Philosophical Problems in Science

Zagadnienia Filozoficzne w Nauce

ISSN 0867-8286 (print format)

e-ISSN 2451-0602 (electronic format)

Editorial Office
Zagadnienia Filozoficzne w Nauce
Wydział Filozoficzny UPJPII
ul. Kanonicza 9, 31-002 Kraków
POLAND

e-mail: info@zfn.edu.pl
www.zfn.edu.pl

Publisher: Copernicus Center Press Sp. z o.o.
pl. Szczepański 8, 31-011 Kraków POLAND
tel. (+48) 12 448 14 12
e-mail: marketing@ccpress.pl
www.ccpress.pl

Philosophical Problems in Science

Zagadnienia Filozoficzne w Nauce

LXXII ■ 2022

Articles

Artykuły

Review articles

Artykuły recenzyjne

Articles

Artykuły

In search of a common, information-processing, agency-based framework for anthropogenic, biogenic, and abiotic cognition and intelligence

Gordana Dodig-Crnkovic
Chalmers University of Technology,
University of Gothenburg
and Mälardalen University, Sweden

Abstract

Learning from contemporary natural, formal, and social sciences, especially from current biology, as well as from humanities, particularly contemporary philosophy of nature, requires updates of our old definitions of cognition and intelligence. The result of current insights into basal cognition of single cells and evolution of multicellular cognitive systems within the framework of extended evolutionary synthesis (EES) helps us better to understand mechanisms of cognition and intelligence as they appear in nature. New understanding of information and processes of physical (morphological) computation contribute to novel possibilities that can be used to inspire the development of abiotic cognitive systems (cognitive robotics), cognitive computing and artificial intelligence.

Keywords

information, computation, cognition, intelligence, extended evolutionary synthesis, anthropogenic, biogenic and abiotic cognition.

Philosophical Problems in Science (Zagadnienia Filozoficzne w Nauce)
No 72 (2022), pp. 7–36

Information, computation, cognition, intelligence, and evolution of living organisms

The notion of information is used nowadays not only to refer to means of communication between humans, but also to denote data structures utilized for communication by other living organisms, even the simplest ones like single cells as used in the fields of bioinformatics or neuroinformatics.

In what follows we build on the ideas presented in (Dodig-Crnkovic, 2017a):

> a view of nature as a network of info-computational agents organized in a dynamical hierarchy of levels. It provides a framework for unification of currently disparate understandings of natural, formal, technical, behavioral, and social phenomena based on information as a structure, differences in one system that cause the differences in another system, and computation as its dynamics, i.e., physical process of morphological change in the informational structure.

In the current definition of computation as a dynamic of information, computation is taken to be any process of information transformation that leads to behavior, and not only those processes that we currently use to calculate, manually or with a machinery:

> Traditionally, the dynamics of computing systems, their unfolding behavior in space and time has been a mere means to the end of computing the function which specifies the algorithmic problem which the system is solving. In much of contemporary computing, the situation is reversed: the purpose of the computing system is to exhibit certain behaviour. [...] We need a theory of the dynamics of informatic processes, of interaction, and information flow, as a basis for answering

> such fundamental questions as: What is computed? What is a process? What are the analogues to Turing completeness and universality when we are concerned with processes and their behaviors, rather than the functions which they compute? (Abramsky, 2008)

Cognition can be defined as a process of "being in the world" of an agent. For living organisms, cognition is a process of life (perception, internal process control by information, actuation/agency), (Maturana, 1970; Maturana and Varela, 1980; Stewart, 1996). Cognition of an organism is based on the ability to learn from the environment and adapt so as to survive as an individual and as a species, for which organisms use information and its processing (computation).

Intelligence, as capacity for problem-solving within an environment/context, can be seen as one of the features of cognition. It is found in all living organisms as they all possess cognition, from single cells to their complex structures constituting tissues, organs, and organisms in constant interaction with each other and with the environment.

Human intelligence is the object of most of the studies of intelligence. Often it is considered to be a multidimensional phenomenon, that includes both classical problem-solving and decision-making ability (logical-mathematical reasoning), existential intelligence (ability to survive), visual-spatial, musical, bodily-kinesthetic, naturalist, linguistic, interpersonal (social), and intra-personal (ability of inner insight) intelligence. However, the question of cognition and intelligence in non-human animals and other organisms is still controversial in philosophy of mind, psychology and even in some parts of cognitive science (Ball, 2022).

Cognition and intelligence on different levels of organization of life—embodied, embedded, enacted, and extended

With the increasing insights into empirical details of processes and structures of cognition, it is emerging that human cognition and intelligence are based not only on activities of nervous system with neurons and glia cells, but equally importantly results from their interaction with non–neuronal subsystems including immune system and other somatic cells as well as the exchanges of the body with the environment. It comes as no surprise, as the nervous system is in a close interaction with the rest of the body.

Human nervous system is made up of two types of cells: primary neurons and glial cells, and it is divided into two parts: the central nervous system (brain and spinal cord) and the peripheral nervous system (autonomic and somatic nervous systems). The nervous system controls and regulates the activities of organs and systems through neuron feedback, enabling the body to respond to environmental changes (Biotechnology-Accegen, 2022). Through the embodiment, the nervous system also communicates with the external world, including other cognitive agents. The understanding that human cognition results from the activities of different types of cells, and not only nerve cells (neurons), is based on the new recognition of the existence of basal cognition/ minimal cognition / microorganismic cognition and intelligence. Unicellular organisms (single cells) have sensors and actuators and use chemical signaling and transfer of genetic information as a basis for adaptation and learning (Baluška and Levin, 2016; Ng and Bassler, 2009; Witzany, 2011; Ben-Jacob, 2003; Ben-Jacob, Shapira and Tauber, 2006). Cognitive (sensory-based) and intelligent (problem-solving) processes are regulating the state of a single cell which is a building block of multicellular living organisms (Manicka and Levin, 2019).

Thus, recently the ideas of cognition and intelligence have increased in scope (Dennett, 2017) with improved understanding of their underlying mechanisms—from the activity on the level of the human brain, to the processes on the somatic cell level. Single cell need not be a part of a human body to be seen as performing cognitive and intelligent behavior, it could be a unicellular organism or a constituent part of an animal or a plant.

At the same time as new insights have been made into the nature of biological cognition, computational and robotic cognitive systems are being developed with various degrees of cognition and intelligence. Some functions of artificial intelligence surpass human capacities (such as processing parallelism, search, memory, precision, and correctness, and often also speed) while many other capacities are far below the human level, such as common-sense reasoning, or goal-directed agency in the sense of self-preservation and self-organization.

Understanding cognition and intelligence in nature on different levels of organization, because of their fundamentally biological mechanisms is only possible if we see it in the context of evolution. As in all of biology, "nothing makes sense except for in the light of evolution" (Dobzhansky, 1973), and the cognition as a process can only be understood in the light of evolution.

However, new *abiotic approaches to cognition* assume that it is possible to construct cognitive agents from abiotic elements. Artificial (artifactual) intelligence is an attempt to produce intelligent behaviors akin to those shown by living beings (from the beginning specifically in humans) but implemented in non-living substrate. We can compare "cognitive behavior" of abiotic systems with the cognitive behavior of living organisms and see how close they are to each other.

The necessity of the Extended evolutionary synthesis (EES)

Looking at the intelligence of a living organism as a result of information processing and embodied goal-directed behaviors on hierarchy of levels of organization, suggests necessity of understanding of the process of evolution in a broader and more inclusive way than before, where biological agents are seen in their natural environments, from single cells to groups of organisms. A scientific meeting organized in partnership with the British Academy by Denis Noble, Nancy Cartwright, Patrick Bateson, John Dupré and Kevin Laland presented and discussed those important *New trends in evolutionary biology in biological, philosophical, and social science perspectives* (Royal Society, 2016).

That emerging view of evolution is called Extended Evolutionary Synthesis (EES), which is a new interpretation of the theory of evolution based on the latest scientific knowledge about life and its changes, emphasizing fundamental mechanisms of constructive development and reciprocally causal nature between an organism and its environment (Schwab, Casasa and Moczek, 2019) More on Extended Evolutionary Synthesis can be found in (Laland et al., 2015), presenting EES and its structure, assumptions, and predictions, and (Müller, 2017a,b) explaining why an extended evolutionary synthesis is necessary. Svensson (2018) argues:

> The Extended Evolutionary Synthesis (EES) will supposedly expand the scope of the Modern Synthesis (MS) and Standard Evolutionary Theory (SET), which has been characterized as gene-centered, relying primarily on natural selection and largely neglecting reciprocal causation.

Evolution is a result of interactions between natural agents, cells and their groups on variety of levels of organization (Jablonka and Lamb, 2014; Laland et al., 2015; Ginsburg and Jablonka, 2019), as Jablonka and Lamb argue in their book *Evolution in Four Dimensions*: *Genetic, Epigenetic, Behavioral, and Symbolic Variation in the History of Life*. These dimensions can be found on different levels of organization of life.

In short, if we want to bring evolutionary theory in coherence with the advancement in other sciences, extended evolutionary synthesis is necessary.

Info-computational lens: agent-based natural information and computation

We use an info-computational lens to approach phenomena of cognition and intelligence. A framework of (Dodig-Crnkovic, 2017c) enables understanding of cognitive systems generated through self-structuring processes of morphological info-computation on the hierarchy of levels in nature from physics, to chemistry and biology, based on agent-centric embodied information and morphological computation. It means that we assume:

- computing nature paradigm, where nature is seen through the lens of information and computation as its dynamics, that is

providing a basis for unification of currently disparate understanding of natural, formal, technical, social and behavioral phenomena;
- an observer-dependent, agent-based reality, that is reality for an agent for which cognition is a result of relational info-computational processes;
- computational interpretation of information dynamics in nature, where computation is physical (morphological) computation;

that enables us to:

- avoid frequent misunderstandings of the inadequate abstract models of computation (as in old computationalism) and focus on embodied morphological computation in physical systems, especially cognitive ones such as living beings;
- suggest the necessity of generalization of the models of computation beyond the traditional Turing machine model and acceptance of "second generation" models of computation capable of covering the whole range of phenomena from physics to cognition (Abramsky);
- understand goal directed behaviors and complexification in living systems through the extended evolutionary synthesis.

The developments supporting info-computational approach, as a variety of naturalism, are found in among others complexity theory, systems theory, theory of computation (natural computing, organic computing, unconventional computing), cognitive science, neuroscience, information physics, agent based models of social systems and information sciences, robotics (especially developmental robotics), bioinformatics and artificial life (Dodig-Crnkovic and Müller, 2011;

Dodig-Crnkovic, 2017c), as well as biosemiotics (Sarosiek, 2021) and Polak-Krzanowski's deanthropomorphized pancomputationalism (Polak and Krzanowski, 2019).

Cognition of a living organism is thus studied as a network of networks of distributed information processing units on variety of levels of organization, from single cells to the whole body including the level of groups of organisms manifest as social cognition.

Natural cognition based on cells processing (computing) information—basal cognition in an extended evolutionary perspective

Despite decades of research into the subject, there is still no agreement about where cognition is found in the living world (Ball, 2022). Is a nervous system needed? If so, why? If not, why not? A new two-part theme issue of *Phil Trans B* on the emerging field of 'Basal Cognition', edited by Pamela Lyon, Fred Keijzer, Detlev Arendt and Michael Levin, explores these questions (Levin et al., 2021; Lyon et al., 2021).

Present increase of knowledge about cellular cognition and new gained details of complex goal-directed behaviors is nicely illustrated by the example of a single-celled predator organism *Lacrymaria olor* ("tears of a swan") hunting down another cell, often used by Michael Levin. *Lacrymaria* has a "neck" a "body" and a "mouth". It beats the hair-like cilia around its "head" and extends its neck up to 8 times its body length, while chasing and finally swallowing another cell. It has no nervous system and no sensors that macroscopic living organisms typically use to chase their prey. How does it manage to identify, follow, and catch the prey? The mechanisms that enable *Lacrymaria* to hunt down another cell, that goal-directedly activate cilia, take care of timing of "mouth" opening and closing are studied in (Weiss, 2020). Likewise, (Coyle et al., 2019) describe how coupled active systems encode an emergent hunting behavior in *Lacrymaria olor*. Even the work of (Mearns et al., 2020) analyzes its hunting behavior, revealing a tightly coupled stimulus-response loop. Furthermore (Wlotzka and McCaskill, 1997) argue that in this case, they observed behavior of a molecular predator and its prey, through coupled isothermal amplification of nucleic acids. In short, research shows that a goal-directed behavior of *Lacrymaria olor*, is a result of a coupled stimulus-response loops. However, importantly, we do not know the meta-level mechanism which activates those loops and makes them goal directed.

Another microorganism under intense study for their goal-directed, efficient learning and adaptive behavior, which are of special interest because of their ability to cause diseases in other organisms, are bacteria. Eshel Ben Jacob have been studying bacterial colonies, their self-organization, complexification and adaptation, smartness, communication and linguistic communication (by chemical languages), social intelligence, natural information processing, and foundations of bacterial cognition (Ben-Jacob, 1998; 2003; 2008; 2009; Ben-Jacob, Becker and Shapira, 2004; Ben-Jacob, Shapira and Tauber, 2006; 2011). Works (Witzany, 2011; Schauder and Bassler, 2001; Waters and Bassler, 2005; Ng and Bassler, 2009) focus on communication (information exchange) mechanisms in bacteria, and especially *quorum sensing*, where group of bacteria make a majority-based decisions. Bacteria colonies and films display various multicellular behaviors, emitting, receiving, and processing a large vocabulary of chemical symbols.

More about experimental methods for study of cell cognition can be found in the work of *The cell cognition project* (Held et al., 2010).

From all the above evidence it is clear that unicellular organisms exhibit basal cognition and intelligence (problem-solving capacities). A fundamental observation connecting this rudimentary biotic cognition and more complex anthropogenic (i.e., human-level, brain-based) cognition, is the following:

> Cognitive operations we usually ascribe to brains—sensing, information processing, memory, valence, decision making, learning, anticipation, problem solving, generalization and goal directedness—are all observed in living forms that don't have brains or even neurons (Lyon et al., 2021).

Similar arguments for biogenic nature of cognition have been presented by (Levin et al., 2021; Yuste and Levin, 2021; Lyon et al., 2021).

Our approach to information-processing mechanisms of cognition, unlike vast majority of artificial cognitive architectures targeting human-level cognition, focus on the development and evolution of the continuum of natural cognitive architectures, from basal cellular architecture up, as studied by (Levin et al., 2021) and already identified by (Sloman, 1984).

The connection between high-level and basal cognition is visible in the role of ion channels and neurotransmitters, studied in nervous cells, but also present in ordinary somatic cells:

> We have previously argued that the deep evolutionary conservation of ion channel and neurotransmitter mechanisms highlights a fundamental isomorphism between developmental and behavioral processes. Consistent with this, membrane excitability has been suggested to be the ancestral basis for

> psychology [...]. Thus, it is likely that the cognitive capacities of advanced brains lie on a continuum with, and evolve from, much simpler computational processes that are widely conserved at both the functional and mechanism (molecular) levels.
>
> The information processing and spatio–temporal integration needed to construct and regenerate complex bodies arises from the capabilities of single cells, which evolution exapted and scaled up as behavioral repertoires of complex nervous systems that underlie familiar examples of Selves (Fields and Levin, 2019).

This biogenic nature of cognition makes it necessary to recognize all living forms, and not only those with nervous systems (Piccinini, 2020), or what is even more frequent only humans, as cognitive systems.

As for the driving mechanisms behind this complexification process in living/cognitive systems, (Fields, Friston et al., 2022, pp.1–2) describe how The Free Energy Principle of Karl Friston can drive neuromorphic development in the fully-general quantum-computational framework of topological quantum neural networks:

> We show how any system with morphological degrees of freedom and locally limited free energy will, under the constraints of the free energy principle, evolve toward a neuromorphic morphology that supports hierarchical computations in which each "level" of the hierarchy enacts a coarse-graining of its inputs, and dually a fine-graining of its outputs. Such hierarchies occur throughout biology, from the architectures of intracellular signal transduction pathways to the large-scale organization of perception and action cycles in the mammalian brain.

Biogenic approach is useful not only for understanding of cognition and intelligence and their evolution in living nature, but also for engineering of artificial systems that need certain level of intelligence, not necessarily the human level, such as nano-bots (Kriegman et al., 2021) or different elements of IoT (Internet of Things).

Cognitive computing and AI—still anthropogenic

Inspired by the behaviors produced by anthropogenic cognition, (Modha et al., 2011) the field of cognitive computing is exploring biomimetic approaches to cognition in abiotic systems (Gudivada et al., 2019) studying cognitive computing systems, their potential and possible futures. In the application domain, e.g. IBM had a cognitive computing project called Systems of Neuromorphic Adaptive Plastic Scalable Electronics (SyNAPSE) (Srinivasa and Cruz-Albrecht, 2012).

The quest for intelligent machines ultimately requires new breakthroughs in computer architecture, theory of computation, computational neuroscience, supercomputing, cognitive science, and related fields orchestrated in a coherent, unified effort.

Cognitive computing, AI and cognitive robotics present attempts to construct abiotic systems exhibiting cognitive characteristics of biotic systems. As a rule, they assume human-level intelligence and human-level cognition, even though biogenic approaches would bring huge benefits. When we acknowledge that cognition in living nature comes in degrees, it is more meaningful to talk about cognition of artifacts, even though the role of cognitive capacities for an artefact is not to assure its continuing existence (unlike in cognition = life

(Stewart, 1996), which gives the evolutionary role to cognition in biotic systems). The difference is that cognitive artifacts are constructed to pursue human goals, not their own intrinsic ones.

Cognition at different levels of organization of a living organism—from cells up

Traditional anthropogenic approach to cognition (Markram, 2012) is looking at cognition and intelligence in humans as the only natural cognitive agents.

Biogenic approaches on the other hand broaden the domain, seeing cognition as an ability of all living organisms (Maturana, 1970; Maturana and Varela, 1980; Stewart, 1996)

More specifically, Maturana and Varela argue:

> A cognitive system is a system whose organization defines a domain of interactions in which it can act with relevance to the maintenance of itself, and the process of cognition is the actual (inductive) acting or behaving in this domain. *Living systems are cognitive systems and living as a process is a process of cognition.* This statement is valid for all organisms, with and without a nervous system (Maturana and Varela, 1980, p.13; cf. Maturana and Varela, 1992).

Cognition is thus a capacity possessed in different forms and degrees of complexity by every living organism. It is entirety of processes going on in an organism that keeps it alive, and present as a distinct agent in the world. A single cell while alive constantly cognizes, that is registers inputs from the world and its own body, ensures its own continuous existence through metabolism and food hunting while avoiding dangers that could cause its disintegration or damage, at the same time adapting its own morphology to the environmental constraints. The entirety of physico-chemical processes depends on the morphology of the organism, where morphology is meant as the form and structure. Work of Marijuán, Navarro and del Moral (2010) presents a study of prokaryotic intelligence and its strategies for sensing the environment.

Multicellularity

Unicellular organisms such as bacteria communicate and build swarms or films with far more advanced capabilities compared to individual organisms, through social (distributed) cognition.

In general, groups of smaller organisms (cells) in nature cluster into bigger ones (multicellular assemblies) with differentiated control mechanisms from the cell level to the tissue, organ, organism and groups of organisms, and this layered organization provides information processing benefits.

Examining the origin of multicellularity (Fields and Levin, 2019) investigates the computational boundary of a "self" and argues that it is bioelectricity that drives multicellularity and scale-free cognition. According to (Fields and Levin, 2019), somatic multicellularity presents a satisficing solution to the prediction-error minimization problem for single cells. From the point of view of information, (Colizzi, Vroomans and Merks, 2020) argue that evolution of multicellularity results from a collective integration of spatial information, while (McMillen, Walker and Levin, 2022) show how to use Shannon information theory (Shannon, 1948) as a tool for integration of biophysical signaling modules. By mapping information flow between cells and pathways, researchers show that information theory supports systems-level view of biological phenomena where molecular reductionism does not work well.

Computationalism is not what it used to be …

… that is the thesis that human cognition and intelligence are Turing machines (Scheutz, 2002). Unlike classical computationalism based on symbol manipulation and Turing Machine model, modern computationalism for modelling of cognitive processes requires new models of computation.

Turing Machine is an abstract logical model of computation equivalent to an algorithm, and it may be used for description of elementary sequential processes in living organisms. However, complex networked physical processes with temporal and other physical resource constraints cannot be adequately modelled as series of sequential logical operations (Turing machines). As Leslie Valiant (2013) succinctly puts it:

> We need computational models for the basic characteristics of life, such as the ability to differentiate and synthesize information, make a choice, adapt, evolve, and learn in an unpredictable world. That requires computational mechanisms and models which are not "certainly, exactly correct" and predefined as Turing machine, but, instead, "probably approximately correct" (PAC).

Computational approaches that are capable of modelling adaptation, evolution and learning are found in the field of natural computation and computing nature (Dodig-Crnkovic, 2014).

Computing, the fourth scientific domain

Info-computational approach incorporates our best current scientific knowledge about the processes in nature, translating them into language of natural information and computation.

The aim of this approach to cognition is to increase understanding of cognitive capacities in diverse types of agents, biological and synthetic, including their ability of learning, and learning to learn (meta-learning) (Dodig-Crnkovic, 2020) as well as their communication and mutual interactions. According to (Denning, 2007), computing can be seen as a natural science. Even more than that, we are witnessing the emergence of a new computing science (Denning, 2010) , connecting natural and formal sciences, adding the dimension of real time and physical constraints to logic and mathematics. As Rosenbloom argues, "Computing may be the fourth great domain of science along with the physical, life and social sciences" (Rosenbloom, 2015). In that new broader, emerging computing science, the Turing Model of computation is a proper subset.

Computing nature and nature inspired computation. Self-generating systems

Complex biological systems must be modeled as self-referential, self-organizing "component-systems" (Kampis, 1991) which are self-generating and whose behavior, though computational in a general sense, goes far beyond Turing machine model. Georg Kampis studied the behavior of self-modifying systems in biology and cognitive science as a basis for a new framework for dynamics, information, computation, and complexity:

> a component system is a computer which, when executing its operations (software) builds a new hardware. [... W]e have a computer that re-wires itself in a hardware-software interplay: the hardware defines the software, and the software defines new hardware. Then the circle starts again (Kampis, 1991, p.223).

Similar position is presented in (Dodig-Crnkovic and Müller, 2011) connecting models of computation from the formal sequential logical machine Turing model to the physical (morphological) concurrent natural computation.

Evolution as generative mechanism for increasingly complex cognitive systems

New insights about cognition and its evolution and development in nature, from cellular to human cognition can be modelled as natural information processing/ natural computation/ morphological computation. In the info-computational approach, evolution in the sense of Extended evolutionary synthesis is a result of interactions between natural agents, cells, and their groups.

Evolution provides generative mechanisms for the emergence of more and more competent living organisms, with increasingly complex natural cognition and intelligence, and those mechanisms can be used as a template for the design and construction of their artifactual, computational counterparts.

Learning from biogenic computing

The concept of biological computation posits that living organisms process information and thus perform computations, and that ideas of information and computation are the key to understanding, modeling, simulation, and control of biological systems. See (Mitchell, 2012) for the exposition of the concept of biological computation, and (Dodig-Crnkovic, 2022) for presentation of cognitive architectures based on natural infocomputation. Cognition as a result of information processing in living agent's morphology, with species-specific cognition and intelligence is described in (Dodig-Crnkovic, 2021).

One of important characteristics of natural computing is its computational efficiency which is becoming increasingly important in the world with pervasive computing and concurrent global warming. The Turing Machine model of computation is not resource-aware, unlike living systems which are constantly optimizing their use of natural resources. Therefore, in the biomimetic approach to cognitive architectures designers are learning from nature how to compute more resource efficiently. Mutual learning between computing, cognitive sciences and neurosciences (Rozenberg and Kari, 2008) leads to improved understanding of how cognition works and develops in nature, and how we can simulate, emulate, and engineer abiotic cognition and intelligence with the properties close to the biotic one.

Morphological computation connecting body, brain, and environment in robotics

The research performed in the diverse fields of soft robotics / self-assembly systems and molecular robotics / self-assembly systems at all scales / embodied robotics / reservoir computing / real neural systems / systems medicine / functional architecture / organization / process management / computation based on spatio-temporal dynamics/ information theoretical approach to embodiment mechatronics / amorphous computing / molecular computing – all connect body, control ("brain") and environment.

In robotics, a brain and body that researchers learn from, sometimes belongs to an octopus, which unlike typical robots has soft body that presents substantially different possibilities from rigid bodies of conventional robots.

Pfeifer and Bongard (2006) were among the first to present a new view of embodied intelligence, arguing that the body shapes the way we think, looking in the first place from the anthropocentric perspective, but the approach applies equally well to biocentric view of cognition. In biologically inspired robotics, embodiment and self-organization are driving forces of evolving intelligence (Pfeifer, Lungarella and Iida, 2007). They are best understood in terms of morphological computation (Pfeifer and Iida, 2005; Hauser, Füchslin and Pfeifer, 2014).

The essential property of morphological computation is that it is defined on a structure of nodes (agents) that exchange (communicate) information. It is thus applied not only in robotics, but generalized to other physical information-processing systems, including living beings (Dodig-Crnkovic, 2013; 2017b; 2018).

Computing Nature and Natural Computation

In his article "Epistemology as Information Theory", Greg Chaitin argues that knowledge should be studied as a result of information processes, thus turning epistemology into study of information:

> And how about the entire universe, can it be considered to be a computer? Yes, it certainly can, it is constantly computing its future state from its current state, it's constantly computing its own time-evolution! And as I believe Tom Toffoli pointed out, actual computers like your PC just hitch a ride on this universal computation! (Chaitin, 2007, p.13)

David Deutsch in his article "What is Computation? (How) Does Nature Compute?" contributes with the similar position in the book "A Computable Universe" by Hector Zenil (2012).

Starting from the above ideas, (Dodig-Crnkovic, 2007) proposes that epistemology can be naturalized through the info-computational approach to knowledge generation. The computing nature framework (naturalist computationalism) makes it possible to describe all cognizing agents (living organisms and artificial cognitive systems) as informational structures with computational dynamics (Dodig-Crnkovic and Müller, 2011; Dodig-Crnkovic and Giovagnoli, 2013; Dodig-Crnkovic, 2014; 2017a; Dodig-Crnkovic and Giovagnoli, 2013; Dodig-Crnkovic, 2017a). Morphological computation in this framework is a process of creation of new informational structures, as it appears in nature, living as non-living. It is a process of morphogenesis, which in biological systems is driven by development and evolution (Dodig-Crnkovic, 2013; 2017b; 2018).

It is worth noting that research on "computing nature" focuses on how physical/ natural/ morphological processes can be interpreted as computation and used to compute, while research on "computable universe" asks the question if we can compute (with our current theories of computing) what we observe as the universe—two different research programs.

Conclusions

New insights from complexity theory, systems theory, theory of computation (natural computing, organic computing, unconventional computing), cognitive science, neuroscience, information physics, agent based models of social systems and information sciences, robotics, as well as bioinformatics and artificial life call for updates in our understanding of cognition and intelligence (Dodig-Crnkovic and Müller, 2011; Dodig-Crnkovic, 2017c).

Traditionally, in the fields of cognitive science, philosophy of mind, cognitive computing and artificial intelligence, cognition and intelligence are assumed to be the abilities of humans. They are described in terms of concepts such as mind, thought, reasoning, logic, etc. However, new understanding of the goal-directed, learning, and adaptive behaviors of all living organisms, from all five kingdoms of life—animal, plant, fungi, protist and monera, from single celled to multi-cellular organisms and their ecologies, all possess level of cognition and intelligence which increases with the complexity of the system.

In this article we present a common framework of info-computation, where computation is physical/morphological computation providing unified approach to anthropogenic, biogenic, and abiotic cognition. The advantage of info-computational approach is that it enables learning of mechanisms of those three types of cognition and intelligence. It also connects different levels of organization as observed in nature.

Cognition and intelligence, coming from the simplest to the most complex in a continuum of natural systems can be source of inspiration for the design and construction of artificial cognitive systems with varying degrees/levels of intelligence, from nano-bots to autonomous cars and android robots.

Bibliography

Abramsky, S., 2008. Information, Processes and Games. In: J. Benthem van and P. Adriaans, eds. *Philosophy of Information*. Amsterdam, The Netherlands: North Holland, pp.483–549.

Ball, P., 2022. *The Book of Minds: How to Understand Ourselves and Other Beings, from Animals to AI to Aliens*. Chicago: University of Chicago Press.

Baluška, F. and Levin, M., 2016. On having no head: cognition throughout biological systems. *Frontiers in Psychology*, 7, p.902.

Ben-Jacob, E., 1998. Bacterial wisdom, Gödels theorem and creative genomic webs. *Physica A*, 248, pp.57–76.

Ben-Jacob, E., 2003. Bacterial Self-Organization: Co-Enhancement of Complexification and Adaptability in a Dynamic Environment. *Phil. Trans. R. Soc. Lond. A*, pp.315–322.

Ben-Jacob, E., 2008. Social behavior of bacteria: from physics to complex organization. *The European Physical Journal B*, 65(3), pp.315–322.

Ben-Jacob, E., 2009. Learning from Bacteria about Natural Information Processing. *Annals of the New York Academy of Sciences*, 1178, pp.78–90.

Ben-Jacob, E., Becker, I. and Shapira, Y., 2004. Bacteria Linguistic Communication and Social Intelligence. *Trends in Microbiology*, 12(8), pp.366–372.

Ben-Jacob, E., Shapira, Y. and Tauber, A.I., 2011. Smart Bacteria. In: L. Margulis, C. Asikainen and W. Krumbein, eds. *Chimera and Consciousness. Evolution of the Sensory Self*. Cambridge Boston: MIT Press.

Ben-Jacob, E., Shapira, Y. and Tauber, A., 2006. Seeking the Foundations of Cognition in Bacteria. *Physica A*, 359, pp.495–524.

Biotechnology-Accegen, 2022. *Nervous System Primary Cells* [Online]. Available at: <https://www.accegen.com/category/nervous-system-primary-cells/> [visited on 10 January 2023].

Chaitin, G., 2007. Epistemology as Information Theory: From Leibniz to ω. In: G. Dodig Crnkovic, ed. *Computation, Information, Cognition – The Nexus and The Liminal*. Newcastle UK: Cambridge Scholars Pub., pp.2–17.

Colizzi, E.S., Vroomans, R.M. and Merks, R.M., 2020. Evolution of multicellularity by collective integration of spatial information. *eLife* [Online], 9:e56349. https://doi.org/10.7554/eLife.56349.

Coyle, S.M. et al., 2019. Coupled Active Systems Encode an Emergent Hunting Behavior in the Unicellular Predator Lacrymaria olor. *Current Biology* [Online], 29, pp.3838–3850. https://doi.org/10.1016/j.cub.2019.09.034.

Dennett, D., 2017. *From Bacteria to Bach and Back: The Evolution of Minds*. New York: Norton & Company.

Denning, P., 2007. Computing is a natural science. *Communications of the ACM*, 50(7), pp.13–18.

Denning, P., 2010. Computing Science: The Great Principles of Computing. *American Scientist*, 98(5), pp.369–372.

Dobzhansky, T., 1973. Nothing in Biology Makes Sense Except in the Light of Evolution. *American Biology Teacher* [Online], 35(3), pp.125–129. https://doi.org/10.2307/4444260.

Dodig-Crnkovic, G., 2007. Epistemology Naturalized: The Info-Computationalist Approach. *APA Newsletter on Philosophy and Computers*, 06(2), pp.9–13.

Dodig-Crnkovic, G., 2013. The Development of Models of Computation with Advances in Technology and Natural Sciences. *6th AISB Symposium on Computing and Philosophy: The Scandal of Computation - What is Computation? - AISB Convention 2013*.

Dodig-Crnkovic, G., 2014. Modeling Life as Cognitive Info-Computation. In: A. Beckmann, E. Csuhaj-Varjú and K. Meer, eds. *Computability in Europe 2014. LNCS*. Berlin Heidelberg: Springer, pp.153–162.

Dodig-Crnkovic, G., 2017a. Computational Dynamics of Natural Information Morphology, Discretely Continuous. *Philosophies* [Online], 2(4), p.23. https://doi.org/10.3390/philosophies2040023.

Dodig-Crnkovic, G., 2017b. *Morphologically Computing Embodied, Embedded, Enactive, Extended Cognition* [Online]. PT-AI 2017. Available at: <https://www.pt-ai.org/2017/posters>.

Dodig-Crnkovic, G., 2017c. Nature as a network of morphological infocomputational processes for cognitive agents. *European Physical Journal: Special Topics* [Online], 226(2). https://doi.org/10.1140/epjst/e2016-60362-9.

Dodig-Crnkovic, G., 2018. Cognition as Embodied Morphological Computation. In: V.C. Müller, ed. *Philosophy and Theory of Artificial Intelligence 2017* [Online], Studies in Applied Philosophy, Epistemology and Rational Ethics. Cham: Springer International Publishing, pp.19–23. https://doi.org/10.1007/978-3-319-96448-5_2.

Dodig-Crnkovic, G., 2020. Natural Morphological Computation as Foundation of Learning to Learn in Humans, Other Living Organisms, and Intelligent Machines. *Philosophies* [Online], 5(3), pp.17–32. https://doi.org/10.3390/philosophies5030017.

Dodig-Crnkovic, G., 2021. Cognition as a Result of Information Processing in Living Agent's Morphology. Species-specific Cognition and Intelligence. *Proceedings of 16th SweCog Conference* [Online], pp.22–25. Available at: <https://www.diva-portal.org/smash/get/diva2:1611781/FULLTEXT01.pdf> [visited on 10 January 2023].

Dodig-Crnkovic, G., 2022. Cognitive architectures based on natural info-computation. In: V.C. Müller, ed. *Philosophy and theory of artificial intelligence 2021*. SAPERE. Berlin, Heidelberg: Springer.

Dodig-Crnkovic, G. and Giovagnoli, R., 2013. Computing nature – A network of networks of concurrent information processes. *Computing Nature* [Online]. Vol. 7. Berlin, Heidelberg: Springer. https://doi.org/10.1007/978-3-642-37225-4_1.

Dodig-Crnkovic, G. and Müller, V.C., 2011. A Dialogue Concerning Two World Systems: Info-Computational vs. Mechanistic. In: G. Dodig Crnkovic and M. Burgin, eds. *Information and Computation*. Singapore: World Scientific Pub Co Inc, pp.149–184.

Fields, C., Friston, K. et al., 2022. *The Free Energy Principle drives neuromorphic development* [Online]. arXiv. Available at: <http://arxiv.org/abs/2207.09734> [visited on 10 January 2023].

Fields, C. and Levin, M., 2019. Somatic multicellularity as a satisficing solution to the prediction-error minimization problem. *Communicative & Integrative Biology* [Online], 12(1), pp.119–132. https://doi.org/10.1080/19420889.2019.1643666.

Ginsburg, S. and Jablonka, E., 2019. *The Evolution of the Sensitive Soul* [Online]. Cambridge, MA: MIT Press. https://doi.org/10.7551/mitpress/11006.001.0001.

Gudivada, V., Pankanti, S., Seetharaman, G. and Zhang, Y., 2019. "Cognitive Computing Systems: Their Potential and the Future". *Computer* [Online], 52(05), pp.13–18. https://doi.org/10.1109/MC.2019.2904940.

Hauser, H., Füchslin, R. and Pfeifer, R., 2014. *Opinions and Outlooks on Morphological Computation*. e-book.

Held, M. et al., 2010. CellCognition: time-resolved phenotype annotation in high-throughput live cell imaging. *Nature Methods*, 7(9), pp.747–754.

Jablonka, E. and Lamb, M., 2014. *Evolution in Four Dimensions: Genetic, Epigenetic, Behavioral, and Symbolic Variation in the History of Life. Revised Edition. Life and Mind: Philosophical Issues in Biology and Psychology*. Cambridge, MA, USA: A Bradford Book. MIT Press.

Kampis, G., 1991. *Self-Modifying Systems in Biology and Cognitive Science: A New Framework for Dynamics, Information, and Complexity*. Amsterdam: Pergamon Press.

Kriegman, S., Blackiston, D., Levin, M. and Bongard, J., 2021. Kinematic self-replication in reconfigurable organisms. *Proceedings of the National Academy of Sciences* [Online], 118(49), e2112672118. https://doi.org/10.1073/pnas.2112672118.

Laland, K.N. et al., 2015. The extended evolutionary synthesis: its structure, assumptions and predictions. *Proceedings of the Royal Society B: Biological Sciences* [Online], 282(1813), p.20151019. https://doi.org/10.1098/rspb.2015.1019.

Levin, M., Keijzer, F., Lyon, P. and Arendt, D., 2021. Basal cognition: multicellularity, neurons and the cognitive lens, Special issue, Part 2. *Phil. Trans. R. Soc. B*, 376(20200458).

Lyon, P., Keijzer, F., Arendt, D. and Levin, M., 2021. Basal cognition: conceptual tools and the view from the single cell - Special issue, Part 1. *Phil. Trans. R. Soc. B*, 376(20190750).

Manicka, S. and Levin, M., 2019. The Cognitive Lens: a primer on conceptual tools for analysing information processing in developmental and regenerative morphogenesis. *Philosophical Transactions of the Royal Society B*, 374(1774).

Marijuán, P.C., Navarro, J. and del Moral, R., 2010. On prokaryotic intelligence: Strategies for sensing the environment. *BioSystems* [Online], 99(2), pp.94–103. https://doi.org/10.1016/j.biosystems.2009.09.004.

Markram, H., 2012. The Human Brain Project. *Scientific American*, 306(6), pp.50–55.

Maturana, H., 1970. *Biology of Cognition*. (Biological Computer Laboratory Research Report BCL 9). Urbana, IL: University of Illinois.

Maturana, H. and Varela, F., 1980. *Autopoiesis and Cognition: The Realization of the Living*. Dordrecht Holland: D. Reidel Pub. Co.

Maturana, H.R. and Varela, F.J., 1992. *The Tree of Knowledge: The Biological Roots of Human Understanding*. Rev. Boston, MA: Shambala.

McMillen, P., Walker, S.I. and Levin, M., 2022. Information Theory as an Experimental Tool for Integrating Disparate Biophysical Signaling Modules. *Int. J. Mol. Sci.* [Online], 23(9580). https://doi.org/10.3390/ijms23179580.

Mearns, D.S. et al., 2020. Deconstructing Hunting Behavior Reveals a Tightly Coupled Stimulus-Response Loop. *Current Biology* [Online], 30(1), 54–69.e1–e9. https://doi.org/10.1016/j.cub.2019.11.022.

Mitchell, M., 2012. Biological computation. *Computer Journal*, 55(7), pp.852–855.

Modha, D.S. et al., 2011. Cognitive Computing. *Communications of the ACM*, 54(8), pp.62–71.

Müller, G.B., 2017a. Correction to 'Why an extended evolutionary synthesis is necessary'. *Interface Focus*, 7(20170065).

Müller, G.B., 2017b. Why an extended evolutionary synthesis is necessary. *Interface Focus*, 7(2017001520170015).

Ng, W.-L. and Bassler, B., 2009. Bacterial quorum-sensing network architectures. *Annual Review of Genetics*, 43, pp.197–222.

Pfeifer, R. and Bongard, J., 2006. *How the Body Shapes the Way We Think – a New View of Intelligence*. MIT Press.

Pfeifer, R. and Iida, F., 2005. Morphological computation: Connecting body, brain and environment. *Japanese Scientific Monthly*, 58(2), pp.48–54.

Pfeifer, R., Lungarella, M. and Iida, F., 2007. Self-organization, embodiment, and biologically inspired robotics. *Science*, 318, pp.1088–1093.

Piccinini, G., 2020. *Neurocognitive Mechanisms: Explaining Biological Cognition*. Oxford: Oxford scholarship online.

Polak, P. and Krzanowski, R., 2019. Deanthropomorphized Pancomputationalism and the Concept of Computing. *Foundations of Computing and Decision Sciences* [Online], 44(1), pp.45–54. https://doi.org/10.2478/fcds-2019-0004.

Rosenbloom, P., 2015. *On Computing: The Fourth Great Scientific Domain*. Cambridge, MA: The MIT Press.

Royal Society, 2016. *New trends in evolutionary biology: biological, philosophical and social science perspectives* [Online]. Available at: <https://royalsociety.org/science-events-and-lectures/2016/11/evolutionary-biology/>.

Rozenberg, G. and Kari, L., 2008. The many facets of natural computing. *Communications of the ACM* [Online], 51, pp.72–83. https://doi.org/10.1145/1400181.1400200.

Sarosiek, A., 2021. The role of biosemiosis and semiotic scaffolding in the processes of developing intelligent behaviour. *Philosophical Problems in Science (Zagadnienia Filozoficzne w Nauce)* [Online], (70), pp.9–44. Available at: <https://zfn.edu.pl/index.php/zfn/article/view/535>.

Schauder, S. and Bassler, B., 2001. The languages of bacteria. *Genes & Dev.*, 15, pp.1468–1480.

Scheutz, M., 2002. *Computationalism New Directions*. Cambridge, MA: MIT Press.

Schwab, D.B., Casasa, S. and Moczek, A., 2019. On the Reciprocally Causal and Constructive Nature of Developmental Plasticity and Robustness. *Frontiers in Genetics* [Online], 9(735). https://doi.org/10.3389/fgene.2018.00735.

Shannon, C.E., 1948. A mathematical theory of communication. *The Bell System Technical Journal* [Online], 27(July 1928), pp.379–423. https://doi.org/10.1145/584091.584093.

Sloman, A., 1984. The structure of the space of possible minds. In: S. Torrance, ed. *The Mind and the Machine: philosophical aspects of Artificial Intelligence*. Chichester: Ellis Horwood, pp.35–42.

Srinivasa, N. and Cruz-Albrecht, J., 2012. Neuromorphic adaptive plastic scalable electronics: analog learning systems. *IEEE Pulse* [Online], 3(1), pp.51–56. https://doi.org/10.1109/MPUL.2011.2175639.

Stewart, J., 1996. Cognition = life: Implications for higher-level cognition. *Behavioral Processes*, 35, 311–326.

Svensson, E.I., 2018. On Reciprocal Causation in the Evolutionary Process. *Evol Biol*, 45, pp.1–14.

Valiant, L., 2013. *Probably Approximately Correct: Nature's Algorithms for Learning and Prospering in a Complex World*. New York, NY, USA: Basic Books.

Waters, C.M. and Bassler, B., 2005. Quorum Sensing: Cell-to-Cell Communication in Bacteria. *Annual Review of Cell and Developmental Biology*, 21, pp.319–346.

Weiss, J., 2020. *Single-celled Lacrymaria olor Hunts Down Another Cell* [Online]. Available at: <https://www.youtube.com/watch?v=sq6Y54mxjOg>.

Witzany, G., 2011. Introduction: Key Levels of Biocommunication of Bacteria. In: G. Witzany, ed. *Biocommunication in Soil Microorganisms* [Online]. Vol. 23. Berlin, Heidelberg: Springer, pp.1–34. https://doi.org/10.1007/978-3-642-14512-4_1.

Wlotzka, B. and McCaskill, J.S., 1997. A molecular predator and its prey: coupled isothermal amplification of nucleic acids. *Cell Chemical Biology*, 4(1), pp.25–33.

Yuste, R. and Levin, M., 2021. New Clues about the Origins of Biological Intelligence. A common solution is emerging in two different fields: developmental biology and neuroscience. *Scientific American* [Online]. Available at: <https://www.scientificamerican.com/article/new-clues-about-the-origins-of-biological-intelligence/> [visited on 10 January 2023].

Zenil, H., 2012. *A Computable Universe: Understanding Computation & Exploring Nature As Computation*. Ed. by H. Zenil. Singapore: World Scientific Publishing Company/Imperial College Press.

But seriously: what do algorithms want? Implying collective intentionalities in algorithmic relays—a distributed cognition approach

Javier Toscano
Center for Advanced Internet Studies

Abstract

Describing an algorithm can provide a formalization of a specific process. However, different ways of conceptualizing algorithms foreground certain issues while obscuring others. This article attempts to define an algorithm in a broad sense as a cultural activity of key importance to make sense of socio-cognitive structures. It also attempts to develop a sharper account on the interaction between humans and tools, symbols and technologies. Rather than human or machine-centered analyses, I draw upon sociological and anthropological theories that underline social practices to propose expanding our understanding of an algorithm through the notion of 'collective intentionalities'. To make this term clear, a brief historical review is presented, followed by an argumentation on how to incorporate it in an integral perspective. The article responds to recent debates in critical algorithm studies about the significance of the term. It develops a discussion along the lines of cognitive anthropology and the cognitive sciences, therefore advancing a definition that is grounded in observed practices as well as in modeled descriptions. The benefit of this approach is that it encourages scholars to explore cognitive

Philosophical Problems in Science (Zagadnienia Filozoficzne w Nauce)
No 72 (2022), pp. 37–65

structures via archaeologies of technological assemblages, where intentionalities play a defining role in understanding socio-structured practices and cognitive ecologies.

Keywords
algorithm studies, distributed cognition, collective intentionalities, socio-computing infrastructures, cognitive anthropology.

Initial Definitional Attempts

Oversimplified definitions of an algorithm are currently available and frequently used, but an algorithm is neither a recipe nor a rigidly constrained and procedural formulation. Limited conceptions that represent them as a sort of entity or thing, a series of steps that need to be applied, or a simple technique that homogenizes a process, lead to weak understandings of the deeper processes, transactions and dynamics that are at stake. Indeed, an algorithm can be a problem-solving device, and this feature in itself can become a point of entry to a more complex analysis. After all, for engineers and computer scientists, "an algorithm is an abstract, formalized description of a computational procedure" (Dourish, 2016, p.3). But even if sleek and apparently elegant, the problem with this definitional reduction is twofold.

On the one hand, it concentrates on processes that happen inside computational machines. This makes the description not only machine-centered, but also introduces a misunderstanding. After all, articulating a notion of code in the early days of computing history, pioneer logicians Newell, Simon and Shaw wrote in a seminal paper that

> the appropriate way to describe a piece of problem solving behavior is in terms of a program [...]. Computers come into the picture only because they can, by appropriate programming, be induced to execute the same sequences of information processes that humans execute when they are solving problems (Newell, Simon and Shaw, 1958, p.151).

In this sense, as programs, algorithms need not be thought of merely as machine drivers. And as we will see, getting rid of this idiosyncratic constraint would allow us to spot algorithms everywhere, as cultural artifacts (Finn, 2017, p.15; Seaver, 2017).

On the other hand, it is also helpful to recall that many current and historical algorithms have not implied as part of their problem-solving process to attain their objective in a neat, simple and efficient form. As a matter of fact, some algorithms aim only at keeping a solution in tension, without giving it away for everyone at every time (puzzles or riddles), while others simply produce contemplative outcomes, or even explicit nonsense (some art or literary pieces in the tradition of Dadaism, for instance). Solutions need not only be effective and efficient, they can also be creative, entertaining, experimental, contestatory, convoluted, tortuous or even purposely enigmatic or mistaken. Without reflecting on the diversity of possible outcomes and their consequences, computer scientists have usually produced equivalences between physical realities and formal symbolic systems, which have minimized the variety of "solutions" of the human world. Obviously this is not "wrong". It is what is expected from our computational machines under the dominant social normativity. But our understanding should not mistake an effect for a cause. And of course, this does not bring us closer to a precise definition of an algorithm in the broader sense that is implied here, even if it makes clearer the scope of the task.

If we take these initial considerations into account, we can see that, in order to look for a definition of an algorithm that describes both what engineers do when they program a computer, and what users do when they tinker and apply that program, or simply invent parallel procedures for the problem they attempt to solve through any other physical technology, we need to take a different approach. First of all, we need to recognize that an algorithm is an attempt to bring something into the material world (an idea, a calculation, a previous experience). Clearly, this does not mean that every symbol will have a physical manifestation, but that symbols are intermediaries, pieces that attempt to make a translation between the ideal and the material. For as Lev Vygotsky (1978) explains, what we conventionally call tools and what we conventionally call symbols are two aspects of the same phenomenon. According to him, mediation through tools could be seen as more outwardly oriented, while mediation through signs could be seen as more inwardly oriented, toward "the self", but both aspects emerge in every cultural artifact.

If we apply this notion of artifact mediation to our search, we can see that, whether through direct tools or indirect symbols, an algorithm implies an iterative interaction with technology, or in other words, a practice of recursive intertwining between humans and the technologies they produce. Yet, for Vygotsky, an interaction with symbols or tools is not simply functional, in the sense in which a subject manipulates an object at will to achieve a task. Instead, cultural artifacts regulate interactions with one's environment and with oneself. But this is not innocuous: cultural mediation has a recursive, bidirectional effect; mediated activity simultaneously modifies both the environment and the subject. Cultural mediation influences behaviors, synthetizes experiences from forbearers, and prepares children to acquire specific sets of accumulated memories, as knowledge (see here also Connerton, 1989). This co-constitution is what Malafouris, along a series of cognitive examinations, has termed as metaplasticity (Malafouris, 2010; 2013; 2015).

In the end, cultural mediation—or the ability to think and operate through cultural artifacts—produces historical modes of thinking—i.e. ideologies—and styles of cognition that affect how we learn, think and represent our environment and ourselves. This is what can be described as the notion of distributed cognition (Cole and Engeström, 1993; Gallagher, 2005; 2013). Hollan et al. (2000, p.177) provide a definition of this process: “Distributed cognition means more than that cognitive processes are socially distributed across the members of a group. It is a broader conception that includes phenomena that emerge in social interactions as well as interactions between people and structure in their environments.” The notion of distributed cognition has been a common hypothesis in linguistics and psychology ever since the writings of Vygotsky were published and made available in different translations since the 1970s, but it is by no means the standard model. There are many critics that maintain that, even if aided through different tools, thinking happens basically inside the brain (Adams and Aizawa, 2008; 2010; Loh and Kanai, 2016), or they present situations in which thinking is affected from “outside” factors (which Clowes terms as “the impact thesis” (Clowes, 2019)). We will not deal here with those arguments, since they appear to have a strong need for an essentialist form of conceptualizing cognition.

Moreover, since culture is for any notion of distributed cognition a foundational concept, anthropologists have made major contributions to our understanding of both the implementation of culturally mediated forms of cognition and the various ways in which the heterogeneity of culture supports and requires the distribution of cognition. One of these anthropologists was Clifford Geertz. For Geertz, individuals submit themselves to governance by symbolically mediated programs for producing artifacts, organizing social life, or expressing emotions. In this recurring process that reaches every layer of an individual's life, humankind determines, if unwittingly, "the culminating states of its own biological destiny" (Geertz, 1973, p.48). He states, in a formulation that evokes the Vygotskian approach:

> [S]ymbols are thus not mere expressions, instrumentalities, or correlates of our biological, psychological, and social existence; they are prerequisites of it. Without men, no culture, certainly; but equally, and more significantly, without culture, no men ().

Geertz's formulation found strong empirical evidence among theoretical biologists, for whom the connection between culture and biology implied more than a simple correlation. As Laland et al. (2000, p.131) later would claim: "cultural traits, such as the use of tools, weapons, fire, cooking, symbols, language, and trade, may have played important roles in driving hominid evolution in general and the evolution of the human brain in particular" (see also Dunbar, 1993; or Aiello and Wheeler, 1995). Nonetheless, when Geertz and other social scientists started confining everything under the domain of "culture", throughout the 1980s, the concept became too broad and lost its specific, explanatory power. As Nick Seaver (2017, p.4) writes: "Its implicit holism and homogenizing, essentialist tendencies seemed politically problematic and ill suited to the conflictual, changing shape of everyday life." As a response, one of the most resourceful attempts in the social sciences to overcome the difficulties brought about by an all-encompassing concept—which was nonetheless useful as a theoretical compass on a structural level—was to turn to the study of practicesand symbolic interactions (Bourdieu, 1972; Certeau, 1984;

Blumer, 1986). Consequently, many sociologists and anthropologists turned from a vision of a frame culture as a unified domain, to the multiplication of sites and cultures, where they could study and map emerging symbolic orders, sometimes coordinated, sometimes conflicting, out of which to make sense of the different layers of social life. This approach left behind the deterministic tone of previous explanations, with their emphasis on rules, models and texts, and began focusing instead on strategies, interests, improvisations and interactional occurrences. Recovering this emphasis, and back to our line of inquiry, Seaver (2017, p.5) provides a description of an algorithm that is worth mentioning:

> Like other aspects of culture, algorithms are enacted by practices which do not heed a strong distinction between technical and non-technical concerns, but rather blend them together. In this view, algorithms are not singular technical objects that enter into many different cultural interactions, but are rather unstable objects, culturally enacted by the practices people use to engage with them.

Seaver highlights the relational aspect of processes, enacted by practices rooted in cultural codes, therefore avoiding both a subject-centered perspective as well as a machine-centered view. In that sense, his definition is in line with a number of interesting theories and methodologies that have emerged in sociology and science and technology studies over the past two decades, for example: actor-networks (Callon, 1986; Latour, 1992; 2005), sociotechnical ensembles (Bijker, 1999), object-centered socialities (Knorr Cetina, 1997), relational materialities (Law, 2004), constitutive entanglements (Orlikowski, 2007) or object-oriented ontology (Harman, 2002; Bryant, 2010), as well as the approach of cognitive ecology (Hutchins, 2010) and material engagement theory (Malafouris, 2005; 2013) in the cognitive sciences.

These theories challenge and transcend conventional distinctions between objects and subjects, as well as between social abstractions and material iterations. Furthermore, their particular value lies in their insistence on speaking of the social (e.g. culture) and the material (e.g. nature) in the same register, and on not resorting to a limiting dualism that treats them as separate, even if interacting, phenomena.

Being an anthropologist, Seaver concentrates on the instabilities, the discontinuities, the confusions, the contradictions and the misunderstandings that enable different traditions and enrich human life. However, his view can be further explored, since it lacks a reasonable explanation of how, despite being categorized as "unstable objects", algorithms may appear as robust, reliable and even intrinsically repeatable. In other words, how do procedural patterns are sustained, despite variance; how consistencies emerge to enable traditions; when are recurrences broken up and when are they maintained? These inquiries are relevant because algorithms are something more than people executing socially available recipes and tweaking them with a personal taste. Algorithms are clusters of affordances and patterns that emerge in every process of *recursive* intertwining between humans and technologies. In that sense, they could be seen as material or immaterial scripts that link mental states with both material procedures and technocultural resources, enacted as a cultural practice to accomplish a specific task (effectively or not). And yet, in this description, mental states need not pertain to a single individual. Actually, if they would really belong to a unique individual (someone looking for a unique solution to his/her own problems, desires or needs) they would be socially illegible. But shouldn't this call for the inference of collective mental states? And what would that entail? The issue demands a deeper inspection, and we will now turn to it.

The mental and the notion of collective intentionalities

In order to inspect closer how humans and technologies interact through material or immaterial procedures linking mental states to real-world conditions, we need to acknowledge what we mean by mental states, and how they emerge as techno-cultural practices out of which specific patterns can be traced. This will require a short detour to explain some basic conceptions, but by the end of this explanation we will have a clearer landscape of the categories at stake.

A mental state can best be delineated by the notion of intentionality. Intentionality is a complex philosophical concept that emerged with Medieval Scholasticism through Medieval Islamic philosophy, but was later retaken and developed in phenomenological circles, starting from the 19th Century. Franz Brentano's work is usually set as a point of departure for contemporary analyses. In his writings, intentionality is set as an attribute of an individual's mind, which adheres to mental contents, as opposed to attributes of the real world, such as extension and duration, which can be predicated of existing objects. Brentano takes on the discussion from St. Thomas Aquinas, who established that the object which is thought is intentionally in the thinking subject, the object which is loved in the person who loves, the object which is desired in the person desiring, etc. In that sense, intentionality is clearly something that can be predicated of inexistent phenomena, but which has an effect on our own conceptions, desires and beliefs. Brentano (1995, p.68) writes:

> Every mental phenomenon is characterized by [. . .] the intentional (or mental) inexistence of an object, and what we might call, though not wholly unambiguously, reference to a content, direction toward an object (which is not to be understood here

> as meaning a thing), or immanent objectivity. Every mental phenomenon includes something as object within itself [. . .] We can, therefore, define mental phenomena by saying that they are those phenomena which contain an object intentionally within themselves.

At this point, intentionality was described as a clear attribute of mental activity, independent of a real world, but clearly related to it, and decisive to ascribe it meaning. This trait was important because it offered a form of cognizing reality without relying on the Kantian formulation that attempted to align (individual) sensations and (social) concepts. In other words, it created a model where things could be cognized beyond a thick web of structured epistemological pre-conceptions. This is precisely what encouraged Husserl's enthusiasm, as inscribed in his motto "Back to the things themselves!" (*Zurück zu den Sachen selbst*!). For as Merleau-Ponty (2005, p.xix) writes:

> What distinguishes intentionality from the Kantian relation to a possible object is that the unity of the world, before being posited by knowledge in a specific act of identification, is 'lived' as ready-made or already there.

However, intentionality in this early stage also made a clear difference between the inner, mental world, and the outer, objective reality. In that sense, it was still trapped in the fundamental dualism that characterized the positivist style of thinking in the late 19th and early 20th centuries. This dynamic has been sufficiently deconstructed, especially within the theories that were mentioned in the previous section, and there is no need to discuss it further. A second problem is that this early notion of intentionality also posited a very clearly delimited "self" for whom an intention (and communication of that intention) is transparent. The precise refutation of this point can be extensive, and it can also run through diverging lines, but for synthetic aims, we can resort back to the Vygotskian approach and understand the “self” as a symbol and a cultural artifact. Actually, both Vygotsky and a contemporary anthropologist of him, G.H. Mead, worked along the lines of a similar hypothesis, which has been termed the “social genesis of the self” (Glock, 1986), and which implied both the process of internalisation (through education in the child) and the genesis of linguistic meaning. For Mead (1972, p.164), for instance, “[t]he process out of which the self arises is a social process which implies interaction of individuals in the group, implies the preexistence of the group.” Accordingly, he adds:

> the self appears in experience essentially as a “me” with the organization of the community to which it belongs. This organization is, of course, expressed in the particular endowment and particular social situation of the individual [. . .]. He is what he is in so far as he is a member of this community, and the raw materials out of which this particular individual is born would not be a self but for his relationship to others in the community of which he is a part (Mead, 1972, p.200).

Following the Vygotsky/Mead hypothesis, there cannot even be a “direct” connection between an individual and her experience, because this connection is mediated through language, by which a “self” appears as some type of thing. In other words, the emergence of a “self” is an effect, or a functional construction, of a subject that has learned how to enunciate and use the particle “I” under a given set of socially-sanctioned, grammatical rules. This brings us to a rather interesting situation on the cognitive side. For if the self is a social construction, what is to be done with what we call “the mental”? Is the link between both notions merely a deficient attribution, or is it a faulty causal connection? Mead describes the mental as an emergent phenomenon, which involves a relationship to the character of things:

> Those characters are in the things, and while the stimuli call out the response which is in one sense present in the organism, the responses are to things out there. The whole process is not a mental product and you cannot put it inside of the brain. Mentality is that relationship of the organism to the situation which is mediated by sets of symbols (Mead, 1972, pp.124–125).

This turns irrelevant the attribution of mentality to the self. On the same grounds, a causal connection between them can only be inferred as inexistent. Instead, both are equally emergent effects of a given symbolic mediation. Mead's description of the mental (thatcognitive relationship of an organism to a situation, mediated by symbols) is the backbone to the definition of an algorithm that was proposed on the previous section. It is also a touchstone in the tradition of cognitive anthropology that has been associated with the idea of cognitive ecologies (Douglas, 1986; Lave, 1988; Connerton, 1989; Hutchins, 1995; 2010), as well as in traditions of cognitive sciences that inquire into models of an embodied, embedded, extended and/or an enactive social mind (Clark, 1997; 2003; 2015; Clark and Chalmers, 1998; Gallagher, 2005; 2013; Gallagher and Miyahara, 2012). Gallagher (2013, p.4), for instance, describes the mental in this way:

> If we think of the mind not as a repository of propositional attitudes and information, or in terms of internal belief-desire psychology, but as a dynamic process involved in solving problems and controlling behavior and action—in dialectical, transformative relations with the environment—then we extend our cognitive reach by engaging with tools, technologies, but also with institutions. We create these institutions via our own (shared) mental processes, or we inherit them as products constituted in mental processes already accomplished by others.

Indeed, breaking the causal link between the mind and the self allow us to see the dense and emergent network of affordances and enactions that constitute cognitive phenomena. But how do intentionalities come back into the picture? For Brentano, intentionalities were so much as the mark of the mental, i.e. the defining quality of an in-existent, psychological phenomenon. But if the mind is not any more located in an inner, private world, should we just simply do without them? Quite the opposite. As a matter of fact, intentionalities play a stronger role within a distributive cognition approach. But we need to refine the conceptual frame to see how this can be integrated into a comprehensive explanation.

An intentionality is not a purpose, nor a design or an intention to do something, although the notions are closely related. Actions are intentional, for example, not only because there is a will behind them, but also because they follow a goal or a project. If I am hungry and I do not have anything to eat at home, I can go out to a supermarket to buy groceries in order to cook, or to a restaurant, or even to a place where food is distributed if my economic means are limited. These, among others, are available modes of action, connected to material and technical functions, social behaviors and actionable symbolic networks. But we know that there used to be a time when, if hungry, people could go out hunting or foraging, and the relevant social programs were there to support those activities. Intentionalities are attached then to historical norms, cultural repertoires, social habits, communal values, rituals and many other forms and forces that can be seen to shape an individual's action. For as Brandom (1994, p.61) writes:

> only communities, not individuals, can be interpreted as having an original intentionality. [T]he practices that institute the sort of normative status characteristic of intentional states must be social practices.

In that sense, the social life of an individual consists in a good deal in determining the appropriateness of her own desires and needs as these are articulated to the available social practices, or cultural programs, through inferential reasoning, practical adjustments and other means.

Now, even if at a first look this explanation seems to restrain an individual's agency, by making her guide a certain "intended" action through a given catalogue of socially sanctioned paths, the picture that this model enables is in fact richer and more complex. In a few words, a strict functionalism does not apply (Elster, 1983; Douglas, 1986, p.32ff). As a matter of fact, a model like Malafouris' material engagement theory actually sustains that the distinctive forms of human agency emerge precisely in the practical space afforded by the interactions (Malafouris, 2008; 2015). After all, an individual never "acts" in a void either. And as Cooren et al (2006, p.11) write:

> Agency is not a 'capacity to act' to be defined a priori. On the contrary, it is 'the capacity to act' that is discovered when studying how worlds become constructed in a certain way.

In that sense, intentionalities are sustained in social practices without losing their capacity for an individual's adaptation, expression and further innovation. And as such, they can be acknowledged as *collective intentionalities*, fundamental pieces that connect an individual to a larger collective, without necessarily turning them into a deterministic setup. Collective intentionalities are in that sense something as action-able paths, through which an individual orients and articulates her actions with the resources and experiences of a cultural community, i.e. a community of practice.

Furthermore, collective intentionalities are so relevant that Tomasello (2014) assigns to them, in an appealing hypothesis, a definite role in the evolution of the species, since they allow coordination and cooperation to occur not only simultaneously, but also throughout generations. For this cognitive linguist, collective intentionalities comprise

> not just symbolic and perspectival representations but conventional and 'objective' representations; not just recursive inferences but self-reflective and reasoned inferences; and not just second-personal self-monitoring but normative self-governance based on the culture's norms of rationality (Tomasello, 2014, p.6).

As such, they are the infrastructure of social life, underlying even culture and language through pre-linguistic aims and forces that acquire a given shape. Developing over the foundations of collective intentionalities,

> culture and language, as agent-neutral conventional phenomena [...] provide another setting within which a new form of human sociality can lead to a new form of human thinking, specifically, objective reflective-normative thinking (Tomasello, 2014, p.141).

In that sense, collective intentionalities can be said to be the building blocks of human-symbol/tool interactions. But in the end, if collective intentionalities are not a quality of the objective world—but rather its foundation—where are these to be seen, or how do they emerge and provide tangible samples for interactions to occur? We will tackle the issue in the following section.

Collective intentionalities and algorithms: heuristics and dynamics

The notion of collective intentionality is only such if it retains one condition that was there since Brentano attempted a definition: it is an attribute of a mental state, i.e. a mark of the mental. But we have seen that, in a distributed cognition approach, the mental cannot be exclusively associated with a self; it is rather an articulated web that links individuals to tools and symbols that have been pre-structured by a collective, and are enacted through social practices. So we are presented with an empirical challenge: how to spot an intentionality if it is neither an objective nor a subjective phenomenon in the classical sense? Collective intentionalities are usually "hidden" to the naked eye, sometimes they are by-products of repeated actions, much as a trailing path in the woods which appears after years and years of different individuals walking through it, but sometimes they also stand out in oblique moves. In any case, they comprise the causal loops that run behind collective articulations (making up a good deal of group identities, for example), and these can emerge as latencies, background or naturalized conventions (a specialized analysis in Chant, Hindriks and Preyer, 2014; in relation to this topic see also Toscano, n.d.). The only minimal assumption is that they stand in a threshold, as that which allows community survival without demanding from individuals that they give up on their autonomy (even though the threshold is dynamic, and is not the same for a child as for the elder, or throughout different knowledge capacities and hierarchies).

In the last years, cognitive scientists have developed different models to locate intentionalities via distinctive approaches. The neo-behaviourist Daniel Dennett, for instance, ascribes intentionality to observed rational behavior, and he describes the agent as someone

> who harbors beliefs and desires and other mental states that exhibit intentionality or 'aboutness', and whose actions can be explained (or predicted) on the basis of the content of these states" (Dennett, 1991, p.76).

The approach is clear, and amounts to correlating traces to directions and motivations in a straightforward way. Of course, it is constrained to reading rational behavior and to valuing every action as instrumental to achieve a specific goal. In contrast, a neo-pragmatist view (Brandom, 1994; 2000; Cash, 2008; 2009) proceeds by ascribing intentionality as an explanation and a specific coupling of action to social norms. As Cash (2008, p.101) argues:

> based on the similarity of their movement to the kind of actions, [...] would entitle us to ascribe such intentional states as reasons.

This might be a key aspect in certain contexts, but it is constrained to knowing what the norms to be applied are, and to evaluating if the ensuing pairing of actions to those norms succeed or not. In that sense, they imply the recognition of patterns, and a judgment on their application or continuity, but they also underestimate the value of deviance and disregard a space for individual creativity. Even a third approach, which we can call a neo-interactionist perspective, aims at understanding other's intentionalities not by acknowledging or judging their actions, but by understanding actual or potential interactions with others in socially appropriate ways. As Gallagher and Miyahara (2012, p.135) write in this account:

> we normally perceive another's intentionality in terms of its appropriateness, it's pragmatic and/or emotional value for our particular way of being, constituted by the particular goals or projects we have at the time, or implicit grasp on cultural

> norms, our social status, and so on, rather than as reflecting inner mental states, or as constituting explanatory reasons for her further thoughts and actions.

The neo-interactionist perspective certainly rounds up some of the forms in which intentionalities emerge, but they do not completely revoke the previous explanations, and instead helps compile a catalogue of intentional enactions.

At this point, we can bring back the definition of an algorithm that was proposed in the first section, and mobilize it in an illustrative form. We can thus define an algorithm as

> a recursive script that links collective intentionalities with both material procedures and technocultural resources, enacted as a cultural practice to accomplish a specific task.

We now know what is meant by a collective intentionality. And we can expect how to look for them. But this definition does more than just describe a process. It wants to reflect on the fact that collective intentionalities are not by themselves the structures that sustain a community's culture. It is really their mobilization, in an algorithmic form, which brings them to life. It would therefore be more precise to see an algorithm as an action than as an object, however "unstable" that object would turn out to be. On those grounds, an algorithm should be seen more as an activity, an "algorithmation", a productive emergent pattern that enables connections out of a given networked system or a distinctive cognitive ecology.

Within the algorithmic feedback loops, the individual performs a key computational function. Clearly, since we do not rely on a machine-centered perspective, a corresponding idea of computation must be outlined. For simplicity, we can take over Hutchins view here. He suggests that computation should be regarded as "the propagation of representational states across representational media" (Hutchins, 1995, p.118). In that sense, individuals are the agents transforming representational states for those collective intentionalities through algorithmic procedures, that is, through recursive technical enactions. But in this finely threaded network, the individual is neither the origin nor the final end. And yet, she is not a simple cog in the system either. She is interconnected, interacting, adjusting herself and her environment with this complex and finely tuned mechanism, which we might indeed call at this point a socio-computing infrastructure (Toscano, n.d.). Yet this term cannot imply a fix and immovable architecture, but a dynamic structure where certain accomplishments, and not others, are viable. Laland et al. (2000, p.130), for instance, refer as "niche construction" to the human-made or human customized structures that are essential to the development, production and continuance of certain activities. Jones et al. (1997) identify that same activity as "ecosystem engineering". The notion of socio-computing infrastructure that is proposed here here should be read along those lines, but where the accent on collective intentionalities and a social cognitive activity is deliberate.

Similarly, Laland et al. (2000) propose the idea of an ecological niche, which implies that an organism occupies a distinctive role in each ecosystem. This opens up yet another approach to the task of identifying intentionalities, as a supplement to the ones that were described above. For in certain contexts, defining the ecological niche of an individual can render a better inspection of the collective intentionalities implied in a given system. In other words, in human-machine interactional systems, focusing on the active or operating roles of given individuals as socially-enabled subjective behaviors or social functions can shed light on the specific collective intentionalities at work. This route acknowledges that the individual is relevant, only not on her own, but through her dynamic links (interpretations, associations, appreciations) to a broader community of practice. This can be useful in anthropological cases, but is doubtlessly crucial in historical inquiries and techno-archeological analyses. We can bring a couple of empirical cases from this latter for consideration.

a) Inka's Khipus

If we think of historical socio-computing infrastructures that were lost or disrupted when the groups tied to them ceased to exist, we can acknowledge which were precisely the missing access points that make the reconstruction or re-interpretation of those systems difficult, or sometimes impossible. Two cases can be explored here at some length. As a first case, we can recall the recent decipherment of ancient *khipus* in Peru. Khipus were devices of statistical notation that stemmed during the Inka Empire, but were used until the Spanish colonial period in that South American country. These devices did not employ numerical symbols, but relied instead on cotton strings of different lengths and colors, and were encoded using knots at different places. As Medrano and Urton (2018, p.2).

> the Inkas filled the twists and knots of the khipus with data, including bureaucratic accounting measures such as tax assignments and census counts.

One element is noticeable when approaching the khipu coded system: the people that used it as a statistical artifact did not suddenly disappeared without leaving a trace (as the Maya civilization did, for instance). On the contrary, the khipus coexisted during some time with the European statistical methods of the epoch, which the Spanish had brought with them. In that situation, even if symbolic and abstract operations were readily available, khipus were kept because they implied a material manifestation of different social values, symbols that the people of that particular culture considered relevant information, as opposed to mere abstractions. In other words, khipus enriched merely numerical data: they registered social relations of a highly organic and interpersonal nature, traits that were indifferent to the Spanish accounting methods, which were therefore inadequate for their transmission (Medrano and Urton, 2018, p.12).

In the Inka worldview, khipus were not only statistics, but a representation of a given reality made possible through a material craft. Of course, since the symbols they employed were not easily manageable, the khipus were discontinued after some time. Nobody wrote how they were encrypted, so the key to reading them disappeared. In a sense, khipus were meant not only as notational systems, but also as mnemonic devices for khipu keepers and scribes. When these professionals finally changed the notational system to make their calculations, the mnemonic function ceased to operate. But while still active, these professionals were implementing an algorithmic procedure: they applied a know-how for a given collective intentionality—to count, or calculate, a given state of affairs—and turned it into an objectified device—a social representation—thus computing it. The khipus were finally deciphered through an analogy with an European-style census that was later discovered to match one of these objects with a strict correlation, but also by paying a close attention to Inka's testimonies on economy, politics, religion and other aspects of their civilization that were highly valued, and considered to be worthy of a specific notational foundation.

b) The Voynich Manuscript

Another case is provided by the situation of the *Voynich Manuscript*, kept in the collection of rare books at the Beinecke Library, at Yale. This fifteenth-century codex has not been deciphered until this day for several reasons, many of which are elements that indicate how a socio-computational infrastructure, and with it a specific algorithmic enaction, is put to work. The "book" was written in an unknown script by an unknown author. The impossibility to assign it a context, a precise culture, or even a specific function within a given literary or scientific biography, contribute to see this piece as an example of a radical particularity that highlights its isolatory character. This is just not how a "book" works. Rather than executing a typical communicative intentionality, the *Voynich Manuscript* contradicts its form and function, and appears as a work of madness. The current custodians of the book present it thus: "the manuscript has no clearer purpose now than when it was rediscovered in 1912" (Clemens, 2016). There are no points of access because nobody knows where to begin with. Of course, some facts can be determined: the approximate date of its physical appearance, as well as a list of its owners, all of which tempt the researchers to make some claims based on analogical and normative assumptions, of the kind that cognitive scientist have shown how to bring about. But in the end, the manuscript has been annulled as an informational device, as well as an instrument of contextual cognition. However, it has become a new source of computations, for the curiosity of the researchers has turned it into an object of study, which means that it is being transfigured across different representational media. In any case, without an anchoring fact that stabilizes its meaning, such investigations speak more about our computational procedures than about the content of the "book", so they also tell about our need for conceptual pre-assumptions and our own inabilities to understand even human-made objects when a clear intentionality is not recognized or set onto them.

Conclusions

Algorithms cannot be reductively described as machine drivers or mere coding language. They imply instead a complex cultural activity that involve both material and immaterial interactions. This article has aimed to show how, as part of their particular enaction, they are constructed along collective intentionalities of different sorts. In that sense, algorithms do shape desires, wants and needs, as these are ingrained in distinctive communities. It is indeed through an algorithmic recursiveness that collective cultures flourish and expand. It is also through an individual's tinkering with them that they can give way to adjustments and innovations, provided that the underlying intentionalities—whether as paths, patterns, occurrences or scripts—remain fundamentally recognizable.

In his book *What do Algorithms want*?, Ed Finn finds an ingenuous answer to this complex question: "This is what algorithms want, or what we design them to want: to know us completely" (Finn, 2017, p.82). But this statement is a simplification that requires further clarification itself. Algorithms cannot want something in themselves, but neither do we. Or the other way around: algorithms want what "we" want, or rather: we want through them. Which is not always something evident. After all, "to want" is a cultured habit, which is ingrained in children through upbringing and education. As individuals, we use socially available algorithms to channel pre-linguistic and abstract desires and needs, which only through them acquire a definable form.

So in a way it is true: algorithms want what we design them to want. But we can only design what is culturally available, collectively interpretable, socially desirable. So it is less true that we design all algorithms "to know us completely". In fact, most of the time, the opposite is just the case. In their recursivity, algorithms enact collective intentionalities that are frequently turned into latencies, background or naturalized conventions, and then cease to appear as constructions to us. (Therefore, only in a culture where information extraction is a viable practice, the design of algorithms to extract information from us—what Finn refers as "to know"—will be a logical consequence.) In the end, algorithms imply an articulatory activity: they are collective processes of cultural inscription, through which individuals enact socially available programmatic technologies for a specific, intentional objective.

This article has sought to provide examples on how to approach collective intentionalities, both by recalling how cognitive scientists apply logical inferences to distinguish emergent phenomena, and by turning to historical socio-computing infrastructures to inspect their legibility (or lack thereof) and operation. Evidently, much works needs to be done to deepen a techno-archeological inquiry of this kind, but this article has sought to contribute with some entry points to enrich such analyses in a distinctive way.

Bibliography

Adams, F. and Aizawa, K., 2008. *The Bounds of Cognition*. 1st ed. Malden, MA: Blackwell Publishing.

Adams, F. and Aizawa, K., 2010. Defending the Bounds of Cognition. In: R. Menary, ed. *The Extended Mind* [Online]. The MIT Press, pp.67–80. https://doi.org/10.7551/mitpress/9780262014038.003.0004.

Aiello, L.C. and Wheeler, P., 1995. The Expensive-Tissue Hypothesis: The Brain and the Digestive System in Human and Primate Evolution. *Current Anthropology* [Online], 36(2), pp.199–221. https://doi.org/10.1086/204350.

Bijker, W.E., 1999. *Of Bicycles, Bakelites, and Bulbs: Toward a Theory of Sociotechnical Change*. 3rd ed., *Inside technology*. Cambridge, MA: MIT Press.

Blumer, H., 1986. *Symbolic Interactionism: Perspective and Method*. Berkeley: University of California Press.

Bourdieu, P., 1972. *Outline of a Theory of Practice*. Cambridge, MA: Cambridge University Press.

Brandom, R.B., 1994. *Making It Explicit: Reasoning, Representing, and Discursive Commitment*. Cambridge, MA: Harvard University Press.

Brandom, R.B., 2000. *Articulating Reasons: An Introduction to Inferentialism*. Cambridge, MA: Harvard University Press.

Brentano, F., 1995. *Psychology from an Empirical Standpoint [1874]*. Ed. by O. Kraus and L.L. McAlister (A.C. Rancurello, D. Terrell and L.L. McAlister, Trans.), *International Library of Philosophy*. London; New York: Routledge.

Bryant, L.R., 2010. *Onticology–A Manifesto for Object-Oriented Ontology Part I*. Available at: <https://larvalsubjects.wordpress.com/2010/01/12/object-oriented-ontology-a-manifesto-part-i/> [visited on 18 January 2023].

Callon, M., 1986. Some Elements of a Sociology of Translation: Domestication of the Scallops and the Fishermen of Saint Brieuc Bay. In: J. Law, ed. *Power, action, and belief: a new sociology of knowledge?*, *Sociological review monograph*, 32. London; Boston: Routledge & Kegan Paul, pp.196–233.

Cash, M., 2008. Thoughts and oughts. *Philosophical Explorations* [Online], 11(2), pp.93–119. https://doi.org/10.1080/13869790802015635.

Cash, M., 2009. Normativity is the mother of intention: Wittgenstein, normative practices and neurological representations. *New Ideas in Psychology* [Online], 27(2), pp.133–147. https://doi.org/10.1016/j.newideapsych.2008.04.010.

Certeau, M.d., 1984. *The Practice of Everyday Life* (S. Rendall, Trans.). Berkeley & Los Angeles, CA: University of California Press.

Chant, S.R., Hindriks, F. and Preyer, a.G., eds., 2014. *From Individual to Collective Intentionality: New Essays*. Oxford, New York: Oxford University Press.

Clark, A., 1997. *Being There: Putting Brain, Body, and World Together Again*. Cambridge; London: The MIT Press.

Clark, A., 2003. *Natural-Born Cyborgs: Minds, Technologies, and the Future of Human Intelligence*. Oxford; New York: Oxford University Press.

Clark, A., 2015. What 'Extended Me' knows. *Synthese* [Online], 192(11), pp.3757–3775. https://doi.org/10.1007/s11229-015-0719-z.

Clark, A. and Chalmers, D., 1998. The extended mind. *Analysis* [Online], 58(1), pp.7–19. Available at: <https://www.jstor.org/stable/3328150> [visited on 3 June 2019].

Clemens, R., ed., 2016. *The Voynich Manuscript*. New Haven; London: Yale University Press.

Clowes, R.W., 2019. Screen reading and the creation of new cognitive ecologies. *AI & Society* [Online], 34(4), pp.705–720. https://doi.org/10.1007/s00146-017-0785-5.

Cole, M. and Engeström, Y., 1993. A Cultural-Historical Approach to Distributed Cognition. In: G. Salomon, ed. *Distributed Cognitions: Psychological and Educational Considerations*, *Learning in doing*. Cambridge, MA: Cambridge University Press, pp.1–46.

Connerton, P., 1989. *How Societies Remember* [Online]. 1st ed. Cambridge University Press. https://doi.org/10.1017/CBO9780511628061.

Cooren, F., Taylor, J.R. and Van Every, E.J., 2006. *Communication as Organizing: Empirical and Theoretical Approaches into the Dynamic of Text and Conversation*, *LEA's communication series*. Mahwah, NJ: Lawrence Erlbaum.

Dennett, D.C., 1991. *Consciousness Explained*. Boston: Little, Brown and Company.

Douglas, M., 1986. *How Institutions Think*. 1st ed., *The Frank W. Abrams lectures*. Syracuse, NY: Syracuse University Press.

Dourish, P., 2016. Algorithms and their others: Algorithmic culture in context. *Big Data & Society* [Online], 3(2), pp.1–11. https://doi.org/10.1177/2053951716665128.

Dunbar, R.I.M., 1993. Coevolution of neocortical size, group size and language in humans. *Behavioral and Brain Sciences* [Online], 16(4), pp.681–694. https://doi.org/10.1017/S0140525X00032325.

Elster, J., 1983. *Explaining Technical Change: A Case Study in the Philosophy of Science* [Online], *Studies in Rationality and Social Change*. Cambridge ; London et al.: Cambridge University Press. Available at: <http://archive.org/details/Elster1983ExplainingTechnicalChangeACaseStudyInThePhilosophyOfScienceBook> [visited on 18 January 2023].

Finn, E., 2017. *What Algorithms Want: Imagination in the Age of Computing* [Online]. 1st ed. Cambridge, MA: The MIT Press. https://doi.org/10.7551/mitpress/10766.001.0001.

Gallagher, S., 2005. *How the Body Shapes the Mind*. Oxford; New York: Clarendon Press.

Gallagher, S., 2013. The socially extended mind. *Cognitive Systems Research* [Online], 25-26, pp.4–12. https://doi.org/10.1016/j.cogsys.2013.03.008.

Gallagher, S. and Miyahara, K., 2012. Neo-Pragmatism and Enactive Intentionality. In: J. Schulkin, ed. *Action, Perception and the Brain: Adaptation and Cephalic Expression* [Online], *New Directions in Philosophy and Cognitive Science*. London: Palgrave Macmillan UK, pp.117–146. https://doi.org/10.1057/9780230360792_6.

Geertz, C., 1973. *The Interpretation of Cultures: Selected Essays*. New York: Basic Books.

Glock, H.-J., 1986. Vygotsky and Mead on the Self, Meaning and Internalisation. *Studies in Soviet Thought* [Online], 31(2), pp.131–148. Available at: <https://www.jstor.org/stable/20100086> [visited on 18 January 2023].

Harman, G., 2002. *Tool-Being: Heidegger and the Metaphysics of Objects*. 1st ed. Chicago: Open Court.

Hutchins, E., 1995. *Cognition in the Wild*. Cambridge, Mass.; London: The MIT Press.

Hutchins, E., 2010. Cognitive ecology. *Topics in Cognitive Science* [Online], 2(4), pp.705–715. https://doi.org/10.1111/j.1756-8765.2010.01089.x.

Knorr Cetina, K., 1997. Sociality with Objects: Social Relations in Postsocial Knowledge Societies. *Theory, Culture & Society* [Online], 14(4), pp.1–30. https://doi.org/10.1177/026327697014004001.

Laland, K.N., Odling-Smee, J. and Feldman, M.W., 2000. Niche construction, biological evolution, and cultural change. *Behavioral and Brain Sciences* [Online], 23(1), pp.131–146. https://doi.org/10.1017/S0140525X00002417.

Latour, B., 1992. Where Are the Missing Masses? The Sociology of a Few Mundane Artifacts. In: W.E. Bijker and J. Law, eds. *Shaping technology/building society: studies in sociotechnical change*, *Inside technology*. Cambridge, MA: MIT Press, pp.225–258.

Latour, B., 2005. *Reassembling the Social: An Introduction to Actor-Network-Theory*, *Clarendon lectures in management studies*. Oxford; New York: Oxford University Press.

Lave, J., 1988. *Cognition in Practice: Mind, Mathematics and Culture in Everyday Life* [Online]. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511609268.

Law, J., 2004. *After Method: Mess in Social Science Research*, *International library of sociology*. London; New York: Routledge.

Loh, K.K. and Kanai, R., 2016. How Has the Internet Reshaped Human Cognition? *The Neuroscientist* [Online], 22(5), pp.506–520. https://doi.org/10.1177/1073858415595005.

Malafouris, L., 2005. The Cognitive Basis of Material Engagement: Where Brain, Body and Culture Conflate. In: E. DeMarrais, C. Gosden and C. Renfrew, eds. *Rethinking Materiality: Engagement of Mind with Material World*. Cambridge: McDonald Institute for Archaeological Research, pp.53–61.

Malafouris, L., 2008. At the Potter's Wheel: An Argument for Material Agency. In: C. Knappett and L. Malafouris, eds. *Material Agency: Towards a Non-Anthropocentric Approach* [Online]. Berlin: Springer, pp.19–36. https://doi.org/10.1007/978-0-387-74711-8_2.

Malafouris, L., 2010. Metaplasticity and the human becoming: principles of neuroarchaeology. *Journal of Anthropological Sciences*, 88(4), pp.49–72.

Malafouris, L., 2013. *How Things Shape the Mind: A Theory of Material Engagement*. Cambridge; London: The MIT Press.

Malafouris, L., 2015. Metaplasticity and the Primacy of Material Engagement. *Time and Mind* [Online], 8(4), pp.351–371. https://doi.org/10.1080/1751696X.2015.1111564.

Mead, G.H., 1972. *Mind, Self and Society: From the Standpoint of a Social Behaviorist [1934]*. Ed. by C. Morris. Chicago; London: University of Chicago Press.

Medrano, M. and Urton, G., 2018. Toward the Decipherment of a Set of Mid-Colonial Khipus from the Santa Valley, Coastal Peru. *Ethnohistory* [Online], 65(1), pp.1–23. https://doi.org/10.1215/00141801-4260638.

Merleau-Ponty, M., 2005. *Phenomenology of Perception: An Introduction [1945]* (C. Smith, Trans.), *Routledge classics*. London; New York: Routledge.

Newell, A., Simon, H.A. and Shaw, J.C., 1958. Elements of a theory of human problem solving. *Psychological Review* [Online], 65(3), pp.151–166. https://doi.org/10.1037/h0048495.

Orlikowski, W.J., 2007. Sociomaterial Practices: Exploring Technology at Work. *Organization Studies* [Online], 28(9), pp.1435–1448. https://doi.org/10.1177/0170840607081138.

Seaver, N., 2017. Algorithms as culture: Some tactics for the ethnography of algorithmic systems. *Big Data & Society* [Online], 4(2), pp.1–12. https://doi.org/10.1177/2053951717738104.

Tomasello, M., 2014. *A Natural History of Human Thinking:* Cambridge, MA: Harvard University Press.

Toscano, J., n.d. Intentionalities of the algorithm: historical practices and socio-computing infrastructures. A philosophical account. *Lo Sguardo*, 34(1).

Vygotsky, L.S., 1978. Tool and Symbol in Child Development. In: M. Cole, V. John-Steiner, S. Scribner and E. Souberman, eds. *Mind in Society: The Development of Higher Psychological Processes*. Cambridge: Harvard University Press, pp.19–30.

Modelling interactive computing systems: Do we have a good theory of what computers are?

Alice Martin
ENAC, Université de Toulouse

Mathieu Magnaudet
ENAC, Université de Toulouse

Stéphane Conversy
ENAC, Université de Toulouse

Abstract

Computers are increasingly interactive. They are no more transformational systems producing a final output after a finite execution. Instead, they continuously react in time to external events that modify the course of computing execution. While philosophers have been interested in conceptualizing computers for a long time, they seem to have paid little attention to the specificities of interactive computing. We propose to tackle this issue by surveying the literature in theoretical computer science, where one can find explicit proposals for a model of interactive computing. In that field, the formal modelling of interactive computing systems has been brought down to whether the new interaction models are reducible to Turing Machines. There are three areas where interaction models are framed. The comparison between TMs and interactive system models is at stake in all of them. These areas are namely some works on concurrency by Milner, on Reactive Turing Machines, and on interaction as a new computing

Philosophical Problems in Science (Zagadnienia Filozoficzne w Nauce)
No 72 (2022), pp. 67–109

paradigm. For each of the three identified models, we present its motivation, sum up its account for interaction and its legacy, and point out issues regarding the understanding of computers. The survey shows difficulties for epistemologists. The reason is that these analyses focus on the formal equivalence between interactive models of computation and classic ones. Such a project is different from addressing how a computing machine can be interactive: in other words, which mechanisms allow it.

Keywords

philosophy of computing, models of computation, interactive computing, computing mechanism, computational mechanistic explanation.

Introduction

In the philosophy of computing, we are paying increased attention to a set of new features of computers. This set has led to the introduction of a new label for these computing machines: they are referred to as interactive computing machines (Dodig-Crnkovic, 2011; Goldin, Wegner and Smolka, 2006; Soare, 2013; Van Leeuwen and Wiedermann, 2001; Wegner, 1997). The set of new features can be captured in the following statement made in a 2011 paper by Gordana Dodig-Crnkovic (our highlights):

> Present day computers are very different from the early standalone calculators designed for mechanizing mathematical operations. They are largely *used for communication in world-wide networks* and variety of information processing and knowledge management. Moreover, they play *an important role in the control of physical processes and thus connect to the*

> *physical world*, especially in automation and robotics. [. . .] Computational processes are nowadays *distributed, reactive, agent-based, and concurrent*. The main criterion of success of the computation is not its termination, but its *response to the outside world*, its speed, generality and flexibility; adaptability, and tolerance to noise, error, faults, and damage. (Dodig-Crnkovic, 2011)

Historically, the concept of interaction was introduced by a computer scientist, Milner, in the 1970s-1980s (Milner, 1975; 1982; 1993; 1999). At first, an *interactive computing system* was defined as a system where several threads execute instructions in parallel while being able to synchronize and communicate at certain moments of the execution. Since then, the characteristics of computer systems have continued to evolve, and by "interactive" we refer today to a broader set of properties that can be grouped as follows: the ability to continuously react in time to external events that modify computing execution. This class of computers deserves all our attention since they are ubiquitous. Every computer system today is designed to respond to external events in a predictable way and according to temporal constraints. In any case, what distinguishes this class of so-called interactive computing machines from the classical computer systems that preceded them is that they are no longer purely *transformational systems*. A *transformational system* is a classical computing device that, given a set of inputs, produces a final output after a finite execution. This evolution of computing complicates the answer to *what a computer is*. The question is well-known in the philosophy of computing (Piccinini, 2008; Rapaport, 2018; Smith, 2002). As already noted, many answers to the question distort it and answer the question of *what a computation is*, immediately projecting the field of investigation into the theory of computability:

> A fairly obvious, trivial, and almost-circular definition of 'computer' says that a computer is a machine that computes. The natural next question is: What does it mean to compute? But this shifts the burden of answering our question away from what computers are to the topic of what computation is. Many of the objections to various theories about computers are really objections to what counts as a computation. (Rapaport, 2018).

This leaves us with the specific issue we want to address. We ask whether models of computation for interaction allow us to answer the question of what an actual (necessarily interactive) computer is. Current computers come in various forms and we chose in this paper to restrict our concerns to a delimited notion of *interaction*, as defined in Human-Computer Interaction (Basman et al., 2018; Beaudouin-Lafon, 2006; Dearden and Harrison, 1997; Hornbaek and Oulasvirta, 2017; Myers, 1994), and target a specific set of ubiquitous computing devices—those interacting with humans, e.g., through digital interfaces. We will not elaborate on *analog computing* (Bielecki, 2019) and *natural computation* (Dodig-Crnkovic, 2011; MacLennan, 2003).

To tackle the issue of interactive computing devices, we propose here an approach that, to the best of our knowledge, has not been proposed so far: we want to examine the models of computation proposed in theoretical computer science to think about interactive computing systems. We offer a literature survey where one can find *explicit theories of interaction* [1]. We show that the formal modelling of interactive computing systems has been brought down to whether the new interaction models are reducible to Turing's *a-machines* (Turing, 1937)—we will refer to them as Turing Machines (TMs). Questioning

[1] We insist on our two criteria: *explicit theories of interaction* in *theoretical computer science*. We have in mind the fact that other communities e.g., the engineering community on reactive systems, are related to our topic but they have not conceptualized *interaction* as such.

the theoretical bounds of the Turing Machine in computer science when faced with the existence of interactive computing devices has been explored at least since Milner's work on communicating and mobile systems (Milner, 1993; 1999). To the best of our knowledge, there are three areas where interaction models are framed as such. These areas are some works (i) on concurrency by Milner and his followers (Milner, 1999; 2006), (ii) on Reactive Turing Machines (Andersen, Mørk and Sørensen, 1997; Baeten, Luttik and Tilburg, 2013; Van Leeuwen and Wiedermann, 2001; 2006), and (iii) on interaction as a new non-algorithmic computing paradigm (Goldin, Wegner and Smolka, 2006; Wegner, 1997; Wegner and Goldin, 2003). For each of the three identified models, we:

- present the motivation behind it,
- sum up its account for interaction,
- identify its legacy,
- point out issues regarding the understanding of computers qua that model.

We then want to show how these approaches, which belong to a formal approach, cannot provide an answer to the question of what computers are, and for two reasons. On the one hand, these models of computation have focused their attention on whether interactive models are reducible to models of classical computation—par excellence, the Turing machine. Proving (or not) that an interactive property can be formalized as a computable property in the classical Turing's sense does not answer the question of how an interactive property comes into existence and can be the object of execution. On the other hand, and this is a correlate, these models do not propose a basis for a mechanistic explanation of the very possibility of an interactive computing system. With only formal models of interactive computation, we might run the risk of not offering an adequate conceptualization of current computers. Therefore, we end up proposing to take the distinction seriously between *models of computation* and *mechanistic computational explanations*, as presented by Miłkowski (2011; 2014).

1. Milner introduces a distinction between interactional and computational behavior

1.1 Motivation

Milner was the one who introduced the concept of interaction in computer science. He summarized his motivations in a famous Turing Award speech (Milner, 1993). Milner was concerned with the logical foundations of computing inherited from Turing. He was preoccupied with the idea that computing practices had evolved since the birth of computing, notably in terms of architecture. He took seriously the possibility that the logical foundations dating back to the thirties may not match the growing challenges of his time and may require additional concepts.

Milner (2006) pointed out that the logical foundations of computing offered by Turing (1937) were previous to the first physical computers and that computer science is grounded in *logic* and *engineering*. On the engineering side, computer science was inherited from von Neumann's pioneering work (Aspray, 1990; Godfrey and Hendry, 1993). Only one thing could happen at once in an early von Neumann's computer. Nevertheless, there was more to computing than von Neumann's architecture (Backus, 1978; Milner, 2006). A growing interest in dealing with concurrency in the sixties and seventies made sequential programming less warranted. Therefore, to Milner, the logical foundations of computing were to evolve. The main flaw of the early logical foundations was the reduction of computing processes to the concept of an algorithm, which tends to associate computing with mere calculation without taking concurrent activity into account. Because of the evolution of computing engineering practice, Milner questioned whether the logical grounds of computing should evolve as well. Milner's thesis can be put in a nutshell: "this logical foundation has changed a lot since Turing but harks back to him. To be more precise: (i) Computing has grown into informatics—the science of interactive systems; (ii) Thesis: Turing's logical computing machines are matched by a logic of interaction" (Milner, 2006). Consequently, a theory and new language to express concurrent activity were required: "we must find an elementary model which does for interaction what Turing's logical machines do for computation" (Milner, 2006).

The need to define a new computing theory is first displayed through the evolution of computing practice. To sum up, Milner's motivation and focus were the solving of concurrency issues in distributed systems, with the idea that the evolution of computing practices required new formal tools: "Through the 1970s, I became convinced that a theory of concurrency and interaction requires a new conceptual framework, not just a refinement of what we find natural for sequential [algorithmic] computing" (Milner, 1993).

1.2 Account for interaction

Milner introduced the opposition between *interactional* and *computational* behaviour. Introducing the concept of interaction, Milner (1975; 1982; 1983) referred to concurrent message passing between agents. Milner's work coincided with Petri's (1980) new model of concurrent processes, which generally intended to describe concurrency in information systems.[2] To Milner, interaction is more *expressive* than a TM, but it still describes an *effective* procedure. Milner did not assert *equivalence* between an interactive model and a TM, but he introduced the topic (Milner, 1999) and left it unanswered. Four main differences between old (computational) and new (interactional) computing are made striking by Milner. First, in Milner's words, a Turing Machine prescribes a behaviour to be executed. In contrast, new computing requires the description of an information flow between several system components. Second, old computing is characterized by a *hierarchical design* when current practice involves *heterarchical* phenomena in the computing system. Third, in new computing, the designer cannot predict when agents will be triggered or the overall behaviour of the computing system. Fourth, the user is not merely looking for an end result in new computing practice. There is more than a mathematical function to evaluate, as it used to be in old computing. The user instead interacts with the system, and the look for an end result is replaced by continuing interaction. Having taken stock of the evolution of computing practice on the engineering side, Milner examines its consequences on the logic foundations of computing. The pi-calculus and his work on the equivalence with automata, known as bisimulation, achieved this reflection on interactive processes (Milner, 1993; 1999) with a formalism.

---

[2] Concurrency theory emerged from Dick Karp's early work in the 1960s, grew with (Petri, 1980) and later work on transition systems (Glabbeek and Plotkin, 2004; Nielsen, Plotkin and Winskel, 1981), and has now developed into a mature theory of reactive systems (Harel and Pnueli, 1985) with diverse network models (for an overview, see Lee and Sangiovanni-Vincentelli, 1998; Lee and Neuendorffer, 2006)

1.3 Legacy

Milner's work on interaction has become a founding block in automata theory and concurrency theory. It installed the notion of a transition system as the prime mathematical model to represent discrete behavior (Arbach et al., 2015; Baldan, Corradini and Montanari, 2001; Glabbeek and Plotkin, 2004; Nielsen, Plotkin and Winskel, 1981). It also showed that language equivalence was not the correct notion when comparing automata for interactive systems. Instead, it should be replaced by a notion of behavioral equivalence or bisimilarity (Milner, 1999). The pi-calculus has inspired research to derive a language from it. The *Pict* (Pierce and Turner, 2000) programming language is an example.

Milner's work is foundational and served as a reference for anyone after him, reflecting on the need for a new framework dedicated to new emerging computing practices. Milner insists on an essential reminder that we would like to consider. When modelling, the engineering practice matters and is to be articulated with the logical foundations of the model, should it involve elaborating a new framework. Famously, Wegner and Goldin acknowledge that Milner was the first to introduce the idea that classic models of computation were insufficient. They argue that Milner did not state clearly whether computation in concurrent communicating systems (CCS) and the pi-calculus were reducible to Turing machines and algorithms (Wegner and Goldin, 2003). If one goes and looks at Milner's Turing Award Speech, it seems true that classical computation translates into an interactive calculus. However, it is not stated whether any formula in the pi-calculus can be expressed in a classical calculus like the lambda-calculus.

1.4 Issues for an account of current computers

Given the account of current computers that we are looking for, we see two limits in the lessons drawn from Milner. First, we are looking for an explanation of the interactive computing phenomena at stake in a computer. Therefore, the relation between layers of abstraction, from the computational to the physical, is crucial. However, to Milner, the physical layer of the machine is not of much interest, and the calculus of CCS needs to abstract away from the physical. As Milner puts it, informatics is about virtual symbols: "physical systems tend to have permanent physical links; they have fixed structure. But most systems in the informatic world are not physical; their links may be virtual or symbolic" (Milner, 1999). From our perspective, abstracting away from the physical world comes at some cost for an explanation. A complete explanation of a computing system can hardly be provided in details within a single understandable abstraction, since a computing sytems is extremely multi-layered (Lee, 2020; Nisan and Schocken, 2005). Therefore, an explanation of a computing system is necessarily a trade-off between understandability and overwhelming details. As we will flesh out in the last section 4 by referring to Miłkowski's work (Miłkowski, 2011; 2016), a good computational explanation must link the formal story and the blueprint of the computing mechanism. Such articulation is not told in a formal theory of concurrent processes. Second, this first story of interactive systems restricts them to concurrent systems, which is only one dimension of interest when describing what current computers do. There are *at least* two core dimensions left aside: what makes possible timing instructions and the connection between physical processes inside and outside the computing system.

2. Reactive TMs: extending the original model

2.1 Motivation

More recently, a literature domain focused on a "Reactive Turing machine" has emerged (Andersen, Mørk and Sørensen, 1997; Baeten, Luttik and Tilburg, 2012; 2013; Luttik and Yang, 2016; Van Leeuwen and Wiedermann, 2006). It reminds us that the purpose of Turing's *a-machine* model was to propose a formal account of what is *computable by effective means* (algorithmically computable). This formalization was achieved before the realization of the first digital computers. In a way reminiscent of Milner, the question is whether the TM model still fits computing practices decades later. The strategy chosen is to see whether *extensions* of the original TM are sufficient to describe new computing practices and whether the obtained model is still equivalent to a TM. The strategy founds its frame within computability theory and reflects on its scope. This literature domain that proposes extensions of the Turing machine to account for interactive computing systems may be traced back to seminal works on a "Universal reactive machine" (Andersen, Mørk and Sørensen, 1997). In that respect, although pointing at the specificity of interactional behaviour, the main framework still relates to Turing's. Baeten, Luttik, and van Tilburg (Baeten, Luttik and Tilburg, 2013) are looking for a model of interactive computation, extending the classical TM with a process-theoretical notion of interaction related to Milner's previous work. The strategy involves questioning the relationship between such extensions and the Church-Turing thesis. As a reminder, the Church-Turing thesis states that a computable function by effective means is computable by a Turing machine. The community interested in Reactive Turing machines asks the following question: can the Church-Turing thesis also be extended? Van Leeuwen and Wiedermann (2001) focus on the possible extension of the Church-Turing thesis to account for interactive computing: "We will motivate the need for a reconsideration of the classical Turing machine paradigm and formulate an extension of the Church-Turing thesis" (Van Leeuwen and Wiedermann, 2001).

What is at stake is whether the Church-Turing thesis holds given warranted new models of computation: "Is the Church-Turing thesis as we know it still applicable to the novel ways in which computers are now used in modern information technology? Will it hold for the emerging computing systems of the future?" (Van Leeuwen and Wiedermann, 2001). The Church-Turing thesis originally did not entail a claim about computing in general (what computers do and will do) but only about *effective computation*. Therefore, it does not follow that we should ask the Church-Turing thesis for answers on what computing is. Replacing a question about computing with a question about computation is the mark of a specific formal perspective within the frame of computability theory. Understanding *computing* and its evolution from a formal perspective consist of questioning *what can be computed* and seeing if there is another notion of computation than effective computation *in the sense of Church-Turing*.

2.2 Account for interaction

The starting point in the Reactive TM community is a standard current computer designed as a distributed system interacting with an environmental agent: a *site machine*. Starting from this model, the reflection on interaction aims at showing the equivalence between this site machine and a Turing machine augmented by some functions. The conclusion is that a site machine computer computes effectively and yet requires a TM with new functions, thus requiring an extension of Church-Turing's thesis. There are effectively computable functions that TMs, in a strict sense, cannot compute. One crucial dimension that the community wants to account for is particularly relevant to us: "In order to mimic site machines, a Turing machine must have a mechanism that will enable it to model the change of hardware or software by an operating agent" (Van Leeuwen and Wiedermann, 2001). To make interaction with an external agent possible, the model needs to integrate a way of entering new, external, and possibly non-computable information into the machine. This is precisely what oracles do. The authors prefer a more general notion: an *advice function*. The model of a Reactive TM (also called a TM with advice) is considered expressive and definitionally equivalent to an Oracle Turing Machine.

Van Leeuwen and Wiedermann identify three key elements that should be integrated all together within the frame of algorithmic computability: "non-uniformity of programs", "interaction of machines", and "infinity of operation". By the non-uniformity of programs, the authors refer to the fact that current programs on a personal computer are no longer fixed but evolve, are upgraded, and their data remain in memory even when the machine is not running. By interaction, they intend to contrast a TM, where all input data are present before the start of the computing procedure, with a modern computer, where continuous streaming of data via input ports is going on. The third mentioned characteristic, the infinity of operation, refers to distributed and mobile communication systems. These systems are to be seen as dynamic networks of many entities sending and receiving signals in unpredictable ways that are to be synchronized. To accommodate the original TM model, Leeuwen and Wiedermann propose to define "Interactive Turing machines with advice." Integrating an "advice" function amounts to entering new, external, and non-computable information into the machine, which requires using oracles (Balcázar, Díaz and Gabarró, 1995; Rogers, 1987). This way, a TM with advice resembles site machines and I/O automata in being equipped with input and output ports. To the authors, formal tools to support interaction and infinite computations are already available. As for interaction, they refer to already well-known and developed literature on the theory of concurrent processes, the programming of parallel processes, communication protocols, and distributed algorithms. As for infinite computations, Leeuwen and Wiedermann understand them from the language-theoretic viewpoint in the theory of omega-automata (Staiger, 1997; Thomas, 1990).

2.3 Legacy

This approach to extending the Turing machine and the Church-Turing thesis is at the junction between Milner's work and Wegner's (presented in the coming section 3). It makes the junction in that it begs the question of a new paradigm. Milner had not formulated his theory of interaction in such radical terms, but Wegner goes further. The Reactive Turing Machine community asks whether the mentioned required extensions lead to a new computing paradigm: "The experience with present-day computing confronts us with phenomena that are not captured in the scenario of classical Turing machines" (Van Leeuwen and Wiedermann, 2001). The computations carried out on Turing machines with advice are said to be "more powerful" than classic computations on a-machines. The authors insist that this claim does not go against the Church-Turing thesis. To Leeuwen and Wiedermann, like other physical systems (Pour-El, 1999), TMs with advice or oracle Turing machines do not fit the concept of a finite algorithm that can be computed by means of a TM. The conclusion pushes towards a paradigm shift:

> What makes them non-fitting under the traditional notion of algorithms is their potentially endless evolution in time. This includes both interaction and non-uniformity aspects. This gives them the necessary infinite non-uniform dimension that boosts their computational power beyond that of standard Turing machines (Van Leeuwen and Wiedermann, 2001).

The authors ensure that such a paradigm shift does not put into question the original Church-Turing thesis because their proposal for interactive computation does not involve solving undecidable problems (Van Leeuwen and Wiedermann, 2001) using effective computation. The work seems to have served as a pivotal point in structuring the debate on a model of interactive computation around its implications for the Church-Turing thesis. This is evidenced by the objections formulated against Wegner's work which pushes further the concept of interaction and the need for a new paradigm: a proposal of this kind had fallen under objections framed within the theory of computability.

2.4 Issues for an account of current computers

The project is focused on extending the original TM to make it "reactive". The proposed level of abstraction cannot account for the mechanisms that make the proposed extensions possible. We can take a closer look at the type of description presented in this formal framework to account for an interactive scenario:

> The computational scenario of an interactive Turing machine is as follows. The machine starts its computation with empty tapes. It is driven by a standard Turing machine program. At

> each step, the machine reads the symbols appearing at its input ports. At the same time, it writes some symbols to its output ports. Based on the current context, i.e., on the symbols read on the input ports and in the 'window' on its tapes, and on the current state, the machine prints new symbols under its heads, moves its windows by one cell to the left or to the right or leaves them as they are, and enters a new state. Assuming there is a move for every situation (context) encountered by the machine, the machine will operate in this manner forever. Doing so, its memory (i.e., the amount of rewritten tape) can grow beyond any limit. At any time $t > 0$, we will also allow the machine to consult its advice, but only for values of at most t (Van Leeuwen and Wiedermann, 2001).

If we look for a mechanistic explanation of computing, we need some elements to be unpacked beyond a formal account to make sense of the quoted scenario above. For example, we need to account for how reading and writing on the ports are possible. It presupposes that the interactive computing system can wait, pause, and react depending on the arrival or absence of new data. What allows such behavior? It presupposes some mechanisms allowing the system either to be interrupted by environmental processes or to check the new incoming values steadily.[3] In other words, given the initial question ("what is an interactive computer?"), some phenomena cannot be accounted for within the frame of an extended Turing machine. The way oracles work remains at a level of abstraction too remote from the minimal causal blueprint we need for our purpose.

[3] More on these mechanisms and on the limitations of oracles can be found in (Martin, Magnaudet and Conversy, n.d.)

3. Going beyond TMs? Wegner's new paradigm

3.1 Motivation

A strong motivation for Wegner's view on interaction is to overcome the Strong Church-Turing thesis (CTT) that he takes to prevent us from fully admitting a new paradigm in computer science. A paper fleshes out in detail clarifications against the CTT:

> The classical view of computing positions computation as a closed-box transformation of inputs (rational numbers or finite strings) to outputs. According to the interactive view of computing, computation is an ongoing interactive process rather than a function-based transformation of an input to an output. Specifically, communication with the outside world happens during the computation, not before or after it. This approach radically changes our understanding of what computation is and how it is modelled. The acceptance of interaction as a new paradigm is hindered by the Strong Church-Turing Thesis (SCT), the widespread belief that Turing Machines (TMs) capture all computation, so models of computation more expressive than TMs are impossible (Goldin and Wegner, 2008).

In other words, the strong CTT stipulates that a TM could solve all computational problems and could compute anything that any computer can compute. Wegner argues that Turing himself would have denied it, referring to Turing's famous paper (Turing, 1937), as he did not only introduce TMs (calling them automatic machines, or a-machines) but did also introduce choice machines (*c-machines*), extending TMs by allowing a human operator to make choices during the computation. Turing did not view *c-machines* as reducible to

TMs, suggesting other forms of computation might exist. Goldin and Wegner also like to remind us that the CTT applies only to the computation of functions rather than to all computations:

> Function-based computation transforms a finite input into a finite output in a finite amount of time, in a closed-box fashion. By contrast, the general notion of computation includes arbitrary procedures and processes—which may be open, nonterminating, and involving multiple inputs interleaved with outputs (Goldin and Wegner, 2008).

For the sake of clarity, Goldin and Wegner propose to formulate the assumptions of the CTT in their proper formulation free of extrapolation (Goldin and Wegner, 2008) explicitly:

- i. "All algorithmic problems are function-based."
- ii. "All function-based problems can be described by an algorithm."
- iii. "Algorithms are what early computers used to do."
- iv. "TMs serve as a general model for early computers."
- v. "TMs can simulate any algorithmic computing device."
- vi. "TMs cannot compute all problems, nor can they do everything that real computers can do."

One reason the strong CTT is "impossible" (Eberbach, Goldin and Wegner, 2004) is that no computable function would determine, given some finite amount of a priori information, all the real-world factors that are necessary to ensure the safe arrival of a car at its destination. An assertion to the contrary would endow TMs with the power to predict the future. Therefore, Wegner introduced *interaction* as a new paradigm, based on an empiricist approach (Wegner, 1995), to broaden algorithmic problem-solving. The reason is that Wegner and his followers take computing machines to be about physical processes, chaotic in nature (Siegelmann, 1995), requiring demanding precision to be controlled (Hartmanis, 1994). Superposed layers of abstractions allow us to describe and control those physical and chaotic computing machines. The challenge is then to bridge the gap between all those layers of abstraction, starting with the lowest physical level. A typical problem we want to solve with computers but not computable in the classic sense would be, e.g., the problem of driving home:

> the problem of driving home from work is computable—by a control mechanism, as in a robotic car, that continuously receives video input of the road and actuates the wheel and brakes accordingly. This computation, just as that of operating systems, is interactive, where input and output happen during the computation, not before or after it (Goldin and Wegner, 2008).

Goldin and Wegner argue that such a notion of computation does find its counterpart neither in the theory of computation nor in the concurrency theory. The motivation that goes hand in hand with this discussion against the strong CTT is a reflection on algorithms and the scope of algorithmic problem-solving. Knuth has given a classic definition for algorithms: “An algorithm has zero or more inputs, i.e., quantities which are given to it initially before the algorithm begins” (Knuth, 1968). Following a recipe (Knuth, 1968), for example, does not actually involve algorithmic problem-solving. To know how to mix the ingredients properly, one needs to adapt to dynamic variables and feedback, such as humidity conditions and the progressive evolution of the texture of the paste that are not pre-given values before execution. To Wegner, that kind of feedback does not belong to the function-based mathematical worldview. The problem of driving home from work, like baking following a recipe, is also among those problems that Knuth meant to exclude from his definition.

3.2 Account for interaction

This leads us to Wegner's account for interaction:

> Computational problem solving requires open testing of assertions about engineering problems beyond closed-box mathematical function evaluation. Therefore, we have proposed interactive computing as an empiricist model that expands computational problem solving from algorithmic TM models and functional input-output to broader concepts of interleaved dynamic streams and observable interaction with the environment (Wegner and Goldin, 2006).

In Wegner's perspective, interactions are more powerful than TMs with finite initial inputs. TMs with oracles and unbounded (dynamically extensible) input streams model more accurately interactive systems than traditional Turing machines. Interactive systems react dynamically to external events. They are also related to the passage of external time. By delaying the binding time of inputs so that they can occur during the computation (rather than only at the beginning) and modelling reactive processes (Manna and Pnueli, 1992) by infinite computations (Thomas, 1990), the modelled entities are extended from algorithms to persistent objects and concurrent processes (Milner, 1999).

Wegner wonders if Milner himself avoided questioning whether the computation in CCS and the pi-calculus went beyond Turing machines and algorithms (Wegner and Goldin, 2003). The question could remain whether Wegner takes interaction as a super-calculus/super-algorithm or as a radical shift from TMs. In other words, to what extent is "interaction more powerful than algorithm" (Wegner, 1997)? In fact, Wegner's claim is sharp. In contrast with Milner, Wegner's focus is not on concurrency between computing processes. Instead, he focuses on the complexity of the triggering of external events outside the machine: "Interactive systems are grounded in an external reality both more demanding and richer in behaviour than the rule-based world of non-interactive algorithm" (Wegner, 1997). He strikes the difference between closed and opened systems, the latter being impossibly wholly described. This impossibility makes interactive systems mathematically problematic: they lack completeness.

> The comfortable completeness and predictability of algorithms is inherently inadequate in modelling interactive computing tasks and physical systems. The sacrifice of completeness is frightening to theorists who work with formal models like Turing machines [. . .]. But incomplete behaviour is comfortably familiar to physicists and empirical model builders. Incompleteness is the essential ingredient distinguishing interactive from algorithmic models of computing and empirical from rationalist models of the physical world (Wegner, 1997).

From this, Wegner concludes that computing systems should not be thought of as algorithms but as *interfaces*, *views*, and *modes of use*, definable as behaviours to be specified. Consequently, an ontological question is also at stake: in what terms should the external world be modelled: as atomic objects and events? As processes and flow? Formally, Wegner's account of interaction has led to the development of Persistent Turing machines (PTMs), a model of sequential computation, and the result that multi-stream interaction machines (MIMs) are more expressive than sequential interaction machines (SIMs) (Goldin, 2000; Goldin, Smolka et al., 2004). Wegner and Goldin trace back the idea that interaction is not expressible by or reducible to algorithms at the closing conference on the 5th-generation computing project in the context of logic programming. Reactiveness of logic programs, realized by the commitment to a course of action, was shown to be incompatible with logical completeness (Wegner and Goldin, 1999).

3.3 Legacy

Wegner's work has been criticized, the main objection being that interaction machines can be proved equivalent to TMs. The objections are focused on the defence of the Church-Turing thesis (Cockshott and Michaelson, 2007; Prasse and Rittgen, 1998), and assume that introducing an interactive computing paradigm denies the results of Church and Turing's work. But this assumption cannot be taken for granted: no one denies that TMs and lambda calculus account for *effective computation*. Both formalisms define the intuitive notion of an *algorithm*. The Church-Turing thesis will only be shaken once someone presents an alternative formal account of an effective procedure. Due to semantic ambiguities, some have interpreted Wegner's work as challenging the Church-Turing thesis. First, Wegner characterizes interaction as *more powerful* than algorithms and TMs. What "powerfulness" precisely refers to is unclear. We will say more about this in the next section (section 4).

Second, there seems to be another semantic ambiguity or alleged identity between "computing" and "computation": "Wegner (and Eberbach) say that it is impossible to describe all computations by algorithms. Thus, they do not accept the classic equation of algorithm and effective computation" (Cockshott and Michaelson, 2007). In the former quoted sentence, a core assumption uses interchangeably "computation" and "computing". But Wegner means that it is impossible to describe everything in computing by algorithms. By "computing", he is referring to what computers do broadly, not to Turing computation in a narrow sense. Therefore, the conclusion made in the quoted sentence does not follow: the identity between an effective computation and an algorithm is not put into question by Wegner.

3.4 Issues for an account of current computers

We are interested in the way Wegner broadens the notion of *interaction*. It is not strictly referring to communicating processes within a computing machine. Possible complex interactions with the environment and the dynamic between inputs and outputs during execution are considered. However, although debunking the focus of the CTT by stating that interaction is more *powerful* and *expressive* than algorithms, Wegner's work is enclosed in a field of discussion framed by the theory of computability. Furthermore, we still need a way of describing the very mechanisms we are interested in to be provided with a mechanistic account of current computers. This is no surprise since Wegner's work aims primarily to reflect on the theoretical limits of classic mathematical tools, e.g., on notions like *completeness*.

4. Why the interactive models identified do not provide us with an answer

We have reviewed the conceptualization of interactive systems in theoretical computer science. We want to defend that these approaches cannot answer the epistemic question asked by philosophers about what current computers are. There are two reasons for this. First, as we have seen, these conceptualizations focus on whether a formal model for interaction is irreducible to a Turing machine and, if so, whether this is a threat to the Church-Turing thesis. This deprives

	Interaction as concurrent communicating systems	**Interaction as extended Turing Machines**	**Interaction as a new paradigm**
Motivation	Provides new logical foundations to fit new engineering challenges, especially concurrency	Extends the TM model to account for interactive devices	Debunks the strong Church-Turing thesis Discusses the scope of algorithmic solving Prones the need for a new computing paradigm
Account for interaction	Information flow Heterarchical design No complete prediction about overall behavior No end-result Process calculi	External data needed during computation Non-uniformity of programs Interaction with agents Infinity of operations Interactive machines are TMs with advice	Computers have rich interaction with the environment during computing execution, but this processing is not merely algorithmic
Uses and criticisms	First conceptualization of interaction Legacy for automata theory	Inspires the need for a new paradigm Puts at the forefront the Church-Turing thesis	Controversy about the powerfulness of the TM
Issues for an account of interactive computing	Definition of interaction restricted to specific properties: concurrency and communication	Formal oracles cannot account for the physical possibility of entering new data	Issues about powerfulness and expressiveness constrict the debate in the realm of computability theory

Table 1: Sum-up: an overview of explicit theories of interactive computing systems in theoretical computer science

us of a level of description to explain the mechanisms that allow a computing system to be interactive. We propose to detail here in section 4 the problems posed by the debate on reducibility. We end the section by mentioning a distinction currently offered in the literature that highlights the limits of a formal approach. It is a distinction, mostly worked by Miłkowski, opposing *mechanistic computational explanation* and *model of computation*.

4.1 Unclear stance towards interaction and Turing reducibility

The first problem with the focus on Turing reducibility in the accounts for interaction is that the stance is not always clear-cut. Milner's work leaves us with the following question: to what extent are the new "logical foundations" for interaction distinct from the classic framework? Irreducibility is not stated in the speech for the Turing Award. There is a simple translation of lambda-calculus into pi-calculus, which is faithful to computational behaviour. Thus, pi-calculus supports functional programming at a higher level of explanation. However, it is unclear whether any behaviour expressed in the pi-calculus can be translated into a classic calculus. In a more recent book, *The Space and Motion of communicating agents* (2009), Milner introduces bigraphs as another formalism for interactive systems. Bigraphs are proven to have the same expressiveness as Turing machines. It looks like Milner proposes to revise the principle of Occam's razor and praise the plurality of formalisms, models, and frames of explanation:

> I reject the idea that there can be a unique conceptual model, or one preferred formalism, for all aspects of something as large as concurrent computation, which is in a sense the whole of our subject — containing sequential computing as a well-behaved special area. We need many levels of explanation: many different languages, calculi, and theories for the different specialisms (Milner, 1993).

It looks like interaction is the new "basic notion":

> Now, what are the new particles, parts of speech, or elements which allow one to express interaction? They lie at the same elementary level as the operation of a Turing machine on its tape, but they differ. For much longer than the reign of

> modern computers, the basic idiom of algorithm has been the asymmetric, hierarchical notion of operator acting on operand. But this does not suffice to express interaction between agents as peers; worse, it locks the mind away from the proper mode of thought (Milner, 2006).

As for the work on extended Turing Machines, does it involve that interaction is something else, something irreducible to TMs? Does interaction amount to a classical model of computation with extended computational power? The latter claim is possibly controversial by revising the Church-Turing thesis. In the end, it looks like interaction is still understood in reference to the classical framework (our italics): "examples of interactive [. . .] indicate that the classical Turing machine paradigm should be *revised (extended)* in order to capture the forms of computation that one observes in the systems and networks in modern information technology" (Van Leeuwen and Wiedermann, 2000).

Criticisms against Wegner show that the criterion of powerfulness is ambiguous when evaluating a model for a computing system. Does powerfulness refer to computational power, involving that an interactive model can express uncomputable functions in Turing's sense? Or does it refer to the expression of more phenomena? Such ambiguity could support some misunderstanding about interaction.

In any case, the literature review on explicit theories of interaction shows that arguments about the powerfulness and equivalence of the interactive and classic models systematically arise.

4.2 Powerfulness and expressiveness: possible ambiguities

Ambiguities around the concepts of powerfulness and expressiveness likely make the debate need clarification. Indeed, there are at least two ways of understanding them. In any case, the powerfulness of a model refers to its expressiveness, which is a semantic property. Expressiveness refers to *what can be expressed* by a given model. If one thinks of a model as a formal language, let us say that expressiveness relates to all the possible sentences one can make in that language.

In a first sense, powerfulness and expressiveness can be understood strictly within computability theory. In that case, the two notions are used when evaluating a mathematical framework supporting the formalization of semantics. What is called "powerfulness" refers to *computational power*, and expressiveness refers to a formal criterion evaluating which functions can be expressed. *Turing completeness* is then a possible evaluation criterion for expressiveness, for instance.

Let us say that among the things that could be expressed in a model are functions (*set A*) and other things than functions (*set B*). Within each set, some sets include more than others. Within *set A*, the set of hypercomputations is more expressive than the set of computable functions since it includes the uncomputable ones. That is a way to be more expressive: expressing more functions. However, framing expressiveness and powerfulness as possibly only about computable functions would seem odd to engineers and computer scientists familiar with other formalisms than those related to computability theory. Nevertheless, objections about interaction theories frame the debate in reference to computability theory.

In a second sense, one can consider the powerfulness and expressiveness of a model *outside the strictly formal computability framework*. Since a model must represent, according to specific objectives, a phenomenon of reality or, say, a system, we can understand the powerfulness of a model as a good match between the model and what is modeled.

Therefore, in that broader sense, a model is expressive, given some purpose, if and only if it describes all phenomena required for that given purpose. In that case, the value of the model and concerns about its expressiveness depend on stated goals. From an engineering perspective, for example, a model is valuable to the extent that it allows engineers to think of future systems design easily. In this case, the value of the model could be evaluated, e.g., in terms of usability (effectiveness, efficiency, and satisfaction (ISO, 2018)). From a scientific perspective, the aim is to make good predictions about a system. The two perspectives are rarely used in isolation since good engineering design requires some science, and good science often relies today on some engineering (Lee, 2017). From the perspective of the philosophy of science and given scientific explanation standards, a good model for a phenomenon rightly describes the mechanisms at stake (Glennan, 2002; Machamer, Darden and Craver, 2000; Miłkowski, 2016). Of course, other possible values for models, from other perspectives, could be found.

To go back to Wegner (Goldin, 2000; Wegner, 1995; 1997; 1998), we argue that this distinction between a narrow and broad sense of expressiveness clarifies criticisms made against him.

In a broad sense, one can interpret Wegner's new paradigm as follows: Wegner considers his interactive model more expressive than a TM by having his model describe *other things than Turing computations*. Wegner's model could then describe more phenomena than a TM. It would not go against the Church-Turing thesis, which remains valid to account for algorithmic problem-solving through effective procedures.

But in a narrow sense of expressiveness, one can interpret (wrongly, we think) the possibility of a new paradigm as follows. Wegner and the tenants of Reactive Turing Machines could think of their interactive model as more expressive than a TM, allowing their model to execute *more functions*, even some of them being uncomputable functions in the sense of the Church-Turing thesis, solving the halting problem. In that case, the claim would indeed be controversial. The bold claim would be the following: a TM is not only providing an account for algorithmic problem-solving through effective procedures but it could also be extended to account for other non-algorithmic processes, solving the uncomputable. Interaction would be some super-calculus, extending the calculative power of the original TM to account for interaction. It would be satisfactorily modeled with a TM, only given more calculation power. It would go down the track of Accelerating Machines or Super-Turing Machines, able to calculate more than Turing's computable functions Copeland, 2002; Copeland and Shagrir, 2011; MacLennan, 2009.

We argue that a theory of interaction does not need to embrace the hypercomputation view. Part of an interaction model could be reduced to the classical TM, but some extra elements needed to express interaction cannot be reduced to an a-machine. That does not mean interactive models have super computational power to solve undecidable problems. It simply means interactive systems do things that a TM cannot do. It is possible to admit they do other things *without implying they compute uncomputable functions*.

4.3 What formal models of computation cannot do: providing a mechanistic explanation of computing

So, do we have a good theory about interactive computers? Do we understand what they are? A natural and common way to go is to reduce the question of what interactive computers are to what interactive computation is. Initially, the first models of computation emerged through computability theory. They served as answers to an abstract mathematical problem, namely the formalization of the intuitive notion of an algorithm. They had nothing to say about computers, as computers did not even exist at the time. Since the computability era, models of computation like the Turing Machines have been exported outside their original scope to serve as a basis for theoretical computer science. Some models of computation (Turing Machines) have even helped to reflect on computers. It is no surprise since computers were thought to be precisely the kind of machines that implement computations. Models of computation have then evolved, accounting for new desired properties to be integrated within the classical framework. In computer science, what makes a model of computation valuable is related to the formal properties it expresses. Once those formal properties are at hand, they allow further procedures to be acted upon them, especially system verification and certification. In the end, models of computation serve as tools used to support and verify a system's design. These models belong to a particular abstraction level: they do not intend to model the system as a whole and the way it works. They focus on verifiable properties, upon which proofs that guarantee the outputs of the system are built. Verifying formal properties is different from investigating why the system behaves the way it does. They are two different tasks. The former task (verification) belongs to applied mathematics. It describes abstract computations through formal models by focusing on specific properties. The latter (understanding computing behaviour) is the question the philosopher begs when asking what a computer is. It requires something else than task-oriented formalizations of properties abstracted away from any physical mechanism. Philosophers of computing need to make sense of the overall behaviour, which requires combining other levels of abstraction. The reason is that an account of computing behaviour calls upon the description of how computation can be carried out: in other words, it requires the description of execution on some computer architecture. Computations and their models belong to a level of abstraction independent from implementation detail. Computations, as already coined, are "medium-independent" (Klein, 2020). On the contrary, to have a model of some execution belongs to a lower level of abstraction, where minimal references to the devices that allow the execution is made. There is no need to dig into fine-grained implementation details to make sense of computing behaviour in mechanistic terms.

The formal debate on model equivalence and powerfulness leaves us needing more building blocks to figure out an explanation for interactive computing: what makes it possible, and what mechanisms support it? A helpful distinction here capturing why we lack the right tools is a recent distinction in the literature between models of computation (formal) and mechanistic explanations of computing. It deserves attention in the context of understanding interactive computing. Questioning model equivalence belongs to formal mathematics; it does not aim at providing a mechanistic account of the computing phenomena. Interactive models of computation propose an upper layer of abstractions to formalize specific properties but do not hint at how interactive computation is carried out. We suggest we need to adopt a different explanatory focus, departing from the perspective adopted by models of computation and understanding *how interactive computation can be executed*.

Such lessons have just started to be drawn. They have motivated, for example, distinctions between computational models and computational explanations (Klein, 2020) or between models of computation and computational mechanisms (Miłkowski, 2014)). The lesson drawn is that formal models of computing systems do not provide us with the appropriate and complete level of description to build an explanation, which is expected to identify the relevant mechanisms at stake. More precisely, an explanation for computing phenomena requires bridging a high-level description of a computation and its blueprint (Miłkowski, 2011; 2016). The approach is based on the standard of mechanistic explanation in science, coupled with the idea that a computational process is intrinsically mechanistic:

> Computational explanations, according to the mechanistic account are constitutive mechanistic explanations: they explain how a mechanism's computational capacity is generated by the orchestrated operation of its component parts. To say that a mechanism implements a computation is to claim that the causal organization of the mechanism is such that the input and output information streams are causally linked and that this link, along with the specific structure of information processing, is completely described. (Miłkowski, 2014).

If one is looking for a mechanistic explanation of a computing process, Miłkowski argues that a model of computation may be insufficient. The reason something is missing is that a model of computation is not strongly equivalent to a mechanism:

> There are two ways in which computational models may correspond to mechanisms: first, they may be *weakly equivalent* to the explanandum phenomenon, in that they only describe the input and output information, or *strongly equivalent*, when they also correspond to the process that generates the output information. (Miłkowski, 2016)

The difference between strong and weak equivalence captures a difference in causal completeness. The formal models of computation are on the side of models that are weakly equivalent to a mechanism: "formal models cannot function as complete causal models of computers. For example, to repair an old broken laptop, it is not enough to know that it was (idealizing somewhat) formally equivalent to a universal Turing machine." (Miłkowski, 2016). An example helps to flesh out the need for such distinction and turns again to the Turing machine:

> Turing machines were not invented to be implemented physically at all, but some people still build them for fun. [...] Imagine a physical instantiation of a trivial logical negation Turing machine, built of, say, steel and rubber and printing symbols on paper tape. Its alphabet of symbols consists of "F" and "T." If the machine finds "T" on its tape, it rewrites it to "F" and halts; if it finds "F," it rewrites it to "T" and halts. Let us suppose that the machine's head is so old and worn out that it tears the paper tape during the readout. As a result, no symbol will appear. [...] Only when we describe the Turing machine literally, as a causal system that has a particular causal blueprint (engineering specifications of how it is built), can we causally predict such a breakdown. [...] Why are breakdowns and malfunctions so important? They help us discover the causal complexity of the system. [...] an abstract model of computation will not predict all the possible outcomes of the breakdown, as it abstracts away from a number of the system's causal characteristics. So it will not tell us what is going to happen with the head; it will only say that the computation will no longer be correct. (Miłkowski, 2011)

Thus, Miłkowski invites us to consider a new project in the philosophy of computing: "it is necessary to acknowledge the causal structure of physical computers that is not accommodated by the models used in computability theory" (Miłkowski, 2011). To the best of our knowledge, such a project to account for interactive computing has still not been carried out to flesh out the mechanisms at stake. If philosophers of computing were to proceed in that direction, two criteria for a good explanation of a computer proposed by Miłkowski could offer some guidance. First, such an explanation should be complete, in the sense of a complete causal model where causally relevant parts and operations are specified (Miłkowski, 2014). Second, a good explanation for computing should explain the competence of the system: "By providing the instantiation blueprint of the system, we explain the physical exercise of its capacity, or competence, abstractly specified in the formal model" (Miłkowski, 2014). For example, it would be necessary to be able to explain in mechanistic terms what the behaviour of an oracle corresponds to. This would be equivalent to explaining which mechanisms allow data arrival, launching, interrupting, or pausing machine processes.

Conclusion

We started from the need to update a question in the philosophy of computing: *what is a computer*? Today's computers are highly interactive, so the question can be rephrased more precisely: *what is an interactive computing system*? It is common to understand computers in terms of existing models of computation, hypothesizing that a computer is primarily a machine that carries out computation. Therefore, the working hypothesis has traditionally answered the initial question by asking what computation is. As already noted, this shift should not be taken for granted. There are, however, and the length of the paper does not allow it, historical and epistemological reasons for this shift that have been described and discussed (Daylight, 2014; Haigh and Priestley, 2020; Mol, 2018). We have chosen in this paper to ask ourselves if the shift is relevant in the case of interactive computing: do we understand what an interactive computer is by questioning the formal models proposed in theoretical computer science for interaction? Our literature review shows that there are better paths. There are two reasons for this. First, the conceptualization of interactive systems in theoretical computer science has focused on their comparison with the Turing machine (and sometimes other classical models), putting forward formal questions about powerfulness and equivalence of models that do not clarify the singularity of interactive systems from an epistemic point of view (rather than formal). There is an inherent difficulty in looking for an explanation of a computing phenomenon in a formal model: it needs more bricks to describe the mechanisms at stake, at a level of abstraction operating the junction between a high-level formalism and the blueprint. This work does not lead us to an aporia but to a research program in the philosophy of computing: we must produce the right level of explanation for interactive computing.[4] This implies an identification of the mechanisms at play that make possible the interaction between processes within the computing machine (whether there are to be thought of as physical or computational processes, or a mix of both[5]) and the environment. The components of such a mechanism are to be identified and described at a level of abstraction that allows a satisfactory reference to the implementation.

---

[4] More considerations on such a research program are fleshed out in a forthcoming paper (Martin, Magnaudet and Conversy, n.d.)

[5] The distinction between computational and physical processes is out of the scope of this paper but more on this can be found e.g., in (Kycia and Niemczynowicz, 2020).

Bibliography

Andersen, H.R., Mørk, S. and Sørensen, M.U., 1997. A universal reactive machine. In: A. Mazurkiewicz and J. Winkowski, eds. *Concur '97: concurrency theory. concur 1997* [Online]. Vol. 1243, Lecture notes in computer science. Berlin; Heidelberg: Springer, pp.89–103. https://doi.org/10.1007/3-540-63141-0_7.

Arbach, Y., Karcher, D., Peters, K. and Nestmann, U., 2015. Dynamic causality in event structures. In: S. Graf and M. Viswanathan, eds. *Formal techniques for distributed objects, components, and systems* [Online]. Vol. 9039, Lecture notes in computer science, pp.83–97. https://doi.org/10.1007/978-3-319-19195-9_6.

Aspray, W., 1990. *John von neumann and the origins of modern computing*. Cambridge, MA: MIT Press.

Backus, J., 1978. Can programming be liberated from the von neumann style? *Communications of the ACM*, 21(8), pp.613–641.

Baeten, J.C., Luttik, B. and Tilburg, P.V., 2012. Turing meets milner. *Concur'12: proceedings of the 23rd international conference on concurrency theory* [Online], Lecture notes in computer science. Berlin; Heidelberg: Springer, pp.1–20. https://doi.org/10.1007/978-3-642-32940-1_1.

Baeten, J.C., Luttik, B. and Tilburg, P.V., 2013. Reactive turing machines. In: O. Owe, M. Steffen and J. Telle, eds. *Fct 2011: fundamentals of computation theory* [Online], Lecture notes in computer science. Berlin; Heidelberg: Springer, pp.348–359. https://doi.org/10.1016/j.ic.2013.08.010.

Balcázar, J.L., Díaz, J. and Gabarró, J., 1995. *Structural complexity i.* [Online]. Second Edition, *Texts in Theoretical Computer Science. An EATCS Series*. Berlin; Heidelberg: Springer. https://doi.org/10.1007/978-3-642-97062-7.

Baldan, P., Corradini, A. and Montanari, U., 2001. Contextual petri nets, asymmetric event structures, and processes. *Information and Computation* [Online], 171(1), pp.1–49. https://doi.org/10.1006/inco.2001.3060.

Basman, A., Tchernavskij, P., Bates, S. and Beaudouin-Lafon, M., 2018. An anatomy of interaction: co-occurrences and entanglements. *Conference* *companion of the 2nd international conference on art, science, and engineering of programming* [Online]. Association for Computing Machinery, pp.188–196. https://doi.org/10.1145/3191697.3214328.

Beaudouin-Lafon, M., 2006. Human-computer interaction. In: D. Goldin, S.A. Smolka and P. Wegner, eds. *Interactive computation: the new paradigm* [Online]. Berlin; Heidelberg: Springer, pp.227–254. https://doi.org/10.1007/3-540-34874-3_10.

Bielecki, A., 2019. *Models of neurons and perceptrons: selected problems and challenges* [Online]. Vol. 770, *Studies in Computational Intelligence*. Springer International Publishing. https://doi.org/10.1007/978-3-319-90140-4.

Cockshott, P. and Michaelson, G., 2007. Are there new models of computation? reply to wegner and eberbach. *Computer Journal* [Online], 50(2), pp.232–247. https://doi.org/10.1093/comjnl/bxl062.

Copeland, B.J., 2002. Hypercomputation. *Minds and Machines* [Online], 12, pp.461–502. https://doi.org/10.1023/A:1021105915386.

Copeland, B.J. and Shagrir, O., 2011. Do accelerating turing machines compute the uncomputable? *Minds and Machines* [Online], 21, pp.221–239. https://doi.org/10.1007/s11023-011-9238-y.

Daylight, E.G., 2014. A turing tale. *Communications of the ACM* [Online], 57(10), pp.36–38. https://doi.org/10.1145/2629499.

Dearden, A.M. and Harrison, M.D., 1997. Abstract models for hci. *International Journal of Human Computer Studies* [Online], 46(1), pp.151–177. https://doi.org/10.1006/ijhc.1996.0087.

Dodig-Crnkovic, G., 2011. Significance of models of computation, from turing model to natural computation. *Minds and Machines* [Online], 21, pp.301–322. https://doi.org/10.1007/s11023-011-9235-1.

Eberbach, E., Goldin, D. and Wegner, P., 2004. Turing's ideas and models of computation. In: C. Teuscher, ed. *Alan turing: life and legacy of a great thinker* [Online]. Berlin; Heidelberg: Springer, pp.159–194. https://doi.org/10.1007/978-3-662-05642-4_7.

Glabbeek, R.V. and Plotkin, G., 2004. Event structures for resolvable conflict. *29th international symposium on mathematical foundations of computer* [Online]. Vol. 3153, Lecture notes in computer science (including subseries lecture notes in artificial intelligence and lecture notes in bioinformatics). Berlin; Heidelberg: Springer, pp.550–561. https://doi.org/10.1007/978-3-540-28629-5_42.

Glennan, S., 2002. Rethinking mechanistic explanation. *Philosophy of Science* [Online], 69(S3), pp.342–353. https://doi.org/10.1086/341857.

Godfrey, M.D. and Hendry, D.F., 1993. The computer as von neumann planned it. *IEEE Annals of the History of Computing* [Online], 15(1), pp.11–21. https://doi.org/10.1109/85.194088.

Goldin, D., 2000. Persistent turing machines as a model of interactive computation. In: K. Schewe and B. Thalheim, eds. *Foundations of information and knowledge systems. foiks 2000* [Online]. Vol. 1762, Lecture notes in computer science. Springer Berlin Heidelberg, pp.116–135. https://doi.org/10.1007/3-540-46564-2_8.

Goldin, D., Smolka, S.A., Attie, P.C. and Sonderegger, E.L., 2004. Turing machines, transition systems, and interaction. *Information and Computation* [Online], 194(2), pp.101–128. https://doi.org/https://doi.org/10.1016/j.ic.2004.07.002.

Goldin, D. and Wegner, P., 2008. The interactive nature of computing: refuting the strong church-turing thesis. *Minds and Machines* [Online], 18, pp.17–38. https://doi.org/10.1007/s11023-007-9083-1.

Goldin, D., Wegner, P. and Smolka, S.A., 2006. *Interactive computation: the new paradigm* [Online]. Berlin; Heidelberg: Springer. https://doi.org/10.1007/3-540-34874-3.

Haigh, T. and Priestley, M., 2020. Historical reflections von neumann thought turing's universal machine was 'simple and neat.' but that didn't tell him how to design a computer. *Communications of the ACM* [Online], 63(1), pp.26–32. https://doi.org/10.1145/3372920.

Harel, D. and Pnueli, A., 1985. On the development of reactive systems. In: K. Apt, ed. *Logics and models of concurrent systems* [Online]. Vol. 13, Nato asi series. Berlin; Heidelberg: Springer, pp.477–498. https://doi.org/10.1007/978-3-642-82453-1_17.

Hartmanis, J., 1994. Turing award lecture on computational complexity and the nature of computer science. *Communications of the ACM* [Online], 37(10), pp.37–43. https://doi.org/10.1145/194313.214781.

Hornbaek, K. and Oulasvirta, A., 2017. What is interaction? *Proceedings of the 2017 chi conference on human factors in computing systems* [Online]. New York, NY: Association for Computing Machinery, pp.5040–5052. https://doi.org/10.1145/3025453.3025765.

ISO, 2018. Ergonomics of human-system interaction. part 11: usability: definitions and concepts. *ISO 9241-11:2018*.

Klein, C., 2020. Polychrony and the process view of computation. *Proceedings of the 2018 biennial meeting of the philosophy of science association. part ii* [Online]. Vol. 87, 5, pp.1140–1149. https://doi.org/10.1086/710613.

Knuth, D.E., 1968. *The art of computer programming, vol.1: fundamental algorithms*. Boston; etc.: Addison-Wesley.

Kycia, R. and Niemczynowicz, A., 2020. Information and physics. *Philosophical Problems in Science (Zagadnienia Filozoficzne w Nauce)* [Online], (69), pp.237–252. Available at: <https://zfn.edu.pl/index.php/zfn/article/view/513>.

Lee, E.A., 2017. Fundamental limits of cyber-physical systems modeling. *ACM Transactions on Cyber-Physical Systems* [Online], 1(1), pp.1–26. https://doi.org/10.1145/2912149.

Lee, E.A., 2020. *Plato and the nerd* [Online]. MIT Press. https://doi.org/10.7551/mitpress/11180.001.0001.

Lee, E.A. and Neuendorffer, S., 2006. Concurrent models of computation for embedded software. *System-on-Chip: Next Generation Electronics* [Online]. https://doi.org/10.1049/PBCS018E_ch7.

Lee, E.A. and Sangiovanni-Vincentelli, A., 1998. A framework for comparing models of computation. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems* [Online]. https://doi.org/10.1109/43.736561.

Luttik, B. and Yang, F., 2016. On the executability of interactive computation. In: A. Beckmann, L. Bienvenu and N. Jonoska, eds. *Pursuit of the universal. cie 2016* [Online]. Vol. 9709, Lecture notes in computer science. Cham: Springer, pp.312–322. https://doi.org/10.1007/978-3-319-40189-8_32.

Machamer, P., Darden, L. and Craver, C.F., 2000. Thinking about mechanisms. *Philosophy of Science* [Online], 67(1), pp.1–25. https://doi.org/10.1086/392759.

MacLennan, B., 2003. Transcending turing computability. *Minds and Machines* [Online], 13, pp.3–22. https://doi.org/10.1023/A:1021397712328.

MacLennan, B., 2009. Super-turing or non-turing? extending the concept of computation. *International Journal of Unconventional Computing*, 5(3-4), pp.369–387.

Manna, Z. and Pnueli, A., 1992. *The temporal logic of reactive and concurrent systems* [Online]. Springer. https://doi.org/10.1007/978-1-4612-0931-7.

Martin, A., Magnaudet, M. and Conversy, S., n.d. *Minds and Machines*.

Milner, R., 1975. Processes: A Mathematical Model of Computing Agents. In: H.E. Rose and J.C. Shepherdson, eds. *Studies in Logic and the Foundations of Mathematics* [Online]. Vol. 80, *Logic Colloquium '73*. Elsevier, pp.157–173. https://doi.org/10.1016/S0049-237X(08)71948-7.

Milner, R., 1982. Four combinators for concurrency. *Proceedings of the annual acm symposium on principles of distributed computing* [Online], Podc '82. New York, NY: Association for Computing Machinery, pp.104–110. https://doi.org/10.1145/800220.806687.

Milner, R., 1983. Calculi for synchrony and asynchrony. *Theoretical Computer Science* [Online]. https://doi.org/10.1016/0304-3975(83)90114-7.

Milner, R., 1993. Elements of interaction: turing award lecture. *Communications of the ACM* [Online], 36(1), pp.78–89. https://doi.org/10.1145/151233.151240.

Milner, R., 1999. *Communicating and mobile systems: the π-calculus*. Cambridge University Press.

Milner, R., 2006. Turing, computing and communication [Online]. In: ed. by D. Goldin, S.A. Smolka and P. Wegner. Berlin; Heidelberg: Springer, pp.1–8. https://doi.org/10.1007/3-540-34874-3_1.

Milner, R., 2009. *The space and motion of communicating agents*. Cambridge: Cambrdige University Press. https : / / doi . org / 10 . 1017 / CBO9780511626661.

Miłkowski, M., 2011. Beyond formal structure: a mechanistic perspective on computation and implementation. *Journal of Cognitive Science*, 12, pp.359–379.

Miłkowski, M., 2014. Computational mechanisms and models of computation. *Philosophia Scientiae* [Online], 18(3), pp.215–228. https://doi.org/10.4000/philosophiascientiae.1019.

Miłkowski, M., 2016. A mechanistic account of computational explanation in cognitive science and computational neuroscience. In: V.C. Müller, ed. *Computing and philosophy: selected papers from iacap 2014* [Online]. Vol. 375. Springer International Publishing, pp.191–205. https://doi.org/10.1007/978-3-319-23291-1_13.

Mol, L.D., 2018. Turing machines. *Stanford Encyclopedia* [Online]. Available at: <https://plato.stanford.edu/entries/turing-machine/>.

Myers, B., 1994. Challenges of hci design and implementation. *Interactions* [Online]. https://doi.org/10.1145/174800.174808.

Nielsen, M., Plotkin, G. and Winskel, G., 1981. Petri nets, event structures and domains, part i. *Theoretical Computer Science* [Online], 13(1), pp.85–108. https://doi.org/10.1016/0304-3975(81)90112-2.

Nisan, N. and Schocken, S., 2005. *The elements of computing systems. building a modern computer from first principles*. Cambridge: MIT Press.

Petri, C., 1980. Introduction to general net theory. In: W. Brauer, ed. *Net theory and applications* [Online]. Vol. 84, Lecture notes in computer science. Berlin; Heidelberg: Springer. https://doi.org/10.1007/3-540-10001-6_21.

Piccinini, G., 2008. Computers. *Pacific Philosophical Quarterly* [Online], 89(1), pp.32–73. https://doi.org/10.1111/j.1468-0114.2008.00309.x.

Pierce, B.C. and Turner, D.N., 2000. Pict: a programming language based on the pi-calculus. *Proof, language and interaction: essays in honour of robin milner*. Cambridge, MA: MIT Press, pp.455–494.

Pour-El, M.B., 1999. The structure of computability in analysis and physical theory: an extension of church's thesis. In: E. Griffor, ed. *Handbook of computability theory, studies in logic and the foundations of mathematics.* [Online]. Amsterdam: Elsevier, pp.449–471. https://doi.org/10.1016/S0049-237X(99)80029-9.

Prasse, M. and Rittgen, P., 1998. Why church's thesis still holds. some notes on peter wegner's tracts on interaction and computability. *The Computer Journal* [Online], 41, pp.357–362. https://doi.org/10.1093/comjnl/41.6.357.

Rapaport, W.J., 2018. What is a computer? a survey. *Minds and Machines* [Online], 28(3), pp.385–426. https://doi.org/10.1007/s11023-018-9465-6.

Rogers, H., 1987. *Theory of recursive functions and effective computability*. Cambridge, MA: MIT Press. https://doi.org/10.2307/3614588.

Siegelmann, H.T., 1995. Computation beyond the turing limit. *Science* [Online], 268(5210), pp.545–548. https://doi.org/10.1126/science.268.5210.545.

Smith, B.C., 2002. The foundations of computing. In: M. Scheutz, ed. *Computationlism: new directions*. Cambridge, MA: MIT Press.

Soare, R.I., 2013. Interactive computing and relativized computability. In: *Computability: turing, godel, church, and beyond* [Online]. Ed. by B.J. Copeland. Cambridge, MA: MIT Press. Chap. 9, pp.214–271. https://doi.org/10.7551/mitpress/8009.003.0010.

Staiger, L., 1997. Omega-languages. In: *Handbook of formal languages* [Online]. Ed. by G. Rozenberg and A. Salomaa. Berlin; Heidelberg: Springer, pp.339–387. https://doi.org/10.1007/978-3-642-59126-6_6.

Thomas, W., 1990. Automata on infinite objects. In: *Formal models and semantics* [Online]. Amsterdam: Elsevier. Chap. 4, pp.133–191. https://doi.org/10.1016/b978-0-444-88074-1.50009-3.

Turing, A., 1937. On computable numbers, with an application to the entscheidungsproblem. *Proceedings of the London Mathematical Society* [Online], s2-42(1), pp.230–265. https://doi.org/10.1112/plms/s2-42.1.230.

Van Leeuwen, J. and Wiedermann, J., 2000. On the power of interactive computing. In: J. van Leeuwen et al., eds. *Theoretical computer science: exploring new frontiers of theoretical informatics* [Online]. Vol. 1872, *TCS 2000. Lecture Notes in Computer Science*. Berlin; Heidelberg: Springer. https://doi.org/10.1007/3-540-44929-9_48.

Van Leeuwen, J. and Wiedermann, J., 2001. Beyond the turing limit: evolving interactive systems. In: L. Pacholski and P. Ružička, eds. *Sofsem 2001:*

theory and practice of informatics [Online]. Vol. 2234, Lecture notes in computer science. Berlin; Heidelberg: Springer, pp.90–109. https://doi.org/10.1007/3-540-45627-9_8.

Van Leeuwen, J. and Wiedermann, J., 2006. A theory of interactive computation. *Interactive computation: the new paradigm* [Online]. Berlin; Heidelberg: Springer, pp.119–142. https://doi.org/10.1007/3-540-34874-3_6.

Wegner, P., 1995. Interaction as a basis for empirical computer science. *ACM Computing Surveys (CSUR)* [Online], 27(1), pp.45–48. https://doi.org/10.1145/214037.214092.

Wegner, P., 1997. Why interaction is more powerful than algorithms. *Communications of the ACM* [Online], 40(5), pp.80–91. https://doi.org/10.1145/253769.253801.

Wegner, P., 1998. Interactive foundations of computing. *Theoretical Computer Science* [Online], 192(2), pp.315–351. https://doi.org/10.1016/S0304-3975(97)00154-0.

Wegner, P. and Goldin, D., 1999. Interaction as a framework for modeling. In: G. Goos et al., eds. *Conceptual modeling: current issues and future directions* [Online]. Vol. 1565, *Lecture Notes in Computer Science*. Berlin; Heidelberg: Springer, pp.243–257. https://doi.org/10.1007/3-540-48854-5_19.

Wegner, P. and Goldin, D., 2003. Computation beyond turing machines. *Communications of the ACM* [Online], 46(4), pp.100–102. https://doi.org/10.1145/641205.641235.

Wegner, P. and Goldin, D., 2006. Principles of problem solving. *Communications of the ACM* [Online], 49(7), pp.27–29. https://doi.org/10.1145/1139922.1139942.

Review articles

Artykuły recenzyjne

Why is neuron modeling of particular philosophical interest?

Paweł Polak
Pontifical University of John Paul II in Krakow

Andrzej Bielecki, *Models of Neurons and Perceptrons: Selected Problems and Challenges*, Springer International Publishing, Cham 2019, pp. 156; Studies in Computational Intelligence Volume vol. 770, DOI: 10.1007/978-3-319-90140-4.

A peculiarity of "philosophy in science" (see Heller, 2019; Polak, 2019) is that the best sources tend to be atypical from the viewpoint of most philosophers: For example, on the one hand, there are works that popularize science, while on the other hand, there are research articles and even specialized monographs. The book discussed here falls into the last category, and it is devoted to modeling neurons and perceptrons. It was written by a mathematician from Kraków, Andrzej Bielecki, who is currently working at the AGH University of Science and Technology.[1] Readers of *ZFN*, as well as the related *Semina Scientiarum* journal, will probably associate him with the philosophical activities that he has practiced within the context of his scientific activities (2016; 2018). Bielecki is an example of a scientist–philosopher

[1] Andrzej Bielecki received an M.Sc. degree in Physics and Mathematics and a Ph.D. in Mathematics, D.Sc. (habilitation) in Mathematics from the Jagiellonian University in Kraków. He obtained a professorship in Computer Science in 2020. His fields of interest includes dynamical systems theory, artificial intelligence, cybernetics, and philosophy of science, and he has written over 120 scientific papers and one textbook.

Philosophical Problems in Science (Zagadnienia Filozoficzne w Nauce)
No 72 (2022), pp. 113–122 •

from the Krakow milieu,[2] and it is worth noting that he develops his philosophical activities, among other things, through his work on the Committee on Philosophy of Science at the Polish Academy of Arts and Sciences in Kraków.

Bielecki's book is published as a volume in the "Studies in Computational Intelligence" series, which is intended for research that contributes to computational intelligence. The book is an in-depth monograph about computationally modeling basic cognitive structures, such as neurons. It comprises five parts that logically present different areas of the subject. The first part, titled "Preliminaries," provides fundamental biological knowledge about neurons and essential information about the basics of artificial neural networks and their applications. The second part is then devoted to the mathematical foundations of modeling, particularly dynamical systems theory. Next, the third part goes into mathematical models of neurons, such as models of entire neurons and models of portions of neurons. The fourth part then focuses on modeling perceptrons, starting with linear perceptrons and ending with nonlinear ones. The final part consists of the appendices.

The author deliberately combines biological and simulation perspectives in his book, aspects that are usually separated into distinct studies within neuron research. This interdisciplinary approach aims to identify new sources of biological inspiration for mathematics and computational modeling. Bielecki also says he chose this approach "because it seems that there are numerous models of biological neural structures that can be the basis for artificial systems and that have not been utilized yet" (Bielecki, 2019, p.3). It is worth adding here

[2] It is worth mentioning that in the book, Bielecki mostly uses examples from research conducted in the Kraków milieu. The use of the cybernetic theory framework could also be interpreted as another sign of the local milieu's influence.

that although it is not explicitly stated in the book, Bielecki's attitude toward interdisciplinarity has resulted from his work in various interdisciplinary research teams that have included biologists and mathematicians.

Bielecki's work provides an essential overview of the contemporary view of neuron modeling, and the included bibliography is a helpful further guide in this area. Here, the reader can find extensive and detailed, yet concisely presented, knowledge about modeling neurons and their networks. By zooming in on this monograph's detailed explanation of the problems of modeling a single neuron, we can quickly realize how simplistic assumptions are often made in projects related to Whole Brain Emulation (WBE). For my part, I regard this as a warning to approach the results of WBE-like projects with extreme caution (e.g. Kycia, 2021). After all, a single neuron itself is still not sufficiently understood (e.g. Bielecki, 2019, p.133), and the complexity of its structure leads one to realize the incredible complexity of the brain, as well as the level of complexity we are trying to master in brain-related research. Even the problem of practical computational complexity in whole-neuron modeling comes up: "It should be stressed that, currently, the computational power of computers is too weak to compose the model of the whole neuron by using models of its parts" (Bielecki, 2019, p.59). The author also notes the need for inter-level studies (i.e., between subneural, neural, and network levels). We should add that if we talk about the emergence of properties at higher levels in philosophy, such topics are consistently overlooked in scientific research.

Bielecki's monograph makes one realize how much effort we should be devoting to discussing the role of simplifying assumptions and idealizations in simulations. Of course, artificial neural networks (ANNs) can be based on greatly simplified models of neurons for technical applications and often succeed at achieving the desired goals. The situation is different in scientific research, however, because it attempts to describe the functioning of neuronal structures, such as the brain, through simulations. Bielecki states this clearly: “In the light of neurophysiological knowledge, the models of the whole neuron are simplified to such an extent that they do not reflect, even approximately, the character of signal processing in the biological neuron.” (Bielecki, 2019, p.57)

It is worth noting that a particular strength in Bielecki’s book is how he does not limit modeling to just standard computer modeling. Indeed, he is also interested in physical (electronic) models that operate on continuous values due to the problems with digital simulations of nonlinear differential equations: “If the model is based on ordinary differential equations, then it can be implemented by using an electronic circuit whose dynamics is described by the same differential equation” (Bielecki, 2019, p.17). Bielecki proposes using a kind of classical analog computation. From a philosophical perspective, this means he does not share the common tacit philosophical assumption among many works that Turing’s computational model can sufficiently describe the real world. For this reason alone, I think that any philosopher who wishes to make a responsible statement about neurons, the brain, and the research about them should read this book. Reflecting on the implications of the knowledge presented here should help the reader to understand how many problematic assumptions we currently make in discussions related to this topic. I would like to share some of my thoughts that were inspired by this book below.

The reviewed monograph brings some exciting contributions to the discussion about simulation methodology in biology. Indeed, the specific issues of biological simulation are worthy of a separate study, which, by the way, Bielecki is currently working on. Nevertheless, the methodological specifics of such simulations are rarely addressed. In Bielecki's book, however, we can find an attractive methodological scheme that has the advantage of being created based on scientific practice. It is therefore an excellent example of "philosophy in science," which in this case is located at the intersection of applied mathematics and biology.

In Bielecki's view, computational modeling begins with biological research (A), which allows us to distinguish relevant structures and processes. The next step then requires biological experiments or observations (B). The crucial properties can only then be determined (C) based on these, enabling a semi-formal description (D) to be formulated. This description can then serve as the basis for creating a formal model (E), which can then act as the basis for constructing a software or hardware implementation (F). According to Bielecki, these final two stages can influence each other, with each acting as the starting point for formulating the other. Finally, it is essential that the results of formal modeling should eventually become the subject of an analysis through a traditional approach (G). Consequently, it may become necessary to modify the experimentation/observation phase (B) or the determination of the crucial properties (C). Such feedback is essential to the computational modeling methodology, but it also indicates how much creative input the scientist has. Models are not mere generalizations of facts, as methodologists once wanted them to be, but rather the result of a complex, looped adaptive process.

Interestingly, the precondition for creating such models—and therefore the need for learning more about complex, or perhaps *more* complex, biological structures—is the ability to perform sufficiently complex calculations, either in digital or analog form. The methodological scheme indicated by Bielecki therefore points to a strongly "non-linear" looped process that occurs during the creation of advanced biological knowledge. It is a case of epistemic bootstrapping, or more precisely, it could be described as epistemic feedback (Weisberg, 2010). Interestingly, an essential argument for considering such a "non-linear logic of scientific development" (to use Heller's words) flows directly from scientific practice. Bielecki, however, is not interested in isolated arguments for and against epistemic bootstrapping. He instead posits the validity of this method based on an analysis of actual scientific practice in biology. It should be emphasized that the reviewed book does not contain detailed philosophical analysis or present a pro and contra discussion of the presented theses but rather seeks to uncover an essential philosophical issue that is entangled with modeling in biology. Nevertheless, meticulous analyses and deliberations about the pros and cons should be the next step in reflecting upon the philosophical issues of biological simulations. Nevertheless, let us highlight that such an endeavor would not be possible without first identifying these issues, and this book plays an important intellectual role by posing important and non-trivial philosophical questions, even if it does so indirectly.

Bielecki's monograph also shows the level of depth in the mathematization of biology that is taking place in research at the cellular and subcellular levels. The author does not apply the slightest hint of persuasion here but rather simply demonstrates the impressiveness of the precise mathematical basis for neuronal modeling. It easily convinces the reader of the deep and practical mathematization of biology that has taken place through computational modeling and the adoption of a cybernetic framework.

Bielecki's remarks about the need to synthesize various modeling approaches are worth special attention: "In this monograph, the cybernetic modeling, the mathematical modeling, and the modeling by using electronic circuits intertwine. [. . .] This is also a specificity of the approach presented in this monograph because these three ways of modeling are usually exploited separately." He also points out this approach's more general, philosophical context: "Since the Enlightenment analytic approach to scientific problems has dominated, and the synthetic approach is, in general, in the state of atrophy. The synthetic mathematical–electronic approach to modeling sub-neural processes, presented in this monograph, tests whether such an approach can be efficient. *The results show that the answer is affirmative* [emphasis added]" (Bielecki, 2019, p.124). Note that I emphasized the final sentence to highlight how the author sees this book as a kind of methodological experiment with a positive result. Indeed, I think this result should be presented to philosophers in more detail to help us understand its methodological soundness, and maybe a separate study on this issue could be appropriate for clarifying Bielecki's ideas.

Now, let me illustrate the conceptual scheme used by Bielecki: It is based on concepts from cybernetics theory, one of the vital mathematical theories that provides the foundation for developing interdisciplinary research and computational modeling. In Poland, cybernetics is still successfully pursued, especially in Kraków at the AGH University of Science and Technology,[3] but contemporary international discussions use somewhat different conceptual systems. A good example is Gordana Dodig-Crnkovic's article in an issue of ZFN (Dodig-Crnkovic, 2022). The deep analogies between the two approaches are surprising. For example, take Bielecki's phrase: "Each type of biological cells, including the simplest bacteria, receives stim-

[3] In private correspondence, the author stated that the most important sources of inspiration on the issue of cybernetics are the works of Tadeusiewicz (1994; 2009), who is a distinguished researcher and the founder of a vivid center of biocybernetics at AGH in Krakow. A further source of inspiration were the works of another Krakow scientist, Mariusz Flasiński (1997; 2016), who is affiliated with Jagiellonian University in Kraków.

uli from its environment and processes the obtained signals" (Bielecki, 2019, p.5). It is close to the info-computational in Dodig-Crnkovic's view, although she uses a specific reference to information theory. It would be worthwhile to analyze the relationship between cybernetics and contemporary information concepts in more depth, because it may be possible to find new, inspiring analogies or more convenient conceptual frameworks.

Finally, let us conclude with the specifics of "philosophy in science," with which I began this review. One of its unique features is that interesting contributions can be rich in philosophical content, even though the word "philosophy" may rarely appear in them, if at all. Andrzej Bielecki's book is an excellent example of this, because he mentions philosophy only twice, and one of those refers to the Enlightenment. Nevertheless, it makes an exciting contribution to understanding the philosophical issues in modern biology.

Abstract

This review article discusses Andrzej Bielecki's book *Models of Neurons and Perceptrons: Selected Problems and Challenges*, as published by Springer International Publishing. This work exemplifies "philosophy in science" by adopting a broad, multidisciplinary perspective for the issues related to the simulation of neurons and neural networks, and the author has addressed many of the important philosophical assumptions that are entangled in this area of modeling. Bielecki also raises several important methodological issues about modeling. This book is recommended for any philosophers who wish to learn more about the current state of neural modeling and find inspiration for a deeper philosophical reflection on the subject.

Keywords

neuron modeling, sub-neuron modeling, computational modeling, analog computation, philosophy in science, philosophy of biology, philosophy of computing.

Bibliography

Bielecki, A., 2016. Cybernetyczna analiza zjawiska życia. *Philosophical Problems in Science (Zagadnienia Filozoficzne w Nauce)* [Online], (61), pp.133–164. Available at: <https://zfn.edu.pl/index.php/zfn/article/view/361> [visited on 27 January 2020].

Bielecki, A., 2018. Epistemologiczne problemy w biologii subkomórkowej: obserwacje, modele matematyczne i symulacje komputerowe. *Semina Scientiarum* [Online], 16, pp.10–23. https://doi.org/10.15633/ss.2482.

Bielecki, A., 2019. *Models of Neurons and Perceptrons: Selected Problems and Challenges* [Online]. Vol. 770, *Studies in Computational Intelligence*. Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-90140-4.

Dodig-Crnkovic, G., 2022. In search of common, information-processing, agency-based framework for anthropogenic, biogenic, and abiotic cognition and intelligence. *Philosophical Problems in Science (Zagadnienia Filozoficzne w Nauce)*, (73), ??–??

Flasiński, M., 1997. "Every Man in His Notions" or Alchemists' Discussion on Artificial Intelligence. *Foundations of Science* [Online], 2(1), pp.107–121. https://doi.org/10.1023/A:1009687513096.

Flasiński, M., 2016. *Introduction to Artificial Intelligence* [Online]. Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-40022-8.

Heller, M., 2019. How is philosophy in science possible? *Philosophical Problems in Science (Zagadnienia Filozoficzne w Nauce)* [Online], (66), pp.231–249. Available at: <https://zfn.edu.pl/index.php/zfn/article/view/482> [visited on 6 October 2021].

Kycia, R., 2021. Information and brain. *Philosophical Problems in Science (Zagadnienia Filozoficzne w Nauce)* [Online], (70), pp.45–72. Available at: <https://zfn.edu.pl/index.php/zfn/article/view/514> [visited on 6 December 2022].

Polak, P., 2019. Philosophy in science: A name with a long intellectual tradition. *Philosophical Problems in Science (Zagadnienia Filozoficzne w Nauce)* [Online], (66), pp.251–270. Available at: <https://zfn.edu.pl/index.php/zfn/article/view/472> [visited on 6 October 2021].

Tadeusiewicz, R., 1994. *Problemy biocybernetyki*. Wyd. 2. Warszawa: Wydawnictwo Naukowe PWN.

Tadeusiewicz, R., ed., 2009. *Neurocybernetyka teoretyczna*. Warszawa: Wydawnictwa Uniwersytetu Warszawskiego.

Weisberg, J., 2010. Bootstrapping in General. *Philosophy and Phenomenological Research* [Online], 81(3), pp.525–548. Available at: <https://www.jstor.org/stable/41057492> [visited on 6 December 2022].

Beyond epistemic concepts of information: The case of ontological information as philosophy in science

Paweł Polak
Pontifical University of John Paul II in Krakow

Roman Krzanowski, *Ontological information: information in the physical world, series: World Scientific series in information studies*, 13, Hackensack, New Jersey: World Scientific 2022, pp. xii+264.

The concept of information plays and important role in science and philosophy, as well as in everyday life, such that it is now hard to imagine that this concept has been only adopted in the late 1940s. Many scientists find it even harder to believe that the concept of information can involve anything other than communication processes, and this is almost certainly due to Claude Shannon's Theory of Communication (TOC) (1949) that entered the canon of unquestionable modern scientific knowledge. Unquestionably accepting Shannon's concept of a measure of information entropy as the definition of information encourages a scholar to treat TOC as scientific dogma.[1] However,

[1] Shannon himself tried to warn against abusing his theory of communication (Shannon, 1956), though apparently unsuccessfully. He called against using the theory as a source of hypotheses in other scientific disciplines. However, the opposite has happened—information metaphors have become unquestionable theoretical core of many modern concepts. This fact should not come as a surprise, because if few people read the original work contenting themselves only with its processed results, it is difficult to suppose that anyone outside a handful of specialists in the history of computing read Shannon's critical remarks very rarely cited in the literature.

Philosophical Problems in Science (Zagadnienia Filozoficzne w Nauce)
No 72 (2022), pp. 123–133

studying the original work of Shannon and Weaver (1964, p.3), we realize that Shannon was primarily interested in communication engineering of digital signals (signal recovery, noise, optimal coding of signal), and the concepts of information and entropy have been borrowed by him from works of physicists like Ludwig Boltzmann, Leo Szilard, and John von Neumann.

After almost 80 years of being around, information is an elusive concept with manifold meanings; Krzanowski (2022) referred to more than 300 definitions of information, As information plays such an important role in contemporary societies, technology and science, it seems only logical that the efforts to clarify the meaning of information should never be abandoned. And this is precisely what Krzanowski's book is about.

However, while most of the published works on information see information through Shannon's lenses the focus of Roman Krzanowski's book (Krzanowski, 2022) is physical information. i.e., information that is not associated with knowledge or communication (Shannon's legacy), and that it is a part of nature as other physical phenomena are; the conceptualization of information has not been widely accepted by the scientific community.

The first significant step was to distinguish between concepts of epistemic information (as found in Shannon's theory) from theories about ontological information. The author's second step was rather than constructing a concept of information *a priori*, to proceed in the spirit of the Kraków school of philosophy in science (Heller, 2019; Polak, 2019; Polak and Trombik, 2022). Krzanowski searches for the meaning of ontological information g attributed to it by reserachers and tries to understand the philosophical basis for using such concept. This is no coincidence, because in the pages of the associated journal "Philosophical Problems in Science/Zagadnienia Filozoficzne w Nauce" (ZFN), the problem of information in science has been discussed from the very beginning to the current day (1978; 1981; 2017; 2020).

The book is divided over seven chapters. The first chapter introduces the several definitions of information, often contradictory, demonstrating difficulty in accurately capturing the essence of this concept. Throughout the chapter the author gradually builds the conceptual base for his thesis and carefully justifies all his decisions. Of course, it is possible to disagree with Krzanowski on many issues, but one must admit that he tries to be very consistent, meaning that the deliberations as a whole constitute a valuable analysis of the concept of information. Even if one disagrees with the author's detailed claims, one would still concede that this book takes on an intriguing intellectual challenge and makes a significant contribution to organizing and illuminating the discussion around the concept of information.

At a time when authors mainly value their own originality, the work of Krzanowski has the characteristics of the best classical philosophy, which built its solutions on critical struggles with the heritage of tradition. It undoubtedly contributes to modern analytic philosophy, but the author's approach is to not simply copy contemporary models but instead creatively draw from various traditions, including Polish analytic philosophy. Although the author is far removed from the theses of Aristotelianism and scholasticism, his perfectly organized, methodical criticism and consistency and his precise argumentation is reminiscent of the style of Thomas Aquinas. More importantly, though, Krzanowski is not pragmatophobic, instead boldly pursuing solutions and seeking his own synthesis in the thicket of proposals. This method certainly sets this book apart from most works on the concept of information.

The crucial concept of ontological information is characterized as "a physical phenomenon" (Krzanowski, 2022, p.6). The author assumes that "this information is perceived as a structure, organization, or form of natural and artificial (artifacts) objects." He also defines this information as being objective and mind-independent while simultaneously clarifying all the concepts involved and trying to provide an argument for every claim. He also warns that ontological information is a metaphysical concept and contributes to contemporary analytical metaphysics.

Krzanowski's analyses start with some intuitions about ontological information. He reconstructs ideas from dispersed quotations, much like how historians of philosophy deal with pre-Socratic philosophy. The methodology is similar because the concept of ontological information frequently manifests itself in the form of dispersed brief remarks.

The collection of these brief remarks by scientists is combined with a careful interpretation and an attempt to reconstruct the philosophical intuitions they contain, the tasks that are the subject of the second chapter. This intriguing and inspiring journey passes through a variety of ideas, culminating in the formulation of the eight main intuitions about ontological information that run through the scientific literature.

The third chapter analyzes the existing philosophical conceptions of ontological information, even though they do not usually refer to it using this term. We can find concepts coined by representatives of different disciplines from various countries, such as Carl von Weizsäcker, Krzysztof Turek, Stefan Mynarski, John Collier, Tom Stonier, Michał (Michael) Heller, Gordana Dodig Crnkovic, César Hildago, Thomas

Nagel, Jacek Jadacki, and Anna Brożek. Krzanowski summarizes these concepts into 11 claims that are explained in detail (Krzanowski, 2022, pp.86–93).

The aforementioned intuitions and claims serve as a foundation for synthesizing the concept of ontological information in the fourth chapter, with Krzanowski ultimately reducing it to three claims:

- (EN) Information has no meaning, but meaning is derived from information by a cognitive agent.
- (PE) Information is a physical phenomenon.
- (FN) Information is responsible for the organization of the physical world.

This is then followed by two corollaries:

- (C1) Information is quanti[FB01?]able.
- (C2) Changes in the organization of physical objects can be denoted as a form of computation or information processing.

After the critical discussion, Krzanowski posits that these three properties and two corollaries are indispensable for the definition and understanding of the concept of ontological information.. This set of properties has a hypothetical status, and this conceptualization is relative to actual science, so it is open to future changes together with the entirety of scientific knowledge.

The fifth chapter is devoted to broadly analyzing the problem of ontological and epistemological aspects of the concept of information. Krzanowski needs to adopt this perspective to further clarify the concept of ontological information. The analysis shows that while both concepts of information are required to account for the full spectrum of interpretations for information, ontological information appears to be more fundamental because it can serve as a carrier of epistemic information.

In the following chapter, Krzanowski moves onto applications and interpretations of ontological information. He critically discusses the concept of an “infon” and data as basic concepts for defining information. The author claims that his conceptualization is more fundamental and better explains the source of epistemic information. Furthermore, Krzanowski attempts to resolve the dilemma of the contradictory abstract and concrete natures for information, and this is another original and inspiring aspect of his book. The example application of ontological information is Krzanowski’s original concept of Minimal Information Structural Realism (first introduced in Krzanowski, 2017). Also of interest is the consideration about possibly applying his conceptualization of ontological information to Popper’s Three Worlds and Mark Burgin’s General Theory of Information. Finally, Krzanowski applies Perzanowski’s ontology to build an ontological foundation for the concept of ontological information.

The final chapter gathers together all the observations from the book, summarizes the key findings and conclusions, and brings up some selected criticisms of ontological information. (Krzanowski probably intentionally avoids the most common dogmatic critiques, regarding them as not being suitable for philosophical consideration.) Finally, through nine questions, the author reveals some perspectives for future research into ontological information. Each question opens up a new field that could be a subject of a new study.

It is worth mentioning that the book has been carefully prepared from an editorial perspective. It has not, however, been spared from minor inaccuracies, such as the fact that Krzysztof Turek received his doctorate from the Pontifical Academy of Theology in Kraków, which only transformed into the Pontifical University of John Paul II in Kraków many years later. In addition, qualifying Jacek Jadacki as a computer scientist (Krzanowski, 2022, p.45) rather than a philosopher and pianist is not only untrue—it is also inconsistent with the rest of the work. Nevertheless, these minor glitches do not significantly impair this important consideration of ontological information. Unfortunately, many more, albeit minor, errors can be found in the bibliography, especially in the Polish titles of works. This may hinder any search for the cited works in databases. Moreover, in some cases, works that have long been out of print, even five years ago, are still marked as being in print.

Assessing philosophical import of the book we may begin by noting that the book is relatively new, but published works have already used the ideas within it. Work that is worthy of mentioning here is that of the philosophy of information specialist Mark Burgin (Burgin and Mikkilineni, 2022). The ideas are also reflected in this issue of *Philosophical Problems in Science* / ZFN (Mścisławski, 2022). We should also emphasize here that the ideas presented in the reviewed book have resulted from the longer, critical reflection conducted by Krzanowski. This is especially true of the concepts of physical information and information, which were originally treated as being synonymous but have now been distinguished in the book, and this division has been well justified.

After reading the book, many questions can be raised, but to be fair to the author, they should be posed with great precision and care. There are certainly questions about whether the book finally resolves the problem of defining information or whether it finally explains the nature of information, but these would be misplaced questions. Indeed, it would be unacceptable for science to achieve these goals through *a priori* considerations, so if we adopt the scientific perspective, we must accept that we cannot provide definitive answers. Nevertheless, this does not mean that we must remain mute on the subject. On the contrary, we can say much about how the concept of information functions in modern science. Of course, Krzanowski's book only addresses the issue of ontological information, because philosophers have paid far too little attention to it. Indeed, the fame of Shannon's work on the theory of communication (often interpreted as the theory of information) has all too often led to an atrophy of criticism and a limited vision for the nature of information. For this reason, the reviewed book makes a valuable contribution to the discussion, and it not only reconstructs the concept of ontological information that is actually used in science but also critically evaluates it.

The book formulates a set of properties of ontological information. This is the first attempt of its kind, and most importantly, it does not start from the author's arbitrary ideas but rather tries to deal with the thicket of intuitions and conceptualizations put forward in scholarly publications. The task is hard as scientists often hide their ignorance behind imprecise statements. After all, they are not professional philosophers, and in this task, which is secondary to their research, they may easily fall fowl of various errors or inaccuracies. This is perhaps why Einstein pointed out that one should pay attention to what scientists actually do rather than what they say about it. If this is indeed the case, then I would like to propose an important area to develop research into the concept of ontological information, namely to investigate how it is actually used in scientific research.

Thus, it becomes necessary to go beyond the mere declarations and conceptualizations made by scholars and look at how the concept is used in explanatory structures and other aspects of research practice, at least if such contexts can be identified. Nevertheless, it should be noted that at the time of writing, such a task was essentially impossible given the scarcity of studies using the concept of ontological information.

Of course, the author's assertion that the concept of ontological information is a metaphysical concept should be borne in mind. Thus, the proposed line of research is also a proposed case study of how metaphysics interacts with the sciences using the example of the modern concept of information. Such studies are also lacking, yet they could be of value to all philosophers of science who do not share the extreme anti-metaphysical position.

We can hope that this reviewed book and the study areas suggested for continuing research should have some impact on science. By stripping certain ideological trappings from the concept of information, broadening the perspectives (escaping Shannon's shadow) and making some necessary clarifications, scientists should certainly be able to develop new lines of research more effectively. Besides, the first harbingers of change have already appeared, such as the work flowing from the Kraków scientific community (Bielecki and Schmittel, 2022), which is based on the work of Krzanowski. Let us hope that other works using such a "purified" concept of ontological information in scientific practice will soon appear, because it will open up a new yet important step in the philosophy of information, namely research into ontological information in scientific explanatory structures.

Acknowledgments The author thanks Łukasz Mścisławski for the in-depth discussions of the concept of information. His article (Mścisławski, 2022) published in this issue of ZFN is the first critical analysis of the concepts presented by R. Krzanowski; Łukasz Mścisławski's ideas significantly contributed to a this review.

Abstract

This review article discusses Andrzej Bielecki's book *Models of Neurons and Perceptrons: Selected Problems and Challenges*, as published by Springer International Publishing. This work exemplifies "philosophy in science" by adopting a broad, multidisciplinary perspective for the issues related to the simulation of neurons and neural networks, and the author has addressed many of the important philosophical assumptions that are entangled in this area of modeling. Bielecki also raises several important methodological issues about modeling. This book is recommended for any philosophers who wish to learn more about the current state of neural modeling and find inspiration for a deeper philosophical reflection on the subject.

Keywords

neuron modeling, sub-neuron modeling, computational modeling, analog computation, philosophy in science, philosophy of biology, philosophy of computing.

Bibliography

Bielecki, A. and Schmittel, M., 2022. The Information Encoded in Structures: Theory and Application to Molecular Cybernetics. *Foundations of Science* [Online], 27(4), pp.1327–1345. https://doi.org/10.1007/s10699-022-09830-8.

Burgin, M. and Mikkilineni, R., 2022. Is Information Physical and Does It Have Mass? *Information* [Online], 13(11), p.540. https://doi.org/10.3390/info13110540.

Heller, M., 2019. How is philosophy in science possible? *Philosophical Problems in Science (Zagadnienia Filozoficzne w Nauce)* [Online], (66), pp.231–249. Available at: <https://zfn.edu.pl/index.php/zfn/article/view/482> [visited on 6 October 2021].

Krzanowski, R., 2017. Minimal Information Structural Realism. *Philosophical Problems in Science (Zagadnienia Filozoficzne w Nauce)* [Online], (63), pp.59–75. Available at: <http://www.zfn.edu.pl/index.php/zfn/article/view/396> [visited on 16 May 2018].

Krzanowski, R., 2020. Why can information not be defined as being purely epistemic? *Philosophical Problems in Science (Zagadnienia Filozoficzne w Nauce)* [Online], (68), pp.37–62. Available at: <https://zfn.edu.pl/index.php/zfn/article/view/494> [visited on 30 November 2021].

Krzanowski, R., 2022. *Ontological Information: Information in the Physical World* [Online]. Vol. 13, *World Scientific series in information studies*. Hackensack, New Jersey: World Scientific. https://doi.org/10.1142/12601.

Mścisławski, Ł., 2022. Is information something ontological, or physical or perhaps something else? Some remarks on R. Krzanowski approach to concept of information. *Philosophical Problems in Science (Zagadnienia Filozoficzne w Nauce)*, (73), ??

Polak, P., 2019. Philosophy in science: A name with a long intellectual tradition. *Philosophical Problems in Science (Zagadnienia Filozoficzne w Nauce)* [Online], (66), pp.251–270. Available at: <https://zfn.edu.pl/index.php/zfn/article/view/472> [visited on 6 October 2021].

Polak, P. and Trombik, K., 2022. The Kraków School of Philosophy in Science: Profiting from Two Traditions. *Edukacja Filozoficzna* [Online], (2(74)). https://doi.org/10.14394/edufil.2022.0023.

Shannon, C.E., 1949. *The mathematical theory of communication*. Urbana: University of Illinois Press.

Shannon, C.E. and Weaver, W., 1964. *The Mathematical Theory of Communication*, *Illini Books*, IB-13. Urbana: University of Illinois Press.

Turek, K., 1978. Filozoficzne aspekty pojęcia informacji. *Philosophical Problems in Science (Zagadnienia Filozoficzne w Nauce)*, (1), pp.32–41.

Turek, K., 1981. Rozważania o pojęciu struktury. *Philosophical Problems in Science (Zagadnienia Filozoficzne w Nauce)*, (3), pp.73–95.