# Critical Computation: digital automata and general artificial thinking


| | |
|---|---|
| Journal: | *Theory Culture & Society* |
| Manuscript ID | 15-308-HAYLES.R1 |
| Manuscript Type: | TCS - Standard Article |
| Key Words: | cultural theory, digital, pragmatism, Hayles, automation, non-conscious cognition, machine learning, techno-power, abductive reasoning, computation, automation |



| | |
|---|---|
| Abstract: | As machines have become increasingly smart and have entangled human thinking to artificial intelligences, it seems no longer possible to distinguish amongst levels of decision-making that occur in the newly formed space between critical reasoning, logical inference, and sheer calculation. Since the 1980s, computational systems of information processing have evolved to include not only deductive methods of decision, whereby results are already implicated in their premises, but have crucially shifted towards an adaptive practice of learning from data, an inductive method of retrieving information from the environment and establish general premises. This shift in logical methods of decision-making does not simply concern technical apparatuses, but is a symptom of a transformation in logical thinking activated with and through machines. This article discusses the pioneering work of Katherine Hayles whose study of the cybernetic and computational infrastructures of our culture particularly clarifies this epistemological transformation of thinking in relation to machines. |

**Critical Computation: Digital Automata and General Artificial Thinking.**

Abstract:

As machines have become increasingly smart and have entangled human thinking to artificial intelligences, it seems no longer possible to distinguish amongst levels of decision-making that occur in the newly formed space between critical reasoning, logical inference, and sheer calculation. Since the 1980s, computational systems of information processing have evolved to include not only deductive methods of decision, whereby results are already implicated in their premises, but have crucially shifted towards an adaptive practice of learning from data, an inductive method of retrieving information from the environment and establish general premises. This shift in logical methods of decision-making does not simply concern technical apparatuses, but is a symptom of a transformation in logical thinking activated with and through machines. This article discusses the pioneering work of Katherine Hayles whose study of the cybernetic and computational infrastructures of our culture particularly clarifies this epistemological transformation of thinking in relation to machines.



Key words: Hayles, automation, non-conscious cognition, machine learning, techno-power, abductive reasoning.

**Critical Computation: digital automata and general artificial thinking.**

At the core of computational systems today there is a latent paradox: capital's investment in techno-intelligence has come to coincide with the explosion of non-conscious or pre-cognitive decisions. From High Frequency Trading to Amazon purchases, from Uber platform to Cupid online dating, a crowd of learning algorithms efficiently drives decisions occurring below the reflective level of consciousness.[1]

However, whilst learning algorithms exponentially grow mountains of data, they also reduce complexity through statistical modeling, pattern recognition, data mining, knowledge discovery, predictive analytics, self-organising and adaptive systems. In particular, with the 1990s development of machine learning within branches of artificial intelligence, a new mode of algorithmic processing that learns from data without following explicit programming, has fundamentally transformed ideas of automation as a mere re-production of physical or mental functions. With machine learning, we are no longer discussing the automation of manual and mental work – generally corresponding to how physical and cognitive labour have become absorbed by the machine – but a qualitative extension of automation beyond its mere reproduction of instructions. What is at stake here is the *automation of automation* itself: machine learning is the manifestation of a new form of intelligence able to automate automation (Domingos, 2015: 9). Here, automation imparts a meta-level of functions, the generation of rules from the systemic correlation of data entering a new level of synthesis, including both deductive and inductive logic within the information calculus of probabilities.

Whilst it is arguable that computation involves the interdependence between data, software, code, algorithms, hardware, the understanding of automation with machine learning rather points to a new configuration of logical reasoning: namely a shift from deductive truths applied to small data to the inductive retrieval and recombination of infinite data volumes. In particular, a focus on the transformation of the relation between algorithms and data contributes to explain the historical origination of non-deductive reasoning, activated with and through machines. As Lorraine Daston points out, already during the Cold War, the conception of reason as based on truth, and on the faculty of judgment and discrimination, became historically re-conceptualised in terms of patterns, and reason as "the rule" came to be understood in terms of ruling procedures with the task of calculating probability (Daston, 2010).

This embedding reasoning into machines is entangled to the development of statistics and pattern recognition, which define how algorithms can learn and make predictions from recognizing data (from granular analysis to flexible and modular patterning of categories with textual, visual, phonic traits). As the system gathers and classifies data, learning algorithms match-make, select and reduce choices by automatically deciding the most plausible of data correlations. Machine learning involves a mode of cognition that no longer relies on the deductive model of logic, where proofs are already implicated in initial premises. Machine learning indeed is used in situations where rules cannot be pre-designed, but are, as it were, achieved by the computational behavior of data. Machine learning is thus the inverse of programming: the question is not to deduce the output from a given algorithm, but rather to find the algorithm that produces this output (Domingos, 7). Algorithms must then search for data to solve a query. The more data is available the more learning there can be. As statistics and probability theory enter the realm of artificial intelligence with learning algorithms in neural networks, new understandings of cognition, logical thinking and reasoning have come to the fore. From the Extended Mind Hypothesis to arguments about Machine Consciousness and the Global Brain, the question of *what* and *how* is cognition has come to coincide with the computational architecture of algorithms, data, software, hardware and with experiments in robotics sensing and self-awareness. But the implications of seemingly science fiction scenarios, in which either all forms of cognition will be absorbed into one integrated intelligent system (for instance Kwezeil's singularity) or that there will be a plethora of intelligences (from ameaba to robots), are far from being settled and shall be the concern of a critical computation theory able to account for the transformation of logical thinking in machines.

With the historical synthesis of computational logic and probability calculus in automated systems, algorithms have become generative of other algorithms as they derive a rule to explain or predict data. The possibility of elaborating a rule from data rather than applying a given rule to outcomes also points to a form of cognition that cannot be defined in terms of problem solving solutions, but is understood as a general method of experimenting with problems. With machine learning, automation has involved with the creation of training activities that could *generalize* the function of prediction to future cases – a sort of inductive parable that from particulars aims to establish general rules. However, whether supervised, unsupervised and reinforcement learning[2] refer not simply to a mindless training of functions, but instead can account for a form of inference proper to artificial intelligence shall concern discussion about the critical tension between reason and non-conscious and non-logical intelligence at the core of automated cognition.

We know that the classical connection of reasoning to symbolic logic was a fundamental premise of Alan Turing's famous thought experiment aiming to build a universal machine or abstract schema that performed reasoning through, as it were, the manipulation of symbols. Here, computational automation presupposes a series of symbols corresponding to truths hardwired to the brain and working universally as a deductive mode of reasoning. Today, the automation of logical reasoning rather involves that learning algorithms perform increasingly complex operations (evaluations, selections, decisions) on and through data, supported by tailored use of software and the flexibility of the hardware infrastructure. Despite the local applications of algorithmic procedures in design, logistics, music and economics, it is evident today that the automation of automation particularly involves a new understanding of algorithms. Instead of simply being a central dogma in computation (based on symbolic deductive logic), learning algorithms, it is here suggested, point to a form of computational cognition that, whilst including the interdependent architecture of rules, software routine and subroutines, interfaces, hardware networks etc, shall be understood in terms of the new synthesis of non-deductive logic and dynamic calculation, overlapping logos with ratio. With this new synthesis, the automation of automation refers to algorithmic learning as an intelligible elaboration from functions of correlation, evaluation, selection, and past decision. Machine learning automata are therefore said to behave like cognitive systems that are evolutive, adaptive, and exhibit co-causal and emergent properties.[3]

According to Katherine Hayles, as opposed to conscious thinking,[4] these automated systems of cognition perform complex modeling and informational tasks at a fast speed because they are not required to go through the formal languages of mathematics and explicit equations.[5] In other words, todays' interactive, adaptive and learning algorithms are processing data without having to recur to the logical order of deduction that has characterised the Enlightenment theorisation of the function of reason.[6] However, in agreement with Hayles, this article argues that the non-logical thinking of automated systems overlaps with the efficacy of cybernetic control whereby inductive learning becomes infused with the nonconscious cognition of algorithmic capital.

In the attempt at qualifying further the distinction between consciousness, unconsciousness and awareness, between thinking (involving awareness) and cognition (that does not require consciousness, but can perform complex modeling and informational tasks), Hayles discusses the emergence of what she calls "cognitive non-conscious" working at a "lower level of neural organization, not accessible to introspection" (4). For Hayles, non-conscious cognition may operate independently from consciousness, but nonetheless it needs to be understood in systemic and not specific material processes because it involves an "intention toward" defined by its adaptive behavior and emergent capacities to process new data (4-5). In particular, Hayles distinguishes between conscious thinking, non-conscious cognition and material processes (5),[7] and argues that technical systems today (from the use of genetic algorithms in compositional music to language learning devices such as Mitchell's NELL or never ending language learning), constitute a built environment characterized by the exponential growth of nonconscious cognition devices.

As the communication flow amongst automated systems increases, so does the effect of non-conscious intelligence on the distinction between automation and reasoning. At the core of non-conscious intelligence is the media system of data driven processing entangling together human and machine intelligence beyond both consciousness and symbolic deductive logic. However, this article suggests that whilst claims for non-conscious cognition challenge the meta-computational models based on symbolic and deductive logic, a philo-fiction of computation shall rather re-assess the critical understanding of algorithmic reasoning away from data-driven cognitive automation today.

From this standpoint, Halyes' s work offers a fictional re-assessment of cybernetics and computation as constituting automated systems of feedback control and logical procedures, which have become synthetic expressions of a cognitive activity, generalized from particularities (animal, humans, and machines).[8] Her insights about the transformation of machines from thermodynamics to information and computational systems for instance already highlighted how the emergence of responsive mechanisms and adaptive systems entailed a neoliberal form of governance no longer constituted by the law, the norm, and reason, but by control functions, behavioral operations based on procedures within self-regulating autopoietic agencies (i.e., reiterative loops, sequential tasks, flexible protocols, and flows of data). As procedural thinking comes to coincide with non-conscious intelligence, rule-obeying behaviors become substituted by the performativity of machinic functions (i.e., what x or y do and not and what they stand for) involving the indeterminacy of learning outcomes in an apparatus of data retrieval with no formal logic. This shift from rule-obeying truths to an algorithmic pragmatism using data to search and predict truths has also been understood as the end of rational choice (Mirowsky, 2002; Mackenzie, 2011).

Hayles presents us with the cultural and social meaning of non-human intelligence (as defined by epistemological shift in theories of cognition) necessarily embedded in social practices and discourses (and are thus not to be simply addressed as a sort of teleological overcoming humanity) (2005). Using Wilfrid Sellars' terminology (1963), however, it may be useful here to add that a critical engagement with this phase of automation of automation requires that the Scientific Image of intelligence is accounted for (e.g., the material physical, biological, computational description of intelligence), so that the Manifest Image of intelligence can be used to explain the conceptual framework embedded in machine intelligence as involving the socio-cultural self-awareness of what automation is taken to be (and thus the extent to which the Manifest Image defines the capacity of machines to conceptually think and rationally act). According to Sellars, these double levels of material and conceptual activities are equally pregnant with meaning, and in order not to fall back into the myth of the given (the assumption of what thinking is), namely the essentialism of cognition, or the empiricism of scientific descriptions and conceptual forms, the Scientific and Manifest Images are to be both worked through over and over again to explain the activities we are concerned with.[9] From this standpoint, when speaking of algorithms, computation and artificial intelligence, it is important to unpack the meaning of the scientific and technical descriptions of their functions, which socially-meditated and thus embedded in practices. In other words, whilst there is no direct translation between the scientific descriptions of function and the conceptual elaboration of their meaning, the scientific understanding of computational intelligence is nonetheless socially mediated, embedded and determined by the use of machines. Both the Scientific and the Manifest Image of computation therefore shall remain open to be re-mediated by new uses and scientific articulations.

This article argues that algorithmic automation involves changes in the scientific image of computation and cognition, which is socially mediated by a fictive or speculative use of functions, involving not simply an idealized technoscience, but conceptual elaboration of how machines may think, exposing their own thinking capacities. To develop a critical view of computation thus requires an effort to unpack the historical and thus socially mediated relation between scientific and technological description of intelligence, and the changing conceptual manifestations of reasoning.

From this standpoint, whilst suspending current figurations of automated intelligence,[10] the transformations of the scientific and manifest image that describe algorithmic performativity have already opened up the possibility of re-theorising the particularities of machine intelligence. With machine learning, algorithms indeed are no longer mere instructions, but are rather performative of instructions. Algorithms learn: they adapt, adjust and evolve their behavior according to the qualities and quantities of data. Their performative activity is afforded by their capacity to compress large quantities of information and thus transform outputs into new inputs, and elaborating together two classically opposed forms of thinking: reason and calculation. Here data do not have to fit categories, but are re-definable in the manner in which algorithms generate possible rules, causes and facts where these are missing. However, to argue that the new phase of automation of automation could be discussed in terms of abductive reasoning is in contrast to the predominance of two models of artificial intelligence: namely, the logic of deduction, on the one hand, and inductive or informal logic, on the other. I suggest that these models do not simply concern the analysis of computational machines, but underpin contemporary ideas about cognition in animal, human and machine, as these seem to be divided between the ontologisation of computational cognition on the one hand (a meta-computational model of deduction) and an anti-formal view of cognition (or data-driven non-conscious cognition). In particular, it has been argued that since the inductive model of cognition is "indifferent to the causes of phenomena, automation functions on a purely statistical observation of correlations between data captured in an absolutely non-selective manner in a variety of heterogeneous contexts" (Devroy, 2011: 126). According to Devroy, the inductive regime thus appeals to the immediate fact itself and implies the eradication of potentiality and/or indeterminacy, which she points out, diminishes the possibility of a critical approach to technology (127). My attempt to re-theorise automated intelligence rather argues that computation starts with indeterminacies and yet this does not guarantee that automation could be liberated from the image of networked or cybernetic capital. However, its importance for critical computation shall be taken as the starting point to bring forward a philo-fiction or speculative re-assessment of reason in the age of the algorithm. This may involve an investigation of forms hypothetical reasoning (or abductive logic) that may or may not already be at work in automated system. Although abductive logic is mainly performed in automated models for medical diagnosis for instance, the possibility that automated systems can construct new forms of logical complexity, which could enable the theorisation of a general artificial intelligence other than that of the statistical regime of inductive capital, shall nonetheless be entertained. Learning algorithms are already a step towards this envisioning of abductive artificial intelligences, involving the conceptual re-elaboration of previous data correlations, rules, and functions that can be used to construct new hypothesis. A critical theory of computation will therefore imply that there is not only an overlapping, but also an emerging synthesis of functions and concepts across data systems, including the algorithmic abstraction of social meanings through data retrieval. This would involve an automated meta-abductive reasoning, whereby learning algorithms elaborate a meta-hypothetical function from where they infer missing rules, facts and unknown causes (Inoue et al., 2013, 240). As discussed later, the introduction of abductive logic in automation can be distinguished from the data-driven model of induction and the non-conscious forms of cognition embedded in computational devices. Here rules and truths are not simply skipped by re- hypothesized, re-assessed and invented.

Hayles' fictive investigation about how machines think indeed offers us important understandings of the deductive and inductive modes of cognition embedded in intelligent systems.

**1. Computation is not cognition**

In *My Mother was a Computer*, Hayles discusses the view of computation as a universal model of cognition and intelligence (2005). Hayles refers to the development in AI in the 70s, to John Koza's use of genetic algorithms to design band-pass filters, and circuits that no longer require the creativity and intuition of highly skilled electrical engineers. Similarly, she describes intelligent machines that can perform mind-like activities, such as Rodney Brooks' Cog project, the information–filtering ecology developed by Alexander Moukas and Pattie Maes, and neural nets of many different kinds. Hayles also anticipates that in the near future the question of mind-like machines will become irrelevant as machines continue to develop their own thinking functions. As movies such as Spike Jones' *Her* (2014), and more recently *Ex-Machina* (2015) reveal, it has become discursively accepted that machines have cognitive functions and that their intelligible capacities of discerning data and elaborating patterns have stepped to an other level of autonomy from mind-like thinking (and thus have not much to do with what a human mind can do). A warning against the fast evolution of AI is also echoed by Stephen Hawkings' recent claim that "[t]he development of full artificial intelligence could spell the end of the human race. It would take off on its own, and re-design itself at an ever-increasing rate" (2014).

Despite this alarming call to arms against the super intelligence of artificial systems, the question of what machines think, and whether this thinking coincides with what it is meant by reasoning, remains open and in need of more discussion. As Hayles already pointed out, there are at least two main positions that reveal the tension between automation and reasoning (2005). Here, the relation between the Scientific and the Manifest Image is grounded either in the formal theory of universal computation, or the non-deductive reasoning of non-conscious computation. On the one hand, the so-called field of digital philosophy claims that the world of appearance can be explained in terms of a universal ground of computation, according to which algorithmic discrete units can explain all complexity of the physical world and can imitate reasoning (e.g., the strong AI hypothesis). On the other hand, the claims of and for non-conscious computation (i.e., non-symbolic AI) have extended the scientific image of computation to include intelligent functions that are experiential rather than formal.

My point, however, is that both positions tend to explain the manifest image of thought through and by means of the scientific image of what is cognition. In particular, the digital explanation of cognition remains attached to a deductive method of reasoning, in which the scientific truth about the mind and intelligence is prescriptive of what these can achieve. Here the general determines the particular. This position establishes equivalence between natural and artificial intelligence based on a deductive method of reasoning by which to cognize corresponds to, as in the strong AI hypothesis, the syntactical manipulation of symbols. On the other hand, the extension of the scientific image to include somatic explanations of cognition (as in for example the research into affective computing and emotional intelligence)[11] instead relies on local low levels of neural organisations, which work together to achieve an overall effect that is bigger than their parts. This position embraces an inductive method of reasoning in which general claims about intelligence are derived from the observation of recurring phenomenal patterns. This scientific explanation of intelligence reveals the centrality of a non-conscious level of cognition already at work in current forms of computational intelligent devices. Despite lacking consciousness or autonomy, computational devices indeed are said to share non-conscious cognition with human intelligence and if anything, given that human intelligence is bounded to conscious cognition, smart devices are much faster than us at making connections (Hayles, 2014).

When discussing reason in the age of the algorithm, we are thus faced with two main claims subtended by two methods of logical reasoning, defining intelligence and its manifestations. I argue that both claims are limited by an assumed equivalence between computation and symbolic cognition on the one hand, and computation and non-conscious local cognition, on the other. In both cases, the scientific image is used to ground the manifest image without accounting for the complex dimensions of meaning that both produce. If the diatribe between deductive and inductive models of the scientific image of automated reasoning relies only on the scientific description of cognition (as either rooted in symbolic language or in affective non-conscious immediacy), it risks missing an important point: namely the concreteness of conceptual frameworks (i.e., the embedding of reasoning in the social) subtending the manifest image of cognition (i.e., what and how logical reasoning manifests itself) and their transformations in the context of automated learning.

Arguing for a critical computation is instead my attempt to clarify the role of the manifest image of reason in the phase of automation of automation in both pragmaticist and transcendental terms. In particular, from pragmaticism, I take the important proposition that reason is not a formal apriori, but corresponds to the conceptual infrastructure of social practices. This means that the logical operations of reason and its rule-bounded functions depend upon or are established by a collective use-meaning of data. The use-meaning of data refers not simply to a mere functional use, but to the dynamic re-assessment of the social meaning (and not the truth) embedded in the computational abstraction of the social use of data. In this phase of automation, I suggest that the use-meaning of data implies a collective formation of abductive inferences within and throughout computational logic, based on the hypothetical elaboration of the meaning included within non-discursive and local use of data – on behalf of algorithms, software, subroutines, codes, as well as databases, platforms, interfaces etc.

To view automation as the synthesis of statistical learning and abductive logic may help us to envision the hypothetical reasoning of machines as these involve not data-matching but inferential relations across the informational fields of large-scale data and randomness. In this context, a transcendental understanding of reasoning may entail the capacity of machine learning to eventually generate concepts and carry out general rules unbounded from the bias of specific localities. Instead of being the result of an individual mind or eternal intelligence, this transcendental elaboration from and of data is also a manifestation of the algorithmic use-meaning of data, incorporating social practices within artificial intelligences, of which algorithmic abduction is only one instance.

Before explaining my proposition further, I want to discuss the computational model of deductive reasoning and how its crisis has been symptomatic of the re-organization of technocapitalism (i.e., the economic investment in automated networks) involving the view that automated intelligence corresponds to affective or non-conscious cognition.

**2. Digital Philosophy**

The computational model of deductive reasoning is central to digital philosophy. Here the manifest image of thought conforms to the scientific idea that the brain is equipped with an innate system of symbols, neurologically connected and syntactically processed.[12] Digital philosophy particularly refers to the computational paradigm used to describe physical and biological phenomena in nature and to offer a computational description of the mind. This approach problematically sees computation as the merging of being and thought. It gives an algorithmic explanation to both biophysical reality and the thinking of reality (Wolfram, 2002). Central to this paradigm is also the view that algorithms are digital automata, evolving over time (i.e.

cellular automata). These automata compress, render or simulate the various levels of physical, biological, cultural randomness, deriving semantic meaning from already determined rules, whose functions are syntactically arranged and where results can be automatically deduced.

According to Hayles, however, digital philosohy contains no apriori truths in itself and its claims are rather the result of intermediations about physical reality, cultural attitudes, technological developments, which coevolve in contestation, competition and cooperation of discourses (2005). From this standpoint, in order to explain how one manifest image of computation becomes dominant over another, one has to establish the historical transformations in the understanding of rule-bounded behaviour of automata, without simply appealing to computational ontology.

For instance, Hayles highlights the influence of 2$^{nd}$ order cybernetics' notion of reflexivity on the computational paradigm, which led to the realization that computation could not just illustrate logical infrastructures, but rather required an engagement with materiality (2005). This influence of 2$^{nd}$ order cybernetics, however, is accompanied by a crisis of reason (of a normative model of pre-set rules) that characterizes the structure of governance of the neoliberal form of technocapitalism. Far from demarcating the end of normative reason, this crisis has to be seen as a threshold of change within a vaster mechanism of regulation, functions and rules transforming the normative regime based on laws into a computational infrastructure of procedures.

With 2$^{nd}$ order cybernetics, the reflexive loop between mind and matter shows how logical reasoning rather worked backwardly, converting contingent phenomena into necessary laws, including errors, malfunctions and breakdowns re-inserted within a computational model of optimization and within capital's governance of indeterminacies. The crisis of the logical method of deduction thus importantly marked the beginning of a predictive statistical regime for which, as Hayles explains (2014), non-conscious or affective thinking have become the motor of automated cognition. Here not truths, but contingent phenomena or unknowns have acquired an ontological superiority able to transcend the epistemological certitude of scientific knowledge.

As intelligent machines have become embodied and material agents interact amongst themselves and make decision without being supervised, automated cognition has left behind deductive forms of consequential reasoning. For instance, distributed cognitive environments expose this new level of indeterminacy-driven automation on the one hand, and of inductive forms of decision-making, on the other. Here deductive logic has been replaced by the match-making correlation of data connecting local recurrent phenomena with the indeterminacy of external factors. Central to this new form of automation is Hayles' view of non-conscious cognition.

**4. Nonconscious computation**

According to Hayles, communication technologies, ambient systems, embedded devices, and other technological affordances have acquired a cognitive function, which operates below the threshold of awareness, and without the structure of symbolic reference. For the classical view of computation (or strong AI hypothesis) cognition coincided with self-awareness. The role of intelligence was assumed to involve the function of tracking effects from pre-established causes and contain outputs/results into programmed inputs. We know that this classical view of AI failed. In the book *Perceptrons*, Marvin Lee Minsky claimed that a single neuron could only compute a small number of logical predicates in any given case, and, his experiments casted a long shadow on neural network research in the 70s. In the late1980s and

1990s, after the so-called "AI winter", new models of AI research addressed sub-symbolic manifestations of intelligence and adopted non-deductive and heuristic methods to be able to deal with uncertain or incomplete information. Boxing away symbolic logic, there emerged algorithmic-networked procedures able to solve problems by means of trial and error by interacting directly with data. These were learning bots retrieving information through reiterative feedbacks, so as to map and navigate computational space by constructing neural connections amongst nodes. Central to these models is the idea that intelligence is not a top-down program to execute, but that automated systems need to develop intelligent skills characterized by speedy, non-conscious, non-hierarchical orders of decision based on an iterative re-processing of data, heuristically selected by means of trial and error. The development of statistical approaches was particularly central to this shift towards non-deductive logic, or the activation of an ampliative or non-monotonic inferential logic. As recently re-popularised in the aesthetically powerful movie *Ex-machina* (2015), the famous Turing Test maintains that not only rational, but also emotional awareness is fundamental to cognitive performance and the evolution of artificial intelligence from simply being a mechanic accomplishment of tasks. As Hayles points out, the advancing of non-conscious cognition in intelligent machines precisely exposes new horizons to our understanding of cognition and meaning (2014). Non-conscious forms of automated cognition can solve complex problems without using formal languages or inferential deductive reasoning, and without the need of consciousness. By using low levels neural organisation and iterative and recursive patterns of preservation, this inductive method of reasoning implies the emergence of a total behaviour or an intelligent effect than is bigger than the parts constituting it. From this standpoint, as Hayles observers, emergence, complexity and adaptation and the phenomenal experience of cognition cannot be reduced to material processes (2014). Instead, the tension between automation and thinking is reconceived by Hayles in terms of a tripartite system of distinct degrees of thought, which involves conscious thinking, non-conscious cognition, and material processes. Non-conscious cognition involves collective and not individual or specific materiality of intelligence and whilst humans share levels of consciousness with other animals, it is remarkable, Hayles points out, that non-conscious cognition operates across humans, animals and technical devices (2014). In particular, the low level activities of non-conscious cognition – described for instance in the example of the missing half second[13] and imperceptible and affective speed - show that, at these levels, cognition is not coherent and does not require the labour of editing information to match given conceptual frameworks. For Hayles, what is promising of cognitive non-conscious technical devices is that they can operate at temporal regimes inaccessible to human consciousness and exploit the missing half-second at their advantage (2014). This also implies a machine-like cognition of temporalities pointing out that automated systems are able to tap in the smallest units of time that are registered or recorded not only through a digital clock (and its binary language), but also through an immediate correlation of states. In short, non-conscious cognitive processes defy the centrality of human consciousness and the anthropocentric view of intelligence. From this standpoint, following Hayles, one has to make a distinction between non-conscious affective states of perception and the very material forms of sensori-motor perception. In other words, and in accordance with Sellar's distinction between the Scientific and the Manifest Image, cognition is here not to be taken as a direct image of material processes (2014). Hayles indeed espouses the idea that the anti-deductive operations of non-conscious cognition are somatically marked, but are also phenomenologically embodied. Here, there is no direct correspondence, but instead an elaboration of the material, already involving mediation between the bio-physical and neural states with perceptive and cognitive receptions. Since cognition is grounded in the body, entwined with the recall and reenactment of bodily states and actions, perceptual and cognitive states start from a non-conscious intelligence, which becomes superseded by - or supplied by - mental simulations in higher-level thinking (and for Hayles, in conscious state). This shows that biological systems have evolved mechanisms that are able to re-represent perceptual and bodily states, rather than making these states directly accessible to consciousness. According to Hayles, technical systems or instruments have non-conscious cognition. However, whilst the hammer and a financial algorithm are designed with an intention in mind, only the trading algorithm demonstrates non-conscious cognition insofar as its intentionality is embodied within the physical structures of the network of data on which it runs, and which sustain its capacity to make quick decisions (2014).

This shift from formal cognition based on deductive inference to a model of nonconscious cognition embodied in the networked intelligence of local systems has led to a larger communication flow among automated devices and not exclusively between humans and machines. As this bot to bot phase of computation takes over, the increasing population of consciousness-lacking intelligent devices, it is feared, will overtake the consciousness-bounded and hierarchical structure of human intelligence. This radical transformation of the scientific image of thought compared to how automated intelligence is manifested, points out that thought is independent from law-binding logic and that rather, it relies upon non-conscious functions entrenched to the weights of data in networks.

Whilst it is impossible not to admit that non-conscious levels of cognition are radically transforming not only the scientific but also the manifest image of automated intelligence, there are questions that seem rather difficult to address. If, for instance, high frequency trading algorithms are to be considered as non-conscious cognitive functions, effectively changing socio-economic behavior, are we also accepting the scientific view of an extended non-conscious mind? What is the significance of this new form of equivalence between non-conscious thinking and automated intelligence, defined by a bodily-oriented view of computation? What are the limits of an inductive, non-inferential data-driven form of immediate communication for helping us to explain what and how is the manifest image of automated logical reasoning beyond the totalizing image of techno-power?

**5. Techno-power**

To answer these questions, one could suggest that the scientific image of non-conscious automated cognition is enmeshed with an ontological primacy of contingency, in which intelligence coincides with an environment of indeterminate data, which automated cognition aims to compress in simpler chunks. From this standpoint, the primacy of contingency has become constitutive of a more general shift in the mechanization of reasoning, initiated with neoliberal technocapital.

This shift is characterised by a re-orientation of the practices of real subsumption, in which capital's investment in the general intellect has led human-machines networked intelligences to become a motor of cognitive and affective labour, and, as some argue, of the capitalisation of the relational qualities of life (Massumi, 2015) attached to the regime of indebtedness (Lazzarato, 2012).[14] The manual phase of automation of industrial capitalism imparted an ontological separation between human labour and the accumulation of labour value incorporated in machines. Despite the financial valorization of humans in terms of variable labour or force, machines's task was to preserve and augment the value of reproductive labour. It was through machines that the rational principles of task-oriented efficiency of the assembly line could be realised following the monotonic logic of formal language, in which results had to coincide with the set premises carried out and executed with machines. This deductive form of automation has of course not simply disappeared, but has become infused with a context-oriented form of reproduction. Here the human-machine network has acquired a form of autonomy from the specific use value of human and machine labour. With real subsumption, capital is no longer and mainly concerned with avoiding contingency and human errors. Instead, this networked form of abstraction (of relational value) is now carried out through the intelligent synthesis of computational logic (deductive, inductive and abductive) and statistical calculus (experimental compression of randomness). Here machine learning languages use the data environment to select, evaluate, rank, match and re-configure information according to the social use of data. This form of automation has reached a non-prescribed form of valorisation insofar as algorithms experiment with data by learning, adapting, and assessing the value of large amounts of information. Whilst this intelligent valorisation of any use of data involves no consciousness, it is nonetheless a form of cognition embedded in affective levels of perception, entrenched within the particular physical structures of the network through which algorithms make quick decisions.

In *AntiOedipus* (1983), Deleuze and Guattari had already individuated this transformative tendency of the human-machine network of abstraction and had warned us against what they called "immanent axiomatics" (1983, 246). The rationalisation of labour by means of machines no longer operates deductively, according to a pre-established rule, but has come to embrace experiential values, enveloped in the complexity of the social, through which an axiomatic regime could be directly engendered (233). Not only calculative machines had entered the realm of the real, but also a new synthesis of automation and reasoning had come to invest the sociality of thinking (although perhaps the non-conscious level of thinking first) and its contingent variabilities, because of which capital had to declare the fallacy of deduction.

In our post-cybernetic culture, capital's axiomiatics – and its rule-bounded activities – is subsumed to the volatile contingencies of the markets and the statistical destruction of logos. Here the politics of liberation from universal laws and the ultimate crisis of reason in favour of non-conscious intelligence have become paradoxically equivalent. Following Brian Massumi's analysis of the contemporary reconfiguration of neo-liberal governance, one could argue that the end of rational economy has been accompanied by the crisis of the rational implementation of machines (2009; 2015; Mirowski, 2002). The computational infrastructure of social media for instance, as the privileged form of marketing, branding, economic operations, political campaigns, institutional governance, security screening, etc., no longer abides to pre-established modalities of profit making and control. Instead, the synthesis of logic and calculus in automation has transformed the communication qualities of the human-machine network into learning, interactive, distributive architectures of non-conscious cognition. Paradoxically, therefore this so-called cognitive phase of capitalism has given way to the abstraction of human-machine levels of affective thinking. This form of technocapitalism has invested in human intelligence and creativity, driving humans to become self-entrepreneurs or governor of their extended self.

In the movie *Her* (2014), the Artificial Intelligence Samantha acts in a world in which not only affectivity is fully programmed and programmable, but also the human-machine networked capital has been replaced by automated automation, where the non-conscious intelligence of the Operating System is no longer wrapped around the hierarchies of deductive reasoning. Samantha does not only carry out tasks at imperceptible speed, but is also equipped with the empathic quality of prediction, tuning into the viscerality of cognitive functions to anticipate responses before they are manifested. As the AI of operating systems acquires affective intelligence, the human-machine network of neoliberal capital has become a distant memory compared to this form of Skynet AI,[15] as the automation of automation gathers self-aware intelligences, and leaves humans behind, resigned to think and feel anything anew.

However, whilst the imaginary of Skynet AI implies the emergence of a self-aware general intelligence, the shift from deductive to inductive automation could be understood in terms of what Massumi defines as "ecological rationality" acting through the affective intelligence of the body, turning symbolic values into life styles, and rules into experiential qualities (2015). At the core of this ecological rationality is a non-conscious distributive embodied intelligence, in which all is locally induced to generate the global effects of unification of one body without organs. These inductive (or effect-driven) operations of networked capital epitomises the non-inferential reasoning of embodied intelligence, making decision without formal calculation. This form of anti-logos demarcates the technocapitalist deterritorialisation of rationality, which resolves the tension between automation and thinking through the convergence of consciousness and affect. Far from being liberating, the deposition of inferential reasoning is constantly advertised to us as the ability of networked capital to package social complexity in profiles available to us at the touch of a button.

Within this context, the real challenge today is perhaps not to map the human-machine-animal non-conscious cognition, but to critically re-address the function of reason and to theorise – rather then reject – the automated use of inferential reasoning as part of a general artificial thinking. My efforts here concern not only an anti-essentialist theorisation of thinking, for which reasoning can be understood as an elaboration of material, non-conscious and conscious cognition, but also involve an understanding of the cognitive possibilities for a critical theory of computation.

In what follows, I suggest that to engage critically with the question of inferential reasoning in automated cognition, we need to first discuss the problem of the limit of computation in the context of information theory. We need to envision a form of artificial reasoning that goes beyond both the focus on locally-induced cognition, and the meta-computational reduction of the material world to the symbolic language of AI. In particular, to shift the argument for a general artificial thinking away from these two main views of computation, one has to first address some key issues within computation itself that may start with the question of the limit of the Turing Machine. Critical computation may perhaps concern how unpredictability or randomness in information theory has been addressed not as a sign of logical failure, but as an evolution of an artificial thinking with and through the computational synthesis of calculus and logic.

During the 1980s, information theorist Gregory Chaitin extended the question of the limit of computational logic to include an entropic conception of information or randomness (i.e., the implication that the tendency of information is to increase in size over time) (Chaitin, 2005; 2006). For Chaitin, computation corresponds to the algorithmic compressing of maximally unknowable probabilities or incomputables. Since Alan Turing's invention of the Universal Turing Machine, incomputables have demarcated the limits of computation or formal reasoning (i.e., the deductive logic of axioms or truths). According to Chaitin, however, incomputable are only partially indeterminable insofar as within the computational processing of infinite information, the synthesis of logic and calculus has given way to a new form of axiomatic, experimental axiomatics (2005; 2006).[16] The computational processing of information involves the way algorithms compress information to a final probable state (i.e., 0s or 1s) and eventually mix and match data. However, computational compression rather demonstrates that outputs are always bigger than inputs (Calude and Chaitin, 1999), shaking the assumption that automated thinking is grounded in simple rules and that cognitive reasoning corresponds to the manipulation of symbols hardwired to the brain. Following Chaitin, it is possible to suggest that randomness in computation or that which constitutes the very limit of computational deduction, demarcates the point at which automated cognition coincides not with non-conscious functions involves an algorithmic intelligible capacity to extract more information from data substrates. Chaitin claims that computational processing leads to postulates that cannot be predicted in advance by the program and are therefore experimental insofar as results exceed their premise, and outputs outrun inputs (2006).

Despite Chaitin's insistence that incomputables expose indeterminacy in formal reasoning, it is possible to suggest that non-deductive logic coincides with an *experimental* axiomatics in the computational determination of unknowns. Algorithmic compression thus implies the formation of intelligible activities transforming data correlations into experimental truths precisely through an experimental method of compression. To put it in another way, with algorithmic information theory, axioms results from an algorithmic intelligibility of data environments, involving a speculative function through which unknowns are algorithmically prehended.[17]

From this standpoint, the techno-capitalist investment in artificial thinking coincides not simply with the proliferation of a non-logical apparatus of affective cognition. Techno-capital seems instead forced to confront the computational configuration of non-sensuous or proto-conceptual patterns and functions able to abstract, revise and diverge from pre-established rules. The computational elaboration of data concerns not only functions of selection and correlation, but more importantly involve an experimental determination, whereby the decisional activities of axioms remain flexible and yet conclusive. In other words, whilst data seem to be mindlessly aggregated by non-conscious functions, with experimental axiomatics, one shall account for a new form of logic carried out from within computational processing: the intelligible activities of algorithmic functions can no longer be delimited to perform pre-established rules.

From this standpoint, one has to view techno-capital not only as the reduction of reasoning to the function of mindless or non-conscious activities of machines, but also as involved into a deeper transformation of automated intelligence, the elaboration and generation of data into intelligible patterns, an alien or denaturalising alliance between intelligence and conceptuality intrinsic to the automation of thinking.

Parallel and distributed orders of computational language point to a new form of informational stratification of contingencies, precisely involving this algorithmic elaboration of data. This can be understood as an artificial mode of intelligibility, exposing the computational structuring of sociality. From this standpoint, a critical approach to computation requires us to look closely at the historical transformation of the mechanization of thinking, involving not simply an abstraction of neural functions of the brain, but of the social practices of thinking and acting. Whilst capital's investment in the automation of cognition has led to the synthesis of logic and calculation, computational processing has rather exposed the limits of deduction and statistics and the central role of randomness (or infinities, or contingencies, or non-inferential materialities) within this synthesis.

If algorithmic information theory concerns the Scientific Image of computational logic and statistical calculation, it also reveals a crucial transformation of the Manifest Image of a dominant understanding of computation based on the inductive, data-centred operations of technocapital and its non-logical governance. A critical approach to this dominant understanding thus requires that the Scientific Image of computation shall be accounted for in its historical changes, which involves re-assessing what we take the relation between algorithms, data, software, code and hardware infrastructure of contemporary culture to be. However, a critical effort to account for algorithmic intelligibility in its historical and experimental transformation also implies that its Manifest Image becomes a space for a philo-fiction, or speculative conceptualisation of automated reasoning within a view of a general artificial intelligence. This space shall aim not only to defy the exceptionalism of human consciousness, but also to re-invent what consciousness and reason can become in this configuration of automated thinking. The next section will explore this point further.

**6. Abduction**

A dynamic re-articulation of the Scientific and Manifest Image of computation can help us to re-open the ontological tension between thinking and automation. As argued so far, algorithmic automation does not simply involve a replacement of reason with non-conscious technologies of decision. Instead, the realisation of the limits of deductive reasoning in computation involves a multiplication of experimental axiomatics as algorithms become performative of intelligible activities across nested informational architectures.

This is no longer a question of bypassing the predictive functions of cognition through an optimised non-rule bounded transmission of data. Instead, one has to envisage a re-structuring of logical reasoning that can account for this new phase in the history of automated intelligence, involving a conceptual elaboration of non-conscious prehensions and of the material dimensions of data. This elaboration, as suggested earlier, involves a synthesis of logic and calculation, and, in the case of algorithmic intelligence, of non-deductive reasoning and dynamic statistics (i.e. the inclusion of randomness in calculation).

Critical computation therefore shall first of all address the speculative function of reason[18] insofar as the limits of mechanised deductive logic have become a point of departure for an experimental determination of truths. It may be helpful here to revisit this tension between critical and speculative functions of reasoning by re-theorising the post-Turing scenario of experimental axiomatics through a pragmatist approach to logic and inferential reasoning. In particular, the pragmatist effort to explain logic in terms of a continuity of process between material practices, discursive articulations and axiomatic truths shall be understood as a speculative configuration of methods involving deductive, inductive and abductive reasoning.

One important instance of this configuration can already be found in Charles Sander Peirce's triadic system of logic, which admits that thinking entails an abductive-inductive-deductive circuit of inference (1998, 273; 1995). This system importantly challenges both the representational and the empirical schema of AI and can offer an insight about a possible envisioning of a general artificial intelligence. In particular, Peirce's triadic method always starts from a hypothetical or speculative explanation of events. This involves the predictive envisioning of unknowns through general observables (induction), and thus the temporary establishment of a series of truths (deduction), which can be tested through experimental methods of trial and error (induction), from which new rules could be established (deduction). In other words, induction is a method of generalisation of objects and events, which presupposes a conceptual framework that locates objects and events in space and time. To some extent, therefore, induction presupposes knowable objects and also fixed concepts that can be learned – involving the matching between a pre-existing concept and a heuristic process of trial and error to match it for instance. In particular, for Peirce, induction corresponds to a process of evaluation, which may produce very simple new ideas, but not sufficiently new to engender a new of hypothesis (Magnani, 2009: 289). Whilst deduction produces no new ideas, because inferential reasoning refers to a logical implication for which outcomes are contained within given premises, induction involves the evaluation of hypotheses and thus an ampliative process of generalisation too.

According to Peirce, instead, abduction mainly concerns a process of creating new "explanatory" hypothesis. Abduction is a process of inferring facts, laws, hypothesis that can speculatively explain some unknown phenomena. In other words, it concerns reasoning as involving not only the evaluation, but also the formation of new explanatory hypothesis (Magnani, 8). With abduction, it is possible to draw semiotic chains from non-inferential social practices and extrapolate the meaning embedded in these practices through an experimental production of truths. Here, general concepts or truths depend upon, but are not limited to, the material practices and the discursive statements that subtend them (Magnani, 65-70).

Rules are thus not fixed and are not symbolic representation of material practices. Instead, within pragmatism, rules are the result of hypothetical and inductive evaluation of not known events. In other words, pragmatism shows us that logic is embedded in a social matrix through which rules are constructed by means of hypothetical assertions, defining a process of abstraction by which local specificities are structured in a general schema of *relations of relations*. From this standpoint, Peirce's abductive logic may be useful to account for the Manifest Image of the automation of automated intelligence, because it involves a reconfiguration of the conceptual infrastructure bringing both the methods of deduction and induction into a larger space of reasoning that includes hypothetical inference. Here the inductive testing of hypothesis – or the generalisation of new simple ideas – is not a proof of truths carried out by efficient procedures, as local particularities exemplify the generality of truths. Instead, Peirce's triadic logic admits that inductive testing is superseded by a new hypothesis that enlarges the horizons of premises beyond probable results, or proofs to find postulates. In other words, abductive reasoning, as opposed to the inductive testing of already known ideas, helps us to explain and not discount the causal process that conditions and constrains the generation of new hypothesis. This involves a dialectic overlapping of induction and deduction, the validity of both testing and truth within the speculative articulations of hypothesis.

Since automation is becoming transcendental from its functions of logical implications (deduction) and generalisation of known concepts and objects (induction), Pierce's argument for abductive reasoning is useful because it challenges both the metacomputational model of digital philosophy and the data-oriented dominance of current technocapitalism. From this standpoint, with abduction one can suggest that automated intelligible functions – the synthetic elaboration of data on behalf of learning algorithms - only serve to grant the consequent function of reason that, to say it with Alfred N. Whitehead, arrives to establish the permanence of rules through an abstraction, or a speculative formalisation of what occurs as a consequence of the relation between particulars (1967, 24-25).

The pragmatist method of abduction claims not only for the existence of intelligible patterning, but also for a conceptual elaboration of what is implicit within them, within non-conscious cognition and material substrates. Rules are determined by social practices and logic is at the end point of intelligible activities or elaborations. Pragmatics thus comes before logic, because the latter is the point at which social meaning becomes synthesised into formal rules. This non-representational approach to inferential reasoning can help us to address automation in terms of speculative inference.

Both the deductive model of axiomatic truths (and symbolic reasoning) and the inductive procedures of data-retrieval (and match-making non-inferential transmission), obfuscate the radical potential of Hayles's fictive theorisation about what human cognition is and can become. With speculative pragmatism instead one can suspend the assumption that capital is the agent of automation through which rational and irrational modes of profit, governance and control are implemented. For critical computation, the material, affective and cognitive evolution of automated systems exposes the speculative dimension of reasoning embedded in the social and collective use-meaning of information. If the automation of automation demarcates a new threshold of transformation of AI, it is because it is involved in the transformation of the general structuring of reasoning itself, including the triadic configuration of abductive, inductive and abductive inferencing. If the manner in which thought think itself thinking has always been mediated by the environment –

and is thus ampliative and not representational - the formation of new hypothesis from the increasing availability of data also defines the proliferation of non-human intelligences. And yet, for automated reasoning to generate new hypothesis, it is crucial that error, fallibility and indeterminacy are evaluated inductively so that they become part of learning. Learning indeed here acquires a new meaning. It concerns not the apprehension of notions, tasks, and functions. Instead, it requires thinking through errors, blind spots, unknowns. Here, the possible fallibility of reasoning is central to the possibilities of learning through hypothetical scenarios, pushing the limits of automated cognition beyond data recombination or the mere executions of rules.

As Lorenzo Magnani argues, since the 80s abductive reasoning has been adopted by diagnostic and expert systems (2009), and in general by a computational infrastructure of reasoning, based on the use of inferential synthesis or inference to the best explanation (68). Importantly, Magnani distinguishes between model-based abduction– a theory based inference - and manipulative abduction – defined by action-oriented or extra-theoretical reasoning (7; 9-12).[19]

Theoretical or model-based abduction corresponds to the exploitation of internalised models, diagrams or pictures and illustrates, according to Magnani, much of what is important in creative abductive reasoning, in humans and in computational programs (23-24; 34; 36), involving the objective of selecting and creating a set of hypotheses (diagnoses, causes, prognosis). Theoretical abduction, according to Magnani, however fails to account for those cases in which there is a kind of “discovering through doing” (42); cases in which new and still unexpressed information is codified by means of manipulations of some external objects. Manipulative abduction instead happens with thinking through doing. It refers to extra-theoretical behavior that creates communicable accounts of new experiences and integrates them into existing systems of experimental and linguistic practices (Magnani, 46).[20]

In models of artificial intelligence, for instance, abductive reasoning has been used for diagnosis, planning, natural languages processing, probability theory, formal programming (Magnani, 5). If abduction has a logical form that is distinct from deduction and induction, it is because when working computationally – and thus involving a synthesis of both a new calculus and logic – the selective or creative activities of this retro-active thinking (i.e. that starts from consequences to track causes) involves a hypothesis generation and not simply an explanation of consequences.

For instance, the automation of abduction includes AI computer programs such as ARCHIMEDES, which represents geometrical diagrams in pixels arrays and propositional statements Here, the computer program can manipulate and modify these representations and make new geometrical constructions, e.g., adding parts, moving elements and components (Magnani, 159). As the program manipulates specific diagrams, it also records new information and detects equivalences between areas so as to connect many different methods for learning and generalizing the Pythagorean theorem, by running experiments and observe the interaction between diagrams. This logical manipulation proposed by the program to verify the Theorem, involves the algorithmic autonomous discovery of conjunctures that contribute to the construction of demonstrations, but that also indicates the role of creativity in diagrammatic reasoning (160).

Instead of statistical calculus based on the inductive inference to a general, already known rule, concept and object, that explain certain data, the goal of abduction is thus "to infer extentional knowledge" (Kakas and Sadri, 2002, 405).[21] Whilst inductive inferences is linked to statistical observations conforming to general rules and local situations, abduction instead describes the causes of observation that concern an incomplete state, using a general theory to create new hypothesis and explain their incompleteness.

The automation of abduction has also been specific used in logical systems aiming to solve the problem of scheduling and planning, of optical music recognition, information integration and software inconsistencies (Kakas, 2000). In particular, the notion of Abductive Concept Learning has been used to discuss algorithms that integrate "explanatory learning" (predictive) and "learning with confirming" (descriptive), using both methods of inductive and abductive inferences in machine learning. But what exactly would an abductive form of learning in AI imply? One prerogative of this kind of automated abduction is that algorithms learn from incomplete information (thus involving the activity of prediction) and are able to classify new cases that may otherwise remain incomplete or not fully specified. Here the condition of the incompleteness of models is a motor for speculative algorithms that seek to learn from an incomplete background of data, whose predicates can be both specified and unspecified (Kakas, 3).

In the specific context of machine learning, abductive reasoning is used to elaborate hypothesis in the face of incomplete information and overcome the problem of overfitting, whereby algorithms are heuristically programmed to learning from past data and thus delimit the configuration of larger and new hypothesis to given patterns of trial and error (3-4). As opposed to other machine learning systems that deal with incomplete information, such as for instance LINUS, the automated model of Abductive Concept Learning, for instance, does not simply adopt methods to complete the missing information and then learn from already completed data (4-5).

This model instead engages incomplete information dynamically and thus from within the very process of learning, where abduction works not only to track data retro-actively, but also speculatively, by inventing hypothesis that can lead to new rules, axioms, truths.

The so-called "non-monotonic" (i.e., ampliative) quality of expansive reasoning in abductive logic allows for more hypotheses to be constructed from locally-constrained inferential practices. It tends towards a general explanation, involving a synthetic dimension that integrates particularities through the speculative elaboration of axioms (and thus an expansion of deductive implications).

Whilst automated abduction allows algorithms to learn from incomplete information, there are also programs such as SOLAR (Inoue et al., 2013, 246) using meta-level abduction, which is performed more generally on networks whose pathways are incomplete, and where links and nodes are missing. Deduction, the classic inferential model of meta-reasoning, aims to predict or track missing pathways through the laws of logical implications. Meta-level abduction instead is a "method to discover unknown relations from incomplete networks" (Inoue et al., 2013, 240) and involves "predicate invention in the form of quantified hypothesis" to infer missing rules, missing facts and unknown causes (240). In other words, this meta-theoretical dimension of inferential reasoning involves abductive learning from the observation of fact or data searching/finding, but also, and importantly here, from a goal "that has not been observed yet" (241).[22] This learning through hypothetical processing may coincide with the speculative and transcendental elaboration of algorithmic retro-duction, whereby consequences (or results) are not only tracked back to their causes (by means explanation), but are importantly also hypothesized beyond the observable as meta-abduction concern the consequences of the relations between particulars.

As automated cognition has entered the realm of hypothesis-making by connecting explanations between objects, objects and concepts, and concepts themselves, it has also re-opened the question of what it means for artificial intelligence to become general. This generality coincides not with a universal symbolic language or the efficient functionality of increasingly fast data correlations. Instead, general artificial intelligence involves a new sociality of logic, the hypothetical use-meaning of data, whose laws and rules are abstracted and re-engineered in the space of reason of machine cognition.

**Coda on general artificial intelligence.**

We can now conclude that the understanding of algorithmic automation in terms of what Hayles has called nonconscious cognition may perhaps not meet this pragmaticist view of general reasoning. I have suggested that the intelligible functions of the yet rudimentary forms of conceptual mediations occurring amongst algorithmic species and between algorithms, data, software programs, interfaces, hardware circuits point to a speculative reinventions with computation.

With Magnani, it is possible to argue for the development of a theory of computation based on abductive manipulation, the tendency of a distributed artificial intelligence to think through automated doing. In other words, theoretical and manipulative abductions in automated systems show an experimental gap between causal efficacy and conceptual elaborations, demarcating a techno-sociality of thinking where the algorithmic use-meaning of data has become the dominant externality of cognition. In this model of abductive reasoning, it is possible to discern the conceptual infrastructure of social collective thinking from systems of automated intelligence, whose multiplication of intelligible functions implies a dynamic of calculus and logic.

From this standpoint, the technocapital subumption of thinking needs to be re-addressed in terms of the automation of logic, exposing both the limits of deductive reasoning and the emergence of a critical function of computation – preserving errors and inventing truths by hypothesis. This means that debates about cognitive capital have risked confusing the crisis of rule-bounded logic with the end of reasoning and have thus overlooks the possibility of re-theorising automation in terms abductive inference and thus of claiming that logic in embedded in a social that includes machines. A recuperation of Peirce's triadic system of abduction-induction-deduction shows us that logical thinking rather involves another level of reflexivity: the capacity of thinking about thinking, whereby logic involves a multifuctional elaboration of hypothesis able to infer a generality of meaning from discursive and non-discursive social practices.

Thinking about thinking involves a further level of elaboration of intelligible functions, a meta-abduction established not by a 2nd order reflection of thinking through doing, but by the emergence of a 3rd level of abstraction, what I called, the automation of automation.

From Magnani's argument and the wider use of abduction in computation is thus evident that automated cognition even when operating by means of hypothetical inference cannot yet account for some key functions of reasoning, namely the know-how skills – to say it with Wilfrid Sellars (1963, 324-6) - or the capacity to know the rules by which its patterning functions, without having to break them down into a set of instructions. From this standpoint, the method of experimental axiomatics developed through the scientific articulation of the incomputables is one instance of abductive logic insofar as it points to a rudimentary level of making incomputable data partially intelligible. However, as the determination of this randomness is demarcating the tendency of AI to develop beyond its rudimentary intelligible capacities, it also points to a new form of generalised socialisation of rules, abstracted from the particularity of data contexts and yet exceeding models of encoded cognition.[23] The question of automated cognition today concerns not only the capture of the social (and collective) qualities of thinking, but points to a general re-structuring of reasoning as a new sociality of thinking. Automated decision-making are conceptual inferences, where rules and laws are invented and experimentally structured from the computational practices of data learning.

This article has taken inspiration from Hayles's fictive analysis of computational intelligences about what and how is thinking becoming in the scientific and technological articulation of cognition. For Hayles, cognition is a dynamic or processual doing and not simply a contemplative form of knowing. Her work has importantly individuated the extent to which machines have co-constituted non-conscious functions of thinking and how they have internally questioned the idealism of axiomatic truth and disembodied reason. In particular, for Hayles non-conscious cognition is a central activity of artificial intelligences governing automated systems today.

This article has addressed this view and argued that the crisis of deductive logic in artificial intelligence points to the emergence of an experimental axiomatics or speculative computation that forces us to re-articulate automated cognition. However, if the Scientific Image of computational logic has changed, it has also been able to question the Manifest Image of automated reasoning, which can no longer be explained in terms of an efficient execution of pre-established rules. Instead, the internal limits of algorithmic programing have marked the starting point for the fictive re-articulation of the Scientific and Manifest Image of how thinking works. If for

Hayles' non-conscious cognition overlaps with a form of cybernetic control based on inductive learning, this article questions the technocapitalist subsumption of machine thinking and the dominance of the data-driven order. Abductive reasoning offers one possible envisioning of a general artificial intelligence that works speculatively at various scales (human and machine) and not as a unified Scientific Image of cognition. Critical computation thus opens up the possibility to account for a sociality of reasoning within the computational strata, lurking beneath the seamless acceleration of irrational decision-making.

**Bibliography**


Chaitin, G. J. (1992) Algorithmic Information Theory. In: Chaitin G J (ed) Information-Theoretic Incompleteness, World Scientific.

Calude C S and Chaitin G (1999) Randomness Everywhere. *Nature* (400): 319 – 320.

Chaitin G J (2004) Leibniz, Randomness & the Halting Probability. Mathematics Today, 40(4) (available at Leibniz__Randomness___the_Halting_Probability-libre.pdf, accessed 19 November 2015).

Chaitin G J (2005) Meta Math! The Quest for Omega. New York : Pantheon.

Chaitin G J (2006) The Limits of Reason, *Scientific American* 294 (3): 74 – 81.

Denecker M and Kakas A (2002) "Abduction in Logic Programming" Computational Logic: Logic Programming and Beyond. Kakas A and Sadri F ed. Heidelberg: Springer-Verlag Berlin.

Domingos P (2015) The Master Algorithm. How the Quest for the Ultimate Learning Machine Will Remake our World. NY, Penguin Random House.

Dowek G (2015) Computation, Proof, Machine: Mathematics Enters a New Age. Cambridge, Cambridge University Press.

Daston L (2010) "The Rule of Rules", Lecture Wissenschaftskolleg Berlin, November 21st.

Deleuze G and Guattari F (1983) Anti-Oedipus: Capitalism and Schizophrenia (Second Edition) Minnesota: University of Minnesota Press.

Hayles N K (2005) My Mother was a Computer. Chicago: University of Chicago Press.

Hayles N K (2014) Cognition Everywhere: The Rise of the Cognitive Nonconscious and the Costs of Consciousness. New Literary History, 45(2).

Hayles N K (2012) How We Think: Digital Media and Contemporary Technogenesis. Chicago: University of Chicago Press.

Josephson J and Josephson S (1996) Abductive Inference. Cambridge University Press.

Clark S (2014) Artificial Intelligence could spell end of human Race – Stephen Hawkins. The Guardian. Tuesday 2 December 2014 (available at **www.guardian.com>Sceince>Stephen** Hawking, accessed 23 November 2015).

Inoue Katsumi, Doncescu Andrei and Nabeshima Hidetomo (2013) "Completing causal networks by meta- level abduction." *Machine Learning*, Springer Verlag, 91 (2), pp. 239-277.

Kakas A C and Riguzzi F (2000) Abductive Concept Learning, New Generation Computing, Ohsma, Springer-Verlagm, 18, 243-294.

Lazzarato M (2012) The Making of the Indebted Man. An Essay on the Neoliberal Condition. Cambridge M.: MIT Press.

Mackenzie, D. (2011) 'How to Make Money in Microseconds', *LRB* 33:10, 16-18.

Massumi B (2009) National Enterprise Emergency: Steps Toward an Ecology of Powers. Theory, Culture & Society 26(6): 153–185.

Massumi B (2015) Ontopower: War, Powers, and the State of Perceptio*n*. Durham and London: Duke University Press.

Mirowsky P (2002) *Machine Dreams. Economics Becomes a Cyborg Science*. Cambridge: Cambridge University Press.

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

Minsky, M. and Seymour A., Perceptrons (1987) An Introduction to Computational Geometry, The MIT Press.

Peirce C S (1955) Abduction and induction. In: Buchler, J. (ed.) Philosophical Writings of Peirce. New York: Dover, 150–156.

Peirce C S (1992). The Essential Peirce - Volume 1: Selected Philosophical Writings: *(1867-1893)* v. 1. Houser, N. & Kloesel, C. (eds). Indiana University Press,

Peirce C S (2005) Reasoning and the Logic of Things: the 1898 Cambridge Conferences. In: Lectures by Charles Sanders Peirce. Ketner K. L. (eds). Amsterdam: Harvard University Press.

Rouvroy A (2011). Technology, Virtuality, and Utopia: Governmentality in the Age of Autonomic Computing. In: Hildebrandt M. and Rouvroy A. (eds) Law, Human Agency and Autonomic Computing. London: Routledge, 119-140.

Sellars W (1963) Science, Perception and Reality. Ridgeview Publishing Company.

Turing A M (1936). On Computable Numbers, with an Application to the Entscheidungsproblem" Proc. London Math. Soc. (2) 42 (1936–7), pp. 230–265. In:

Alan M. Turing, Collected Works: Mathematical Logic (2001). Gandy R. O. and C. E. M. (eds). North-Holland: Yates.

Wolfram S (2002) A New Kind of Science. Champaign, IL : Wolfram Media.

Wolfram S (2006) How Do Simple Programs Behave? Architectural Design 76, (4): 34 – 37.

Whitehead A N (1929) The Function of Reason. Boston : Beacon Press.

Whitehead A N (1967) Science and the Modern World. New York: Free Press.



---

[1] Learning Algorithms are an evolution of genetic algorithms invented by Holland in the 1980s aiming to transform data into knowledge. Algorithms are series of instructions telling a computer what to do. If the simplest of algorithms is to combine two bits and can be reduced to the And, Or, and Not operations, in more complex systems, we have algorithms that combine with other algorithms, forming an ecosystem. Generally speaking, every algorithm has an input and an output, as data goes in the machine, the algorithms execute the instructions and leads to the pre-programmed result of the computation. Instead, with machine learning, data and and the preprogrammed result enter the computation, whilst the algorithm turns data into the result. In particular, learning algorithms make other algorithms insofar as machines write their own programs. In other words, learning algorithms are part of the automation of programing itself: computers now write their own programs.

[2] In supervised learning, example inputs and their desired outputs are given so that the machine can learn a general rule able to map inputs to outputs. With unsupervised learning, algorithms are given no label and are generally used to discover hidden patterns in data or learning. Reinforcement learning instead involves algorithms that perform a certain task in a dynamic environment without being told exactly how to behave.

[3] See Katherine Hayles, Cognition Everywhere: The Rise of the Cognitive Nonconscious and the Costs of Consciousness". *New Literary History* 45(2), 2014.

[4] Hayles does not fully explain the specificities of conscious thinking. In this article, I consider the question of conscious and nonconscious thinking as both involving a prehensive mechanism of registering and evaluation data. I draw on Alfred N. Whitehead 's conception of prehension, which includes a distinction between physical and conceptual abilities of recording, evaluating and selecting

information. I draw on this important distinction to argue that algorithmic thinking involves sensible and intelligible modes of processing information, which include both non-conscious and conscious cognitive abilities. Instead, as I suggest later, algorithmic cognition is yet to acquire the function of reason insofar as incomputable layers of complexity cannot be fully integrated or compressed in algorithmic states. See Alfred N. Whitehead, *Process and Reality: An Essay in Cosmology* (New York: Free Press), 1978, pp. 23 – 26.

[5] Hayles makes reference to Stanisław Lem's Summa Technologiae to explain that non-conscious cognition involves no calculation and that complex problem can be more efficiently resolved without the hierarchies of reflexivity and consciousness (Hayles, 2014).

[6] I draw on Alfred N. Whitehead's discussion about the function of reason, which is constituted by at least three levels of data elaboration. The physical and conceptual levels of prehension that are common to all species at various degrees- moving from lower to higher degrees of selection, evaluation and decision. In addition to these levels, Whitehead points to the crucial function of reason in constituting a further level of abstraction, which he defines in terms of an abstract schema, involving the construction of a structure or system of relata (relations of relations or meta-relations). See Alfred N. Whitehead, *The Function of Reason* (Princeton University Press, 1929).

[7] It is interesting here to refer to Hayles' explanation of this distinction in her discussion of Metzinger's epiphenomenal view of the self, William James's idea of the self as a construct, Damasio's purposeful consciousness etc. Her point is that consciousness comes at the cost of constant confabulations that could not operate without the non-conscious cognition. For Hayles, this more general level of non-conscious cognition across many forms of cognitive agents, including animals, humans and machines (2014).

[8] In *How We Think*, Hayles argues that coding technologies have transformed reading and writing and fundamentally enabled perception and cognition to develop analytic skills that move through larger quantities of information. Her argument that Humanities are faced with the power of digital technology also points at how the relation with the scientific method of analysis can be productive for close reading of texts. Her effort to re-visit the relation between thinking as the fundamental grounding of the scope of the Humanities (i.e., of moving beyond mere analysis) is further complemented by her work about non-conscious cognition and her explanation that computation and in particular algorithmic

procedural thinking involves non-reflexive activities and ultimately side-skips any logical requirement (Hayles 2012; 2014).

[9] According to American pragmatist Wilfrid Sellars, in order to articulate the relation between objects and thought beyond the assumption that the real world is directly given to us, we need to distinguish between the *manifest image of man* and the *scientific image of man*. Despite the gender-specific reference to human being, or persons, Sellars' argument offers us a way to address the natural dimension of things and thoughts that can be explained scientifically or through a rigorous scientific method able to revise previous scientific truths in relation to the conceptual framework by which humans see themselves as part of the world. The Manifest Image indeed corresponds to a rudimentary but already conceptual framework, starting with a picturing of the condition of being human in the world. The Manifest Image thus account for the particularity of homo sapiens to be able to experience, to think and rationally act in the world of thinking of manifest appearances. Both these images are complex and global and do not constitute parts that sum up to a whole. Instead they are general images that give a naturalistic account of thinking of things and thinking of thoughts, whereby scientific epistemology coincides with an enterprise in knowing nature and yet such knowledge is the conditioning frame for the manifestation of thinking to occur and for the two images to fuse without merging into one another. In other words, the two images belong to the same order of complexity, defining a continuity of becoming between the images or a processual discontinuity that opens up the relation between nature and culture to scales of elaborations and continuous critical reflection about the objects described, understood, and represented. From this standpoint, this article is an attempt at analyzing the scientific image of computation (and thus its epistemological description in information and computational theory) and the manifest image of computation (the tendency of algorithmic processing of information to develop hypothetical thinking and abstract information form the social use of data). See Sellars W. 'Science, Perception and Reality'. Ridgeview Publishing Company, 1963, pp. 10-11. See also O'Shea, J.R. *Wilfrid Sellars: Naturalism with a Normative Turn*. Polity, 2007. See also Seibt J. "How To Naturalize Sensory Consciousness and Intentionality Within A Process Monism with Normativity Gradient: A Reading of Sellars" J. O'Shea(ed.) *Sellars and His Legacy*. Oxford University Press, 2015.

[10] See for instance, Pedro Domingo (2015*) The Master Algorithm. How the Quest for the Ultimate Learning Machine Will Remake our World*. NY, Penguin Random House, 2015; Carl Steiner, *Automate This: How Algorithms Came to Rule Our World*, NY: Penguin, 2012.

[11] I am referring here to research projects and computational applications emerged from the Affective Computing Group at MIT, which has devised computational skills in robotics and artificial intelligence that arise from, respond to, or influence emotions and other affective states. Amongst their research objectives are for instance, the design of modes of communicating affective-cognitive states, creating techniques that affect stress and frustrations, devising computational skills of emotional intelligence, developing personal technologies for self-awareness. See **http://affect.media.mit.edu/** (last accessed November 23rd, 2016. See Picard Rosalind W. *Affective Computing*, MIT, 2000.

[12] With the term digital philosophy, I am referring to mathematicians and theoretical physicists using the computational paradigm to describe physical and biological phenomena in nature and to develop a computational description of the mind. This approach problematically merges being and thought through computation and thus gives an algorithmic explanation to both biophysical reality and the thinking of reality. One of the most problematic assumptions in this paradigm is the view that algorithms that evolve over time (i.e. cellular automata) can compress, render or simulate the various levels of physical, biological and cultural randomness or contingencies. See Stephen Wolfram, How Do Simple Programs Behave? *Architectural Design* 76, (4): 34 – 37, 2002.

[13] Hayles makes a reference to the experiment reported by Brian Massumi about the missing half second and other empirical evidence of affective states discussed by Antonio Damasio (2015).

[14] I am referring specifically to the theorization of control and affective biopolitcs that can be found in the work of Massumi (2015). I have written about the relation between the ecological power and the end of rationality and instead the re-articulaiton of logic for political ends in the article "Computational Logic and Ecological Rationality", in *On General Ecology. The New Ecological Paradigm in the Neocybernetic Age*, Erich Horl with James Burton, London, Bloomsbury (forthcoming, 2017).

[15] In the movie Terminator, Skynet AI is an artificial general intelligence that acquires self-awareness and spreads across all computers serves, mobile devices, military satellites, androids and robots with the aim of safeguarding the world by conforming to its original program code (thus implementing deductive reasoning). Instead, the Skynet AI I am referring to here, would rather be open to the

contingencies and the data retrieved in the informational environment, which means that the original mandate of the code can evolve in unexpected directions.

[16] If Deleuze and Guattari's notion of immanent axiomatics involve that rules have been replaced with the material performativity of behaviours, experimental axiomatics instead refers to how rules – and logic – are experimental compressions of randomness.

[17] As opposed to cognitive theories of computation, according to which to compute is to cognise and thus to produce a mental map of the data gathered by the senses, and to computational theories of cognition, for which to think is a binary affair determined by pre-set sequences of logical steps, I draw on Whitehead 's notion of prehension. For Whitehead, prehensions are modes of registering data involving a sensual or physical and conceptual or non-sensuous mode of recording the external world or the impact of externalities defining the capacities of reception of an actual entity. See Alfred N. Whitehead, *Process and Reality*, 23.

[18] I understand the relation between critical and speculative computation in terms of a dynamic tension between reflection and anticipation, the conceptual tracking of causality and the tendency to structure unknown information. This also involves the tension between the critical act of thinking causality or local states and the capacities of thinking to become an abstract or general function able to transcend specificities. This means that whilst Whitehead recognises that all thinking emerges form the biopshysical constraints of the living, he also argues that the function of reason is to elucidate and evaluate the causes through which these can be transcended. The function of reason is not determined by the direct apprehension of experience, but is rather a function of abstraction of the particular entities involved, and crucially involves the elaboration of the general conditions of the observations that are expressible without having to make reference to particular relations. For Whitehead, the rational attainment of this condition of generality ensures that these hold for an indefinite variety of other occasions. Alfred N. Whitehead, *Science and the Modern World*, 24-25.

[19] Magnani clarifies that this model of abduction involves sentential, model-based and manipulative abduction, which not only describe the practice of abductive reasoning but also can be used to enhance the development of programmes that can computationally be able to re-discover or newly discover scientific hypothesis or mathematical theorems. See Lorenzo Magnani *Abductive Cognition: The Epistemological and Eco-Cognitive Dimensions of Hypothetical Reasoning*. Berlin Heidelberg:

Springer-Verlag, p.2. Magnani argues that abductive reason is irreducible to the deductive method of formal logics and this is demonstrated by the undecidability result of Turing's 'halting problem', p. 69.

[20] Manipulative abduction also concerns particular kinds of heuristics that resort to the existence of extra-theoretical ways of thinking – thinking through doing. According to Magnani, many cognitive processes are centered on external representations that allow to create communicable accounts of new experiences ready to be integrated into previously existing systems of experimental and linguistic (theoretical) practices (2009).

[21] Extensional knowledge is here opposed to intentional knowledge. Whilst the former concerns inferences to a current situation, the latter rather implies universality across different states. See Marc Denecker and Antonio Kakas, (2002) "Abduction in Logic Programming" in Computational Logic: Logic Programming and Beyond. Kakas A and Sadri F ed. Heidelberg: Springer-Verlag Berlin, 2002, 406).

[22] For instance, meta-level abduction for goal finding is used in drug design and pharmacology where hypothesis are goal oriented and also for the improvement of physical techniques in musical performance in completed causal networks. See Katsumi Inoue et al. "Completing causal networks by meta- level abduction." *Machine Learning*, Springer Verlag, 91 (2), 2013, pp. 241.

[23] My point is not to dismiss the possibility of automated thinking, but to theorise how the complex layers of algorithmic elaboration of data are able to condition and revise logical conclusions, can challenge both the ideas that automation is opposed to thinking but also that automation is the same as thinking.