’s emphasis on education that furnishes the justification of the conceptualization of machines as “domesticated ignorant entities”, that is proposed in this article.

The section

- *Conceptualising computational machines as physical systems*

shows that conceptualising machines as dynamically active in distributed physical entities of various kinds suitably transformed so that data can be *encoded* and *decoded* to obtain appropriate results further sheds light on a naturalised eco-cognitive perspective. To take advantage of this conceptualisation of machines as physical systems it is also important to illustrate a comparison between physical computing and physics, showing that the first is the reverse of the second, as described in the subsection

- *The illuminating comparison between physical computing and physics*

and that at least from the classical perspective of computational machines as digital machines, the physical entities exploited for computation are technological artefacts, see subsection

- *Computational physical entities are technological artefacts*

Moreover, it is within this intellectual framework that it is important to quote the recent attention in computer science devoted to the importance of the simplification of cognitive and motor tasks caused in organic entities thanks to morphological features.

Through eco-cognitive computationalism, we can definitely acknowledge that the concept of computation evolves over time as a result of historical and contextual factors, and we can construct an epistemological view that depicts the “emergence” of new types of computations that exploit new substrates. This new viewpoint demonstrates how the computational domestication of ignorant entities might result in the emergence of novel unconventional cognitive embodiments (morphological computing, DNA computing, quantum computing, etc., as illustrated by a recent area of research called “natural computing”.

Furthermore, in this article, the key terms are referred to as 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 selects, this is important to be noted) 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 like for example 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 characterised 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 another important concept introduced in this article. Instead, this openness characterizes unlocked human cognition in terms of higher types of abductive reasoning, both in 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, an "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. Alan Turing: Conceptualising Machines as Unorganised Brains that Can Be Computationally Organised Mimicking Human Education

Eco-cognitive computationalism considers computation as something active in physical entities that have been adequately changed into what I call "cognitive mediators", in which data may be encoded and decoded to generate beneficial outcomes. We must state right away that, even if cognition is viewed as computational, eco-cognitive computationalism does not equate computation to digital computation, that is the processing of digits following appropriate rules. We will see that other substrates can carry computations: it is reductive to conceptualize computation in digital terms.

I think that to clarify the concept of computation it is necessary to provide a new analysis of the related concepts of information and cognition. No static and unchangeable definitions of the concepts of information, cognition, and computation will be provided because eco-cognitive computationalism takes into account the importance of historical and variable aspects, which lead to constant amendments of the status of their meanings.

As the title "Alan Turing conceptualised machines as unorganised brains that can be computationally organised mimicking human education" of the present section illustrates, the foundational conceptualization that led to the birth of modern computer science with the final establishment of both the Universal Logical Computing Machine and the Universal Practical Computing Machine (this last concept obviously refers, for example, to the digital laptop that we use every day) is due to some Turing's speculations. Turing himself describes a kind of proto-eco-cognitive computationalism that sees the evolutionary emergence of information, meaning, and the primitive forms of cognition as the outcome of a complicated eco-cognitive intertwining and simultaneous *coevolution*, temporally oriented, of the states of brain, body, an external surroundings. Indeed Turing explicitly states – adopting a philosophical approach – that the brain of a child is an "unorganised machinery", which may be organised by adequate "interference training". The organization might lead to the machine being modified into a universal machine or something similar.

Turing adds that "From the standpoint of evolution and genetics, this depiction of the cortex as an unstructured apparatus is highly satisfying" (Turing, 1969, p. 16), a statement that is actually endowed with a speculative value: indeed I should add that Turing's depiction of an infant's brain as an unorganised machine is mainly theoretical, and it primarily serves as a useful heuristic in his creative cognitive processes. Indeed, we know from fundamental neuroscientific studies that the brain is instead a highly ordered system. Turing further observes that babies' cortices are "socially" constructed through language. The proposed random building of the baby brain, in this view, does not address neurobiological issues but rather the idea of an unorganised brain as fundamentally characterised by a lack of training, i.e., a lack of indications received through the senses from the external (including social) environment, such as language. As a result, the term "unorganised machinery" has a distinct meaning: indeed Turing says "It is conceivable that the same machine might be regarded by one man as organised and by another as unorganised" (Turing, 1969). Further details are illustrated in Magnani, (2022b), chapter one.

Turing provides the conceptual framework necessary to demonstrate how, by replicating the aforementioned procedure, the Universal Logical Computing Machine and the Universal Practical

Computing Machine may be created. "Computationalizing" a machine is a matter of *education*, exactly as in the case of the human brain. In this respect, the dual innovation is realised as the externalisation (a kind of domestication through education) of computational faculties in those artefactual physical entities that compute for some human or artefactual agents: those computers that I dubbed "mimetic minds" (Magnani, 2006), taking advantage of the viewpoint supplied by Turing.[1]

It is spontaneous for Turing to use the notion of education as it is occurring in the case of the growth of babies' brains to develop the new concept of computing: physical machines may also be *educated* to generate specific types of behavior. In this perspective, education and programming coincide. Yes, programming is similar to human education in that it allows us to computationally domesticate machines in order to get suitable results. The new notion of the (Universal) Logical Computing Machine (LCM) belongs to this intellectual framework together with the companion one of the (Universal) Practical Computing Machines (PCM), which is the implementation of the first in a digital computer.

3. Conceptualising Computational Machines as Physical Systems

In the first chapter of my book (Magnani, 2022b) I have illustrated some aspects of the relationships between the concepts of information, cognition, and computation, attempting to answer the question "Is cognition computation?" We may say yes, it is, but it is not merely computing, thus cognition and computation are not interchangeable. We can affirm that both information processing and computation are intertwined with cognition, and a lot of studies have been realised to clarify the various roles and types of computation and information processing implicated in cognition, even if such research is doomed to become obsolete, at least from an eco-cognitive point of view. I prefer to assert, from a naturalistic epistemological perspective, that, given the fact that the concept of computation changes when new research results are ready, and so it is sensible to semantic variations, the other two concepts, which are already illustrated in themselves by precarious definitions, that are linked to changes of knowledge, technologies, and cultural ideas, also change.

A recent perspective I especially favor tries to elucidate the notion of digital computation by clearly demonstrating and emphasising that digital computation is implemented in *physical systems*.[2]

My eco-cognitive computationalism does not try to provide a firm and unchangeable description of the concepts of information, cognition, and computing. Instead, I propose an epistemological view that depicts how we can grasp modifications and expansions of their meanings through the example (and the illustration of the "emergence") of what can be called the *new* "computational domestication" of physical entities, thanks for example to the new exploitation for computation of new substrates, such as in the case of morphological computing. To that end, I shall take the viewpoint on physical computation presented (Horsman et al., 2014), which is especially suited to my concerns:

1. a computer is a physical system having real constituent elements and internal interactions that allow it to transition from one physical state to another;
2. as a result, I agree with Horsman et al., who claim that physical computing is "the use of a physical system to predict the outcome of an abstract evolution" (Horsman et al., 2014, p. 14). I have to add that it is worth noting that in computations, we are not dealing with a physical system that has to be explained (as in physics), but rather with an abstract object that we wish to evolve using the physical system itself. When we recognize that a physical entity is being exploited in such a way, physical evolution executes a computation that is then

[1] In chapter one of (Magnani, 2022b) the idea of mimetic mind is discussed in detail.

[2] These investigations provide the advantage of avoiding the so-called pancomputationalism, that is the philosophical view that attributes computational capacities to all natural entities, that is, the belief that every physical system is a digital computing system that can be defined in computational terms, which I consider metaphysically biased; cf. especially the part "Pancomputationalism naturalized" of chapter two of (Magnani, 2022b).

interpreted as such by a human or another artificial agent, (together with its cognitive content);
3. standardly a computer is a technology created by applying scientific concepts to precision-engineer physical systems to get certain expected results required specs (see Section 3.2 below for further information);
4. when a special link between abstract mathematical/logical entities and the physical ones is realised, we can conclude that a physical system (for example a digital computer) is computing;
5. contrarily to the previous observations concerning computation as physical, and beyond their use in conventional (digital) and eccentric cases (morphological, quantum, DNA computing, etc.) the concept of *computation*, and its related system properties, *information*, have been imported into other scientific areas into the aim of making intelligible and "explain" various natural processes such as photosynthesis and the consciousness, and a considerable quantity of contemporary cross-disciplines has proposed the conclusion that "everything is information", "the universe is a quantum computer", or "everything computes" (Horsman et al., 2014. p. 2). Clearly, many scholars believe that by characterising the cosmos and everything in it as a computer, the notion of *physical computation* becomes *empty*. Cf. also the section "Pancomputationalism 'naturalized" of chapter two of (Magnani, 2022b).[3]

3.1. The Illuminating Comparison between Physical Computing and Physics

Following the perspective suggested by the already cited Horsman et al. and as quickly indicated in the previous section, physical computing may be advantageously seen as the opposite of mathematical physics: *a physical system is used to predict the result of an abstract dynamics*, whereas in physics, *an abstract model is used to predict physical dynamics*. Indeed, physics is engaged in abstractly modeling physical systems thanks to the exploitation of abstract mathematical theories to predict the outcome of physical evolution.

Again: if physical computing is the opposite of mathematical science, and is related to the exploitation of *a physical system to predict the outcome of abstract dynamics*, then a physical system that lacks this predictive aspect is not a computer; similarly, a cluster of mathematical equations is a bad physical model if it lacks predictive power. As a result, in physical computation, we wish to take an abstract thing, a computation, and represent it into a physical entity. We face a *representational* process, which allows for a "comparison between physical processes and mathematical described computations" (Horsman et al., 2014). On the contrary, in physics, just to make an example, an electron is represented as a wave function, which serves as a map between physical and abstract spaces: in a computation, "the initial impetus is not a physical system that needs to be described, but rather an abstract object that we wish to evolve. An abstract problem is the reason why a physical computer is used" (cit., p. 10). As a result, we begin with the problem of a reversed representation right away.

3.2. Computational Physical Entities Are Technological Artefacts

The creation of what Turing referred to as Universal "Practical" Computing Machines is absolutely linked to the role of suitable technologies capable of "domesticating"[4] physical artefacts to act in expected ways: after all, in order to possess appropriate physical computations, we need technical help, to build physical artefactual entities capable of completing the work.

It is surely an astonishing result the fact that Turing's abstract machine, as a Discrete-State Machine (DSM) (Longo, 2009), has been implemented in a physical machine, thanks to the

[3] The Actor model (Hewitt, 2012) is an example of an alternative unorthodox view: computation is conceptualised as scattered in space, where computational devices communicate asynchronously and the entire computation is not in any well-defined state. Reacting to a message an Actor may make local choices, produce other Actors, send new messages, and determine how to reply to the next message received. Turing's model is a subset of the Actor Model. According to this viewpoint, the Internet conducts unconventional computations and appears to be governed by super-recursive algorithms (Dodig-Crnkovic, 2011).

[4] The so-called "domestication of physical entities" will be described in the following section 4.

building of artificial technological processes that can evolve through discrete states, taking advantage of the well-known silicon electronic devices. We also know that at least starting from Turing, computer science explained that physical artefactual digital entities are created to process information and knowledge reorganising them in little boxes, bits, and pixels, that is in discrete packages, so avoiding fuzziness, contingency, and chaotic deterministic aspects, in which a fluctuation/variation below the measuring interval produces multiple and different evolutions for a system.

Turing emphasizes in his concept of Discrete State Machine (DSM), already in 1950, that this effect is potentially avoidable: the Turing machine is an abstract (logical, as he put it) computer with a Laplacian behavior; as a DSM, it manages values that are invariably discrete. The notion of a program, as well as its mathematical properties connected to its implementation, are deterministic in the Laplace sense: the application of rules (or, in the case of Laplacian physics, equations) implies predictability. A scientific theory aims to reach success when confirmed practically exempt from huge anomalies, explanatorily satisfactory, and standardly capable of predicting the behavior of a lot of processes, it represents the best example of the prototype of the success of a progressive scientific research program, as illustrated by (Lakatos, 1970). These kinds of theories are used in engineering and so are fruitful epistemic tools capable of engineering physical entities, that are in such a way "domesticated" to obtain a desired reliable behavior.

4. Conceptualising Machines as the Fruit of a Domestication of Ignorant Substrates

It is well-known that today's conventional computing substrates are highly engineered silicon devices whose processes are based on the stable knowledge of those scientific theories that also contributed to originating them. A current trend in computer science and AI is related to the aim of exploiting more and more physical substrates capable of supporting computational performances and considerably reliable in granting good outcomes of the related processing of data.

We have said in the first section of this article that physical computing is the opposite of mathematical science, the most important task is to "educate" a physical entity, beyond digital computation, to predict the outcome of abstract dynamics and not the contrary (that is the case of an abstract tool capable of predicting a physical behavior, an equation for example). As a result, understanding the physical system's predictive capabilities is critical. Some cases do not fit this requirement: a stone or a piece of marble will not be suitable to perform in time modifications of themselves capable of matching a reliable and repeatable computational passage from initial states represented in them to final states that should be picked up later on, but other unsuspected entities are not so "unusable" to computational aims…

4.1. Biological Systems Compute Information Units in Unconventional Ways

In the context of the so-called morphological computation, bodies that are "ignorant" become mimetic bodies, that is, bodies that may imitate different cognitive mediating activities. I use the term "mimetic" in the same way that Turing uses it, as I have illustrated in the first sections of this article: "mimicking education, we should hope to modify the machine until it could be relied on to produce definite reactions to certain commands" (Turing, 1969) Turing adds. Analogously, we may acquire computational results mimicking morphological features: a classic example is provided by (Nakajima et al., 2013; Nakajima et al., 2015), who "purely" mimicked part of an octopus morphology using an artificial model of a soft robotic arm motivated by it. Considering recent morphological computation breakthroughs, we can safely infer that we are confronted with new epistemological and technological dimensions of what I have named distributed computation (Magnani, 2022b): to achieve novel cognitive outputs and outcomes, computing is more and more being dispersed over a wide range of entities, props, instruments, bodies, and gadgets. The potential of morphological computation principles in robot design might lead to a new generation of robots with more adaptability and fewer control parameters. The possible suggestions derived from morphological computation research in robot design might result in a new variety of robots with improved flexibility and a smaller set of control parameters.

From the perspective of my eco-cognitive paradigm, an intriguing instance of the exploitation of biological organisms to conduct computation must be highlighted. Biological beings may compute unique forms of informational units, units that can be found in unexpected and unusual computing systems, that are not comparable to the ones I discussed in the preceding section on morphological computation. I am referring to the contribution of the so-called DNA computers, which are a relatively well-known type of non-silicon-oriented computation. Molecular computing is of course an area of study situated at the intersection of several approaches that developed techniques and methods to "domesticate" molecules in order to obtain the desired computation, or manufacture the desired product, or govern the functioning of a given molecular system. (Rozenberg et al., 2012) usefully observes that the core notion underlying molecular computing is that data may be encoded as (bio)molecules, such as DNA strands, and then transformed using molecular science methods. So to speak, a molecular program is a collection of molecules that will play a particular role, that is, the one of running the program represented by this collection when placed in a suitable substrate. It is a tradition of studies that also promoted and enhanced the possibility of making intelligible some central aspects of nanosciences, such as for example self-assembly.

To summarize, DNA computing is based on the notion that molecular biological processes may be used to execute arithmetic and logical operations on data recorded as DNA strands (Hagiya et al., 2016). DNA bio-operations can be employed for computations, resulting in a biocomputation consisting of a series of bio-operations: it is obvious to highlight the fact that DNA-encoded information differs significantly from silicon-based electronically encoded information, and bio-operations clearly vary from electronic computer procedures.[5] Sensors, processors, and actuators, all formed of molecular devices, make up a molecular robot, which reacts autonomously to its surroundings by checking it, making choices with its computers, and performing actions on it. The intelligent controllers of such molecular robots should thus be molecular computers. To create our "DNA computers", as Kary (2013) calls them, the most crucial comparison that had to be established to unveil the secret of DNA computing, was the one between two procedures, one biological and the other mathematical: (a) the result of applying simple operations (copying, splicing, etc.) to initial information encoded in a DNA sequence produces the complicated structure of an organic being, and (b) the result f(w) of applying a computable function to an argument w can be discovered using a mixture of fundamental simple functions to w.

The prosperity of research on computational domestication of new substrates, as demonstrated by morphological computation, DNA, and quantum computing, among other examples, also adumbrates some philosophical/epistemological concerns intended to show some skepticism regarding the consistency of ultimate viewpoints such as pan-cognitivism, pan-informationalism, and pan-computationalism, which are relatively popular in the current debate about computational philosophy. After all, it seems plausible to state that nature does not compute in itself, and information is not an intrinsic characteristic of it. Also, cognition cannot be seen as a constitutive attribute, neither of organic nor of inorganic entities. I think, following a naturalistic approach, it is only the huge "semiotic" human activities that create a world of computational domestications of ignorant entities, a world of meanings carried by cognitive representations in which the life of informational aspects is rich and fundamental.[6]

[5] Further details concerning this difference can be usefully found in (Kari et al., 2012). On further taking computational advantage of several artificial molecular "machines", encompassing some composed of DNA or RNA to construct robotic systems cf. (Hagiya et al., 2016, p. 6).

[6] More details on these last aspects that regard computational philosophy can be found in (Magnani, 2018a, 2018b, 2022b).

5. Deep Learning Cognitive Strategies Are Locked: An Abductive Eco-Cognitive Perspective

Seeing some recent deep learning programs provided by the successful developments of AI further reveals new interesting aspects of computation. 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 19x19 board. In March 2016, Google also defeated Lee Sedol, a 9 dan player considered the best champion on 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 generating new strategies and so phenomenologically simulating human beings' performances. This deep learning program "learned" a lot by attending thousands of games played by humans, such as the one played by Lee Sedol, thanks to the so-called "reinforcement learning": the program in turn plays repeatedly against itself to further strengthen its own deep neural networks.[7]

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 ways that agents adopt when dealing with other agents to the variegated related intertwined or collective cognitive actions. In turn, ecological thinking (or ecological rationality) (Gigerenzer & Selten, 2002; Raab & Gigerenzer, 2005; Gigerenzer & Brighton, 2009) 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 AI considerably shared view that sees strategies as a composition of successive selections of appropriate heuristics.

I think that it is in the perspective of the studies on abductive cognition (Magnani, 2009) that we can suitably and profitably consider what I just called strategic cognition stressing the contrast between what I called *locked* and *unlocked strategies* (Magnani, 2019). Some more words have to be added to illustrate some aspects of the concept of abduction. A hundred years ago, Charles Sanders Peirce (Peirce, 1866–1913) 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. In the context of scientific discovery, 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 by using classical syllogistic logic) – the abductive inference rule corresponds to the well-known fallacy called affirming the consequent. Many reasoning conclusions 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, 1958), and in the area 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.[8] Six volumes (monographs and collections) are currently available (Josephson and Josephson, 1994, Flach and Kakas, 2000, Kuipers 2000, Magnani, 2001, Gabbay and Woods, 2005, Aliseda, 2006, Walton, 2004 and three special issues of international journals (*Philosophica*,

[7] AlphaGo Zero is a version of DeepMind's Go software AlphaGo; the recent AlphaZero further enriches AlphaGo Zero, cf. `https://en.wikipedia.org/wiki/AlphaZero, and learns by its own played games.`

[8] A list of the classical bibliography on abduction is given in (Magnani, 2001).

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.[9] Finally, I have to quote the two volumes of the *Handbook of Abductive Cognition* (Magnani, 2023) I recently edited, a first comprehensive reference guide to the interdisciplinary field of abductive cognition that offers a logical, cognitive, epistemological, and philosophical analysis of abductive reasoning, covering hypothetical, creative, diagnostic, model-based, and adversarial inferences.

Before further describing the details of the concept of abduction, I invite the reader to consider this anticipation of mine: deep learning machines (and so AlphaGo/AlphaZero programs) are characterised by locked strategies, a fact that strongly impacts the kind of creativity that is realised 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 that important concept. I introduced two kinds of abduction, *selective* (Magnani, 2001) – 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), I already quoted above. Not only, 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 thanks to 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 originate new data – that were not available before – and new heuristics capable of enhancing the ways agents adopt for solving 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. (Magnani, 2009)). In the case of deep learning machines (and in the case of the games we are considering in this article) it is patent we face cases of manipulative abduction: the cognitive processes are intrinsically related to the manipulation of the stones of the game Go and several cognitive embodied moments are at stake, together with the needed visualization of the entire external context, the competitor, etc.

5.1. Abductive Cognition and AlphaGo/AlphaZero

It is well-known that abduction research has increased knowledge about creative cognition, also when dealing with the simple case of a novel move in a Go game. The two concepts I have introduced during my studies on the abduction of knowledge-enhancing abduction and eco-cognitive openness (Magnani, 2017) are very appropriate to examine the locked and unlocked abductive strategies I have introduced above. As I already said, locked and unlocked strategies are fundamental 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, that 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/AlphaZero, which aim at mechanising 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

[9] General classical considerations on abduction in science and AI can also be found in Gooding (1996); Josephson and Josephson (1994); Kuipers (1999); Thagard (1988); Shrager and Langley (1990).

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 (Magnani, 2019). Indeed, abduction is also occurring at a sub-symbolic level, both in humans and machines, and so no surprise to see deep-learning machines capable of leading to abductive results. Not only that, we have to remember that abduction is also characterised by multimodality (for example, it can be performed by exploiting diagrams), as I have described in (Magnani, 2009; Magnani, 2017). We have also 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 always presented various methods, formalisms, and algorithms capable of leading to the construction of programs related to abductive performances.[10]

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

6. Natural, Artificial, and Computational Games

6.1. Locked and Unlocked Strategies in Natural and Artificial Environments

Go is a game played by human agents and AlphaGo/AlphaZero is an automatic computational deep learning program that also plays that game and that is capable of competing with humans. We can say that Go is already something "artificial" because it is created by human agents that established rules, a specific board, and other material entities, that is stones. AlphaGo/AlphaZero is even more artificial, a fruit of technological creativity, an engineered product fruit 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 (I will soon describe them) are at play, as wonderfully exemplified by the phenomenological tradition, a game encompassing embodied aspects and distributed cognitions, and also visual, kinesthetic, and motor sensations (Magnani, 2018c). This "natural" game is characterised 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, just to make an example, by a specific (unchangeable in its structure) Go board, which indeed renders cognitive strategies *locked*. A fixed board, determined stones and rules, render the scenario strongly constrained, and locked.

6.2. Reading Ahead as an Abductive Engine

Reading ahead, as the Go players standardly say, is the generation of anticipations aimed at being very strong, and reliable (both deeply-minded or intuitive). They involve

1. groups of potential moves to be selected and their possible outcomes. A present scenario at time, shown by the board, "adumbrates" – Husserlian concept! – a posterior possible more productive scenario at time , 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, the interaction between distant stones, keeping the whole board in mind during local

[10] 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 (Ramoni et al., 1992).

fights, and other issues that involve the overall game. It is therefore possible to allow a tactical loss when it confers a strategic advantage".[11]

All these strategies, even multiple and variegated, are "locked": indeed the entities of each scenario remain the same during the whole process of the game. The only modifications that are allowed are the ones concerning the number of possible stones at play and their position 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 his strategies – internal sources not ineluctably connected to the previous time spent playing with 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 less degree of closure that in the case of the computational program AlphaGo. In humans, strategies are locked with respect to the external fixed scenario, but more open concerning the mental references to other available extended strategic endowments; in AphaGo/AlphaZero 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/learned 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.[12]

7. 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 (Magnani, 2017), ranging from the performances 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 also an automatic game player such as AlphaGo/AlphaZero 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 we can reach exceptional selective or creative optimal abductive results, that is we need what I have called (Magnani, 2017) *optimization of eco-cognitive situatedness*, in which eco-cognitive openness is fundamental Magnani (2016). An optimization of eco-cognitive situatedness is particularly necessary in the case of creative abduction: in a recent book (Magnani, 2022a) I have also emphasised 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 10 below I will 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, 2007) would have called rigid and finite definitory rules, and by finite material entities, which would fixedly work as cognitive mediators capable to govern agents' reasoning.

It is in (Magnani, 2015, 2016, 2017) that I stressed the relevance in good abductive cognition towards hypotheses of what 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 optimisation of situatedness, which can only be built by a *maximization of the changeability* of the starting data which characterizes the abductive cognitive processes: inputs have to be maximally expanded, reconstructed, or changed and the same

[11] Cf. Wikipedia, entry Go (game) `https://en.wikipedia.org/wiki/Go_(game)`.

[12] I have to note that my notion of locked strategy is not related to the standard nomenclature of the game theory, for more synthetic information see, for example, Wikipedia entry Go (game) `https://en.wikipedia.org/wiki/Go_(game)`.

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*.[13] In rundown, abductive processes to hypotheses – in an appreciable quantity of cases, for example in science – are highly *information-sensitive* and confront with a stream of information and data that has to be continuous and fittingly promoted and upgraded when needed.

8. The Physics of Eco-Cognitive Openness

8.1. 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* (Vitiello, 2001, 2012, 2019b, 2019a), related to the extension of the quantum model of the brain to the dissipative dynamics (Prigogine and Stengers, 1984). Indeed, in this perspective two aspects 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 there can be built a full set (a space) of other states of nonzero energy. 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" (Vitiello, 2012, p. 316). In 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 (Vitiello, 2012, pp. 318-319).

Consequently, the order is a *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.[14]

As I have already said the mechanism of spontaneous breakdown of symmetry is responsible for the order, as the product of *self-organising dynamics*: "The process of symmetry breaking is triggered by some external input; the 'choice' of the specific symmetry pattern which is actually realised is, on the contrary, 'internal' to the system. Therefore one speaks of self-organising 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" (Vitiello, 2012).

Given the fact that in the quantum model of the brain, a particular memory is associated to a specific degree of ordering (that is to a specific value of the vacuum code), and given the fact the

[13] I have furnished more cognitive and technical details to explain in (Magnani, 2016) and in the recent book (Magnani, 2022a).

[14] 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.

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. Not only, it is important to say that, differently than in the case of a computer, *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.[15] Furthermore, the brain state evolution is a "story" of the "irreversible" intertwining of it 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 degrees of freedom. The doubled degrees of freedom refers 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 result characterised as the "time-reversed copy" (the Double) of the brain.[16] The environment is constantly reverberated in a "model" in the brain,[17] 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 characterised 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" (Vitiello, 2012) and 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 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 be also 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 (Hutchins, 2005, 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 stability of the system. 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:

[15] It is well-known that isolation of an individual tends to generate various pathologies, not only at the psychological level.

[16] The physical model also explains that the existence of the first cannot be independent of the existence of the second, and vice versa. The system "brain/environment" is treated as the closed system "brain and its Double".On the relationship with the Double as a route to consciousness and on the role of objectiveness of the external world as the primary, necessary condition for consciousness to exist, see (Vitiello, 2012, p. 328–335);; consciousness would be rooted in a restless dialog – entanglement – of the self with its Double.

[17] Cognitive science also stressed this fact when Brooks once observed that at the root of the more basic forms of cognition, it can be hypothesised that the "world serves as its own best model" (Brooks, 1991).

> […] 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 (Vitiello, 2012, p. 326).

8.2. Abductive Errors Vindicated

A dissipative brain elaborates the relationship with its surroundings through cyclic processes of trial and error, that is 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)" (Vitiello, 2012).[18] The dissipative brain is an *erratic* brain (Minati and Vitiello, 2006), in which the uniqueness of its identity and 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" (Peirce, 1866–1913): this means that human abductive cognition is "akin to the truth" but not exempt from failures, making mistakes is 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 do 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 dysfunction), of the new most highly creative ones.

Turing machines are not allowed to make 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 fluctuations processes of the dissipative brain are affected by unpredictability. An unpredictability is constitutively linked to mistakes as (Minati and Vitiello, 2006, p. 71) further observe:

> Sometimes mistakes are useful for introducing 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's 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.

9. Big Data: Huge but Locked

In subsection 8.1 above I have quoted an interesting passage, written by Vitiello, regarding the so-called dissipative brain. Here again the text: "The process of symmetry breaking is triggered by some external input; the 'choice' of the specific symmetry pattern which is actually realised is, on the contrary, 'internal' to the system. Therefore one speaks of self-organising dynamics: ordering is an inner (spontaneous, indeed) dynamical process. […] in the brain, contrary to the computer

[18] Fluctuating background jeopardizes the precise determination of the trajectories from initial conditions in a totally unpredictable way.

case, ordering is not imported from the outside, it is the outgrowth of an 'internal' dynamical process of the system" (Vitiello, 2012, pp. 316–317).

If we accept that in the brain, contrary to 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 into the machine from the outside. This order is of course governed by the related software but it 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 itself has previously generated, we face the computational manipulations of data that consist of a huge quantity of Go games, games, and 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 5 above – character of the informational scenario that feeds the system aking for more openness in the reservoir of data.

It is not a 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-contextualization* of facts from their context of origin to the aim of improving the quality of their eventual computational manipulation. Just to make an example, in the case of biological data "One of the main tasks of database curators is to decontextualize 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-contextualization of most types of biological data" (Leonelli, 2014, 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 task 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 grant writing" (Leonelli, 2014, p. 4).[19]

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), 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 in 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" (Leonelli, 2020).[20] A serious criticism of the abductive capacities of computational performances on large databases is contained

[19] Further details can be found in (Leonelli, 2020).

[20] On the general problem of discoverability and its discontents, which encompasses the present one regarding the curation of big data, see my already quoted book (Magnani, 2022a).

in an article written by (Calude and Longo, 2017). 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.[21]

10. Locked Strategies Limit Creativity

We have seen above that optimization of situatedness is related to unlocked strategies. Locked strategies, that characterize 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 subsection 6.2 above. I have said that in the game Go stones, board, and rules are fixed and so predetermined. We have to add that, in general, the presence of rigid and unchanging scenarios in natural and artificial world, even if obviously related to the presence of stabilities which in turn give origin to habits of behavior, surely tends 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 his reasoning strategies[22] 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 parts 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, all the remaining aspects remain immutable.[23] 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 of 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.[24]

[21] A recent interesting discussion about 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 (Kawamleh, 2021). 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 instead described by (Vallverdú, 2020). Finally, the readers that 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 (Numerico, 2021), that especially stresses the fact that more or less reliable predictions generated by computational systems become *de facto* "prescriptions" capable to surreptitiously modify human behaviors.

[22] Many interesting examples are illustrated in the (Magnani and Bertolotti, 2017).

[23] 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.

[24] Some data regarding the history of the so-called automated scientific discovery in AI are illustrated in (Magnani, 2009, chapter two, section 2.7 "Automatic Abductive Scientists").

Many managers are prospecting an exploitation of deep learning programs to aid science resolve relevant real-world problems in healthcare but also in other fields such as scientific reasoning. Many implementations of business thanks to the 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`), [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 did 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. As I have already anticipated, (Calude and Longo, 2017, 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.

Conclusion

In the first part of this article, taking advantage of what I called eco-cognitive computationalism, I have illustrated various ways of conceptualising machines, starting from the fundamental role Turing assigned to the process of "education" of a machine, paralleling it to the education of human brains, that wonderfully favored the invention of the Logical Universal Machine. It is this Turing's emphasis on education that furnishes the justification of the conceptualization of machines as "domesticated ignorant entities", that I advanced in this article. I have described that conceptualising machines as dynamically active in distributed physical entities of various kinds suitably transformed so that data can be *encoded* and *decoded* to obtain appropriate results sheds further light on my eco-cognitive perspective. I also quoted the "morphological computation" (and other unconventional types of computation), in which the importance of the simplification of cognitive and motor tasks caused in organic entities by the morphological features is computationally exploited: I have said that ignorant bodies can be computationally domesticated to render an intertwined computation simpler, resorting to that "simplexity" of animal embodied cognition, which represents one of the main quality of organic agents. Finally, I stressed that we cannot reach a stable and reliable conceptualization of machines because the concept of computation changes, depending on historical and contextual causes, and I have illustrated the "emergence" of new kinds of computations that take advantage of the exploitation of new substrates, also sketching some aspects of the so-called "natural computing".

In the second part of this article, I have illustrated the concepts of locked and unlocked strategies, abduction, and optimization of eco-cognitive openness, as key concepts capable of highlighting the fundamental epistemological and cognitive differences between deep learning computational cognitive performances (I considered the example of the program Alphago/AlphaZero) and the 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 the 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 a direct product of the "internal" open dynamical process of the system.

Acknowledgements

Parts of this chapter are excerpted with slight modifications from the two articles Human abductive cognition vindicated: computational locked strategies, dissipative brains, and eco-cognitive openness, *Philosophies* 2022, 7, 15 (`https://doi.org/10.3390/philosophies7010015`) and Conceptualising machines in an eco-cognitive perspective, *Philosophies* 2022, 7, 94. (`https://doi.org/10.3390/philosophies7050094`) (both (CC BY) license (`http://creativecommons.org/licenses/by/4.0/`)). For the informative critiques and interesting exchanges that assisted me to enrich my analysis of the issues treated in this article, I am obligated to John Woods, Atocha Aliseda, Woosuk Park, Giuseppe Longo, Gordana Dodig-Crnkovic, Luís Moniz Pereira, Paul Thagard, Joseph Brenner, to the two reviewers, and to my collaborators Selene Arfini and Alger Sans Pinillos.

References

Aliseda, A. (2006). *Abductive Reasoning. Logical Investigations into Discovery and Explanation*. Springer, Heidelberg/Berlin.

Brooks, R. A. (1991). Intelligence without representation. *Artificial Intelligence*, *47*, 139–159. https://doi.org/10.1016/0004-3702(91)90053-M

Calude, C. S. & Longo, G. (2017). The deluge of spurious correlations in big data. *Foundations of Science*, 22(3), 595–612. https://doi.org/10.1007/s10699-016-9489-4

Dodig-Crnkovic, G. (2011). Significance of models of computation, from Turing model to natural computation. *Minds and Machines*, *21*(2), 301–322. https://doi.org/10.1007/s11023-011-9235-1

Flach, P., & Kakas, A. (Eds.). (2000). *Abductive and inductive reasoning: Essays on their relation and integration*. Kluwer Academic Publishers.

Gabbay, D. M. & Woods, J. (2005). *The Reach of Abduction*. North-Holland, Amsterdam.

Gigerenzer, G., & Brighton, H. (2009). Homo heuristicus: why biased minds make better inferences. *Topics in cognitive science*, *1*(1), 107–143. https://doi.org/10.1111/j.1756-8765.2008.01006.x

Gigerenzer, G. & Selten, R. (2002). *Bounded Rationality. The Adaptive Toolbox*. The MIT Press, Cambridge, MA.

Gooding, D. (1996). Creative rationality: Towards an abductive model of scientific change. *Philosophica*, 58(2), 73–102. https://doi.org/10.21825/philosophica.82334

Hagiya, M., Aubert-Kato, N., Wang, S., & Kobayashi, S. (2016). Molecular computers for molecular robots as hybrid systems. *Theoretical Computer Science*, 632, 4–20. https://doi.org/10.1016/j.tcs.2015.11.002

Hanson, N. R. (1958). *Patterns of Discovery. An Inquiry into the Conceptual Foundations of Science*. Cambridge University Press, London.

Hewitt, C. (2012). What is computation? Actor model versus Turing's model. In Zenil, H., editor, *A Computable Universe. Understanding and Exploring Nature as Computation*, pages 159–185, Singapore. World Scientific.

Hintikka, J. (2007). *Socratic Epistemology. Explorations of Knowledge-Seeking by Questioning*. Cambridge University Press, Cambridge.

Horsman, C., Stepney, S., Wagner, R. C., & Kendon, V. (2014). When does a physical system compute?. *Proceedings. Mathematical, physical, and engineering sciences*, *470*(2169), 20140182. https://doi.org/10.1098/rspa.2014.0182

Hutchins, E. (2005). Material anchors for conceptual blends. *Journal of Pragmatics, 37*(10), 1555–1577. https://doi.org/10.1016/j.pragma.2004.06.008

Josephson, J.R. & Josephson, S. G., editors (1994). *Abductive Inference. Computation, Philosophy, Technology*. Cambridge University Press, Cambridge.

Kari, L. (2013). DNA computing based on insertions and deletions. oai:CiteSeerX.psu:10.1.1.297.3308.

Kari, L., Seki, S., & Sosík, P. (2012). DNA computing: Foundations and implications. In Rozenberg, G., Bäck, T., and Kok, J., editors, *Handbook of Natural Computing*, *3*, 1073–1128. Springer, Cham, Switzerland. https://doi.org/10.1007/978-3-540-92910-9_33

Kawamleh, S. (2021). Can machines learn how clouds work? The epistemic implications of machine learning methods in climate science. *Philosophy of Science*, *88(*5), 1008–1020. https://doi.org/10.1086/714877

Kuipers, T. (2000). *From Instrumentalism to Constructive Realism. On Some Relations between Confirmation, Empirical Progress and Truth Approximation*. Kluwer Academic Publisher, Dordrecht.

Kuipers, T. A. F. (1999). Abduction aiming at empirical progress of even truth approximation leading to a challenge for computational modelling. *Foundations of Science*, *4*, 307–323. https://doi.org/10.1023/A:1009688227714

Lakatos, I. (1970). Falsification and the methodology of scientific research programs. In Lakatos, I. and Musgrave, A., editors, *Criticism and the Growth of Knowledge*, pages 365–395. The MIT Press, Cambridge, MA.

Leonelli, S. (2014). What difference does quantity make? On the epistemology of Big Data in biology. *Big Data and Society*, 1(1). https://doi.org/10.1177/2053951714534395

Leonelli, S. (2020). Scientific research and big data. In Zalta, E.N., editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, Summer 2020 edition.

Longo, G. (2009). Critique of computational reason in the natural sciences. In Gelenbe, E. and Kahane, J.-P., editors, *Fundamental Concepts in Computer Science*. Imperial College Press/World Scientific, London.

Magnani, L. (2001). *Abduction, Reason, and Science. Processes of Discovery and Explanation*. Kluwer Academic/Plenum Publishers, New York.

Magnani, L. (2006). Mimetic minds. Meaning formation through epistemic mediators and external representations. In Loula, A., Gudwin, R., and Queiroz, J., editors, *Artificial Cognition Systems*, pages 327–357. Idea Group Publishers, Hershey, PA.

Magnani, L. (2009). *Abductive Cognition. The Epistemological and Eco-Cognitive Dimensions of Hypothetical Reasoning*. Springer, Heidelberg/Berlin.

Magnani, L. (2015). The eco-cognitive model of abduction: Ἀπαγωγή now: Naturalizing the logic of abduction. *Journal of Applied Logic*, 13, 285–315. https://doi.org/10.1016/j.jal.2015.04.003

Magnani, L. (2016). The eco-cognitive model of abduction. Irrelevance and implausibility exculpated. *Journal of Applied Logic*, *15,* 94–129. https://doi.org/10.1016/j.jal.2016.02.001

Magnani, L. (2017). *The Abductive Structure of Scientific Creativity. An Essay on the Ecology of Cognition*. Springer, Cham, Switzerland.

Magnani, L. (2018a). Computationalism in a dynamic and distributed eco-cognitive perspective. In Dodig-Crnkovic, G. and Burgin, M., editors, *Philosophy and Methodology of Information,* Vol. 1, pages 365–288, Singapore. World Scientific.

Magnani, L. (2018b). Eco-Cognitive Computationalism: From Mimetic Minds to Morphology-Based Enhancement of Mimetic Bodies. *Entropy*, *20*(6), 430. https://doi.org/10.3390/e20060430

Magnani, L. (2018c). Playing with anticipations as abductions. strategic reasoning in an eco-cognitive perspective. *Journal of Applied Logic – IfColog Journal of Logics and their Applications*, *5*(5), 1061–1092. Special Issue on "Logical Foundations of Strategic Reasoning" (guest editors W. Park and J. Woods).

Magnani, L. (2019). AlphaGo, locked strategies, and eco-cognitive openness. *Philosophies*, 4(1):8.

Magnani, L. (2022a). *Discoverability. The Urgent Need of an Ecology of Human Creativity*. Springer, Cham, Switzerland.

Magnani, L. (2022b). *Eco-Cognitive Computationalism. Cognitive Domestication of Ignorant Entities*. Springer, Cham, Switzerland.

Magnani, L., editor (2023). *Handbook of Abductive Cognition*. Springer, Cham, Switzerland. Two volumes.

Magnani, L. & Bertolotti, T., editors (2017). *Handbook of Model-Based Science*. Springer, Switzerland.

Minati, G. & Vitiello, G. (2006). Mistake making machines. In Minari, G., Pessa, E., and Abram, M., editors, *Systemics of Emergence: Research and Development*, pages 67–78. Springer, Heidelberg.

Nakajima, K., Hauser, H., Kang, R., Guglielmino, E., Caldwell, D., & Pfeifer, R. (2013). A soft body as a reservoir: Case studies in a dynamic model of octopus-inspired soft robotic arm. *Frontiers in Computational Neuroscience*, *7*, 91. https://doi.org/10.3389/fncom.2013.00091

Nakajima, K., Hauser, H., Li, T., & Pfeifer, R. (2015). Information processing via physical soft body. *Scientific Reports*, *5*(1). https://doi.org/10.1038/srep10487

Numerico, T. (2021). *Big data e algoritmi*. Carocci, Rome.

Peirce, C.S. (1866–1913). *Collected Papers of Charles Sanders Peirce*. Harvard University Press, Cambridge, MA. vols. 1–6, Hartshorne, C. and Weiss, P., eds.; vols. 7–8, Burks, A. W., ed. (1931-1958).

Prigogine, I. & Stengers, I. (1984). *Order out of Chaos. Man's New Dialogue with Nature*. Bantam, London.

Raab, M. & Gigerenzer, G. (2005). Intelligence as smart heuristics. In Sternberg, R. J. and Prets, J. E., editors, *Cognition and Intelligence. Identifying the Mechanisms of the Mind*, pages 188–207. Cambridge University Press, Cambridge, MA.

Ramoni, M., Stefanelli, M., Magnani, L., & Barosi, G. (1992). An epistemological framework for medical knowledge-based systems. *IEEE Trans. Syst. Man Cybern., 22*, 1361-1375. https://doi.org/10.1109/21.199462

Rozenberg, G., Bäck, T., & Kok, J., editors (2012). *Handbook of Natural Computing*. Springer, Cham, Switzerland.

Shrager, J. & Langley, P., editors (1990). *Computational Models of Scientific Discovery and Theory Formation*. Morgan Kaufmann, San Mateo, CA.

Thagard, P. (1988). *Computational Philosophy of Science*. The MIT Press, Cambridge, MA.

Turing, A. M. (1969). Intelligent machinery [1948]. In *Machine Intelligence*, volume 5, pages 3–23. Edinburgh University Press, Edinburgh. Edited by B. Meltzer and D. Michie.

Vallverdú, J. (2020). Approximate and situated causality in deep learning. *Philosophies*, 5(1), 2. https://doi.org/10.3390/philosophies5010002

Vitiello, G. (2001). *My Double Unveiled. The Dissipative Quantum Model of Brain*. John Benjamins, Oxford.

Vitiello, G. (2012). The dissipative brain. In Globus, G., Pribram, K.H., and Vitiello, G., editors, *Brain and Being. At the boundary between science, philosophy, language and arts*, pages 317–338. John Benjamins, Amsterdam.

Vitiello, G. (2019a). Brain, mind, and the ontological prejudice. *Activitas Nervosa Superior*, *61*, 112-115. https://doi.org/10.1007/s41470-019-00037-0

Vitiello, G. (2019b). Mind and matter. two entangled parallel time-lines, one reconstructing the past in remembering, the other extrapolating into the future in predicting. In de Barros, J.A. and Montemayor, C., editors, *Quanta and Mind. Essays on the Connection between Quantum Mechanics and Consciousness*, pages 103–113. Springer, Cham, Switzerland.

Walton, D. (2004). *Abductive Reasoning*. The University of Alabama Press, Tuscaloosa, AL.