

Critical Computation: Digital Automata and General Artificial Thinking

Luciana Parisi

Goldsmiths, University of London

Theory, Culture & Society

2019, Vol. 36(2) 89–121

© The Author(s) 2019

Article reuse guidelines:

sagepub.com/journals-permissions

DOI: 10.1177/0263276418818889

journals.sagepub.com/home/tcs



Abstract

As machines have become increasingly smart and have entangled human thinking with artificial intelligences, it seems no longer possible to distinguish among 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 establishing 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.

Keywords

abductive reasoning, automation, Hayles, machine learning, non-conscious cognition, techno-power

At the core of computational systems today there is a latent paradox: capital's investment in techno-intelligence has led to the explosion of non-conscious or pre-cognitive decisions. With high-frequency trading, Netflix and Amazon recommendation algorithms, with Uber and Air B&B live platforms and micro-targeted online dating sites, cognitive capital seems to have turned the subsumption of the 'general intellect', and thus of social intelligence, into a crowd of learning algorithms efficiently

Corresponding author: Luciana Parisi. Email: l.parsi@gold.ac.uk

Extra material: <http://theoryculturesociety.org/>

driving decisions without the support of consciousness.¹ This automation of the general intellect, based on the frequency of data use and content, defines a mediatic infrastructure of statistical modelling, pattern recognition, data mining, knowledge discovery, predictive analytics, self-organizing and adaptive systems. In particular, with the 1990s development of machine learning within branches of artificial intelligence (AI), the automation of cognition has introduced a new mode of algorithmic processing that learns from data without following explicit programming. The increasing adaptation of machine learning systems across financial, military, governmental and educational systems is fundamentally challenging notions of automation classically intended as mere reproduction 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 in the form of fixed capital. Instead, this qualitative extension of automation beyond the mechanical reproduction of instructions involves an overcoming of automation itself, whereby algorithmic rules now generate or construct patterns from the re-assemblage of data. What is at stake here is the *automation of automation*: the automated generation of new algorithmic rules based on the granular analysis and multimodal logical synthesis of increasing volumes of data. In particular, machine learning has been said to define the manifestation of a new form of intelligence able to automate automation (Domingos, 2015: 9). Here, the automation of the intellect does not simply imply the subsumption of social values through a new rationalization of social thinking. The automation of automation instead concerns a meta-level of algorithmic function, whereby social thinking is not only organized by machines, but is algorithmically engendered through neural networked layers that eventuate new meaning of artificial thinking. The automation of automation therefore points out that the subsumption of the intellect in capital's valorization of automated cognition relies upon the social meaning of artificial thinking implied within the technoscientific descriptions of intelligence.

This article argues that changes in the scientific image of computation and cognition stem from a socially mediated understanding of artificial thinking involving not a symbolic representation of ideas but a dynamic logic of algorithmic learning. These are historical changes in the scientific and technological descriptions of intelligence stemming from the computational theorization of the limits of reason, and post-Second World War experiments with the automation of reasoning in machines. Katherine Hayles' view of this shifted meaning of automated intelligence in terms of non-conscious cognition² points out that cognitive systems perform complex modelling and informational tasks at the fastest speed without abiding by the formal languages of mathematics or explicit equations.³

In the attempt to qualify further the distinction between consciousness, unconsciousness and awareness, thinking (involving awareness) and cognition (that does not require consciousness, but can perform complex modelling and informational tasks), Hayles discusses the emergence of what she calls the 'cognitive non-conscious' working at a 'lower level of neural organization, not accessible to introspection' (2014: 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 behaviour and emergent capacities to process new data (2014: 4–5). In particular, Hayles distinguishes between conscious thinking, non-conscious cognition and material processes (2014: 5),⁴ 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 non-conscious cognition devices.

In other words, Hayles addresses the changing meaning of how machines think in terms of today's interactive, adaptive and learning algorithms that skip the logical order of deduction, which was central to the Enlightenment theorization of the function of reason.⁵ In agreement with Hayles, this article argues that the non-logical thinking of automated systems overlaps with the efficacy of a cybernetic calculus whereby control and prediction rely on inductive learning. Here cybernetic control becomes infused with the non-conscious algorithms of cognitive capital.

Hayles (2005) presents us with epistemological shifts in theories of cognition, which, she suggests, are necessarily embedded in social practices and discourses (and are thus not to be simply addressed as a sort of teleological overcoming of humanity). To further account for this question of machine thinking, however, this article extends this epistemological articulation of artificial thinking by borrowing Wilfrid Sellars' (1963) theorization of the scientific and manifest image. I argue that the scientific image of intelligence (e.g. the material physical, biological, computational description of intelligence) is mediated by the manifest image of intelligence involving the socio-cultural self-awareness of a form of artificial thinking that admits the capacity of machines to think conceptually and act rationally. According to Sellars, these double levels of material and conceptual activities are equally pregnant with meaning. In order not to fall back into the myth of the given (the assumption that thinking merely coincides with its neurological descriptions), namely the essentialism of cognition, or the empiricism of scientific descriptions and conceptual forms, both the scientific and manifest images are to be worked through to explain the relation between the material and the mental activities we are concerned with.⁶ From this

standpoint, when speaking of algorithms, computation and AI, this article argues that it is important to address scientific and technical descriptions as socially mediated meanings. In other words, while there is no direct translation between the scientific descriptions of the functions of computation and the conceptual manifestation of their meaning, it is not possible to admit that the scientific understanding of computational intelligence is not socially mediated, embedded and determined by the socio-technical meanings of artificial thinking. From this standpoint, a critical articulation of how machines may think is already implied in the collective conceptions of automated cognition, which are re-directing Hayles' distinction between non-conscious and conscious cognition towards the image of the automation of automation.

While it is arguable that computation involves interdependence between data, software, code, algorithms, and hardware, the automation of automation instantiated within new forms of machine learning, for instance, has emerged from a shift in computational models of logical reasoning: namely, from deductive truths applied to small data to the inductive retrieval and recombination of infinite data volumes. In particular, the transformation of the relation between algorithms and data contributes to explaining the historical origination of non-deductive reasoning, activated with and through machines. As Lorraine Daston (2010) points out, already during the Cold War the conception of reason as based on truth, and on the faculty of judgement and discrimination, became historically reconceptualized in terms of patterns, and reason as 'the rule' came to be understood in terms of ruling procedures with the task of calculating probability.

This embedding of reasoning into machines is entangled with the development of statistics and pattern recognition, which define the socially mediated manifest image of algorithms as learning machines making predictions by recognizing data (through granular analysis, flexible and modular patterning of categories with textual, visual, phonic traits). As the system gathers and classifies data, learning algorithms therefore match-make, select and reduce choices by automatically deciding the most plausible of data correlations. Machine learning indeed is used in situations where rules cannot be pre-designed, but are, as it were, achieved by the computational behaviour 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, 2015: 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 AI 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, critical thinking today needs to be

concerned with more general questions about what cognition is and how it has come to coincide with the computational architecture of algorithms, data, software and hardware, and with experiments in robotics sensing and self-awareness. However, the implications of automated cognition, central to the critique of cognitive capital, are far from being settled and will be the concern of a critical computation theory addressing the specificity of this manifest image of algorithmic thinking. For instance, the possibility of elaborating a rule from data retrieval rather than applying a given rule to outcomes points to a form of cognition that cannot be defined in terms of problem-solving, but will be understood as a general method of experimenting with problems. In particular, with machine learning, automation has involved the creation of training activities that *generalize* the function of prediction to future cases – a sort of inductive parable that, from particulars, aims to establish general rules. Here, in the case of supervised, unsupervised and reinforcement learning algorithms,⁷ a critical computation will refer not simply to mindless training, but rather offer an enquiry into forms of inference characterizing this artificial thinking. This enquiry will navigate the tension between theories of reason vis-a-vis the emergence of non-conscious intelligence in automated cognition.

This article suggests that a critical view of computation requires an effort to unpack this tension to account for indeterminacy in conditions of knowledge that both constrain and enable the scientific and manifest image of algorithmic thinking. From this standpoint, if indeterminacy is central to the epistemological possibilities of algorithmic thinking beyond deductive logic, the automation of automation will be seen not as a mindless execution of rules or a form of unconscious cognition, but as a critical mode of artificial thinking. 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 but re-hypothesized, re-assessed and invented. Although abductive logic is mainly performed in automated models for medical diagnosis, the possibility that automated systems can construct new forms of logical complexity, which could enable the theorization of a general artificial intelligence other than that of the statistical regime of inductive capital, will nonetheless be entertained. Learning algorithms are already a step towards this envisioning of abductive artificial intelligences, involving conceptual re-elaboration from data correlations, rules and functions that can be used to construct new hypotheses. This amounts to an automated meta-abductive reasoning, whereby learning algorithms elaborate a meta-hypothetical function through which they infer missing rules, facts and unknown causes (Inoue et al., 2013: 240).

Despite the local applications of algorithmic procedures in design, logistics, music and economics, it is evident today that the automation

of automation rather involves a cultural transformation in the conceptualization of reasoning with and through machine thinking. This is also a transformation in the meaning of cognitive capital increasingly relying on the automation of learning, and of the intelligible elaboration of new forms of data correlation, evaluation, selection and decisions. Machine learning automata are understood to behave like cognitive systems that are evolutive, adaptive, and exhibit co-causal and emergent properties (Hayles, 2014).

From this standpoint, Hayles' work already offers a reassessment of cybernetics and computation as central to automated systems of feedback control and logical procedures, which have exposed the changing meaning of cognitive activities, generalized from particularities (animals, humans and machines).⁸ Her insights about neoliberal forms of governance no longer being constituted by the law, the norm and reason, but by control functions, behavioural operations based on procedures within self-regulating autopoietic agencies (i.e. reiterative loops, sequential tasks, flexible protocols and flows of data), point to the shifted meaning of artificial thinking. As rule-obeying behaviours become substituted by the performativity of machinic functions (i.e. what *x* or *y* do and do not do, and what they stand for), the indeterminacy of learning outcomes has also become central to the epistemological critique of the end of reason. This shift from rule-obeying truths to an algorithmic pragmatism, using data to search for and predict truths, has also been understood as the end of rational choice (MacKenzie, 2011; Mirowski, 2002).

From this standpoint, while suspending current figurations of automated intelligence (Domingos, 2015; Steiner, 2012), the transformations of the scientific and manifest image that describe algorithmic performativity have already opened up new meanings of artificial thinking. With machine learning, algorithms indeed are no longer mere instructions, but are rather performative of instructions. Algorithms learn: they adapt, adjust and evolve their behaviour according to a qualitative synthesis of vast quantities of data. Their performative activity is afforded by their capacity to compress large quantities of information and thus transform outputs into new inputs, involving a new synthesis of reasoning and calculation. Here data do not have to fit categories, but are redefinable 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 here an attempt at discussing a critical theory of computation that questions the predominance of two models of AI in the techno-capital valorization of automated cognition: 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 ontologization 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' (Rouvroy, 2011: 126). According to Rouvroy, the inductive regime thus appeals to the immediacy of data retrieval, which eradicates potentiality and/or indeterminacy, limiting the possibility of a critical approach to technology (2011: 127).

My attempt to re-theorize automated intelligence draws from these views but also argues that the crisis of deductive logic is mediated by new meanings of artificial thinking stemming from the scientific image of experimental axiomatics, which has indeterminacy at its core. I suggest that as the scientific image of computational logic has changed (from Turing to post-Turing descriptions of intelligence), it has also questioned the manifest image and thus exposed the changing meaning of automated reasoning. Here artificial thinking no longer coincides with the efficient execution of pre-established rules. The internal limits of logic in computation have rather pushed the epistemological view of artificial thinking beyond deductive and inductive models. Drawing on Hayles' theorization of non-conscious cognition as a form of inductive learning, this article questions the assumption that techno-capital always already subsumes any mode of machine thinking, and ultimately of automation. Instead, a critical view of artificial thinking is an attempt at reducing the dominance of data-driven systems of retrieval and transmission, deprived of any hierarchical logic, to only one form of automated cognition through which capital is extending social subjection. And yet, capital investment in machine intelligence will also be questioned with and through the epistemological proliferation of multimodal logics (and thus the socially mediated meaning of artificial thinking) that expose the possibilities of automated reasoning beyond the function of fixed capital. From this standpoint, this article argues that abductive reasoning offers one possible envisioning of a general artificial thinking that accounts for multimodal logic and does not simply mirror one specific image of automated cognition.

Computation Is Not Cognition

In *My Mother Was a Computer*, Hayles (2005) discusses the view of computation as a universal model of cognition and intelligence. Hayles refers to the development in AI in the 1970s, 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 Jonze's *Her* (2013), and more recently Alex Garland's *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 up to another 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 Hawking's (2014) 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'.

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 is meant by reasoning, remains open and in need of more discussion. As Hayles (2005) has already pointed out, there are at least two main positions that reveal the tension between automation and reasoning. 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 by means of the scientific image of what cognition is. In particular, the digital explanation of cognition remains attached to a deductive model 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 the research into affective computing and emotional intelligence, for example)⁹ instead relies on local low levels of neural organization, 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 bound to conscious cognition, smart devices are much faster than us at making connections (Hayles, 2014).

When discussing the power of algorithmic decision underpinning the mediatic infrastructure of the political, cultural and social infrastructure today, we are thus faced with the dominant view of two modes of logical reasoning, defining intelligence and its manifestations. On the one hand, the reduction of reasoning to the computational view of cognition based on the manipulation of symbols, and, on the other, the anti-cognitivist argument that computational decisions act below cognition at the local level of non-logical communication. 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 social embedding of reasoning) 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 the automation of automation in both pragmaticist and transcendental terms. In particular, from pragmatism, I take the important proposition that reason is not a formal *a priori*, but corresponds to the conceptual infrastructure of social practices. This means that the logical operations of reason and its rule-bound 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 reassessment 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 and so on.

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 reorganization of techno-capitalism (i.e. the economic investment in automated networks) involving the view that automated intelligence corresponds to affective or non-conscious cognition.

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. 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 (2005), however, digital philosophy contains no a priori 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. 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 (2005) highlights the influence of second-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. This influence of second-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 techno-capitalism. 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 second-order cybernetics, the reflexive loop between mind and matter shows how logical reasoning rather worked in a backwards way, 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 (2014) explains, 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 among themselves and make decisions 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.

Non-conscious 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* (Minsky and Seymour, 1987), Marvin Lee Minsky claimed that a single neuron could only compute a small number of logical predicates in any given case, and his experiments cast a long shadow on neural network research in the 1970s. In the late 1980s 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 among 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-making heuristically selecting data 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-popularized in the aesthetically powerful movie *Ex Machina* (dir. Alex Garland, 2014), 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 mechanical accomplishment of tasks. As Hayles (2014) points out, the advance of non-conscious cognition in intelligent machines precisely exposes the new meanings of our understanding of cognition. 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 of neural organization, iterative and recursive patterns of preservation, this inductive method of reasoning implies the emergence of a total behaviour or an intelligent effect that is bigger than the parts constituting it. From this standpoint, as Hayles (2014) observes, emergence, complexity and adaptation, and the phenomenal experience of cognition cannot be reduced to material processes. 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 intelligence, nor specific materiality of intelligence and, while humans share levels of consciousness with other animals, it is remarkable, Hayles (2014) points out, that non-conscious cognition operates across humans, animals and technical devices. In particular, the low-level activities of non-conscious cognition – described for instance in the example of the missing half-second,¹⁰ at speeds so fast as to be imperceptible and affective speeds – 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 (2014), what is promising regarding cognitive non-conscious technical devices is that they can operate in temporal regimes inaccessible to human consciousness and exploit the missing half-second to their advantage. This also implies a machine-like cognition of temporalities, pointing out that

automated systems are able to tap into 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 Sellars' (2014) distinction between the scientific and the manifest image, cognition is here not to be taken as a direct image of material processes. Hayles indeed espouses the idea that the anti-deductive operations of non-conscious cognition are somatically marked, but are also phenomenologically embodied, and mediated by meaning. Here, there is no direct correspondence, but instead an elaboration of the material, involving the mediation between the biophysical and neural states with perceptive and cognitive receptions. Since cognition is entwined with the recall and re-enactment 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 a 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, while 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 (Hayles, 2014).

This shift from formal cognition based on deductive inference to a model of non-conscious 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-bound logic and that, rather, it relies upon non-conscious functions entrenched in the weight of data in networks.

While 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 the meaning of artificial thinking, there are questions that are to be addressed. If, for instance, high-frequency trading algorithms

are to be considered as non-conscious cognitive functions, effectively changing socioeconomic behaviour, 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 the manifest image of automated logical reasoning is pushing beyond the totalizing image of techno-power?

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 into 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 techno-capital.

This shift is characterized by a re-orientation of the practices of real subsumption, in which capital's investment in the general intellect has led human-machine networked intelligences to become a motor of cognitive and affective labour, and, as some argue, of the capitalization of the relational qualities of life (Massumi, 2015) attached to the regime of indebtedness (Lazzarato, 2012).¹¹ The manual phase of automation of industrial capitalism imparted an ontological separation between human labour and the accumulation of labour value incorporated in machines. While human labour has become valorized in terms of variable labour or force, the machines' task was rather to absorb, preserve, accumulate and reproduce the value of labour within itself. It was through machines that the rational principles of task-oriented efficiency of the assembly line could be realized 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 mainly concerned with avoiding contingency and human errors. Instead, this networked form of abstraction (of relational value) is now sustained by 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 reconfigure information according to the social use of data. This form of automation has reached a non-prescribed

form of valorization insofar as algorithms experiment with data by learning, adapting and assessing the value of large amounts of information. While this intelligent valorization 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 *Anti-Oedipus* (1983), Deleuze and Guattari had already identified this transformative tendency of the human-machine network of abstraction, and had warned us against what they called 'immanent axiomatics' (1983: 246). The rationalization 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 (1983: 233). Not only had calculative machines 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 axiomatics – and its rule-bound activities – are 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 (2009, 2015; see also Mirowski, 2002) analysis of the contemporary reconfiguration of neoliberal governance, one could argue that the end of rational economy has been accompanied by the crisis of the rational implementation of machines. The computational infrastructure of social media, for instance, as the privileged form of marketing, branding, economic operations, political campaigns, institutional governance, security screening and so on, no longer abides by 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 techno-capitalism has invested in human intelligence and creativity, driving humans to become self-entrepreneurs or governors of their extended self.

In the movie *Her* (dir. Spike Jonze, 2013), the artificial intelligence Samantha acts in a world in which not only is affectivity fully programmed and programmable, but also 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

operate at speeds so fast as to be imperceptible, but is also equipped with an empathic quality of prediction, tuning into the viscosity 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,¹² as the automation of automation gathers self-aware intelligences and leaves humans behind, resigned to not being able to think and feel anything anew.

However, while 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 (2015) defines as 'ecological rationality' acting through the affective intelligence of the body, turning symbolic values into lifestyles, and rules into experiential qualities. 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 epitomize the non-inferential reasoning of embodied intelligence, making decisions without formal calculation. This form of anti-logos demarcates the techno-capitalist deterritorialization of rationality, which resolves the tension between automation and thinking through the convergence of consciousness and affect. Far from being liberating, the deposing 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 human-machine-animal non-conscious cognition, but to critically re-address the function of reason and to theorize – rather than reject – the automated use of inferential reasoning as part of a general artificial thinking. My efforts here concern not only an anti-essentialist theorization of thinking, for which reasoning can be understood as an elaboration of material, non-conscious and conscious cognition, but also involve a re-articulation of the critical possibilities 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 the problem of unpredictability or randomness in information theory is not a sign of logical failure but of the transformation of the scientific image of the relation between ratio 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 (2005, 2006), however, incomputables 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.¹³ 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 also 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 in the brain. Following Chaitin, it is possible to suggest that randomness in computation, as that which constitutes the very limit of computational deduction, demarcates the point at which automated cognition coincides not with non-conscious functions, but with algorithmic intelligibility, extracting more information from data substrates. Chaitin (2006) 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.

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, the algorithmic intelligibility of data environments involves a speculative function through which unknowns are computationally prehended.¹⁴

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 to be forced to confront the computational configuration of non-sensuous or proto-conceptual patterns that are 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, while data

seem to be mindlessly aggregated by non-conscious patterns, the scientific image of experimental axiomatics rather asks us to account for a new meaning of artificial thinking embedded in the intelligible activities of algorithmic prehensions.

From this standpoint, one has to view techno-capital not only as the reduction of reasoning to the non-conscious activities of machines but also as involved in a deeper transformation of automated thinking, namely exposing an alien or denaturalizing process of reasoning with and through machines.

Parallel and distributed orders of computational language point to a new form of informational stratification of contingencies, precisely involving this algorithmic processing of data. This can be understood as an artificial mode of intelligibility that works through the computational structuring of social thinking. From this standpoint, a critical approach to computation requires us to look closely at the historical transformation of the automation of thinking, involving not simply an abstraction of neural functions of the brain, but of the social practices of thinking and acting. While 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 techno-capital and its non-logical governance. A critical approach to this dominant understanding thus requires that the scientific image of computation should be accounted for in its historical changes, which involves reassessing 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 conceptualization of automated reasoning, within a view of a general artificial thinking. This space will aim not only to defy the exceptionalism of human consciousness but also to reinvent what consciousness and reason can become in this configuration of automated thinking. The next section explores this point further.

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 may not

simply involve a replacement of reason with non-conscious technologies of decision. Instead, the realization 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 optimized non-rule-bound 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 will first of all address the speculative function of reason¹⁵ insofar as the limits of automated 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-theorizing 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 tripartite configuration of methods involving deductive, inductive and abductive reasoning.

One important instance of this configuration can already be found in Charles Sander Peirce's (1998: 273; see also his 1995) triadic system of logic, which admits that thinking entails an abductive-inductive-deductive circuit of inference. 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 thinking. In particular, Peirce's triadic method always starts from a hypothetical or speculative explanation of events. This involves first 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 generalization 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 confirm it for instance. In particular, for Peirce, induction corresponds to a process of evaluation, which may produce very simple new ideas, but ones that are not sufficiently new to engender a new

hypothesis (Magnani, 2009: 289). While 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 generalization too.

However, according to Peirce, veritable reasoning will include abduction as this mainly consists in creating new 'explanatory' hypotheses. Abduction is a process of inferring facts, laws, hypotheses that can speculatively explain some unknown phenomena. In other words, it defines reasoning not simply in terms of evaluation, but also as the formation of new explanatory hypotheses (Magnani, 2009: 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, 2009: 65–70).

Rules are thus not fixed and are not symbolic representations 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 the methods of both deduction and induction into a larger space of reasoning that includes hypothetical inference. Here the inductive testing of hypotheses – or the generalization of new simple ideas – is not a proof of truths actualized 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 hypotheses. This involves a dialectic overlapping of induction and deduction, the validity of both testing and truth within the speculative articulations of hypotheses.

Since automation is becoming transcendental because of its functions of logical implications (deduction) and generalization of known concepts and objects (induction), Peirce's argument for abductive reasoning is useful because it challenges both the meta-computational 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 the part of learning algorithms – only serve to grant the consequent function of reason that, to put it in Alfred N. Whitehead's terms (1967: 24–5), arrives to establish the permanence of rules through an abstraction or a speculative formalization of what occurs as a consequence of the relation between particulars.

The pragmatist method of abduction makes a claim not only for the existence of intelligible patterning but also for a conceptual elaboration of what is implicit within patterns, 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 synthesized 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 of non-inferential transmission) obfuscate the constructive potential of Hayles' theorization about what is at stake with an artificial form of cognition. Similarly, these insights can contribute to suspending 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 social structuring of reasoning itself, including the triadic configuration of abductive, inductive and deductive inferencing. If the manner in which thought thinks itself thinking has always been mediated by the environment – and is thus ampliative and not representational – the formation of new hypotheses from the increasing availability of data also defines the proliferation of non-human intelligences. And yet, for automated reasoning to generate new hypotheses, 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 primarily the cognition of notions, tasks and functions. Instead, it requires apprehension through errors, blind spots, unknowns. Here, the possible fallibility of reasoning points out that Hayles' view of non-conscious cognition is central to abductive possibilities of learning because it is involved in the construction of hypothetical scenarios, pushing the limits of automation beyond data recombination or the mere execution of rules.

As Lorenzo Magnani (2009) argues, since the 1980s abductive reasoning has been adopted by diagnostic and expert systems, and in general by

a computational infrastructure of reasoning based on the use of inferential synthesis or inference to the best explanation (2009: 68). Importantly, Magnani distinguishes between model-based abduction – a theory-based inference – and manipulative abduction – defined by action-oriented or extra-theoretical reasoning (2009: 7, 9–12).¹⁶

Theoretical or model-based abduction corresponds to the exploitation of internalized models, diagrams or pictures and illustrates, according to Magnani, much of what is important in creative abductive reasoning, in humans and in computational programs (2009: 23–4, 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’ (2009: 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 behaviour that creates communicable accounts of new experiences and integrates them into existing systems of experimental and linguistic practices (Magnani, 2009: 46).¹⁷

In models of artificial intelligence, for instance, abductive reasoning has been used for diagnosis, planning, natural languages processing, probability theory and formal programming (Magnani, 2009: 5). If abduction has a logical form that is distinct from deduction and induction, it is because, when working computationally, the selective or creative activities of this retroactive thinking (i.e. that starts from consequences to track causes) involves 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 pixel arrays and propositional statements. Here, the computer program can manipulate and modify these representations and make new geometrical constructions, for example adding parts, moving elements and components (Magnani, 2009: 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 observing 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 indicate the role of creativity in diagrammatic reasoning (Magnani, 2009: 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 extensional knowledge’ (Denecker and Kakas, 2002: 405).¹⁸ While inductive inferences are 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 hypotheses and explain their incompleteness.

The automation of abduction has also been specifically used in logical systems aiming to solve problems of scheduling and planning, of optical music recognition, information integration and software inconsistencies (Kakas and Riguzzi, 2000). In particular, the notion of Abductive Concept Learning defines algorithms that integrate ‘explanatory learning’ (predictive) and ‘learning with confirming’ (descriptive), using methods of both 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 and are thus predictive, 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 and Riguzzi, 2000: 3).

In the specific context of machine learning, abductive reasoning is used to elaborate hypotheses in the face of incomplete information and overcome the problem of overfitting, whereby algorithms are heuristically programmed to learn from past data and thus delimit the configuration of larger and new hypotheses to given patterns of trial and error (Kakas and Riguzzi, 2000: 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 (Kakas and Riguzzi, 2000: 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 retroactively but also speculatively, by inventing hypotheses 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).

While 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 hypotheses' to infer missing rules, missing facts and unknown causes (2013: 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' (2013: 241).¹⁹ 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 of explanation) but are also, importantly, hypothesized beyond the observable.

As automated cognition has entered the realm of hypothesis-making by connecting explanations between objects, objects and concepts, and concepts themselves, it has also reopened 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 Thinking

We can now conclude that the understanding of algorithmic automation in terms of what Hayles has called non-conscious cognition may perhaps not meet this pragmaticist generalization of reasoning. However, I have suggested that Hayles' insights into the new meaning of cognition, as embedded in the scientific image of non-conscious decisions, already offer us an argument about the epistemological transformation of thinking in relation to machines. In particular, the neuro-biological descriptions of the relation between non-conscious cognition as bodily markers and consciousness as the re-presentation of bodily states strongly challenge the manifest image of reason coinciding with the model of deductive logic. From this standpoint, Hayles' discussion of non-conscious cognition already points to the conceptual mediations involved in the relation between distinct species of algorithms and between algorithms, data, software programs, interfaces and hardware circuits. In short, Hayles' view already paves the way for a critical computation that challenges the meaning of cognition by addressing the dynamic relation between the scientific and the manifest image of thinking. One crucial contribution to critical computation is Hayles' (2017: 22) articulation of biological and technical modes of cognition involving a process of interpretation that are context-bound, and thus connects information with meaning. It is precisely through the focus on the relation between information and meaning, and between distinct scientific descriptions of

cognition (from evolutionary to computational and neuro-biological theories), that Hayles' work offers a re-reading of the epistemological distinction between human and non-human cognition. In her effort to articulate together distinct scales of cognition that could account for a general artificial thinking, that she calls 'planetary cognitive ecology' (2017: 11), Hayles (2017: 174) specifically argues that computational media are cognitive systems that interact with human cognitive capabilities at the level of sensation, of the cognitive non-conscious and of modes of awareness (including both consciousness and the unconscious). In other words, her visions address how computational media are transforming the cognitive possibilities of the space of reason.

This article engages further with these possibilities and focuses on logical reasoning in machines beyond the dominant models of deduction and induction. It argues that the scientific image of the cognitive non-conscious is central to the capitalization of affective states now absorbed within the computational form of fixed capital and also subtends the dominance of a manifest image whereby logical reasoning has been replaced by automated correlations of data. Critical computation instead aims to trace the transformation – not disappearance – of logic and reason in automated systems of cognition.

This article has suggested that the theoretical and manipulative logic of abductions in automated systems shows the triadic configuration of a complex space of reason in the gaps between causal efficacy (the non-conscious fast correlations among all forms of data) and the experimental finality of algorithmic processing that includes the abductive-inductive-deductive logical reasoning reconfiguring causality beyond a linear sequence of given causes and effects. This is also to argue that the algorithmic use-meaning of data, more importantly, entails a transformation of the manifest image of reason that exposes how a new techno-social culture of thinking is embedded in the externality of cognition. While it is possible to discern the manifest image of this social cognition from the scientific image of automated intelligence involving the dynamic synthesis of logic and calculus, the article argues that the limits of deductive reasoning will be rather addressed as a symptom of the emergence of a critical function of and within the self-determination of computation as the dominant space of reason. Here the fallacy of reasoning corresponds to the point of departure for a computational generation of hypotheses, a speculative function within the automation of cognition.

Without taking into account this epistemological transformation in machine thinking, debates about cognitive capital risk overlooking the crucial realization within techno-capital that the condition of automating rule-bounded logic required the alienation of reason, that is the origination and expansion of the space of reasoning beyond the logic of deduction and induction. Similarly, by overlooking the possibility of a critical

re-theorizing of reason from within the automation of cognition as an engine through which to expose the dynamic tension between the scientific and manifest image of artificial thinking, it is not possible to account for an epistemological alternative to the given opposition between reason and automation. 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 logical reasoning involves a multifunctional elaboration of hypotheses 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 second-order reflection of thinking through doing, but by the emergence of a third level of abstraction, what I called the automation of automation.

From Magnani's (2009) argument and the wider use of abduction in computation it 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 distinction between the know-how and the knowing that capacities – to put it in Wilfrid Sellars' terms (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 incomputables is one instance of abductive logic insofar as it points to a rudimentary level of making incomputable data partially intelligible. However, the determination of this randomness is demarcating the tendency of AI to develop beyond its rudimentary intelligible capacities and points to a generalized socialization of rules, abstracted from the particularity of data contexts and yet exceeding models of encoded cognition.²⁰ 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 already involves within itself a mode of conceptual inferences, where rules and laws are invented and experimentally structured from the social dimensions of computational learning.

This article has taken inspiration from Hayles' analysis of computational intelligences about *what* – and *how* – thinking is 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 identified 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.

Since the scientific image of computational logic has changed, it has also questioned 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 programming have marked the starting point for a critical 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 techno-capitalist subsumption of machine thinking and the dominance of the data-driven order. Abductive reasoning offers one possible envisioning of a general artificial thinking that works speculatively at various scales (human and machine) and does not represent a unified scientific image of cognition. Critical computation argues for the theorization of a sociality of reasoning within the computational strata lurking beneath the seamless acceleration of irrational decision-making.

Notes

1. Learning algorithms are an evolution of genetic algorithms invented by Holland in the 1980s aiming to transform data into knowledge (see Holland, 1975). 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 this leads to the pre-programmed result of the computation. Instead, with machine learning, data and the pre-programmed result enter the computation, while 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. Hayles does not fully explain the specificities of conscious thinking. In this article, I consider the question of conscious and non-conscious thinking as both involving a prehensive mechanism of registering and evaluation data. I draw on Alfred N. Whitehead's (1978: 23–6) 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.
3. Hayles makes reference to Stanisław Lem's *Summa Technologiae* to explain that non-conscious cognition involves no calculation and that complex problems can be more efficiently resolved without the hierarchies of reflexivity and consciousness (Hayles, 2014: 200).
4. It is interesting here to refer to Hayles' (2014) 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

and so on. Her point is that consciousness comes at the cost of constant confabulations that could not operate without non-conscious cognition. For Hayles (2014), this more general level of non-conscious cognition exists across many forms of cognitive agents, including animals, humans and machines.

5. I draw on Alfred N. Whitehead's (1929) 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).
6. According to American pragmatist Wilfrid Sellars (1963), 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 accounts for the particularity of *Homo sapiens* to be able to experience, to think and to act rationally in the world of thinking of manifest appearances. Both these images are complex and global and do not constitute parts that add 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 to analyse 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 (1963: 10–11); see also O'Shea (2007) and Seibt (2015).
7. 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.

8. In *How We Think*, Hayles (2012) 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 revisit 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-steps any logical requirement (Hayles, 2012, 2014).
9. I am referring here to research projects and computational applications that 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. Among 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, and developing personal technologies for self-awareness. See <http://affect.media.mit.edu/> (accessed 23 November 2016) and Picard (2000).
10. 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 (2000).
11. I am referring specifically to the theorization of control and affective biopolitics that can be found in the work of Massumi (2015). I have written about the relation between ecological power and the end of rationality and instead the re-articulation of logic for political ends in Parisi (2017).
12. In the movie *Terminator*, Skynet AI is an artificial general intelligence that acquires self-awareness and spreads across all computers servers, 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 be open to the contingencies and the data retrieved in the informational environment, which means that the original mandate of the code could evolve in unexpected directions.
13. If Deleuze and Guattari's notion of immanent axiomatics means that the rules have been replaced with the material performativity of behaviours, experimental axiomatics instead refers to how rules – and logic – are experimental compressions of randomness.
14. As opposed to cognitive theories of computation, according to which to compute is to cognize 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 Whitehead (1978: 23).

15. 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 while Whitehead recognizes that all thinking emerges from the biophysical 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. See Whitehead (1967: 24–5).
16. Magnani clarifies that this model of abduction involves sentential, model-based and manipulative abduction, which not only describes the practice of abductive reasoning but also can be used to enhance the development of programmes that can computationally have the ability to rediscover or newly discover scientific hypotheses or mathematical theorems. See Magnani (2009: 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' (2009: 69).
17. 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 (2009), many cognitive processes are centred on external representations that allow for the creation of communicable accounts of new experiences ready to be integrated into previously existing systems of experimental and linguistic (theoretical) practices.
18. Extensional knowledge is here opposed to intentional knowledge. While the former concerns inferences to a current situation, the latter rather implies universality across different states. See Denecker and Kakas (2002: 406).
19. For instance, meta-level abduction for goal finding is used in drug design and pharmacology, where hypotheses are goal oriented, and also for the improvement of physical techniques in musical performance in completed causal networks. See Inoue et al. (2013: 241).
20. My point is not to dismiss the possibility of automated thinking, but to theorize 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.

References

- Calude CS and Chaitin GJ (1999) Randomness everywhere. *Nature* 400: 319–320.
- Chaitin GJ (1992) Algorithmic information theory. In: Chaitin GJ (ed) *Information-Theoretic Incompleteness*. Singapore: World Scientific.

- Chaitin GJ (2004) Leibniz, randomness and the halting probability. *Mathematics Today* 40(4). Available at: <https://arxiv.org/pdf/math/0406055.pdf> (accessed 18 December 2018).
- Chaitin GJ (2005) *Meta Math! The Quest for Omega*. New York: Pantheon.
- Chaitin GJ (2006) The limits of reason. *Scientific American* 294(3): 74–81.
- Clark S (2014) Artificial intelligence could spell end of human race – Stephen Hawkins. *The Guardian*, 2 December. Available at: <https://www.theguardian.com/science/2014/dec/02/stephen-hawking-intel-communication-system-astrophysicist-software-predictive-text-type> (accessed 23 November 2015).
- Damasio A (2000) *The Feeling of What Happens: Body and Emotion in the Making of Consciousness*. New York: Mariner Books.
- Daston L (2010) The rule of rules. Lecture at Wissenschaftskolleg Berlin, 21 November.
- Deleuze G and Guattari F (1983) *Anti-Oedipus: Capitalism and Schizophrenia*, 2nd edn. Minneapolis, MN: University of Minnesota Press.
- Denecker M and Kakas A (2002) Abduction in logic programming. In: Kakas A and Sadri F (eds) *Computational Logic: Logic Programming and Beyond*. Heidelberg: Springer-Verlag.
- Domingos P (2015) *The Master Algorithm. How the Quest for the Ultimate Learning Machine Will Remake our World*. New York: Penguin Random House.
- Dowek G (2015) *Computation, Proof, Machine: Mathematics Enters a New Age*. Cambridge: Cambridge University Press.
- Hayles NK (2005) *My Mother was a Computer*. Chicago: University of Chicago Press.
- Hayles NK (2012) *How We Think: Digital Media and Contemporary Technogenesis*. Chicago: University of Chicago Press.
- Hayles NK (2014) Cognition everywhere: The rise of the cognitive nonconscious and the costs of consciousness. *New Literary History* 45(2).
- Hayles NK (2017) *Unthought: The Power of the Cognitive Nonconscious*. Chicago: University of Chicago Press.
- Holland JH (1975) *Adaptation in Natural and Artificial Systems*. Cambridge, MA: MIT Press.
- Inoue K, Doncescu A and Nabeshima H (2013) Completing causal networks by meta-level abduction. *Machine Learning* 91(2): 239–277.
- Josephson J and Josephson S (1996) *Abductive Inference*. Cambridge: Cambridge University Press.
- Kakas AC and Riguzzi F (2000) Abductive concept learning. *New Generation Computing* 18: 243–294.
- Lazzarato M (2012) *The Making of the Indebted Man: An Essay on the Neoliberal Condition*. Cambridge, MA: MIT Press.
- MacKenzie D (2011) How to make money in microseconds. *London Review of Books* 33(10): 16–18.
- Magnani L (2009) *Abductive Cognition: The Epistemological and Eco-Cognitive Dimensions of Hypothetical Reasoning*. Berlin: Springer-Verlag.
- 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 Perception*. Durham, NC: Duke University Press.

- Minsky M and Seymour A (1987) *Perceptrons: An Introduction to Computational Geometry*. Cambridge, MA: MIT Press.
- Mirowski P (2002) *Machine Dreams: Economics Becomes a Cyborg Science*. Cambridge: Cambridge University Press.
- O'Shea JR (2007) *Wilfrid Sellars: Naturalism with a Normative Turn*. Cambridge: Polity.
- Parisi L (2017) Computational logic and ecological rationality. In: Horl E and Burton J (eds) *On General Ecology: The New Ecological Paradigm in the Neocybernetic Age*. London: Bloomsbury.
- Peirce CS (1955) Abduction and induction. In: Buchler J (ed) *Philosophical Writings of Peirce*. New York: Dover, pp. 150–156.
- Peirce CS (1992) *The Essential Peirce*, vol. 1: *Selected Philosophical Writings (1867–1893)*, edited by Houser N and Kloesel C. Bloomington: Indiana University Press.
- Peirce CS (2005) Reasoning and the logic of things: The 1898 Cambridge Conferences. In: *Lectures by Charles Sanders Peirce*, edited by Ketner KL. Amsterdam: Harvard University Press.
- Picard RW (2000) *Affective Computing*. Cambridge, MA: MIT 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, pp. 119–140.
- Seibt J (2015) How to naturalize sensory consciousness and intentionality within a process monism with normativity gradient: A reading of Sellars. In: O'Shea J (ed) *Sellars and His Legacy*. Oxford: Oxford University Press.
- Sellars W (1963) *Science, Perception and Reality*. Atascadero, CA: Ridgeview Publishing Company.
- Steiner C (2012) *Automate This: How Algorithms Came to Rule Our World*. New York: Penguin.
- Turing AM (2001 [1936]) On computable numbers, with an application to the Entscheidungsproblem. In: *Alan M. Turing, Collected Works: Mathematical Logic*, edited by Gandy RO and Yates CEM. Amsterdam: North-Holland. (First published in: *Proceedings of the London Mathematical Society* s2–42 (1): 230–265).
- Whitehead AN (1929) *The Function of Reason*. Boston, MA: Beacon Press.
- Whitehead AN (1967) *Science and the Modern World*. New York: Free Press.
- Whitehead AN (1978) *Process and Reality: An Essay in Cosmology*. New York: Free Press.
- Wolfram S (2002) *A New Kind of Science*. Champaign, IL: Wolfram Media.

Luciana Parisi researches the philosophical consequences of technology in culture, aesthetics and politics. She is a Reader in Critical and Cultural Theory at Goldsmiths, University of London, and co-director of the Digital Culture Unit. She is the author of *Abstract Sex: Philosophy, Biotechnology and the Mutations of Desire* (Continuum Press, 2004) and *Contagious Architecture: Computation, Aesthetics and Space* (MIT

Press, 2013). She currently researching the history of automated reason and the transformation of logical thinking in machines.

This article is part of the *Theory, Culture & Society* special issue on ‘Thinking with Algorithms: Cognition and Computation in the Work of N. Katherine Hayles’, edited by Louise Amoore.