STUDIES
IN APPLIED
PHILOSOPHY,
EPISTEMOLOGY
AND
RATIONAL
ETHICS

Gordana Dodig-Crnkovic
Raffaela Giovagnoli (Eds.)

SAPERE

Computing Nature

Turing Centenary Perspective

Springer

Studies in Applied Philosophy, Epistemology and Rational Ethics 7

For further volumes:
http://www.springer.com/series/10087

Advisory Board

Gordana Dodig-Crnkovic
and Raffaela Giovagnoli (Eds.)

Computing Nature

Turing Centenary Perspective

Springer

Editors
Gordana Dodig-Crnkovic
School of Innovation, Design and Engineering
Mälardalen University
Västerås
Sweden

Raffaela Giovagnoli
Faculty of Philosophy
Pontifical Lateran University
Vatican City
Italy

ISSN 2192-6255 ISSN 2192-6263 (electronic)
ISBN 978-3-642-37224-7 ISBN 978-3-642-37225-4 (eBook)
DOI 10.1007/978-3-642-37225-4
Springer Heidelberg New York Dordrecht London

Library of Congress Control Number: 2013933717

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Contents

Computing Nature – A Network of Networks of Concurrent Information Processes

Gordana Dodig-Crnkovic[1] and Raffaela Giovagnoli[2]

[1] Department of Computer Science and Networks, Mälardalen University, Sweden
gordana.dodig-crnkovic@mdh.se
[2] Faculty of Philosophy. Lateran University, Rome, Italy
raffa.giovagnoli@tiscali.it

1 Introduction

The articles in the volume Computing Nature present a selection of works from the Symposium on Natural/Unconventional Computing at AISB/IACAP (British Society for the Study of Artificial Intelligence and the Simulation of Behaviour and The International Association for Computing and Philosophy) World Congress 2012, held at the University of Birmingham, on the occasion of the centenary of Alan Turing's birth.

This book is about nature considered as the totality of physical existence, the universe. By physical we mean all phenomena - objects and processes - that are possible to detect either directly by our senses or via instruments. Historically, there have been many ways of describing the universe (the cosmic egg, the cosmic tree, the theistic universe, the mechanistic universe) while a particularly prominent contemporary model is the computational universe.

One of the most important pioneers of computing, Turing, seen by Hodges [1] as natural philosopher, can be identified as a forerunner and founder of the notion of computing nature and natural computing through his ideas about morphological computing, "unorganized" (neural-network type) machines and "oracle" machines. Turing's impact on the development of computing can be seen as two-fold: laying down the foundations of the theory of computing by his Turing Machine model he provided such powerful paradigm that soon led to the belief that it is all we can do when it comes to computing. But, "There are assumptions underlying the paradigm which constrain our thinking about the realities of computing", as Cooper in this volume rightly observed. On the other hand, his work on natural computing points towards the development in different directions. It is obvious from Turing's own research that he did not consider Turing Machine model the only possible way of computation.

After many decades of development, present day computers are distinctly different from the early stand-alone calculating machines that Turing helped construct, that were designed to mechanize computation of mathematical functions. Computers today are networked and largely used for world-wide communication and variety of information processing and knowledge management. They are cognitive tools of extended mind (in the sense of Clark and Chalmers) used in social interactions and they provide ever growing repositories of information. Moreover, computers play an

G. Dodig-Crnkovic and R. Giovagnoli (Eds.): *Computing Nature*, SAPERE 7, pp. 1–22.
DOI: 10.1007/978-3-642-37225-4_1

important role in the control of physical processes and thus connect directly to the physical world in automation, traffic control, robotics and more. Apart from classical engineering and hard-scientific domains, computing has in recent decades pervaded new fields such as biology and social sciences, humanities and arts – all previously considered as typical soft, non-mechanical and unautomatable domains.

Computational processes running in networks of networks (such as the internet) can be modeled as distributed, reactive, agent-based and concurrent computation. The main criterion of success of this computation is not its termination, but its behavior - response to changes, its speed, generality and flexibility, adaptability, and tolerance to noise, error, faults, and damage. Internet, as well as operating systems and many database management systems are designed to operate indefinitely and termination for them would be an error. We will return to the topic of concurrent computing and its relationship with Turing machine model of computation in more detail later on.

One of the aims of this book is to show the state of the art of developments in the field of natural/unconventional computation which can be seen as generalization and enrichment of the repertoire of classical computation models. As a generalization of the traditional algorithmic Turing Machine model of computation, in which the computer was an isolated box provided with a suitable algorithm and an input, left alone to compute until the algorithm terminated, natural computation models *interaction* i.e. communication of computing processes with each other and with the environment. In natural systems, computation is information processing that can proceed on both symbolic and sub-symbolic (signal-processing) level. For human cognitive processes it means that not only the execution of an algorithm can be seen as computation, but also learning, reasoning, processing of information from sense organs, etc.

Hewitt [2] characterizes the Turing machine model as an *internal* (*individual*) framework and the Actor model of concurrent computation as an *external* (*sociological*) model of computing. This tension between an (isolated) individual *one* and (interacting) social *many* resonates with two articles from this volume: Cottam et al. who distinguish ”conceptual umbrella of *entity* and *its ecosystem*” and Schroeder's view that “Information can be defined in terms of the categorical opposition of one and many, leading to two manifestations of information, selective and structural. These manifestations of information are dual in the sense that one always is associated with the other.” Here information is directly related with computation defined as information processing. [3]

The frequent objection against the computational view of the universe, elaborated by Zenil in this volume, is that ”it is hard to see how any physical system would not be computational.” The next frequently mentioned issue is: if the universe computes, what are the input and the output of its computation? This presupposes that a computing system must have an input from the outside and that it must deliver some output to the outside world. But actor system [2] for example needs no input. Within pancomputationalist framework, the whole universe computes its own next state from its current state [4]. As all of physics is based on quantum mechanical layer of information processing, zero-point (vacuum) oscillations can be seen as constant input for the computational network of the universe. What causes different processes in the universe is the interaction or exchange of information between its parts. The universe is a result of evolution from the moment of big-bang or some other primordial state, through the complexification of the relationships between its actors by computation as a process of changes of its informational structure. Physical forces are established through particle exchanges (message exchanges) which necessarily connect particles into a web of physical interactions which are manifestation of natural laws. The whole of the universe is in the state of permanent flow, far from steady state, which results in forming increasingly complex structures, [5]. So much on the input-output objection.

As to the objection that not all of the universe can be computational, as it is a too powerful a metaphor, [6] it is essential to keep in mind the layered architecture of the computing nature, as not all of computation is the same – computation is proceeding on many scales, on many levels of hierarchical organization. Moreover, in tandem with computation, universe is described by information, representing its structures. *Given that computation follows physical laws*, or *represents/implements physical laws*, generative model of the universe can be devised such that some initial network of informational processes develops in time into increasingly complex (fractal, according to Kurakin, [5]) information structures.

The parallel could be drawn between natural computationalism and atomic theory of matter which is another general theory which implies that all of matter is made of atoms (and void). We may also say that all of physics (structures and processes) can be derived from elementary particles (and void that is an ocean of virtual particles which for short time, obeying Heisenberg uncertainty relations, pop into existence and quickly thereafter disappear). This does not make the world a soup of elementary particles where no differences can be made, and nothing new can emerge. Those basic elements can be imagined as neodymium ball magnets from which countless structures can be constructed (in space and time, through interactions).

Unified theories are common and valued in physics and other sciences, and natural computationalism is such a unified framework. It is therefore not unexpected that physicists are found among the leading advocates of the new unified theory of informational and computational universe – from Wheeler, via Feynman, to our contemporaries such as Fredkin, Lloyd, Wolfram, Goyal and Chiribella. For the articles of latter two physicists on the topic of informational universe, see the special issue of the journal *Information* titled *Information and Energy/Matter* [7] and the special issue of the journal *Entropy* titled *Selected Papers from Symposium on Natural/ Unconventional Computing and its Philosophical Significance* [8].

Conceptualizing the physical world as a network of information networks evolving through processes of natural computation helps us to make more compact and coherent models of nature, connecting non-living and living worlds. It presents a suitable basis for incorporating current developments in understanding of biological, cognitive and social systems as generated by complexification of physicochemical processes see Deacon [9] through self-organization of molecules into dynamic adaptive complex systems that can be understood as morphogenesis, adaptation and learning—all of which can be understood as computation (information processing), [50].

2 Re-Conceptualizing of Nature as Hierarchically Organized Network of Networks Computational Architecture

2.1 Natural Hierarchy

> “If computation is understood as a physical process, if nature computes with physical bodies as objects (informational structures) and physical laws govern process of computation, then the computation necessarily appears on many different levels of organization. Natural sciences provide such a layered view of nature. One sort of computation process is found on the quantum-mechanical level of elementary particles, atoms and molecules; yet another on the level of classical physical objects. In the sphere of biology, different processes (computations = information processing) are going on in biological cells, tissues, organs, organisms, and eco-systems. Social interactions are governed by still another kind of communicative/interactive process. If we compare this to physics where specific “force carriers” are exchanged between elementary particles, here the carriers can be complex chunks of information such as molecules or sentences and the nodes (agents) might be organisms or groups—that shows the width of a difference.” [3]

Searching for a framework for natural computation and looking at nature from variety of perspectives and levels of organization, Cottam et al. in this book address the general question of hierarchy in nature and point to Salthe who ”*restricts the term hierarchy to two forms: the scale (or compositional) hierarchy and the specification (or subsumption) hierarchy.” However, they find that for the description of ntural systems a third form they name* the representation or model hierarchy *is most suitable. The central tenet is the birational ecosystemic principle: “Nature seen through sciences brings all of Science under a generalized umbrella of* entity and its ecosystem*, and then characterizes different types of entity by their relationships with their relevant ecosystems.”* (emphasis added)

In spite of suggested tree-structure with representation on top, followed by model hierarchy with subsequent compositional and subsumption hierarchy, the authors emphasize the movement between the bottom and the top. Parts define the whole, which once established, affect its parts. As a case in point, they provide an example of a (Natural) model hierarchy for a tree represented at different scales: ”{a tree described in terms of atoms}, {a tree described in terms of molecules}, {a tree described in terms of cells}… up to {a tree described in terms of branches}, {a tree as itself – a tree}”. Here inter-scale interfacing and consequently digital-analog interfaces are discussed and it is pointed out that naturally-hierarchical multi-scale organisms function qualitatively differently from a digital computer. The article concludes with the hope that this birational ecosystemic hierarchical framework will be capable of providing a new definition of computation, closer to physical processes in nature.

2.2 Cognitive Level of Information Processing

In a hierarchy of organizational levels in nature the most complex level of information processing is cognitive level and it subsumes all lower levels that successively emerge from their antecedent lower levels. Lindley in this volume addresses the problems encountered in the development of engineered autonomous and intelligent systems caused by *exclusively linguistic models of intelligence*. The alternative proposed is "taking inspiration more directly from biological nervous systems". This approach is argued to be able to go "far beyond twentieth century models of artificial neural networks (ANNs), which greatly oversimplified brain and neural functions". This implies study of computation as information processing in neural and glial systems in order to *implement* "asynchronous, analog and self-* architectures that digital computers can only *simulate*." (emphasis added) The difference between physical process as it appears in nature and its computational simulation is essential when we not only talk about them and represent them them, but also use them as computational resources in AI systems.

Continuing on the level of neural systems, Phillips's paper addresses the important topic of coordination of concurrent probabilistic inference. Adaptively organized complexity of life builds on information processing and in cognitive agents with neural systems also on inference. The paper discusses the theory of Coherent Infomax in relation to the Theory of free energy reduction of probabilistic inference. Coherent Infomax shows how neural systems combine local reliability with context-sensitivity and here we recognize the leitmotif from several other papers: individual in relation to the social, or agent and its eco-system.

Basti and Dodig-Crnkovic recognize significant role Turing played as a pioneer of natural computing, especially in the field of morphological computing and neural networks (unorganized machines).

Bull et al. in this book address Turing's unorganized machines as models of neural systems. Turing in his 1948 paper [10] made an essential insight about the *connection of social aspects of learning and intelligence*. From the contemporary perspective of natural computing we see networks as information processing mechanisms and their role in intelligence is fundamental. Suggesting that natural evolution may provide inspiration for search mechanisms to design machines, Bull et al investigate Turing's dynamical representation for networks of vesicles (membrane-bound compartments with Belousov-Zhabotinsky mixture) used as liquid information processing system. Communication between vesicles via chemical signals - excitations propagating between vesicles, was seen as imitation or cultural information communication. The authors hope that this may provide "a useful representation scheme for unconventional computing substrates".

Arriola-Rios, Demery, Wyatt, Sloman and Chappell contribute to this book with a study of object representation in animals (especially parrots) and robots from segregate information about physical objects. This work helps better understanding of the mechanisms of information processing on the cognitive level. Information in a cognizing agent forms internal representations dependent on the way of its use, or the way of the interaction of the agent with the environment. Information could be compressed and re-used for interpretations, and identification of causal relationships and functions. It is described "how a selection of key elements from the environment could be used as a basis for an object representation, and considered possible underlying exploration strategies for gathering information by observing natural behaviour." Particular analysis is devoted reasoning about deformable objects through key frames.

Still on the cognitive level of an agent processing information through its physical structures, it is instructive to connect to Cooper's observation about Turing's model.

> "We might have been told at some point that it was devised as a disembodying model of machine computation. Not so, of course. 2012 has made everyone aware of the very specific physicality of the computing situation that Turing was modelling, the predominantly women `computers' following instructions. (…) The underlying physicality may be highly complex. But such things as the human computer's aches and pains, her feelings of hunger or boredom, are factored out of the process."

Yet, on a more basic level, "her feelings of hunger or boredom" were part of her being alive which made it possible for her to compute. If we want to construct such self-sustained, intelligent, adaptive computers capable of not only following instructions but even creating new algorithms, we might need to take the boredom and joy and other human characteristics into our broader model of computing. Those are qualities that may fuel creativity, even though they act as disturbance when performing lengthy mechanical calculations. Nature uses both mechanical and creative computing in cognitive agents.

3 The Unreasonable Effectiveness of Mathematics in the Natural Sciences (Except for Biology). Mathematicians Bias and Computing Beyond the Turing Limit

Mathematician's contribution to the development of the idea of computing nature is central. Turing as an early proponent of natural computing put forward a machine model that is still in use. How far can we hope to go with Turing machine model of computation?

In the context of computing nature, living systems are of extraordinary importance as up to now science haven't been able to model and simulate the behavior of even the simplest living organisms. "The unreasonable effectiveness of mathematics" observed in physics (Wigner) is missing for complex phenomena like biology that today lack effective mathematical models (Gelfand), see Chaitin [11].

Not many people today would claim that human cognition (information processing going on in our body, including brain) can be adequately modeled as a result of computation of one Turing machine, however complex function it might compute. In the next attempt, one may imagine a complex architecture of Turing machines running in parallel as communicating sequential processes (CSPs) exchanging information. We know today that such a system of Turing machines cannot produce the most general kind of computation, as truly asynchronous concurrent information processing going on in our brains. [4]

However, one may object that IBM's super-computer Watson, the winner in man vs. machine "Jeopardy!" challenge, runs on contemporary (super)computer which is claimed to be implementation of the Turing machine. Yet, Watson is connected to the Internet. And Internet is not a Turing machine. It is not even a network of Turing

Machines. Information processing going on throughout the entire Internet includes signaling and communication based on complex asynchronous physical processes that cannot be sequentionalized. (Hewitt, Sloman) As an illustration see Barabási et al. article [12] on parasitic computing that *implements computation on the communication infrastructure of the Internet, thus using communication for computation.*

Zenil in this volume examines the question: "What does it mean to claim that a physical or natural system computes?" He proposes a behavioural characterisation of computing in terms of a measure of programmability, which reflects a system's ability to react to external stimuli. To that end Zenil investigates classical foundations for unconventional computation.

Cooper in his chapter "What Makes a Computation Unconventional?" investigates the relationships between method and matter, process and embodiment. He addresses the phenomenon of emergence, which is recurrent theme of this book, but in Coopers approach emergence is related to unconventional, higher type computation:

> "Although Stephen Kleene provided formal content to the notion of higher type computation via a series of papers spanning over 30 years (1959 - 1991), the physical relevance of his take on the topic needs to be clarified. A forthcoming book on "Computability At Higher Types" by John Longley and Dag Normann is eagerly anticipated. The intuition is that computational unconventionality certainly entails higher type computation, with a correspondingly enhanced respect for embodied information."

Hernandez-Espinosa and Hernandez-Quiroz, starting from the old computationalism defined as the belief that the human mind can be modeled by Turing Machines, analyze Wolfram's Principle of Computational Equivalence based on his studies of cellular automata – the claim that "any natural (and even human) phenomenon can be explained as the interaction of very simple rules." The next step in cellular automata models may be to replace present basic simple rules of cellular automata with more elaborate ones. Instead of synchronous update of the whole system; they can be made asynchronous networks of agents, placed in layered architectures on different scales etc. Here we recognize the basic idea of generative science which is to generate apparently unanticipated and infinite behaviour based on deterministic and finite rules and parameters reproducing or resembling the behavior of natural and social phenomena. As an illustration see Epstein, [13].

If we want to generalize the idea of computation so to be able to encompass more complex operations than mechanical execution of an algorithm, simulating not only a person executing strictly mechanical procedure, but the one constructing a new theory, we must go back to underlying mathematics.

While Cooper in this volume asks *"To what extent can the explanatory power of the mathematics clarify key issues relating to emergence, basic physics, and the supervenience of mentality on its material host?"* Dodig-Crnkovic and Burgin investigate the explanatory power of mathematics. They analyze methodological and philosophical implications of algorithmic aspects of unconventional/natural computation that extends the closed classical universe of computation of the Turing machine type. The new extended model constitute an open world of algorithmic constellations, allowing increased flexibility and expressive power, supporting constructivism and creativity in mathematical modeling and enabling richer understanding of computation, see [14].

3.1 Hypercomputation - Beyond the Turing Limit

Hypercomputation is the research field that formulated the first ideas about the possibility of computing beyond Turing machine model limits. The term hypercomputation was introduced by Copeland and Proudfoot [15]. The expression "super-Turing computation" was coined by Siegelman and usually implies that the model is physically realizable, while hypercomputation in general typically relies on thought experiments. Present volume offers two contributions that sort under hypercomputation, written by Franchette and Douglas.

Franchette studies the possibility of a physical device that hypercomputes by building an oracle hypermachine, which would be a device able to use external information from nature in order to go beyond Turing machines limits. The author also addresses an analysis of the verification problem for oracle hypermachines.

Douglas in his contribution presents a critical analysis of Siegelmann Networks.

3.2 Physical Computation “In Materio” - Beyond the Turing Limit

Several authors at the Symposium on Natural/Unconventional Computing at AISB/IACAP World Congress 2012 (Stepney, Cooper, Goyal, Basti, Dodig-Crnkovic) underlined the importance of physical computing, or as Stepney [16] termed it, “computation in materio”. Along the same lines, Cooper in his article Turing's Titanic Machine? [17] diagnoses the limitations of the Turing machine model and identifies the ways of overcoming those limitations by introducing:

- Embodiment invalidating the `machine as data' and universality paradigm.
- The organic linking of mechanics and emergent outcomes delivering a clearer model of supervenience of mentality on brain functionality, and a reconciliation of different levels of effectivity.
- A reaffirmation of experiment and evolving hardware, for both AI and extended computing generally.
- The validating of a route to creation of new information through interaction and emergence.

Related article by the same author, The Mathematician's Bias and the Return to Embodied Computation, in [18], analyses the role of physical computation vs. universal symbol manipulation.

The theme of embodied computation is addressed in this volume by Hernandez-Quiroz and Padilla who examine actual physical realizability of mathematical constructions of abstract entities - a controversial issue and important in the debate about the limits of the Turing model. The authors study a special case of physical realizability of the enumeration procedure for rational numbers via Cantor's diagonalization by an Ising system.

3.3 Higher Order Computability - Beyond the Turing Limit

One of the main steps towards the new paradigm of natural/unconventional computing is to make visible host of myths which are surrounding the old paradigm and helping it to survive. One of those myths is that our modern computers with all their programming languages are just diverse implementations of Turing machines. However, as Kanneganti and Cartwright already argued twenty years ago:

> "Classic recursion theory asserts that all conventional programming languages are equally expressive because they can define all partial recursive functions over the natural numbers. This statement is misleading because programming languages support and enforce a more abstract view of data than bit strings. In particular, most real programming languages support some form of higher-order data such as potentially infinite streams (input and output), lazy trees, and functions." Kanneganti and Cartwright [19]

Kleene was a pioneer of higher order computability as he "opened the frontiers of computability on higher type objects in a series of papers first on constructive ordinals and hierarchies of number-theoretical predicates and later on computability in higher types." Soare [20]

Also Cooper [21] underlines the importance of higher-order computational structures as characteristic of human thinking. This can be connected to higher-order functional programming, which means, among others, programming with functions whose input and/or output may consist of other functions.

> "Kreisel [21] was one of the first to separate cooperative phenomena (not known to have Turing computable behaviour), from classical systems and proposed [22] (p˙143, Note 2) a collision problem related to the 3-body problem as a possible source of incomputability, suggesting that this might result in "an analog computation of a non-recursive function (by repeating collision experiments sufficiently often)". This was before the huge growth in the attention given to chaos theory, with its multitude of different examples of the generation of informational complexity via very simple rules, accompanied by the emergence of new regularities (see for example the two classic papers of Robert Shaw [33], [32]). We now have a much better understanding of the relationship between emergence and chaos, but this still does not provide the basis for a practically computable relationship." Cooper [21] (emphasis added)

4 Concurrent Computing and Turing Machine Model

4.1 Bi-directional Model Development of Natural Computation

Turing machine (originally named "logical calculating machine") model of computation was developed by Turing in order to describe a human (at that time called" a computer") executing an algorithm:

> "It is possible to produce the effect of a computing machine by writing down a set of rules of procedure and asking a man to carry them out. Such a combination of a man with written instructions will be called a 'Paper Machine'. A man provided with paper, pencil, and rubber, and subject to strict discipline, is in effect a universal machine." Turing [10]

The underlying logic of Turing's "logical calculating machine" is fully consistent standard logic. Turing machine is assumed always to be in a well defined state. [2] In contemporary computing machinery, however, we face both states that are not well defined (in the process of transition) and states that contain inconsistency:

> "Consider a computer which stores a large amount of information. While the computer stores the information, it is also used to operate on it, and, crucially, to infer from it. Now it is quite common for the computer to contain inconsistent information, because of mistakes by the data entry operators or because of multiple sourcing. This is certainly a problem for database operations with theorem-provers, and so has drawn much attention from computer scientists. Techniques for removing inconsistent information have been investigated. Yet all have limited applicability, and, in any case, are not guaranteed to produce consistency. (There is no algorithm for logical falsehood.) Hence, even if steps are taken to get rid of contradictions when they are found, an underlying paraconsistent logic is desirable if hidden contradictions are not to generate spurious answers to queries." Priest and Tanaka [22]

Open, interactive and asynchronous systems have special requirements on logic. Goldin and Wegner [23], and Hewitt [2] argue e.g. that computational logic must be able to model interactive computation, and that classical logic must be robust towards inconsistencies i.e. must be paraconsistent due to the incompleteness of interaction.

As Sloman [24] argues, concurrent and synchronized machines are equivalent to sequential machines, but some concurrent machines are asynchronous, and thus not equivalent to Turing machines. If a machine is composed of asynchronous concurrently running subsystems, and their relative frequencies vary randomly, then such a machine cannot be adequately modeled by Turing machine, see also [4].

Turing machines are discrete but can in principle approximate machines with continuous changes, yet cannot implement them exactly. Continuous systems with non-linear feedback loops may be chaotic and impossible to approximate discretely, even over short time scales, see [25] and [2]. Clearly Turing machine model of computation is an abstraction and idealization. In general, instead of idealized, symbol-manipulating models, more and more physics-inspired modeling is taking place.

Theoretical model of concurrent (interactive) computing corresponding to Turing machine model of algorithmic computing is under development. (Abramsky, Hewitt, Wegner) From the experience with present day networked concurrent computation it becomes obvious that Turing machine model can be seen as a special case of a more general computation. During the process of learning from nature how to compute, we both develop computing and at the same time improve understanding of natural phenomena.

> “In particular, the quantum informatic endeavor is not just a matter of feeding physical theory into the general field of natural computation, but also one of using high-level methods developed in Computer Science to improve on the quantum physical formalism itself, and the understanding thereof. We highlight a seemingly contradictory phenomenon: passing to an abstract, categorical quantum informatics formalism leads directly to a simple and elegant graphical formulation of quantum theory itself, which for example makes the design of some important quantum informatic protocols completely transparent. It turns out that essentially all of the quantum informatic machinery can be recovered from this graphical calculus. But in turn, this graphical formalism provides a bridge between methods of logic and computer science, and some of the most exciting developments in the mathematics of the past two decades“ Abramsky and Coecke [25]

The similar two-way process of learning is visible in biocomputing, see Rozenberg and Kari [26]. As we already mentioned “the unreasonable effectiveness of mathematics in the natural sciences” does not (yet) apply to biology, as modeling of biological systems attempted up to now was too crude. Living systems are essentially open and in constant communication with the environment. New computational models must be interactive, concurrent, and asynchronous in order to be applicable to biological and social phenomena and to approach richness of their information processing repertoire.

Present account of models of computation highlights several topics of importance for the development of new understanding of computing and its role: natural computation and the relationship between the model and the physical implementation, interactivity as fundamental for computational modeling of concurrent information processing systems (such as living organisms and their networks), and new developments in logic needed to support this generalized framework. Computing understood as information processing is closely related to natural sciences; it helps us recognize connections between sciences, and provides a unified approach for modeling and simulating of both living and non-living systems. [4]

4.2 Concurrency and Actor Networks in Nature All the Way Down

In his article: What is computation? Concurrency versus Turing's Model, Hewitt [2] makes the following very apt analysis of the relationship between Turing machines and concurrent computing processes:

> “Concurrency is of crucial importance to the science and engineering of computation in part because of the rise of the Internet and many-core architectures. However, concurrency extends computation beyond the conceptual framework of Church, Gandy [1980], Gödel, Herbrand, Kleene [1987], Post, Rosser, Sieg [2008], Turing, etc. because there are effective computations that cannot be performed by Turing Machines. In the Actor model [Hewitt, Bishop and Steiger 1973; Hewitt 2010], computation is conceived as distributed in space where computational devices communicate asynchronously and the entire computation is not in any well-defined state. (An Actor can have information about other Actors that

> it has received in a message about what it was like when the message was sent.) Turing's Model is a special case of the Actor Model." Hewitt [2] (emphasis added)

According to natural computationalism/pancomputationalism [4] every physical system is computational, but there are many different sorts of computations going on in nature seen as a network of agents/actors exchanging "messages". The simplest agents communicate with simplest messages such as elementary particles (with 12 kinds of matter and 12 kinds of anti-matter particles) exchanging 12 kinds of force-communicating particles. Example from physics that we can recast into actor model is Yukawa's theory of strong nuclear force modeled as exchange of mesons (as messages), which explained the interaction between nucleons. Complex agents/actors like humans communicate through languages which use very complex messages for communication. Also, exchange of information causes change of actors. Those changes are simple in simple actors such as elementary particle that can change its state (quantum numbers) and in complex agents with memory, communication results in substantial changes in agents' way of response.

Natural computational systems as networks of agents exchanging messages are in general asynchronous concurrent systems. Conceptually, agent-based models and actor models are closely related, and as mentioned, understanding of interactions between agents in interaction networks fits well in those frameworks.

Physical Computing - New Computationalism.
Non-symbolic vs. Symbolic Computation

It is often argued that computationalism is the opposite of connectionism and that connectionist networks and dynamic systems do not compute. This implied that human mind as a processes powered by human brain as a network of neurons cannot be adequately modeled in computational terms. However, if we define computation in a more general sense of natural computation, instead of high level symbol manipulation of Turing machine, it is obvious that *connectionist networks and dynamical systems do compute*. Computational modeling of cognitive processes requires computing tools that contain not only Turing Machine model but also connectionist network models. That is also the claim made by Scheutz in the Epilogue of the book Computationalism: New Directions [27], where he notices that:

> "Today it seems clear, for example, that classical notions of computation alone cannot serve as foundations for a viable theory of the mind, especially in light of the real-world, real-time, embedded, embodied, situated, and interactive nature of minds, although they may well be adequate for a limited subset of mental processes (e.g., processes that participate in solving mathematical problems). Reservations about the classical conception of computation, however, do not automatically transfer and apply to real-world computing systems. This fact is often ignored by opponents of computationalism, who construe the underlying notion of computation as that of Turing-machine computation." Scheutz [27] p. 176

Classical computationalism was the view that classical theory of computation (Turing-machine-based, universal, and disembodied) might be enough to explain cognitive phenomena. New computationalism (natural computationalism) emphasizes that embodiment is essential and thus physical computation, hence natural computationalism.

The view of Scheutz is supported by O'Brien [28] who refers to Horgan and Tienson [29] arguing that "cognitive processes, are not governed by exceptionless, representation-level rules; they are instead the work of defeasible cognitive tendencies subserved by the non-linear dynamics of the brains neural networks."

Dynamical characterization of the brain is consistent with the analog interpretation of connectionism. But dynamical systems theory is often not considered to be a computational framework. O'Brien [28] notices that *"In this sense, dynamical systems theory dissolves the distinction between intelligent and unintelligent behaviour, and hence is quite incapable, without supplementation, of explaining cognition. In order for dynamical engines to be capable of driving intelligent behaviour they must do some computational work: they must learn to behave as if they were semantic engines."*

O'Brien and Opie [30] thus search for an answer to the question how connectionist networks compute, and come with the following characterization:

> "Connectionism was first considered as the opposed to the classical computational theory of mind. Yet, it is still considered by many that a satisfactory account of how connectionist networks compute is lacking. In recent years networks were much in focus and agent models as well so the number of those who cannot imagine computational networks has rapidly decreased. Doubt about computational nature of connectionism frequently takes the following two forms.

1. (W)hile connectionists typically interpret the states and activity of connectionist networks in representational terms, closer scrutiny reveals that these putative representations fail to do any explanatory work, and since there is ''no computation without representation'' (Pylyshyn 1984, p. 62), the connectionist framework is better interpreted non-computationally.
2. "the connectionist networks are better characterized as dynamical systems rather than computational devices."

In the above denial of computational nature of connectionist models the following confusions are evident.

1. Even though it is correct that there is "no computation without representation", representation in this context can be any state of activation in a cognizing agent that causes the agent to "recognize" the information. It can be a dynamical state induced in the agents' brain as a consequence of perception and that dynamical state, even though it has no apparent resemblance of the source of information, is causally connected to it.
2. Dynamical systems compute and their computation in general is natural computation. One of the central questions in this context is the distinction between symbolic and non-symbolic computing. Trenholme [31] describes the relationship of analog vs. symbolic simulation:

> "Symbolic simulation is thus a two-stage affair: first the mapping of inference structure of the theory onto hardware states which defines symbolic computation; second, the mapping of inference structure of the theory onto hardware states which (under appropriate conditions) qualifies the processing as a symbolic simulation. Analog simulation, in contrast, is defined by a single mapping from causal relations among elements of the simulation to causal relations among elements of the simulated phenomenon." Trenholme [31] p.119. (emphasis added)

Both symbolic and sub-symbolic (analog) simulations depend on causal/analog/physical and symbolic type of computation on some level but *in the case of symbolic computation it is the symbolic level where information processing is observed*. Similarly, even though in the analog model symbolic representation exists at some high level of abstraction, it is the physical agency of the substrate and its causal structure that define computation (simulation).

Basti in this volume suggests how to "integrate in one only formalism the physical ("natural") realm, with the *logical-mathematical* ("computation"), studiying their relationships. That is, the passage from the realm of the *causal* necessity ("natural") of the physical processes, to the realm of the *logical* necessity ("computational"), and eventually representing them either in a sub-symbolic, or in a symbolic form. This foundational task can be performed, by the newborn discipline of *theoretical formal ontology."* Proposed formal ontology is based on the information-theoretic approach in quantum physics and cosmology, the information-theoretic approach of dissipative QFT (Quantum Field Theory) and the theoretical cognitive science.

Freeman offers an accurate characterization of the relationship between physical/sub-symbolic and logical/symbolic level in the following passage:

> *"Human brains intentionally direct the body to make symbols, and they use the symbols to represent internal states.* The symbols are outside the brain. Inside the brains, the construction is effected by spatiotemporal patterns of neural activity that are operators, not symbols. *The operations include formation of sequences of neural activity patterns that we observe by their electrical signs.* The process is by neurodynamics, not by logical rule-driven symbol manipulation. *The aim of simulating human natural computing should be to simulate the operators. In its simplest form natural computing serves for communication of meaning. Neural operators implement non-symbolic communication of internal states by all mammals, including humans, through intentional actions. (…)* I propose that symbol-making operators evolved from neural mechanisms of intentional action by modification of non-symbolic operators." [32] (emphasis added)

Consequently, our brains use non-symbolic computing internally in order to manipulate relevant external symbols/objects!

4.3 Physical Computation/Natural Computation vs. Turing Machine Model

So in what way is physical computation/natural computation important vis-à-vis Turing machine model? One of the central questions within computing, cognitive science, AI and other related fields is about computational modeling (and simulating) of intelligent behaviour. What can be computed and how? It has become obvious that we must have richer models of computation, beyond Turing machine, if we are to adequately model and simulate biological systems. What exactly can we learn from nature and especially from intelligent organisms?

It has taken more than sixty years from the first proposal of the test Turing called the ”Imitation Game", as described in Turing [33] p. 442, to the Watson machine winning Jeopardy. That is just the beginning of what Turing believed one day will be possible - a construction of computational machines capable of generally intelligent behavior as well as the accurate computational modeling of the natural world. So there are several classes of problems that deserve our attention when talking about computing nature.

To ”*compute*” nature by any kind of computational means, is to model and/or simulate the behaviors of natural systems by computational means. Watson is a good example. We know that we do not function like Watson or like chess playing programs that take advantage of brute force algorithms to search the space of possible states. We use our ”gut feeling” and ”fingertip-feeling”/ ”fingerspitzengefühl” and they can be understood as embodied, physical, sub-symbolic information processing mechanisms we acquire by experience and use when necessary as automatized *hardware-based, automatic* recognition tools.

To *compute nature* means to interpret natural processes, structures and objects as a result of natural computation which is in general defined as information processing. This implies understanding and modeling of physical agents, starting from the fundamental level of quantum computing via several emergent levels of chemistry, biology, cognition and extended cognition (social, and augmented by computational/ information processing machinery).

At the moment we have bits and pieces of the picture – *computing* nature, that is computational modeling of nature and computing *nature*, that is nature understood in itself as a computational network of networks.

5 The Relationship between Human Representation, Animal Representation and Machine Representation

We would like to highlight the relevance of the relationship between human representation and machine representation to show the main issues concerning “functionalism” and “connectionism”. We propose to discuss the notion of “representation” because an important challenge for AI is to simulate not only the “phonemic” and “syntactic” aspects of mental representation but also the “semantic” aspect. Traditionally, philosophers use the notion of “intentionality” to describe the representational nature of mental states namely intentional states are those that “represent” something, because mind is directed toward objects. We think that it is important to consider the relevance of “embodied cognition” for contentful mental states (see, for instance, the classical thought experiment of the “Chinese room” introduced by Searle to criticize the important results of the Turing test, [34]).

The challenge for AI is therefore to approximate to human representations i.e. to the semantic content of human mental states. There are two competing interpretations of mental representations relevant for AI. The first focuses on the discreteness of mental representations and the second focuses on their inter-relation [35]. The first corresponds to the symbolic paradigm in AI, according to which mental representations are symbols. Proponents of the symbolic representation point on a semantic that rests on the relation between tokens of the symbol and objects of representation. The intentional mechanism functions in a way that the content of a symbol does not depend on the content of other symbols. In this sense, each symbol is discretely conferred with its intentional content. The second corresponds to connectionism in AI, according to which mental representations are distributed patterns. Proponents of this view intend the way in which a mental representation is conferred with its intentional content as mediated by relations with other representations. The virtue of connectionism as presented in the neural networks resides in the fact that the categories represented admit borderline cases of membership. As regards the composition of mental representations, it reveals itself to be the complex, contextually modulated interaction of patterns of activation in a highly interconnected network. We aim to describe the main aspects of the two approaches to make clear: the mechanisms characterizing the different way by which representations are conferred with their intentional content; the nature and structure of the categories represented and the ways in which mental representations interact.

The task to consider the similarity between human and artificial representation could involve the risk of skepticism about the possibility of "computing" this mental capacity. If we consider computationalism as defined in purely abstract syntactic terms then we are tempted to abandon it because human representation involves "real world constrains". But, a new view of computationalism could be introduced that takes into consideration the limits of the classical notion and aims at providing a concrete, embodied, interactive and intentional foundation for a more realistic theory of mind [27]. We would like to highlight also an important and recent debate on "digital representation" [36] that focus on the nature of representations in the computational theory of mind (or computationalism). The starting point is the nature of mental representations, and, particularly, if they are "material". There are authors such as Clark who maintain that mental representation are material [37] while others like Speaks think that thought processes use conventional linguistic symbols [38]. The question of digital representation involves the "problem of physical computation" [39] as well as the necessity of the notion of representation [40] so that we only have the problem of how to intend the very notion of representation [41, 42]. But, there is also the possibility of understanding computation as a purely physical procedure where physical objects are symbols processed by physical laws on different levels of organization that include "every natural process" in a "computing universe" [43]. In this context, we need a plausible relation between computation and information. Info-computational naturalism describes the informational structure of the nature i.e. a succession of level of organization of information. Morphology is the central idea in the understanding of the connection between computation and information. It proceeds by abstracting the principles via information self-structuring and sensory-motor coordination. The sensory-motor coordination provides an "embodied" interaction with the environment:

information structure is induced in the sensory data, thus facilitating perception, learning and categorization.

Among the possibilities to compute human representational processes, Basti in this volume proposes a natural account from the field of formal ontology. In particular, he implements the so-called "causal theory of reference" in dynamic systems.

We think it is necessary to find a plausible philosophical strategy to consider the capacities that are common to human and machine representation (Giovagnoli in this volume). Analytic Pragmatism that is represented by the American philosopher Brandom [44] suggests relevant ideas to describe human, animal and artificial capacities for representing the external world. It is easier to start with the human case and so to describe discursive practices and to introduce norms for deploying an autonomous vocabulary, namely a vocabulary of a social practice (science, religion etc.). These norms are logical and are at the basis of an "inferential" notion of representation. But, inference in this sense, recalling Frege, is material. Brandom refuses the explanation of representation in terms of syntactical operations as presented by "functionalism" in "strong" artificial intelligence (AI or GOFAI). He does not even accept weak AI (Searle), rather he aims to present a "logical functionalism" characterizing his analytic pragmatism. According to Brandom, we are not only creatures who possess abilities such as to respond to environmental stimuli we share with thermostats and parrots but also "conceptual creatures" i.e. we are logical creatures in a peculiar way and we need a plausible view to approach human capacities.

Very interesting results are offered by Arriola-Rios and Demery et al. who discuss in this book how salient features of objects can be used to generate compact representations in animals and robots, later allowing for relatively accurate reconstructions and reasoning. They would like to propose that when exploration of objects occurs for forming representations, it is not always random, but also structured, selected and sensitive to particular features and salient categorical stimuli of the environment. They introduce how studies into artificial agents and into natural agents are complementary by emphasizing some findings from each field.

Along this line, Bull, Holley, De Lacy Costello and Adamatzky present initial re-sults from consideration of using Turing's dynamical representation within uncon-ven-tional substrate – networks of Belousov-Zhabotinsky vehicles – designed by an imitation based i.e. cultural approach. Over sixty years ago, Alan Turing presented a simple representation scheme for machine intelligence namely a discrete dynamical system network of two-input NAND gates. Since then only a few other explorations of these unorganized machines are known. As the authors underscore in their paper, it has long been argued that dynamic representations provide numerous useful features, such as an inherent robustness to faults and memory capabilities by exploiting the structure of their basins of attraction:

> "For example, unique attractors can be assigned to individual system states/outputs and the map of internal states to those attractors can be constructed such that multiple paths of similar states lead to the same attractor. In this way, some variance in the actual path taken through states can be varied, e.g., due to errors, with the system still responding appropriately. Turing appears to have been thinking along these lines also".

6 Conclusions, Open Problems and Future Work

"It turns out to be better to use the world as its own model." Brooks [45]

As already argued, we enjoy and appreciate what Wigner named "*the unreasonable efficiency of mathematics in natural sciences*" [46] – except for biology. Time is right to address biology at last and try to find out how best to use computation to model and simulate behavior of biological systems. In this context it can never be overemphasized that*:"nothing in biology makes sense except in the light of evolution"* – an insight made by the evolutionary biologist Dobzhansky [47]. In order to model (simulate) evolution we need generative models. As demonstrated by e.g. Epstein [13] and Wolfram [48], such models are capable of producing complex behaviors starting from simple structures and processes (rules).

Of all biological phenomena, cognition (the ability of living organisms to process information beyond simple reactivity) seems to be the most puzzling one, as in more complex organisms it is related to phenomena such as mind, intelligence and mental (thought) processes. Cognition in highly developed organisms indeed looks like a miracle if one does not take into account that it took several billions of years in nature to develop through the process of evolution from simplest forms to increasingly complex ones. Through the reverse engineering of evolution we are learning how organisms function through computational models such as *Human Brain/Blue Brain project*. At the same time we learn to compute in novel and more powerful ways, such as developed in the IBM's project on *Cognitive Computing*.

For the future work on computing nature it remains to reconstruct the process of evolution of life in terms of information and computation (as information processing), starting from the process of abiogenesis i.e. the transition from amino acids to first living organisms. Of special interest are the evolution of nervous systems and brains in animals and thus the development of complex cognitive capacities, such as intelligence. This understanding of the evolution and the development in terms of information and computation will lead to improved understanding of underlying mechanisms of morphological computing as information self-structuring [48].

Here follows the list of some important questions to answer in the framework of natural computation (information processing in physical systems).

- *Generative modeling of the evolution and development of physical structures of the universe,* starting with minimum assumptions about primordial universe in terms of information and computation, based on actor (agent) networks exchanging information (messages/particles).
- *Generative modeling of hierarchical structure of emergent layers of organization* in physical systems in terms of natural computing. Modeling of the process in which the whole constraints its parts and showing how its (higher level) properties emerge.
- *Understanding and describing of the evolution and development of living organisms* on earth within the framework of natural computation (morphological computation, self-organization of informational structures through computational processes – concurrent computational processes, modeled as above. [49]

– *Understanding intelligence and consciousness,* in terms of information and computation. Explaining how representations (symbolic level) emerge from sub symbolic information processes. Understanding how exactly our brains process information, learn and act in terms of information and natural computation on different levels of organization. Working out the connections between connectionist networks/dynamic systems and symbol manipulation, sub-symbolic and symbolic information processing.
– *Explaining how physics connects to life* and how the fact that we evolved from physical matter defines the ways we interact with the universe and form our concepts and actions (observer problem in epistemology), continuation of the project started by Deacon [9].
– *Uncovering details of info-computational mechanisms* involved in DNA control of cellular processes.
– *Application of natural computation to program nano-devices* and universally programmable intelligent matter. [50]
– *Answering questions for which natural computationalism is especially suitable framework,* such as: why is the genetic difference between humans and other animals much smaller than we imagined before genome sequencing? How does the evident difference between humans and apes developed, given our social communication system as computational infrastructure that acts as a basis of human social intelligence established by natural computing?

From all above proposed research a richer notion of computation will emerge, which in its turn will help in the next step to better address natural phenomena as computations on informational structures. As Penrose in the foreword to [18] states:

> "(S)ome would prefer to define "computation" in terms of what a physical object can (in principle?) achieve (Deutsch, Teuscher, Bauer and Cooper). To me, however, this begs the question, and this same question certainly remains, whichever may be our preference concerning the use of the term "computation". If we prefer to use this "physical" definition, then all physical systems "compute" by definition, and in that case we would simply need a different word for the (original Church-Turing) mathematical concept of computation, so that the profound question raised, concerning the perhaps computable nature of the laws governing the operation of the universe can be studied, and indeed questioned."

With this new idea of natural computation generalizing current Turing model of computation, nature indeed can be seen as a network of networks of computational processes and what we are trying is to compute the way nature does, learning its tricks of the trade. So the focus would *not* be *computability* but *computational modeling*. How good computational models of nature are we able to produce and what does it mean for a physical system to perform computation, where computation is implementation of physical laws.

It is evident that natural computing/ computing nature presents a new natural philosophy of generality and scope that largely exceed natural philosophy of Newton's era, presented in his Philosophiae Naturalis Principia Mathematica. Natural computation brings us to the verge of a true paradigm shift in modeling, simulation and controlling the physical world, and it remains to see how it will change our understanding of nature and especially living nature including humans, their societies and ecologies.

References

1. Hodges, A.: Turing. A Natural philosopher. Phoenix (1997)
2. Hewitt, C.: What is computation? Actor Model versus Turing's Model. In: Zenil (2012)
3. Dodig Crnkovic, G.: Physical Computation as Dynamics of Form that Glues Everything Together. Information 3(2), 204–218 (2012)
4. Dodig-Crnkovic, G., Müller, V.: A Dialogue Concerning Two World Systems: Info-Computational vs. Mechanistic. In: Dodig-Crnkovic, G., Burgin, M. (eds.) Information and Computation. Information Studies, World Scientific Publishing Co. (2009), http://arxiv.org/abs/0910.5001
5. Kurakin, A.: The self-organizing fractal theory as a universal discovery method: the phenomenon of life. Theoretical Biology and Medical Modelling 8(4), 1–66 (2011), http://www.tbiomed.com/content/8/1/
6. Piccinini, G.: Computation in Physical Systems. In: Zalta, E.N. (ed.) The Stanford Encyclopedia of Philosophy (Fall 2012 edn.), http://plato.stanford.edu/archives/fall2012/entries/computation-physicalsystems/
7. Special issue of the journal Information on Information and Energy/Matter, http://www.mdpi.com/journal/information/special_issues/matter
8. Special issue of the journal Entropy: Selected Papers from Symposium on Natural/ Unconventional Computing and its Philosophical Significance, http://www.mdpi.com/journal/entropy/special_issues/unconvent_computing
9. Deacon, T.W.: Incomplete Nature: How Mind Emerged from Matter. W. W. Norton Company, New York (2011)
10. Turing, A.: Intelligent Machinery. In: Ince, D.C. (ed.) Mechanical Intelligence, pp. 107–127. Elsevier, Amsterdam (1948)
11. Chaitin, G.: Mathematics, Biology and Metabiology, Talk (2009), http://www.cs.auckland.ac.nz/~chaitin/jack.html
12. Barabási, A.-L., Freeh, V.W., Jeong, H., Brockman, J.: Parasitic computing. Nature 412, 894–897 (2001)
13. Epstein, J.: Generative Social Science: Studies in Agent-Based Computational Modeling. Princeton Studies in Complexity. Princeton University Press, Princeton (2007)
14. Dodig Crnkovic, G., Burgin, M.: Unconventional Algorithms: Complementarity of Axiomatics and Construction. Entropy 14(11), 2066–2080 (2012), http://www.mdpi.com/journal/entropy/special_issues/unconvent_computing, doi:10.3390/e14112066
15. Copeland, J., Proudfoot, D.: Alan Turing's forgotten ideas in computer science. Scientific American (April 1999)
16. Stepney, S.: The Neglected Pillar of Material Computation. Physica D: Nonlinear Phenomena 237(9), 1157–1164 (2008)
17. Cooper, S.B.: Turing's Titanic Machine? Communications of the ACM 55(3), 74–83 (2012)

18. Zenil, H. (ed.): A Computable Universe, Understanding Computation & Exploring Na-ture As Computation. World Scientific Publishing Company/Imperial College Press (2012)
19. Kanneganti, R., Cartwright, R.: What is a universal higher-order programming lan-guage? In: Lingas, A., Carlsson, S., Karlsson, R. (eds.) ICALP 1993. LNCS, vol. 700, pp. 682–695. Springer, Heidelberg (1993)
20. Soare, R.I.: Computability and Recursion. The Bulletin of Symbolic Logic 2(3) (1996)
21. Cooper, S.B.: How can nature help us compute? In: Wiedermann, J., Tel, G., Pokorný, J., Bieliková, M., Štuller, J. (eds.) SOFSEM 2006. LNCS, vol. 3831, pp. 1–13. Springer, Heidelberg (2006)
22. Priest, G., Tanaka, K.: Paraconsistent Logic. In: Zalta, E.N. (ed.) The Stanford Encyclopedia of Philosophy (Winter 2004 edn.), `http://plato.stanford.edu/archives/win2004/entries/logic-paraconsistent/`
23. Goldin, D., Wegner, P.: Paraconsistency of Interactive Computation. In: PCL 2002 (Workshop on Paraconsistent Computational Logic), Denmark (2002)
24. Sloman, A.: The irrelevance of Turing machines to AI. In: Scheutz, M. (ed.) Computationalism: New Directions, pp. 87–127. MIT Press, Cambridge (2002), `http://www.cs.bham.ac.uk/research/cogaff/00-02.html#77`
25. Abramsky, S., Coecke, B.: Physics from Computer Science. Int. Journal of Unconventional Computing 3(3), 179–197 (2007)
26. Rozenberg, G., Kari, L.: The many facets of natural computing. Commun. ACM 51, 72–83 (2008)
27. Scheutz, M. (ed.): Computationalism: New directions. MIT Press, Cambridge (2002)
28. O'Brien, G.: Connectionism, Analogicity and Mental Content. Acta Analytica 22, 111–131 (1998)
29. Horgan, T., Tienson, J.: Connectionism and the Phylosophy of Psychology. MIT Press, Cambridge (1996)
30. O'Brien, G., Opie, J.: How do connectionist networks compute? Cogn Process 7, 30–41 (2006)
31. Trenholme, R.: Analog Simulation. Philosophy of Science 61(1), 115–131 (1994), `http://www.jstor.org/stable/188292`
32. Freeman, W.J.: The Neurobiological Infrastructure of Natural Computing: Intentionality. New Mathematics and Natural Computation 5, 19–29 (2009)
33. Turing, A.M.: Computing machinery and intelligence. Mind LIX, 433–460 (1950), `http://cogprints.org/499/0/turing.html`
34. Searle, J.: Chinese room argument, 4, Scholarpedia (2009), `http://www.scholarpedia.org/article/Chinese_room_argument`
35. Carter, M.: Minds and Computers. Edimburgh University Press, Edimburgh (2007)
36. Müller, V.C.: Representation in Digital Systems. In: Briggle, A., et al. (eds.) Current Issues in Computing and Philosophy. IOS Press (2008)
37. Clark, A.: Material Symbols. Philosophical Psychology 19, 291–307 (2006); Sedivy, S.: Minds: Contents without Vehicles. Philosophical Psychology 17, 149–179 (2004)
38. Speaks, J.: Is mental content prior to linguistic meaning? Nous 40, 428–467 (2006)
39. Shagrir, O.: Why we view the brain as a computer. Synthese 153, 393–416 (2006)
40. Fodor, J.A.: The mind-body problem. Scientific American 244, 114–123 (1981)
41. O'Brien, G.: Connectionism, analogicity and mental content. Acta Analitica 22, 111–131 (1998)

42. Basti, G., Perrone, A.L.: Intentionality and Foundations of Logic: a new Approach to Neurocomputationalism. In: Kitamura, T. (ed.) What should be computed to understand and model Brain Function? From Robotics, Soft Computing, Biology and Neuroscience to Cognitve Philosophy, pp. 239–288. World Publishing, Singapore (2001)
43. Dodig-Crnkovic, G.: Epistemology naturalized: The info-computationalist approach. APA Newsletter on Philosophy and Computers 6, 9–14 (2007)
44. Brandom, R.B.: Between Saying and Doing: Towards an Analytic Pragmatism. Oxford University Press (2008)
45. Brooks, R.A.: Intelligence Without Representation. Artificial Intelligence 47, 139–159 (1991)
46. Wigner, E.: The Unreasonable Effectiveness of Mathematics in the Natural Sciences. In: Communications in Pure and Applied Mathematics, vol. 13(I). John Wiley & Sons, Inc., New York (February1960)
47. Dobzhansky, T.: Nothing in Biology Makes Sense Except in the Light of Evolution. American Biology Teacher 35 (March 1973); reprinted in Evolution versus Creationism, J., Zetterberg, P. (ed.) ORYX Press, Phoenix (1983)
48. Dodig-Crnkovic, G.: Wolfram and the computing nature. In: Zenil, H. (ed.) Irreducibility and Computational Equivalence. ECC, vol. 2, pp. 351–370. Springer, Heidelberg (2013)
49. Dodig-Crnkovic, G.: The Info-computational Nature of Morphological Computing. In: Müller, V.C. (ed.) Theory and Philosophy of Artificial Intelligence. SAPERE, vol. 5, pp. 59–68. Springer, Heidelberg (2012)
50. MacLennan, B.: Molecular Combinatory Computing for Nanostructure Synthesis and Control. In: Proceedings of IEEE Nano 2003, San Francisco, August 12-14 (2003)

A Framework for Computing Like Nature

Ron Cottam, Willy Ranson, and Roger Vounckx

The Living Systems Project, Dept. of Electronics and Informatics,
Vrije Universiteit Brussel, Pleinlaan 2, 1050 Brussels, Belgium
life@etro.vub.ac.be
http://life.etro.vub.ac.be

Abstract. We address the context within which 'Natural' computation can be carried out, and conclude that a birational ecosystemic hierarchical framework would provide for computation which is closer to Nature. This presages a major philosophical change in the way Science can be carried out. A consequence is that all system properties appear as intermediates between unattainable dimensional extremes; even existence itself. We note that Classical and Quantum mechanical paradigms make up a complementary pair. What we wish to do is to bring *all* of Science under a generalized umbrella of *entity* and its *ecosystem*, and then characterize different types of entity by their relationships with their relevant ecosystems. The most general way to do this is to move the ecosystemic paradigm up to the level of its encompassing logic, creating a complementary pair of conceivably *different* logics – one for the entity we are focusing on; one for the ecosystem within which it exists – and providing for their quasi-autonomous *birational* interaction.

Keywords: natural computation, modeling, ecosystem, birationality, hierarchy, entropy, life.

1 Introduction

Does Nature compute? By matching computational models to Natural phenomena we are effectively saying yes. But we cannot *know* for certain. All we can say is that this is how it looks. Processes in nature can be far more complex than the simple functional computation of a Turing machine – with asynchronous parallelism; recursively communicating subroutines; … Computation is a logical process by which the configuration of data may be modified. Consequently it is devoid of meaning per se, and relies on evaluation at the higher context-dependent level of information processing to confirm its procedural correctness. Our conventional concept of computation is an empirically-aligned invented paradigm, like any other Scientific paradigm, which must therefore be principally referred to Natural processes, and whose validity may be challenged at any time. Computation becomes 'more Natural' when its currently-adopted paradigm is challenged by another, more effectively Nature-referred one, and not when a previously unemployed Natural mechanism is adopted to carry out 'computation' as it is defined within the *current* paradigm.

G. Dodig-Crnkovic and R. Giovagnoli (Eds.): *Computing Nature*, SAPERE 7, pp. 23–60.
DOI: 10.1007/978-3-642-37225-4_2

So how can we compute like Nature? This is arguably what we are trying to do already. But is the framework within which we compute sufficient? We ourselves would say no. We manage well enough for systems which are close to linearity and equilibrium, but outside of that we quickly run into difficulties. So what can we change? Quantum mechanics [1] and ecosystemics [2] have led the way – both of them insisting that we cannot in general observe a system without changing it. But we have not followed. Maybe we should. This is what we will address.

What is Natural computation? And as a starting point, what is conventional computation? Conventionally, we presume that given a relevant process of 'computation', we can to some degree know precisely and accurately both an initial data configuration (at time t_1) and a resultant data configuration (at time t_2). The difficulty is always to justify the presumption. *If* we can model the process in a formally defined environment, then we reduce it to a parallelized[1] or serialized set of logical and *accessible* incremental steps whose 'correctness' is defined by the axioms of the particular formal environment. Even so, both the *belief* aspect of systemic axioms and Gödel's incompleteness theorems [3] leave us with a lack of complete justification.

Having at least partially justified our model, we then move to comparative empirical testing in a probably closed domain where accurate measurements are available. This is all well and good, but it would be a waste of time if that is all we require; the point of the exercise is clearly to obtain computational results *outside* the empirical test-domain. To do this, we need to be able to extrapolate our 'justification' to these disparate regions of interest. But how can we then be certain of the 'correctness' of our initial-final computational relationship? Ideally, we need to view the combination of test-domain and disparate regions from a *higher level*, within which we can characterize their overall degree of correspondence. Clearly, we have now trapped ourselves in an apparently infinite series of higher and higher levels of reference, without any final justification at all being available.

This is the kind of impasse in which Rosen [4] found himself when trying to set up his (M,R)-system model of an organism. His initial attempts to make *efficient cause* [5] entirely internal to the organism led to an infinite series of creator, creator-creator, creator-creator-creator… entities. His solution was to point out that recursion could occur from a later creator to an earlier one, effectively truncating the infinite series by turning it back on itself (see [6] for an extensive discussion of this aspect). Can we solve our present difficulty in a similar way? It is instructive to invert Rosen's [4] initial argument. Rather than defining first that efficient cause *must* be internal, which then leads to his *closed* solution for an organism[2], we can point out that in an internally self-consistent closed system all of its static characteristics can be successfully evaluated (while retaining the *belief* and *Gödel* constraints we referred to earlier). As a consequence, in our computational quandary, we can only solve our justification difficulty with any degree of success if we are operating within a *closed* system,

---

[1] Neglecting here for simplicity the overhead involved in splitting a process into justifiable parallel paths.

[2] Note that Rosen [4] avoided the implication of an organism being an *at least* partially open system by ignoring the input of elements from the environment in his final model (see [6]).

where ultimately our local computational result is justifiably derived from the way it appears to be at higher levels of criticism.

It is of value to note two controlling aspects here. Firstly, if we are attempting to justify local computational success by progressively relating it to larger and larger domains, Einstein's [7] relativity comes to our aid in effectively closing off the system we are working with from the rest of the Universe – within a given timescale. Secondly, Rosen's (M,R)-system is effectively a model for Natural computation, and its principal *relational* characteristics are associated with *change*, and not stasis!

Our computational model now includes *justification* as a prime consideration, and this within an approximately closed *hierarchical* representation. As such, *local* justification is obtained from *global* characteristics,[3] and the entire hierarchy is in a constant state of flux between its different interrelated levels: Natural computation cannot provide temporally independent static solutions!

Much of what we will have to say addresses the manner in which we habitually describe, model or *point to* the elemental 'nuts and bolts' of our surroundings, and whether our conventional Scientific habits are sufficient, or even 'fit for purpose'. In this, our *birational* approach is closely related to the conventionally mono-rational discipline of *semiotics* [8-9] – or the study of signs – and sign processes (*semiosis*), such as analogy, metaphor, designation, signification, symbolism, indication and communication. In accordance with linguistics, semiotics exposes the character and behavior of Natural objects and phenomena in terms of *syntactics* – the relationships between signs themselves – *semantics* – the relationships between signs and their denotata, or 'what they refer to' – and *pragmatics* – the relationships between signs and the effects they have, or the contexts within which they are effective [10].

The central thrust of our argumentation will particularly address the semantic relationships between models and their denotata, and the pragmatic relationships between denotata, or *denoted entities*, and their contexts, or *denoted ecosystems* in the terminology we will adopt. By ecosystem we refer to the environment within an entity exists or thrives.[4] As such, the same environment may appear differently to different organisms, for example, as they will each be subject to different specific sets of signs, or influences, and not to others. Agamben [12] has given as an example of Jakob von Uexküll's *biosemiotics* [11] that a tick's parasitism of another organism is controlled by a set of particular environmental signals, ranging from "the odor of butyric acid" to its new host's "precise temperature", while these may be entirely irrelevant to other organisms.

General public attention to this pragmatic relationship between entity and ecosystem derives initially from developments starting in the 1960s – best illustrated by the publication of Rachel Carson's book "Silent Spring" [13] – although this was predated not only by quantum mechanics, but also by the appearance of second-order cybernetics in the mid- to late-1900s [14]. The aspect of entity-ecosystem characteristics which will be of most importance to us in this chapter will be the *complementarity* between entity

---

[3] This corresponds to a basic character of non-relativistic post-Newtonian classical mechanics, where local and global effects are presumed to coincide.

[4] Referred to by Jakob von Uexkull [11] as the entity's *umwelt*.

and ecosystem. This complementarity implies that entity and ecosystem cannot be simplistically separated, and that their *symbiosis* influences properties of both of them. Critically, for a complementary pair, 'the whole is not necessarily equal to the sum of the parts', although the common assumption [5] that 'the whole' is automatically '*more* than the sum of the parts' does not necessarily hold [15]. Deacon [16] rephrases this as 'the whole is *less* than the sum of the parts', but attention to Root-Bernstein and Dillons' [17] arguments suggests that any evolutionary process results in a combination of both of these – more of some, but less of other.

Our approach in this chapter will be to address the general character of *representation* of complex systems, and to reduce it to a form which is amenable to Natural computation. Our target is to bring all of Science under a generalized umbrella of *entity* and its *ecosystem*. This, then, will permit us to characterize different types of entity by their relationships with their relevant ecosystems, while retaining characteristics they hold in common. The most general way of doing this is to move the ecosystemic paradigm up to the level of its encompassing logic [18], thus creating a complementary pair of conceivably *different* coupled logics – one for the entity we are focusing on; one for the ecosystem within which it exists – and providing for their quasi-autonomous *birational* interaction.[5] An important comparison to this interpretation of Nature can be found in Brenner's [19] book "Logic in Reality", which puts forward the Principle of Dynamic Opposition, instantiating the necessary co-existence of a characteristic and its opposite or contradiction. This aspect of "Logic in Reality" supports the central focus of Deacon's [16] book "Incomplete Nature": that of *absence*:

> "*The concepts of information, function, purpose, meaning, intention, significance, consciousness, change and human value are intrinsically defined by their fundamental incompleteness*" [20].

We are of the opinion that Deacon's *absence* corresponds, at least approximately, to the *ecosystem* we refer to here.

A central aspect of our development here will be the concept of scale. Scale is most usually associated with linear dimension, or *size* – an analogue we will often adopt here. In accordance with Matsuno's [21] description of observation as a mutual measurement, we will refer to differences in scale in terms of differences in *perceptional capacity* [22]. Most particularly, we will relate differences in scale to differences between the various more-or-less complex models which may be used to represent an entity or a system, as we will illustrate later.

---

[5] Throughout this paper we will use the term *logic* for the set of static operational rules which apply to a specific entity in a particular context, and *rationality* for the context-dependent logically-determined path through which the entity progresses, either as a consequence of its initial state or quasi-autonomously towards a predetermined goal. This restriction to *quasi*-autonomy corresponds to a belief that system (re)organization is always driven by contextual influences, and that there are no instances of 'pure' *self*-organization. Otherwise, the birational framework we propose would have limited meaning, as 'pure' self-organization would weaken, if not completely destroy, entity-ecosystem complementarity.

Natural systems tend towards hierarchy. Multiple individual scales or levels are separated by complex regions whose character resembles a less-than-formal interpretation of '=' in '1+2=3' [18]. Inter-scale communication and scale isolation are both necessarily partial [23]. Decision-making processes are nominally data-destructive: global data judged to be unhelpful is progressively thrown away until a simple *localized* choice remains.[6] Analogously, a *denoted entity* hierarchy is reductive towards localization. Somewhat surprisingly, the set of intervening complex regions we referred to above make up a second, *denoted ecosystem* hierarchy, which is expansive – or 'reductive' in its own way – towards nonlocalization [24]. Each of these two hierarchies embodies one of the two rationalities we created earlier. At any extant scale there is a complementary pair of models of *denoted entity* and *denoted ecosystem* – one related to post-Newtonian classical representation, the other to quantum representation – both evaluated at that scale [25].

We believe that computation within this *birational* framework can be much closer to Nature than conventional *mono-rational* approaches, most particularly for complex and living systems. Conventional formal systems rely on a single homogeneous logic, and progression through a hierarchy corresponds to either top-down or bottom-up transit. Our own creative processes use *both* of these modes, usually invoked alternately, to guide our progress most efficiently towards suitable conclusion. We propose that Scientific or philosophical investigations should *always* be formulated *birationally*, in a way which is related to the Western interpretation of yin-yang [26] – as a *complementary pair* rather than an alternation of opposites.

2 Philosophical Relationships

A fascinating aspect of this birational approach is that representations and properties now *always* exist as intermediates between pairs of ideal extremes. Quantum logic, for example, no longer *replaces* post-Newtonian classical logic; it *complements* it [27], consequently identifying all *real* entities as compromises between the two. This philosophically non-traditional *included middle* is identical to that of the philosophical logic of Stéphane Lupasco [28], and to the implications of Brenner's "Logic in Reality" [19].

The fundamental postulate of "Logic in Reality" (LIR), its Principle of Dynamic Opposition, states that

1) every real complex process is accompanied, logically and functionally, by its opposite or contradiction (Principle of Dynamic Opposition), but only in the sense that when one element is (predominantly) present or actualized, the other is (predominantly) absent or potentialized, alternately and reciprocally, without either ever going to zero; and

2) the emergence of a new entity at a higher level of reality or complexity can take place at the point of equilibrium or maximum interaction between the two.

---

6 This is part of the reason why a digital computer needs access to memory.

A necessary concept is the categorial non-separability of, for example, individuality and non-individuality; part and whole; subjectivity and objectivity in relation to the experiment-experimenter pair.

The six Axioms of Life in Reality (LIR) are:

LIR1: (Physical) Non-Identity: There is no A at a given time that is identical to A at another time. This formulation is essentially that of Leibniz.

LIR2: Conditional Contradiction: A and non-A both exist at the same time, but only in the sense that when A is primarily actual, non-A is primarily potential, and vice versa, alternately and reciprocally.

LIR3: Included (Emergent) Middle: An included or additional third element or T-state emerges from the point of maximum contradiction at which A and non-A are equally actualized and potentialized, but at a higher level of reality or complexity, at which the contradiction is resolved .

LIR4: Logical Elements: The elements of the logic are all representations of real physical and non-physical entities, processes and systems none of which can be totally identical to another.

LIR5: Functional Association: Every real logical element e – objects, processes, events – always exists in association, structurally and functionally, with its anti-element or contradiction, non-e; in physics terms, they are conjugate variables. This Axiom applies to the classical pairs of dualities, e.g., identity and diversity.

LIR6: Asymptoticity: No process of actualization or potentialization of any element goes to 100% completeness

These six axioms form a whole which is very close to the proposition we will make here. LIR1, for example, corresponds the implications of Einstein's [7] relativity in an extensive system. LIR3 and LIR6 correspond to our observation above that quantum logic complements post-Newtonian classical logic, "*consequently identifying all real entities as compromises between the two*". The context-dependence of Life in Reality corresponds to the entity-ecosystemic relationship we have described above, unlike the situation in conventional logic systems. However, two differences in approach should be noted. Firstly, our representation relates to hierarchical systems consisting of numerous different clearly identifiable scales or levels. Secondly, in our proposition, we do not specify a functional association of every real logical element with its anti-element or contradiction (c.f. LIR4 above), but a functional association of every entity with its Natural ecosystem. In mono-rational non-hierarchical terms, however, this would reduce to LIR4. We would prefer to denote Lupasco and Brenners' included middle as the *exclusive middle*, to emphasize our proposition's equivalence to the implications of Brenner's LIR6. It is then understandable that the measurement of a particle's/quantum-wave-packet's properties with suitable equipment will indicate *mixed* properties [29] and not uniquely particulate or wave characteristics!

Science is often described as the child of Aristotelian pragmatism [5], but this leaves no functional place for the abstraction of *models*. In a birational description, Science automatically has two complementary aspects: the Aristotelian pragmatism of measurement and the Platonic abstraction of models.

The complementary ecosystemic paradigm relates an individual to its global environment. In computational terms, this can be correlated, for example, with Turing's advances in code-breaking [30], through the insistence on relating an individual coded message to the globality of its 'environment' of place, date, time, weather, … and proceeding by the elimination of consequently discovered contradictions.

In a birational framework *existence itself* 'becomes' a derivative of localization and nonlocalization, and the informational entropy associated with living systems becomes a compromise between that of two complementary kinds of entropy: one nominally inversely proportional to the conventional definition of *order*, and one nominally proportional to it.[7]

3 Outline

The remainder of this paper is organized as follows. We begin by establishing the context within which we will proceed ('4 Setting the Stage'). Our first focus will be on 'size-dependent' Natural properties ('5 Scale in Nature') and their organization and representation in large Natural systems ('6 Representation and Natural Hierarchy'). Next we will explore how these properties are interrelated ('7 Inter-scale Interfacing') and the nature of their classification ('8 Digital versus Analogue in Hierarchy'). This will lead us to considerations of system cohesion ('9 Quasi-stability'). We will draw comparisons between Natural- and digital- information processing ('10 Organisms and Computers') and address the lack of universal presence of hierarchy in Nature ('11 Hierarchy and Opportunism'). Our main focus will be on the establishment of a general *entity-ecosystem* framework ('12 Birational Complementarity') and we will end with a reflection on its implications ('13 Unavoidable Consequences and Conclusions').

Although this proposal may appear at first sight to be anti-Scientific, such is not the case: it simply addresses an extension of the current Scientific point of view and of its inclusivity. As such it is

> "*part of an emerging dynamic-informational paradigm in which all currently accepted wisdom is to be questioned. The concepts of the role of intermediates in complex systems, a principle of complementarity in a new systems paradigm and above all the need for a new, non-standard but encompassing logical framework for describing birational interactions are essential in order to, finally, break through a number of 'brick walls' in philosophy, logic and the sciences of life and mind*".[8]

This development holds out the exciting prospect of conceivably overcoming the impasse of complexity currently facing artificial intelligence and finally opening the way to developing *real* non-biological intelligent systems.

---

[7] Note that this dual entropy is not directly equivalent to that invoked by Landsberg [31-32], which is related to both *informational* entropy and *thermodynamic* entropy.

[8] Joseph Brenner: private communication.

4 Setting the Stage

For us to address the current framework for computation we must first describe it. Scientific measurement and consideration take place in and around particular models, which are usually constructed in terms of a limited set of *independent parameters*, and are usually derived from previous conceptual forms through intuition or inspiration. The validity of a model is checked by comparing its predictions with *multi-parametric models* in other domains, resulting in an always provisional conclusion of acceptability or of definitive rejection. At any typical point in the temporal evolution of Science there is an overall coherence between models in different investigative domains: they will all be more or less related to a consistent grounding – the current paradigm, whether this be purely Scientific, or humanist, or religious in character.

At somewhat irregular intervals there is an upheaval leading to change in the grounding paradigm, usually following a realization that information derived from the application of specific models to measurement data cannot be resolved within the current paradigm. After its three centuries of predominance, for example, the post-Newtonian classical paradigm began to be questioned around the turn of the 19th century following the discovery of what became known as the 'ultraviolet catastrophe' [33]. This led to the development of quantum mechanics in the first decades of the 20th century as a way of resolving the problem [34].

However, a description of Scientific structure in terms of just model and paradigm does not tell the whole story. Both of these are subject to the all-encompassing constraints of logic. As a joint epistemological/ontological hierarchy, not only do contemporary models exist within a current paradigm; a current paradigm itself exists within an over-arching logic structure. At this point we should emphasize that although human endeavor is most usually coupled to a single quasi-universal mono-logic, there is no reason why this should always be the case. Human language and discourse are so closely integrated with consequent mono-rationality that it is difficult to imagine any other way they could evolve. However, there is no obviously fundamental reason why a change in the grounding logic system should be philosophically excluded. Ecosystemics itself exhibits non-homogenous logic in the way that an individual species is related to the multiple other species of its environment. As John Kineman has pointed out,[9] if you remove the species of *bear* from an ecosystem, the remaining uninhabited niche or '*bear-hole*' is not exactly equivalent to the missing species, as the bears' absence immediately modifies relationships between all the other species present.

Our overall position is that now, at the beginning of the 21st century, a new 'paradigm shift' is necessary to resolve the apparent exclusion of *life* from biology and Science in general. Rosen [4] has pointed out that living systems survive and propagate through continuity in the relationships between their distinguishable parts, and not predominantly through the local characteristics of those parts themselves. The problem area is principally that of *reductionism*, whose over-application has led Rosen to describe biology as "*the study of the dead*".

---

[9] John Kineman: private communication.

Since the nineteen sixties there has been a successful revolution in approaches to living Nature [35], through the development of a view relating organisms to their entire relevant environments – the development of *ecosystemics*. Until now, however, this revolution has been confined to living Nature, leaving the 'hard' or 'precise' sciences of physics and chemistry mostly untouched. We believe that it is now time for a new paradigmatic revolution which will extend the ecosystemic approach to all of Science in a way that makes it possible to integrate life into Science in general, while leaving the 'exact' sciences *exact*! The most obvious way to do this is to move the ecosystemic paradigm up from the paradigmatic level to that of logic itself, thus moving a large part of the paradigmatic complementary inter-dependence up to the logic level and creating a complementary *ecosystemic* pair of logics to replace conventional mono-rationality [27]: see Figure 1(a) and Figure 1(b).

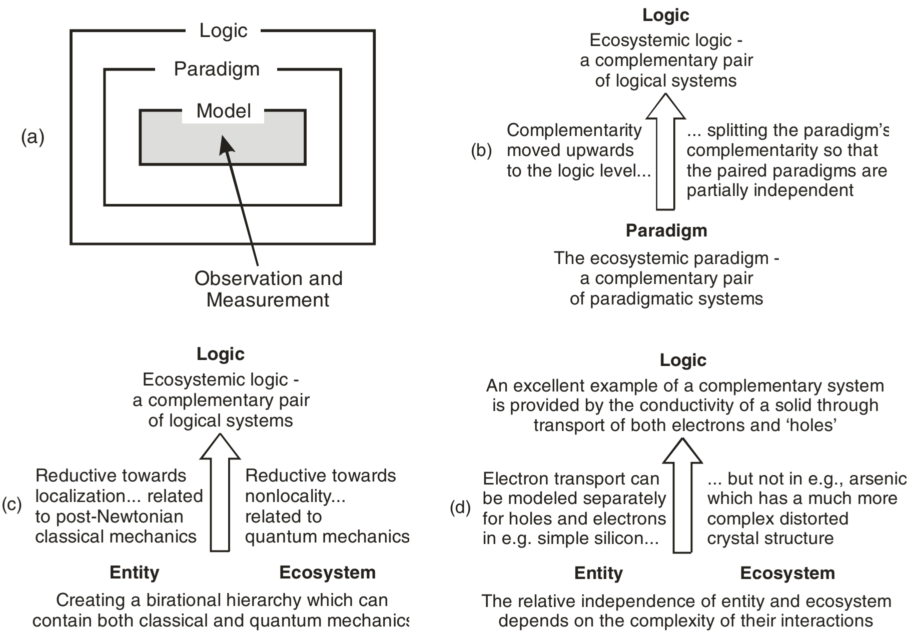


Fig. 1. Moving the ecosystemic paradigm up to the logic level: (a) the conventional placing of the ecosystemic paradigm; (b) the ecosystemic paradigm moved up to the logic level; (c) the resulting characters of the two remaining quasi-independent parts of the ecosystemic paradigm; (d) an example of coupling through complexity: the transport of electric current in a solid

Interestingly, this attributes very particular characteristics to the remaining parts of the original ecosystemic paradigm itself. Rather than remaining simply *organism* and *ecosystem*, these two now come to represent post-Newtonian-*classical* and *quantum* mechanical descriptions, respectively [18]: see Figure 1(c). This places these two ostensibly discrete and independent paradigms in a complementary framework of relationships! It is comparatively easy to see why this should be the case. Post-Newtonian classical mechanics is based on the primacy of localization: it is specified in terms of precision and accuracy of local measurement. Quantum mechanics, on the other hand, is principally 'nonlocal' and probabilistic in its derivation and operation.

An excellent example of complementarity between opposites, in a way which is related to Brenner's [19] LIR 'A and non-A' description, is the conduction of electricity in solids (Figure 1(d)). Silicon has a simple *cubic* (zinc blende) crystal structure, and its electrical conductivity is describable in two isotropic separate parts, one related to an entity – the electron – and one related to its environment – the 'hole' which is left where an electron is *absent*. Arsenic, however, has a more complex *trigonal* (distorted cubic) crystal structure, and the two components of electron and hole conductivity cannot be so easily separated [36]. This is reminiscent of the 'bear and bear-hole' we referred to earlier, as a 'hole' in arsenic is not isotropically equivalent to the absence of an electron.

Stepping a little to one side for a moment, we should explain what we mean by three of the important terms we will rely on. We will regularly refer to *digital* as opposed to *analogue* characters. While occasionally the reference will be to styles of computer or computation, the sense we will more often imply is a more general distinction between a differentiated (discretized) context and an undifferentiated (integrated) continuous one. The third, somewhat obscure term is *nonlocal*. Einstein's [7] relativistic arguments limit the speed of communication between any two real separated entities to the (measured) velocity of light – some 300 million meters per second. *Nonlocality* describes communication across space which transcends this limitation, which is nominally impossible in reality. However, Bennett et al. [37] have proposed that instantaneous nonlocal communication of this kind is in fact possible, but that the decoding of such an 'instantaneous' message *must* rely in addition on a second message passing between transmitter and receiver in 'real-time' (or 'real-speed') – thus eliminating the at-first-sight apparently 'faster-than-light' transmission of information.[10]

The clearest example of the apparent *difference* between post-Newtonian-*classical* and *quantum* mechanical descriptions and of their paradigmatic *coupling* is that of *light*. We can describe optical phenomena within both representational schemes: in the post-Newtonian classical paradigm, light is represented as *waves*;[11] in the quantum paradigm, it is represented as *particles*. These two representations lie at opposite ends of a spectrum of *size*:[12] 'waves' only exist as spatial extensions; 'point-like particles' lack spatial extension, being zero-dimensional, so do not take up any space [39]. A fundamental aspect of post-Newtonian classical mechanics is the unstated assumption that there is an instantaneous correlation between local and global properties. This apparently, and surprisingly, corresponds to the quantum mechanical characteristic of nonlocality, but it is a false nonlocality which takes no account of relativistic communication limitation.

---

[10] This scheme, referred to by Bennett et al. [37] as 'teleportation' after the TV series *Star Trek*, does in fact work, and is in use for the uninterceptable transmission of coded information (e.g. Shields and Yuan 2007).

[11] Note that Newton proposed a corpuscular, or particulate theory of light, but that this was superseded by the wave description principally due to Huygens [38].

[12] Note that for simplicity, we are describing the situation here only in terms of *spatial* extension: the word *size* we use can also refer to functional complication, or 'extension'.

Similarly, the quantum mechanical description of light as localized photons, or particles, is a false description, as the probability wave of any 'localized particle' automatically spreads over the entire Universe! Post-Newtonian classical mechanics permits the existence of single-frequency optical waves, but any single frequency wave must also spread over the entire Universe, because a phase change at any location implies a simultaneous analogous phase change at every other location. This injects nonlocality into post-Newtonian classical mechanics! In a related manner, although photonic probability waves touch everywhere simultaneously, a photon's position in 'space-time' can be determined through attention to its constraints, injecting localization into quantum mechanics. We maintain that these two paradigms of post-Newtonian-*classical* and *quantum* mechanics are opposite (complementary) 'faces of the same coin', and that they always coexist in any descriptive exercise.

This, then, is the first stage of our argument. Our intention is to replace conventional mono-rationality by a complementarily-coupled birational framework, derived from the ecosystemic paradigm, within which the conventional disciplines of Science – physics; chemistry; biology – may be successfully embedded without degrading their present performances. Conventional modeling relies heavily on reductionism in attempting to describe Natural phenomena, with the result that it cannot effectively deal with living organisms whose character is principally determined by their subsystemic interrelationships. The primary aspect of our derived ecosystemic birational framework is the interrelationship between its constituent paradigms [40]. This makes it an ideal support for representations of life itself, as an integral part of Nature rather than an uncommon emergence from it, to which end we reject any initial assumption of fundamental difference between organic and inorganic, between organisms and crystals, or between living and non-living. Consequently, a basic premise of our approach is that, in common with Nature itself, any distinction between living and non-living must automatically drop out of our subsequent modeling rather than being imposed from the start.

5 Scale in Nature

In general, Natural systems exhibit different scales. By scale, we mean 'size'-related differences in properties. A distinction is frequently made between *scale* – in terms of spatially structural aspects – and *level* – in terms of functional complication or complexity. However, as will become clear, this distinction is irrelevant for the interpretation of Nature which we will present, and we will consequently use the words interchangeably where this is convenient.

It is notable that although organisms often present a number of distinguishably different scales, inorganic entities exhibit few, and the differences between them, although observable, are less marked [41]. The number of scales exhibited by an organism ranges from a minimum for primitive forms, for example *two* for some slime molds/fungi[13] [43], up to an apparently limiting upper number for mammals.

[13] Note that we attribute *two* scales to an entity with a single level of internal structure and an overall outward appearance [42].

Dinosaurs and mice exhibit similar extant scales, whose number depends on the framework within which it is established: for example *ten* (elementary particle, atom, molecule, molecule group, organelle, cell, tissue, organ, organ system, organism) from a simplified 'Physics' point of view [44], *five* (cell, tissue, organ, organ system, organism) in a simple biological framework [45], or *four* (hadron, atom, cell, memon) in Jagers op Akkerhuis's [46] *operator hierarchy*.

In his booklet "What is Life?: the Physical Aspect of the Living Cell" Schrödinger [47] attempted to couple the characteristics of organisms and crystals by introducing the idea of an 'aperiodic crystal' that contained genetic information in its configuration of covalent chemical bonds. This stimulated research which culminated in discovery of the character of DNA [48]. A closer connection between organisms and crystals can be found in the structure of some fundamental biochemicals, for example the lipid *pdmpg*, which exhibits an almost crystalline regularity in its structure (see Figure 2). In any case, the major defining characteristic of any scale[14] system is the nature and quantity of information which is transferred between its different scales. In an organism the informational differences between scales can be enormous; in a crystal they are minimal [41].

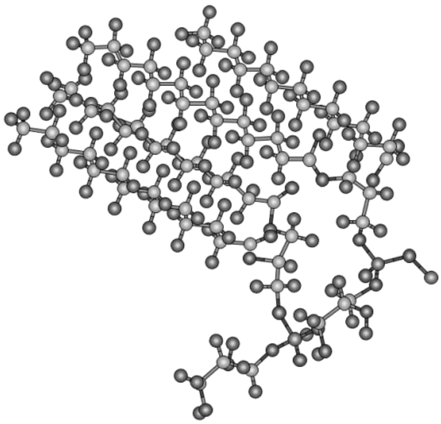


Fig. 2. The almost crystalline structure of the lipid *pdmpg*

A vital aspect of any multi-scaled system is the difficulty of establishing a conventional externalized 3rd person point of view or description. If there were *no* difference in properties between the different scales, then

1. it would be very easy to formulate an accurate 3rd person description, because
2. in that case there would not be any different scales *at all*,

which is nearly the case for a crystal, where macro forms follow on almost directly from microscopic atomic arrangements. Such is far from the case for organisms,

[14] Note that we will use the word *scale* intentionally as an adjective in place of the more grammatically correct *scalar*, to avoid confusion between references to the *scale* phenomena we are describing and the mathematical entity of a *scalar* (as a noun).

where cross-scale informational differences can be extreme. Consequently, a different 1st person description will be associated (at least) with each scale, and it will be impossible to generate an *accurate* 3rd person 'summation' of these from an external viewpoint.[15] If for no other reason, this makes it impossible to accurately 'understand' what is going on in another person's *mind* at any particular moment [40]!

6 Representation and Natural Hierarchy

We concur with Lupasco [28] and Brenner [19] that the very nature of reality is a concurrent operation of actual and potential, and that

> "*when one element is (predominantly) present or actualized, the other is (predominantly) absent or potentialized, alternately and reciprocally, without either ever going to zero*" (see the Principle of Dynamic Opposition above).

This logic is independent of the means of representation employed, and consequently Brenner and Lupascos' interpretation of reality sits at the top of the representational tree-structure shown in Figure 3. Following on from this starting point, we may address Nature in a variety of ways, with a variety of different possible formulations or representations. However, having introduced the presence of an organism's different organizational scales, we must now reflect on how they are coupled together to present the unified appearance we perceive as a 3rd person interpretation from outside[16]. Natural systems tend towards *hierarchy* [50]. This character of 'hierarchy' is closely related to the correspondingly named structure which is applied to business enterprises, where different levels of organization – such as chief executive (CE), top-level managers, middle management, foremen and workers – are coupled together to provide a hopefully efficient coordinated activity.

In a Natural context the overall organization is much more coherent and self-consistent than in a business – from both top-down and bottom-up perspectives – and there is no single predominant level which is comparable to a business's CE. Consequently, we will always portray a hierarchy 'on its side', with the 'top' (i.e. the equivalent of CE) at the right hand side. Salthe [44], [50-51] has extensively described the properties of hierarchy in Nature. However, he restricts the term *hierarchy* to two forms: the scale (or *compositional*) hierarchy and the specification (or *subsumption*)

---

[15] Note the word 'accurate' here. A 3rd person view is always possible... what is at issue is its justifiable closeness to what is apparently being viewed. Science deals with this by cross-referencing between different domains, to establish what is a 'good' model and what is not.

[16] Ultimately, such a reflection results in the expression of a concept referred to as *hyperscale*, but for the sake of brevity we will not address this aspect here, as it is not directly relevant to our argument. An extensive description of hyperscale and its implications may be found in a number of previous publications [25], [27], [40], [49].

hierarchy. However, we find that a third form – the *representation* or *model* hierarchy – is most suitable for describing the properties of Natural systems.[17]

It is our opinion that a *representation* or *model* hierarchy is the parent of Salthe's two hierarchical children, and we believe that these are two reduced formalizations of the representation hierarchy itself, as indicated in Figure 3. As such, there may even be other as-yet unnoticed reduced hierarchical formulations, and even intermediate forms between them, as sketched out in Figure 3. The *representation* or *model* hierarchy provides the link between Brenner-Lupasco representation 'at the top of the tree' and Salthe's compositional and subsumption hierarchies lower down.

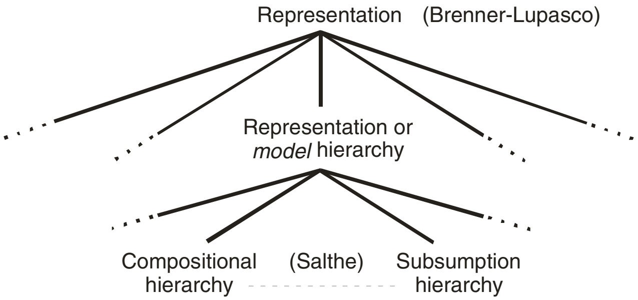


Fig. 3. The tree-structured relationship between this work, Brenner (2008) and Lupasco's (Brenner 2010) philosophical logics, and Salthe's approach to hierarchy. The model hierarchy provides a necessary link in the representational chain, as it exhibits both the Dynamic Opposition of Brenner's and Lupasco's high-level logical representation and the lower-level hierarchical aspects of Salthe's interpretations.

So, what exactly is a representation or model hierarchy? Figure 4 provides an easy-to-grasp example, in the form of a set of models of an electronic diode. The models show a progression from left to right from the simplest representation (a), corresponding to a binary digital model, to the most detailed representation (f), corresponding to the digital approximation of an analogue set of empirical measurements.[18] In common with many series of progressively more and more 'accurate' representations of a complex system, the intermediate series here shows one change – between (c) and (d) – where there is a step function in the style of modeling – from piecewise-linear to exponential. This hierarchical series appears to be a combination of Salthe's [44] compositional and subsumptive forms. The set of individuals (a), (b), (c) and the set of individuals (d), (e), (f) both appear to be compositional series (added elements are 'a part of' the whole). The two types of set, of piecewise-linear representation

[17] For the moment we will refer to this concept as a *representation* or *model* hierarchy, as convenient. Ultimately, we will be drawn to address it as a fundamental character of Nature – a *Natural* hierarchy. All these three names are valid.

[18] We have adopted here the most intuitive way to present this series, starting from the simplest (a) on the left and moving towards the most complex (f) on the right. It is important to note that in all of our illustrations of hierarchy the *highest level* (i.e. the 'simplest' representation) is on the *right hand side*, not as shown here on the left.

(a)-to-(c) and exponential representation (d)-to-(f), appear to be subsumptive (each set is representatively 'a kind of' the whole).

A further example of a (Natural) model hierarchy is that of a tree (see Figure 5(a)), where the differently scaled models are {a tree described in terms of atoms}, {a tree described in terms of molecules}, {a tree described in terms of cells}… up to {a tree described in terms of branches}, {a tree as *itself* – a tree}. Our generalized portrayal of a (Natural) model hierarchy is presented in Figure 5(b) for this same example of *a tree*, where each vertical line represents a differently scaled multi-parametric model in terms of the quantity of information needed to describe it (which would, of course, be far more extensive for the population of workers than for the CE; and more extensive for 'a tree as atoms' than for 'the tree as itself').

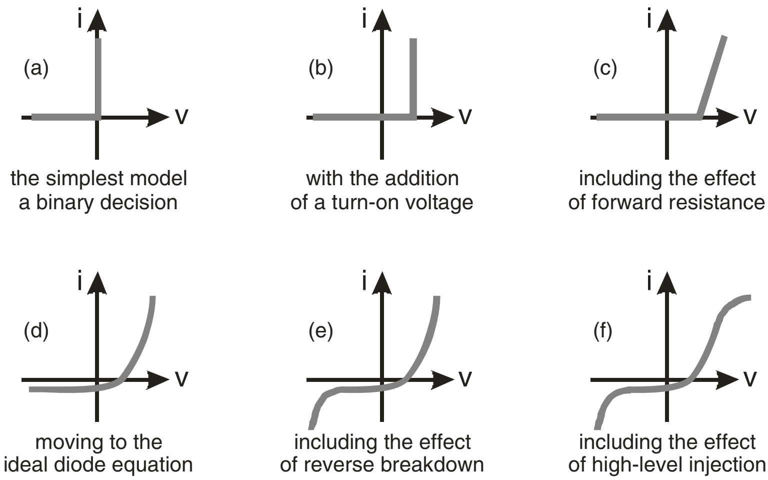


Fig. 4. The sequence of electronic diode models making up an example of a model hierarchy. The physical origin of the differently-modeled effects is of no particular consequence here, and the individual models (a) to (f) are referred to in the accompanying text.

It can be seen from Figure 5(a) that this series is both compositional (atoms-molecules-cells …) and subsumptive (a tree 'as itself' subsumes 'a tree as branches' and 'a tree as molecules'…).

It will be evident from Figure 5 and from the description we have given that the character, or at least the quantity of information at each organizational scale, will be different. This is a major problem: how can adjacent but dissimilar multi-parametric scales be self-consistently correlated? Surprisingly, we encounter the same difficulty even with a simple arithmetic equation such as 1+2 = 3, where we lose information on going from left to right, as we drop one of the 2 digits we started with (i.e. from '1+2' to '3'), and we consequently lose its degrees of freedom.[19] If we try subsequently to

[19] The number of degrees of freedom of a system or of one of its elements can be defined as the minimum number of coordinates which are required to specify its state. If we describe the state of a non-rotating element in three-dimensional space, for example, it will possess three translational degrees of freedom.

return towards the left hand side from the single ‘3’, it will become obvious that we have no way at all of knowing whether we started initially on the left with 3, or 1+2, or 2+1, or even 1+1+1 !

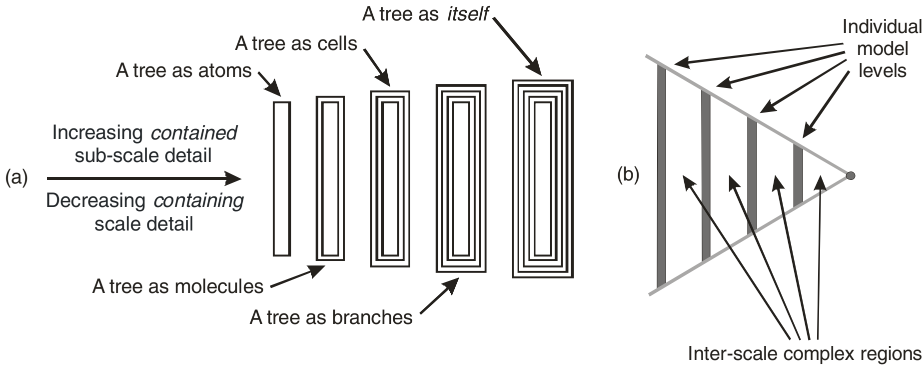


Fig. 5. (a) A tree, as another simple example of a multi-scale model hierarchy; (b) The general representation of a (Natural) model hierarchy for this same example of *a tree*. Each vertical line represents a differently scaled model in terms of the quantity of information needed to describe it without taking into account the subsumed information of the lower scales – which means that the most complex description at the right hand side of (a) is here the least extensive. The inter-scale regions indicated are multiply fractal and complex. The net result is close to being a ‘subsumptive’ hierarchy which is formulated in terms of scale![20]

In general, therefore, if we ‘travel’ from a lower extensively detailed multi-parametric scale to a higher one whose description requires fewer parameters, we must remember that it will no longer be possible *from there* to accurately determine the characteristics of the lower scale. This corresponds to the general rule that it is not possible in a model to simultaneously maximize accuracy, precision and generality [52]. As a consequence, individual scales will necessarily be partially isolated from each other. However, they will still be partially inter-communicating and directly or indirectly interacting, so that *approximately* complete and *apparently* self-consistent correlation can be carried out between them. This is a fundamental property of a Natural hierarchy: individual scales are partially isolated from each other, and only partially communicating with each other.

7 Inter-scale Interfacing

The regions between adjacent scales in both Figures 5(a) and 5(b) appear to be virtually impossible to successfully model, even by some kind of extremely simplified approximation.[21] It may be that future developments in Natural computation will

---

[20] Stan Salthe: private communication.

[21] These inter-scale regions are archetypically chaotic from a post-Newtonian classical viewpoint, and the accurate evaluation of any model would therefore require infinite computational precision to avoid consequences such as the ‘butterfly effect’.

resolve this issue, but it may also be the case that Nature itself cannot successfully compute these regions on a purely local basis. Gutowitz and Langton [53], for example, have suggested that the phenomenon of *critical slowing down* at phase changes could conceivably be linked to

> "*fundamental limits on physical systems' abilities to effectively 'compute' their own dynamics*".

The '=' we are used to encounter in expressions such as '1+2 = 3' is logically defined in its arithmetic context as a statement that the left and right sides of the 'equation' are indistinguishable. This is clearly not the case: the character string '1+2' is easily distinguishable from the string '3'. This is one of the drawbacks of conventional mathematics in its application to real-world problems: even though extensively useful, it is far too short-sighted in its formality to comprehensively represent even the simplest of circumstances. If we try blindly to rely on '1+2 = 3' to represent a left-to-right summation of apples, we can end up on the right with *one very big apple*! Even so, will this super-apple be equivalent to '1+2' apples in terms of width, or of weight, or of color, or of taste? We have no way of knowing. From a wider perspective which takes account of unspecified properties, '1+2 = 3' is primarily a *hierarchical* relationship: the two sides of the 'equation' characterize different *scales*. The only way we can rely on this kind of equation is to provisionally close our eyes to reality and trust the abstract formal nature of mathematical definitions. Then, in some *but not all* contexts, we will be successful: 1 apple + 2 apples can happily result in the 3 apples we would like, for example – if we disregard disturbing properties such as degrees of freedom.

The coordinating relationship between any pair of adjacent scales of a real Natural entity, therefore, will be very different from the abstract equivalence of '='. Communication between the two must be bi-directional, but this bi-directionality will be *unavoidably* asymmetrical. It must take account of differences in properties between the two scales; most particularly those which apparently disappear or emerge during passage through the intervening region. Above all, it must be *aware* of the entity's more global context, which it must be able to adapt to. The resulting overall relationships which characterize the inter-scale regions, therefore, must be *appropriately* applicable in *any* conceivable local circumstances. This makes these regions complex in the manner described by Rosen [4]:

> "*A system is simple if all of its models are simulable. A system that is not simple, and that accordingly must have a nonsimulable model, is complex.*"

Consequently, the only way to *completely* represent one of the complex inter-scale regions we refer to in all of its temporal detail would be by employing an infinite number of formalizations [4], which seems somewhat impractical! Sometimes a model's simple computational rules can produce amazingly rich complex behavior, but if the inter-scale regions could indeed be simplistically and logically modeled then adjacent scales would collapse into each other, and there would be no hierarchy at all [25].

It is also impossible to manageably digitize these inter-scale regions, as completely accurate conversion would require an infinite number of digital bits. This makes the inter-scale regions archetypically analogue[22] in character [54]. In reality, the situation is even far less obvious than it at first appears. If we examine either Figure 5(a) or 5(b) of our example of a tree, the extreme right hand side of the illustration corresponds to a digitized representation of *whether the tree exists or not*: the tree's 'reality' is modeled as *true* or *false*, equivalently to the '*1*' or '*0*' of conventional computing notation. If we now move progressively towards the left of Figure 5(a) or 5(b), we pass through a scaled sequence of multi-parametric representations, which through {a tree described in terms of branches}, {a tree described in terms of cells}, {a tree described in terms of molecules}, {a tree described in terms of atoms} approaches more and more closely a detailed analogue description of the tree. So, the extreme right hand side of Figures 5(a) and 5(b) corresponds to the simplest *digital* representation of the tree, by indicating its reality; the extreme left hand side corresponds to a *digital* simulation of an *analogue* representation of the tree (the extreme left hand side of the sequence is equivalent to a digital representation in terms of an infinite number of bits, and it cannot consequently be practically distinguished from an analogue representation: the distinction between 'analogue space-time' and 'quantized space-time' is similarly problematic[23]). This illustrates the principal character of a multi-scale Natural hierarchy: it provides an *interface* between *analogue* and *digital* representations [58] – between existential integration and differentiation! We submit that Natural hierarchical interfacing of this kind constitutes *the* defining feature of our Universe. There are far-reaching consequences of this proposed conceptualization – most particularly with respect to the historical evolution of the Universe.

8 Digital versus Analogue in Hierarchy

Conventional interpretation depicts the origin of the Universe as a homogeneous, isotropic, undifferentiated state of incredibly high energy density, which precedes space and time [59].[24] Following the Big Bang, later stages of Universal expansion develop differentiation into a multiplicity of more ordered and consequently less communicating entities, whose configuration in the absence of life progressively decays through time towards the uniform physical entropy of 'heat death' [60]. In its most complete form, Natural hierarchy delivers the *temporal* interface between the

---

[22] Note that to avoid confusion we will consistently use the spelling 'analogue' here, implying, as we logically should, that the *analog* referred to in digital information processing systems and the *analogue* referred to in inter-dimensional comparison are directly related, if not precisely the same.

[23] We have here avoided any further discussion of the potential analogue or quantized nature of space-time itself. We do not believe, maybe surprisingly at this point, that such a debate would radically alter the basic ideas we present here. Reference to Dodig-Crnkovic [55], Kurzweil [56] and Lloyd [57] and further reading of our argument will support this belief.

[24] There are now a number of far more esoteric versions of pre-Big Bang conditions, but until now they all remain as conjecture.

infinite analogue energy of the pre-Big Bang condition and the digital character of evolved multiple differentiations. This characterizes Natural hierarchy as the major descriptor of continuum-to-discrete, integration-to-differentiation and analogue-to-digital processes in the Universe. To be consistent, therefore, we should not only always resort to our generalized representation to clarify the properties of any multi-scaled entity or phenomenon; we should also 'insert' it into *any* context where analogue and digitized structures or representations are physically or functionally adjacent.

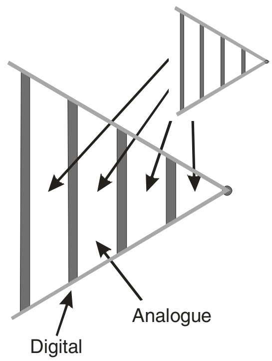


Fig. 6. The consequences of inserting our entire hierarchy between any digital-analogue adjacency

Figure 6 indicates the consequences of inserting our entire hierarchy between any digital-analogue adjacency in an extended rendering of Figure 5(b). Between each pair of multi-parametric (and therefore digitized) model scales lie complex (analogue) regions: we must insert *at least* a version of our entire hierarchy between the scales at every stage, as shown in Figure 6. But this will not be enough. We have now, in each case, implanted between *analogue* and *digital* a set of digitized sub-scales, and are faced with the frightening prospect that we should *again* insert *at least* a version of the entire hierarchy *between these*… and so on, *ad infinitum*! We end up with an infinite sequence of scales, sub-scales, sub-sub scales, sub-sub-sub-scales,… whose repetition towards smaller and smaller size resembles the *self-similarity* of fractals [61]. This progressive finer and finer accumulation of 'detail repeating itself' not only makes the inter-scale regions infinitely (digitally) fractal, but also makes them digital reproductions, or *digital facsimiles*, of analogue complexity. It unavoidably, and somewhat confusingly, makes the entire Natural hierarchy we are describing a primarily *digital facsimile* of what initially appeared to be a purely analogue view of our surroundings!

9 Quasi-Stability

The picture we now have of the constitution of a Naturally-hierarchical entity or system is one of multiple constituent elements which are distributed throughout multiple dissimilar scales of organization. A degree of cohesion between these different elements is necessary to prevent the entity or system fragmenting into a disorganized collection of lower level components [62], and this cohesion must be communicated throughout the entity or system, at the very least at its lowest organizational level. For an organism, for example, there must be enough communicated cohesion at the level of the individual cells. If the organism expands beyond a particular size the local communicational overhead required to maintain cohesion will become excessive, and the only alternative to fragmentation is to reorganize in a way that reduces this local overhead [40]. The generation of a new higher scale can achieve this by replacing a multiplicity of local communications by a reduced number of higher-level long range ones. More correctly, any reconfiguration to prioritize long range communication will only take place *if* it will result in lower overall communicational overhead. This is a Natural analogue of the technique used in large scale computer chip manufacture of grouping communications between different regions of the chip to increase speed of operation and save energy [63].

If our represented multi-scale Natural entity or system is to be at least provisionally stable it must be able to *inter-correlate* its different *partially-isolated* scales effectively – most particularly and most consistently at its lowest levels of organization. The solution to this conundrum, however, is comparatively simple. The greater the stability of the entire entity or system, the greater the proportion of information crossing between scales which is quasi-static and structure-supporting in character. This is very clear in the case of the long-term stability of an inorganic crystal – for example, gallium arsenide [41] – where nearly all of the inter-scale transfer of information is structural, and information content is virtually identical across scales. It must, however, *also* be the case for an organism, as a means of maintaining its quasi-stability. We can conceptually split any inter-scale information transfer into two functionally-different parts. One of these is the quasi-static structure-supporting information we have already referred to; the other is any remaining non-quasi-static information, which although being unexceptional from the 'point of view' of the 'transmitting' scale, will be *novel* to the 'receiving' scale. Although neither of these two parts can be rigidly defined, we find ourselves once more in a domain where Aristotelian pragmatism lies between Platonic dimensional extremes: the biological reality of an organism's quasi-stability lies in the middle ground between the two extremes of stability-supporting static information and stability-weakening novel information.

10 Organisms and Computers

In his book "Life Itself: a Comprehensive Inquiry into the Nature, Origin, and Fabrication of Life", Rosen [4] has extensively addressed the criteria for existence of a living organism in terms of its indispensable internalization of Aristotle's [5] *efficient cause*, or 'means of creation or construction'. His ultimate depiction is of an integrated graph of feedback or 'looping-back' processes, each representing one of an organism's internal functions of Metabolism, Repair and reproduction – his (M,R)-system. Although open to a degree of criticism [6] his graph confirms, unsurprisingly but somewhat realistically, that an organism is indeed only *quasi*-stable, corresponding to our conclusions in terms of cross-scale information transport. Rosen [4] has suggested in reference to his book's figure [10C.6] that:

> "*Any material system possessing such a graph as a relational model (i.e. which realizes that graph) is accordingly an organism. From our present perspective, we can see that [10C.6] is not the only graph that satisfies our conditions regarding entailments; there are many others. A material realization of any of them would likewise, to that extent, constitute an organism.*"

The Natural model hierarchy we are describing is capable of reproducing *at least* the relational functionality of Rosen's figure [10C.6], but it also enables us to progress much farther than Rosen's (M,R)-system implies [6], [64]. From the character of inorganic crystals we have referred to, and the conceptual split we have proposed between stability-supporting static information and stability-weakening novel information, it is clear that the former *quasi-static* cross-scale informational relationships specify nothing other than a non-living system, while the latter *novel* or closely scale-related informational relationships are what make an organism *alive*!

The last few decades have seen a wide-ranging growth of research into *artificial intelligence*, *artificial life* and *artificial consciousness* – all three of these addressing the essential nature of living organisms through the medium of digital computation. The historical move from analogue to digital computation was driven by a wish for the generality of von Neumann computers and for computational precision as a surrogate for accuracy. Unlike analogue circuitry, digital circuits permit wide variation in component tolerances without degrading overall performance. There is, however, a current resurgence of interest in analogue computation, particularly in the context of morphological computation and robotics. It is important to note that although digital and analogue computers each realize enormous benefits in some areas of information processing, they both suffer from critical disadvantages when compared to an organism.

While the evolution of digital systems has followed a path which is conceptually parallel to that of Natural organisms [65], they implement radically different logical structures, and this makes for a fundamental dissimilarity in the ways they can operate and in what they can achieve. Conventional digital computers consist of a (very) large array of simple Boolean logic gates. These are connected together in a complicated manner to provide the extensive abstract functionality to which we are now accustomed. However, each gate is a physically *real* entity, and takes a certain period of time to provide an output that correctly corresponds to its current inputs. To avoid consequent logical errors, a repetitive clock signal tells all of the gates to wait for a predefined 'settling time' before passing on their outputs to subsequent gate inputs. It is most instructive to describe the computer clock signal with respect to an altogether different characteristic: it completely isolates *each gate* from *all of the others*, except in the ways that the computer was conceived or programmed. Consequently, a digital computer is *only* capable of exhibiting *local* control or phenomena: it is incapable of generating or accessing *any* larger-scale or global properties. This, then, makes it impossible to generate in a digital computer any phenomenon which depends on globally-coupled properties, such as *intelligence*, *life* or *consciousness* [18]. A further disadvantage of Boolean implementation is that every gate in a digital computer is at the same non-hierarchical level [25]. This means that the bigger the computer's gate network, the slower the computer will perform (thus the race during the last few decades to increase computer clock-signal frequency or speed).

A Naturally-hierarchical multi-scale organism functions in a completely different manner from a digital computer. The partial isolation of its scales means that most of its information processing is being carried out at its lowest scales, and the highest, most abstract scales can operate in *real time* quasi-independently of the lower ones. This is an enormous advantage to an organism, as it means that its large complicated and complex system can react comparatively rapidly to threatening external stimuli [18]. This, then, is an area where analogue computation has the advantage over digital, because its operating speed is unrestricted by the imposition of a delaying clock signal. Computation routinely takes the form of a reconfiguration of analogue empirical data in support of our analogue style of comprehension. Analogue computers do not need to encode and decode data into binary format, and they can be dramatically faster than digital ones at computing continuous functions. However, analogue computers are subject to numerous electrical limitations, such as component inaccuracies and instabilities, the noise floor of their signals, the finite nature of an electron's charge, microelectronic parasitic effects, temperature issues and non-linearities.

Computers of either kind can, of course, be used to control complex systems (e.g. a nuclear power station) at both their local and global levels, but the conceptual distinction is in the mind of the control-system designer, not inherently in the computer's make-up. For example, neural network software can indeed carry out a distributed form of processing, but this is a combination of many *local* processes and it has *no* non-local character. Although it may be convenient to describe a digital computer as a number of different operating levels, an individual processing gate neither knows nor cares if it is processing a part of the operating system, the user interface, an application …

Chalmers [66] claims that continuous systems would need to exploit infinite precision to exceed the powers of discrete systems. However, the reverse is also the case: discrete systems would need to exploit an infinite number of computational bits to exceed the feasible precision of continuous systems! Natural hierarchy implements the advantages of *both* digital and analogue processing, as does the brain. The most obvious differences between an organism and a computer are an organism's self- properties,[25] like self-organization, self-repair, self-reproduction and self-control, which are totally absent from conventional computers. However, it is the Natural hierarchical character of an organism which is responsible for the occurrence of all of these self- capabilities.

11 Hierarchy and Opportunism

Following on from our submission that Natural systems tend towards hierarchy it might be expected that we would find hierarchical organization literally everywhere.

[25] Note, once more, that nominally 'self-' properties are *always* contextually or environmentally mediated.

But such is not the case, and we must explain why. It should be noted that we have only submitted that Natural systems *tend* towards hierarchy – not that they necessarily achieve it. To see why this is so we must look at the implications of *Evolution*.

Darwin's [67] 19th century depiction of the Evolution of organisms specified the three essential features of *variation*, *reproduction* and *selection*. Conventionally, his *variation* is associated with DNA mutation, his *reproduction* corresponds to our usual description of reproduction, and his *selection* is carried out by the survival or death of members of a species through environmental influences, cooperation or competition. We consider that these three essential features are the result of the evolution of Evolution *itself*, from a simple integral form more reminiscent of that normally associated with simple chemical interactions to a later 'crystallization' corresponding to Darwin's differentiated description. This then suggests a continuity of evolution from the Big Bang up to the present day, during which every development of our Natural environment has been subject to prior influence.

Evolution is opportunistic, if nothing else. Nature has not been created 'in one go': it has evolved from state, to state, to state on the basis of what went before. This means that the characteristics which we observe today do not necessarily correspond to a 'rational' construction. This is very clear from the current physiognomies of organisms. Referring to the progressive evolution of physical characteristics, Sigmund [68] has pointed out that:

> "*What serves for thermoregulation is re-adapted for gliding; what was part of the jaw becomes a sound receiver; guts are used as lungs and fins turn into shovels. Whatever happens to be at hand is made use of.*"

This is the real character of Evolution. We ourselves, for example, have a backbone which originally evolved to support our hanging internal organs while we moved around on four limbs. Fortunately, not all the consequences of Evolution are disadvantageous. Much of evolution can be associated with the distribution or exchange of *autonomies* between members of a species, internal scales of an organism, or even whole parts of an organism. Collier [62] has provided a beautiful example of this kind of relationship, in his suggestion that our brains have gained informational autonomy by ceding supportive biological autonomy to our bodies.

In an opportunistic Evolutionary environment, we cannot expect that current Natural states or organisms will always correspond to the tendency towards hierarchy we have suggested. Even so, much of our surroundings *does* correspond to a hierarchical model. For example, there is evidence that parts of our brains operate hierarchically [69], while other parts do not [70].

12 Birational Complementarity

We now come to the central issue we wish to address. The study of ecosystemics leads us to the assertion that *birational* concepts are more general than conventional *monorational* approaches in their applicability to the disciplines of physics, chemistry and biology. We consider that Science should be reformulated to take account of what Nils

Bohr believed to be an all-pervasive complementarity [71]. We do not imagine, however, that a *binary* ecosystemic complementarity is the ultimate scenario towards which we are heading: Nature operates through multiple concurrent complementarities, and in moving towards a binary representation we are simply taking the first step towards a hoped-for future intellectual multi-complementarity more closely related to the informational integration of consciousness [18], [72-73]. It could be argued that more complicated complementarities could be looked at pairwise – as a multiplying-up of Brenner's LIR 'A' and 'non-A' – but we believe that in a realistic setting the interactions between different elemental pairs would *themselves* interact, producing a more complex arrangement. This is in any case an important feature of Natural hierarchy, as indicated in Figure 7(a), where bottom-up ('emergent' [74]) and top-down ('slaving' [75]) inter-scale processes interact with each other [58]. It also indicates a weakness in the traditional circular representation of cybernetic feedback 'loops', as shown in Figure 7(b), which fails to take into account this vital characteristic.

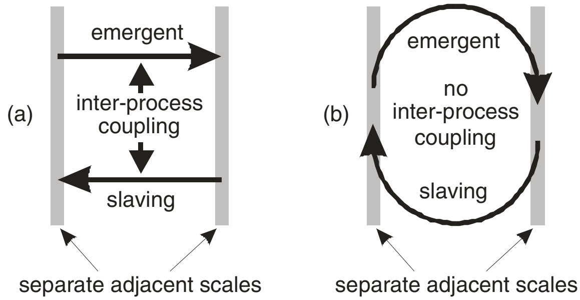


Fig. 7. (a) Parallel inter-scale process coupling, and (b) representational circularity

We have earlier shown [6] that Rosen's (M,R)-systems can be represented as a complementary pair of extremes (Figure 8(a)), and that the consequence is a mid-region, found only in organisms, where the four processes of {software flow}, {hardware flow}, {induction of software flow} and {induction of hardware flow} coexist (Figure 8(b)). One of these, {induction of hardware flow}, corresponds to the teleodynamic organization referred to by Deacon [16], where

> "*one physical system is capable of influencing other physical systems via something that is merely virtual*" [20].

Deacon [16] describes the emergence of morphodynamics (often referred to as 'self-organization') from homeodynamics (thermodynamics), and that teleodynamics (life, evolution, semiosis…) can emerge from the interactions of morphodynamic processes. This mirrors Rosen's [4] (M,R)-systems, where the (teleodynamic) organism 'emerges' from the interaction of (morphodynamic) metabolism, (morphodynamic) repair and (morphodynamic) reproduction. The means by which Rosen finally 'encloses' efficient cause is the identification of the individual parts of *B* in Figure 8(a) as processors themselves – corresponding to one of Deacon's [16] 'intrinsic constraints'.

Figure 5(b) portrays the general representation of a Natural model hierarchy, with discrete (digitized) scale models separated by (analogue) complex regions. Figure 6 does not in any way contradict this portrayal; it merely catalogues more clearly the complexity of the inter-scale regions. However, we now note an entirely unexpected feature of Natural hierarchy. Close attention to the comprehensive assembly of relationships between the individual scales, the individual complex regions and the unified nature of the denoted entity indicates that the set of inter-scale complex regions forms a previously unnoticed *second* hierarchy, whose different 'scales' are interleaved with those of the initial one (Figure 9) [18].

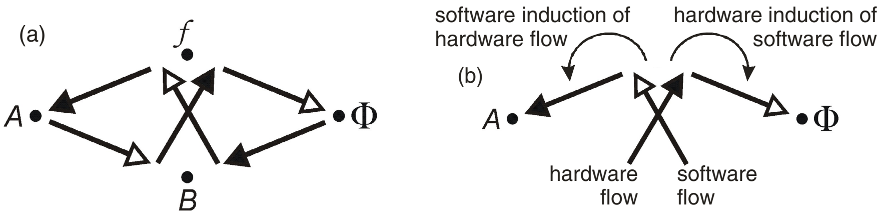


Fig. 8. (a) A symmetrical representation of Rosen's [4] (M,R)-systems, showing the induction of software flow on the left (a classical 'machine') and its complement, the induction of hardware flow, on the right. Empty-headed arrows represent software flow, and solid-headed arrows its hardware induction. ***f*** is the functor driving the metabolism (M), ***A*** represents the environmental input of material and ***f-A-B*** corresponds to the metabolic activity; **Φ** is the repair (R) functor, recreating ***f*** from ***B***. (b) A characterization of the mid-region of the representation, where {software flow}, {hardware flow}, {induction of software flow} and {induction of hardware flow} co-exist

An apparent contradiction must now be resolved. We have suggested that the inter-scale complex regions are Rosennean in nature – thus archetypically *analogue* – and that they would therefore require an infinite number of models to completely formally represent them, or an infinite number of bits for accurate digital representation. But we have also concluded that our Natural hierarchy is a primarily *digital facsimile* of an *analogue* view of our surroundings. However, this seemingly analogue view is based on the multi-parametric *digitized* form of modeling we habitually and *apparently* necessarily adopt in the human pursuit of representational accuracy. We say *apparently*, because it is only our reliance on *conventional* techniques of computation that makes it necessary to fragment the complementary characteristics of Natural phenomena in order to represent them in a tractable form. We have indicated in a previous publication [76] that organisms appear to resort to a more performant chaos-based scheme of information processing. Chaotic systems have the ability to explore their phase spaces and generate new information in a manner which is more related to their characteristic Lyapounov exponents [77] than to their characteristic processing-element size, and this makes it possible to increase information-processing density rather than merely quantity [76].

On close examination, our original formulation of a multi-scale Natural hierarchy does indeed turn out to be primarily *digital* in character, because its multi-parametric models are constructed through the fragmentation of phenomenological properties.

The complex inter-scale regions, however, should be assessed from an entirely different viewpoint, as their properties are essentially inseparable. This particular *analogue* viewpoint is totally inaccessible from a conventional Scientific platform, as it is 'strange' to traditional modeling and conventional computation: it is characteristic of a *holistic* approach, rather than a *reductive* one. We cannot stress this point too strongly: the *second* hierarchy of complex regions is *analogue* in nature, as opposed to the digital character of the *first* one. The first *scale* hierarchy is *reductive* towards localization at the right hand side of Figure 9. The second, *inter-scale* complex region hierarchy, however, is *expansive* towards the left hand side, but as Figure 9 implies this is equivalent to saying that it is *reductive towards nonlocality* [24], [27].

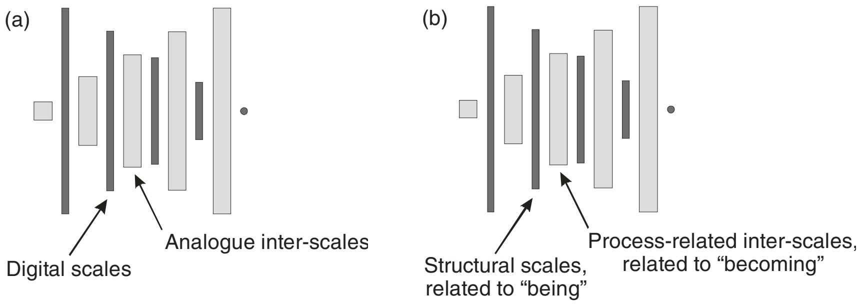


Fig. 9. (a) The analogue inter-scale regions form a second hierarchy which is partially independent of the first scale hierarchy, and (b) association of the digital scales with the *structural* term 'being', and the analogue inter-scales with the *process-related* 'becoming'

The complex contained region (see Figure 5) associated with a specific scaled containing model of a denoted entity in Figure 9 constitutes the ecosystem from which that scale will appear to have emerged. This means that at every scale of a Natural hierarchy there is a combination of *containing* information and *contained* information (Figure 5(a)) – the extant *model* and its associated *complex region*, respectively – which makes up a {*denoted-entity model – denoted-ecosystem model*} pair, and which *completely* describes the denoted entity at that scale. Transition towards the right hand side of Figure 9 is accompanied by a progressive *reduction* in the containing information from model to model, corresponding to an *increase* in the amount of hidden contained information: transition towards the left *increases* the containing information, and the contained information *decreases*. To a first approximation the total information at each scale will be the same, The intimate inseparability of the two hierarchies of Figure 9 creates a singular birational framework within which *any and all* Natural entities and phenomena can be embedded. We believe that computation within this *birational* framework can be closer to Nature than conventional mono-rational approaches, most particularly for complex and living systems. We conclude that Scientific or philosophical investigations should *always* be formulated birationally, in a way which is related to the Western interpretation of *yin-yang* [26] – as a complementary pair rather than an alternation of opposites.[26]

As indicated in Figure 9, the two hierarchies possess completely different characters. It should be noted that a model does not necessarily constitute static properties alone, and that both *structure* and *process* have implications for each other. We have earlier indicated a belief that the dynamic 'material' from which a Natural hierarchy is 'constructed' is a complement of these two reductively-separated aspects of structure and process, and we have referred to their complement by the term *struccess* [78]. We have also expressed the opinion that this 'material' has also a limited character of awareness [18]. The digital-representation model scales of a Natural hierarchy themselves are nominally structural with procedural implications; the analogue inter-scale regions are nominally procedural with structural implications, thus the entire dual hierarchy invokes both static and dynamic facets of reality. The birational character of Natural hierarchy reflects Brenner's [19] Principle of Dynamic Opposition, and we will see later that his 'actual' and 'potential' can be located at its heart.

Transition from the right hand side of Figure 9 towards the left involves passing through sequentially-scaled models whose digital facsimiles contain more and more information, and whose representations approximate better and better the analogue properties they attempt to reproduce. However, we must address the *real time accessibility* of these models with an eye on the information-processing timescales they consequently impose. Models towards the right hand side of Figure 9 contain limited amounts of information, and they are therefore conducive to rapid evaluation. If the entity denoted by our hierarchy is an organism, then these simplified models facilitate rapid response to an external threat. Models towards the left hand side of Figure 9 contain more detailed information, which would demand more time-consuming processing. This is comparable to the potential of our own brains, where a rapid reactive system through the amygdala (related to 'fear-learning' [79]) is available to bypass the more accurate but slower processing of the cerebral cortex.

Ultimately, such a birational Natural dual hierarchy describes the past, current or future emergence or evolution of any differentiated entity from its Universal origin in the Big Bang, and the farther we progress towards the left hand side of Figure 9 the more we approach *perfect nonlocality*. Conversely, the more we approach the right hand side of Figure 9 the more we approach the simple closed-off non-communicating character of *pure localization* or formal logic.[27] These two extremes of perfect nonlocality and pure localization provide inaccessible spatial dimensional extremes between which *any* differentiated entity *must* and *can only* exist: they are Platonic in their perfection or purity. This makes our usual conception of *existence* itself a relative derivative of these two absolute models of *nonlocality* and *localization*!

---

[26] Brenner's LIR view is that they are indeed opposites, but not mutually exclusive ones, and that from their dynamic opposition emerges a third term, which is their conjunction, the *Tao*.

[27] Godel's incompleteness theorems [3] indicate, as might be expected, that the description we are presenting is itself incomplete, as any model involving Rosennean complexity must be.

The model scales which appear in our first Natural hierarchy each constitutes the totality of information which is required to completely describe the denoted entity at that scale *as viewed from outside*. This descriptive or *containing* information, however, is progressively less and less able to *accurately* represent in detail the entire denoted entity as we move farther and farther towards the right hand side of Figure 9. Each scaled model, therefore, is necessarily an incomplete description, and any information which has been automatically subsumed or *contained* during its generation is omitted. This contained information is effectively 'invisible' at that scale of modeling, and it can be likened to the *hidden variables* [80] which appear, for example, in David Bohm's treatment of Quantum Mechanics in a realistic setting, where he states that:

> "*In the enfolded (or implicate) order, space and time are no longer the dominant factors determining the relationships of dependence or independence of different elements. Rather, an entirely different sort of basic connection of elements is possible, from which our ordinary notions of space and time, along with those of separately existent material particles, are abstracted as forms derived from the deeper order. These ordinary notions in fact appear in what is called the 'explicate' or 'unfolded' order, which is a special and distinguished form contained within the general totality of all the implicate orders*" [81]

While Bohm's stated relationship between *implicate* and *explicate* orders is less than fashionable, we believe that his explicate order corresponds to the *denoted entity* of our first (single) Natural hierarchy: it consequently appears in various forms in a multi-scale representation. His implicate order *at a specific scale* would then be the entirety of that scale's *pre-emergent* information, and his explicate order would be the scale's *post-emergent* model description, or *containing* information. But *where* is the post-emergent *contained* information located in our birational hierarchy? We submit that the total hidden or *contained* information in a birational Natural hierarchy constitutes the totality of the complex inter-scale regions. But, *for a specific scale*, will the contained information be its pre-emergent complexity, to its left in Figure 9, or its post-emergent complexity, to its right in Figure 9? Well, neither and both!

We pointed out earlier that a Natural hierarchy constitutes an *interface* between analogue and digital representations, and that in order to constitute a complex region we must insert *at least* one facsimile of our entire hierarchy between adjacent scales of the hierarchy *itself*. More precisely, if we are to identify the extant scales of a Natural hierarchy as digital representations and the complex regions as analogue ones, we must establish the scale sequence illustrated in Figure 10(a): transition from digital scale A to higher digital scale B requires first a *digital-to-analogue* hierarchical interface, then the expected analogue region, and lastly a hierarchical *analogue-to-digital interface.*

While we might intuitively expect the *contained* information of scale A to constitute uniquely a pre-emergent complexity to its left, prior creation of scale A will also have generated to its right a potentiality for creating scale B approximately equivalent to the A-to-B analogue region indicated in Figure 10(a). The subsequent emergence of scale *B* will recursively adapt, or *slave* [75] scale *A* and its associated information (and indirectly all other extant scales). Following this emergence of scale *B*, the complete post-emergent contained information associated with scale *A* will then be approximately as shown: it will neither be the entire complex region to the left of *A*, nor the entire region to its right, but 'part of each' (Figure 10(b)). Although Brenner's [19] 'actuality' may be identified with the extant scales – *A* and *B* here – it remains difficult to uniquely associate a specific region of the hierarchy with his 'potentiality'. Given that in an ecosystemic system 'a bear' is not equal and opposite to 'a bear-hole', this should come as no surprise.

A major characteristic of the representation shown in Figure 10(b) is that the digital and analogue layers are no longer completely separated: each of digital and analogue encroaches on the other, thus providing a progressive change between them. This corresponds to an early proposition [58] for a structure for coupling comparative analogue processing to digital decision-making in a thresholded multi-scale system, where the threshold progressively changed from analogue to digital between comparison and decision-making.

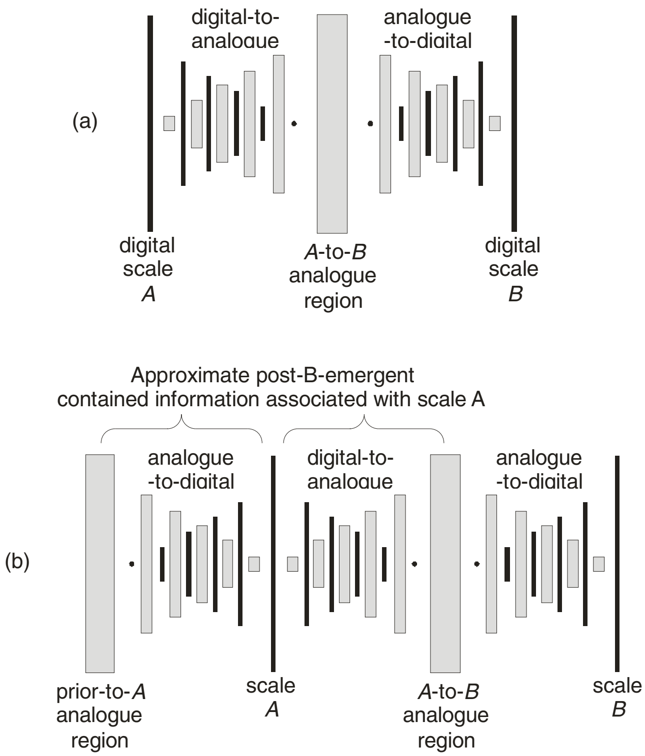


Fig. 10. (a) The sequence of elements on transiting from lower scale *A* to higher scale *B*; (b) approximate contained information associated with scale *A* after the emergence of scale *B*

As usual, it is much easier to describe the character of a static state than to specify how it comes into being. As Pirsig [82] suggests:

> "*A Dynamic advance is meaningless unless it can find some static pattern with which to protect itself from degeneration back to the conditions that existed before the advance was made. Evolution can't be a continuous forward movement. It must be a process of ratchet-like steps in which there is a Dynamic movement forward up some new incline and then, if the result looks successful, a static latching-on of the gain that has been made; then another Dynamic advance; then another static latch*".

Deacon [16] puts forward a similar proposal: that teleological processes result in constraint stabilization and support the development of life through entropy 'ratcheting'.

A major disadvantage of mono-rational approaches to Natural multi-scale systems is that relationships between the different scales must be 'inserted' from outside. The *birational* hierarchy we have described takes account of static properties, but it also intimately integrates both structure and processes of change through an overall context-dependent partiality. Not only are the individual scales partially integrated and partially autonomous, so are the individual sub-hierarchies themselves: change in *any* property of *any* part of a birational Natural hierarchy results in partial correlated changes in *every* other property. We believe that computation within this *birational* framework can be far closer to Nature than conventional mono-rational approaches, most particularly for complex and living systems.

13 Unavoidable Consequences and Conclusions

The imposition of birational ecosystemic principles on the conventionally mono-rational domains of Science and Philosophy results in unexpected and unavoidable consequences. Principal among these is a change in our perception of existence, which in a mono-rational scheme must be *absolute*. Not so in an ecosystemic context, where entities or phenomena 'exist' in a context-dependent intermediate Aristotelian condition *relative* to the spatiotemporally-dimensional Platonic extremes of *pure localization* and *perfect nonlocality*. These two extremes themselves 'exist' as a pair of complementary paradigms related to post-Newtonian-*classical* and *quantum* mechanics, respectively, which are found to be of equivalent importance rather than a discarded imperfect classical precursor and a currently-adopted undisputable quantum 'reality'. Existence *itself* is fragmented between numerous organizational scales which are recursively coupled by facsimiles of the entire bi-paradigmatic structure itself to create a multiply-fractal 'living' hierarchy. Nowhere is there a functionally isolated entity or phenomenon; only *degrees* of interconnectedness, from the weak to the strong, but *never* zero or complete.

Our original mission was to bring *all* of Science under a generalized umbrella of *entity* and its *ecosystem*, and then characterize different types of entity by their relationships with their relevant ecosystems. This can be achieved within the birational ecosystemic hierarchical framework we have described, where any and all definable entities and processes are contextually-dependent and consequently more or less coupled or autonomous.

The overall representational scheme we envisage is as follows:

As in our Figure 3; at the highest, most abstract level sits the potential-actual couple advocated by Lupasco [28] and Brenner [19]. All representation is derivative of this principle. *Reality* can be represented in terms of a multitude of conceptions, or *denotable entities*, both simple and complex: matter; energy; entropy; particles; waves; biology; hierarchy; Science; life; society; politics; theology… , and in each case potentiality and actuality are inaccessible and balanced, and reality emerges from potentiality and actuality as an *included middle*,[28] contrary to the more traditionally philosophical position of excluded middle.

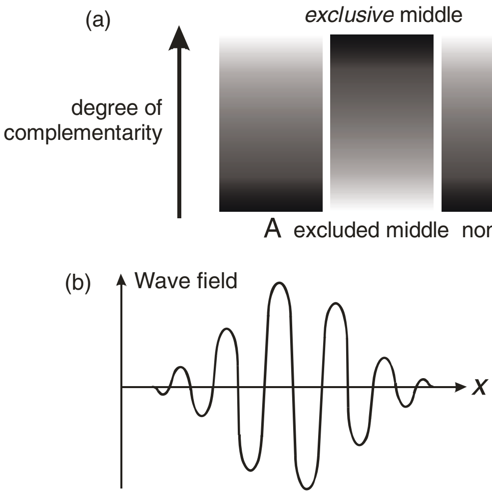


Fig. 11. (a) The transition from excluded to *exclusive middle* depending on the degree of complementarity, illustrated in terms of Brenner's coupled {A and non-A}. The borders between the middle and the extremes are unrealistically portrayed as abrupt: they could more be realistically represented in terms of a recursive application of Dempster-Shafer probability, and (b) the form of a quantum-mechanical photonic wave-packet, exhibiting neither zero (particulate) linear extension *x* nor infinite (wave) extension, but its *exclusive middle* state between the two of them

Our own preferred description here would be of an *exclusive middle*, whose predominance and distinction from the classical excluded middle depends on the degree of complementarity between the potential-actual extremes (see Figure 11). A notable distinction between Brenner's {A and non-A} and our own formulation is that we are

[28] (With regard to) the "included middle", this is one and perhaps not the best translation of the Lupasco idea, developed by Nicolescu [83], of *le tiers inclus*. Nicolescu prefers 'included third', which carries the flavor of an ontologically real entity, emerging from real contradictorial interactions. (Joseph Brenner: private communication).

portraying a reality as an emergence from its ecosystem, where 'the bear' and 'the bear-hole' are not exactly equal and opposite.

Next we must choose a representative medium. Our choice is clearly the Natural hierarchy, although numerous other formulations could be imagined. In Figure 3 we have portrayed the hierarchical representation of Salthe's [44] hierarchies. This brings to the fore an important aspect of Natural hierarchy, illustrated in Figure 12(a), which we must emphasize. The formalizations which can be said to exist at a particular scale of representation are emergences from the ecosystem at that scale. An even minor change in the represented *denoted entity* can bring to the fore other, possibly very different emergences: change in the expected character of hierarchy, for example, could result in completely different styles of formalization from Salthe's [44] compositional and subsumptive modes (see, for example, the note we made earlier about the tree represented in Figure 5). It is not even necessarily the case that the positive or negative sense of elements at a given scale must be self-consistent. Figure 12(b) suggests that the recognizable thermodynamic and informational elements of *entropy* can fit into a similar hierarchical representation, even though their senses are opposite.

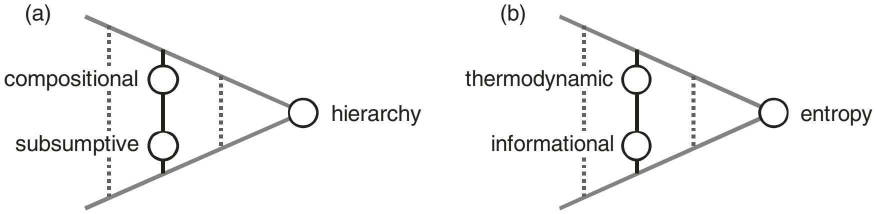


Fig. 12. (a) The hierarchical constituents of hierarchy – Salthe's compositional and subsumptive hierarchies, and (b) the hierarchical constituents of entropy – thermodynamic and informational entropy

And what happens when we represent hierarchically the *denoted entity* of *hierarchy itself*? Our entire scheme – including Brenner's potential-actual couple – becomes recursive, in the same way that progression from lower towards higher scales of a Natural hierarchy is recursive. In fact, there is no difference: the Natural hierarchy itself includes its own recursivity, and the sensitivity to definition of a scaled denoted entity leads to a process of evolution.

We noted earlier that an ecosystemic approach could be related to Turing's advances in code-breaking [30], through the insistence on relating an individual coded message to the globality of its 'environment' of place, date, time, weather,… but also that it is impossible to generate in a (monorational) digital (conventional universal) computer any phenomenon which depends on globally-coupled properties, such as *intelligence*, *life* or *consciousness* [18]. From an unconventional-computation point of view these global properties are all regarded as the outcomes of Natural computation, and the hope is that advances in Natural styles of computation will make possible the generation of globally-coupled properties. Birational ecosystemic hierarchy holds out the exciting prospect of fabricating more generally applicable computational machines, conceivably bypassing the 'brick wall' of complexity currently facing artificial intelligence and opening the way to the development of real non-biological intelligent systems.

The relationship between life and informational entropy provides an excellent example of the way in which the birational approach creates new viewpoints. Informational entropy is most usually referred to as the inverse of order. Given equal numbers of black and white balls, the most ordered arrangement is taken to be with all the black balls together and all the white balls together, as in Figure 13(a). Reasonably, the least ordered, most entropic arrangement would then be the alternation of black and white balls shown in Figure 13(b). However, this latter is really an alternate kind of order, which is found in the atomic arrangement of crystalline gallium arsenide, for example, where atom types alternate through the crystal [41].

If we associate informational entropy with the inverse of order, we now need *two different* kinds of entropy to associate with the two different kinds of order – informational-entropy.1 and informational-entropy.2. This provides a good example of a change in the *denoted entity* (entropy) which results in a change in the number of extant sub-types – see Figure 14, compared to Figure 12(b).

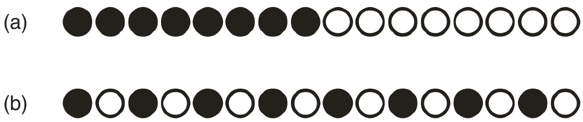


Fig. 13. (a) A conventional two-component ordered state; (b) the alternate ordered state

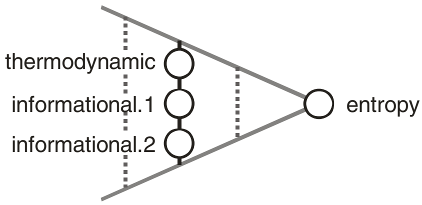


Fig. 14. Modification in the definition of the *denoted entity* of entropy results in multiplication of the number of sub-types, now including two types of informational entropy

Here again we meet with a pair of dimensional extremes enclosing reality.[29] If we start from one of these two extreme ordered states and move towards the other we would expect the initial state's associated entropy to progressively rise, and the other state's associated entropy to progressively fall away (Figure 15). We suggest that life colonizes the mid-region between the two kinds of order where the total summed entropy may be lowest.

This raises a number of fascinating questions. Life has been described as being located 'at the edge of chaos' [84]. Is the mid-ground of Figure 15 this region? Or does it correspond to Sabelli's [85] *bios* – one step beyond chaos? Deacon [16] has

[29] Neither of these two ordered states is accurately accessible in reality: the sequences would need to be infinitely long, while in reality the 'observational window' is always much smaller than this.

indicated that life is related to a balance between thermodynamic entropy and informational entropy. Does this mid-ground correspond to the informational entropy required in his proposal?

Complementarity is the order of the day, whether this is unformalizably-intimate, as in recursively-functional organisms, or formalizably-distant, as in the loosely-tractable 'inorganic' subjects of conventional Science. Complementarity may be approximated in various ways, from its most simplistic representation as a pair of opposites to the extreme complexity which apparently characterizes *life*. Science habitually resorts to a two-part representation: first the institution of a set of non-recursive orthogonal parameters, whether dependent or independent, then the establishment of a formal, preferably linear relationship between them.[30] This effectively ejects recursively-functional living organisms from Science's purview; thus Rashevsky [86] and Rosens' [4] particular focus on the *interrelational* characteristics of life.

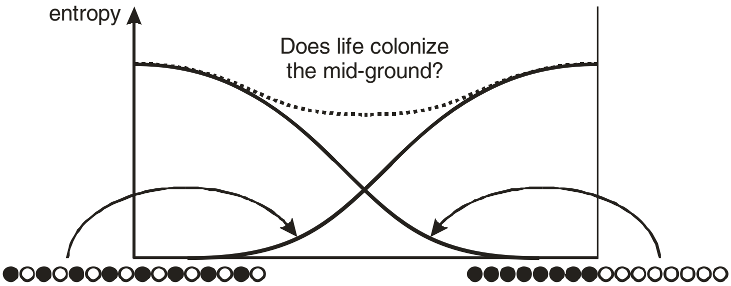


Fig. 15. The progression of entropy on moving away from each of the extreme ordered states, and the hypothesized colonization by life of the mid-ground where entropy is lowest

The birational ecosystemic association is fundamentally one of combined existential differentiation and integration, neither one nor the other, where predominance depends unavoidably on context. Differentiating segregation leads to the *reductionism* of conventional Science; unifying integration leads to the *holism* of human relations. Either of these can be embedded in the general hierarchical scheme we have presented. We believe that computation within this *birational* framework can be far closer to Nature than conventional mono-rational approaches, most particularly for complex and living systems.

Acknowledgement. The authors wish to thank Gordana Dodig-Crnkovic, Joseph Brenner and Stan Salthe for their constructive and helpful comments in connection with this chapter.

[30] An excellent example of this approach can be found in electronics, where a resistive component is first characterized in terms of the institution of non-interactive orthogonality between two functionally-dependent parameters – voltage V and current I – and then the component's resistance R is established by the formal relation R = V/I.

References

1. Liboff, R.L.: Introductory Quantum Mechanics. Addison Wesley Longman, Boston (2002)
2. Cronan, C.S.: Introduction to Ecology & Ecosystems Analysis. Shaw-Ferguson Environmental Publications, Orono (1996)
3. Berto, F.: There's Something about Gödel: The Complete Guide to the Incompleteness Theorem. John Wiley and Sons, Hoboken (2010)
4. Rosen, R.: Life Itself. Columbia University Press, New York (1991)
5. Hardie, R.P., Gaye, R.K.: Aristotle: Physics, `http://etext.library.adelaide.edu.au/a/aristotle/physics/`
6. Cottam, R., Ranson, W., Vounckx, R.: Re-mapping Robert Rosen's (M,R) Systems. Chemistry and Biodiversity 4, 2352–2368 (2007)
7. Einstein, A.: Relativity: The Special and General Theory. Translated by Robert W. Lawson. Henry Holt and Company, New York (1920)
8. Chandler, D.: Semiotics for Beginners, `http://www.gregeckler.com/images/pdf/AD235/SemioticsForBeginners.pdf`
9. Cottam, R., Ranson, W., Vounckx, R.: A Diffuse Biosemiotic Model for Cell-to-Tissue Computational Closure. BioSystems 55, 159–171 (2000)
10. Fisch, M.H., Ketner, K.L., Kloesel, C.J.W. (eds.): Peirce, Semeiotic and Pragmatism. Indiana University Press, Bloomington (1986)
11. von Uexküll, T.: The Sign Theory of Jakob von Uexküll. In: Krampen, M., Oehler, K., Posner, R., Sebeok, T.A., von Uexküll, T. (eds.) Classics of Semiotics, pp. 147–179. Plenum Press, New York (1987)
12. Agamben, G.: The Open: Man and Animal. Translated by Kevin Attell. Stanford University Press, Stanford (2004)
13. Carson, R.: Silent Spring. Houghton Mifflin, Boston (1962)
14. von Foerster, H.: Cybernetics of Cybernetics. University of Illinois Press, Urbana (1974)
15. Koffka, K.: Perception: An Introduction to the Gestalt-Theory. Psychological Bulletin 19, 531–585 (1922)
16. Deacon, T.W.: Incomplete Nature: How Mind Evolved from Matter. W. W. Norton & Co., New York (2012)
17. Root-Bernstein, R.S., Dillon, P.F.: Molecular Complementarity I: the Complementarity Theory of the Origin and Evolution of Life. Journal of Theoretical Biology 188, 447–479 (1997)
18. Cottam, R., Ranson, W.: A Biosemiotic View on Consciousness Derived from System Hierarchy. In: Pereira Jr., A., Lehmann, D. (eds.) The Unity of Mind, Brain and World: Current Perspectives on a Science of Consciousness. Cambridge University Press, Cambridge (2012)
19. Brenner, J.E.: Logic in Reality. Springer, Berlin (2008)
20. Brenner, J.E.: The Logical Dynamics of Information: Deacon's "Incomplete Nature". Information 3, 1–36 (2012)
21. Matsuno, K.: The Internalist Stance: a Linguistic Practice Enclosing Dynamics. Annals of the New York Academy of Sciences 901, 322–349 (2000)
22. Cottam, R., Ranson, W., Vounckx, R.: Towards Cross-modeling between Life and Solid State Physics. In: Simeonov, P.L., Smith, L., Ehresmann, A.C. (eds.) Integral Biomathics: Tracing the Road to Reality, pp. 85–95. Springer, Berlin (2012)

23. Cottam, R., Ranson, W., Vounckx, R.: Diffuse Rationality in Complex Systems. In: Bar-Yam, Y., Minai, A. (eds.) Unifying Themes in Complex Systems, vol. II, pp. 355–362. Westview Press, Boulder (2004)
24. Cottam, R., Ranson, W., Vounckx, R.: Localisation and nonlocality in computation. In: Holcombe, M., Paton, R. (eds.) Information Processing in Cells and Tissues, pp. 197–202. Plenum Press, New York (1997)
25. Cottam, R., Ranson, W., Vounckx, R.: Autocreative Hierarchy I: Structure - Ecosystemic Dependence and Autonomy. SEED J. 4, 24–41 (2004)
26. Wong, E.: Harmonizing Yin and Yang. Shambhalla Publications, Boston (1997)
27. Cottam, R., Ranson, W., Vounckx, R.: Sapient Structures for Intelligent Control. In: Mayorga, R.V., Perlovsky, L.I. (eds.) Toward Artificial Sapience: Principles and Methods for Wise Systems, pp. 175–200. Springer, New York (2008)
28. Brenner, J.E.: The Philosophical Logic of Stéphane Lupasco. Logic and Logical Philosophy 19, 243–284 (2010)
29. Mittelstaedt, P., Prieur, A., Schieder, R.: Unsharp Particle-Wave Duality in a Photon Split Beam Experiment. Foundations of Physics 17, 891–903 (1987)
30. Erskine, R., Smith, M. (eds.): The Bletchley Park Codebreakers. Biteback Publishing Ltd., London (2011)
31. Landsberg, P.T.: Is Equilibrium Always an Entropy Maximum? J. Stat. Phys. 35, 159–169 (1984)
32. Landsberg, P.T.: Can Entropy and "Order" Increase Together? Phys. Lett. A 102, 171–173 (1984)
33. Kuhn, T.: Black-Body Theory and the Quantum Discontinuity: 1894–1912. Clarendon Press, Oxford (1978)
34. Planck, M.: The Theory of Heat Radiation. Translated by Masius, M.P., 2nd edn. Blakiston's Son & Co., Philadelphia (1914)
35. Hagen, J.B.: An Entangled Bank: The Origins of Ecosystem Ecology. Rutgers University Press, New Brunswick (1992)
36. Jeavons, A.P., Saunders, G.A.: Correlation Between the Galvanomagnetic Effects and the Fermi Surface of Arsenic. Physics Letters A 27, 19–20 (1968)
37. Bennett, C.H., Brassard, G., Crepeau, C., Jozsa, R., Peres, A., Wootters, W.: Teleporting an Unknown Quantum State via Dual Classical and EPR Channels. Phys. Rev. Lett. 70, 1895–1899 (1993)
38. Huygens, C.: Traité de la Lumiere. Pieter van der Aa, Leiden (1690)
39. Udwadia, F.E., Kalaba, R.E.: Analytical Dynamics: a New Approach. Cambridge University Press, New York (1996)
40. Cottam, R., Ranson, W., Vounckx, R.: The Mind as an Evolving Anticipative Capability. J. Mind Theory, 39–97 (2008)
41. Cottam, R., Saunders, G.A.: The Elastic Constants of GaAs from 2K to 320K. J. Phys. C Solid State 6, 2015–2118 (1973)
42. Cottam, R., Ranson, W., Vounckx, R.: A Biologically Consistent Hierarchical Framework for Self-referencing Survivalist Computation. In: Dubois, D.M. (ed.) Computing Anticipatory Systems: CASYS 1999 – 3rd International Conference, AIP Conference Proceedings 465, pp. 252–262. AIP, Woodbury (2000)
43. Railsback, B.: Slime Molds (the Mycetozoa or Fungus Animals), `http://www.gly.uga.edu/railsback/1122SlimeMolds.html`
44. Salthe, S.N.: Hierarchical Structures. Axiomathes 22 (2012) (Online first March 9, 2012: accessed July 5, 2012)

45. Westbrook, G.: Levels of Organization, `http://utahscience.oremjr.alpine.k12.ut.us/sciber00/7th/cells/sciber/levelorg.htm`
46. Jagers op Akkerhuis, G.: The Operator Hierarchy: a Chain of Closures linking Matter, Life and Artificial Intelligence. Radboud University Press, Nijmegen (2010)
47. Schrödinger, E.: What is Life? The Physical Aspect of the Living Cell. Cambridge University Press, Cambridge (1944) (reprinted 1992)
48. Watson, J.D., Berry, A.: DNA: the Secret of Life. Albert A. Knopf, New York (2004)
49. Cottam, R., Ranson, W., Vounckx, R.: Living in Hyperscale: Internalization as a Search for Unification. In: Wilby, J., Allen, J.K., Loureiro-Koechlin, C. (eds.) Proceedings of the 50th Annual Conference of the International Society for the Systems Sciences, paper #2006-362, pp. 1–22. ISSS, Asilomar (2006)
50. Salthe, S.N.: Development and Evolution: Complexity and Change in Biology. MIT Press, Cambridge (1993)
51. Salthe, S.N.: Evolving Hierarchical Systems. Columbia University Press, New York (1985)
52. Levins, R.: Evolution in Changing Environments. Princeton University Press, Princeton (1968)
53. Gutowitz, H.A., Langton, C.G.: Mean Field Theory of the Edge of Chaos. In: Morán, F., Merelo, J.J., Moreno, A., Chacon, P. (eds.) ECAL 1995. LNCS, vol. 929, pp. 52–64. Springer, Heidelberg (1995)
54. Cottam, R., Ranson, W., Vounckx, R.: Life and Simple Systems. Sys. Res. Behav. Sci. 22, 413–430 (2005)
55. Dodig-Crnkovic, G.: Investigation into Information Semantics and Ethics of Computing. Mälardalen University Press Dissertations #33, Mälardalen (2006)
56. Kurzweil, R.: The Age of Intelligent Machines. MIT Press, Cambridge (1992)
57. Lloyd, S.: Programming the Universe: A Quantum Computer Scientist Takes on the Cosmos. Alfred A. Knopf, New York (2006)
58. Langloh, N., Cottam, R., Vounckx, R., Cornelis, J.: Towards Distributed Statistical Processing – Aquarium: A Query and Reflection Interaction Using Magic: Mathematical Algorithms Generating Interdependent Confidences. In: Smith, S.D., Neale, R.F. (eds.) ESPRIT Basic Research Series, Optical Information Technology, pp. 303–319. Springer, Berlin (1993)
59. Singh, S.: Big Bang: the Origin of the Universe. HarperCollins, New York (2004)
60. Adams, F.C., Laughlin, G.: A Dying Universe: the Long-term Fate and Evolution of Astrophysical Objects. Rev. Mod. Phys. 69, 337–372 (1997)
61. Mandelbrot, B.: Fractals and Chaos. Springer, Berlin (2004)
62. Collier, J.D.: Autonomy in Anticipatory Systems: Significance for Functionality, Intentionality and Meaning. In: Dubois, D.M. (ed.) Computing Anticipatory Systems: CASYS 1998 - 2nd International Conference, AIP Conference Proceedings 465, pp. 75–81. AIP, Woodbury (1999)
63. Hardesty, L.: Chips as Mini Internets: Data-routing Techniques that Undergird Internet could Increase Efficiency of Multicore Chips, `http://www.sciencedaily.com/releases/2012/04/120410145958.htm`
64. Mikulecky, D.C.: Editorial. Chemistry and Biodiversity 4, 2269–2271 (2007)
65. Golshan, K.: Physical Design Essentials: an ASIC Design Implementation Perspective. Springer, New York (2007)
66. Chalmers, D.J.: The Conscious Mind: In Search of a Fundamental Theory, pp. 330–331. Oxford University Press, Oxford (1996)

67. Darwin, C.: On the Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life. John Murray, London (1859)
68. Sigmund, K.: Games of Life: Explorations in Ecology, Evolution, and Behavior. Oxford University Press, New York (1993)
69. Zhuo, Z., Cai, S.-M., Fu, Z.-Q., Zhang, J.: Hierarchical Organization of Brain Functional Networks during Visual Tasks. Phys. Rev. E 84, 031923 (2011)
70. Rossion, B., Dricot, L., Goebel, R., Busigny, T.: Holistic Face Categorization in Higher Order Visual Areas of the Normal and Prosopagnosic Brain: Toward a Non-hierarchical View of Face Perception. Front Hum. Neurosci. 4, 225 (2010)
71. Katsumori, M.: Niels Bohr's Complementarity: Its Structure, History, and Intersections with Hermeneutics and Deconstruction. Boston Studies in the Philosophy and History of Science, vol. 286. Springer, New York (2011)
72. Tononi, G.: An Information Integration Theory of Consciousness. BMC Neuroscience 5, 42 (2004)
73. Schroeder, M.J.: The role of information integration in demystification of holistic methodology. In: Simeonov, P.L., Smith, L.S., Ehresmann, A.C. (eds.) Integral Biomathics: Tracing the Road to Reality, vol. 106, pp. 283–296. Springer, Heidelberg (2012)
74. Lewes, G.H.: Problems of Life and Mind (First Series), vol. 2, p. 412. Trübner, London (1875)
75. Haken, H.: The Science of Structure: Synergetics. Prentice Hall, New York (1984)
76. Cottam, R., Ranson, W., Vounckx, R.: Emergence: Half a Quantum jump? Acta Polytech Sc. Ma 91, 12–19 (1998)
77. Nicolis, G., Prigogine, I.: Exploring Complexity, p. 21. Freeman, New York (1989)
78. Cottam, R., Ranson, W., Vounckx, R.: Life as its Own Tool for Survival. In: Allen, J.K., Hall, M.L.W., Wilby, J. (eds.) Proceedings of the 43rd Annual Conference of the International Society for the Systems Sciences, paper #99268, pp. 1–12. ISSS, Asilomar (1999)
79. LeDoux, J.E.: Brain Mechanisms of Emotion and Emotional Learning. Current Opinion in Neurobiology 2, 191–197 (1992)
80. Bohm, D., Hiley, B.J.: The Undivided Universe. Routledge, London (1993)
81. Bohm, D.: Wholeness and the Implicate Order. Routledge, London (1980)
82. Pirsig, R.M.: Lila: an Enquiry into Morals, p. 176. Bantam Press, New York (1991)
83. Nicolescu, B.: Manifesto of Transdisciplinarity. State University of New York Press, Albany (2002)
84. Langton, C.G.: Computation at the Edge of Chaos: Phase Transitions and Emergent Computation. Physica D 42, 12–37 (1990)
85. Sabelli, H.C.: Bios: a Study of Creation. World Scientific Publishing Co., Singapore (2005)
86. Rashevsky, N.: Topology and Life: In Search of General Mathematical Principles in Biology and Sociology. Bulletin of Mathematical Biophysics 16, 317–348 (1954)

The Coordination of Probabilistic Inference in Neural Systems

William A. Phillips

Department of Psychology, University of Stirling and Frankfurt Institute of Advanced Studies
wap1@stir.ac.uk.

Abstract. Life, thought of as adaptively organised complexity, depends upon information and inference, which is nearly always inductive, because the world, though lawful, is far from being wholly predictable. There are several influential theories of probabilistic inference in neural systems, but here I focus on the theory of Coherent Infomax, and its relation to the theory of free energy reduction. Coherent Infomax shows, in principle, how life can be preserved and improved by coordinating many concurrent inferences. It argues that neural systems combine local reliability with flexible, holistic, context-sensitivity. What this perspective contributes to our understanding of neuronal inference is briefly outlined by relating it to cognitive and neurophysiological studies of context-sensitivity and gain-control, psychotic disorganization, theories of the Bayesian brain, and predictive coding. Limitations of the theory and unresolved issues are noted, emphasizing those that may be of interest to philosophers, and including the possibility of major transitions in the evolution of inferential capabilities.

Keywords: Probabilistic inference, Coherent Infomax, free energy reduction.

1 Introduction

Many forms of organised complexity have arisen during nature's long journey from uniformity to maximal disorder, despite the ever present forces of noise and disorder. Biological systems are able to create and preserve organised complexity, by, in effect, making good predictions about the likely consequences of the choices available to them. This adaptively organised complexity occurs in open, holistic, far-from-equilibrium, non-linear systems with feedback. Though usually implicit, probabilistic inference is crucial, and useful inference is only possible because the laws of physics are sufficiently reliable. The endless variety of individual circumstances and the prevalence of deterministic chaos and quantal indeterminacy make many things uncertain, however; so, to thrive, biological systems must combine reliability with flexibility.

It is in neural systems that the crucial role of probabilistic inference is most obvious. Helmholtz correctly emphasized the centrality of unconscious inference to perception, and many examples of its use for contextual disambiguation can be given [1]. Furthermore, it has also now been explicitly shown how such unconscious inference may also be central to reinforcement learning, motor control, and many other biological processes [2].

G. Dodig-Crnkovic and R. Giovagnoli (Eds.): *Computing Nature*, SAPERE 7, pp. 61–70.
DOI: 10.1007/978-3-642-37225-4_3

Better formalisation of these issues is clearly needed, so Section 3 outlines an elementary neurocomputational perspective that uses information theory measures to shed light on them, i.e. the theory of Coherent Infomax [3, 4, 5]. A major advantage of that theory is that, in addition to being formally specified and simulated in large artificial neural networks, it has wide-ranging empirical roots, being related, often in detail, to much empirical data from neuroanatomy, cellular and synaptic physiology, cognitive psychology, and psychopathology. Section 4 briefly discusses relations between this theory and that of free energy reduction [2], to which it has deep connections, and which has been applied to an even wider range of phenomena than has Coherent Infomax. Finally, in Section 5, difficulties of the theory and unresolved issues of possible philosophical interest are briefly discussed. First, the following section outlines some of the difficult conceptual problems to be solved by theories of neuronal inference.

2 Theories of Neuronal Inference and Difficult Problems That They Must Solve

The preceding arguments suggest several issues on which we need to make progress. What is organised complexity? What are the capabilities and constraints of various forms of inductive inference, e.g. classical versus Bayesian [6], conscious versus unconscious [7]? How is reliability combined with flexibility, i.e. how is information about reliable generalities combined with information about individual particularities? How is localism combined with holism? What forms of learning and processing does neural inference require, and how are they implemented at the synaptic, local circuit, and systems levels? Do biological capabilities for probabilistic inference evolve towards forms of inference with greater accuracy, generality, or abstraction? Information theory measures such as Shannon entropy and free-energy have been applied to these issues, but how can they be tested and what do they contribute to our understanding?

Several theories of probabilistic inference in neural systems have been proposed, including the Bayesian brain [8, 9], predictive coding [10], efficient coding and Infomax [11, 12, 13], and sensorimotor integration [14]. It has been argued that all can be unified via the principle of least variational free energy [2, 15, 16]. The free energy principle is formulated at the level of the interaction of the system with its environment – and emphasizes Bayes optimal inference using hierarchical architectures with backward as well as forward connections. As free energy theory offers a broad synoptic view of neuronal inference the Coherent Infomax theory will be compared to that.

The theory of Coherent Infomax stresses the necessity of avoiding information overload by selecting only the information that is needed. This necessity arises not only from requirements of computational tractability, but also from an unavoidable property of noisy high-dimensional spaces. As dimensionality increases the number of possible locations in that space increases exponentially, with the consequence that nearly all events occur at novel locations. Probabilistic inference based on prior events then becomes impossible. This problem is well-known within the machine learning community, where it is referred to as the ‘curse-of-dimensionality’. It may be avoided by selecting only the information that is ‘relevant’; but how? Coherent Infomax suggests a solution: select information that reveals latent statistical structure in the available data. Useful combinations of datasets between which to seek predictive relations may be found by genetic search prescribing gross system architectures combined with the learning algorithms of Coherent Infomax, as outlined in the following section.

3 The Theory of Coherent Infomax: A Brief Outline

An unavoidable consequence of the curse-of-dimensionality is that large amounts of data must be divided into subsets that are small enough to make learning feasible. If they were processed independently, however, then relations between the subsets would be unobservable. Success in finding useful relations would then be completely dependent upon the original division into subsets, but that is unlikely to be adequate unless the crucial relations were already known. Coherent Infomax responds to this dilemma by dividing data at each level of an interpretive hierarchy into many small subsets, and searching for variables defined on them that are predictably related across subsets. This strategy allows for endlessly many ways in which the data can be divided into subsets and linked by modulatory coordinating interactions between them.

These considerations suggest minimal requirements for local neural processors performing such inference. They must have a subset of inputs within which latent variables may be discovered and compressed into fewer dimensions. These are referred to as driving, or receptive field (RF), inputs. They must also receive inputs conveying information about the activity of other processors with which they are to seek predictive relations. These are referred to as contextual field (CF) inputs. They control the gain of response to the driving RF inputs but cannot by themselves drive processor activity, because, if they did, that would contradict the strategy for avoiding the curse-of-dimensionality. Given this constraint, local processors can have a rich array of contextual inputs, far richer than the array of driving inputs within which they seek predictable variables.

The theory of Coherent Infomax has grown from combining such considerations with much empirical data from several relevant disciplines [3, 4, 5]. Only a brief outline is given here. For full formal presentations see the original publications. The theory uses three-way mutual information and conditional mutual information to show how it is possible in principle for contextual inputs to have large effects on the transmission of information about the primary driving inputs, while transmitting little or no information about themselves, thus influencing the transmission of cognitive content, but without becoming confounded with it. Guided by neuroanatomy, the gross system architecture assumed is that of at most a few tens of hierarchical layers of processing, with very many specialized but interactive local processors at each stage. Feed forward connections between layers are driving, whereas a larger number of lateral and feedback connections provide coordinating gain-control as shown in Figure 1. Minimally, the function of local processors is to select and compress that information in their driving receptive field (RF) input that is relevant to the current task and situation, as indicated by the contextual field (CF) input that modulates transmission of RF information.

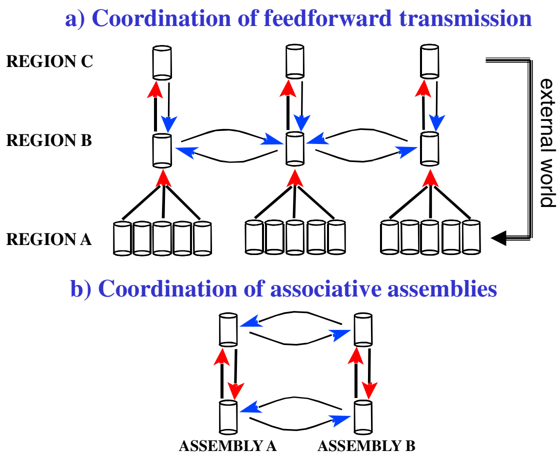


Fig. 1. Examples of system architectures that could be built from the local neural processors of Coherent Infomax, shown here as small cylindrical columns. Though only a few are shown in each region, in useful applications, as in mammalian cerebral cortex, there would be very many in each region. Receptive field connections, shown by thick lines, provide the input from which information is to be selected and compressed. Coordinating contextual field connections, shown by thin lines, control the gain of response, and provide the inputs with which predictive relations are to be sought.

This is formalized as an objective function describing the signal processing work to be done, as shown in Figure 2 by arrows associated with each of the four components of the output H(X).

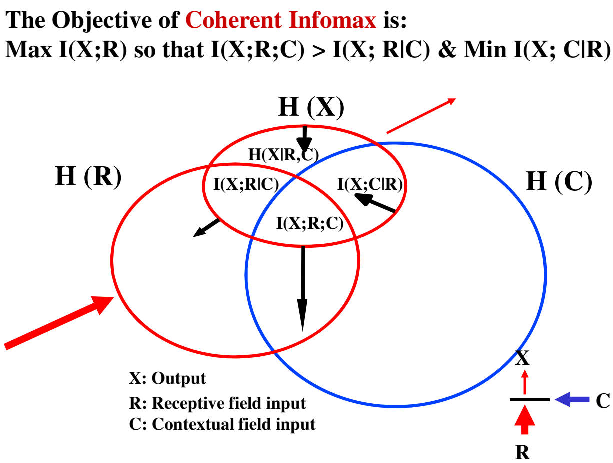


Fig. 2. The objective of local processors in Coherent Infomax. The ovals show the Shannon entropy in each of three probability distributions. Information flow through the local processor is shown in the small icon, bottom right. Contextual entropy can be greater than the other two because it is not to be transmitted in the output. Thus, it enables narrowly focussed receptive field processing to operate within a broad context.

Outward pointing arrows show components that should be increased, with priority being shown by arrow length. Inward pointing arrows show components that should be decreased. In short, the objective is to maximise the information transmitted about the receptive field input subject to the constraints of substantial data reduction while emphasizing the mutual information between receptive field input and contextual field input and minimizing any information transmitted specifically about the context. To show how that objective could be met in neural systems, a biologically plausible activation function for idealized local neural processors was formulated to include the required gain-control, and a learning rule for modifying the synaptic strengths of the connections between these local processors was derived analytically from the objective function. What most impressed us about the consequent learning rule is that, although it was deduced formally from the objective function, assuming none of the physiological evidence concerning the dependence of synaptic plasticity on current and prior activity, it is broadly in agreement with that evidence. The theory of Coherent Infomax thus shows how it is possible for neural systems to perform probabilistic inference in a way that combines reliability with flexibility, and localism with holism, while making useful inference feasible by selecting only information that is relevant, and thus avoiding the curse-of-dimensionality. It has guided studies of common neurobiological foundations for cortical computation [17], dynamic coordination in the brain [18], cognitive impairments in schizophrenia [19], and of relations between probability theory, organised complexity and brain function [20].

4 Relations to the Theory of Free Energy Reduction

The current growth of interest in inference and prediction as possible keys to a fundamental understanding of neuronal systems is seen in the many groups working on ‘predictive coding’ and the ‘Bayesian brain’ as cited in Section 2. Those theories do not usually make use of gain-control or context to select the relevant information to be coded and used, however, and rarely show explicitly how the curse-of-dimensionality can be avoided. One theory that may be able to do so, however, is that proposing a unifying brain theory based on ideas from statistical physics and machine learning [2]. This has already received deep philosophical examination, and been found to have considerable interest from that perspective [21], even though it still needs further development. It interprets many aspects of neural structure and function as having evolved to reduce Helmholtz free-energy using a form of predictive coding in which ascending activities predicted by feedback descending from higher levels in the hierarchy are suppressed. In contrast to this, Coherent Infomax proposes that activities predicted by contextual input can be amplified. Thus, the form of predictive coding used in free energy theory seems to imply effects of context that are in opposition to those of Coherent Infomax. Furthermore, the theory of free energy reduction is formulated at the level of an agent in an environment with distal causes and parameters that are hidden from the agent; Coherent Infomax is formulated at the level of local neural processors operating within a large population of other such processors, with which they can communicate either directly or indirectly.

There are at least three grounds for thinking that these two theories are not in essence opposed, however. First, both theories imply that the fundamental objective of neuronal dynamics is to reduce any differences between predicted and observed probability distributions. Indeed, it may even be possible to unify the two perspectives by formulating the objective of Coherent Infomax as the maximisation of predictive success and of free energy reduction as the minimisation of prediction failure (Phillips and Friston, in preparation). Such a common goal could be described as maximising the transmission of information that is relevant to the context, or alternatively as reducing uncertainty about sensory inputs given the contextual constraints. Second, the two theories may be complementary, rather than opposed, because Coherent Infomax emphasizes lateral connections between streams of processing dealing with distinct datasets, while also including some downward connectivity, whereas the theory of free energy reduction emphasizes downward connections, while also including some lateral connectivity. Third, it has been argued that predictive coding theories can be made formally equivalent to theories based on evidence for amplifying effects of top-down attentional inputs [22]. This was done by reorganising the computations required for predictive coding, and assuming that suppressive effects of prediction operate on intra-regional signals, rather than on inter-regional signals. Furthermore, a detailed model of that form of predictive coding argues that it is compatible with much of the neurobiological evidence [23]. These studies therefore suggest that some form of predictive coding may be compatible with both Coherent Infomax and the theory of free energy reduction. Deeper examination of relations between those two theories is therefore a major task for the future.

5 Unresolved Issues and Difficulties of the Theory

The conceptual depth and empirical scope of the free energy and Coherent Infomax theories raises many unresolved and controversial issues, some of which may have philosophical significance. There is time here to mention only a few, and each in no more than speculative and flimsy outline.

First, is any unified theory of brain function possible? As a recent philosophical examination of the free energy theory shows this is an issue of lasting debate, with the ‘neats’ saying ‘Yes’, and the ‘scruffies’ saying ‘No’ [21]. As the issue cannot be resolved by failing to find any unifying theory, it can only be resolved by finding one. Some are happy to leave that search to others, on the assumption that Darwinian evolution is the only unifying idea in biology. Even if true that need not deter the search for unifying principles, however, because it can be argued that free energy theory both formally specifies what adaptive fitness requires and shows how neural systems can meet those requirements (Friston, personal communication).

Second, another crucial issue concerns the possibility of major transitions in the evolution of inferential capabilities. Seven major transitions in the evolution of life have been identified [24], such as the transition from asexual to sexual reproduction. Only one of those concerned cognition, i.e. the transition to language. Major transitions in the evolution of inferential capabilities prior to language are also possible, however, and it is crucial to determine whether this is so because empirical studies of inferential capabilities will be misinterpreted if they are assumed to reflect a single strategy, when instead they reflect a mixture of strategies, either across or within species. One way in which work of the sort discussed here could contribute to this issue is by proposing various possible inferential strategies. They could range from those with requirements that are easier to meet but with severely limited capacities, through intermediate stages of development, to those having more demanding requirements but with enhanced capabilities. Some possible transitions are as follows: from predictions only of things that are directly observable to estimates of things not directly observable; from generative models averaged over various contexts to those that are context specific; from hypotheses determined by input data to those that are somehow more internally generated; from probabilistic inference to syntactic structure, and, finally, from hypothesis testing to pure hypothesizing freed from testing. Within stages marked by such transitions there would still be much to be done by gradual evolutionary processes. For example, context-sensitive computations can make astronomical demands on computational resources, so they will be useful only if appropriate constraints are placed on the sources and size of contextual input, as already shown for its use in natural language processing [25]. Thus, even given the ability to use contextual information, the search for useful sources of contextual input could still be a lengthy process, even on an evolutionary timescale, and produce much diversity.

Third, how can apparently simple objectives, such as specified by Coherent Infomax and free energy theory, help us understand the overwhelming evidence for wide individual differences in cognitive style and capabilities? To some extent answers to this question are already available as it has been shown that within human cognition there are wide variations in context-sensitivity across sex and occupation [26], culture [27], schizotypy [28], and developmental stage [29]. The use of these theories to help us understand the diversity of cognitive capacities both within and between species is in its infancy, however.

Fourth, why are there several different neurobiological mechanisms for gain-control? Earlier work done from the Coherent Infomax perspective, both in relation to normal and psychotic cognition [19], emphasized only NMDA synaptic receptors for the predominant excitatory neurotransmitter glutamate, but we now realize that several other gain-control mechanisms are also important, particularly at the level of micro-circuitry involving inhibitory inter-neurons. The various uses, capabilities and limitations of these different mechanisms for gain-control remain to be determined.

Fifth, as Coherent Infomax is formulated at the level of local neural processors that operate only within a population of other such processors, are they not doomed to imprisonment in such a ‘Chinese room’, with no hint of a world beyond? As Fiorillo argues, neuroscience must be able to ‘take the neurons perspective’ [30], but how can that be done without thereby losing contact with the distal world beyond? Coherent Infomax suggests an answer to this dilemma, first, by being formulated explicitly at the level of the local neuronal processor, and, second, by searching for predictable relations between diverse datasets. Discovery of such interdependencies implies the existence of distal causes that produce them. The more diverse the datasets the more distal their common origins are likely to be. This can be seen as a neurocomputational version of Dr Johnson's refutation of idealism when he kicked a stone and said "I refute it thus". A distal reality is implied both by the agreement between what he sees and what he feels, and by his successful prediction of the outcome of his actions. Though this argument seems plausible to me, I am not a philosopher, so it may be in need of closer philosophical examination.

Sixth, coherence, as conceived within this theory, depends upon the long-term statistics of the co-occurrence of events defined at the highly specialized level of receptive fields, which convey information only about fragments of the current state as a whole, so how can episodic capabilities that deal with unique events, such as working memory and episodic memory, be included within such a conception? My working assumption is that these capabilities are closely related to the syntactic grammars of language and schematic structures. Though syntactic and statistical conceptions of cognition have long been contrasted, there is no fundamental conflict between them because, as many studies have shown, grammars can be acquired by statistical inference. The use of such grammars to create novel but relevant patterns of activity seems to be, in essence, close to what the theory of Coherent Infomax has to offer, but I know of no attempt to explore that possibility.

Seventh, how can attention and consciousness be included within these theories? Within Coherent Infomax, attention is assumed to operate via the contextual inputs, which are purely modulatory as required. One psychophysical study of texture perception by humans used the formal information theoretic measures of the theory, and, indeed, in that case attention had the complex set of properties predicted [30]. That one study has not been followed-up, however, and though it has promise, far more needs to be done. The theory of free energy reduction has also been related in detail to attention [31], but in that case also far more is needed.

Eighth, can the dynamics of biological systems be described as maximising a formally specified objective without implying that they have a long-term objective? This question is distinct from the much debated issue contrasting descriptive and prescriptive formulations. Instead, it concerns the temporal course of evolution. Is it progressive or not? Evolutionary biologists are divided on this issue, but Coherent Infomax implies that it can be progressive, provides a conceptual measure of the progress, i.e. as increasing organised complexity, and suggests ways in which neuronal systems contribute to that progress [20]. We can then think of life at the ecological and species levels, not as 'evolved to reproduce', but as 'reproducing to evolve'; i.e. in the direction of the formally specified objective. From that perspective we can think of our own individual efforts as directed, not merely towards survival, but as directed towards whatever organised complexities we choose to create.

Acknowledgements. I thank Andy Clark, Karl Friston, and Jim Kay for insightful discussions of issues addressed in this paper.

References

1. Phillips, W.A., von der Malsburg, C., Singer, W.: Dynamic coordination in brain and mind. In: Von der Malsburg, C., Phillips, W.A., Singer, W. (eds.) Dynamic Coordination in the Brain: From Neurons to Mind. Strüngmann forum report, vol. 5, ch.1, pp. 1–24. MIT Press, Cambridge (2010)
2. Friston, K.J.: The free-energy principle: a unified brain theory? Nat. Rev. Neurosci. 11, 127–138 (2010)
3. Phillips, W.A., Kay, J., Smyth, D.: The discovery of structure by multi-stream networks of local processors with contextual guidance. Network-Comp. Neural Systems 6, 225–246 (1995)
4. Kay, J., Floreano, D., Phillips, W.A.: Contextually guided unsupervised learning using local multivariate binary processors. Neural Networks 11, 117–140 (1998)
5. Kay, J., Phillips, W.A.: Coherent infomax as a computational goal for neural systems. B. Math. Biol. 73, 344–372 (2011), doi:10.1007/s11583-010-9564-x
6. Jaynes, E.T.: Probability Theory: The Logic of Science, Edited by G. Larry Bretthorst. Cambridge University Press, Cambridge (2003)
7. Engel, C., Singer, W.: Better than Conscious? Strüngmann forum report, vol. 1. MIT Press, Cambridge (2008)
8. Yuille, A., Kersten, D.: Vision as Bayesian inference: analysis by synthesis? Trends Cogn. Sci. 10(7), 301–308 (2006)
9. Knill, D.C., Pouget, A.: The Bayesian Brain: the role of uncertainty in neural coding and computation. Trends Neurosci. 27, 712–719
10. Rao, R.P., Ballard, D.H.: Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. Nat. Neurosci. 2(1), 79–87 (1999/2004)
11. Barlow, H.B.: Inductive inference, coding, perception, and language. Perception 3, 123–134 (1974)
12. Barlow, H.B.: Possible principles underlying the transformations of sensory messages. In: Rosenblith, W. (ed.) Sensory Communication, pp. 217–234. MIT Press, Cambridge (1961)
13. Linsker, R.: Perceptual neural organization: some approaches based on network models and information theory. Annu. Rev. Neurosci. 13, 257–281 (1990)
14. Wolpert, D.M., Diedrichsen, J., Flanagan, J.R.: Principles of sensorimotor learning. Nat. Rev. Neurosci. (in press)
15. Friston, K., Kilner, J., Harrison, L.: A free energy principle for the brain. J. Physiol. Paris 100(1-3), 70–87 (2006)
16. Friston, K.J., Stephan, K.E.: Free-energy and the brain. Synthese 159, 417–458 (2007)
17. Phillips, W.A., Singer, W.: In search of common foundations for cortical computation. Behav. Brain Sci. 20, 657–722 (1997)
18. von der Malsburg, C., Phillips, W.A., Singer, W. (eds.): Dynamic coordination in the brain: from neurons to mind. Strüngmann forum report, vol. 5. MIT Press, Cambridge (2010)
19. Phillips, W.A., Silverstein, S.M.: Convergence of biological and psychological perspectives on cognitive coordination in schizophrenia. Behav. Brain Sci. 26, 65–138 (2003)
20. Phillips, W.A.: Self-organized complexity and Coherent Infomax from the viewpoint of Jaynes's probability theory. Information 3(1), 1–15 (2012), doi:10.3390/info3010001
21. Clark, A.: Whatever next? Predictive brains, situated agents, and the future of cognitive science. Behav. Brain Sci. (in press)
22. Spratling, M.W.: Predictive-coding as a model of biased competition in visual attention. Vis. Res. 48, 1391–1408 (2008)

23. Wacongne, C., Changeaux, J.-P., Deheane, S.: A neuronal model of predictive coding accounting for the mismatch negativity. J. Neurosci. 32, 3665–3678 (2012)
24. Szmatháry, E., Maynard Smith, J.: The major evolutionary transitions. Nature 374, 227–232 (1995)
25. Ginter, F., Boberg, J., Jarvinen, J., Salakoski, T.: New techniques for disambiguation in natural language and their application to biological text. J. Mach. Learn. Res. 5, 605–621 (2004)
26. Phillips, W.A., Chapman, K.L.S., Berry, P.D.: Size perception is less context-sensitive in males. Perception 33, 79–86 (2004)
27. Doherty, M.J., Tsuji, H., Phillips, W.A.: The context-sensitivity of visual size perception varies across cultures. Perception 37, 1426–1433 (2008)
28. Uhlhaas, P.J., Phillips, W.A., Mitchell, G., Silverstein, S.M.: Perceptual grouping in disorganized schizophrenia. Psychiatry Research 145, 105–117 (2006)
29. Doherty, M.J., Campbell, N.M., Tsuji, H., Phillips, W.A.: The Ebbinghaus illusion deceives adults but not young children. Developmental Science, 1–8 (2009), doi:10.1111/j.1467-7687.2009.00931.x
30. Fiorillo, C.D.: On the need for a unified and Jaynesian definition of probability and information within neuroscience. Information 3, 175–203 (2012)
31. Phillips, W.A., Craven, B.J.: Interactions between coincident and orthogonal cues to texture boundaries. Percep and Psychophys 62, 1019–1038
32. Feldman, H., Friston, K.J.: Attention, uncertainty, and free-energy. Frontiers in Human Neuroscience 4, 215 (2010)

Neurobiological Computation and Synthetic Intelligence

Craig A. Lindley

Intelligent Sensing and Systems Laboratory, CSIRO ICT Centre, Hobart, Australia
craig.lindley@csiro.au

Abstract. When considering the ongoing challenges faced by cognitivist approaches to artificial intelligence, differences in perspective emerge when the synthesis of intelligence turns to neurobiology for principles and foundations. Cognitivist approaches to the development of engineered systems having properties of autonomy and intelligence are limited in their lack of grounding and emphasis upon linguistically derived models of the nature of intelligence. The alternative of taking inspiration more directly from biological nervous systems can go far beyond twentieth century models of artificial neural networks (ANNs), which greatly oversimplified brain and neural functions. The synthesis of intelligence based upon biological foundations must draw upon and become part of the ongoing rapid expansion of the science of biological intelligence. This includes an exploration of broader conceptions of information processing, including different modalities of information processing in neural and glial substrates. The medium of designed intelligence must also expand to include biological, organic and inorganic molecular systems capable of realizing asynchronous, analog and self-* architectures that digital computers can only simulate.

Keywords: Artificial intelligence, neuroscience, natural computing.

1 Introduction

Alan Turing [27] provided the definitive challenge for research in artificial intelligence (AI) of creating a computer program that could not be distinguished in communication via a remote interface from a human operator. This challenge has had the great advantage of providing a constrained and measurable problem for artificial intelligence, which more generally suffers from being highly unconstrained [11]. That is, AI seeks to make machines more intelligent, which immediately raises questions of what intelligence is and how it might be detected or measured. The focus of the Turing test on textual discursive capability has encouraged a symbolic AI paradigm that emphasizes the definition of formalized linguistic representations and the logic of high level cognitive operations that are involved in verbal and textual discourse. The Turing test meshes well with Newell and Simon's physical symbol system hypothesis [17] that: "A physical symbol system has the necessary and sufficient means for general intelligent action." In this conception, the foundations of discursive intelligence become the foundations of general intelligence.

G. Dodig-Crnkovic and R. Giovagnoli (Eds.): *Computing Nature*, SAPERE 7, pp. 71–85.
DOI: 10.1007/978-3-642-37225-4_4

While this approach may lead to systems that can pass the Turing test within limited contexts, as a general paradigm of intelligence it has severe limitations, as summarized by Lindley [15]. Indeed, following the argument of [15], not only is the Turing test limited to symbolic discursive intelligence, the foundation of Turing's challenge, computing machinery, is a narrow and historically situated understanding of machines that unnecessarily constrains the historical development of synthetic intelligences. In the age of nanotechnology and biotechnology, the distinction between machines and biological organisms breaks down. This suggests that the realization of intelligence by design can shift towards foundations in the design of self-replicating, self-assembling and self-organizing biomolecular elements or analogs capable of generating cognizing systems as larger scale assemblies, analogous to the neurobiological substrate of human cognition. That is, the paradigm of biomolecular engineering implies the construction of human level intelligence (HLI), not from the top-down by the manipulation of symbols, but from the bottom-up by the synthesis of neural architectures starting at the level of molecular engineering.

Here this bottom-up approach will be referred to as synthetic intelligence (SI), by analogy to synthetic biology, characterized by the OpenWetWare group as "A) the design and construction of new biological parts, devices, and systems, and B) the re-design of existing, natural biological systems for useful purposes" (http://syntheticbiology.org/, accessed on 11 September, 2012). SI may be contrasted with symbolic AI (SAI) which is largely based upon top-down analysis of higher level cognitive functions with the aim of deriving abstract symbolic models that can be expressed independently of mechanisms by which representation and logical inference may be automated.

While SI can be motivated by reflection upon the limited success of SAI, it finds a strong and obvious demonstration in human neural systems, of which only functions pertaining to abstraction and symbol manipulation are readily captured by SAI. For SI, abstraction and symbol manipulation are functions that need to be achieved from the bottom-up, in the context of more extensive non-symbolic neural functions such as perception, orientation, control of movement, automation of vital functions, impulsive and goal-oriented behaviour generation, drive arbitration and contextualization, etc.. However, the achievement of SI as a product of engineering is, of course, yet to be demonstrated. Moreover, the pursuit of SI immediately raises the question of how well we understand neurobiology. One of the issues to consider in this is the degree to which our understanding of neurobiology is conditioned by historically situated metaphors and current technologies, just as [15] describes the metaphorical construction of AI in terms of intelligence as computation, and of robotics as the reinvention of the (typically) human form in the media of twentieth century electromechanical engineering and industrial technology.

A central question here is that of the nature and role of information processing in understanding the functions of neurobiology, especially those functions that may appear most relevant to the realization of HLI, and the implications of this for how SI may be achieved. Critical to this is the degree to which information processing is being used metaphorically or literally. A metaphor can be defined as "a figure of speech in which a word or phrase is applied to an object or action that it does not literally denote in order to imply a resemblance"

(http://www.collinsdictionary.com/dictionary/english/metaphor, accessed 11 January 2012). In neurobiology it is often very unclear when descriptions in terms of computation or information processing are metaphorical and when they are literal. The degree of literalness or metaphor depends upon the degree to which specific models of computation capture more or less fundamental aspects of the operation of the brain associated with the realization of HLI. For SI, which seeks to realise designed intelligence, the distinction between the literal and the metaphorical, and what falls within the gap, can be critical to success or failure.

2 Neurobiology: Systems, Signals and Processes

Understanding neural signal transmission and processing in the brain requires understanding at many different spatial scales, including those of ion channels (at a scale around 1 pm), signaling pathways (1 nm), synapses (1 μm), dendritic subunits (10 μm), neurons (100 μm), microcircuits (1 mm), neural networks (1 cm), subsystems (10 cm) and the whole nervous system (1 m) [26]. For understanding intelligence, a key question is: at what levels of this structural and functional hierarchy are information processes critical to intelligence carried out? The answer is not obvious, since a top-down, cognitive account of intelligence might identify intellectual functional capabilities that could be realized by a high level implementation providing a layer of abstraction in the place of lowers levels of the hierarchy of neural signal processing; this is exactly what a high-level, rule-based model of problem-solving does. For bottom-up, neurobiological accounts, the question amounts to where to identify the bottom (are abstract neurons enough, or do we need to model synapses, ion channels, molecular interactions, etc.). Conversely, what aspects of the physiology, structure and operation of the brain that are not captured by the concept of information processing may nevertheless be critical to achieving HLI? Or turning the last question around, which concept(s) of information processing are critical to achieving HLI? Or which explanations of neurobiological function need to be expressed in which language of information processing and/or computation to explicate the critical foundations of HLI?

Symbolic AI has focussed upon behaviour, what may be inferred from behaviour regarding functional types and capacities, and the derivation from language constructs of more formalized models (e.g. taxonomies, propositions, rules, etc) of linguistic forms, their syntax and semantics. Subsymbolic AI has focussed upon simplified models of neurons and neural networks characterized by different learning rules and topologies. Since HLI has yet been approached by any kind of AI system, it must be asked if these approaches are adequate, or whether comparison with biological brains and nervous systems can reveal structures, functions, processes or principles that have not yet been used in AI that may nevertheless be critical for the achievement of artificial or synthetic HLI. Since SI is concerned with the bottom-up creation of intelligence, the discussion here will focus upon the lower layers of the hierarchy, i.e. the levels of simple neural interconnections and below (ignoring larger scale circuits, network topologies and subsystems).

The most ubiquitous image of the brain is that of a kind of wet computer with circuitry consisting of a vast network of interconnected neurons transmitting electrical signals among themselves, with each neuron summing weighted inputs and issuing an output signal if the sum of inputs exceeds a threshold (the integrate-and-fire model of neurons). This view is a simple embodiment of the neuronal doctrine [28], which places neurons and their synaptic interconnections at the centre of brain and nervous system functionality. This is really tautologous in the case of the nervous system, since the nervous system is the network of neurons and their interconnections. Less obviously, however, it may be asked what contributions the neural system makes to intelligence, compared with the contributions of physiological information processing that is not carried out by neurons. This requires a consideration of signal processing by neural and non-neural systems.

A single bipolar neuron cell consists of a cell body from which there extend dendritic trees surrounding the cell body, and an elongated axon that also leads to a branching end structure. Dendrites accept inputs from other neurons in the form of neurotransmitters, via synapses that often occur on small projections referred to as dendritic spines. Any given input can have an additive or subtractive effect upon the summation of inputs at the neuron body, with different inputs having different strengths. When a sufficient balance of additive over subtractive inputs is received, the cell body accumulates enough of a potential difference between the inner and outer surfaces of its surrounding membrane for ion channels embedded in the membrane to open, leading to the movement of ions between the inside and the outside of the cell. This movement of ions cascades along the cell axon as an action potential, a voltage spike providing a signal that is transmitted via the terminal axonal branches to the dendrites of other neurons. After the passage of such an electrochemical pulse, following a brief refractory period during which no further action potentials can be generated, the ionic balance across the neuron cell membrane returns to the rest potential. The details of this process are covered in many neuroscience texts (e.g. [2], [25], [13]).

Significant contributions to action potential generation are made by astrocytes, oligodendrocytes and Schwann cells. Astrocytes have a role in regulating the concentration of extracellular K+. Oligodendrocytes (in the central nervous system) and Schwann cells (in the peripheral nervous system) provide insulating myelin wrapping to the axons of some neurons. Where the myelin wrapping occurs, there is no ion-containing extracellular fluid on the outer surface of the axon, and no electrical potential can travel along these segments. At the gaps between myelinated sections, the nodes of Ranvier, there are high densities of voltage-gated ion channels, so that when an action potential creates a sideways spread of Na+, the action potential jumps from one node of Ranvier to the next, creating a highly accelerated propagation of action potentials and consequently much faster speed of signal transmission than in unmyelinated axons. The speed of transmission of an action potential in the fastest, myelinated, cells of the peripheral nervous system is about 150 ms^{-1}, a speed that is two million times slower than the transmission of an electric signal along a wire or a pulse of light along an optical fibre. The metaphor of neural transmission as an electrical signal is highly misleading in this respect; an action potential is measured across and propagates along the neuron cell membrane due to a cascading flow of charged particles, a process of electrochemical diffusion.

The primary signal transmission connections between neurons occur at synapses. A synapse consists of a synaptic terminal on the presynaptic side, a synaptic cleft, which is a gap of ~20 nm width, and a postsynaptic membrane on the receiving side. A subset of a wide range of possible neurotransmitters is issued from synaptic terminals of an activated neuron. The neurotransmitters move across the synaptic cleft to receptors on dendrites on the post-synaptic side. Hence action potentials as such are not directly transmitted from one neuron to another (in most cases), the inter-neuron connection being mediated by neurotransmitters passing across the synaptic cleft and into receptor proteins in the post-synaptic membrane. Hence chemical synapses provide electrical and physical isolation between interconnected neurons. However, some neurons do have electrical synaptic interconnections, gap junctions that are channels allowing the passage of ions for direct propagation of electrical signals, but also allowing the passage of larger molecules, thereby creating metabolic coupling in addition to electrochemical coupling between neurons [28]. There are also anterograde connections from post-synaptic neurons to presynaptic neurons, typically realised by gaseous neurotransmitters such as nitric oxide.

Neurotransmitters, as well as hormones secreted by neurons, are not limited to local effects, but can diffuse more widely through extracellular space, thereby bypassing the dendritic/axonal network. These broad diffusion processes can be referred to as Volume Transmission (VT) processes. Processing within the dendritic/axonal network can be referred to as wiring transmission (WT) [28]. WT is rapid (from several microseconds to a few seconds), highly localised, signals pass between two cells, and the effects are phasic (i.e. event-related). VT is slow (from seconds to minutes/hours), global, has one-to-many signals, and the effects are tonic (extended over numerous events). VT may be the consequence of synapse leakage, open synapses, ectopic release (i.e. neurotransmitters released from the surface away from synapses), etc..

WT networks formed by neurons and their interconnections are the main information processing structure of the brain posited by the neuronal doctrine. Hence it is this interconnected structure of weighted links, integrators and action potential generators that has been regarded as the primary system of information processing and computation in the brain. This is also the model that has been adopted by simple artificial neural network models derived from the neuronal doctrine during the latter half of the twentieth century. The neuronal doctrine accommodates increasing sophistication in the understanding of how biological neurons and their networks function. Synapses are highly plastic, their plasticity being a major mechanism of learning, with synapses between co-active neurons being strengthened while synapses between neurons that are rarely active at the same time deteriorate. Most of the adult human brain does not undergo any significant new neuron creation (with the notable exception of the olfactory epithelium), but dendrites and dendritic connections, ion channels, dendritic spines and synapses undergo continuous ongoing changes. By these mechanisms, cortical neuron populations of a sufficient size appear to be capable of learning any function of an arbitrary number of dimensions (e.g. [7]).

Information processing by the propagation of action potentials through neuron networks, and the integrate-and-fire operation of individual neurons represents one level of computation in the brain. Processing of information within the dendritic trees of single neurons has recently been proposed to represent local forms of computation within the dendritic structure, together with back propagation of spikes from the soma via dendrites, spike generation in dendritic spines and shafts, and bistable dynamics [16]. Computations conducted within dendral structures may include simple arithmetic, from simple to complex logic functions, filtering, and even integrate and fire functions within dendral substructures, creating a two-layered 'neuron' model (similar to simple classic ANNs) within a single neuron. The extended structure of dendritic trees means that different spatiotemporal input patterns can have different effects on neuron firing, allowing for computation of directional selectivity, in retinal and audio processing [16]. Moreover, Rall and Shepherd [19] proposed that two neuronal populations (excitatory and inhibitory) could communicate via direct synapses between their dendrites, without involving axonal propagation (see also [20]). While these forms of computation may be observed within dendritic structures, as London and Hausser [16] note, the key question is the extent to which the brain takes advantage of these building blocks to perform computations. For SI the question is that of the extent to which these mechanisms may contribute to HLI. Hameroff [12] even suggests that dendritic cross-connections provide the foundations of consciousness.

The changes in ion channels, dendritic and axonal tree growth and interconnections, and synaptic processes in response to network activity can be seen as another level of computation, and one more fundamentally associated with neural plasticity. It is also possible to model the internal processes of cells in terms of information processing and/or computation (e.g. [26]). This includes the effects of neurotransmitter reception, continuous metabolic processes, and interactions between these two. Hence computation/information processing occurs at the intra-cellular level, as well as at the WT and VT levels.

This picture of the neuronal doctrine makes information processing, or computation, within the brain complex and multi-levelled. In general there about 500 different types of human neurons. There are about 100 billion neurons in a single brain, each of which is connected to 1,000-10,000 others with over 200,000 km of axons [18]. Hence the WT network is highly complex, even without considering detailed mechanisms of intercellular communication, dendritic processing, synaptic processing, plasticity and intra-cellular processes.

However, neurons only constitute about 10% of brain cells. The rest consist of glial cells [28], of which 80% are astrocytes. For most of the time since the discovery of glial cells in the late 19th century they have been regarded as secondary support cells for neurons, e.g. providing nutrients and mopping up excess neurotransmitters. However, research over the last couple of decades has radically revised this understanding. It is now known that astroglia are the stems cells from which neurons differentiate. Those that remain as astrocytes form networks connected via gap junction bridges that provide intercellular communication, providing transfer paths for ions, metabolic factors and second messengers throughout the central nervous system (CNS). Astrocytes also engage in long distance communication by calcium wave propagation initiated by stimulation of neurotransmitter receptors in the astroglial cell membrane [28]. Astroglia appear to express all known forms of neurotransmitters, which can influence neuron activity, and they possess numerous ion channels that can be activated by extracellular and intracellular activity, such as the activity of neighbouring neurons [28]. Hence neurons and astroglia appear to form parallel and intercommunicating systems of signal transmission and processing. Glial cells also determine the differentiation, microarchitecture, synaptogenesis, and death of neurons and neural structures. Verkhratsky and Butt [28] hypothesize that neuronal networks are specialised for fast communication (i.e. metaphorically, they provide a kind of internet within the CNS), while astroglia provide the most substantial information processing, integration and storage functions of the brain. Evidence for the significance of glia is found in their dramatic increase, both in absolute numbers and relative to the numbers of neurons, on a phylogenetic scale, reaching the greatest complexity in the human brain [28].

One further information processing system within neurobiological systems that will be mentioned in this paper is the system of hormones that also interacts with the processes described above. Hormones are chemicals secreted by specific groups of cells that are carried by the bloodstream to other parts of the body where they act on other cells to produce specific physiological effects [2]. Neurosecretory, or neuroendocrine, cells in the hypothalamus are almost the same as neurons, except that they do not release neurotransmitters, but instead they secrete hormones into the blood stream [2]. The effects of hormones on the body include reproductive development and rhythms, water and salt balance, growth, the secretion of other hormones, metabolic rate, emotional arousal, inflammation reactions, digestion and appetite control [2]. Hormones constitute a VT system in the terms used by Verkhratsky and Butt [28]. Hormones act gradually, change the intensity or probability of behaviours, are influenced (in type and quantity released) by environmental factors, have a many-to-many relationship with cells, organs and behaviours, are secreted in small amounts and released in bursts, may vary rhythmically in levels, may be mutually interacting, and are graded in strength (unlike the digital nature of neuronal action potentials [2]). Hormones, neural systems, behaviours and their consequences, are highly interactive and integrated. Hence an understanding of the processes and consequences of the nervous system, including the achievement of HLI, requires understanding the parallel signalling and information processing system mediated by hormones.

The specific role of neural processing can be highlighted by comparing i) the most complex organisms lacking a neural system with ii) the simplest organisms having a neural system. i) are the Porifera, or sponges, while the most basic examples of ii) are members of the phylum Cnidaria, which includes sea anemones, corals, jellyfish and hydrozoa.

Poriferans are sessile, suspension-feeding, multicellular animals lacking true tissues and having cells that are all capable of changing form and function [4]. Feeding is achieved with the aid of flagellate cells that circulate water through a system of water canals. The cells of a sponge are continuously mobile. Flagellations used to circulate water are not coordinated at a detailed level, although overall activity levels of a sponge can vary. Larger scale movements include closing their oscula (large water vents) or ostia (smaller water input pores), constricting water canals, and reversing the direction of water flow. These changes can be triggered by water-born particle size, or by direct tactile stimulation. In some cases the whole sponge can slowly change its location (e.g. at a rate of 4 mm per day). There is no evidence that Poriferans have neurons, action potentials or specialized sense organs. Rather, the spread of behaviour appears to be by direct mechanical interaction among adjacent cells, and possibly by the diffusion of chemical messengers. Larger scale movements are achieved by contractile myocytes acting independently but otherwise forming analogs of muscle tissues. The conduction of contractions is typically less than 0.04 cm sec^{-1} and is always unpolarized and diffuse. Conduction speeds in hexactinellids can achieve 0.22 cm sec^{-1}, which is considered by Lawn et al [14] to be too slow for a nervous system, but too fast for simple chemical-like diffusion, raising the possibility of the existence of a primitive non-neural enhanced messaging system. The Poriferan A. queenslandica has been shown to have a nearly complete complement of synaptic genes, without the coordination of regulatory circuitries required to express these genes in a regulated neuronal form [5]. Hence the development of neural systems from more primitive Poriferan ancestors may have evolved from the development of mechanisms for localizing gene expression in the specialized form of neurons.

Organisms manifesting the simplest forms of nervous system such as the Cnidaria, show a significant enhancement of speed, complexity and coordination in sensory discrimination and behaviour generation, compared with the Porifera. Taking anemones (order Actiniaria) as a more detailed example of Cnidaria, movements are achieved by muscle-like tissues acting in longitudinal, axial and radial directions to create retraction, elongation, and contraction (at the mouth and base) [4]. Movements include those of the whole body, with some species locomoting via pedal disc manipulations, somersaults or inch-worm type movements, or swimming via flexing the body or thrashing tentacles. Sensors include mechanoreceptors in the form of cilia from individual cells, which also appear to function as chemoreceptors. Cnidaria show general sensitivity to light. For some polyps this is not associated with any specific sense neurons, most likely being mediated by epidermal neurons, while more sophisticated Cnidaria have radially distributed ocelli (simple eyes). Cnideria do not have centralized nervous systems or brains, but have diffuse, decentralized nerve nets that include sensory neurons, networks of intermediate neurons with some clustered ganglia providing coordination functions, and motor neurons for activating movements [22]. Cnidarian nerve junctions are mediated by synapses, while those responsible for triggering rapid movements needed for swimming are mediated by gap junctions forming direct electrochemical connections. Cnidarian sensory neurons are the most primitive among animals, and most of their neurons and synapses are unipolar and hence bidirectional. This means that a sufficiently strong stimulus will result in the spread of pulses in all directions from the point of stimulation. Motile Cnidarians have more sophisticated neural systems than sessile species, some having a separate and faster network of bipolar neurons. In addition to a neural system, Cnidarians also have a slow and highly diffuse impulse conduction system, involving epidermal cells and muscle elements, similar to that of sponges.

The comparison of Cnidaria with Porifera highlights major benefits arising from a neural system, including at the most primitive level of enhanced sensory capabilities and capacities for rapid motion coordinated across the large scale structure of an organism's body. Increasing levels of complexity within neural systems introduce localized concentrations of neurons in the form initially of nerve ganglia, ultimately leading to the evolution of centralized vertebrate brains. Sensory neurons for senses of all types provide inputs to nervous systems in the form of voltage spike trains having frequencies proportional to the intensity of the stimuli. Motor neurons similarly active muscle cells via spike trains where the density, frequency and persistence of spikes determines the degree and duration of muscle activation. A critical and as yet unanswered question is that of which subset of the forms of computation that biological neural networks realize are sufficient for HLI, and does HLI have any critical dependencies upon other levels of biological information processing? Once this question is definitively answered, implementation of a homologous system in any alternative medium may be sufficient for achieving HLI.

3 Neurophysiology and Computer Systems, Essential Differences

A fundamental question in the quest for synthetic HLI is that of which levels of abstraction or description represent the lowest level necessary for the realisation of HLI. Human levels of intelligence are poorly defined and poorly constrained. Cognitive science and cognitive psychology have made some progress in the top-down decomposition of human intelligence into functional parts and facilitating capacities. Knowledge of how cognitive constructs map onto brain structures and processes at lower physiological levels are being increasingly provided by correlation studies with neuroimaging, lesion studies, etc.. But since there are as yet no convincing demonstrations of synthetic or artificial HLI, it is not yet clear where the lower limits of necessary functional and structural decomposition are. There is also a conceptual question of the degree to which constructs such as computation, information and information processing, and which particular understandings of these constructs, are helpful (or not) in creating a working model of the relationship between physiological processes at different levels of abstraction/description and the achievement of HLI. Answers to this question provide foundations for considering which technologies may provide suitable media for the achievement of the required forms of computation and information processing. However, specific differences between biological brains and computational machines as we know them may make machines as such incapable of achieving the intelligence demonstrated by biological brains, irrespectively of issues of abstract computation and information models.

Potter [18] presents a number of differences between natural intelligence (NI) and AI, suggesting features that could, at least in principle, potentially make AI more brain-like. These include:

- brains don't have a CPU, they are highly distributed; NI uses lots of parallelism
- biological memory mechanisms are not physically separable from processing mechanisms
- biological memories are dynamic and continually reshaped by recall
- the brain is asynchronous and continuous; resulting phase differences can encode information. As noted by Crnkovic [6], the asynchrony of brain processing means that it does not conform to Turing computation.
- brains do not separate hardware from software; i.e. computation and information processing are not abstracted from the physical level, and the physical level is continuously changing (e.g. in mechanisms of plasticity noted above)
- NI thrives on feedback and circular causality. The nervous system is full of feedback at all levels, including the body and the environment in which it lives; it benefits in a quantifiable way from being embodied and situated.
- NI uses lots of sensors
- NI uses lots of cellular diversity
- delays are part of the computation. The brain computes with timing, not Boolean logic.

Further differences may also be noted, including:

- the brain is analog, where computers are digital. The digitisation of atemporal quantities leads to quantisation errors, and the quantisation of time leads both to potential quantisation errors and aliasing (the appearance of high frequency content in the form of illusory low level frequency components) although it is unclear how critical these errors are in functions underlying the achievement of HLI.
- neural systems are coextensive with the human body. This leads to layered and partially hierarchical control, achieved to a degree in some AI architectures (e.g. [3]).
- biological power distribution is decentralised and coextensive with information processing; hence human metabolic processes are implicated in information processing
- brains and nervous systems are intimately integrated with other bodily systems and processes. It may be reasonable to suggest that more abstract cognitive functions (e.g. abstract problem solving, mathematics) are understandable without considering parallel or underlying physiological processes. But even the most abstract brain operations are in practice constrained by factors in their physiological substrate (e.g. successful high level reasoning requires energy and sleep).
- much of the organization of the brain is topographic, from sensory processing to the contralateral organisation of the cerebral hemispheres
- physical locations can matter. A good example of this is the use of differential timing of the arrival of aural information via separate dendrites as a cue for sound localisation (see [16])
- NI systems are built from the bottom up in processes that are self-assembling, self-organizing, and adaptively self-maintaining (characterised by Crnkovic [6] as self-* processes), based upon a ubiquitous (genetic) instruction set that is expressed in ways that vary according to highly local conditions and their recursively embedded contexts
- the foundations of NI have not been designed, but have evolved

The last two points are critical. There is currently no demonstration proof of the achievement of HLI in a way in which its mechanisms and contents are fully comprehensible within human consciousness. However, to achieve an SI that is capable of HLI should not require directly building a fully capable system; it is only in more primitive biological animals (e.g. insects) that specific neurons and neural connections are genetically pre-programmed. Rather, as in the human case, detailed neural architecture (i.e. at the level of differentiating individual neurons and growing their interconnections) can be based upon initiating processes of self-assembly and self-organisation that can create a sufficiently complex microstructure to achieve an adaptive, learning and growing nascent SI. The nascent SI must be capable of maturing through self-organisation in interaction with its environment to full HLI and beyond, just as in biological HLI.

4 Neurophysiological Processing as Information Processing

In considering the nature of the brain as an information processing system, it is necessary to be clear about what kind of information processing system it is, and according to what understandings of the term information. A widely used formal definition of information was first formulated by Shannon [23]. Shannon's concept of information can be summarized in the following way: if there are n possible messages, then n is a measure of the information produced by the selection of one message from the set, when all messages are equally likely. That information can be expressed by $\log_2 n$. This represents a number to the base 2 which can be represented by a sequence of bits (binary digits) of length $\log_2 n$, where any specific bit pattern of length $\log_2 n$ can represent a particular message among the set of n possible messages.

To contextualize this characterization of information, Shannon and Weaver [24] describe three levels of communication problems: "A. How accurately can the symbols of communication be transmitted? B. How precisely do the transmitted symbols convey the desired meaning? C. How effectively does the received meaning affect conduct in the desired way?" The mathematical theory of communication is concerned with A. This conception of information has been used in many analyses of neural system function, providing methods of measuring probability distributions, supporting analysis of information bottlenecks, and providing a view of cortical systems as systems that maximize information [1]. Information maximization includes maximizing the richness of representations, heuristic identification of underlying causes of an input, to provide economies of space, weight and energy, and as a reasonable heuristic for describing models [1]. Potential disadvantages of the use of mathematical information theory include the need for vast amounts of data for generating reliable probability distributions, the need for independent sources of the usefulness of an encoding scheme, the uncertain nature of neural encoding of information, and the assumed stationarity of probability distributions known to an information receiver [1]. Nevertheless, it has been suggested that the overall goal of the brain and individual neurons is to minimise uncertainty, which corresponds with the maximization of information. The 'neurocentric' approach of Fiorillo ([8], [9]) proposes that the activity of individual neurons can be fully described in Bayesian terms grounded in information theory, where a single neuron integrates information from molecular sensors to reduce uncertainty about the state of its world. In this case, the state of the world is the local environment of the neuron, where information input to the neuron is a property of biophysical mechanisms from the level of single molecules and up, rather than being inferred by a scientific observer from external properties of the environment. Hence the computational goal of the nervous system is the minimization of uncertainty (of the brain about its world) exclusively based upon the information and mechanics of the system, a view closely related to Friston's [10] theory of free energy minimization but with an internal rather than external view of information sources and their resulting probabilities.

Fiorillo [8] emphasizes that the neurocentric approach uses probabilities only to describe the biophysical information of a neuron. "There is no physical step that must occur within the nervous system to "calculate" probabilities from information. Probabilities are a quantitative property of information in much the same way that mass is a quantitative property of matter. Likewise, information is an intrinsic property of matter and energy. Information follows the rules of physics, and Bayesian principles allow us to quantify information using probabilities." Crnkovic [6] goes further than this, proposing that "Information may be considered the most fundamental physical structure. Info-computationalist naturalism understands the dynamical interaction of informational structures as computational processes." This view goes much further than Shannon's view of information, which is essentially an epistemological one, to one that equates the epistemological with the ontological. A further development of this view is morphological computation, where biological computation is conceived as a "computational mechanism based on natural physical objects as hardware which at the same time acts as software or a program governing the behavior of a computational system" [6]. Morphological computation captures many of the essential differences between biological brains and semiconductor-based computers noted earlier in this paper, since physically interacting objects realize computations that are naturally analog, continuous, parallel, built from the bottom up, and topographic by virtue of the spatial embeddedness of causal interactions. Semiconductor-based computers achieve this in their hardware to realize an abstract layer of information processing that does not achieve morphological computing in its software, since the information model of computer software is based upon abstract and non-situated logical and arithmetic operations (non-situated in the sense that the meaning of strings of 1s and 0s and their manipulations is established by human readings and is not in general dependent upon their physical connections to the context, which are primarily limited to transformation into noisy thermal energy).

5 Conclusions: Implications for Synthetic Intelligence

All of the hierarchical levels of the human nervous system have been simulated in detail using digital computer technology [26]. However, both cognitive and neural simulations have had limited success to date and certainly fall far short of HLI.

Lindley [15] has argued that cognitive approaches in particular do not appear to be promising in isolation. Here it may be added that an essential problem for cognitivist AI is an implicit Cartesian dualism, where AI has focussed upon modelling the mind, and the symbol grounding problem, a central problem for AI, is the computationalist's version of the problem for dualists of the nature of the 'mechanism' by which the mind and the body interact. In the spirit of Rorty [21], it is possible to shift the emphasis away from the nature of mind towards the physiological foundations of the generation and use of mentalist language. The focus then shifts to the physiological foundations of human intelligence, the simulation of which has also not yet demonstrated anything remotely approaching HLI. However, the physiological project has two major advantages over the neo-Cartesian cognitivist project. Firstly, as considered in this paper, neural computation includes many broad frontiers of ongoing knowledge development, including cellular, sub-cellular and molecular processes, the role of dendritic computation, the role of astroglia, and the embedded interdependencies between brain systems, bodies and their contexts. Simply put, we understand biological intelligence so incompletely that it provides an ongoing wealthy source of methods, principles and foundations yet to be comprehensively understood, let alone transferred into the design of intelligence as an engineered artifact.

The second great advantage of the physiological project as an engineering project is that it is no longer limited to twentieth century engineering media. It is possible to apply molecular regulation, transgenic and viral techniques to selectively modify neurons and neural populations to generate "designer dendrites" [16] and other neural structures having specific computational properties and topologies. It is also possible to explore the creation of brain system analogs in different chemistries, including organic and inorganic systems different from 'wild' neural biochemistry. These systems can implement the analog, asynchronous and necessarily self-* characteristics of biological nervous systems in ways that are not possible with digital simulations. In implementing the morphological computing foundations of biological intelligence, these systems can be realizations of synthetic intelligence, rather than simulations or metaphors.

This is not an argument for any specific technique or a proposal of any particular solution to the challenge of engineering HLI. Rather, it is a call for reframing the AI quest, for it to escape the confines of 20th century technological metaphors and seek new avenues of enquiry derived from contemporary neuroscience and biochemistry. This may involve the implementation of neurological or biological signal processing, information processing and computation models using digital information processing and computing technologies, as practiced in computational neuroscience, computational cell biology, etc.. It may also embrace the more direct exploration of the molecular and cellular foundations of HLI using the methods of cell biology, genetic modification, and molecular synthesis. While these techniques are frontiers of rapid ongoing research in their own right, they represent a radical alternative for the project of artificial intelligence, an alternative perhaps better called synthetic intelligence for its proximity to synthetic biology and the implication of a bottom-up synthetic process. The proof of any approach to engineering HLI is in the demonstration of HLI by an engineered system. The failure of SAI demonstrates that SAI as traditionally conceived does not work. It is reasonable to assume that this follows from its choice of the wrong levels of symbolic abstraction as the foundation of intelligence, and the resulting ineffectiveness of the models of symbol processing applied to the manipulation of symbols of this kind. Neural processing models are widely understood as a viable alternative to SAI, but ANN approaches have also reached the area of diminishing returns in terms of knowledge generated for research effort expended. The only way forward is a shift to the rapidly expanding tools and theories of biomolecular science and engineering, those tools and theories that are actively advancing our understanding of the only successful model of HLI that we have, i.e. human intelligence.

References

1. Baddeley, R.: Introductory Information Theory and the Brain. In: Baddeley, R., Hancock, P., Földiák, P. (eds.) Information Theory and the Brain. Cambridge University Press (2000)
2. Breedlove, S.M., Watson, N.V., Rosenzweig, M.R.: Biological Psychology, 6th edn. Sinauer Associates, Inc. Publishers, Sunderland (2010)
3. Brooks, R.A.: Cambrian Intelligence. MIT Press (1999)
4. Brusca, R.C., Brusca, G.L.: Invertebrates, 2nd edn. Sinauer Associates (2003)
5. Conaco, C., Bassett, D.S., Zhou, H., Arcila, M.L., Degnan, S.M., Degnan, B.M., Kosik, K.S.: Functionalization of a Protosynaptic Gene Expression Network. In: Proceedings of the National Academy of Sciences, June 20 (2012)
6. Crnkovic, G.D.: Info-computationalism and Morphological Informational Structure. In: 1st Annual Conference on Integral Biomathics, August 29-31. Stirling University, Scotland (2011)
7. Eliasmith, C., Anderson, C.H.: Neural Engineering: Computation, representation and dynamics in neurobiological systems. MIT Press (2003)
8. Fiorillo, C.D.: A new approach to the information in neural systems. In: 1st Annual Conference on Integral Biomathics, August 29-31. Stirling University, Scotland (2011)
9. Fiorillo, C.D.: Towards a general theory of neural computation based on prediction by single neurons. PLoS ONE 3, e3298 (2008)
10. Friston, K.: The free energy principle: a unified brain theory? Nat. Rev. Neurosci. 11, 127–138 (2010)
11. Granger, R.: Engines of the Brain: The Computational Instruction Set of Human Cognition. AI Magazine 27, 15–32 (2006)
12. Hameroff, S.: The "conscious pilot"—dendritic synchrony moves through the brain to mediate consciousness. Journal of Biological Physics 36(1) (2009), doi:10.1007/s10867-009-9148-x
13. Kandel, E.R., Schwartz, J.H., Jessell, T.M.: Principles of Neural Science, 4th edn. McGraw-Hill, New York (2000)
14. Lawn, I.D., Mackie, G.O., Silver, G.: Conduction system in a sponge. Science 211, 1169–1171 (1981)
15. Lindley, C.A.: Synthetic Intelligence: Beyond A.I. and Robotics. In: 1st Annual Conference on Integral Biomathics, August 29-31. Stirling University, Scotland (2011)
16. London, M., Hausser, M.: Dendritic computation, Annu. Rev. Neurosci. 28, 503–532 (2005)

17. Newell, A., Simon, H.A.: Computer Science as Empirical Inquiry: Symbols and Search. CACM 19, 113–126 (1975)
18. Potter, S.M.: What Can AI Get from Neuroscience? In: Lungarella, M., Iida, F., Bongard, J.C., Pfeifer, R. (eds.) 50 Years of AI, Festschrift. LNCS (LNAI), vol. 4850, pp. 174–185. Springer, Heidelberg (2007)
19. Rall, W., Shepherd, G.M.: Theoretical reconstruction of field potentials and dendrodendritic synaptic interactions in olfactory bulb. J. Neurophysiology 5, 884–889 (1968)
20. Rinzel, J.: Distinctive Roles for Dendrites in Neuronal Computation. SIAM News 40(2) (March 2007)
21. Rorty, R.: Philosophy and the Mirror of Nature. Princeton University Press, Princeton (1979)
22. Ruppert, E.E., Fox, R.S., Barnes, R.D.: Invertebrate Zoology, 7th edn., pp. 111–124. Brooks/Cole (2004)
23. Shannon, C.E.: A Mathematical Theory of Communication. Bell System Technical Journal 27, 379–423, 623–656 (1948)
24. Shannon, C.E., Weaver, W.: The Mathematical Theory of Communication. The University of Illinois Press, Urbana (1949)
25. Squire, L.R., Bloom, F.E., Spitzer, N.C., du Lac, S., Ghosh, A., Berg, D.: Fundamental Neuroscience, 3rd edn. Academic Press (2008)
26. Sterratt, D., Graham, B., Gillies, A., Willshaw, D.: Principles of Computational Modelling in Neuroscience. Cambridge University Press (2011)
27. Turing, A.M.: Computing Machinery and Intelligence. Mind (59), 433–460 (1950)
28. Verkhratsky, A., Butt, A.: Glial Neurobiology: A Textbook. John Wiley and Sons Ltd. (2007)

A Behavioural Foundation for Natural Computing and a Programmability Test*

Hector Zenil

Institut d'Histoire et de Philosophie des Sciences et des Techniques
(Paris 1 Sorbonne-Panthéon/ENS Ulm/CNRS), Paris, France
hector.zenil-chavez@malix.univ-paris1.fr

Abstract. What does it mean to claim that a physical or natural system computes? One answer, endorsed here, is that computing is about programming a system to behave in different ways. This paper offers an account of what it means for a physical system to compute based on this notion. It proposes a behavioural characterisation of computing in terms of a measure of programmability, which reflects a system's ability to react to external stimuli. The proposed measure of programmability is useful for classifying computers in terms of the apparent algorithmic complexity of their evolution in time. I make some specific proposals in this connection and discuss this approach in the context of other behavioural approaches, notably Turing's test of machine intelligence. I also anticipate possible objections and consider the applicability of these proposals to the task of relating abstract computation to nature-like computation.

Keywords: Turing test, computing, nature-like computation, dynamic behaviour, algorithmic information theory, computationalism.

Faced with the question of computation, it may be tempting to go along with the formal mathematical position and simply invoke Turing's model. This paper doesn't need to do this, though its author couldn't be more wholehearted in granting the beauty and generality of the universal Turing machine model, which, it will be argued, is also a natural foundation for unconventional (and natural) computation.

To date the study of the limits of computation has succeeded in offering us great insight into this question. The borderline between decidability and undecidability has provided an essential intuition in our search for a better understanding of computation. One can, however, wonder just how much can be expected from such an approach, and whether other, alternative approaches to understanding computation may complement the knowledge and intuition it affords, especially in modern uses of the concept of computation, where objects or events are seen as computations in the context of physics.

One such approach involves not the study of systems lying "beyond" the uncomputable limit (the "Turing limit"), but rather systems at the farthest reaches of the computable, in other words the study of the minimum requirements for universal computation. How easy or complicated is it to assemble a machine that is Turing universal?

* Invited Talk at the *Symposium on Natural/Unconventional Computing and its Philosophical Significance*, *AISB/IACAP Alan Turing World Congress 2012*.

G. Dodig-Crnkovic and R. Giovagnoli (Eds.): *Computing Nature*, SAPERE 7, pp. 87–113.
DOI: 10.1007/978-3-642-37225-4_5

This minimalistic bottom-up approach is epitomised by Wolfram's programme [38] in its quest to study simple programs, a programme initiated by Minsky [23] and to which several authors have contributed (see [39] for an excellent survey). The underlying question is how pervasive and ubiquitous the computational property of universality is in computational and natural systems. From the various results concerning small universal computing systems, we now know that generating universality takes very little, indeed that it seems to be the case that it is more difficult to design a non-trivial non-Turing-complete computer language than a Turing-complete one. Thus it seems natural to believe that computation and universality are not exclusive to digital computers.

This paper is organised as follows. In Section 1, the foundations of natural computation are discussed, taking as a starting point Turing's case—argued in relation to digital computation—for the disembodied essence of natural computation. In Section 2, the behavioural approach to natural computation will be introduced, based on notions of algorithmic complexity, and with an analogy drawn between it and Turing's pragmatic approach to machine intelligence. In Section 3, a characterisation and taxonomy of computation (and of computers) based on the compression-based approximation of a system's algorithmic complexity is advanced and, finally, in Section 4 possible objections are analysed, also in light of the way in which they can be transferred between Turing's test and the definition of nature-like computation adopted herein.

1 A Classical Foundation for Unconventional Computation

A compiler written between computational systems, hence a mapping between symbols and states, is the usual way of proving in a technical fashion that one system is equivalent to another in computational power (hence that *it computes*). A legitimate question that arises is whether we need this technical apparatus to define computation. The problem can be phrased in the words of M. Conrad [7] *In the real world, little if anything is known of the primitive operations or symbols of a system.*

One strong criticism of the idea that natural objects (including the universe) compute is that the question and answer become meaningless, as it is hard to see how any physical system would not be computational [28,31]. One concept that Turing did not advance (although he suggested taking into account the percentage of people acknowledging the success or failure of his machine intelligence test [37]), but that is very much in the spirit of another of his seminal contributions (the relativisation of computation, in his notion of degrees of computation [36]), is a metric of intelligence, one where passing or failing is beside the point, but which tells us how close or far we are from intelligent behaviour.

This paper advances a metric of approximative, asymptotic and limit behaviour, not for intelligence, but for computation, one that identifies objects to which some degree of computation can be assigned on the basis of how they behave, and particularly on the basis of whether they can be programmed. It thereby places programmability at the centre of our definition of computation and so avoids representationalism.

1.1 A Behavioural Approach to Computation

Among the most important of Turing's contributions to AI was his test of machine intelligence [37], devised as a response to the question of whether computers could think. The Turing test is a pragmatic behavioural approach to the problem of assigning intelligence to objects (see Fig. 1). In the spirit of Turing, one may ask whether objects other than electronic computers compute, in particular natural objects and natural processes. This question ultimately leads to the more general question of whether the universe itself computes (also known as "pancomputationalism"), and if so how. Some speculative answers have been given, but in this presentation we take a more pragmatic and behavioural approach to the question, in the spirit of Turing's approach to intelligence.

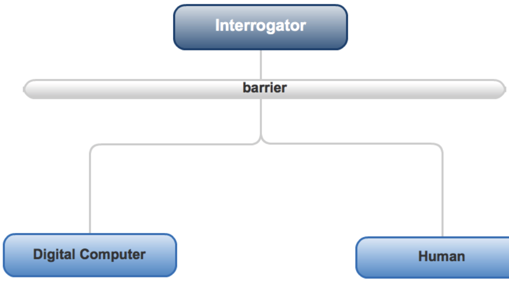


Fig. 1. The basic elements of Turing's test of intelligence

When Alan Turing was thinking about AI he believed "that in about fifty years' time it will be possible to programme computers, with a storage capacity of about 10^9, to make them play the imitation game so well that an average interrogator will not have more than a 70 percent chance of making the right identification after five minutes of questioning. . . . I believe that at the end of the century the use of words and general educated opinion will have altered so much that one will be able to speak of machines thinking without expecting to be contradicted."

Most would agree that Turing's faith hasn't exactly been vindicated, perhaps because of the way in which the definition of intelligence has changed over time, indeed every time that some task requiring intelligence has been successfully executed by a computing machine, from crunching numbers faster than humans to faring better at chess, and more recently, performing some rather complicated games on TV shows. I think we live in a time where it has finally become common practice to treat objects other than electronic and human computers as computing objects, and so I shall address the ineluctable generalisation of the concept of computation beyond the realm of digital computers, and more specifically its extension to natural systems. If Turing's claim were to be revised, with *objects computing* being substituted for "machines thinking", the prediction seems right on target: "I believe that at the end of the century the use of words and general educated opinion will have altered so much that one will be able to speak of [all kinds of objects computing] without expecting to be contradicted."

1.2 Digital Computation as Natural Computation

Turing's most important contribution to science is his definition of universal computation, integral to his attempt to mechanise the concept of a computing machine. A universal (Turing) machine is an abstract device capable of carrying out any computation for which an instruction can be written. More formally, given a fixed description of Turing machines, we say that a Turing machine U is universal if for any input s and Turing machine M, $U(\langle M \rangle, s)$ halts if M halts on s and outputs $M(s)$; and does not halt if $M(s)$ does not (where $\langle M \rangle$ means the codification of M in bits so that it can be fed to a Turing machine U that accepts binary inputs). In other words, U is capable of running any Turing machine M with input s.

The fact that we need hardware and software is an indication that we need a programmable substratum that can be made to compute something for us, but Turing's main contribution vis-à-vis the concept of computational universality is that data and programs can be stored together in a single memory without any fundamental distinction. One can always write a specific-purpose machine with no input to perform any computation, and one can always write a program describing that computation as the input for a (universal) Turing machine, so in a sense there is a non-essential distinction between program and data.

It is clear that one can derive a fundamental kind of natural computation from Alan Turing's seminal concept of universal computation. Turing points out [37] that given that Babbage's computer did not use electrical power, and that because Babbage's and all digital computers are in some fundamental sense equivalent, electricity cannot be a fundamental property of computation. Neither is it the carrier. In other words, Turing universality disembodies computation, uncoupling it from any physical substratum. This doesn't mean that one can carry out computations without physical elements, but rather that the nature of the physical elements is not very relevant except insofar as it bears upon the (important) question of resources (capacity, speed). A programmer uses memory space and cpu cycles in a regular computer to perform a computation, but this is by no means an indication that computation requires a computer (say a PC), only that it needs a substratum. The behaviour of the substratum is the underlying property that makes something a computation.

The main difference between a digital electronic computer and a natural system that possibly computes, is that the former was designed for the purpose, and hence one can easily identify all its elements and have recourse to them when establishing a definition of computation. For natural systems, however, there is little hope that even if their elements were to be identified, one could define their states in a way that captured all their convolutions well enough to establish that they possessed some property of computation. This situation is not that different from the undecidability of the halting problem, but it is in some sense more general. For digital computation, the undecidability of the halting problem means that if one wished to know whether a computation would eventually halt, one would have no other option than to run it and wait and see (possibly for an infinite length of time). In natural systems, the halting problem is closer to the reachability problem, that is, the question of whether a system will reach a certain configuration. By reduction to the halting problem, this can also be proven to be undecidable. The halting and reachability problems are in a strong sense behavioural and subjective in nature, as the behaviour of a system has to be determined by waiting, witnessing and recording it so that it can be understood in retrospect. If for Turing machines M, the function that computes M cannot in general be found, there is little hope of ever finding or even defining the function of a natural system. Hence one has to give up on trying to define computation for natural systems using elements such as states or functions.

We know that systems that nobody ever designed as computers are able to perform universal computation, for example Wolfram's Rule 110 [38,6] (in the rulespace of the so-called elementary cellular automata [38]), and that this, like other remarkably simple systems, is capable of universal computation (e.g. Conway's game of Life [2] or Langton's ant [19]). These systems may be said to readily arise physically, not having been deliberately designed. There is, however, no universal agreement as regards the definition of what a computer may or may not be, or as to what exactly a computation might be, even though what computation is and what a computer is are well grasped on an intuitive level.

Now we would like a concept of computation associated with natural and physical phenomena that we can measure and build on. We want a metric of computation that allows us to identify what is a computer and what is not. We want to be able to distinguish what computes from what does not. And we want a metric that we can use.

2 A Turing Test-Inspired Approach to Computation

As for Turing's test of intelligence, where one needs to accept that humans think if the test is to make sense, the reader must first accept that digital computation is performed in nature and that nature is capable of digital computation, even if only by the digital computers constructed by humans for precisely such a (general) purpose. Human behaviour is to the Turing test what digital computation is to this behavioural approach to natural computation. The argument can be rendered more succinctly thus: Electronic computers compute, electronic computers are physical objects, physical objects are part of the universe, a part of the universe is therefore capable of computation. Computers can be seen as the result of the re-programming of a part of the universe to make it compute what we want it to compute. This means that the question is not exactly whether the universe is capable of digital computation but rather whether the universe *only* performs computation, and if so, what kind of computation. I aim to provide a behavioural definition of computation that admits a wider definition of the notion of 'computation'. Notice that I am replacing the question of whether a system is capable of digital computation with the question of whether a system can behave like a digital computer and whether a digital computer can exhibit the behaviour of a natural system. So the approach is still classical in this sense, but purposely neutral with regard to the ontological issue. Also notice again the similarity with Turing's approach to machine intelligence. Turing chose to sometimes speak of "imitation" instead of "behaviour". "Imitation", however, seems to carry connotations of intentionality (see Subsection 4.2.3), and I am not very comfortable with the suggestion that a natural system may have a will to, or may purposefully imitate another system, especially if it is forced to do so artificially (although imitation is quite common in nature, where, for example, some animals mimic the behaviour of other animals to avoid being preyed upon).

To make sense of the term "computation" in the contexts I'm interested in (modern views of physics), I propose a behavioural notion of nature-like computation (similar in spirit to the coinage "physics-like computation" [29,33]) that is compatible with digital computation but meaningful in broader contexts, independent of representations and possible carriers. This will require a measure of the degree of programmability of a system based on a compressibility index which is ultimately rooted in the concept of algorithmic complexity. I ask whether two computations are the same if they look the same and I try to answer with a specific tool possessing the potential to capture a notion of qualitative behaviour.

In [9], a similar approach, but this time to the question of life, is audaciously put forward, also in the spirit of Turing's test. The idea is to recognise living systems by the way they behave and communicate through the signals transmitted between biological cells. This approach uses a biological interrogator to ask not what life is but rather when an artificial cell can be said to be alive.

Table 1. Comparison of Turing tests for intelligence, life [9] and computation

	Turing test for intelligence [37]	*Turing test for life* [9]	*Turing test for computation*
Imitated property	Thought	Cellular functions	Programmability
Subjects in question	Computing machines	Biological and artificial cells	any object
Embodiment of property	Human intelligence	Biological life (metabolism, evolution, etc)	Digital computers
Probing mechanism	Questions/answers mediated by natural language	Questions/answers mediated by physicochemical language (chemical potentials, mechanical, transduction, signalling, etc.)	Behavioural evaluation (sensitivity to external stimuli, behavioural differences, etc.) mediated by a lossless compression algorithm.

The behavioural approach takes Turing's disembodied concept of universal computation independent of substratum to its logical limit, its central question being whether one can program a system to behave in a desired way. This is again close to Turing's test in which the interrogator cannot directly see the individual replying, because intelligence is not a property that requires the possessor to have a "skin" (in the words of Turing himself [37]), for example, or to be a human being for that matter (Turing's approach), just as computation doesn't require electricity, or for that matter a digital computer (this approach). This approach that bases itself on the extent to which a system can be programmed tells us to what degree a given system resembles a computer. As the interrogator we will use a lossless compression algorithm that manifests properties of an observer, such as some type of subjectivity and finite resources. As suggested by Sutner [33], it is reasonable to require that any definition of computation in the general sense, rather than being a purely logical description (e.g. in terms of recursion theory), should capture some sense of what a physical computation might be. Sutner adds "A physical system is not intrinsically a computer, rather it is necessary to interpret certain features of the physical system as representing a computation." This obliges Sutner to take into consideration the observer and the act of interpretation of a physical system.

In many ways, this account of computation can be derived from the negation of Piccinini's 4th. feature (*the wrong things do not compute*) [27], which I think is dogmatic and gets in the way of the extension of the notion of computation to cover natural computation. Among the things that Piccinini rules out as objects that possibly compute are planetary systems, hurricanes and digestive systems. In fact, Piccinini himself seems to have some difficulty ([27], p. 508) justifying how a digestive system is not computational. For insofar as a legitimate mechanistic account can be given of a digestive system, that would mean that it possesses precisely the sorts of properties and components that are taken into consideration in determining whether or not a system counts as a computer. I will argue that one doesn't need to axiomatically rule out such systems as computing or not. I will avoid making claims about whether or not such systems compute, because the approach advanced herein is above all a pragmatic approach designed to have applications (in fact it was first developed as a tool for the investigation of dynamical properties of computer programs and not primarily as a philosophical account).

On the other hand, the behavioural account defended herein does satisfy Piccinini's 3rd requirement (*the right things compute*). Piccinini's requirements 2 (*Explanation*) and 6 (*Taxonomy*) are at the core of this proposal connecting programmability and computation and providing a grading system based on behaviour. Piccinini's requirement 5 (*Miscomputation*) doesn't seem very relevant to this proposal, and even if it were, to this author this feature doesn't seem essential to computation, for it is hard to see how a computational system can miscompute other than in the eyes of the observer. Indeed Piccinini himself sees this as troublesome in an account of computation, as it violates requirement 1. In fact, weak (i.e. observer dependent) miscomputation is pervasive in nature; I think nature amply manifests this kind of "miscomputation". In summary, I reject requirement 1 (the basis of Piccinini's account), satisfy requirements 2, 3, and 6, particularly 2 and 6 at which I think this proposal excels. And concerning requirement 4, I remain neutral, not to say unconvinced, although I can acknowledge a form of *weak miscomputation*, that is a computation that does not go in the way the observer expects it to. This approach allows a taxonomy of computation.

2.1 Algorithmic Complexity as an Approximative Measure of Programmability

The traditional connection between behaviour and computation has tended toward explaining behaviour as computation [17] or computation as emulating brain activity [22], but this author has no knowledge of explorations in the direction of explaining computation as behaviour.

This paper proposes an alternative behavioural definition of computation based on whether a system is capable of reacting to the environment—the input—as reflected in a measure of *programmability*. This will be done by using a phase transition coefficient previously defined in an attempt to characterise the evolution of cellular automata and other systems. This transition coefficient measures the sensitivity of a system to external stimuli and will be used to define the susceptibility of a system to being (efficiently) programmed, in the context of a nature-like definition of computation.

Turing's observer is replaced by a lossless compression algorithm, which has subjective qualities just like a regular observer, in that it can only partially "see" regularities in data, there being no perfectly effective compression algorithm in existence. The compression algorithm will look at the evolution of a system and determine, by means of feeding the system with different initial conditions (which is analogous to questioning it), whether it reacts to external stimuli.

The compressed version of the evolution of a system is an approximation of its algorithmic (Kolmogorov) complexity defined by [18,4]:

$$K_T(s) = \min\{|p|, T(p) = s\}$$

That is, the length of the shortest program p that outputs the string s running on a universal Turing machine T) [18,4]. A technical inconvenience of K as a function taking s to be the length of the shortest program that produces s is its non-computability, proven by reduction to the halting problem. In other words, there is no program which takes a string s as input and produces the integer $K(s)$ as output. This is usually taken to be a major problem, but one would expect a universal measure of complexity to have such a property. The measure was first conceived to define randomness and is today the accepted objective mathematical measure of complexity, among other reasons because it has been proven to be mathematically robust (in that it represents the convergence of several independent definitions). The mathematical theory of randomness has proven that properties of random objects can be captured by non-computable measures. One can, for example, approach K using lossless compression algorithms that detect regularities in order to compress data. The value of the compressibility method is that the compression of a string as an approximation to K is a sufficient test of non-randomness. If the shortest program producing s is larger than $|s|$ the length of s, then s is considered to be random.

Based on the principles of algorithmic complexity, one can use the result of the compression algorithms applied to the evolution of a system to characterise the behaviour of the system [40] by comparing it to its uncompressed evolution. If the evolution is too random, the compressed version won't be much shorter than the length of the original evolution itself. It is clear that one can characterise systems by their behaviour [40]: if they are compressible they are simple, otherwise they are complex (random-looking). The approach can be taken further and used to detect phase transitions, as shown in [40], for one can detect differences between the compressed versions of the behaviour of a system for different initial configurations. This second measure allows us to characterise systems by their sensitivity to the environment: the more sensitive the greater the variation in length of the compressed evolutions. A classification places at the top systems that can be considered to be both efficient information carriers and highly programmable, given that they react succinctly to input perturbations. Systems that are too perturbable, however, do not show phase transitions and are grouped as inefficient information carriers. The efficiency requirement is to avoid what is known as Turing tarpits [26], that is, systems that are capable of universal computation but are actually very hard to program. This means that there is a difference between what can be achieved in principle and the practical ability of a system to perform a task. This approach is therefore sensitive to the practicalities of programming a system rather than to its potential theoretical capability of being programmed. What if, instead of trying to draw a crystal clear line between what is and is not a computer, one were to define a measure of (*"computedness"*)? I propose the following approach as a first approximation to *programmability*.

Let C be an approximation to K (given that K is non-computable) by any means, for example, by using lossless compression algorithms or using the coding theorem technique we presented in [12]. Let's define the function f as the variability of a system M as the result of fitting a curve ϕ (by (linear) regression analysis) to the data points produced by different runs of increasing time t' (for fixed n) up to a given time t, of the sums of the differences in length of the approximations to Kolmogorov complexity (C) of a system M for inputs i_j, $j \in \{1, \ldots, n\} \in E$, divided by $t(n-1)$ (for the sole purpose of *normalising* the measure by the system's "volume," so that one can roughly compare different systems for different n and different t). With E an enumeration of initial inputs for M. The following expression is a more formal attempt to capture this first step:

$$f(M,t,n) = \phi\left(\frac{\sum_{j=0}^{n-1}|C(M_1(i_j))-C(M_1(i_{j+1}))|}{1(n-1)}, \ldots, \frac{\sum_{j=0}^{n-1}|C(M_t(i_j))-C(M_t(i_{j+1}))|}{t(n-1)}\right) \quad (1)$$

That is the sum of the differences of the compressed lengths of M for different initial conditions i_j. $M_t(i)$ is a system M running for time t and initial input configuration i. At the limit $\mathbb{C}_t^n$ captures the behaviour of M_t for $t \to \infty$, but the value of $\mathbb{C}_t^n$ depends on the choices of t and n (we may sometimes refer to $\mathbb{C}$ as assuming a certain t and n), so one can only aim to capture some average or asymptotic behaviour, if any (because no convergence is guaranteed). $\mathbb{C}_t^n$ is, however, an indicator of the degree of programmability of a system M relative to its external stimuli (input i). The larger the derivative, the greater the variation in M, and hence in the possibility of programming M to perform a task or transmit information at a rate captured by $\mathbb{C}_t^n$ itself (that is, whether for a small set of initial configurations M produces a single significant change or does so incrementally). Now the second step is to define the asymptotic measure, that is the derivative of f with respect to time, as a system's programmability (first basic definition):

$$\mathbb{C}_t^n(M) = \frac{\partial f(M,t,n)}{\partial t} \quad (2)$$

For example, as is shown in [40], certain elementary cellular automata rules that are highly sensitive to initial conditions and present phase transitions which dramatically change their qualitative behaviour when starting from different initial configurations can be characterised by these qualitative properties. A further investigation of the relation between this transition coefficient and the computational capabilities of certain known (Turing) universal machines has been undertaken in [42]. We will refrain from exact evaluations of $\mathbb{C}$ to avoid distracting the reader with numerical approximations that may detract from our particular goal in this paper. Other calculations have been advanced in [43] and [44].

2.2 A Behavioural Approach to Computation

The following are first approaches to definitions connected to the qualitative behaviour of *computational* systems:

Approximate variability (the number of possible different evolutions of a system): Let $U_1, U_2, \ldots$ be an enumeration of inputs to a system M. We are interested in the question of how different the evolution of $M(U_i)$ is to the evolution of $M(U_j)$, in particular the maximum difference.

Programmability: The capability of a system to change, to react to external stimuli (input) in order to alter its behaviour. Programmability, then, is a combination of variability and external control.

Computational universality: Maximum programmability.

Efficient programmability: Maximum variability changes reached in polynomial time (of a *small* degree).

Efficient universal computation: Universality with measurable variations detected in polynomial time (of a *small* degree).

Notice how close this approach is to Turing's test for intelligence. This is a kind of generalisation of the Turing test: *computation is what behaves as such*, and it does so if it can be programmed.

The following assertions follow (a technical paper with formal definitions is in preparation):

- A system U is capable of computation if $\mathbb{C}_t^n(U) > 0$ for $t, n > 0$.
- A 0-computer is not a computer in any intuitive sense because it is not capable of carrying out any calculation.
- A system capable of (Turing) universal computation has a non-zero $\mathbb{C}$ limit value (see [42]). (A non-zero $\mathbb{C}$ value, however, doesn't imply Turing universality.)
- A system U capable of Turing computational universality asymptotically converges to $\lim \mathbb{C}_t^n(U) = 1$ for $t, n \to \infty$.

The use of a general lossless compression algorithm is comparable with the role of an interrogator in Turing's test (see Fig. 2). To the compression algorithm the carrier of the computation is irrelevant as long as it can be represented in some form such that it can serve as input when running said compression algorithm. On the other hand, a compression algorithm is resource bound, in that it cannot implement in a finite time all the tests that can effectively detect all possible regularities in the data. This means that the algorithm is somehow subjective; it will first resort to what strikes it as the most obvious patterns to use to compress the data. Yet the algorithm does this in a sophisticated way, with a greater likelihood of success than a human compressor, as it is systematic and implements general methods. Lossless compression algorithms can also be set to run for a longer time to attempt more methods of compression, just as a human observer would devise more methods of compression given more time.

So a system S is provided with a random input i (a "question") and the lossless compression algorithm evaluates the reaction of the system (then mapping the input i

to a numerical value $C(S(i))$, the compressed length of $S(i)$ using the compression algorithm C). Just as observers would do for regularity appreciations (or answer evaluations), different compression algorithms may retrieve different compression lengths of $S(i)$, as they may differ in the way they compress. This compressed value is not completely arbitrary, as there is some objectivity in a strong desirable sense. This is because lossless compression is a sufficient test of non-randomness, meaning that if a lossless compression algorithm C is able to compress $S(i)$ then the Kolmogorov complexity of $K(S(i))$ cannot be greater than $(C(S(i)))$. On the other hand, no C' algorithm can compress $S(i)$ such that $(C(S(i))) < K(S)$ by definition of K, so the values of a compression algorithm C are not completely arbitrary (or subjective).

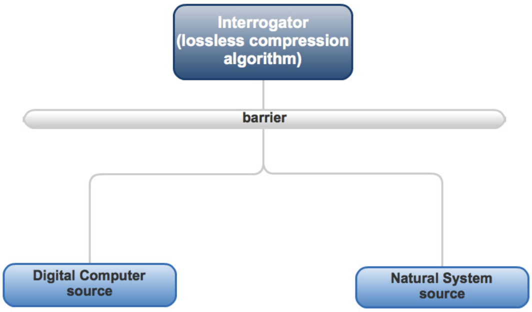


Fig. 2. The Turing-test inspired approach to the question of computation as a behavioural test undertaken by a lossless compression algorithm in the role of the answer evaluator. Notice that the natural system can be a human being or anything else.

One may challenge the configuration depicted in 2 as lacking a true questioner, given that the compression algorithm evaluates the answers but does not formulate the questions, meaning that the test, unlike Turing's, is not self-contained. This is a very good and legitimate point, but thanks to Turing, it is not very well founded. This is because from Turing we know that a system S with input i can be rewritten as a new system $S'(\langle S\rangle, i)$, that is a new system S' encoding S with input i. One can actually do this not just for a single input, but for any number of inputs, even an infinite number of inputs such as in an enumeration. Let E be an enumeration for S and p_E the program that produces E (we know that the program exists by definition). Then $S'(\langle S\rangle, \langle p_E\rangle)$ such that S' behaves like S and uses p_E to feed S with an infinite number of inputs (just as S for i, S' may not halt). So in some strong sense the system is neutral even to having all the questions at once or not.

3 A Taxonomy of Computation

The measure proposed in 2.1 can be used to dynamically define computation based on the *degree of programmability* of a system. The advantage of using the transition coefficient $\mathbb{C}$ is that it is indifferent to the internal states, formalism or architecture of a computer or computing model; it doesn't even specify whether a machine has to be digital or analog, or what its maximal computational power must be. It is only based on the behaviour of the system in question. It allows us to minimally characterise the concept of computation on the basis of behaviour alone.

Now we can attribute the property of computation to natural and physical objects, hence arriving at a measure of *Nature-like computation*, and distinguish between the computational attributes of physical objects depending on their programmability.

Our proposal has many similarities to Piccinini's mechanistic approach, yielding a hierarchy of computing objects. But while he puts calculators and (specific-purpose) computers in different categories, I don't see any essential reason to do so. He places the concept of programmability at the centre of the discussion, as I do, but all in all our approaches are very different. His mechanistic approach doesn't seem particularly suitable for natural computation. At a more fundamental level, Piccinini's approach differs from this approach in that he seems to attribute importance to the physical implementation of a computation and to its physical components, whereas this is not a matter of interest here. Unlike Piccinini, I do not think that the property of computing is an objective feature of a system.

A program can be defined as that which turns a general-purpose computer into a special-purpose computer. This is not a strange definition, since in the context of computer science a computation can be regarded as the evolution undergone by a system when running a program. However, while interesting in itself, and not without a certain affinity with our approach, this route through the definition of a general-purpose computer is a circuitous one to take to define computation. For it commits one to defining computational universality before one can proceed to define something more basic, something which ideally should not depend on such a powerful (and even more difficult-to-define) concept. Universal computation is without a doubt the most important feature of computation, but every time one attempts to define computation in relation to universal computation, one ends up with a circular statement [computation is (Turing) universal computation], thus merely leading to a version of a CT thesis.

As Piccinini suggests in [27], a Turing universal computer, and indeed a human being, can do more than follow one algorithm. They can follow any algorithm, which is typically given to them in the form of instructions. "More generally, a human can be instructed to perform the same activity (e.g. knitting or playing the piano) in many different ways. Any machine that can be easily modied to yield different output patterns may be called 'programmable'. In other words, 'being programmable' means being modiable so as to perform relatively long sequences of different operations in a different way depending on the modication."

If everyday things like fridges or lamps can be deemed computational, then it's hard to see how any physical system whatsoever is not computational (this relates to Putnam's realisation theorem, see Subsection 4.2.6). We can now meaningfully ask the question whether a lamp or a fridge is or isn't a computer, without trivialising the question itself or any possible answer. A lamp's output, for example, can be described by two different behaviours (in this case, traditionally identified as states), that is, on and off, which are triggered by external input (via a switch). Even if the lamp can be considered to react to external stimuli, it is very limited in its behaviour, and the space of its initial configurations is finite and small (it has only two possible initial configurations). Hence the slope of the differences of the behavioural evolution in time is very close to 0. A lamp is therefore a very limited computer with $\mathbb{C}$ value very close to 0. If one wished to rule out lamps or fridges as computing devices one would only need to define a threshold beyond which a system can be said to compute and beneath which it would not be said to compute. With a definition of programmability one can expect to be able to construct a *hierarchy of* computing *objects* (see Table 2), with digital general-purpose computers placed correctly (at the top of the hierarchy of computers), while other objects that we may consider (non) computing objects can be found at or near the bottom. It is clear that the threshold is at the level of specific-purpose computers, given that we may want to include in the definition of computation entities that compute simple functions such as—only—the successor function, or the sum of 2 integers, while we may not be able to assign any computing capabilities to a specific-purpose "computer" capable of—only—"computing" the identity function.

Table 2. A primitive hierarchical view of computation according to the first approximation of computation based on the coefficient $\mathbb{C}$ with customisable threshold δ is considered a computer if $\mathbb{C} > \delta$, otherwise it is not. The symbol "$>>$" is for systems for which (assuming they operate as they usually do, e.g. a fully capable human brain) no mistake about their computational capabilities can be made based on their degree of programmability approached by $\mathbb{C}$. That is, their $\mathbb{C}_t^n$ value is strictly greater than δ for any δ for t and n that are run for long enough (that is, long enough to be greater than δ).

Object	$\mathbb{C}$ *value*	*Threshold flag* ($\mathbb{C} > \delta$?)
General-purpose digital (electronic) computer	$\mathbb{C} >> \delta > 0$	Yes
Human brains	$\mathbb{C} >> \delta > 0$	Yes
Specific-purpose computers (e.g. calculators, successor machine)	$\mathbb{C} \geq 0 < \delta$	Yes/No
Lamps	$\mathbb{C} \sim 0 < \delta$	No
Rocks	$\mathbb{C} = 0 < \delta$	No

Brains and digital (Turing universal) computers can show great variation for two different random inputs, potentially even for two arbitrarily close inputs (according to same sensible distance of inputs), but for systems with low $\mathbb{C}$ this is different. For example, a lamp has only two possible "random" inputs: *on* and *off*, and the same number of outputs. For rock-like systems (including rocks themselves), the rock looks the same disregarding the possibly thinkable inputs for a rock. Fig. 3 shows a rock-like behaviour of an elementary cellular automaton.

According to Piccinini, we distinguish computers from most other things because, at the very least, computers are more versatile than other computing mechanisms. He thus attributes a measure of positive versatility to the concept of computation (or the computer). "Computers can do arithmetic but also graphics, word processing, Internet

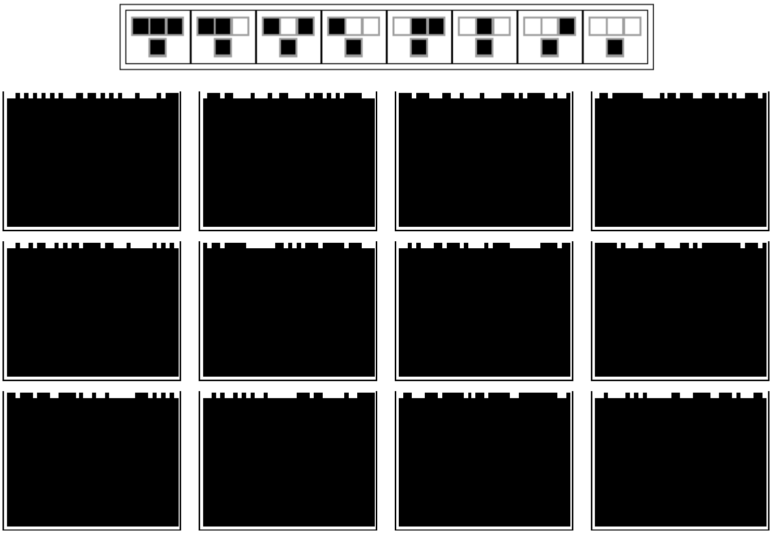


Fig. 3. Example of a "rock-like" behaviour by an elementary cellular automaton [38] with Rule 4 for several "random" initial configurations and evolving from top to bottom. The evolutions are preceded by an icon (top) illustrating the rule that the system follows for every possible cell configuration.

browsing, and a myriad other things", Piccinini says. And he adds: "Computer versatility calls for an explanation" [27]. Some objects, such as abaci, have parts that need to be moved by hand. They may be called computing aids, as Piccinini does. Of course abaci would have very small, if not zero, $\mathbb{C}$ values with no human intervention, and therefore can be flagged as non-computers, even for small δ threshold value.

This account does justice to digital computers and the practices of computer scientists and computability theorists. On the one hand, digital computers, calculators, both universal and nonuniversal Turing machines, and finite state automata, are examples of computation under the proposed definition. These can be recognised as computers, and universal digital computers can be placed at the top of the hierarchy of computational systems. On the other hand, the definition also places the concept of *programmability* at the centre of the practice of computer science, but through algorithmic complexity one can also define higher classes of computation based on Turing degrees, given that abstract machines that can solve the halting problem behaviourally perform a computation that cannot be carried out by a Turing machine that may not halt. That is, the oracle machine does halt, but it does not halt for every possible computation; it has its own new halting problem of a higher degree, and so on, hence building up the arithmetical hierarchy without need of explicit descriptions of states or functions.

It is clear that computers are not the only programmable mechanisms. So are brains, as are many other natural systems that we can now control and direct to perform certain tasks that they were not supposed to be naturally capable of (e.g. through genetic engineering). A computer is a system that can be modified to compute in different ways. I think one, if not the most important features of brains and computers is that they can be reprogrammed (in different but analogous ways). Along the lines of Fodor's conclusion [14], but with no need of state representation, if the brain is a programmable system, then it is a computer under this behavioural approach. This is paradoxical because according to Fodor's slogan "no computation without representation," according to which, in order for a system to be ascribed computational status, it needs to be construed as representing information in some way. According to the view proposed here, we should neither reject or accept this dictum because a natural occurring process can be assigned a computational value if and only if it can be programmed, regardless of whether it represents anything. As pointed out by Dresner [13], a measurement-theoretic representation typically is accompanied by a uniqueness theorem that states how all the homomorphisms from the given empirical structure to the numerical one relates to each other (that is, can be obtained from each other). I will provide some clues of how to do this in the answer to possible objections in Section 4.

Beyond formalisms, the present account of computers and computation is used to formulate a rigorous taxonomy. According to this behavioural approach, all Turing machines that compute a function other than identity are computers, and all that do so are universal Turing machines. It encompasses minds and computers while excluding almost everything else, investing minds and computers with a special status. One may think of some possible counterexamples. Think of the billiard ball computing model. It is designed to perform as a computer and can therefore be trivially mapped onto the states of a digital computer. Yet it is a counterexample of what the semantic account sets out to do, viz. to cordon off minds and computers (believed capable of computation) from things like billiard balls, tables and rocks (believed to be incapable of computation). The billiard ball computing model, as a system, however, is identified as computational in this behavioural approach, without further ado.

4 Addressing Possible Objections

Despite avoiding representationalism, which is one advantage of this approach, we find that certain objections to Turing's test, including some addressed by Turing himself, can serve as objections to the behavioural approach to computation, and that possible objections to the behavioural approach to computation can also serve as objections to the Turing test. Nevertheless, we claim that the behavioural approach can provide useful tools for natural computation, and we will use it as a basis for a set of measures capturing different properties of the dynamic behaviour of natural systems, measures drawing on concepts from algorithmic information theory and compressibility. The objections are not thoroughly addressed here, as each may require a paper of its own, but I sketch some possible responses to explore.

4.1 Technical Objections

Let me first address some possible technical objections before turning to the philosophical ones. These and other objections deserve careful scrutiny, but there is no reason to address them all in depth here.

4.1.1 The Assumption That Compressibility Can Capture Different Behaviours

One assumption that the first approach to a definition of programmability makes is that compression algorithms are able to distinguish between different behaviours. From the proposed definition we derive the differences in the compressed lengths of the evolution of a system. But it may not be clear whether the length of the compressed version for a given initial configuration of a system can differ from the length of the compressed version of the evolution of a system for a different initial configuration that yields apparently different behaviour. The problem can be stated as follows: Imagine that one has two very different processes generating different data files, but that the lengths of their compressed versions using, for example, gzip, are the same. It may seem that our approach is suggesting that both processes are behaviourally the same, even when, apart from the coincidence in the compressed lengths of their respective outputs, they may in fact be completely different.

$K(s)$, however, is *upper semi-computable*[21]; there is a sequence of lossless compression algorithms approximating $K(s)$: $C_1(s) \geq C_2(s) \geq C_3(s) \geq \ldots \geq K(s)$. That is, one can find a sequence of compression algorithms that asymptotically approaches K. $K(s)$ cannot then be greater than the most compressed version of s. The invariance theorem [4] in the theory of algorithmic information guarantees that the outputs can be distinguished from one another at the limit, no matter how close they are to each other, by a compression algorithm approaching K, and up to a bounded degree of precision (which can be large, but increasing t eventually overcomes it). More formally, the invariance theorem states that if $C_U(s)$ and $C_{U'}(s)$ are the shortest programs generating s using the universal Turing machines U and U' respectively, their difference will be bounded by an additive constant independent of s.

It is easy to see that the underlying concept is that since both U and U' are universal Turing machines, one can always write a general translator (a compiler) between U and U' such that one can run either Turing machine and get one or another complexity value, simply adding the constant length of the translator to the result.

This means that eventually, if two processes are essentially different in the sense of algorithmic complexity, they will have different C values from some time t on up to K. The caveat that a system may be characterised in an imprecise fashion still applies, but the invariance theorem guarantees that the approach is sound theoretically, even if in practice it may sometimes be misleading, in a way that we are used to with compression algorithms that may not "see" regularities in a file (e.g. a file containing the digits of π).

It is worth noticing that two different evolutions produced by the same rule system, such as a cellular automaton, may not necessarily have the same Kolmogorov complexity (in fact it is unlikely they will if they appear different) because the system in question is $S(i)$ and not S alone, that is S for the initial configuration i (e.g. Rule 30 elementary cellular automaton [38] starting from a black cell is a different system than Rule 30 starting from a repetition of ten times 01). From Turing's universality, we know that $S(i)$ can always be rewritten as S', that is a system with empty input that behaves like S for input i, where it is clear that $S \neq S'$, and this difference is ultimately captured by the difference between $K(S)$ and $K(S')$, that is the lengths of the shortest programs producing S and S'.

4.1.2 The Choice of Enumeration of Initial Configurations

The interrogator plays an important part in this Turing-based approach, which is why the initial input configurations are key—their role is analogous to that of the interrogator questioning the system. In general, one can always tamper with an enumeration E to make a system behave in a certain way for a limited period of time, as one can always run a system and then pick initial conditions for which the system behaves in a certain way, proceeding to design another enumeration E' for which the first E'_t members are members of E but sorted from $t = 1, ..., n$ such that the system behaves in a desired way for the first n elements. So how sensitive to the choice of initial input enumeration is the Turing-test inspired approach to the problem of natural computation? One can make n as large as one wishes, but the limit behaviour of a system will always go beyond n. Does this guarantee that from some point on (e.g. n) the system will start behaving "naturally"? Imagine that one knew that a system behaved in a certain way for even length initial configurations. One could then design a E such that all initial configurations are of even length. But E has to reach every possible initial configuration in finite time, so there is no way to design E so that it would run all even length inputs and then all odd length inputs in a finite time. There is no way to fool the limit analysis of the behaviour of a system by tampering with the initial configurations for more than a finite number of inputs.

The general question of the appropriate enumeration of inputs for a system is worth exploring, especially for natural systems, given that it is not always clear what the enumeration of inputs for a natural system might be (questions arise, for example, about continuous-value parameters that may need to be discretised in order for a compression algorithm to analyse). One obvious problem is that of "encrypted systems". What if an efficient programmable computing system looks intentionally random and inefficient? Say one Turing universal system (e.g. Rule 110 [38,6]) behaving like another random-looking system (such as Rule 30 in the same rulespace). It is still Rule 110, but the question is whether one would be able to identify and program Rule 110 if it is behaving like Rule 30. It may be that one can only know it is Rule 110 if one knows the decrypting function, so the compression algorithm can be fooled. This is related to who can pass a "stupidity test", that is a system that is so smart that it knows how to look stupid, or to really is (one cannot pass, however, an intelligence or computation test without being intelligent or being able to compute.). The question of "encrypted systems" occurring in nature is important to address. But this is certainly related to a feature I think is desirable in this behavioural approach, that of observer-dependent subjectivity (Subsection 4.2.4) and to the question of the enumeration of initial conditions (Subsection 4.1.2) and the question of some sort of minimal need for representation (Subsection 4.2.1).

4.2 Foundational Objections

It is interesting to see how some objections serve at once as arguments against the Turing machine intelligence approach and this natural computational approach, while others do not (e.g. the Mathematical Objection (Searle [32], Penrose [25]) doesn't seem obviously to apply to the question of computation). Other examples are the theological and the consciousness arguments, which work against both machine intelligence and natural computation by endowing humans and natural things with *qualia*, which are said not to be concomitants of the domain of digital computation. The objections work differently however, because in the case of machine intelligence they are meant to "safeguard" the essence of the human being, endowing it with irreproducible qualities such as consciousness, while in the case of natural computation they work to "safeguard" the nature of digital computation. The advantage of my approach as compared to Turing's is that there are fewer people willing to defend machines than humans, though heated debates are carried out in both directions. The current tendency in computation, however, is greater openness to the possibility that objects and systems other than electronic computers compute.

4.2.1 Some Representation Is Needed

It is interesting to note that one needs some representation of the output of a system before feeding the compression algorithm (see Fig. 4). What about the introduction or the simplification of complexity in the encoding process from the language of the system to the language of a digital computer implementing the lossless compression algorithm?

This is indeed the case, and it implies that there is some communication and mapping between the natural system and the digital computer implementing the lossless compression algorithm, but this mapping is of a very different nature from the mapping of states or functions among systems. Is this representation always possible?

On the one hand, one can always discretise data. On the other hand we know that a discrete language can always be translated into binary. So in a technical sense this is always possible. This is related to the previous discussion of the question of whether a universal system could emulate a random-looking system to hide its programmability capabilities, and what this would mean.

The proposal advanced herein is, however, different to the requirement of a strong form of representationalism, where knowing the states of a system to put them into relation with another is needed, which is fully dependent on the complete (and unlikely) knowledge of the states of a natural system. Here, however, it is only needed a weak

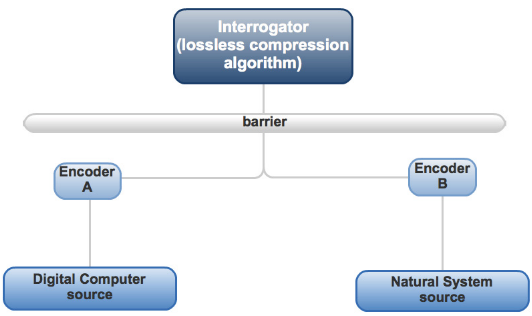


Fig. 4. What is the nature of the encoders? They work in both directions encoding "questions" properly for each system, and feeding the lossless compression algorithm in the right format. Encoders A and B may be of very different nature. Simple encoders always seem a possibility, but questions about their implementation and role remain legitimate.

translation of one system output language into another, represented by the encoders boxes in Fig. 4. Encoders should not be seen as a drawback, we deal with them all the time in computing in the form of compilers. In systems development different programming languages are used for different purposes in different places of a system. Even if one can simulate each other as Turing complete, very few non-trivial applications today are fully developed in a single programming language.

4.2.2 Programmability

A second immediate reaction is whether placing programmability at the centre of a definition of computation is too strong as an assumption. For one may think of artificial and natural systems that may not look *programmable*, yet one would be ready to grant can compute (e.g. discrete neural networks). I think this objection arises from a conflation between the standard meaning of programming and the behavioural one I am advancing here. While it is true that for many artificial and natural systems a concept of programmability is difficult to determine, the concept of programmability advanced in this paper is about whether one can, by any means, make a system behave in a way other than the way it was already behaving. In this sense, for example, a logic circuit or a batch process may not qualify as a computation if these are unable to react to external stimuli or if the observer is unable to witness such an interaction if it happens in the design or the launch of a computing process.

4.2.3 Human-Machine and Intentionality Objections

When Julien Offray de La Mettrie [35] took Descartes' method to what he claimed was its logical conclusion in his L'homme-machine, the argument was that Descartes' attempt to defend the theory of a human soul by relegating mechanical behaviour to animals in fact acted against humans. For if animals were capable of feeding, moving and interacting with other animals, strictly speaking, there was nothing to prevent human behaviour from being seen as a consequence of mechanical behaviour. In the Turing test we see a similar reversal of the argument, where it is not the machine's intelligence that is questioned but rather the intelligence of the human being, not because the questioners harbour the suspicion that humans may not be intelligent but because the mechanisms that drive human intelligence may turn out to be of the same order as those that drive computers today.

Searle advances the problem of intrinsic meaning or "intentionality" [32]. Harnad [16] defines it as the symbol grounding problem. I consider this objection weak in our context (though it remains to be further explored), because if assumed, the definition of computation is rendered meaningless in the physical context (we know we can program certain natural things, these things would be considered computers when computing for us, and not when not). For the Turing test, some "intention" is desirable, as Turing is trying to convince his reader that there is no argument in principle for a machine to fail an intelligence test if it increasingly improves its performance when imitating intelligent human behaviour. Also it is clear that electronic computers, back in Turing's time as well as today, are assembled for the purpose of computing, hence no harm is done by assuming some intentionality.

Dennett has suggested [11] that it would seem that explicit representation is not necessary for the explanation of propositional attitudes. For example, during a game of chess with a computer program, attitudes such as "It thinks that the queen should be moved to the left" are often attributed to the computer. Yet no one would suggest that the computer actually thinks or believes, in the way we do. I think it is clear how this behavioural approach to computation is compatible with this view, and neutral on intentionality questions, as it is only interested in the ways a system seems to behave and not whether it "really" does so (meaning it intended to do so, whether we are concerned with computers or with natural systems, including the brain).

4.2.4 The Observer-Oriented Objection

One immediate reaction to, and a possible objection to this approach, concerns the applicability of such a behavioural (observer-oriented) definition, given the possibly arbitrary choice of δ (see e.g. Table 2). According to certain arguments, computation is observer-relative, either in the sense that many physical systems implement many computations (Putnam [28]), or in the sense that almost all physical systems implement all computations (Searle [31]). Some physical objects, for example, may be seen to implement any computation of whatever complexity. Thus the walls in Searle's [31], implement his wordprocessing program. Since the physical description of an object underdetermines its computational description in this way, computation is deemed observer-relative [31].

This is of course a legitimate objection, which also applies to other behavioural approaches to other notions, such as the notion of intelligence, and to the Turing test. I have suggested (Section 2.1), however, that a measure of limit behaviour is possible, and that even if δ is very large, one can always overcome it over time for systems that are indubitably computing devices according to the programmability approach (computers and brains), while one can always contrive to have trivial devices such as lamps and rocks not pass as computers, leaving a flexible space in between for systems that may or may not, subjectively, be considered computers.

Its dependence on programming language or universal Turing machine has traditionally been considered one of the drawbacks of Kolmogorov complexity. In this approach we actually take advantage of this property of Kolmogorov complexity, as it assorts with a behavioural approach to computation that cannot but be observer-(or machine-) relative. This is because the Kolmogorov (program-size or algorithmic) complexity only makes sense once a universal Turing machine or Turing-complete language is fixed. On the other hand, because Kolmogorov complexity (K) is uncomputable (another commonly identified drawback), or more precisely, upper semicomputable, it is what the compression algorithm "observes" that approaches K that we will turn to our advantage in capturing the qualitative behaviour of a computational system in order to quantitatively measure it.

Piccinini argues that any reasonable definition of computation should be objective. I don't think, however, that this should be a sine qua non of a reasonable account of computation, nor that failure to meet this objectivity criterion makes an account vacuous or trivial. In fact I think computation is intrinsically user/observer oriented, both in practice and in theory. In practice, computation is mostly, if not entirely about programming systems. On the one hand, programming systems is intentional (driven by the desire to make a computer behave in a particular way), even if intentionality is not essential to computation. On the other hand, theory prescribes subjectivity in various ways. The halting problem can be read as an observer-relative property of computational systems, given that one cannot, in general, ever know whether a computation will halt except by running a system for a number of steps—which depends on the willingness of the observer to wait, if it doesn't halt before the specified number of steps. The problem is not exclusive to halting, but extends to reachability in general, that is, the question of whether a system will reach a certain configuration. Universal computation is subjective in the sense that one has to decide to stop a computation and deal with the fact that one may never know whether such a computation will ever halt or reach a certain configuration.

4.2.5 The Halting Problem Prescribes Subjectivity at All Levels

The halting problem is the problem of deciding whether a computation will halt or not. The halting problem implies that computations can be divided into 2 categories: reducible and irreducible, that is computations that are simple enough to be determined to halt or never halt, and computations for which the only option is to run them and wait for them to halt, which may obviously take an infinite amount of time. Irreducible computations can also be classified into 2 kind of computations: computations that never halt and therefore not even running the computation will help and computations that halt in time t but there is no way to know t but by running the computation for at least t. Clearly this characterisation incorporates an important role of the observer in that there exists computations for which one can only know whether they will halt by running them, and introduces a subjective component, namely the fact that the observer has to decide a runtime cutoff that is willing to wait before making an informed assumption about the (non-)halting characteristic of a computation.

Now one can see how an observer is important in the account of computation even for the most classical case of the unsolvability of the halting problem. This is even more evident when considering other phenomena, such as reachability, that is whether a computation will reach certain configuration, in which for some computations only an observer willing to run and witness the computation may answer.

The undecidability of the halting problem affects all theoretical and practical notions related to computation. For example, in Kolmogorov complexity one can never say whether an object is random (one can say whether an object is simple if it has been compressed but not the converse). This doesn't make algorithmic complexity useless. In fact it is this observer-relative property (with respect to compression algorithms that may or may not "see" regularities in the data) that the measure is most useful—for all kinds of applications, including classification of animal species and languages by compressibility, detection of genetic sequences, fraud and plagiarism detection. In finite Kolmogorov complexity, finite randomness is in the eye of the beholder, in the sense that any finite sequence can always be part of a random or non-random string. Hence the quality of being random is observer-dependent, just as it is in the case of the halting problem.

I think that it is denying the role of the observer that makes the intuitive notion of computation vacuous or trivial. The observer plays an essential role in the definition of computation. This is made explicit in our definition of computation, for the purposes of generalising and characterising natural computation.

Under this approach, computation is observer-relative (in agreement with many authors who endorse computationalism), just as intelligence is observer-relative in the case of the Turing test. We find that certain objections to Turing's test, including some addressed by Turing himself, can serve as objections to this behavioural approach to computation (we will address some of them), and that possible objections to the behavioural approach to computation can also serve as objections to the Turing test. Nevertheless, we claim that the behavioural computation approach can provide useful tools for natural computation, and we will use it as the basis for a toolkit of quantitative measures—based on concepts from algorithmic information theory and compressibility—capturing different qualitative properties of natural systems.

Paradoxically, the behavioural approach does not explain a system's behaviour, at least not in full, for we can explain part of a system's behaviour once a first behavioural analysis is performed, but not in the way we would be led to expect if we followed Smith or Piccinini, for it is not intended to be a theory of computing, nor does it set out to fully account for the causes of a system's behaviour, only for its apparent behaviour. The approach proves to have applicability and to provide insight into the properties of dynamical systems about whose internal states one could potentially have no information, nor any clue as to the possible mappings between a natural and an abstract computational system. But it also works well for systems which we know and whose internal states we can study in full detail, producing all manner of mappings to other models of computation, as we have shown using cellular automata and the way in which the measures based on this behavioural approach allow us to characterise phase transitions or rates of information transfer from a purely behavioural perspective.

Take the example of having to calculate the Lyapunov exponents of a natural system. Even if the system can be described as a dynamical system for which orbits can be described, this already assumes that one is able to represent such dynamics. Of course the behavioural approach also assumes that one can capture the behaviour of the system, but it does not assume full knowledge of the precise evolution of the system. In fact one can to some extent analyse a system in an instant of time without having to go through intermediate times (this will of course impact the final result, as it improves in direct proportion the more one observes the system).

If the observer is essential to the definition of computation, one has to acknowledge that there is no sense to the most general question of whether the universe computes, because no definition of the universe allows for external stimuli (external to the universe), nor for the output of the universe to reside outside it for an observer to evaluate.

4.2.6 Does Implementation Matter?

The question of the implementation of computation seems not to have been taken seriously until critics of computationalism brought forward certain arguments to the effect that a great many physical systems implement many, if not all, computations. Such arguments have been presented by Putnam [28] and Searle [31]. According to Putnam (the eponymous Putnam's Realization Theorem), "for every ordinary open system S, for every finite state automaton M (without input and output), for any number n of computational steps of the automaton M, and for every real-time interval I (divisible into n subintervals) S realizes n computational steps of M within I". And according to Searle (what is sometimes called Searle's Thesis), "for any program and for any sufficiently complex (physical) object, there is some description of the object under which it is implementing the program."

Along the lines of the question asked by Chalmers [5], what makes a rock compute something (or nothing) rather than everything? It seems that, at least prima facie, what (abstract) computability and (concrete) computation have in common is some logical description, only the characterisation of the latter isn't exhausted by a purely logical description, so implementation does matter. And it does matter in my approach, given that the rock may potentially be capable of any computation (think of using its particles to build a more programmable device), but it does not do so at the level at which it must be described as a rock, and if we look at it through a Turing-test inspired lens and attempt to make it behave in one way or another, i.e. program it to behave differently for different external stimuli (see Fig. 3 for a "rock-like" behaviour of an abstract system).

4.2.7 Laws Have No Distinguished Character

It has been suggested [24] that I am assigning a special status to physical laws, or to computer programs for that matter. This is an understandable objection but in fact it represents a misconception of my position. The misunderstanding resides in the conclusion that by connecting laws to computer programs as opposed to data, I give physical laws a special, immortal and unchanging status. Computer programs, however, can be written in bits. And, as I have explained in Section 1.2, Turing proved that computer programs and data are not essentially different; one can always exchange one for the other. That is, it is possible to write the transition table of a Turing machine in the form of an input for a universal Turing machine, or to build a transition table (a Turing machine) from the computer program description.

In algorithmic probability there is only one strong assumption regarding the distribution of objects. What Levin's universal distribution is supposed to indicate is the probability of a string being generated by a program, but one has to make an assumption as regards the distribution of programs in order to talk about *picking a random program*. And that is the only possible uninformed assumption–the uniform distribution. That is, any program of the same length is equally likely to occur as a product of chance. But apart from this one is free to interchange programs. There is nothing special about physical laws. They can be seen as highlighting or summarising a regularity in the data (the world), and data can change, hence physical laws may do so as well.

4.2.8 The Question of Scale

In the real world, things are constituted by smaller elements unless they are elementary particles. One therefore has to study the behaviour of a system at a given scale and not at all possible scales, otherwise the question becomes meaningless, as elements of a physical object are molecules, and ultimately atoms and particles that have their own behaviour, about which too the question about computation can be asked. This means that a $\mathbb{C}$-computer may have a low or null $\mathbb{C}$ at some scale but contain $\mathbb{C}'$-computers with $\mathbb{C}' > \mathbb{C}$ at another scale (for which the original object is no longer the same as a whole). A setup in which $\mathbb{C}' \leq \mathbb{C}$ is actually often common at some scale for any computational device. For example, a digital computer is made of simpler components, each of which at some macroscopic level but independently of the interconnected computer, is of lower behavioural richness and may qualify for a $\mathbb{C}$ of lower value. In other words, the behavioural definition is not additive in the sense that a $\mathbb{C}$-computer can contain or be contained in another $\mathbb{C}'$-computer such that $\mathbb{C} \neq \mathbb{C}'$.

In the physical world, under this qualitative approach, things may compute or not depending on the scale at which they are studied. To say that a table computes only makes sense at the scale of the table, and as a $\mathbb{C}$-computer it would have a very limited $\mathbb{C}$, that is a very limited range of behaviour, for it can hardly be programmed to do something else.

The behavioural definition is not immune to scale. Something may or may not compute at a certain level of description but it may compute at another more macro- or more microscopic level of description. But the concept of the object is also not scale invariant (we call things by different names when we change scale, e.g. we call the constituents of a rock atoms, or the aggregation of H_2O in liquid form water).

4.2.9 Batch Process Objection

A batch process is the execution of a program on a computer without the need of any external intervention. This kind of system would go unnoticed by this proposed behavioural approach given the insensitivity of such a system to any external stimuli, as it is programmed to perform a task without interacting with anything else until it stops and produces some output, if any. During this time the process may look as if it were doing nothing, but this is merely appearance, and there are ways for the observer to ascertain that it is in fact computing, at the lowest level by its external resource consumption and release, such as energy and heat (which one could also manipulate to make the process change behaviour, for example, stop the process), and at another level, by monitoring the process for a long-enough time. The batch process instance is only valid as an objection between the time $t = 1$ when the process is actually initiated (it has to), and $t = n - 1$, because at least at one time $t = 0$ or $t = n$ (if it halts and produces an output) some interaction with the outside is expected to happen. So while some computers may fail to be identified by the behavioural definition, the limit behaviour definition seems to be immune to this objection, except insofar as it may for all (proper) purposes consider something that may be computing as not computing because it is disconnected from the external world in which the observer lives.

4.2.10 The Contingency of Quantum Mechanics

Using algorithmic probability (AP) S. Lloyd claims [20]:

> I would suggest, merely as a metaphor here, but also as the basis for a scientific program to investigate the computational capacity of the universe, that this is also a reasonable explanation for *why the universe is complex. It gets programmed by little random quantum fluctuations*, like the same sorts of quantum fluctuations that mean that our galaxy is here rather than somewhere else.
>
> (S. Lloyd, 2002)

We don't know whether AP can be adapted to a quantum version but we do know that there is no need for *quantum fluctuations* to generate algorithmic structure [12] that Lloyd was trying to explain on the basis of quantum mechanics.

The strong assumption in the context of classical computation and classical mechanics is *determinism*. The wave-function collapse in quantum mechanics and the problem of measurement may challenge determinism at two different levels, but otherwise classical mechanics prescribes determinism in the (macroscopic) universe. Classical (Newtonian) mechanics guarantees the deterministic output (the problem is to generate the same input). Running a computation twice with the same input generates the same output through exactly the same path just as would do a classical system following the rules of classical mechanics (that in practice this is not possible is due to the problem of limited accuracy of the measurement of the initial conditions).

4.2.11 Connections to Computational Complexity

In the light of this research now one can find an interesting connection of the measure $\mathbb{C}$ to traditional computational complexity where one is concerned with the needed resources for a computation to be carried out. $\mathbb{C}$ provides clues on whether a system may be Turing universal but not on whether a system may not be universal, because universality requires variability and sensitivity to external stimuli to program a computation. Also $\mathbb{C}$ is greatly influenced but not directly related to universality given that universality will guarantee that $\lim \mathbb{C}_t^n = \infty$ for $t, n \to \infty$, but a positive value $\mathbb{C}$ does not guarantee universality, it guarantees sensibility which in this context is a measure of the capability of the system to be programmed to do different (even if limited) computations by transferring information from the input to the output. But $\mathbb{C}$ ultimately depends on the way in which $\mathbb{C}$ is calculated for a finite number of initial configurations and a finite number of steps, hence systems that may compute at a slow pace may be misclassified for some t and n small enough. $\mathbb{C}$ can be, however, thought as also measuring efficiency of a system to be programmed. So one can relativise this concept introducing time complexity classes. So one can say that a system with $\mathbb{C}$ value that grows in linear time is efficient, but it is not efficient if it grows in logarithmic time.

5 Concluding Remarks

This paper has addressed the problem of recognising computation. It partially fulfils some of the requirements that according to several authors any definition of computation should meet (e.g. [30], [27]), while I have made the case that some properties are not needed and should not be required or expected, especially in the novel context of natural computation and artificial biology.

Computational models can be very useful even when not every detail about a system is known. The aim of systems biology, for example, is to understand the functional properties and behaviour of living organisms, while the aim of synthetic biology is to design, control and program the behaviour of living systems, even without knowing the details of the biological systems in question. Along the lines of Turing's intelligence test, this approach seems to be useful for investigating qualitative properties of computing systems in a quantitative fashion, and since it places programmability at the centre of computation it serves as a possible foundation for natural computation.

Acknowledgements. I would like to thank Marcin Miłkowski for some references and the organisers of the symposium *Natural/Unconventional Computing and its Philosophical Significance* for their kind invitation to speak at the AISB/IACAP World Congress

2012—Alan Turing 2012. I also wish to thank the FQXi for the mini-grant awarded by way of the Silicon Valley Foundation under the title “Time and Computation”, which this project studies in connection to behaviour (mini-grant no. 2011-93849 (4661)).

References

1. Ausländer, S., Ausländer, D., Müller, M., Wieland, M., Fussenegger, M.: Programmable single-cell mammalian biocomputers. Nature (2012)
2. Berlekamp, E.R., Conway, J.H., Guy, R.K.: Winning Ways for your Mathematical Plays. AK Peters Ltd. (2001)
3. Blanco, J.: Interdisciplinary Workshop with Javier Blanco: Ontological, Epistemological and Methodological Aspects of Computer Science. University of Stuttgart, Germany (July 7, 2011)
4. Chaitin, G.J.: On the length of programs for computing finite binary sequences: Statistical considerations. Journal of the ACM 16(1), 145–159 (1969)
5. Chalmers, D.J.: Does a Rock Implement Every Finite-State Automaton? Synthese 108, 310–333 (1996)
6. Cook, M.: Universality in Elementary Cellular Automata. Complex Systems 15, 1–40 (2004)
7. Conrad, M.: The Price of Programmability. In: Herken, R. (ed.) The Universal Turing Machine, A Half-Century Survey. Springer (1994)
8. Berlekamp, E., Conway, K., Guy, R.: Winning Ways for your Mathematical Plays, vol. 2. Academic Press (1982)
9. Cronin, L., Krasnogor, N., Davis, B.G., Alexander, C., Robertson, N., Steinke, J.H.G., Schroeder, S.L.M., Khlobystov, A.N., Cooper, G., Gardner, P.M., Siepmann, P., Whitaker, B.J., Marsh, D.: The imitation game—a computational chemical approach to recognizing life. Nature Biotechnology 24(10) (2006)
10. Davis, M.: Universality is Ubiquitous, Invited Lecture. In: History and Philosophy of Computing (HAPOC 2011), Ghent, November 8 (2011)
11. Dennett, D.C.: Brainstorms: Philosophical Essays on Mind and Psychology. MIT Press (1981)
12. Delahaye, J.-P., Zenil, H.: Numerical Evaluation of the Complexity of Short Strings: A Glance Into the Innermost Structure of Algorithmic Randomness. Applied Math. and Comp. 219, 63–77 (2012)
13. Dresner, E.: Measurement-theoretic representation and computation-theoretic realisation. The Journal of Philosophy cvii(6) (2010)
14. Fodor, J.: The Language of Thought. Harvard University Press (1975)
15. Fredkin: Finite Nature. In: Proceedings of the XXVIIth Rencotre de Moriond (1992)
16. Harnad, S.: The Symbol Grounding Problem. Physica D 42, 335–346 (1990)
17. Hodgkin, A., Huxley, A.: A quantitative description of membrane current and its application to conduction and excitation in nerve. Journal of Physiology 117, 500–544 (1952)
18. Kolmogorov, A.N.: Three approaches to the quantitative definition of information. Problems of Information and Transmission 1(1), 1–7 (1965)
19. Langton, C.G.: Studying artificial life with cellular automata. Physica D: Nonlinear Phenomena 22(1-3), 120–149 (1986)
20. Lloyd, S.: Computational capacity of the Universe. Physical Review Letters 88, 237901 (2002)
21. Li, M., Vitányi, P.: An Introduction to Kolmogorov Complexity and Its Applications, 3rd edn. Springer (2008)

22. McCullock, W., Pitts, W.: A Logical Calculus of Ideas Immanent in Nervous Activity. Bulletin of Mathematical Biophysics 5(4), 115–133 (1943)
23. Minsky, M.: Computation: Finite and Infinite Machines. Prentice Hall (1967)
24. Parsons, W.T.: Patterns in the Noise: Physics as the Ultimate Environmental Science. FQXi's, Essay Contest, Questioning the Foundations (2012)
25. Penrose, R.: The Emperor's New Mind. Oxford University Press (1989)
26. Perlis, A.J.: Epigrams on Programming. SIGPLAN Notices 17(9), 7–13 (1982)
27. Piccinini, G.: Computers. Pacic Philosophical Quarterly 89, 32–73 (2008)
28. Putnam, H.: Representation and Reality. MIT Press, Cambridge (1988)
29. Margolus, N.: Physics-like Models of Computation. Physica 10D, 81–95 (1984)
30. Scott, D.S.: Outline of a mathematical theory of computation. Technical Monograph PRG-2. Oxford University Computing Laboratory, England (November 1970); Theoretical Computer Science 412, 183–190 (2011)
31. Searle, J.R.: The Rediscovery of the Mind. MIT Press (1992)
32. Searle, J.R.: Is the Brain a Digital Computer. In: Philosophy in a New Century, pp. 86–106. Cambridge University Press (2008)
33. Sutner, K.: Computational Processes, Observers and Turing Incompleteness
34. Turner, R.: Specification. Minds and Machines 21(2), 135–152 (2011)
35. Thomson, A.: Machine man and other writings. Cambridge University Press (1996)
36. Turing, A.M.: Systems of logic based on ordinals. Proc. London Math. Soc. 45 (1939)
37. Turing, A.M.: Computing Machinery and Intelligence. Mind 59, 433–460 (1950)
38. Wolfram, S.: A New Kind of Science. Wolfram Media (2002)
39. Woods, D., Neary, T.: The complexity of small universal Turing machines: a survey. Theor. Comput. Sci. 410(4-5), 443–450 (2009)
40. Zenil, H.: Compression-based investigation of the behaviour of cellular automata and other systems. Complex Systems 19(2) (2010)
41. Zenil, H., Soler-Toscano, F., Joosten, J.J.: Empirical Encounters With Computational Irreducibility and Unpredictability. Minds and Machines 222(3), 149–165 (2012)
42. Zenil, H.: On the Dynamic Qualitative Behaviour of Universal Computation. Complex Systems 20(3) (2012)
43. Zenil, H.: What is Nature-like Computation? A Behavioural Approach and a Notion of Programmability. Philosophy & Technology (2013), doi:10.1007/s13347-012-0095-2
44. Zenil, H.: Programmability for Natural Computation and the Game of Life as a Case Study. J. of Experimental & Theoretical Artificial Intelligence (forthcoming)

Alan Turing's Legacy: Info-computational Philosophy of Nature

Gordana Dodig-Crnkovic

Mälardalen University, Department of Computer Science and Networks,
School of Innovation, Design and Engineering, Västerås, Sweden
gordana.dodig-crnkovic@mdh.se

Abstract. Alan Turing's pioneering work on computability, and his ideas on morphological computing support Andrew Hodges' view of Turing as a natural philosopher. Turing's natural philosophy differs importantly from Galileo's view that the book of nature is written in the language of mathematics (The Assayer, 1623). Computing is more than a language used to describe nature as computation produces *real time physical behaviors*. This article presents the framework of Natural info-computationalism as a contemporary natural philosophy that builds on the legacy of Turing's computationalism. The use of info-computational conceptualizations, models and tools makes possible for the first time in history modeling of complex self-organizing adaptive systems, including basic characteristics and functions of living systems, intelligence, and cognition.

Keywords: Alan Turing, Morphological computing, Natural computing, Info-computationalism.

1 Turing and Natural Philosophy

Andrew Hodges [1] describes Turing as a Natural philosopher: "He thought and lived a generation ahead of his time, and yet the features of his thought that burst the boundaries of the 1940s are better described by the antique words: natural philosophy." Turing's natural philosophy differs from Galileo's view that the book of nature is written in the language of mathematics (The Assayer, 1623). Computation is not just a language of nature; it is the way nature behaves. Computing differs from mathematics in that computers not only calculate numbers, but more importantly they can produce real time physical behaviours.

Turing studied a variety of natural phenomena and proposed their computational modeling. He made a pioneering contribution in the elucidation of connections between computation and intelligence and his work on morphogenesis provides evidence for natural philosophers' approach. Turing's 1952 paper on morphogenesis [2] proposed a chemical model as the basis of the development of biological patterns such as the spots and stripes that appear on animal skin.

Turing did not originally claim that the physical system producing patterns actually performs computation through morphogenesis. Nevertheless, from the perspective of

G. Dodig-Crnkovic and R. Giovagnoli (Eds.): *Computing Nature*, SAPERE 7, pp. 115–123.
DOI: 10.1007/978-3-642-37225-4_6

info-computationalism, [3,4] argues that morphogenesis is a process of morphological computing. Physical process, though not computational in the traditional sense, presents natural (unconventional), physical, morphological computation.

An essential element in this process is the interplay between the informational structure and the computational process – information self-structuring. The process of computation implements physical laws which act on informational structures. Through the process of computation, structures change their forms, [5]. All computation on some level of abstraction is morphological computation – a form-changing/form-generating process, [4].

In this article, info-computationalism is identified as a new philosophy of nature providing the basis for the unification of knowledge from currently disparate fields of natural sciences, philosophy, and computing. An on-going development in bioinformatics, computational biology, neuroscience, cognitive science and related fields shows that in practice biological systems are currently already studied as information processing and are modeled using computation-theoretical tools [6,7,8].

Denning declares: “Computing is a natural science” [9] and info-computationalism provides plenty of support for this claim. Contemporary biologists such as Kurakin [10] also add to this information-based naturalism, claiming that “living matter as a whole represents a multi-scale structure-process of energy/matter flow/circulation, which obeys the empirical laws of nonequilibrium thermodynamics and which evolves as a self-similar structure (fractal) due to the pressures of competition and evolutionary selection”. [11, p5]

2 Universe as Informational Structure

The universe is, from the metaphysical point of view, "nothing but processes in structural patterns all the way down" [12, p228]. Understanding patterns as information, one may infer that information is a fundamental ontological category. The ontology is scale-relative. What we know about the universe is what we get from sciences, as "special sciences track real patterns" [12, p242]. This idea of an informational universe coincides with Floridi’s Informational Structural Realism [13,14]. We know as much of the world as we explore and cognitively process:

“Reality in itself is not a source but a resource for knowledge. Structural objects (clusters of data as relational entities) work epistemologically like constraining affordances: they allow or invite certain constructs (they are affordances for the information system that elaborates them) and resist or impede some others (they are constraints for the same system), depending on the interaction with, and the nature of, the information system that processes them.” [13, p370].

Wolfram [15] finds equivalence between the two descriptions – matter and information:

“[M]atter is merely our way of representing to ourselves things that are in fact some pattern of information, but we can also say that matter is the primary thing and that information is our representation of that. It makes little difference, I don’t think there’s a big distinction – if one is right that there’s an ultimate model for the representation of universe in terms of computation.” [16, p389].

More detailed discussion of different questions of the informational universe, natural info-computationalism including cognition, meaning and intelligent agency is given by Dodig Crnkovic and Hofkirchner in [17].

3 The Computing Universe – Naturalist Computationalism

Zuse was the first to suggest (in 1967) that the physical behavior of the entire universe is being computed on a basic level, possibly on cellular automata, by the universe itself, which he referred to as "Rechnender Raum" or Computing Space/Cosmos. Consequently, Zuse was the first pancomputationalist (natural computationalist), [18]. Chaitin in [19, p.13] claims that the universe can be considered to be a computer "constantly computing its future state from its current state, constantly computing its own time-evolution account!" He quotes Toffoli, pointing out that "actual computers like your PC just hitch a ride on this universal computation!"

Wolfram too advocates for a pancomputationalist view [15], a new dynamic kind of reductionism in which the complexity of behaviors and structures found in nature are derived (generated) from a few basic mechanisms. Natural phenomena are thus the products of computation processes. In a computational universe new and unpredictable phenomena emerge as a result of simple algorithms operating on simple computing elements such as cellular automata, and complexity originates from the bottom-up emergent processes. Cellular automata are equivalent to a universal Turing Machine. Wolfram's critics remark, however, that cellular automata do not evolve beyond a certain level of complexity; the mechanisms involved do not produce evolutionary development. Wolfram meets this criticism by pointing out that cellular automata are models and as such surprisingly successful ones. Also Fredkin [20] in his Digital philosophy builds on cellular automata, suggesting that particle physics can emerge from cellular automata. For Fredkin, humans are software running on a universal computer.

Wolfram and Fredkin, in the tradition of Zuse, assume that the universe is, on a fundamental level, a discrete system, and is thus suitably modeled as an all-encompassing digital computer. However, the computing universe hypothesis (natural computationalism) does not critically depend on the discreteness of the physical world, as there are digital as well as analog computers. On a quantum-mechanical level, the universe performs computation on characteristically dual wave-particle objects [21], i.e. both continuous and discrete computing. Maley [22] demonstrates that it is necessary to distinguish between analog and continuous, and between digital and discrete representations. Even though typical examples of analog representations use continuous media, this is not what makes them analog. Rather, it is the relationship that they maintain with what they represent. Similar holds for digital representations. The lack of proper distinctions in this respect is a source of much confusion on discrete vs. continuous computational models.

Moreover, even if in some representations it may be discrete (and thus conform to the Pythagorean ideal of number as a principle of the world), computation in the universe is performed at many different levels of organization, including quantum computing, bio-computing, spatial computing, etc. – some of them discrete, others continuous. So computing nature seems to have a use for both discrete and continuous computation, [23].

4 Information Processing Model of Computation

Computation is nowadays performed by computer systems connected in global networks of multitasking, interacting devices. The classical understanding of computation as syntactic mechanical symbol manipulation performed by an isolated computer is being replaced by the information processing view by Burgin, [24]. Info-computationalism adopts Burgin definition of computation as information processing.

In what follows, I will focus on explaining this new idea of computation, which is essentially different from the notion of context-free execution of a given procedure in a deterministic mechanical way. Abramsky summarizes this changing paradigm of computing as follows:

"Traditionally, the dynamics of computing systems, their unfolding behaviour 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 behaviours, rather than the ***functions which they compute?"*** *[25, p483]*

According to Abramsky, there is a need for second generation models of computation, and in particular there is a need for process models such as Petri nets, Process Algebra, and similar. The first generation models of computation originated from problems of formalization of mathematics and logic, while processes or agents, interaction, and information flow are genuine products of the development of computers and Computer Science. In the second generation models of computation, previous isolated systems with limited interactions with the environment are replaced by processes or agents for which interactions with each other and with the environment are fundamental.

As a result of interactions among agents and with the environment, complex behaviour emerges. The basic building block of this interactive approach is the agent, and the fundamental operation is interaction. The ideal is the computational behaviour of an organism, not mechanical machinery. This approach works at both the macro-scale (such as processes in operating systems, software agents on the Internet, transactions, etc.) and on the micro-scale (from program implementation, down to hardware).

The above view of the relationship between information and computation presented in [25] agrees with ideas of info-computational naturalism of Dodig-Crnkovic [3] which are based on the same understanding of computation and its relation to information. Implementation of info-computationalism, interactive computing (such as, among others, agent-based) naturally suits the purpose of modelling a network of mutually communicating processes/agents, see [3,4,5].

5 Natural Computation

Natural computing is a new paradigm of computing which deals with computability in the natural world. It has brought a new understanding of computation and presents a promising new approach to the complex world of autonomous, intelligent, adaptive, and networked computing that has emerged successively in recent years. Significant for Natural computing is a bidirectional research [7]: as natural sciences are rapidly absorbing ideas of information processing, computing is concurrently assimilating ideas from natural sciences.

The classical mathematical theory of computation was devised long before global computer networks. Ideal, classical theoretical computers are mathematical objects and they are equivalent to algorithms, Turing machines, effective procedures, recursive functions or formal languages. Compared with new computing paradigms, Turing machines form the proper subset of the set of information processing devices, in much the same way as Newton's theory of gravitation presents a special case of Einstein's theory, or Euclidean geometry presents a limited case of non-Euclidean geometries, [5].

Natural/Unconventional computing as a study of computational systems includes computing techniques that take inspiration from nature, use computers to simulate natural phenomena or compute with natural materials (such as molecules, atoms or DNA). Natural computation is well suited for dealing with large, complex, and dynamic problems. It is an emerging interdisciplinary area closely related to artificial intelligence and cognitive science, vision and image processing, neuroscience, systems biology and bioinformatics, to mention but a few.

Computational paradigms studied by natural computing are abstracted from natural phenomena such as self-* attributes of living (organic) systems (including -replication, -repair, -definition and -assembly), the functioning of the brain, evolution, the immune systems, cell membranes, and morphogenesis.

Unlike in the Turing model, where the Halting problem is central, the main issue in Natural computing is the adequacy of the computational response (behaviour). The organic computing system adapts dynamically to the current conditions of its environments by self-organization, self-configuration, self-optimization, self-healing, self-protection and context-awareness. In many areas, we have to computationally model emergence which is not algorithmic according to Cooper [26] and Cooper and Sloman [27]. This makes the investigation of computational characteristics of non-algorithmic natural computation (sub-symbolic, analog) particularly interesting.

In sum, solutions are being sought in natural systems with evolutionary developed strategies for handling complexity in order to improve complex networks of massively parallel autonomous engineered computational systems. Research in theoretical foundations of Natural computing is needed to improve understanding of the fundamental level of computation as information processing which underlies all computing.

6 Information as a Fabric of Reality

"Information is the difference that makes a difference. " [29]

More specifically, Bateson's difference is the difference in the world that makes the difference for an agent. Here the world also includes agents themselves. As an example, take the visual field of a microscope/telescope: A difference that makes a difference for an agent who can see (visible) light appears when she/he/it detects an object in the visual field. What is observed presents a difference that makes the difference for that agent. For another agent who may see only ultra-violet radiation, the visible part of the spectrum might not bring any difference at all. So the difference that makes a difference for an agent depends on what the agent is able to detect or perceive. Nowadays, with the help of scientific instruments, we see much more than ever before, which is yet further enhanced by visualization techniques that can graphically represent any kind of data.

A system of differences that make a difference (information structures that build information architecture), observed and memorized, represents the fabric of reality for an agent. Informational Structural Realism [13] [30] argues exactly that: information is the fabric of reality. Reality consists of informational structures organized on different levels of abstraction/resolution. A similar view is defended in [12]. Dodig Crnkovic [3] identifies this fabric of reality (Kantian 'Ding an sich') as potential information and makes the distinction between it and actual information for an agent. Potential information for an agent is all that exists as not yet actualized for an agent, and it becomes information through interactions with an agent for whom it makes a difference.

Informational structures of the world constantly change on all levels of organization, so the knowledge of structures is only half the story. The other half is the knowledge of processes – information dynamics.

It is important to note the difference between the potential information (world in itself) and actual information (world for an agent). Meaningful information, which is what in everyday speech is meant by information, is the result of interaction between an agent and the world. Meaning is use, and for an agent information has meaning when it has certain use. Menant [31] proposes to analyze relations between information, meaning and representation through an evolutionary approach.

7 Info-computationalism as Natural Philosophy

Info-computationalist naturalism identifies computational process with the dynamic interaction of informational structures. It includes digital and analog, continuous and discrete, as phenomena existing in the physical world on different levels of organization.

Our present-day digital computing is a subset of a more general Natural computing. In this framework, computational processes are understood as natural computation, since information processing (computation) is not only found in human communication and computational machinery but also in the entirety of nature.

Information represents the world (reality as an informational web) for a cognizing agent, while information dynamics (information processing, computation) implements physical laws through which all the changes of informational structures unfold.

Computation, as it appears in the natural world, is more general than the human process of calculation modelled by the Turing machine. Natural computing takes place through the interactions of concurrent asynchronous computational processes, which are the most general representation of information dynamics [5].

8 Conclusions

Alan Turing's work on computing machinery, which provided the basis for artificial intelligence and the study of its relationship to natural intelligence, together with his computational models of morphogenesis, can be seen as a pioneering contribution to the field of Natural Computing and the Computational Philosophy of Nature. Today's info-computationalism builds on the tradition of Turing's computational Natural Philosophy. It is a kind of epistemological naturalism based on the synthesis of two fundamental cosmological ideas: the universe as informational structure (informationalism) and the universe as a network of computational processes (pancomputationalism/naturalist computationalism).

Information and computation in this framework are two complementary concepts representing structure and process, being and becoming. Info-computational conceptualizations, models and tools enable the study of nature and its complex, dynamic structures, and uncover unprecedented new possibilities in the understanding of the connections between earlier unrelated phenomena of non-living and living nature [28].

Acknowledgements. The author would like to thank Gerard Jagers Op Akkerhuis, Christophe Menant and Shuichi Kato for helpful comments.

References

1. Hodges, A.: Turing. A Natural philosopher. Phoenix, London (1997)
2. Turing, A.M.: The Chemical Basis of Morphogenesis. Philosophical Transactions of the Royal Society of London 237(641), 37–72 (1952)
3. Dodig-Crnkovic, G.: Information and Computation Nets. Investigations into Info-computational World. In: Information and Computation, pp. 1–96. Vdm Verlag, Saarbrucken (2009)
4. Dodig-Crnkovic, G.: The Info-computational Nature of Morphological Computing. In: Müller, V.C. (ed.) Theory and Philosophy of Artificial Intelligence. SAPERE, vol. 5, pp. 59–68. Springer, Heidelberg (2012)

5. Dodig-Crnkovic, G.: Significance of Models of Computation from Turing Model to Natural Computation. Minds and Machines 21(2), 301–322 (2011)
6. Rozenberg, G., Kari, L.: The many facets of natural computing. Communications of the ACM 51, 72–83 (2008)
7. Landweber, L.F., Kari, L.: The evolution of cellular computing: nature's solution to a computational problem. Biosystems 52(1-3), 3–13 (1999)
8. Van Hornweder, K.: Models and Mechanisms of the Morphogenesis of Biological Structures (Technical Report UT-CS-11-681) (2011), ftp://cs.utk.edu/pub/TechReports/2011/ut-cs-11-681.pdf
9. Denning, P.: Computing is a natural science. Comm. ACM 50(7), 13–18 (2007), http://cs.gmu.edu/cne/pjd/PUBS/CACMcols/cacmJul07.pdf
10. Kurakin, A.: Scale-free flow of life: on the biology, economics, and physics of the cell. Theoretical biology & medical modelling 6(6) (2009)
11. Kurakin, A.: The self-organizing fractal theory as a universal discovery method: the phenomenon of life. Theoretical Biology and Medical Modelling 8(4) (2011), http://www.tbiomed.com/content/8/1/4
12. Ladyman, J., Ross, D., Spurrett, D., Collier, J.: Everything must go: metaphysics naturalised. Clarendon Press, Oxford (2007), http://www.worldcat.org/title/everything-must-go-metaphysics-naturalised/oclc/84150765&referer=brief_results
13. Floridi, L.: A defense of informational structural realism. Synthese 161(2), 219–253 (2008)
14. Floridi, L.: Against digital ontology. Synthese 168(1), 151–178 (2009)
15. Wolfram, S.: A New Kind of Science. Wolfram Media (2002)
16. Zenil, H.: Randomness Through Computation: Some Answers, More Questions. World Scientific Pub. Co. Inc., Singapore (2011)
17. Dodig Crnkovic, G., Hofkirchner, W.: Floridi's "Open Problems in Philosophy of Information", Ten Years After. Information 2(2), 327–359 (2011)
18. Zuse, K.: Rechnender Raum. Friedrich Vieweg & Sohn, Braunschweig (1969)
19. Chaitin, G.: Epistemology as Information Theory: From Leibniz to Ω. In: Dodig Crnkovic, G. (ed.) Computation, Information, Cognition – The Nexus and The Liminal, pp. 2–17. Cambridge Scholars Publishing, Newcastle (2007)
20. Fredkin, E.: Finite Nature. In: XXVIIth Rencotre de Moriond (1992)
21. Lloyd, S.: Programming the universe: a quantum computer scientist takes on the cosmos, 1st edn. Knopf, New York (2006)
22. Maley, C.J.: Analog and digital, continuous and discrete. Philos. Stud 155, 117–131 (2010)
23. Lesne, A.: The discrete versus continuous controversy in physics. Mathematical Structures in Computer Science 17, 185–223 (2007)
24. Burgin, M.: Super-Recursive Algorithms. Monographs, pp. 1–320. Springer-Verlag New York Inc., New York (2004)
25. Abramsky, S.: Information, Processes and Games. In: Benthem van, J., Adriaans, P. (eds.) Philosophy of Information, pp. 483–549. North Holland, Amsterdam (2008)
26. Cooper, S.B., Löwe, B., Sorbi, A.: New Computational Paradigms. In: Cooper, S.B., Löwe, B., Sorbi, A. (eds.) Changing Conceptions of What is Computable. Springer Mathematics of Computing series, vol. XIII. Springer (2008)
27. Dodig-Crnkovic, G., Burgin, M.: Information and Computation. World Scientific Pub. Co. Inc., Singapore (2011)

28. Dodig-Crnkovic, G., Müller, V.: A Dialogue Concerning Two World Systems: Info-Computational vs. Mechanistic. In: Dodig Crnkovic, G., Burgin, M. (eds.) Information and Computation, pp. 149–184. World Scientific (2011)
29. Bateson, G.: Steps to an Ecology of Mind, pp. xxv–xxvi. Ballantine, NY (1972)
30. Sayre, K.M.: Cybernetics and the Philosophy of Mind. Routledge & Kegan Paul, London (1976)
31. Menant, C.: Computation on Information, Meaning and Representations. An Evolutionary Approach. In: Dodig Crnkovic, G., Burgin, M. (eds.) Information and Computation. World Scientific (2011)

Dualism of Selective and Structural Manifestations of Information in Modelling of Information Dynamics

Marcin J. Schroeder

Akita International University, Akita, Japan
mjs@aiu.ac.jp

Abstract. Information can be defined in terms of the categorical opposition of one and many, leading to two manifestations of information, selective and structural. These manifestations of information are dual in the sense that one always is associated with the other. The dualism can be used to model and explain dynamics of information processes. Application of the analysis involving selective-structural duality is made in the contexts of two domains, of computation and foundations of living systems. Similarity of these two types of information processing allowing common way of their modelling becomes more evident in the naturalistic perspective on computing based on the observation that every computation is inherently analogue, and the distinction between analogue and digital information is only a matter of its meaning. In conclusion, it is proposed that the similar dynamics of information processes allows considering computational systems of increased hierarchical complexity resembling living systems.

Keywords: Selective and structural information, Dynamics of information processing, Hierarchic levels of information.

1 Introduction

The concept of information has several very different definitions. In this large variety, only few qualify as correct and intelligible. Too frequently, definitions simply refer to intuitive understanding of the explanatory concepts selected from the vernacular vocabulary. It is quite rare that the formulation of the definition refers to any particular philosophical background. However, there are two clearly distinctive or even competitive tendencies in the understanding of information. One is characterized by explicit or implicit reference to selection, sometimes in alternative form of difference or distinction. The other has the general idea of the form or structure as the focal point of explanation.

The definition of information used in this paper was introduced and extensively analyzed in earlier articles of the author. Its desirable feature is that the both ideas of selection and of structure can be found as alternative and complementary ways of its interpretation.

Moreover, it turns out that the selective and structural manifestations of information are dual in the sense that one always is associated with the other. The dualism is

G. Dodig-Crnkovic and R. Giovagnoli (Eds.): *Computing Nature*, SAPERE 7, pp. 125–137.
DOI: 10.1007/978-3-642-37225-4_7

being used in present article to model and explain dynamics of information processes. Dynamical processes of this type are analyzed in contexts of the two domains, of computation and foundations of living systems, but there is nothing which would limit this model to any particular domain. In conclusion, it is suggested that the similar dynamics of information processes allows considering computational systems of increased complexity resembling living systems.

Due to the scope and limitation of the format of this paper more detailed presentation of the technical issues related to mathematical theory of information developed by the author for the description of the dual concept of information and of information dynamics will be published elsewhere.

2 Dualism of Selective and Structural Information

The concept of information is understood here in the way it was defined in earlier papers of the author [1] as an identification of a variety. Thus, starting point in the conceptualization of information is in the categorical opposition of one and many.

The variety in this definition, corresponding to the "many" side of the opposition is a carrier of information. Its identification is understood as anything which makes it one, i.e. which moves it into or towards the other side of the opposition. The preferred word "identification" (not the simpler, but possibly misleading word "unity") indicates that information gives an identity to a variety, which does not necessarily mean unification, uniformization or homogeneization. However, this identity is considered an expression of unity or "oneness".

There are two basic forms of identification. One consists in the selection of one out of many in the variety (possibly with limited degree of determination which element of the variety is selected), the other in a structure binding many into one (with different degrees of such binding). This brings two manifestations of information, the selective and the structural. The two possibilities are not dividing information into two types, as the occurrence of one is always accompanied by the other, but not on the same variety, i.e. not on the same information carrier. For instance, information used in opening a lock with the corresponding key can be viewed in two alternative ways. We can think about it as a way to make the selection of the key, out of some variety of keys, or we can think about the spatial structure of the key which fits the structure of the lock. In the first case, the variety consists of a collection of keys, in the second the variety consists of the material units (for instance molecules) forming appropriate geometric shape of the key. It can be easily observed that the varieties in this example are related hierarchically. Every element of one variety (keys) is an instance of the other (molecules to be bound into a key). Thus, we can consider selective and structural information as dual manifestations of one concept, with the duality related to objective, structural characteristics of reality.

Coexistence of different manifestations of information justifies introduction of the concept of an information system understood as a complex of varieties (information carriers) whose forms of identification are pair-wise combined through selective-structural duality. Going beyond a pair of information carriers will be considered later in the context of systems in which a hierarchic chain of related pairs can be identified, as for instance in living systems.

As mentioned above, the identification of a variety may differ in the degree. For the selective manifestation this degree can be quantitatively described using appropriate probability distribution and measured using for instance Shannon's entropy, or more appropriate measure when we want to characterize information within the system, not its transmission between systems [2]. For the structural manifestation the degree can be characterized in terms of decomposability of the structure [3].

Selective-structural duality of information is reflected in a variety of contexts. An example of very general character can be found in the way how we form concepts. One way is focusing on the denotation and the selection of objects which we want to include in denotation. Another way is to focus on the connotation determined by the configuration of characteristics which describe it.

Another example can be found in the analysis of scientific or philosophical inquiry. In his philosophical analysis of the methods of science and history Wilhelm Windelband [4] introduced frequently revoked distinction, or even opposition of nomothetic and idiographic methodologies. The former has its starting point in the acknowledgement of the differences, but assumes the existence of similarities which produce grouping within the variety, and therefore it is looking for comparable aspects and serves identification of the subject of study. The latter is assuming the uniqueness of the object of study and therefore is focused on elements which constitute this uniqueness through specific structural characteristics. Although, the distinction is between methodologies of inquiry, not between manifestations of information, association with information is quite evident.

Similar, but much more frequently used distinction in the context of cultural studies has been introduced more than a half century later by Keneth L. Pike [5]. He called his methodological schemata etic and emic methodologies, deriving their names from phonetic and phonemic studies of language. Here too, the distinction is based on the differences in the perspective of the study. In the first case the subject of study is viewed in a comparative manner as a member of a variety in which differences and similarities are used to establish its unique characteristics. In the second case, the subject of the study, whose uniqueness is already assumed, is viewed from the inside with the aim to reconstruct its internal structure.

In these examples, as well as in all instances of the reflection of the selective-structural duality in methodological analysis, it is considered obvious that the choice of a particular method is dictated by the discipline of inquiry. Physics for instance is recognized always as a paradigm of the nomothetic or ethic approach corresponding to selective information. After all, probability distributions describe the state of a system, collective one in classical physics, and individual in quantum physics. But closer look reveals that actually in this domain both methodological positions are omnipresent. It is enough to recall tendency of geometrization in physics continuing beyond the General Relativity Theory, or the special role of the field theory to recognize the presence of the view associated with structural information.

The selective-structural dualism of information can be found not only in the distinction of methodological perspectives in physics. Wave-particle dualism which is understood as a characteristic of physical reality at quantum mechanical level can be interpreted as an expression of the dualism of selective and structural manifestations of information. Corpuscular image of an electron is based on the selection of its position out of a variety of possibilities described by a probability distribution. Wave image is based on the structural characteristics of the space.

The most significant is association of the selective-structural dualism of information with the dualism of function and structure in the foundation studies of living systems, which constitutes the central theme of the work of Humberto Maturana and Francisco Varela [6] on autopoiesis. Here it becomes clear that this dualism is not just a matter of the choice of a method of inquiry, but it is a characteristic of living systems. Function determines structure and structure determines function. Maturana and Varela were looking for the resolution of this convolution in autopoiesis, self-construction of living systems. However, from the point of view of information studies, there is no need to restrict this dualism to living systems, as it is simply reflection of the universal dualism of selective and structural information. Functions of the elements of a system give them identity by distinguishing them from, and giving them their place in the differentiated variety. On the other hand, this distinction is a consequence of the specific structural characteristics that they posses, their internal structure allows them to play specific roles in the system. It is not a matter of the right or wrong perspective of the study, but an inherent feature of all information systems.

Mathematics provides several different examples of dualism which can be very clearly associated with that of selective and structural information. The most fundamental can be traced back to the 19th Century when Felix Klein formulated in his 1872 Erlangen Program the view of geometry as a theory of invariants for the group of transformations of a geometric space. Instead of identification of the objects of geometric studies through analysis of their internal structure, the structure of transformations of the plane or space is selected, and only then geometric objects appear as those subsets of points which are transformed into themselves, although their points may be exchanged. Such an approach, in which instead of inquiry of internal structure of objects, the structure of transformations preserving the identity of these objects (i.e. selection of invariants) is analyzed, has become commonly used in a wide range of mathematical theories leading to the development of the theory of categories and functors.

In the past, the dualism of selective and structural information has been present in information studies only in the form of a competition between two, apparently conflicting views on the "proper" answer to the question "What is information?" [1]. The dominating position focusing on the selective manifestation of information and neglecting the structural one was supported by the practical success of Shannon's quantitative characterization of information in terms of entropy. But the failure in establishing equally successful semantics for information understood exclusively in terms of selection was driving the efforts to shift studies of information to its structural manifestation.

The dual approach achieved through the definition of information used in the present paper has more advantages than just reconciliation between adherents of competing views on information. It also helps to model dynamics of information in processes of evolution or computation.

3 Dynamics of Information in Computing

The definition of information in terms of one-many opposition has been a starting point for author's attempt to formulate a theoretical framework for information [7]. This framework has a static form reminding logical structure, at least in the sense of similarity of the mathematical formalisms. However, the formalism used by the author can be used to model process of information integration which can be interpreted in terms of temporal orientation (input/output) [3].

The change of the level of information integration is not a dynamical process, understood as transformation resulting from the interaction of different information systems. For this reason, information integration, although modelled by a theoretical device called a Venn gate in the earlier papers of the author should not be confused with traditionally understood computation.

What is computation in the present conceptual framework? First, we have to clarify some quite common confusion related to the distinction between analogue and digital computing. The distinction between "analogy and digital" principles, automata, or machines introduced by John von Neumann [8] at the time when first computers were being constructed was referring to the way the numbers are represented, by certain physical quantity, or by "aggregates of digits."

For von Neumann the main issue here was in handling errors. He wrote "Thus the real importance of the digital procedure lies in its ability to reduce the computational noise level to an extent which is completely unobtainable by any other (analogy) procedure."

Of course, von Neumann was right about practical advantages of "digital procedure" in handling errors, but he overlooked what actually constitutes the distinction and why it is important outside of practical considerations of precision. The mistake he made is being perpetuated even now. Of course, the numbers are always represented by physical quantities, even in digital computers. For instance, the typical implementation of computing units associates digit 1 with one physical state and 0 with another physical state. But it is only an interpretation of the distinction between two physical states. Moreover, the positional numerical system used in this interpretation is not based on aggregation of digits, but on very specific and conventional structural rules. "Aggregates of digits" do not exist independently from the physical systems constituting machines or any other computing systems. To that extent everyone will agree with Ralph Landauer [9] that information is physical.

Thus, the actual distinction is in the semantics of information. It is the way how we associate numbers with physical states of the computing machine which decides whether computing is digital or analogue. Information itself is neither one, nor the other. Cat is not becoming more English, when described with the English word "cat".

To avoid going too far beyond the scope of this paper, simplifying assumption will be made that information is associated with the state of the physical system which is used as a computing machine. Then, observables will assign numbers to particular states, giving meaning to information, but we have to remember the lesson from quantum mechanics making clear distinction between the concepts of a state and an observable [10]. As a consequence, every process of computing is a physical process with some dynamic characteristics. Association of numbers with the states of the computing system belongs to the interpretation of information, the same way as in physics observables provide numerical interpretation of the states of a physical systems. Numerical values of observables cannot be identified with states, and therefore cannot be identified with information involved in the process of computing. The same applies to the choice of the numerical system used to represent numbers.

Recognition of the fact that every computation is being carried by some physical information systems justifies the interest in its description as a dynamical process. However, the dynamics of computation does not have to be understood in traditional terms of mechanics. Physicality of computation is just a matter of the ontological status of information systems involved.

We can find some analogy with the status of the Second Law of Thermodynamics, in its interpretation introduced by Boltzmann. We can apply this principle to every sufficiently complex system without any reference to standard physical observables. However, its validity requires that this complex system has the ontological status of a part of the physical reality with all its consequences.

Now, when a justification of our naturalistic perspective is presented we can begin analysis of the process of computing modelled by Turing machines. Once again we have to be careful with traditional way of imagining of the process. Traditional vision of computing is similar to the way people were interpreting mechanical processes before Isaac Newton introduced his Third Principle of Mechanics. In pre-Newtonian vision of the world, every change had to have an active agent (subject) and passive object of the action. Newton recognized that in mechanical phenomena there is no action, but only interaction. The Third Principle states that we cannot distinguish between an agent and recipient of action, as we have always mutual interaction. I cannot claim that my pushing the wall is in any way different from wall's pushing me, as long as we analyze it in terms of mechanics.

From this point of view the interpretation of a head in Turing machine printing a character on the tape is an arbitrary assumption. If we want to consider process of computing in a naturalistic perspective, we can simply talk about mutual interaction in which characters change (or not) on the tape in contact with the head, and the head is changing its state/instruction in contact with the tape. In literally understood physical model of Turing machine, the change of the head may be negligible. But in general we cannot exclude this change from consideration.

More precisely, we could describe Turing machine as a device consisting of two information systems, which in order to retain traditional terminology are called a "tape" and a "head", each consisting of independent components being themselves information (sub)systems. For the tape, components are cells. For the head, subsystems are positions of instructions on the list. At every moment both systems have finite, but unlimited number of engaged components (non-empty cells, or non-empty instructions), and the number of engaged components can grow without restriction.

Each component (cell or position on the list of instructions) is capable to assume one of the finite number of states (possibly different for the components of the tape and the components of the head). For cells on the tape the states are characters in the traditional description of Turing machine. For components of the head (positions on the list of instructions), there is a finite number of choices for an instruction which give the position particular state. Also, we can assume that in the initial step of computation only finite number of positions have nonempty instructions.

Now, we have a crucial and restrictive assumption that these two fundamental information systems can interact only by the contact or interaction of a single pair of active components (which corresponds to the traditional assumption that the head is in the state with one particular instruction, and it can read and act on a single cell).

Experience from the studies of Turing machines suggests that the assumption is not restrictive as long as the difference between one pair of active components is contrasted with clearly defined finite number of pairs. The restrictive character appears when we exclude the possibility of interaction on the scale of all systems.

The process of computing is described as follows. The active cell is changing (or not) its state (character) into one determined by the state of the active component of the head (particular instruction in the position on the list for given state). On the side of the head the change of the instruction depends on the state of the cell (character in the cell). Then both fundamental information systems change their active component. Again this change on the tape depends on the state of active component of the head, the change in the head depends on the state of active cell (character).

Thus, the dynamics of computation considered as an interaction of two information systems consists in the change of current states of both active components, that of the tape and that of the head. The change is a mutually conditioned selection. Also change of the choice of active pair of components is similarly cross related. The crucial point is that the interaction acts as a new information system which cannot be reduced to interacting systems. The variety involved consists of all possible pairs of states which can be selected as an outcome of the step in computation. Another, independent information system consists of all possible selections of the pair of next active pair.

In traditional description of a Turing machine the information regarding dynamics of the process (how components are changing and what the choice of next pair of active components is) is "physically" located in the head or on the tape. For each step of computing, it is located within the instruction as a conditional statement of doing something, if the current tape cell has given state. However, there is nothing that compels us to such model. Equally well we can think that the instruction has form of a character, and what is happening with the tape is a result of the reaction of the tape's active cell to this character, and of selection of the next pair of active components activated externally but conditioned by the states of the pair of present active components.

This machine is little bit more general than Turing's A-machine, as the process allows changes of instructions in the head. This machine could be called a symmetric machine (an S-machine) because the process consists in mutual interaction producing similar type of change. It is being reduced to usual Turing A-machine, if we additionally assume that the instructions in the head are not changing. Of course, this assumption is making Turing A-machine a special case of an S-machine.

There are several natural questions regarding this generalization. For instance, whether for every S-machine there exists an equivalent A-machine producing the same outcome on the tape after computation performed on arbitrary input tape. Another example would be the question about universal S-machines (machine which can produce arbitrary global finite state of the tape, by appropriate choice of the initial global state of the tape, but without any change of the state of the head. However, for general S-machines we have also dual questions regarding configuration of instructions after computation or the minimal number of instructions which produce the same outcome of computation.

At this point we can observe that as long as we are interested in the relationship between computation and fundamental characteristics of life (or living objects), in contrast to traditional studies of computing, it is non-computability which is of special interest. If living objects perform some process of computation, achieving the final stage of computation is a death of the system. Thus, sustainability of life is more likely to be associated with non-computability. However, this issue is outside of the scope of the present paper, since we are more interested in similarities between the two domains, than differences.

For the symmetric Turing machines describing a general dynamic process of the interaction of a pair of complex systems with a restricting assumption that the interaction is in each moment through exactly one pair of active components (mild restriction), we can consider additional distinction between deterministic and nondeterministic machines. The distinction is based on the requirement that the choice of the next pair of active components is strictly determined by the states of the present active components, not random or determined only up to some probability distribution (rather strong restriction).

Even with these two restrictions, symmetric Turing machine gives us a model of information dynamics applicable to a very wide range of information systems.

We know that computation cannot be reduced to one information system. Claude Shannon [11] showed that the head of Turing machine has to have at least two different states. Similar requirement of at least two characters for the tape is obvious. Once we have a variety of two states and choice between them, we have an information system.

Now, the dynamics of the process of computation is revealed in the selective-structural dualism of information. For both fundamental information systems (tape and head considered globally) information is structural. The state of all tape consists of configuration of characters in its cells, but computation is an interaction in which the choice of one out of many states (characters) for the active component (cell) is being made. Similarly, the state of the head is in the configuration of instructions, but in each step of computation one out of many possible choices of instruction is being made. The selection of states and active components is shaping the global structures of the tape and of the head. However, process of local selection is dependent on the global structural characteristics of the tape and the head.

Finally, we could consider an extension of the process of computation using the concept of selective-structural information dualism. While computation considered at the level of active, interacting pair of components refers to the selective manifestation of information (e.g. selection of a character for the cell), each character can be understood as structural manifestation of information, if we can decompose it into a variety of elements with some structure. Corresponding to this structural manifestation, its selective counterpart can be subject to interaction which results in its own dynamics. This way we can consider multi-level symmetric Turing machines, which resemble systems encountered in the study of the foundations of life.

4 Dynamics of Evolution

Before we enter the analysis of evolutionary mechanisms, it is necessary to consider more general issue of control systems. In this domain the most fundamental principle has been formulated by W. Ross Ashby as the Law of Requisite Variety "A model system or controller can only model or control something to the extent that it has sufficient internal variety to represent it" [12], [13]. This principle in the informal, intuitive form and in application to the process of generation, not to the modelling or controlling has been until the end of the 18th Century used as an argument for the hierarchy of beings and the need for supremely intelligent creator acting intentionally to generate them [14].

It seemed obvious that any complex system can be generated only by a system of higher level of organization. This reasoning is based on the assumptions that creation is an action (not interaction) and requires a design. Following the Law of Requisite Variety such a design, i.e. internal model is impossible without higher degree of variety. Evolutionary model of the development of life disposed of the design putting this higher level of variety in the environment. Thus the species are getting increasingly complex by the interaction with the environment, which of course is a carrier of a huge amount of information.

Let's start from a dualistic model of relatively simple mechanism of feedback control. It requires interaction of two information systems. Selection of a state of one of them through interaction is accompanied by the selection of a state of the other, which in turn has its reflection in the structural manifestation of information. This structural manifestation of information in the second system is determining the structural information of the first system. And this corresponds to the modification of the selection of its state.

For instance, using classical example of a governor controlling work of the steam machine, we have two information systems which can be in a variety of states. One is a valve whose state (described by the amount of steam passing through it) decides about the speed of the work of the machine. The other is a pair of balls hanging on the arms rotating around the vertical axis whose rotation is propelled by the machine. Its state (velocity of rotation) is selected by the work done by machine. From the structural point of view, information is manifested by the geometric structures of the systems, diameter of the valve and extension of the arms on which the balls are attached. The higher is extension of arms, the smaller diameter of the valve.

Interaction between the two information systems is as follows. Choice or selection of the amount of steam is determining the choice of the velocity of rotation. But velocity of rotation corresponds to the structural information regarding position of the balls. Position of the balls (structural information) is determining the structural characteristics of the valve. And finally this structural information corresponds to the selective manifestation in form of the amount of steam flowing through the valve.

The governor is a simple case of an artefact invented by humans, originally with the intention to control the speed of work of windmills. There is more complicated situation when we want to explain the dynamics of information in systems which were naturally generated without any intentional design.

We can proceed to the dualistic description of the evolutionary process. Here, in distinction from the earlier example where the function was a result of human invention and the structure followed the needs of implementation, we can encounter confusion which puzzled generations of biologists, but which can be easily resolved within the dualistic perspective.

The mechanism of evolution is usually reduced to natural selection in which the fittest organisms survive and reproduce transmitting and perpetuating their genetic information. The puzzling question is about the meaning of the term “fittest”. Does it have any other meaning beyond the tautological statement that these are organisms which survived and reproduced?

The answer is that the meaning of the term “fittest” is expressing the relationship between two manifestations of information. While naturally, natural selection describes the dynamics of information for selective manifestation in terms of reproduction (which obviously requires survival), the fittest individuals are those whose phenotype has structural characteristics compatible with structural characteristics of the environment.

More generally, we can describe the evolutionary process as such in which two (or more) information systems interact. Interaction is determining the outcome of the selection, and therefore the dynamical view seems more natural in terms of the selective manifestation. However, it is the structural manifestation of information which actually demonstrates the results of evolution. And what is most important, there is no point in asking which manifestation is more important, primary, or true. Dynamics of information has two manifestations, simply because information does.

5 Dynamics of Information in Living Systems

Thus far we were talking about biological evolution of species as a dynamical information process. We were concerned with the question how this process can be understood. There is another, much more difficult question why it occurs, and why in this particular way. To seek the answer, we have to consider more general issue of the dynamics of information in the living systems. Naturally, it is equivalent to the inquiry regarding the question “What is life?” We will consider here only some aspects of this extremely broad and deep problem. Specifically those related to the selective-structural dualism of information. The issues related to the necessity of holistic methodology in the study of life are presented in another article of the author [15].

The main fallacy in answers to the question "What is life?" is in the attempt to explain life by distinguishing one particular process driving and determining all other in the multi-level hierarchical structure of the biosphere. This fallacy is being perpetuated even in most recent publications [16]. The process chosen by the authors of explanations could be photosynthesis (but, what about forms of life which do not depend on it?), metabolism, reproduction with transmission of genetic information, formation of large organic molecules, etc. In each case, authors believe that life can be reduced to one particular level of organization, in analogy to mechanical systems built from basic subcomponents or to the vision of the world built from fundamental particles ("atoms") through their aggregation.

Another problem is in the restriction of attention to what is called a biosphere. In the earliest fundamental answer to the question Erwin Schrödinger [17] pointed at what he called negative entropy of the light coming from sun as the factor driving processes of life. It is also a fallacy perpetuated continuously by generations of authors who change the name of the factor (negentropy, entropy deficit, inhomogeneity, etc.) but do not notice that the light coming to earth does not have high or low entropy. It is a matter of the process in which incoming visible light, for which the atmosphere is transparent, is transformed by living systems and reradiated into cosmic space as infrared radiation of 22 times higher entropy [18]. Thus, it is not that light coming to Earth has low entropy, but that we have complex process which is making this entropy low relative to the outcoming radiation. There is nothing which prevents this infrared radiation to drive processes of life somewhere else, if re-radiated from there longer-wave outcoming radiation could have entropy several times higher than radiation coming from Earth.

Thus, the driving factor is a mechanism which transcends biosphere and which has its source in astronomical phenomena of huge spatial and temporal measures. But this driving factor itself would not produce life processes. It is just a necessary condition for life. It creates conditions allowing generation of information participating in the dynamic processes of life at all of its levels. Life cannot be understood by observing only one of these levels, as it is usually done. We can artificially generate processes from one level in a system of limited complexity, but they cannot continue functioning independently, and this lack of ability to survive excludes considering such a system as living.

Of course, evolution of species, cycles of metabolism, photosynthesis, or reproduction are component processes of life. But neither has privileged or exclusive position. We can ask however about the common features for component processes of life. Here we can find again help in the dualistic perspective on information, and the concept of information definitely is the best candidate to unify description of all life processes.

The main characteristic of life processes consists in enriching information in one system of a smaller variety, i.e. lower informational volume through the dynamic interaction with another of a large volume. This process was already described in a general view of the dynamics of information in the evolution of species. We need in this case generation of a large variety of objects and interaction with the other system which selects some of them (the fittest) whose structural characteristics predestine them to survive. Thus, the collective system is increasing its organization (internal information) not because they have some design, but they fit selective information of the outer system. The crucial point is in inseparable dualism between the two manifestations of information and multi-level character of the total system. The multiple levels can be identified at intra-organismal and at inter-organismal side of the organization of life. Typical approach in determination of the levels is the use of either functional (selective) aspect of bio-dynamics described above in the context of the work of Maturana and Varela, or structural characteristics. However, we should be aware of their dual relationship.

6 Conclusions and Future Work

There are two domains of special interest where dualism of selective and structural information can be used to model dynamics of information, computation and living systems. Although in both cases dynamics is similar, there is a big contrast between the levels of complexity between them. In what here was described as a slightly more general view of Turing machines there are two information systems (tape and head) which are considered at the two levels corresponding to selective and structural information.

Life constitutes an extremely complex system of at least dozens of levels and the number of component information systems exceeding any practical limits of calculation. However, the basic mechanism involving in its description the two manifestations of information is the same as in symmetric Turing machines.

On the other hand, there is nothing which prevents us from designing computational systems of complexity going beyond two levels. This may require more complicated (multilevel) semantics of information (which in traditional Turing machines is typically an association of particular combination of the states of cells with natural numbers). Each cell may be considered a carrier of an information system with its own variety and with dynamical mechanisms of evolution adjusted to the conditioning by higher or lower levels of the hierarchical structure.

The study of such theoretical systems and their practical implementation is of some interest and has a potential wide range of applications.

References

1. Schroeder, M.J.: Philosophical Foundations for the Concept of Information: Selective and Structural Information. In: Proceedings of the Third International Conference on the Foundations of Information Science, Paris (2005), http://www.mdpi.org/fis2005
2. Schroeder, M.J.: An Alternative to Entropy in the Measurement of Information. Entropy 6, 388–412 (2004)
3. Schroeder, M.J.: Quantum Coherence without Quantum Mechanics in Modeling the Unity of Consciousness. In: Bruza, P., Sofge, D., Lawless, W., van Rijsbergen, K., Klusch, M. (eds.) QI 2009. LNCS, vol. 5494, pp. 97–112. Springer, Heidelberg (2009)

4. Windelband, W.: Praeludien: Geschichte und Naturwissenschaft, vol. II, pp. 136–160. Tubingen (1924); Engl. transl. Lamiell, J. T.: History and Natural Science. Theory and Psychology 8, 6–22 (1998)
5. Pike, K.L.: Emic and Etic Standpoints for the Description of Behavior. In: Pike, K.L. (ed.) Language in Relation to Unified Theory of the Structure of Human Behavior, Mouton, The Hague (1966)
6. Maturana, H.R., Varela, F.J.: Autopoiesis and Cognition: The Realization of the Living. Boston Studies in the Philosophy of Science, vol. 42. Reidel, Dordrecht (1980)
7. Schroeder, M.J.: From Philosophy to Theory of Information. International Journal Information Theories and Applications 18(1), 56–68 (2011)
8. von Neumann, J.: The General and Logical Theory of Automata. In: Taub, A.H. (ed.) John von Neumann, Collected Works, vol. V. Pergamon Press (1948)
9. Landauer, R.: Information is Physical. Physics Today. May 23-29 (1991)
10. Jauch, J.M.: Foundations of Quantum Mechanics. Addison-Wesley, Reading (1968)
11. Shannon, C.E.: A Universal Turing Machine With Two Internal States. In: Shannon, C.E., McCarthy, J. (eds.) Automata Studies, pp. 157–165. Priceton University Press, Princeton (1956)
12. Ashby, W.R.: An Introduction to Cybernetics. Chapman & Hall, London (1956)
13. Heylighen, F.: Principles of Systems and Cybernetics: an evolutionary perspective. In: Trappl, R. (ed.) Cybernetics and Systems 1992, pp. 3–10. World Science, Singapore (1992)
14. Schroeder, M.J.: Concept of Information as a Bridge between Mind and Brain. Information 2(3), 478–509 (2011), http://www.mdpi.com/2078-2489/2/3/478/
15. Schroeder, M.J.: The Role of Information Integration in Demysti-fication of Holistic Methodology. In: Simeonov, P.L., Smith, L.S., Ehresmann, A.C. (eds.) Integral Biomathics: tracing the Road to Reality. Proceedings of iBioMath 2011-Am, San Jose, CA, USA, iBioMath 2011-Eu, Paris, France and ACIB 2011, Stirling, UK, pp. 79–96. Springer, Berlin (2012)
16. Sato, N.: Scientific Élan Vital: Entropy Deficit or Inhomogeneity as a Unified Concept of Driving Forces of Life in Hierarchical Biosphere Driven by Photosynthesis. Entropy 14, 233–251 (2012)
17. Schrödinger, E.: What is Life? Cambridge University Press, Cambridge (1945)
18. Schroeder, M.J.: Entropy Balance in Climate Changes. In: Proceedings of the 44-th Coal Science Symposium, Akita (October 2007), http://ci.nii.ac.jp/naid/110006611099/

Intelligence and Reference

Formal Ontology of the Natural Computation

Gianfranco Basti

Faculty of Philosophy. Lateran University, Rome, Italy
basti@pul.it

Abstract. In a seminal work published in 1952, “The chemical basis of morphogenesis”, A. M. Turing established the core of what today we call “natural computation” in biological systems, intended as self-organizing dissipative systems. In this contribution we show that a proper implementation of Turing’s seminal idea cannot be based on diffusive processes, but on the coherence states of condensed matter according to the dissipative Quantum Field Theory (QFT) principles. This foundational theory is consistent with the intentional approach in cognitive neuroscience, as far as it is formalized in the appropriate ontological interpretation of the modal calculus (formal ontology). This interpretation is based on the principle of the “double saturation” between a singular argument and its predicate that has its dynamical foundation in the principle of the “doubling of the degrees of freedom” between a brain state and the environment, as an essential ingredient of the mathematical formalism of dissipative QFT.

Keywords: Morphogenesis, quantum field theory, self-organizing systems, dissipative structures, double saturation, degrees of freedom doubling, chaotic trajectory, chaotic trajectory, cognitive neuroscience.

1 Introduction

1.1 Natural Computation and Algorithmic Computation

Today the *Natural Computation* (NC) is considered as an alternative paradigm to the *Algorithmic Computation* (AC) paradigm in natural sciences and in computer sciences, being the paternity of only the latter one generally ascribed to Alan Mathison Turing (1912-1954) pioneering work. On the contrary, after the publication of his famous seminal paper on algorithmic computation in 1936 [1] based on the notions of Turing Machine (TM) and Universal Turing Machine (UTM), Turing worked for widening the notion of “computation” in the direction of what today we define as “natural computation”.

Before all, he defined the notion of Oracle-machine(s)[1] and of their transfinite hierarchy, in his doctoral work at Princeton, under the Alonso Church supervision, published in 1939 [2].

[1] I.e., a TM enriched with the output of operations not computable by a TM, endowing the TM with the primitives of its computable functions.

G. Dodig-Crnkovic and R. Giovagnoli (Eds.): *Computing Nature*, SAPERE 7, pp. 139–159.
DOI: 10.1007/978-3-642-37225-4_8

Afterward, in 1947, in a lecture given at the *London Mathematical Society* [3], and hence in an unpublished communication for the *National Physical Laboratory* in 1948 [4], he sketched the idea of computational architectures made by undefined interacting elements, that can be suitably trained, so to anticipate the so-called Artificial Neural Networks (ANN) computational architectures.

Finally, in a much more known contribution on a new mathematical theory of morphogenesis, published in 1952 [5], Turing was the first who studied a model of pattern formation via non-linear equations, in the specific case of chemical reaction-diffusion equations simulated by a computer.

This pioneering work on non-linear systems, and their simulation via computers, is, indeed, among all the pioneering works of Turing, the most strictly related with the new paradigm of NC, because of its wide field of application in practically every realm of mathematical and natural sciences, from cosmology and fundamental physics, to thermodynamics, chemistry, genetics, epigenetics, biology, and neurosciences; but also in human sciences, from cognitive and social sciences, to ecology, to economical sciences, to linguistics, …, and wherever a mathematical modeling of empirical data makes sense.

In a recent paper devoted to illustrate the new paradigm of NC in relationship with the old paradigm of AC [6], G. Dodig-Crnkovic emphasizes the main differences between the two paradigms that can be synthesized according to the following, main dichotomies:

1. *Open*, interactive agent-based computational systems (NC)[2] vs. *closed*, stand-alone computational systems (AC);
2. Computation as *information processing and simulative modeling* (NC) vs. computation as *formal (mechanical) symbol manipulation* (AC);
3. *Adequacy* of the computational response via self-organization as the main issue (NC) vs. *halting problem* (and its many, equivalent problems) as the main issue in computability theory (AC).

Of course, such dichotomies must be intended, in perspective, as oppositions between complementary and not mutually exclusive characters of computation models. However, as Dodig-Crnkovic emphasizes, such a complementarity might emerge only when a foundational theory of NC will be sufficiently developed, overall as to the semantic and the logic of NC. The present contribution is devoted precisely to this aim, even though it is necessary to add to the previous list other two essential dichotomic characters of NC, emphasized by Dodig-Crnkovic in other papers, overall the more recent one published on the *Information* journal [7]:

[2] So, she synthesizes this important fundamental character of NC approach: «Agent Based Models are the most important development in this direction, where a complex dynamical system is represented by interacting, in general adaptive, agents. Examples of such systems are in physics: turbulence, percolation, sand pile, weather; in biology: cells organs (including brain), organisms, populations, ecosystems; and in the social sphere: language, organizations, and markets».

1. *Intentional, object-directed, pre-symbolic* computation, based on *chaotic* dynamics in neural computation (NC) vs. *representational, solipsistic, symbolic computation,* based on *linear* dynamics, typical of the early AI approach to the cognitive neuroscience (AC).
2. *Dual ontology* of the energy-information distinction in natural (physical, biological and neural) systems (NC), based at the foundational level on Quantum Mechanics (QM), vs. *monistic ontology,* based on the energy-information equivalence in all natural systems (AC).

1.2 Relevance of the Reference Problem in NC

In this paper, we want to suggest how a foundational approach to NC, overall as to its logical and semantic components, cannot disregard the essential point of how to *integrate in one only formalism* the *physical* ("natural") realm, with the *logical-mathematical* ("computation") one, as well as their relationships. That is, the passage from the realm of the *causal* necessity ("natural") of the physical processes, to the realm of the *logical* necessity ("computational"), eventually representing them either in a sub-symbolic, or in a symbolic form. This foundational task can be performed, by the newborn discipline of *theoretical formal ontology* [8,9,10,11,12], as distinguished from *formal ontology engineering* – an applicative discipline, well established and diffused in the realm of computational linguistics and semantic databases.

Particularly, the distinction between the *formal logic* and the *formal ontology* is precious for defining and solving a foundational misunderstanding about the notion of *reference* that the NC approach had the merit of emphasizing, making aware of it the largest part of the computer science community – and also the rest, we hope, of the scientific community, as far as NC is spreading all over the entire realm of the natural sciences.

In fact, as A. Tarski rightly emphasized since his pioneering work on formal semantics [13], not only the *meaning* but also the *reference* in *logic* has nothing to do with the *real*, *physical world.* To use the classic Tarski's example, the semantic reference of the true statement "the snow is white" is not the whiteness of the crystalized water, but at last an empirical set of data, to which the statement is referring, eventually taken as a *primitive* in a given formal language. In other terms *logic* is always *representational,* it concerns relations among tokens, either at the *symbolic* or *sub-symbolic* level. It has always and only to do with representations, not with real things. This is well emphasized, also, by R. Carnap's principle of the *methodological solipsism* in formal semantics [14], that both (the early) H. Putnam [15] and J. Fodor [16] rightly extended also to the *representationalism* of cognitive science, as far as it is based in the so-called *functionalist* approach of the classic, symbolic AI, and hence of the classic AC paradigm. Finally, this is also the deep reason for what Quine defines as the "impenetrability of reference" beyond the network of equivalent statements, signifying the same referential object in different languages [17].

Now, in AC, any formal theory of reference and truth is faced with the Gödelian limits making impossible a recursive procedure of satisfaction in a semantically closed formal language. What we emphasized also elsewhere [18,19,20], as the core of the reference problem, is that such a recursive procedure, for being complete, would imply the solution of the *coding* problem through a diagonalization procedure; that is, the solution of the so–called "Gödel numbering" problem. In computational terms, the impossibility of solving the coding problem through a diagonalization procedure in AC means that no TM can constitute by itself the "basic symbols" of its own computations. For this reason Tarski rightly stated that, at the level of the propositional calculus, the semantic theory of truth has nothing to say about the conditions under which a given simple ("atomic" in L. Wittengstein's terms) proposition can be asserted. And for this very same reason, in his fundamental paper about *The meaning of "meaning"* [15], Putnam stated that no ultimate solution exists in logic both of the problem of *reference* and, at the level of linguistic analysis, of the problem of *naming*.

In this sense, Putnam stated, we would have to consider ultimately names as *rigid designators* in S. Kripke's sense [21], i.e. in a "one – to – one relationship" with their singular referential objects. However no room exists, also in Kripke's theory of the *partial reference* [22], for justifying formally (algorithmically) the condition of self-reference that the notion of rigid designation supposes. Kripke's modal theory, indeed, uses, in the context of a "three-valued logic", Kleene's genious solution of the partial recursive predicates, for dealing with the problem of the enumeration (*labeling*) of partial functions [23]. By defining the "label" outside the partial domain to be labeled, it avoids inconsistencies and hence undecidabilities, but at the cost of a *substantial arbitrariness* in defining the label. Hence a formal language has always to suppose the existence of names (or numbers) as rigid designators, and cannot give them a non-arbitrary foundation. However what the logic necessity cannot in principle give, the causal necessity could give, as R. Penrose suggested [24].

It is thus evident the necessity of *formal ontology* for formalizing a non-arbitrary approach to the meaning/reference problem in the NC paradigm. That is, it is evident the necessity of *a formal calculus of relations* able to include in the same, coherent, formal framework both "causal" and "logical" relations, as well as the "pragmatic" (real, causal relations with the cognition/communication/computation agents), and not only "syntactic" (logical relations among terms) and/or "semantic" (logical relations among symbols) components of meaningful actions/computations/cognitions.

2 From Formal Logic to Formal Ontology

2.1 Three Eras in the Interpretation of Modal Logic Syntactic Structures

Following [25], we can distinguish three eras in the short history of the modern modal logic.

1. The *first era* is related with the origins of modern modal logic. Starting from 1912, before the publication by Bertrand Russell of Wittegstein's *Tractatus Logico-Philosophicus,* the young American philosopher Clarence I. Lewis denounced in several papers [26,27,28] the limit of using the "material implication" of extensional logic also for the formalization of other types of *deduction/demonstration,* typical of the humanistic disciplines. At the same time, Lewis' vindication of the

oddity of what is today defined as *philosophical logic,* with respect to the *mathematical logic* of the *Principia*, recognized also the power that the axiomatization of mathematical logic will have for the worldwide diffusion of the scientific thought and practice. A similar formalization, according to Lewis, had thus to be developed for what he defined as the *strict implication* typical of metaphysical arguments, in which it is impossible to admit that *true* consequences can be implied by *false* premises, as it is possible by *material implication* of the mathematical logic[3]. In this way, Lewis re-discovered the classical distinction among different ways of defining *necessity* in different linguistic uses (e.g., the *logical* necessity of mathematics is different from the *causal* necessity of ontology, from the *obligation* of ethics and of law, etc.).

2. After Lewis' pioneering work devoted to the intensional interpretations of modal syntactic structures, the *second era* of modal logic development, comprised between '60's and 70's of the last century, is related to the development of Kripke's *formal semantics*, as far as it is based on his brilliant notion of *frame,* as a particular evolution of the mathematical notion of "set". A "frame" indeed is a set of elements *with the complete collection of relations defined on pairs of them*, as we see below. The brilliance of such a notion is related to the fact that the frame notion can be applied, not only to the formalization of intensional models of the modal structures in Lewis' sense, but also to the formalization "from the inside" of extensional, mathematical and algebraic interpretations (models) of the modal structures.
3. All this is related to the so-called *third era* of modern modal logic, from 80's of the last century till now, that is, to *the algebraic interpretation of modal logic,* and of Kripke's relational semantics based on frames. This way back from the philosophical to the mathematical logic, made modal logic an essential tool in *theoretical computer science*, not only for the computer simulation of semantic tasks, but overall for testing "from the inside" the truth and the consistency of mathematical models. Of course, this holds also for the models of *computational physics* and *biology*. This algebraic interpretation is based on two fundamental principles defining the relations between modal logic and mathematical logic:
   (a) The *correspondence principle* between modal formulas defined on models, and first-order formulas in one free-variable of the predicate calculus. This allows the use of modal logic frame semantics, which is a decidable second-order theory, as a meta-logical tool for individuating and testing *decidable* (and hence *computable*) *fragments* in first-order mathematical models, and hence of computer programs too.
   (b) The *duality principle* between modal relational semantics and algebraic semantics, based on the fact that models in modal logic are given not by substituting free variables with constants, like in the predicate calculus semantics, but by *using binary evaluation letters* (0,1) in relational structures (frames) like in algebraic semantics.

[3] In fact, in the concrete existence realm, it is meaningless that an *effect* (= "consequence" in ontological sense) occurs (= it is *true* in ontological sense) if its proper *cause* (= "premise" in ontological sense) does (did) not occur too (= it is *false* in ontological sense).

Modal formal logic is thus fundamental also in our case, i.e., in developing a consistent formal ontology of the *dual (energy-information) ontology* emerging from:

1. The *information-theoretic approach in quantum physics and cosmology* ("It from bit"), in the wider context of a *relational interpretation* of QM [29], perfectly consistent also with its *modal* interpretation [30,31,32,33]. We cannot develop here this point (for an updated synthesis, see [34]).
2. The *information-theoretic approach* of dissipative QFT and its *algebraic* formalism, as a foundational theory of the dissipative structures, applied to the study of the chemical and the biological systems, neural systems included.
3. The *theoretical cognitive science,* since the modal logic furnishes scientists and philosophers with *one only syntactic formalism,* capable of bridging among causal (physical), intensional (psychical), and computational (logical) components of the cognitive agency. This is fundamental for an effective solution *of the reference problem.*

2.2 Intensional Interpretations of Modal Logic

The *modal logic* with all its *intensional* interpretations constitutes the "first era" of its development, i.e., what is today defined as *philosophical logic* [35], as far as it is distinguished from the *mathematical logic*, the logic based on the *extensional* calculus, and the *extensional* meaning, truth, and identity[4].

For our aims, it is sufficient here to recall that formal modal calculus is an extension of classical propositional, predicate and hence relation calculus with the inclusion of some further axioms. Here, we want to recall only some of them — the axioms **N, D, T, 4** and **5** —, useful for us:

N: $<(\mathbf{X}\rightarrow\alpha) \Rightarrow (\Box\mathbf{X}\rightarrow\Box\alpha)>$, where **X** is a set of formulas (language), $\Box$ is the necessity operator, and α is a meta-variable of the propositional calculus, standing for whichever propositional variable *p* of the object-language. **N** is the fundamental *necessitation rule* supposed in any normal modal calculus

D: $<\Box\alpha\rightarrow\Diamond\alpha>$, where $\Diamond$ is the possibility operator defined as $\neg\Box\neg\,\alpha$. **D** is typical, for instance, of the *deontic* logics, where nobody can be obliged to what is impossible to do.

T: $<\Box\alpha\rightarrow\alpha>$. This is typical, for instance, of all the *alethic* logics, to express either the *logic* necessity (determination by law) or the *ontic* necessity (determination by cause).

---

[4] What generally characterizes intensional logic(s) as to the extensional one(s) is that neither the *extensionality axiom* – reducing class identity to class equivalence, i.e., $\mathbf{A}\leftrightarrow\mathbf{B}\Rightarrow\mathbf{A}=\mathbf{B}$ - nor the *existential generalization axiom* – $Pa\Rightarrow\exists xPx$, where *P* is a generic predicate, *a* is an individual constant, *x* is an individual variable – of the extensional predicate calculus hold in intensional logic(s). Consequently, also the Fegean notion of *extensional truth* based on the truth tables does not hold in the intensional, predicate and propositional calculi. Of course, all the "first person" (both singular, in the case of individuals, and plural, in the case of groups), i.e., the *belief* or *intentional* (with *t*) statements, belong to the intensional logic, as J. Searle, from within a solid tradition in analytic philosophy, rightly emphasized [75]. For a formal, deep characterization of intensional logics as to the extensional ones, from one side, and as to intentionality, from the other side, see [76].

4: $\langle\Box\alpha\rightarrow\Box\Box\alpha\rangle$**.** This is typical, for instance, of all the "unification theories" in science where any "emergent law" supposes, as necessary condition, an even more fundamental law.

5: $\langle\Diamond\alpha\rightarrow\Box\Diamond\alpha\rangle$**.** This is typical, for instance, of the logic of metaphysics, where it is the "nature" of the objects that determines necessarily what it can or cannot do.

By combining in a consistent way several modal axioms, it is possible to obtain several *modal systems* which constitute as many syntactical structures available for different intensional interpretations. So, given that **K** is the fundamental modal systems, given by the ordinary propositional calculus **k** plus the necessitation axiom **N,** some interesting modal systems are for our aims are: **KT4** (**S4,** in early Lewis' notation), typical of the physical ontology; **KT45 (S5,** in early Lewis' notation), typical of the metaphysical ontology; **KD45** (**Secondary S5**)**,** with application in deontic logic, but also in epistemic logic, in ontology, and hence in NC as we see.

As we said (see note 4), the extensional notion of truth does not hold in intensional logics, but each of them has its own *truth condition* characterizing it. Generally, the truth condition of a given intensional logic is expressed in *terms of a reflexivity principle,* i.e., a formal scheme that, by applying the proper modal operator of this logic on its argument, is able to give back it as *true.* So, in the *alethic* (either logical or ontological) interpretations of modal structures the necessity operator $\Box p$ is interpreted as "p is true in all possible worlds", while the possibility operator $\Diamond p$ is interpreted as "p is true in some possible worlds". So, the *reflexivity principle* for the necessity operator in its alethic interpretations holds in terms of the axiom **T**, i.e, $\Box p \rightarrow p$. In fact, if p is true in all possible worlds, it is true also in the *actual world* (E.g., "if it is necessary that this heavy body falls (because of Galilei's law), then this body really falls").

This is not true in *deontic* contexts. In fact, "if it is obligatory that all the Italians pay taxes, it does not follow that all Italians really pay taxes", i.e., $\mathbf{O}p \not\rightarrow p$, where **O** is the necessity operator in deontic context. In fact, the obligation operator **O**p must be interpreted as "p is true in all *ideal* worlds" different from the actual one, otherwise $\mathbf{O}=\Box$, i.e., we are in the realm of (meta)physical determinism, where freedom is an illusion and ethics too. The reflexivity principle in deontic contexts, able to make obligations really effective in the actual world, must be thus interpreted in terms of an *optimality operator* $\mathbf{O_p}$ for *intentional agents*, i.e,

$$(\mathbf{O}p\rightarrow p) \Leftrightarrow ((\mathbf{O_p}\,(x,p) \wedge c_a \wedge c_{ni}) \rightarrow p) \tag{1}$$

Where x is an intentional agent, c_a is an acceptance condition and c_{ni} is a non-impediment condition. In similar terms, in *epistemic* contexts, where we are in the realm of representations of the real world we have a different intensional interpretation of modal operators. The interpretations of the two modal epistemic operators **B**(x,p), "x believes that p", and **S**(x,p), "x knows that p" are the following: **B**(x,p) is true iff p is true in the realm of representations believed by x. **S**(x,p) is true iff p is true for all the *founded* representations believed by x. Hence the relation between the two operators is the following:

$$\mathbf{S}(x,p) \Leftrightarrow (\mathbf{B}(x,p) \wedge \mathbf{F}) \tag{2}$$

Where **F** is a *foundation relation,* outside the range of **B,** and hence outside the range of *x* consciousness, otherwise we should not be dealing with "knowing" but only with a "believing of knowing". I.e., we should be within the realm of solipsism and/or of metaphysical nihilism, systematically reducing "science" or "well founded knowledge" to "believing". So, for instance, in the context of a *logicist* ontology, such a **F** is interpreted as a supposed actually infinite capability of human mind of attaining the logical truth [36]. We will offer, on the contrary, a different *finitistic* interpretation of **F** within NC **.** Anyway, as to the reflexivity principle in epistemic context,

$$\mathbf{B}(x,p) \not\to p$$

In fact, believing that a given representation of the actual world, expressed in the proposition *p*, is true, does not mean that it is *effectively* true, if it is not well *founded.* Of course, such a condition **F** — that hence has to be an *onto*-logical condition — is by definition satisfied by the operator **S**, the operator of sound beliefs, so that the reflexivity principle for epistemic context is given by:

$$\mathbf{S}(x,p) \to p \tag{3}$$

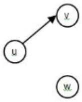
2.3 Kripke's Relational Semantics

The "second era" of modern modal logic is related with Kripke's relational semantic that is an evolution of Tarski formal semantics, with two specific characters: 1) it is related to an *intuitionistic logic* (i.e., it considers as non-equivalent excluded middle and contradiction principle, so to admit coherent theories violating the first one), and hence 2) it is compatible with the *necessarily incomplete character* of the formalized theories (i.e., with the Gödel theorems outcome), and with the *evolutionary character* of natural laws, not only in biology but also in cosmology. In other terms, while in Tarski classical formal semantics, the truth of formulas is concerned with the state of affairs of *one only actual world,* in Kripke relational semantics the truth of formulas depends on states of affairs of worlds different from the actual one (= possible worlds). On the other hand, in contemporary cosmology it is nonsensical speaking of an "absolute truth of physical laws", with respect to a world where the physical laws cannot be always the same, but have to evolve with their referents [37,38].

Anyway, the notion of "possible world" in Kripke semantics has not only a physical sense. On the contrary, as he vindicated many times, the notion of "possible world", as syntactic structure in a relational logic, has as many senses as the semantic models that can be consistently defined on it. In Kripke words, the notion of "possible world" in his semantics has a *purely stipulatory character.* In the same way, in Kripke semantics, like the notion of "possible world" can be interpreted in many ways, so also the relations among worlds can be given as interpretations of the only relation of *accessibility.* In this way, a unified theory of the different intensional interpretations (alethic – ontology included –, deontic, epistemic, etc.) of modal logic became possible, as well as a graphic representation of their relational semantics.

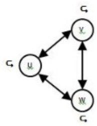
The basic notion for such a graphic representation is the notion of *frame*. This is an ordered pair , **<W,** ***R*>**, constituted by a domain **W** of possible worlds $\{u, v, w\ldots\}$, and a by a two-place *relation* R defined on **W**, i.e., by a set of ordered pairs of elements of **W** ($R \subseteq \mathbf{W}\times\mathbf{W}$), where $\mathbf{W}\times\mathbf{W}$ is the *Cartesian product* of **W** per **W**.

E.g. with $\mathbf{W} = \{u,v,w\}$ and $R = \{uRv\}$, we have:



(4)

According to such a model, the accessibility relation R is only in the sense that v is accessible by u, while w is not related with whichever world. If in **W** all the worlds were reciprocally accessible, i.e., R = $\{uRv, vRu, uRw, wRu, wRv, vRw\}$, then we would have R only included in $\mathbf{W}\times\mathbf{W}$. On the contrary, for having $R = \mathbf{W}\times\mathbf{W}$, we need that each world must be related also with itself, i.e.:

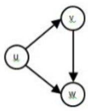


(5)

Hence, from the standpoint of the relation logic, i.e., by interpreting $\{u,v,w\}$ as elements of a class, we can say that this *frame* represents an *equivalence class.* In fact, a *R, transitive, symmetrical* and *reflexive* relation holds among them. Hence, if we consider also the *serial relation:* <(om u)(ex v)(uRv)>[5] where "om" and "ex" are the meta-linguistic symbols, respectively of the universal and existential quantifier, we can discuss also the particular *Euclidean relation* that can be described in a Kripke frame.

The Euclidean property generally in mathematics means a weaker form of the transitive property (that is, if one element of a set has the same relation with other two, these two have the same relation with each other).

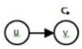
I.e.,<(om u) (om v) (om w) (uRv et $uRw \Rightarrow$ vRw)> :



(6)

Where *et* is the meta-symbol for the logical product.

Hence, for seriality, it is true also <(om u)(om v) ($uRv \Rightarrow vRv$)>:

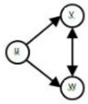


(7)

---

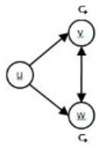
[5] For ontological applications it is to be remembered that seriality means in ontology that the causal chain is always closed, as it is requested in physics by the first principle of thermodynamics, and in metaphysics by the notion of a first cause of everything.

Moreover, <(om *u*) (om *v*) (om *w*) (*uRv* et *uRw* ⇒ *vRw* et *wRv*)>:



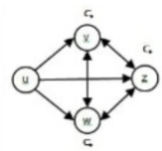
(8)

Finally, if we see at the last two steps, we are able to justify, via the Euclidean relation, a set of *secondary* reflexive and symmetrical relations, so that we have the final frame of a *secondary equivalence* relation among worlds based on an Euclidean relation with a third one:



(9)

Of course, this procedure of *equivalence constitution by a transitive and serial* (=causal) relation can be iterated indefinitely:



(10)

Let us consider now the algebraic interpretation of modal logic applied to the QFT approach to biological and neural systems for an original solution of the reference problem in formal ontology.

3 Dissipative QFT in Biological and Neural Systems and the Formal Ontology of the Reference Problem

3.1 "Coherent States and Coherent Domains in the Physics of the Living Matter"

The title of this sub-section is between quotation marks because effectively it is the title of a recent review paper of the Italian physicist, Giuseppe Vitiello, from the University of Salerno [39]. It synthesizes more than thirty years of research in the QFT widely and universally applied to the study of coherence phenomena in the condensed matter, and extended to the study also of Thermal Field (TF) of dissipative systems, the biological systems and the brain — the "dissipative brain", according to his very effective expression — included.

In fact, it is evident that the vital functions do not depend *only* on the chemical agents (biomolecules) and their interactions at different level of self-organization of the biological matter, but depend also critically on which "organizes the molecular traffic" among the chemical partners. In other terms, each complex vital function consists in an ordered series of single chemical events, according to the chemistry laws. All the chemical interactions, however, (e.g., the van der Waals forces) hold *only for short distances.* The fact that a given molecule arrives in the proximity of the proper receptor, so to make possible the chemical event, can depend neither on the chemical laws, nor on the diffusive processes alone, according to Turing early hypothesis of "morphogenesis" [5], because of the *casual character* of diffusive processes. They indeed, would imply, on one side, a slow temporal dynamics, and, on the other side a series of not appropriate molecular interactions, outside of the "coded" molecular sequence.

The only way for efficiently "canalizing" the molecules, all oscillating according to frequencies depending on quantum physics laws, consists thus in submitting them to electromagnetic fields oscillating according to specific frequencies. Specific molecules can thus recognize each other, also at long distances, and among a multitude of other molecules [40] [41]. The medium in which such oscillating electromagnetic fields occur is the *water*, constituting more than 70% of our bodies, over the 80% of our molecules, in which all the proteins of our bodies are immersed, and in which only the biomolecules are *active*. Now, what characterizes both water molecules and organic molecules is a strong *electrical dipole field.* So, to sum up, the basic hypothesis of QFT applied to living matter is that "at the dynamic fundamental level, *the living matter can be considered as a set of electrical dipoles whose rotational symmetry is broken down*" ([39], p. 16. For the mathematical apparatus of the theory, see [42,43,44,45,46,47]). This is not a reductionist view, because the characterizing properties of living matter, are *macroscopic structures and functions*, with their own laws, *emerging* over the microscopic dynamics generating it. On the contrary, in such a way, the ambiguous notion of *emergence* has, in the context of QFT, a precise connotation, and it is quantitatively well defined. *The emergence of macroscopic properties is given by the dynamic process determining the system ordering.* Of course, any emergence process is related also to a scale change, then, because the dynamic regime responsible of this change is of a quantum nature — because the elementary components have a quantum nature —the resultant system, with its macroscopic properties, is thus a *quantum macroscopic system.*

So, if we consider more closely the *nature* of the correlations among the elementary components in living matter (essentially, the oscillating molecules and their electromagnetic fields), the correlations are essentially *phase correlations,* so that the role of correlation waves is the *fine-tuning* of the elementary oscillations. The "coherence" consists in such a *being in phase.* This implies the immersion of the coherent regions into non-coherent ones, so that their dimensions, because of their dynamic nature, can fluctuate. In this way, other control parameters, such as the temperature, the spatial density of distribution of the material elements, as well as the density of distribution of the electric charges and their fluctuations, can play a fundamental role. Namely, they can determine, either the formation of more extended coherence domains, or, instead, the further fragmentation of them, till their complete destruction, and the recovery of the symmetric "disordered" state.

3.2 Dual Nature of the Dissipative QFT Approach to Biological Systems

Crystals, are, for instance, typical examples of early successful applications of QFT in the realm of non-living, condensed, matter. In crystals, the "order parameter", that is the macroscopic variable characterizing the new emerging level of matter organization, is related to the *matter density distribution*. In fact, in a crystal, the atoms (or the molecules) are "ordered" in well-defined positions, according to a *periodicity law* individuating the crystal lattice. Other examples of such ordered systems, in the non-living realm, are the magnets, the lasers, the super-conductors, etc. In all these systems the emerging properties related to the respective order parameters, are neither the properties of the elementary constituents, nor their "summation", but new properties depending on *the modes in which they are organized*, and hence on *the dynamics controlling their interactions.*

So, any process of *dynamic ordering,* and of *information gain,* is related with a process of *symmetry breakdown,* the symmetry of the disorder of the "quantum vacuum"[6], related to the "third principle of thermodynamics", i.e., with the irreducible motion of particles at the fundamental level and the associated quantum field. In the magnet case, the "broken symmetry" is the rotational symmetry of the magnetic dipole of the electrons, and the "magnetization" consists in the correlation among all (most) electrons, so that they all "choose", among all the directions, that one proper of the magnetization vector.

Finally, whichever dynamic ordering among many objects implies an "order relation", i.e., a *correlation* among them. What, in QFT, at the *mesoscopic/macroscopic* level is denoted as *correlation waves* among molecular structures and their chemical interactions, at the *microscopic* level any correlation, and more generally any interaction, is mediated by *quantum correlation particles.* They are called "Goldstone bosons" or "Nambu-Goldstone bosons" [48,49,50], with mass — even though always very small (if the symmetry is not perfect in finite spaces) —, or *without mass at all* (if symmetry is perfect, in the abstract infinite space). Hence, differently from the *gauge bosons* (the photons γ of electromagnetic field; the gluons g of the strong field, the bosons $W^{\pm}$ and the boson Z of the weak field; and the scalar Higgs boson H^0 of the Higgs field, common to all the precedent ones), which are energy exchange mediators the Goldstone bosons simply vanish when the ordered *modality* of interaction they mediate disappear (e.g., by heating a diamond over 3,545 °C). This is the basis of the fundamental "Goldstone theorem" [51,52]. So, despite the correlation quanta are real particles, observable with the same techniques (diffusion, scattering, etc.) of other particles, not only in QFT of condensed matter, but also in QED and in QCD, wherever we have to reckon with broken symmetries [50], nevertheless they do not exist *outside* the system they are correlating. Also on this regard, a dual ontology is fundamental for avoid confusions and misinterpretations.

[6] In a disordered macrostate, any microstate is equivalent and hence symmetric as to the whole conservation. This is no longer the case, when an ordered macrostate emerges: not all the microstates are equivalent as to the ordered macrostate conservation. Dynamic ordering is thus always related with a symmetry breakdown of the microstate equivalence.

3.3 Doubling of the Degrees of Freedom (DDF) in Dissipative QFT and Its Significance in Cognitive Neuroscience

The Background

As Perrone and myself emphasized in several papers on the physical basis of intentionality [20,18,53,54], only the long-range correlations, which propagate in real-time along wide areas of the brain, and manifest themselves as aperiodic "chaotic" oscillations, can offer a valid dynamical explanation of an intentional act, always involving the simultaneous interaction among emotional, sensory and motor components, located in very far areas of the brain. Such a coordination, that constitutes also the dynamic "texture" of long-term memory phenomena, cannot be explained in terms of the usual axon-synaptic networking, too slow and too limited in space and time, for giving a suitable explanation of this requirement .

On the other hand, Walter J. Freeman and his collaborators, during more than forty years of experimental research by the Neurophysiology Lab at the Dept. of Molecular and Cell Biology of the University of California at Berkeley, not only shared our same theoretical convictions, but observed, measured and modeled this type of dynamic phenomena in mammalian and human brains during intentional acts.

The huge amount of such an experimental evidence found, during the last ten years, its proper physical-mathematical modeling in the dissipative QFT approach of Vitiello and his collaborators, so to justify the publication during the last years of several joint papers on these topics (see, for a synthesis, [55,56]).

To sum up [57], Freeman and his group used several advanced brain imaging techniques such as multi-electrode EEG, electro-corticograms (ECoG), and magnetoencephalogram (MEG) for studying what neurophysiologist generally consider as the *background activity* of the brain, often filtering it as "noise" with respect to the synaptic activity of neurons they are exclusively interested in. By studying these data with computational tools of signal analysis to which physicists, differently from neurophysiologists, are acquainted, they discovered the massive presence of *patterns of AM/FM phase-locked oscillations*. They are intermittently present in resting and/or awake subjects, as well as in the same subject actively engaged in cognitive tasks requiring interaction with the environment. In this way, we can describe them as features of the background activity of brains, modulated in amplitude and/or in frequency by the "active engagement" of a brain with its surround. These "wave packets" extend over coherence domains covering much of the hemisphere in rabbits and cats [58,59,60,61], and regions of linear size of about 19 cm in human cortex [62], with near zero phase-dispersion [63]. Synchronized oscillations of large scale neuron arrays in the b and g ranges are observed by MEG imaging in the resting state and in the motor-task related states of the human brain [64].

DDF in Dissipative QFT of Brain Dynamics

So, what was missing to the Umezawa's pioneering efforts to apply QFT to brain long-term memory dynamics [65] was the mechanism of DDF characterizing the dissipative QFT and its algebraic formalism, developed by E. Celeghini, M. Rasetti, and G. Vitiello during the 90's [66], and explicitly applied by Vitiello himself to the modeling of brain dynamics, but also in any realm of quantum physics, from cosmology, to quantum computing, till chemistry and biology.

In fact, we know that the relevant quantum variables in biological system are the electrical dipole vibrational modes in the water molecules, constituting the oscillatory "dynamic matrix" in which also neurons, glia cells, and the other mesoscopic units of the brain are embedded. The condensation of Goldstone massless bosons (named, in the biological case, Dipole Wave Quanta, DWQ) — corresponding, at the mesoscopic level, to the long-range correlation waves observed in brain dynamics — depends on the triggering action of the external stimulus for the symmetry breakdown of the quantum vacuum of the corresponding brain state. In such a case, the "memory state" corresponds to a coherent state for the basic quantum variables, whose mesoscopic order parameter displays itself at the mesoscopic level, by the amplitude and phase modulation of the carrier signal.

In the classical Umezawa's model [65], however, the system suffered in an "intrinsic limit of memory capacity". Namely, each new stimulus produces the associated DWQ condensation, by cancelling the precedent one, for a sort of "overprinting". *This limit is systematically overcome in dissipative QFT where the many-body model predicts the coexistence of physically distinct amplitude modulated and phase modulated patterns*, as it is observed in the brain. That is, by considering the brain as it is, namely an "open", "dissipative" system continuously interacting with the environment, there not exists one only ground (quantum vacuum) state, like in thermal field theory of Umezawa where the system is studied at equilibrium, but, in principle, infinitely many ground states (quantum vacuum's), so to give the system a potentially infinite capacity of memory. To sum up, the solution of the overprinting problem relies on three facts [67]:

1. In a dissipative (non-equilibrium) quantum system, there are (in principle) infinitely many quantum vacuum's (ground or zero-energy) states, on each of which a whole set of non-zero energy states (or "state space" or "representation states") can be built.
2. Each input triggers one possible irreversible time-evolution of the system, by inducing a "symmetry breakdown" in one quantum vacuum, i.e., by inducing in it an ordered state, a coherent behavior, effectively "freezing" some possible degrees of freedom of the constituting elements behaviors (e.g., by "constraining" them to oscillate on a given frequency), in the same time "labeling" it as the coherent state induced by that input, as an "unitary non-equivalent state" of the system dynamics. In fact, such a coherent state persists in time as a ground state (DWQ are not energetic bosons, are Goldstone bosons) as a specific "long-term" memory state as long as, of course, the brain is coupled with its environment. A brain no longer coupled with its environment is either in a pathological state, or it is directly dead.
3. At this point emerges the DDF principle as a both physical and mathematical necessity of the model. Physical, because a dissipative system, even though in non-equilibrium, must anyway satisfy the *energy balance*. Mathematical, because the 0 energy balance requires a "doubling of the system degrees of freedom". The *doubled* degrees of freedom, say $\tilde{A}$ (the tilde quanta, where the non-tilde quanta A

denote the brain degrees of freedom), thus represent the environment to which the brain is coupled. The environment (state) is thus represented as the "time-reversed *double*" of the brain (state) on which it is impinging. The environment is thus "modeled on the brain", according to the finite set of degrees of freedom the environment itself elicited. Anyway, which are the available degrees of freedom to be elicited for that input depends on the brain itself that, for this reason, is effectively a *self-organizing* system.

Of course, the point 3 represents the essential idea of the "doubling algebra" (algebra/co-algebra) formalism, constituting the mathematical core of the dissipative QFT model that we cannot illustrate here, and for which we refer to [66], and to the wide literature quoted in [57]. Of the DDF we illustrate only, in the final section of this paper its logical relevance, for an original solution of the reference problem, not yet developed till now. For concluding this part, dedicated to the relevance of the dissipative QFT in cognitive neuroscience, I want to emphasize only three final remarks [56,57]:

1. Another success of the dissipative QFT model is that the irreversible time evolution because of the dissipative condition (each coherent state is constituted of "entangled", non-separable, tilde and non-tilde DWQ's), of each "unitary non-equivalent coherent state" can be characterized macroscopically as an *input-labeled* classical chaotic trajectory, in the brain-environment phase space, as it was experimentally observed. I.e., they are trajectories, in the infinite limit: i) bounded and never intersecting with itself; ii) non intersecting with others for different initial conditions; iii) diverging in time also for small differences in the initial conditions. On the other hand, the finite conditions of real systems, the presence of noise and other constraining conditions make possible the phenomena of the "chaotic itinerancy" among different attractors, the fusion of attractors and/or of chaotic trajectories differing for only few degrees of freedom, and other phenomena of "associative memories". The real dynamics so live in a continuous interplay between "noise" and "chaos" for which Freeman invented the neologism of "stochastic chaos" for characterizing the dissipative QFT dynamics of the brain.
2. QFT approach is very different from other approaches to cognitive neuroscience based on QM, in which the quantum effects occur only at the microscopic level. On the contrary, in QFT the effects of quantum events display themselves as a *macroscopic quantum state,* due to the *coherence* of the correlation modes. This makes possible that the interaction between such a mechanism and the electrochemical activity of neurons and synapses, observed by neurophysiologist as the first response to the external stimuli, occurs effectively *only at the macroscopic level*, as the relationship between the *background activity* (memory) and its *ongoing activity* (synapses), in the global interaction between the brain and is environment.
3. Because QFT coherent states are "entangled states" between tilde (environment) and non-tilde (brain) DWQ's, it is evident that also this approach supports the localization of mind and of its logical machinery not "inside" the brain, but *in the dual (energy/information) interplay between the brain and its environment* [67], like all the approaches based on the intentional and not representational theory of mind [18,54,68,69,70,71]. This last remark opens the way to an ontological and hence logical interpretation of the DDF scheme.

3.4 Double Saturation S/P and the Solution of the Reference Problem

To conclude this paper we want to offer for the first time a logical and ontological interpretation of the DDF in brain dynamics as a possible solution of the reference problem, in the direction of the interplay between *physical necessity and logical necessity* that the same notion of NC implies. For this we want to use in a not yet formalized way — that its outside the scope of this paper — the modal logic machinery, developed by Kripke's theory of frames (§1.2), in strict connection with his logical theory of truth (§2.3), in the direction of its algebraic interpretation, applied to the algebra-doubling formalism (co-algebras) of the dissipative QFT.

The first point to recall for understanding this point is that *in any definite description* to be associated to a proper name intended as a *rigid designator* the relationship connecting Subject *S* and Predicate *P* is not of class membership, $\in$, like when we say "Aristotle is *a* philosopher", but of *identity,* =, like when we say "Aristotle is *the* philosopher". The second point to recall is that the notion of *saturation,* today normally used in *modal* model theory for denoting which subset of elements of a given domain *effectively* satisfy a given relation, was introduced in logic by G. Frege for justifying the unity of proposition, where the predicate is the *unsaturated* component and the subject is the *saturated* one.

The solution that the intentional theory of reference suggests is the *double saturation S/P, causally driven by the same referential object.* By such a procedure their logical identity and hence the *referential relation* of the definite description is causally constructed [72,10,54,11].

Thomas Aquinas (1225-1274)[7] depicted in the Middle Age his causal theory of reference in the following way:

> Science, indeed, depends on what is object of science, but the opposite is not true: hence the relation through which science refers to what is known is a *causal* [*real* not *logical*] relation, but the relation through which what is known refers to science is only *logical* [*rational* not *causal*]. Namely, *what is knowable (scibile) can be said as "related", according to the Philosopher, not because it is referring, but because something else is referring to it.* And that holds in all the other things relating each other like the measure and the measured (*Q. de Ver.,* 21, 1. Square parentheses and italics are mine).

In another passage, this time from his commentary to Aristotle book of *Second Analytics*, Aquinas explains the singular reference in terms of a "one-to-one universal" (i.e. Kripke's rigid designators), as opposed to "one-to-many universals" of the generic predications.

---

[7] Historically, he first introduced the notion and the term of "intention" (*intentio*) in the epistemological discussion, in the context of his naturalistic ontology. The approach was hence rediscovered in the XIX century by the philosopher Franz Brentano, in the context of a conceptualist ontology, and hence passed to the phenomenological school, through Brentano's most famous disciple: Edmund Husserl.

It is to be known that here "universal" is not intended as something predicated of many subjects, but according to some adaptation or adequation (*adaptationem vel adaequation*) of the predicate to the subject, as to which neither the predicate can be said without the subject, nor the subject without the predicate (*In Post.Anal.*, I,xi,91. Italics mine).

So, Aquinas' idea is that the predicative statement, when it is denoting a singular object, must be characterized by a "mutual redefinition" between the subject *S* and the predicate *P,* "causally" driven by the referential object itself. DDF mechanism is evidently in an operational, even though unaware, continuity with such Aquinas' theory[73].

On the other hand it is evident that the modal construction of an equivalence relation illustrated step by step in the frames (6)-(9) in §2.3 constitutes a logical description of the DDF principle in dissipative QFT. It is sufficient to interpret *u* as the referential object (environment), *w* as the brain state, *v* as the input state, so that in (6), *uRw*, *uRv,* and *vRw* represent the transitive and serial (= *causal*) relations constituting the initial step of the procedure. Particularly, the relationship *vRw* represents the starting step of DDF in which the input elicits the coherent state (the freezing of the degrees of freedom) in the brain state. In frames (8) and (7) the doubling of the degrees of freedom and the entanglement conditions are, respectively, posed, so to conclude the onto-logical constitution of the transitive-reflexive-symmetrical relations, i.e., the equivalence relation (=*logical*), between *S/P* of a definite description denoting the referential object, we are searching for. Moreover, if we interpret this procedure inside the Kripke theory of truth, as it is natural to do, it is evident that the final frame (9) constitutes an onto-logical depiction of an "unitary in-equivalent state", "labeled" by the referential object *u,* i.e., the "seed" of a new "equivalence class" (see frame (10)).

However, precisely because of this *causal labeling* by the referential object, the theory has no longer that limit of arbitrariness that it has in the original Kripke use of Kleene's partial recursive functions (see above §1.2). In this sense, because the modal equivalence does not generally implies bisimilarity[8] – but bisimilarity is implied only when the specific conditions of the famous van Benthem theorem occur ([25] pp.104ff.), so in our case bisimilarity occurs only when the "doubling" input/output is given in each cognitive agent[9]. This means that the same input causally produces different state-transition sequences (chaotic trajectories) in different cognitive agents, however all equivalent among themselves because *causally* labeled by the same input.

---

[8] We recall here that "computation" in theoretical computer science can be interpreted as a Labeled Transition System (LTS), in the sense that "when we traverse an LTS we build a sequence of state transitions – or to put it another way, we compute" ([25], p.68). So, roughly speaking, bisimilarity between two models *M* and *N* in modal logic means that at each accessibility relation between two states m_i and m_j in *M,* corresponds a relation between n_i and n_j in *N*. So, if we interpret such models as two equivalent programs in dynamic logic (i.e., two "black boxes" producing equivalent outputs for equivalent inputs), their bisimilarity means the further condition of a correspondence between the different "labeled" steps of their execution.

[9] On the other hand, one of the most famous scholars in modal logic, Prof. Yde Venema of the University of Amsterdam, recently demonstrated that the modal logic is the proper logic for co-algebras, just as equation logic is the proper logic for algebras ([77], p. 332).

In this sense, the causal relations from the world *u* (=referential object) onto each of the other worlds (=different, but equivalent definite descriptions of the same object), in the equivalence class of the frame (10), represent the foundation clause **F** of the epistemic logic in its intentional interpretation (see §2.2). Finally, it is evident by such a reconstruction that the localization of a cognitive agency is not "inside the brain", but in the interplay between a brain and its environment.

4 Conclusions

In this paper we showed how the theoretical formal ontology can support a foundational theory of the singular reference, in the context of the NC approach to theoretical computer science, putting in one only relational framework both causal and logical relations. Effectively, this approach satisfies all the NC features listed in §1.1. At the same time, following the *correspondence principle* between modal and mathematical logic (see §2.1), we used frame semantics for individuating as *decidable fragments* particular first order formulas in one free variable: the definite descriptions in QFT modeling of brain-environment dynamics. In this sense, such an approach can offer a foundational modal theory of the logical calculus inside the intentional approach to cognitive neuroscience, till now lacking. It can offer also an adequate start point for developing a NC approach, based on the dissipative QFT model to cognitive functions, as Vitiello himself proposed to develop [67]. Finally, the principle of the input labeling function typical of DDF, offers an original solution to the arbitrariness of the labeling function in Kripke's modal theory of truth, because Kripke's theory is lacking in an intrinsic relation between the labeling function and the definition of the partial domain satisfying the predicate to be labeled (see §1.2). Such an intrinsic relationship naturally exists in DDF approach because the very same causal relation determines, even though in a non-algorithmic way – however forbidden by Gödel theorems –, both the satisfying partial domain of a given predicate, and its label.

References

1. Turing, A.M.: On computable numbers with an application to the Entscheidung problem. Proceedings of the London Mathematical Society 42, 230–265 (1936)
2. Turing, A.M.: Systems of logic based on ordinals. Proceedings of the London Mathematical Society 45, 161–228 (1939)
3. Turing, A.M.: Lecture to the London Mathematical Society on 20 February 1947. In: Ince, D.C. (ed.) Collected Works, I: Mechanical Intelligence, pp. 87–105. North Holland, Amsterdam (1992)
4. Turing, A.M.: Intelligent Machinery, report for the National Physical Laboratory. In: Ince, D.C. (ed.) Collected Works, I: Mechanical Intelligence, p. 106. North Holland, Amsterdam (1992/1948)
5. Turing, A.M.: The chemical basis of morphogenesis. Phil. Trans. R. Soc. London B 237, 37–72 (1952)
6. Dodig-Crnkovic, G.: The significance of models of computation. From Turing model to natural computation. Mind and Machine (2011), doi:10.1007/s11023-011-9235-1

7. Dodig-Crnkovic, G.: Physical computation as dynamics of form that glues everything together. Information 3, 204–218 (2012)
8. Cocchiarella, N.B.: Logic and ontology. Axiomathes 12, 117–150 (2001)
9. Cocchiarella, N.B.: Formal Ontology and Conceptual Realism. Springer, Berlin (2007)
10. Basti, G.: Ontologia formale: per una metafisica post-moderna. In: Strumia, A. (ed.) Il problema dei fondamenti. Da Aristotele, a Tommaso d'Aquino, all'Ontologia Formale, pp. 193–228. Cantagalli, Siena (2007)
11. Basti, G.: Ontologia formale. Tommaso d'Aquino ed Edith Stein. In: Ales-Bello, A., Alfieri, F., Shahid, M. (eds.) Edith Stein, Hedwig Conrad-Martius, Gerda Walter. Fenomenologia della persona, della vita e della comunità, Laterza, Bari, pp. 107–388 (2011)
12. Meixner, U.: The theory of ontic modalities. Ontos Verlag, Frankfurt (2007)
13. Tarski, A.: The Concept of Truth in Formalized Languages. In: Corcoran, J. (ed.) Logic, Semantics, Metamathematics, 2nd edn., Hackett, Indianapolis IN, pp. 152–278 (1983, 1935)
14. Carnap, R.: Testability and meaning. Philosophy of Science 3-4, 419–461 (1936)
15. Putnam, H.: The meaning of 'meaning'. In: Philosophical Papers II: Mind, Language and Reality, pp. 215–271. Cambridge UP, Cambridge (1975)
16. Fodor, J.A.: Metodological solipsism considered as a resarch strategy in cognitive psychology. Behavioral and Brain Sciences 3-1, 63–73 (1980)
17. Quine, W.V.O.: Sticks and stones or the ins and the outs of existence. In: Rounder, L.S. (ed.) On Nature, Boston Univ. Studies in Philosophy and Religion, vol. 6, pp. 13–26. Notre Dame UP, Notre Dame (1984)
18. Basti, G., Perrone, A.L.: Intentionality and Foundations of Logic: a New Approach to Neurocomputation. In: Kitamura, T. (ed.) What Should be Computed to Understand and Model Brain Function? From Robotics, Soft Computing, Biology and Neuroscience to Cognitive Philosophy, pp. 239–288. World Publishing, Singapore (2001)
19. Basti, G., Perrone, A.L.: Le radici forti del pensiero debole. Dalla metafisica, alla matematica, al calcolo. Il Poligrafo and Lateran UP, Padua-Rome (1996)
20. Basti, G., Perrone, A.L.: Chaotic neural nets, computability, undecidability. An outlook of computational dynamics. International Journal of Intelligent Systems 10, 41–69 (1995)
21. Kripke, S.: Naming and necessity. Harvard UP, Cambridge (1980)
22. Kripke, S.: Outline of a theory of truth. The Journal of Philosophy 72-19, 690–716 (1975)
23. Kleene, S.C.: Introduction to meta-mathematics. North Holland, Amsterdam (1952)
24. Penrose, R.: Shadows of the mind. A search for the missing science of consciousness. Oxford UP, Oxford (1996)
25. Blackburn, P., De Rijke, M., Venema, Y.: Modal logic. Cambridge tracts in theoretical computer science. Cambridge UP, Cambridge (2010)
26. Lewis, C.I.: Implication and the Algebra of Logic. Mind 21, 522–531 (1912)
27. Lewis, C.I.: The Calculus of Strict Implication. Mind 23, 240–247 (1914)
28. Lewis, C.I.: A Survey of Symbolic Logic. University of California Press, Berkeley (1918)
29. Rovelli, C.: Relational quantum mechanics. Int. J. Theor. Phys. 35, 1637–1678 (1996)
30. Van Fraassen, B.C.: Quantum mechanics. An Empiricist View. Oxford UP, Oxford (1991)
31. Kochen, S.: A new interpretation of quantum mechanics. In: Mittelstaedt, P., Lahti, P. (eds.) Symposium on the Foundations of Modern Physics, pp. 151–169. World Scientific, Singapore (1985)
32. Dieks, D.: Modal Interpretation of Quantum Mechanics, Measurements, and Macroscopic Behaviour. Physical Review A 49, 2290–2300 (1994)
33. Dieks, D.: Quantum mechanics: an intelligible description of objective reality? Foundations of Physics 35, 399–415 (2005)

34. Dickson, M., Dieks, D.: Modal Interpretations of Quantum Mechanics. In: Zalta, E.N. (ed.) The Stanford Encyclopedia of Philosophy (Spring 2009 edn.) (2009), http://plato.stanford.edu/archives/spr2009/entries/qm-modal/
35. Burgess, J.P.: Philosophical logic (Princeton foundations of contemporary philosophy). Princeton UP, Princeton (2009)
36. Galvan, S.: Logiche intensionali. Sistemi proposizionali di logica modale, deontica, epistemica, Franco Angeli, Milano (1991)
37. Davies, P.: Universe from bit. In: Davies, P., Gregersen, N.H. (eds.) Information and the Nature of Reality. From Physics to Metaphysics, pp. 65–91. Cambridge UP, Cambridge (2010)
38. Benioff, P.: Towards A Coherent Theory of Physics and Mathematics: The Theory-Experiment Connection. Foundations of Physics 35, 1825–1856 (2005)
39. Vitiello, G.: Stati coerenti e domini coerenti della fisica del vivente (Coherent states and coherent domains of the physics of the living matter). La Medicina Biologica 4, 13–19 (2010)
40. Frölich, H.: Long range coherence and energy storage in biological systems. Int. J. of Quantum Chemistry 2, 641 (1968)
41. Popp, F.A., Yan, Y.: Delayed luminescence of biological systems in terms of coherent states. Physics Letters A293, 93–97 (2002)
42. Del Giudice, E., Doglia, S., Milani, M., Vitiello, G.: Spontaneous symmetry breakdown and boson condensation in biology. Phys. Lett. 95A, 508 (1983)
43. Del Giudice, E., Doglia, S., Milani, M., Vitiello, G.: A quantum field theoretical approach to the collective behavior of biologicl systems. Nucl. Phys. B251, 375 (1985)
44. Del Giudice, E., Doglia, S., Milani, M., Vitiello, G.: Electromagnetic field and spontaneous symmetry breakdown in biological matter. Nucl. Phys. B275, 185 (1986)
45. Del Giudice, E., Preparata, G., Vitiello, G.: Water as a free electron laser. Phys. Rev. Lett. 61, 1085 (1988)
46. Del Giudice, E., Vitiello, G.: The role of the electromagnetic field in the formation of domains in the process of symmetry breaking phase transitions. Phys. Rev. A74, 022105 (2006)
47. Vitiello, G.: Coherence and electromagnetic field in living matter. Nanobiology 1, 221 (1992)
48. Nambu, Y.: Quasiparticles and Gauge Invariance in the Theory of Superconductivity. Physical Review 117, 648–663 (1960)
49. Goldstone, J.: Field Theories with Superconductor Solutions. Nuovo Cimento 19, 154–164 (1961)
50. Goldstone, J., Salam, A., Weinberg, S.: Broken Symmetries. Physical Review 127, 965–970 (1962)
51. Itzykson, C., Zuber, G.: Quantum field theory. McGraw-Hill, New York (1980)
52. Umezawa, H.: Advanced field theory: micro, macro and thermal concepts. American Institute of Physics, New York (1993)
53. Basti, G., Perrone, A.L.: Neural nets and the puzzle of intentionality. In: Neural Nets. WIRN Vietri-01. Proceedings of 12th Italian Workshop on Neural Nets, Vietri sul Mare, Salerno, Italy, May 17-19. Springer, Berlin (2002)
54. Basti, G.: Logica della scoperta e paradigma intenzionale nelle scienze cognitive. In: Carere-Comes, T. (ed.) Quale scienza per la psicoterapia? Atti del III Congresso nazionale della SEPI (Society for the Exploration of Psychotherapy Integration). Florence Art Edition, Firenze, pp. 183–216 (2009)

55. Freeman, W.J., Vitiello, G.: Nonlinear brain dynamics as macroscopic manifestation of underlying many-body field dynamics. Physics of Life Reviews 3-2, 93–118 (2006)
56. Freeman, W.J., Vitiello, G.: Dissipation and spontaneous symmetry breaking in brain dynamics. Journal of Physics A: Mathematical and Theoretical 41-30, 304042 (2008)
57. Vitiello, G.: Coherent states, fractals and brain waves. New Mathematics and Natural Computing 5-1, 245–264 (2009)
58. Freenan, W.J.: Origin, structure, and role of background EEG activity. Part 1. Analytic amplitude. Clin. Neurophysiol. 115, 2077–2088 (2004)
59. Freeman, W.J.: Origin, structure, and role of background EEG activity. Part 2. Analytic phase. Clin. Neurophysiol. 115, 2089–2107 (2004)
60. Freeman, W.J.: Origin, structure, and role of background EEG activity. Part 3. Neural frame classification. Clin. Neurophysiol. 116, 111–1129 (2005)
61. Freeman, W.J.: Origin, structure, and role of background EEG activity. Part 4. Neural frame simulation. Clin. Neurophysiol. 117, 572–589 (2006)
62. Freeman, W.J., Burke, B.C., Holmes, M.D., Vanhatalo, S.: Spatial spectra of scalp EEG and EMG from awake humans. Clin. Neurophysiol. 114, 1055–1060 (2003)
63. Freeman, W.J., Ga'al, G., Jornten, R.: A neurobiological theory of meaning in perception. Part 3. Multiple cortical areas synchronize without loss of local autonomy. Intern. J. Bifurc. Chaos 13, 2845–2856 (2003)
64. Freeman, W.J., Rogers, L.J.: A neurobiological theory of meaning in perception. Part 5. Multicortical patterns of phase modulation in gamma EEG. Int. J. Bifurc. Chaos 13, 2867–2887 (2003)
65. Umezawa, H.: Development in concepts in quantum field theory in half century. Math. Japonica 41, 109–124 (1995)
66. Celeghini, E., Rasetti, M., Vitiello, G.: Quantum dissipation. Annals of Phys. 215, 156–170 (1992)
67. Vitiello, G.: The dissipative brain. In: Globus, G.G., Pribram, K.H., Vitiello, G. (eds.) Brain and Being - At the boundary between Science, Philosophy, Language and Arts, pp. 317–330. John Benjamins Pub. Co., Amstedam (2004)
68. Bateson, G.: Mind and nature: a necessary unity. Hampton Press, Princeton (2002)
69. Marturana, H.R., Varela, F.J.: Autopoiesis and cognition. The realization of the living. Reidel, Boston (1980)
70. Clark, A.: Supersizing the mind. Embodiment, action and cognitive extension. Oxford University Press, Oxford (2008)
71. Noë, A.: Out of our heads. Why you are not your brain and other lessons from the biology of consciousness. Hill and Wang Publishers, New York (2009)
72. Perrone, A.L.: A formal scheme for avoiding undecidable problems. Applications to chaotic behavior characterization and parallel computation. In: Andersson, S.I. (ed.) Analysis of Dynamical and Cognitive Systems. LNCS, vol. 888, pp. 9–52. Springer, Heidelberg (1995)
73. Freeman, W.J.: Nonlinear dynamics and the intention of Aquinas. Mind and Matter 6-2, 207–234 (2008)
74. Freeman, W.J.: Foreword. In: Freeman, W.J. (ed.) How brains make up their brains (Japanese Edition) (2012) (in press)
75. Searle, J.R.: Intentionality. An essay in the philosophy of mind. Cambridge UP, New York (1993)
76. Zalta, E.: Intensional logic and the metaphysics of intentionality. MIT Press, Cambridge (1988)
77. Venema, Y.: Algebras and co-algebras. In: Blackburn, P., van Benthem, J.F., Wolter, F. (eds.) Handbook of Modal Logic, pp. 331–426. Elsevier, Amsterdam (2007)

Representation, Analytic Pragmatism and AI

Raffaela Giovagnoli

Faculty of Philosophy, Lateran University, Rome, Italy
raffa.giovagnoli@tiscali.it www.pul.it

Abstract. Our contribution aims at individuating a valid philosophical strategy for a fruitful confrontation between human and artificial representation. The ground for this theoretical option resides in the necessity to find a solution that overcomes, on the one side, strong AI (i.e. Haugeland) and, on the other side, the view that rules out AI as explanation of human capacities (i.e. Dreyfus). We try to argue for Analytic Pragmatism (AP) as a valid strategy to present arguments for a form of weak AI and to explain a notion of representation common to human and artificial agents.

Keywords: Representation, Analytic Pragmatism, Weak AI, Strong AI, Inference.

1 Representation in AI

The notion of "representation" is at the basis of a lively debate that crosses philosophy and artificial intelligence. This is because the comparison starts from the analysis of "mental representations". First, we move by adopting a fruitful distinction between the "symbolic" and the "connectionist" paradigms in AI [1]. This distinction is useful to highlight two different ways of explaining the notion of representation in AI.

An important challenge for AI is to simulate not only the "phonemic" and "syntactic" aspects of mental representation but also the "semantic" aspect. Traditionally, philosophers use the notion of "intentionality" to describe the representational nature of mental states namely intentional states are those that "represent" something, because mind is directed toward objects. The challenge for AI is therefore to approximate to human representations i.e. to the semantic content of human mental states. If we think that representation means to connect a symbol to the object of representation we focus on the discreteness of mental representations. On the contrary, it could be plausible to focus on the inter-relation of mental representations. The first corresponds to the symbolic paradigm in AI, according to which mental representations are symbols. The second corresponds to connectionism in AI, according to which mental representations are distributed patterns [2].

The task to consider the similarity between human and artificial representation could involve the risk of skepticism about the possibility of "computing" this mental capacity. If we consider computationalism as defined in purely abstract syntactic terms then we are tempted to abandon it because human representation involves "real

G. Dodig-Crnkovic and R. Giovagnoli (Eds.): *Computing Nature*, SAPERE 7, pp. 161–169.
DOI: 10.1007/978-3-642-37225-4_9

world constrains". But, a new view of computationalism could be introduced that takes into consideration the limits of the classical notion and aims at providing a concrete, embodied, interactive and intentional foundation for a more realistic theory of mind [3].

Generally speaking, there are several authors who try to conceive a notion of representation that is similar to human mental activity. They criticize Strong AI or GOFAI that rely to a theory of representation attempting to build a mental model by working backwards from sense-impressions and by giving rise to five tensions [4]:

1. mind and world
2. mind and body
3. mental activity and perception
4. plans and behavior
5. abstract ideals and concrete things.

Actually, there are different ways of construing the first opposition, but in AI it has been operationalized by a sharp qualitative distinction between the inside of the machine and the world outside. The obvious consequence is a common idealization to suppose that one's world model is complete in every relevant respect and stays up-to-date automatically. In Agre's words: "Even in domains that involve physical objects, it is common for AI people (and computer scientists in general) to employ the same words to name both the representations in a machine and the things that those representations represent" [5]. The dissociation between mind and body emerges from the typical division of labor in "planning" in AI, where "the mind" generates the plan, and "the body" executes it. Mental activity and perception become conflating as the formal organization of grammatical utterances is privileged upon perceptual activity. The opposition between plans and behavior is originated by an evident difficulty that concerns our complex world. The conception of a plan as a computer program does not capture the knowledge required that is essentially bound to a dynamic world. The last dissociation is directed to Frege and the traditional semantic theory, that aims at capture the content or sense of thoughts and utterances without reference to embodied activities and relationships with which are used. It is agreeable that human representation has to do with concrete reality and for this reason Searle, for instance, provides a thoughtful reinterpretation of the Fregean thoughts, but still remain the problem of how to shift from traditional to natural computing.

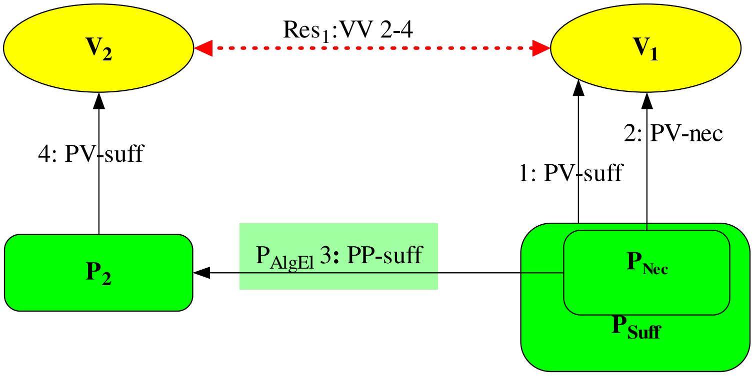
We would like to highlight also an important and recent debate on "digital representation" [6] that focuses on the nature of representations in the computational theory of mind (or computationalism). The starting point is the nature of mental representations, and, particularly, if they are "material". There are authors who maintain that mental representation are material [7] other think that thought processes use conventional linguistic symbols [8]. The question of digital representation involves the "problem of physical computation [9] as well as the necessity of the notion of representation [10] so that we only have the problem of how to intend the very notion of representation [11]. But, there is also the possibility of understanding computation as a purely syntactic procedure or to include "every natural process" in a "computing universe" [12].

2 What Is AP?

The core point of Brandom's original book *Between Saying and Doing* [13] is to describe discursive practices and to introduce norms for deploying an autonomous vocabulary namely a vocabulary of a social practice (science, religion etc.). These norms are logical and are at the basis of an "inferential" notion of representation. But, inference in this sense, recalling Frege, is material [14]. Brandom refuses the explanation of representation in terms of syntactical operations as presented by "functionalism" in "strong" artificial intelligence (AI). He does not even accept weak AI (Searle), rather he aims to present a "logical functionalism" characterizing his analytic pragmatism (AP) [15]. Even though Brandom uses his account of representation to refuse computationalism, his pragmatism is different from the Dreyfus's one, which rests on a non-linguistic know-how (logically and artificially not computable). According to Brandom, we are not only creatures who possess abilities such as to respond to environmental stimuli we share with thermostats and parrots but also "conceptual creatures" i.e. we are logical creatures in a peculiar way.

First, we introduce "practice-vocabulary sufficiency" or "PV-sufficiency" which obtains when exercising a specific set of abilities is sufficient for someone to count as deploying a specified vocabulary [16]. These are for instance "the ability to mean red by the word red" or "the capacity to refer to electrons by the word electrons" (Brandom includes even *intentions* to refer). Together with these basic abilities we must consider the relationship between these and the vocabulary in which we specify them. A second basic meaning-use relation is the "vocabulary-practice sufficiency" or just "VP-sufficiency" namely the relation that holds between a vocabulary and a set of practices-or-abilities when that vocabulary is sufficient to specify those practices-or-abilities. In order to deploy any autonomous vocabulary we must consider the necessity of certain discursive practices defined as "asserting" and "inferring" that, according to Brandom, rule out computationalism [17].

Another basic "meaning-use" relation is the "PV-necessity" that allows the development of more complex relations as exemplified in the following diagram:



There are practices that are PV-necessary to deploy a vocabulary V_1 and are PP-sufficient for practices-or-abilities PV-sufficient to deploy V_2. According to the PV-necessity thesis, there are two abilities that must be had by any system that can deploy an autonomous vocabulary: the ability to respond differentially to some sentence-tokenings as expressing claims the system is disposed to *assert* and the ability to respond differentially to moves relating one set of such sentence-tokenings to another as *inferences* the system is disposed to *endorse*. By hypothesis, the system has the ability to respond differentially to the inference from p (premise) to q (conclusion) by accepting or rejecting it. It also must have the ability to produce tokenings of p and q in the form of asserting.

But what is important is that if we want to sort inferences into good or bad we must focus on conditionals that are PP-necessary to deploy an autonomous vocabulary. What is the relationship between these abilities? By hypothesis, the system has the ability to respond differentially to the inference from p to q by accepting or rejecting it. It also must have the ability to produce tokenings of p and q in the form of asserting. In Brandom's terms "Saying that if something is copper then it conducts electricity is a new way of doing – by saying – what one was doing before endorsing the material inference from "that is copper" to "That conducts electricity". Conditionals make explicit something that otherwise was *implicit* in the practical sorting of non-logical inferences into good and bad. Where before one could only in practice *talk* or *treat* inferences *as* good or bad, after the algorithmic introduction of conditionals one can indorse or reject the inference by explicitly *saying* something, by asserting or denying the corresponding conditionals. What the conditional says explicitly is what one endorsed by doing what one did" [18]. Conditionals are thus the paradigm of *logical* vocabulary to remain in the spirit of Frege's *Begriffschrift*. But, according to Brandom, the meaning-use analysis he provides of conditionals specifies the genus of which logical vocabulary is a species. That genus are ascribed three characteristics:

1. being deployed by practices-or-abilities that are algorithmically elaborated from;
2. practices-or-abilities that are PV-necessary for every autonomous vocabulary (and hence every vocabulary whatsoever) and that
3. suffice to specify explicitly those PV-necessary practices-or-abilities.

Any vocabulary meeting these conditions is called by Brandom "*universal* LX-vocabulary". A crucial consequence of this proposal is that only algorithmic elaboration is required to turn the ability to distinguish material incompatibility into ability to deploy logical *negation*. For example if the ability to distinguish a monochromatic patch is deployed, it (together of the conditional) lets one say *that* two claimable claims are incompatible: "If a monochromatic patch is red, then it is *not* blue".

What are the results of Brandom analytic pragmatism for logic? A first response comes from an observation Brandom formulates in the spirit of Hempel famous essay "The Theoretician's Dilemma" according to which vocabulary and metavocabulary seem of two different kinds pulling against one another. Because Brandom explains logical vocabulary as a species of the genus of conditionals then the dilemma is solved. A further step is to explain why analytic pragmatism is semantically transparent and analytically efficacious. The semantic transparence is due to the fact that we do not need, for example, to use notions such as definitability, translateability, reducibility, supervenience or whatever because there is no interest to the claim that culinary vocabulary supervenes, for instance, on chemical vocabulary, if it turns out we mean that it does so if we can help ourselves to the vocabulary of home economics as an auxiliary in securing that relation. The problem is: how is the contrast between semantic form and content to be drawn so as to underwrite criteria for demarcation for logical vocabulary?

Even Frege's notion of substitution seems not to fulfill this requirement as it does not *provide* but *presuppose* a criterion of demarcation of logical vocabulary. According to Brandom, Frege makes the notion of formality promiscuous because we can pick any vocabulary we like to privilege substitutionally: an inference in good and a claim true in virtue of its *theological* or *geological* form just in case it is good or true and remains so under all substitutions of *non*-theological for *non*-theological vocabulary, or *non*-geological for *non*-geological vocabulary. The sense-dependence in Frege's terms implies that theological and geological formality will not just depend upon but will express an important aspect of the *content* of theological and geological concepts.

The second criteria of *analytical efficacy* means that logic must help in the processes of establishing the semantic relation between vocabularies and we have, according to Brandom, a much more powerful "glue" available to stock together and articulate what is expressed by favored base vocabularies be they phenomenological, secondary-quality or observational (criticism to Russell and Whitehead *Principia*).

Semantic transparence is thus secured by the fact that practices sufficient to deploy logical vocabulary can be algorithmically elaborated from practices necessary to deploy any autonomous vocabulary. The notion of algorithmic elaboration gives a definite sense to the claim that the one set of abilities is in principle sufficient for the other: anyone who can use any base vocabulary already knows how to do everything needed to deploy any universal LX-vocabulary. For analytical efficacy we focus on the fact that logic has an expressive task: to show how *to say* in a *different* vocabulary what can be already be said using the target vocabulary. But logic is PV necessary i.e. logical vocabulary must make it possible to *say* something one could not *say* without it.

Brandom aims at grasping the true nature of human beings namely a problem that is crucial even for AI that tries to go in the direction of a more natural form computationalism. So our question is the following: is our true nature "logical" in virtue of the "fact" that conditionals are the genus of our expressive rationality?

3 Why AP Could Be a Fruitful Strategy to Simulate Representation?

In this conclusive session I'll try to show that the notion of representation described in AP terms presents aspects that are common to human and artificial intelligence.

The PV- and VP-sufficiency thesis suggest that basic practices can be computationally implemented and this description corresponds to the Brandomian interpretation of the Turing test and, consequently, to the refusal of a classical symbolic interpretation in AI (GOFAI) of the notion of human representation. Brandom introduces a pragmatic conception of artificial intelligence or "pragmatic AI" which means that any practice-or-ability P can be decomposed (pragmatically analyzed) into a set of primitive practices-or-abilities such that:

1. they are PP-sufficient for P, in the sense that P can be algorithmically elaborated from them (that is, that a*ll* you need in principle to be able to engage in or exercise P is to be able to engage in those abilities plus the algorithmic elaborative abilities, when these are all integrated as specified by some algorithm); and
2. one could have the capacity to engage or exercise *each* of those primitive practices-or-abilities without having the capacity to engage in or exercise the target practice-or-ability P [19].

For instance, the capacity to do long division is "substantively" algorithmically decomposable into the primitive capacities to do multiplication and subtraction. Namely, we can learn how to do multiplication and subtraction without yet having learning division.

On the contrary, the capacities to differentially respond to colors are not algorithmically decomposable into more basic capacities. This observation entails that there are human but also animal capacities that represent a challenge for strong AI (GOFAI), but nowadays not for new forms of computationalism. Starting from Sellars, we can call them *reliable differential capacities to respond to environmental stimuli* [20] but these capacities are common to humans, parrots and thermostats so that they do not need a notion of representation as symbol manipulation.

Along the line introduced by Sellars, Brandom intends the notion of representation in an "inferential" sense. It is grounded on the notion of "counterfactual robustness" that is bound to the so-called frame problem [21]. It is a cognitive skill namely the capacity to "ignore" factors that are not relevant for fruitful inferences. The problem for AI is not *how* to ignore but *what* to ignore. In Brandom's words: "Since non-linguistic creatures have no semantic, cognitive, or practical access at all to most of the complex relational properties they would have to distinguish to assess the goodness of many material inferences, there is no reason at all to expect that that sophisticated ability to distinguish ranges of counterfactual robustness involving them could be algorithmically elaborated from sorts of abilities those creatures do have" [22].

Nevertheless, we could start by studying what "intelligence" really is by starting from the simplest cases. We can also show the relevance of language for a Turing Robot to steal categories far beyond the temporal and spatial scope of its sensorimotor interactions and data [23]. Harnad proposes the "narrow hybrid approach" to symbol grounding on the basis of sensorimotor interactions with the distal objects of which they are the proximal projections. This sensorimototr capacity is a robotic capacity and aims at capturing instrumental responses or the arbitrary names that successfully sort them according to what is adaptive for the hybrid system. The essential point of

Harnad's proposal is the acquisition of categories by "symbolic theft". Categories can be acquired by "nontoil" through the receipt of verbal information under the conditions that the symbols in the verbal message are already grounded (by sensorimotor toil or indirectly and recursively by previous grounded verbal messages) and that there is someone who already possesses the category and is ready to share it with you.

Harnad makes two important suggestions:

1. Successful sorting capacity must be based on detectable invariance and
2. The invariance can be learned via experience or via hearsay.

The role of language becomes very clear if we consider a useful example: the mushroom/toadstool case [24]: "In a mushroom world I could earn these to important survival categories the hard way, through honest toil, sampling the sensorimotor projections and trying to sort them based on feed back from sometimes getting sick and sometimes getting nourished. Assuming the problem is soluble, that is, that projections are successfully sortable, then if I have the requisite learning capacity, and there is enough time in the day, and I don't kill myself or die of hunger first, I will sooner or later get it right, and the basis of my success will be some sort of invariance in the projections that some internal mechanism of mine has laboriously learned to detect. Let's simplify and say that the invariant is the Boolean rule "if it's white and has red spots, it's a toxic toadstool; otherwise it's an edible mushroom".

Naturally, life becomes easier namely without toil and risk if one could be informed that a "toadstool" is a "mushroom" that is "white" with "red spots". Clearly, one has to know what "mushroom" and "white" and "red" and "spots" were, but, symbolic theft is recursive, though not infinitely regressive (the vocabulary of theft must be grounded directly in honest toil and/or Darwinian theft).

Things become more difficult in the case of the categorization, for instance, of chairs, bears, games and goodness namely the problem is to individuate the shared categories of the sensorimotor projections of all the members of each of these categories. Let's consider Harnad's "Peakaboo Unicorn" [25]: "A Peekaboo Unicorn is a Unicorn, which is to say, it is a horse with a single horn, but it has the peculiar property that it vanishes without a trace whenever sense or measuring instruments are trained on it. So not just in practice, but in principle, you can never see it; it has no sensorimotor projections. Is "Peakaboo Unicorn" therefore a meaningless term?"

It is a meaningful term like "Toadstool" or "Zebra" in the sense that we can give a linguistic description of them. The sentences "a toadstool is a white mushroom with red spots" and "a Zebra is a horse with black and white strips" provide the way you can learn what "toadstool" and "zebra" mean without having to find out the hard way. Again, we need the terms of the sentences to be grounded directly or indirectly. In the case of the Peekaboo Unicorn, it is "horse", "horn", "sense", "measuring instrument" and "vanish" that must be grounded. Harnad's example shows that language gives us resources to give meaning also to abstract entities and this discussion provides arguments for the implementation of "representative" capacities in artificial agents. The "symbol grounding problem" reveals the real challenge for AI (according to the Turing-Test scale) as grounding requires an internal mechanism that can learn by both sensorimotor toil and symbolic theft. A Turing Robot unlike an encyclopedia is "grounded" namely the connection between its symbols and what they are interpretable as is not mediated by an external interpreter. Language has a functional value in humans and robots. It allows it to rely on its proximal projections and the mechanism in between them for grounding.

4 Conclusion

Brandom introduces the notion of "counterfactual robustness" to overcome strong GOFAI, to avoid the primacy of prelinguistic background capacities and skills in weak AI (Searle) and phenomenology (Dreyfus). The notion of representation he introduces could work only if we embrace a peculiar form of inferentialism. Differently, we could think that AP is useful as a philosophical strategy to analyze skills that are common to human, animal and artificial intelligence in a broad sense and also those inferential capacities that are connected with logical laws common to human and artificial agents [26].

Acknowledgments. I would like to thank Gordana Dodig-Crnkovic for the fruitful criticisms and comments on my draft. I am grateful to Gianfranco Basti for the clarification of several essential points. I would also thank the participants to the symposium "Natural/Unconventional Computing" for the interesting and lively discussion.

References

1. Carter, M.: Minds and Computers. Edimburgh University Press, Edimburgh (2007)
2. Carter, ch. 18 (2007)
3. Scheutz, M. (ed.): Computationalism. MIT, New Directions (2002)
4. Agre, P.: The Practical Logic of Computer Work in M. Scheutz (ed.) (2002)
5. Agre, pp. 133–134 (2002)
6. Müller, V.C.: Representation in Digital Systems. In: Briggle, A., et al. (eds.) Current Issues in Computing and Philosophy. IOS Press (2008)
7. Clark, A.: Material Symbols. In: Philosophical Psychology, 19th edn., pp. 291–307 (2006); Sedivy, S.: Minds: Contents without vehicles. Philosophical Psychology 17, 149–179 (2004)
8. Speaks, J.: Is mental content prior to linguistic meaning? Nous 40, 428–467 (2006)
9. Shagrir, O.: Why we view the brain as a computer. Synthese 153, 393–416 (2006)
10. Fodor, J.A.: The mind-body problem. Scientific American 244, 114–123 (1981)
11. O'Brien, G.: Connectionism, analogicity and mental content. Acta Analitica 22, 111–131 (1998)
12. Dodig-Crnkovic, G.: Epistemology naturalized: The info-computationalist approach. APA Newsletter on Philosophy and Computers 6, 9–14 (2007)
13. Brandom, R.: Between Saying and Doing. Oxford University Press, Oxford

14. Brandom, R.: Making It Explicit, Cambridge University press, Cambridge, ch. 2 (1994); Giovagnoli, R.: Razionalità espressiva. Scorekeeping: inferenzialismo, pratiche sociali e autonomia, Mimesis, Milano. In: Giovagnoli, R. (ed.) Prelinguistic Practices, Social Ontology and Semantics, vol. XI(1). Etica & Politica/Ethics & Politics (2004, 2009)
15. Brandom, ch. 2, ch. 3 (2008)
16. Brandom, pp. 74–77 (2008)
17. Brandom, pp. 77–83 (2008)
18. Brandom, pp. 82–83 (2008)
19. Brandom, p. 78 (2008)
20. Sellars, W.: Empiricism and the Philosophy of Mind. Harvard University Press, Cambridge (1957, 1997)
21. Brandom, p. 79 (2008)
22. Brandom, p. 83 (2008)
23. Harnad, S.: Symbol Grounding and the Origin of Language. In: Scheutz (ed.) (2002)
24. Harnad, p. 153 (2002)
25. Harnad, p. 154 (2002)
26. Boley, H.: Semantic Web. Knowledge Representation and Inferencing (2010), `http://www.cs.unb.ca/~boley/talks/DistriSemWeb.ppt` Carter (2007); Evans, R.: Introducing exclusion logic as a deontic logic. In: Governatori, G., Sartor, G. (eds.) DEON 2010. LNCS, vol. 6181, pp. 179–195. Springer, Heidelberg (2010); Giovagnoli, R.: Osservazioni sul concetto di "pratica autonoma discorsiva" in Robert Brandom. Etica & Politica/Ethics and Politics IX(1), 223–235 (2008); Giovagnoli, R.: On Brandom's "Logical Functionalism'". The Reasoner 4 (3) (2010), `http://www.thereasoner.org`

Salient Features and Snapshots in Time: An Interdisciplinary Perspective on Object Representation

Veronica E. Arriola-Rios[1,⋆], Zoe P. Demery[2,⋆],
Jeremy Wyatt[1], Aaron Sloman[1], and Jackie Chappell[2]

[1] School of Computer Science, University of Birmingham,
Edgbaston, Birmingham, B15 2TT, UK
{v.arriola-rios,j.l.wyatt,asloman}@cs.bham.ac.uk
[2] School of Biosciences, B15 2TT, UK
{zxd878,j.m.chappell}@bham.ac.uk

Abstract. Faced with a vast, dynamic environment, some animals and robots often need to acquire and segregate information about objects. The form of their internal representation depends on how the information is utilised. Sometimes it should be compressed and abstracted from the original, often complex, sensory information, so it can be efficiently stored and manipulated, for deriving interpretations, causal relationships, functions or affordances. We discuss how salient features of objects can be used to generate compact representations, later allowing for relatively accurate reconstructions and reasoning. Particular moments in the course of an object-related process can be selected and stored as 'key frames'. Specifically, we consider the problem of representing and reasoning about a deformable object from the viewpoint of both an artificial and a natural agent.

Keywords: Representations, Learning, Exploration, Cognitive Agents, Animal Cognition, Deformable Objects, Affordances, Dynamic Representation, Salient Features.

1 Introduction

The cognitive architecture of any animal or machine (jointly 'agents') has limits, so it cannot contain a perfect model of the dynamic external and internal world, such as about all matter, processes, affordances, or more abstract concepts, like 'mind' or 'spirit'. Every agent receives a particular amount of data through its sensors. How useful that data is in the short or long term depends on the environmental conditions, how accurately the data might be processed into information, and the agent's behavioural response. Frequently, an agent should maximise the amount of meaningful, relevant information it can obtain about its surroundings, while minimising the energy expended, but this is highly dependent on the nature of the agent [1]. This applies not just to a static snapshot of time, but also to a constantly changing world with a past, present and future, where being able to predict events, or select between alternative actions without actually trying

⋆ Joint first authors.

G. Dodig-Crnkovic and R. Giovagnoli (Eds.): *Computing Nature*, SAPERE 7, pp. 171–184.
DOI: 10.1007/978-3-642-37225-4_10

them, may be useful for the agent. So in these circumstances, what are the most useful elements for the agent to store and process in its cognitive architecture and how may they best be coded? Principally, we propose that when an agent gathers information through its senses, often it may form object representations supported by exploration[1].

To date in the field of animal cognition (AC), there has been surprisingly little systematic, quantitative research on exploration, and how it could support learning mechanisms in different agents (see [3] for more discussion). What research there is, has largely been on humans and focussed on Bayesian network learning (e.g. [4]). Among the non-human animal researchers, the focus has been on *what* the different cognitive capacities of different species are, rather than *how* they actually process information to achieve those capacities [5]. For example, the 'trap-tube task' is a typical litmus test for causal understanding of gravity (e.g. [6]). It has revealed a lot about many species, but it is just a binary measure of whether an individual can complete the task or not. No one has fully investigated why one individual can succeed at the task, while another fails – is it something about their different exploratory strategies? Moreover, although quite complex-looking actions can often be performed by agents with simple mechanisms and small neural architectures (e.g. [7]), they may not be able to *generalise* these actions to other similar, but novel circumstances. Thus in this paper, we are concerned with more complex, flexible agents. Another area consistently ignored in AC, but one which may provide answers, is how the senses support exploratory learning (e.g. [8]).

It is a blossoming area in Artificial Intelligence (AI) however. Robots force us to explicitly define the model design, suggesting concrete, testable hypotheses for AC. However, we believe there is not yet a robot/simulation that can flexibly abstract concepts, or generalise knowledge to new situations. AI has looked at different learning mechanisms in isolation with relative success, but few projects have tried combining them into one agent (e.g. [9]). Therefore, AC behavioural experiments can provide realistic biological constraints and inspire more integrative cognitive modelling in AI.

We would like to propose that when exploration of objects occurs for forming representations, it is not always random, but also *structured*, *selective* and *sensitive* to particular features and salient categorical stimuli of the environment. Also that it can follow through three stages of theory formation – the forming, the testing and the refining of hypotheses [10]. Each hypothesis may be specific to a particular group of affordances or processes ('exploratory domains'), but they may also be generalisable to novel contexts. We introduce how studies into artificial agents and into natural agents are complementary [10], by comparing some findings from each field.

First, we will take a top-down approach to explore what some of the general environmental constraints imposed on an agent's system when internalising the world around it may be. Then we will look at some of the possible mechanisms to solve these problems, particularly in the visual domain of object representation. There are several

[1] Cognition does not always rely on internal representations and the degree of detail in any internal representation can vary greatly depending on the situation. For instance, there can be a lack of detail especially when the environment can largely control an agent's behaviour, such as in flocking behaviour or in using pheromone trails. Here alternative, but complementary, mechanisms may be more relevant, such as emergency or embodiment, but in this paper we will not consider these cases[2].

methodological problems in computer vision research, including recognition, tracking and mental imagery [11]. Within robotics, we present an approach where simulations of real objects, calibrated from real-time environmental data, can be used as artificial mental imagery. We have exploited a combination of key features from image analysis, computer graphics and animation, as well as aspects of physical models, to generate an internal representation of a deformable object with predictive capabilities. Finally, we will consider the degree of ecological validity of this model by comparing it with AC behavioural findings about parrots, who are notoriously exploratory and playful throughout their lives.

2 Requirements for the Agent-Environment Interaction

An agent interacting with its surrounding environment often combines perception and analysis with action. It can also be driven by its goals, which can be quite explicit, like foraging for survival, or particular problem-solving tasks. Or they can be quite implicit, such as to gather information by apparently random exploratory behaviour. Shaw [12] suggests, *"The chief end of an intelligent agent is to understand the world around it."* Here, the word 'understanding' implies the agent's ability to make predictions about the world. For this to take place, the agent should be able to detect the consistent properties or salient features in its environmental niche. These properties allow a link to form between the agent and the environment. We will now consider what some of these primordial properties might be (see also [13]).

2.1 Redundancy

Given the inherent limitations of the agent, it will only be possible for it to gain a partial understanding of its surroundings[2]. This partial understanding may not allow the agent to make perfect predictions for all environmental events, so it cannot always be ready to process useful information. As it detects sensory data, it also may not succeed at processing relevant signals. Therefore, we expect there may be errors and inexactitudes at different levels of the agent's perceptual or analytical processes. It may thus be useful for its system to be able to tolerate this margin-of-error. Some agents often have more than one mechanism to find things, solve problems, or to perform actions. The agent could just react according to different layers of data filtering, or it could use one or a combination of different learning mechanisms [10]. While qualitatively different, all of these mechanisms produce similar, valid results. In this sense, we call these different possible mechanisms 'valid ways', and say they are 'redundant'. Therefore, redundancy allows the agent to ignore irrelevant data, or to reconstruct faulty perceptions from new perceptions that convey the same information.

2.2 Consistency

When multiple methods are used to collect or analyse data, they can act in a complementary way, and contribute by providing different information. Alternatively, they may

[2] An artificial agent (e.g. a virtual automaton) in a very simple environment can make perfect predictions; but we are not concerned with these cases.

be superfluous; in which case, they confirm previous findings. For an agent, different methods of perceiving or deducing the same thing should be consistent with each other, if there is enough knowledge. An agent that sees a pen while touching it, should gain tactile information in accordance with the position and surface of the image it sees. If there is a fault in the synchronisation between this visual and tactile information, the agent will not be able to properly integrate this information, or accurately describe the object. This principle is present in human mathematics: different methods used to solve the same equation, *must give the same answer*.

2.3 Persistency

For an agent to be able to make relatively accurate predictions about the environment, there should be at least a few unchanging rules in the environment for a significant period of time. These rules are useful for the agent's internal knowledge and learning mechanisms. The strongest examples can be found in mathematics and physics. In order to develop the cosmological theories of physics, it is necessary to assume that the physical laws that rule at the present time on planet Earth, are the same rules that applied during the Big Bang and in galaxies far beyond ours. Agents should respond in the same way to the environment. During complex actions, an agent may change or modify their goals and plans. Even then though, they should make the changes according to a particular, foreseeable pattern, which may be rooted, for example, in their brain structure. If agents do not follow persistent rules, their behaviour is erratic and unpredictable.

2.4 Regularity

This is the predictable presence of previously perceived features or classes of them, due to a fixed relationship between occurrences[3]. There should be persistent patterns in the environment, allowing at least for partial predictions, particularly when an agent is faced with different causal problems. Causality is a manifestation of regularity, where the partaking elements are not always identifiable, but whose manifestation always entails the same consequence. Thus, agents should have mechanisms capable of detecting these patterns to take advantage of them. Then the environment could be categorised using a finite amount of key features linked by predictive relationships, including elements representing continuous features. For example, a small mathematical equation can describe an infinite parabola.

Sequentiality. This is a particular form of regularity, but in a universe with only one temporal dimension, it becomes especially relevant. Sequentiality is the presence of a series of features of two or more elements that are nearly always perceived in the same total or partial order[4]. The first features can be used to identify the sequence and predict either the following features, or the rules set needed to process them. Some examples include: identify a command and know which actions to execute; analyse the first items of a numerical sequence and predict the next; listen to the first notes of a

[3] This can be present in different dimensions, or in a hierarchical structure.

[4] These may not be contiguous and can include cause-and-effect learning.

song and remember how to sing the rest (which was memorised in advance); identify the beginning of a question and prepare to understand it to look for the answer; or listen to the sound of a prey and prepare to chase.

Structure with Partial Order and Layers. There could also be a succession of *sub*-sequences. The connections here would only allow a few options to follow, such as beginnings of other sub-sequences. This forms a branching structure, which becomes layered, modular, and, in some cases, hierarchical [14]. The maximum length of an existing sequence, and the maximum number of branches that can be remembered and manipulated, impose strict limitations upon what the agent can understand, and the types of patterns it is capable of detecting. However, this structure may allow more complex agents to make abstractions, as concepts formed at one stage could be re-used and refined to repeatedly form ever more complex concepts in multiple ways [1]. This allows for progressively specific and parallel processes (e.g. [15]).

2.5 Experience

For small and well-identified tasks, a largely pre-programmed agent may suffice. Little experience may be needed in a relatively static environment, such as where precocial animals, whose behaviour has been almost completely determined by their genome, just need to survive long enough to reproduce. Other agents are often required to adapt to diverse, dynamic environments, where a lot more learning is required (see [1] for greater discussion). The different extractions of relevant information (Section 2.1) would more likely be processed by mechanisms shaped and influenced by experience. The agent should seek out information to reinforce, evolve and, when possible, prove or disprove its current models, particularly if its expectations are violated. Depending on the needs and the competences of the agent, a specific, relevant subset of experiences would allow specific, relevant features of the individual's niche to be captured (e.g. [16]). We believe there is continual extension of these 'branches' or 'blocks of knowledge' throughout the life of a cognitive agent. At different ages or stages of development, an agent should take in different aspects of the same overheard conversation, for instance, or different aspects of the operation of the same tool.

2.6 Where Does This Leave Us?

All of the above described environmental features/constraints together form a structured universe. Parts of this structure may be perceived and understood by artificial and natural agents. The existence of regularities reduces the information needed to describe a part of the environment, as once enough elements and relationships have been identified, the rest can be inferred. Some animals may have the ability to identify 'valid ways' and describe them as 'formalisms'; sets of rules that can warrant good results when sufficient conditions are met [10]. This is essentially how science operates, particularly logic, mathematics and computer science.

Within the field of AI, some formalisms for 'knowledge representation' focus on the association of symbols to entities (i.e. objects, relationships and processes) in a structured way, such as 'Frame Languages' [17]. However others, like 'First Order Logic', incorporate powerful systems of deduction. These symbolic languages are extremely powerful for discrete reasoning, but they may not be particularly appropriate for describing continuous dynamics, or even for making predictions, such as when objects move through an environment. In AI, it is highly relevant to consider the amount and type of knowledge needed before an agent can be capable of processing it. How much does the agent need to know to be able to predict a few movements of different objects? Can that knowledge be learned from experience, or does it need to be pre-programmed?

In certain contexts, the minimum number of necessary elements to complete a description is known as the number of degrees of freedom. For example, given the generic mathematical formula that describe parabolae, only three points are needed to specify a single, infinite parabola. This principle can be directly applied in computer graphics. By making use of algebraic equations, an infinite amount of shapes can be approximated, represented and reconstructed with just a few polynomials [18]. Furthermore, transformations of these shapes can be encoded with mathematical formulae, thus allowing the representation of physical objects and processes; which can be used to implement a form of mental imagery.

Hence, whether the powerful deductive machinery is available in a natural or an artificial agent, it is important to define how we go from representations of continuous transformations, to discrete objects and events. As with the popular phrase, 'a picture is worth a thousand words', predicate logic may not be able to naturally represent 3D graphical information in a consistent, complete and compact description. It may be possible, however, to extract logical information from graphical simulations when required for symbolic reasoning. Here we give an example of how this could be achieved in AI by combining traditional animation techniques, computer graphics and physics, with symbolic representations.

We believe this approach may be more rigorous than the standard mechanism used in human brains. Humans can recognise things without being able to draw them [19], or use mental imagery without making exact simulations [11] (while our AI system requires them). This shows how we need to better understand the underlying mechanisms of natural agents processing and representing the world around them. Observations of natural exploration behaviour do provide realistic biological constraints on the design of AI models for object representation. We will investigate these issues in AC by running behavioural experiments on parrots, as our exemplar exploratory and adaptive species. Is there evidence of each of the environmental requirements/regularities described above being attended to by the parrots? Does their exploration behaviour suggest underlying strategies for processing and representing the environment?

3 Designing a Representation

3.1 Using Key Frames to Model Deformable Objects

The study of the perception and understanding of the affordances of deformable objects is particularly appropriate to illustrate the points outlined in the section above. The problem of representing solid objects, their properties and their related processes has been studied in great detail in computer graphics [20], and there have been attempts to generate representations using semantic information [21]. Within the first field, there are several good representations for many different types of shapes, most of them based on meshes, splines or quadrics [18]. The motion of objects is simulated with an approach analogous to traditional cartoon animations. There is a set of key frames, where a 'key frame' is a drawn snapshot in time defining the start and end points of any smooth transition, and all of the frames connecting them are called the 'inbetweeners'.

Currently, key frames are identified and drawn by humans; in traditional animation the most skilled cartoonists are responsible for them. Due to the smoothness of the transition between key frames, it is possible for a less-skilled cartoonist to interpolate the inbetweeners. In computer animation, the control points and curves defining the geometry and colours of the scene are set in the key frames. The transitions between key frames are mainly polynomial interpolations, or continuous mathematical transformations of these control elements [22]. To create realistic animations, movements are often captured from real objects. This is a very slow and expensive process [23]. In an attempt to automate the rendering of realistic movements and the inbetweeners, physics engines have been incorporated into the animation packages. They are also present in real-time virtual environments where interaction with a user takes place.

However, the incorporation of physics changes the dynamics of producing an animation slightly. Instead of interpolating between two key frames, the first key frame is given by a human designer and the simulation stops when a given condition is satisfied, thus automatically generating the inbetweeners and the final key frame. Note that predictive capabilities have been attained, and that the simulation is now required to specify the new parameters of the material. This includes mass, young coefficients or spring stiffness, in addition to the method's criteria, such as integration methods or time steps. Correctly estimating these parameters is a difficult problem.

Furthermore, while the simulations may look plausible to the human eye, they may not be physically accurate, so different models are required to simulate different materials and differently shaped objects. A natural agent's brain faces a similar computational problem, yet evolution largely seems to have solved it in a *qualitatively* different way. Humans can reason and make predictions about features of the world, but we probably do not simulate it in the quantitative way a physics engine does. It is still not completely clear how or what exactly the underlying mechanism is in various natural agents. Behavioural experiments can allow us to *infer* what is going on in an animal's mind. However, interpretation of the data is largely based on assumptions and only allows us to make indirect conclusions. Invasive techniques, such as particular neurophysiological or brain imaging methods, only provide partial information about the *content*, or even about the structure or neural representations, in an animal's mind. Thus, if done correctly, AI simulations can be very illuminating. We suggest that an initial list of problems an artificial agent needs to solve are:

1. Generate an internal representation of real deformable objects in the surrounding environment;
2. Identify key frames of the related environmental processes;
3. Interpolate (continuously or discretely) the links between frames;
4. Use previous knowledge to predict future events.

The automation of the animation process provides one solution for the first three points. Traditional animation techniques, however, cannot address the fourth point. The use of physics models and formal logics can address the two last points, but in this case the agent needs to select and calibrate the right model. It is still debatable whether physics models can correctly approximate all the ranges of processes observed in natural environments, given the inherent limitations of mathematical models to model real, complex deformations. Furthermore, there is still no model that integrates all of the points into one agent. Given the huge variety of possible affordances perceived by humans alone, we expect that some form of learning should be used to generate the model(s), which would provide the interpolating link between key frames and aid the artificial agent in making predictions. However, *which* type of learning mechanism is still open to question.

Here we present the advances of a preliminary, physics-based method, where a general (though not completely accurate) model of deformable objects is used [24], and an artificial agent learns to calibrate and use it in the way described above. The next step is to take the key frame representation of the object and extract symbolic ones from it. Then we need to take functions that describe the transformations, associate a symbol to each, and consider that symbol as referring to a categorised process or action. For several cases, this step should be quite straightforward, since the representation has already been discretised, grounded and categorised [25]. Then the already developed, symbolic-level machinery can be applied. Finally, the overall results can be compared with natural exploration behaviour (e.g. of parrots) for ecological validity. Is there evidence of similar mechanisms in natural systems? Is our model biologically plausible?

3.2 Representing the Object's Shape

Kakariki Experiment I: AC Implications for AI Models. When segregating the world around itself, we believe an agent first needs to identify and represent distinct objects. Then the agent needs to understand what the shape of each object means, i.e. its affordances when it interacts with the rest of the world. What are its physical properties? How can some of these properties be encoded in the memory of the agent? For instance, if two key elements are connected by a known relationship, anything in between is already implicitly represented. Contact points and segments of continuous curves can be approximated by lines and polynomials, and delimited by key points. Under this light, it is natural that an agent would be more interested in these points of discontinuity. Indeed, in our first AC experiment, we found that this does seem to be the case, at least for the New Zealand red-fronted parakeet or 'kakariki' (*Cyanoramphus novaezelandiae*).

We chose kakariki as our model animal species for investigating how agents gather and represent environmental information, as they are neophilic and have a high exploratory tendency throughout their lives. Moreover, as with many other parrot species, they are relatively intelligent and have an anatomy adapted to dexterity and fine object manipulation. We presented a series of novel objects of a range of different rigid shapes to 21 kakariki individually and recorded their exploratory behaviour in detail over a 25-minute period. They spent most of the time exploring the corners and indents of the objects, then areas of high curvature second, over smooth surfaces. We would like to suggest this may be because corners and areas of high curvature are more likely to cue useful properties/affordances about different objects.

Related to this finding, it is interesting to consider in AI how Shaw uses information theory to apply the principle of maximising information and predictive capabilities to an image analysis task, and the first result he finds is an edge detector [12]. Similarly, related to the AC finding on relative importance of areas of high curvature, Ravishankar [26] found that it is easier for an artificial agent to recognise deformed objects by placing emphasis on the bending around points of high curvature. It is further compelling that a piece-wise continuous mathematical function is naturally segmented at its points of discontinuity; corners are discontinuities of the derivative of a function of one dimension; edges are discontinuities of the derivative of a function of two dimensions; while points of high curvature (maxima, minima and inflexion points) are points where the first derivatives are zero. It would seem that the same points that mathematicians deem interesting are playing a major role in both natural and artificial agents, as features for object segmentation, categorisation, tracking and, possibly, prediction. Therefore the use of mathematical curves to approximate deformable objects is highly illustrative.

AI Model I: Modelling the Sponge. In one dimension, a way of approximating a continuous curve is by a succession of lines. In two or more dimensions, shapes can be approximated by **meshes**, where each element is 'flat' and defined by nodes and edges. Triangular and hexagonal meshes are widely used. Alternatively, quadrics and polynomials of two or three degrees can be used. They are flexible enough for representing most humanly distinguishable continuous shapes. Polynomials have been used to form **splines**, which are defined by a small set of control points. They can be used to interpolate as much detail of a shape as desired, since the polynomials are continuous, while the connections between them can be discontinuous (e.g. [18]). This is why we considered meshes and splines for our model.

As an experimental example, our model analysed the process of deforming a sponge. In general, compliant materials have the potential to change their shape in nearly an infinite amount of unpredictable ways, therefore understanding deformable objects poses a particularly interesting challenge for both artificial and natural agents. Unlike rigid objects, it is not possible to know in advance all of the elements required for representing the deformation. How many splines or elements in a mesh is required, or what are its degrees of freedom? For some specific objects under controlled circumstances, these possibilities can be restricted, as in medical research with human organs [27]. However, an agent that interacts with an environment populated with unrestricted deformable objects, requires a more general solution. One approach is to automatically generate a hierarchical mesh to represent a few objects in a virtual environment, which adapts as an object deforms [28]. However, this has not yet been directly tried in robotics, where an internal representation needs to match objects in the external environment. This continues to remain an open question even in the AC literature – what would the agent do if an object becomes deformed to a shape unforeseen by the initial representation?

As a tentative first step towards solving this problem, we looked at modelling a spheric robotic finger pushing against a sponge. Please note we are not claiming this model replicates animal vision or reasoning, but it may provide a building block from which to work from. The movement of the robotic finger was blocked by a pencil

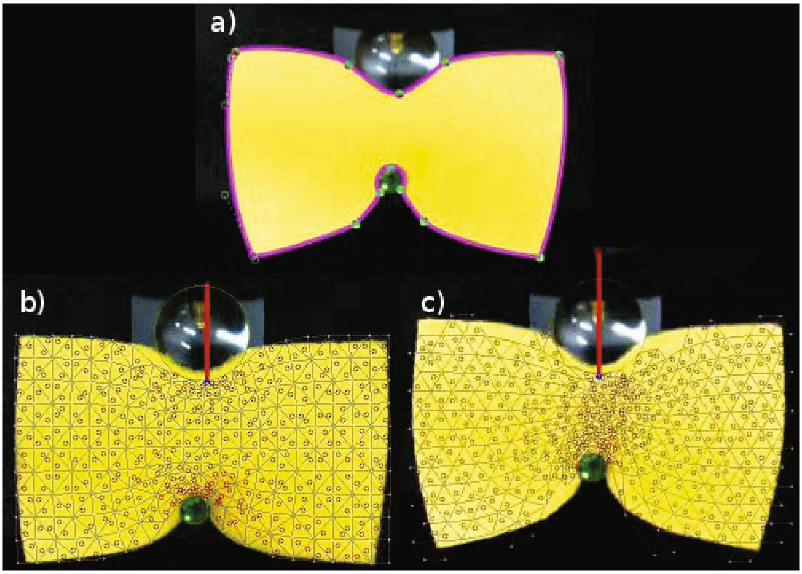


Fig. 1. Top view of an experiment where a robotic finger (sphere at the top) pushes a sponge, perpendicular to its widest axis, against a pencil that serves as an obstacle (green cap). **a)** The contour of a deformed sponge approximated by a series of splines, with the control points placed by a human. **b)** The sponge represented by a rectangular mesh, generated in the first frame before deformation; the mesh configuration was predicted by the physics model. **c)** Hexagonal mesh, similar to (b).

directly opposite. The finger performed a continuous movement against the sponge, while a camera and a force sensor registered the interaction. Figure 1 illustrates the use of splines and meshes to approximate the contour and surface of the sponge as it became deformed.

3.3 Representing the Related Processes

Kakariki Experiment II: More Implications for Models. Once the agent can generate a representation of any object shape it may detect, we believe the next step is for it to understand the related physical processes in the environment. It should identify the key elements and unite them with appropriate functions. How does the object become deformed when interacted with in the world?

We first considered this in the natural dimension in a second AC behavioural experiment. We presented the same kakariki with five cubes of different deformabilities in a random order over five trials over different days. As in the previous experiment, in each trial we allowed them 25 minutes to interact with the objects as they chose and recorded their exploration behaviour in detail.

As we predicted in [10], they initially explored the two extremes the most (i.e. the most rigid and the most deformable cube), but their exploratory 'focus' or 'strategy' changed. So in the second and third trial, the cube of the 'median' or intermediate deformability was explored significantly more than all of the other cubes. Then in the final two trials, the cubes the next interval along (i.e. the second-most deformable cube and the second-most rigid cube) became more of a focus for the kakariki's exploration. In conclusion, the exploration strategy seems to change with time, perhaps as more experience and progressively more specific knowledge is gained about the deformability of objects and different object categories. We would like to suggest that the kakariki had a exploration strategy that allowed them to gain more information about the *process* of deforming an object.

AI Model II: Modelling the Deformation. Simultaneously, we wanted to consider what the design of this internal strategy/learning mechanism could be for an artificial agent. In the AI example of deforming the sponge, the following key frames can be identified:

1. **The finger starts moving.** At this point the force sensor detects only some noise, but the command to move has been given and the vision (camera) begins to detect changes between frames, i.e. that the position of the finger is changing. Thus, the first key frame would contain the finger separated from the sponge and the pencil.
2. **The finger touches the sponge.** At this point the force sensor detects an abrupt increase in one direction. Visually, collision detection routines begin to detect a contact between the circle (i.e. the finger), and one or two triangles in the mesh (i.e. the sponge).
3. **The finger stops moving.** No more changes are detected.

Notice that these coarse key frames are the frames where things change in a very noticeable manner. It is possible to connect frames 1 and 2 by using a function that describes the simple linear translation of the circle (finger). Between frames 2 and 3, the same translation function applies to the circle, but also the physics model gets activated to deform the mesh (sponge) as the circle pushes it. These two functions can predictively describe the observed movements. At frame 3, no function or model is required anymore, because the execution of the command is over and there is no more movement. The scene has ended. From this perspective, the whole process/action can productively be segmented into smaller actions. The internal representation of each frame can be formed by tracing back the activation and deactivation of the required *mechanisms*. Now each segment can be re-represented by a single symbol. The whole sequence can be described as something like: displace finger; push sponge; stop. The agent can then choose between thinking of the command it executed (e.g. translate), or the changes in the sponge (detected through vision or touch), or combinations of both.

There are precedents to doing this type of segmentation, such as in the work by [21]. Here the agent, *Abigail*, analyses a simple circle-and-sticks simulation of ping-pong. Even for this highly simplified world, it was not trivial to unequivocally detect the points of discontinuity that establish the beginning and end of an action. However, Siskind was not quite using our concept of segmentation in modelling, which is just an extension of the idea of a polynomial connecting two control points. Even though the use of splines to approximate curves is a widely used technique, there is not a general technique that can automatically generate a spline from scratch to approximate any curve. It is a brand new research field; to investigate the use of models for interpolation between frames, segmenting and understanding actions.

4 Conclusion

By studying both artificial and natural agents, we can provide a fuller account of how, when necessary, an individual can efficiently represent objects and their related processes in the environment from the huge number of sensory signals they receive. In this light, we can also consider what the requirements posed by the external environment may be upon the finite brain of the agent. Thus, we have briefly discussed two behavioural experiments on parrot exploration of novel objects to give us an insight into what the biological constraints might be on an AI model for representing deformable objects.

Specifically, we have described how a selection of key elements from the environment could be used as a basis for an object representation, and considered possible underlying exploration strategies for gathering information by observing natural behaviour. These key elements are connected through functions, which indicate how to obtain the value of other points. The same mechanism could be used to represent processes and actions, by identifying key frames, and finding the correct physics model to interpolate between frames. It is possible to segment a complex interaction between the agent and the environment into individual actions, by detecting: the commands given; discontinuities in the sensory signals; and the intervals of application of each mechanism. Each of these individual actions could then be represented by symbols. These symbols are grounded in the environment through the selected key elements. It is straightforward to use these symbols for traditional problem-solving tasks, as in [14]. We have further provided evidence that natural agents seem to similarly focus their exploration behaviour on key environmental elements, such as corners, edges and areas of high curvature. Likewise, at least with parrots, individuals seem to attend first to extreme exemplars of particular object properties, including deformability/rigidity, but this exploration strategy becomes gradually refined with time. However, we cannot yet confirm if this parrot exploration is due to similar underlying mechanisms as those presented in our AI model. In conclusion, we have presented an interesting *preliminary* analysis of some of the forms of object representation that may be useful to intelligent natural agents in certain contexts, and demonstrated these capabilities in working computer models.

Acknowledgements. We would like to thank our colleagues in the Intelligent Robotics Lab and the Cognitive Adaptations Research Group for their help throughout this project. This work was made possible by our funders, the CONACYT, Mexico, the ORSAS Scheme and the BBSRC, UK.

References

1. Chappell, J., Sloman, A.: Natural and artificial meta-configured altricial information-processing systems. International Journal of Unconventional Computing 3(3), 211–239 (2007)

2. Calvo, P., Gomila, T. (eds.): Handbook of cognitive science: An embodied approach. Elsevier, London (2008)
3. White, R.W.: Motivation reconsidered: The concept of competence. Psychological Review 66(5), 297–333 (1959)
4. Schulz, L., Sommerville, J.: God does not play dice: Causal determinism and preschoolers' causal inferences. Child Development 77(2), 427–442 (2006)
5. Thomas, R.: Investigating cognitive abilities in animals: unrealized potential. Cognitive Brain Research 3(3-4), 157–166 (1996)
6. Visalberghi, E., Tomasello, M.: Primate causal understanding in the physical and psychological domains. Behavioural Processes 42, 189–203 (1997)
7. McCrone, J.: Smarter than the average bug. New Scientist 190(2553), 37–39 (2006)
8. Demery, Z.P., Chappell, J., Martin, G.R.: Vision, touch and object manipulation in Senegal parrots *Poicephalus senegalus*. Proceedings of the Royal Society B 278(1725), 3687–3693 (2011)
9. Markram, H.: The blue brain project. Nature Reviews Neuroscience 7(2), 153–160 (2006)
10. Demery, Z., Arriola-Rios, V.E., Sloman, A., Wyatt, J., Chappell, J.: Construct to understand: learning through exploration. In: Proceedings of the International Symposium on AI-Inspired Biology, pp. 59–61 (2010)
11. Mesulam, M.: From sensation to cognition. Brain 121, 1013–1052 (1998)
12. Shaw, J.M.: Unifiying Perception and Curiosity. PhD thesis, University of Rochester, Department of Computer Science, The College Arts and Sciences. Rochester, New York (2006); Supervised by Professor Dana H. Ballard
13. Chappell, J., Demery, Z.P., Arriola-Rios, V., Sloman, A.: How to build an information gathering and processing system: Lessons from naturally and artificially intelligent systems. Behavioural Processes 89(2), 179–186 (2012)
14. Arriola, V.E., Savage, J.: Knowledge Acquisition and Automatic Generation of Rules for the Inference Machine CLIPS. In: Gelbukh, A., Kuri Morales, Á.F. (eds.) MICAI 2007. LNCS (LNAI), vol. 4827, pp. 725–735. Springer, Heidelberg (2007)
15. Karmiloff-Smith, A.: Beyond Modularity: a developmental perspective on cognitive science. MIT Press, Cambridge (1999)
16. von Bayern, A.M., Heathcote, R.J., Rutz, C., Kacelnik, A.: The role of experience in problem solving and innovative tool use in crows. Current Biology 19(22), 1965–1968 (2009)
17. Minsky, M.: A framework for representing knowledge. In: Winston, P. (ed.) The Psychology of Computer Vision, p. 81. McGraw-Hill, NY (1975)
18. Montagnat, J., Delingette, H., Ayache, N.: A review of deformable surfaces: topology, geometry and deformation. Image and Vision Computing 19(14), 1023–1040 (2001)
19. Barber, B.: How to Draw Everything. Arcturus (2009)
20. Hoffmann, C.M., Rossignac, J.R.: A road map to solid modeling. IEEE Transactions on Visualization and Computer Graphics 2, 3–10 (1996)
21. Siskind, J.M.: Grounding language in perception. Artificial Intelligence Review 8, 371–391 (1994)
22. Parent, R. (ed.): Computer Animation. Morgan Kaufmann (2009)
23. Menache, A.: Understanding Motion Capture for Computer Animation. Elsevier, London (2011)
24. Teschner, M., Heidelberg, B., Muller, M., Gross, M.: A versatile and robust model for geometrically complex deformable solids. In: Proceedings of Computer Graphics International (CGI 2004), pp. 312–319 (2004)

25. Steels, L.: The symbol grounding problem has been solved. so what's next? In: de Vega, M. (ed.) Symbols and Embodiment: Debates on Meaning and Cognition, Oxford University Press, Oxford (2008)
26. Ravishankar, S., Jain, A., Mittal, A.: Multi-stage contour based detection of deformable objects. In: Forsyth, D., Torr, P., Zisserman, A. (eds.) ECCV 2008, Part I. LNCS, vol. 5302, pp. 483–496. Springer, Heidelberg (2008)
27. McInerney, T., Terzopoulos, D.: Deformable models in medical image analysis: a survey. Medical Image Analysis 1(2), 91–108 (1996)
28. Morris, D., Salisbury, K.: Automatic preparation, calibration, and simulation of deformable objects. Computer Methods In Biomechanics And Biomedical Engineering 11(3), 263–279 (2008)

Toward Turing's A-Type Unorganised Machines in an Unconventional Substrate: A Dynamic Representation in Compartmentalised Excitable Chemical Media

Larry Bull, Julian Holley, Ben De Lacy Costello, and Andrew Adamatzky

Unconventional Computing Group
University of the West of England, Bristol BS16 1QY, U.K.
Larry.bull@uwe.ac.uk

Abstract. Turing presented a general representation scheme by which to achieve artificial intelligence – unorganised machines. Significantly, these were a form of discrete dynamical system and yet such representations remain relatively unexplored. Further, at the same time as also suggesting that natural evolution may provide inspiration for search mechanisms to design machines, he noted that mechanisms inspired by the social aspects of learning may prove useful. This paper presents initial results from consideration of using Turing's dynamical representation within an unconventional substrate - networks of Belousov-Zhabotinsky vesicles - designed by an imitation-based, i.e., cultural, approach. Turing's representation scheme is also extended to include a fuller set of Boolean functions at the nodes of the recurrent networks.

1 Introduction

In 1948 Alan Turing produced an internal paper where he presented a formalism he termed "unorganised machines" by which to represent intelligence within computers (eventually published in [39]). These consisted of two main types: A-type unorganised machines, which were composed of two-input NAND gates connected into disorganised networks (Figure 1); and, B-type unorganised machines which included an extra triplet of NAND gates on the arcs between the NAND gates of A-type machines by which to affect their behaviour in a supervised learning-like scheme. In both cases, each NAND gate node updates in parallel on a discrete time step with the output from each node arriving at the input of the node(s) on each connection for the next time step. The structure of unorganised machines is therefore very much like a simple artificial neural network with recurrent connections and hence it is perhaps surprising that Turing made no reference to McCulloch and Pitts' [29] prior seminal paper on networks of binary-thresholded nodes. However, Turing's scheme extended McCulloch and Pitts' work in that he also considered the training of such networks with his B-type architecture. This has led to their also being known as "Turing's connectionism" (e.g., [10]). Moreover, as Teuscher [36] has highlighted, Turing's unorganised machines are (discrete) nonlinear dynamical systems and therefore have the potential to exhibit complex behaviour despite their construction

G. Dodig-Crnkovic and R. Giovagnoli (Eds.): *Computing Nature*, SAPERE 7, pp. 185–199.
DOI: 10.1007/978-3-642-37225-4_11

from simple elements. The current work aims to explore the use of Boolean dynamic system representations within networks of small lipid-coated vesicles. The excitable chemical Belousov-Zhabotinsky (BZ) [42] medium is packaged into the vesicles which form the simple/elementary components of a liquid information processing system. The vesicles communicate through chemical "signals" as excitation propagates from vesicle to vesicle. Initial experimental implementations which use micro-fluidics to control vesicle placement have recently been reported [25].

This paper begins by considering implementation of the basic two-input NAND gates using the vesicles and then how to design networks of vesicles to perform a given computation. In particular, a form of collision-based computing (e.g., [1]) is used, along with imitation programming (IP) [8], which was also inspired by Turing's 1948 paper, specifically the comment that "*Further research into intelligence of machinery will probably be very greatly concerned with 'searches' …. [an example] form of search is what I should like to call the 'cultural search' … the search for new techniques must be regarded as carried out by the human community as a whole*" [39]. Kauffman [22] introduced a form of dynamical Boolean network which uses any possible Boolean function at each node – random Boolean networks (RBN). The use of other well-known Boolean functions within the networks is subsequently explored here, again through collision-based computing. Performance from the extension to more realistic signal propagation times within the networks is also explored.

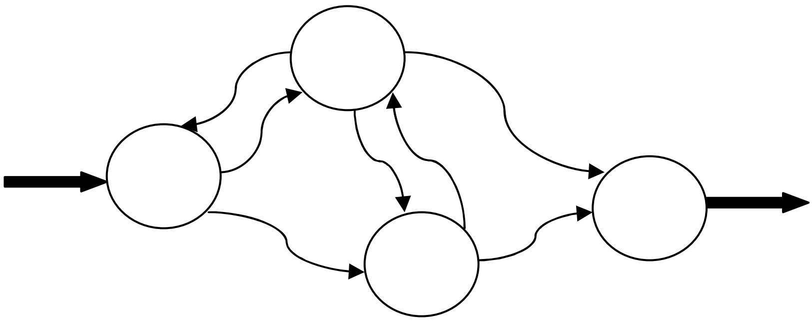


Fig. 1. Simple example A-type unorganised machine consisting of four two-input NAND gate nodes (N=4), with one input (node 1) and one output (node 4) as indicated by the bold arrows

2 Background

"The machine is made up from a rather large number N of similar units. Each unit has two input terminals, and has an output terminal which can be connected to input terminals of (0 or more) other units. We may imagine that for each integer r, 1 ≤ r ≤ N, two numbers i(r) and j(r) are chosen at random from 1..N and that we connect the inputs of unit r to the outputs of units i(r) and j(r). All of the units are connected to a central synchronising unit from which synchronising pulses are emitted at more or less equal intervals of time. The times when these pulses arrive will be called 'moments'. Each unit is capable of having two states at each moment. These states

may be called 0 and 1. The state is determined by the rule that the states of the units from the input leads come are to be taken at the previous moment, multiplied together and then subtracted from 1" [39].

A-type unorganised machines have a finite number of possible states and they are deterministic, hence such networks eventually fall into a basin of attraction. Turing was aware that his A-type unorganised machines would have periodic behaviour and he stated that since they represent "*about the simplest model of a nervous system with a random arrangement of neurons*" it would be *"of very great interest to find out something about their behaviour"* [39]. Figure 2 shows the fraction of nodes which change state per update cycle for 100 randomly created networks, each started from a random initial configuration, for various numbers of nodes *N*. As can be seen, the time taken to equilibrium is typically around 15 cycles, with all nodes in the larger case changing state on each cycle thereafter, i.e., oscillating (see also [36]). For the smaller networks, some nodes remain unchanging at equilibrium on average; with smaller networks, the probability of nodes being isolated is sufficient that the basin of attraction contains a degree of node stasis. However, there is significant variance in behaviour.

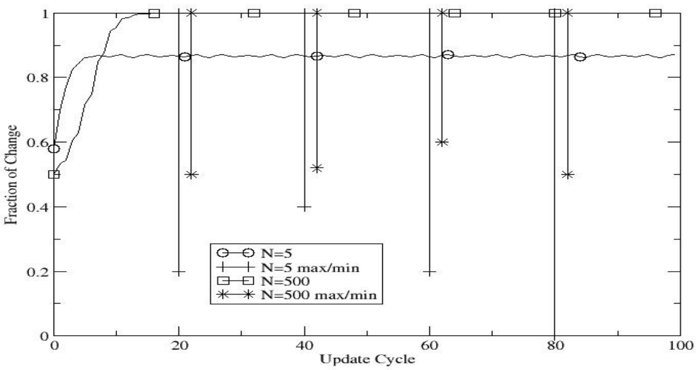


Fig. 2. Showing the average fraction of two-input NAND gate nodes which change state per update cycle of random A-type unorganised machines with various numbers of nodes *N*. Error bars show max. and min. values from 100 trials.

Turing [39] envisaged his A-type unorganised machines being used such that they " *... are allowed to continue in their own way for indefinite periods without interference from outside*" and went on to suggest that one way to use them for computation would be to exploit how the application of external inputs would alter the (dynamic) behaviour of the machine. This can be interpreted as his suggesting individual attractors are used to represent distinct (discrete) machine states and the movement between different attractors as a result of different inputs a way to perform computation. Note this hints at some of the ideas later put forward by Ashby [6] on brains as dynamic systems.

Teuscher [36] used a genetic algorithm (GA) [18] to design A-type unorganised machines for bitstream regeneration tasks and simple pattern classification. Bull [8] used IP to design simple logic circuits, such as multiplexers, from them. Here the unorganised machine had an external input applied, was then updated for some number of cycles, e.g., sufficient for an attractor to be typically reached, and then the state of one or more nodes was used to represent the output. More generally, it is well-established that discrete dynamical systems can be robust to faults, can compute, can exhibit memory, etc. (e.g., see [23][41]).

Given their relative architectural simplicity but potential for complex behaviour, A-type unorganised machines appear to be a good candidate (dynamic) representation to use with novel computing substrates. Their use for a chemical computing system is considered here. It can be noted that Turing (e.g., [40]) was also interested in chemical reaction-diffusion systems, for pattern formation not computation.

3 Chemical Computing

Excitable and oscillating chemical systems have been used to solve a number of computational tasks such as implementing logical circuits [34], image processing [26], shortest path problems [33] and memory [31]. In addition chemical diodes [5], coincidence detectors [15] and transformers where a periodic input signal of waves may be modulated by the barrier into a complex output signal depending on the gap width and frequency of the input [32] have all been demonstrated experimentally.

A number of experimental and theoretical constructs utilising networks of chemical reactions to implement computation have been described. These chemical systems act as simple models for networks of coupled oscillators such as neurons, circadian pacemakers and other biological systems [24]. Ross and co-workers [16] produced a theoretical construct suggesting the use of "chemical" reactor systems coupled by mass flow for implementing logic gates neural networks and finite-state machines. In further work Hjelmfelt et al. [17] simulated a pattern recognition device constructed from large networks of mass-coupled chemical reactors containing a bistable iodate-arsenous acid reaction. They encoded arbitrary patterns of low and high iodide concentrations in the network of 36 coupled reactors. When the network is initialized with a pattern similar to the encoded one then errors in the initial pattern are corrected bringing about the regeneration of the stored pattern. However, if the pattern is not similar then the network evolves to a homogenous state signalling non-recognition.

In related experimental work Laplante et al. [27] used a network of eight bistable mass coupled chemical reactors (via 16 tubes) to implement pattern recognition operations. They demonstrated experimentally that stored patterns of high and low iodine concentrations could be recalled (stable output state) if similar patterns were used as input data to the programmed network. This highlights how a programmable parallel processor could be constructed from coupled chemical reactors. This described chemical system has many properties similar to parallel neural networks. In other work Lebender and Schneider [28] described methods of constructing logical gates using a series of flow rate coupled continuous flow stirred tank reactors (CSTR) containing a bistable nonlinear chemical reaction. The minimal bromate reaction involves the oxidation of cerium(III) (Ce^{3+}) ions by bromate in the presence of bromide and sulphuric acid. In the reaction the Ce^{4+} concentration state is considered as "0" "false" ("1""true") if a given steady state is within 10% of the minimal (maximal) value. The reactors were flow rate coupled according to rules given by a feedforward neural network run using a PC. The experiment is started by feeding in two "true" states to the input reactors and then switching the flow rates to generate "true"-"false", "false"-"true" and "false"-"false". In this three coupled reactor system the AND (output "true" if inputs are both high Ce^{4+}, "true"), OR (output "true" if one of the inputs is "true"), NAND (output "true" if one of the inputs is "false") and NOR gates (output "true" if both of the inputs are "false") could be realised. However to construct XOR and XNOR gates two additional reactors (a hidden layer) were required. These composite gates are solved by interlinking AND and OR gates and their negations. In their work coupling was implemented by computer but they suggested that true chemical computing of some Boolean functions may be achieved by using the outflows of reactors as the inflows to other reactors, i.e., serial mass coupling.

As yet no large scale experimental network implementations have been undertaken mainly due to the complexity of analysing and controlling many reactors. That said there have been many experimental studies carried out involving coupled oscillating and bistable systems (e.g., see [35][11][7][21]). The reactions are coupled together either physically by diffusion or an electrical connection or chemically, by having two oscillators that share a common chemical species. The effects observed include multistability, synchronisation, in-phase and out of phase entrainment, amplitude or "oscillator death", the cessation of oscillation in two coupled oscillating systems, or the converse, "rhythmogenesis", in which coupling two systems at steady state causes them to start oscillating [13].

Vesicles formed from droplets of BZ medium (Figure 3), typically just a few millimetres in diameter, exhibit many properties which may be considered as rudimentary for possible future molecular processing systems: signal transmission, self-repair, signal gain, self-organisation, etc. Their potential use for computation has begun to be explored through collision-based schemes (e.g., [3][4][19][20]). This paper considers their use within a dynamic representation using a collision-based scheme.

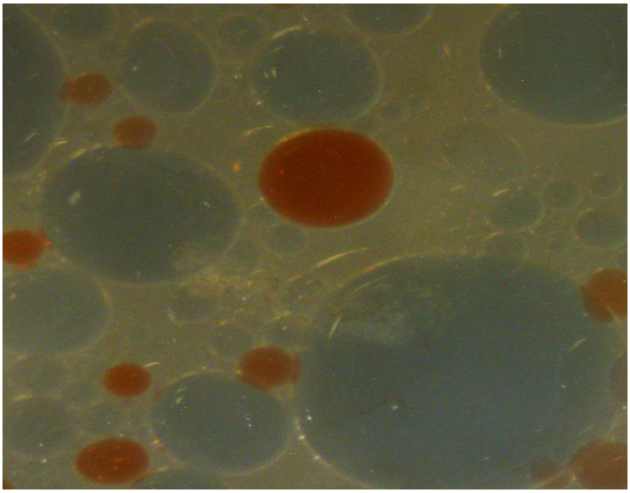


Fig. 3. Showing the BZ droplet vesicles

Collision-based computing exploits the interaction of moving elements and their mutual effects upon each other's movement wherein the presence or absence of elements at a given point in space and time can be interpreted as computation (e.g., see [2] for chemical systems). Collision-based computing is here envisaged within recurrent networks of BZ vesicles, i.e., based upon the movement and interaction of waves of excitation within and across vesicle membranes. For example, to implement a two-input NAND gate, consider the case shown in Figure 4: when either input is applied, as a stream of waves of excitation, no waves are seen at the output location in the top vesicle - only when two waves coincide is a wave subsequently seen at the output location giving logical AND. A NOT gate can be constructed through the disruption of a constant Truth input in another vesicle, as shown.

A-type unorganised machines can therefore be envisaged within networks of BZ vesicles using the three-vesicle construct for the NAND gate nodes, together with chains of vesicles to form the connections between them. Creation of such chains is reported in the initial experimentation with micro-fluidics noted above [25]. As also noted above, it has recently been shown that IP is an effective design approach with the dynamic representation.

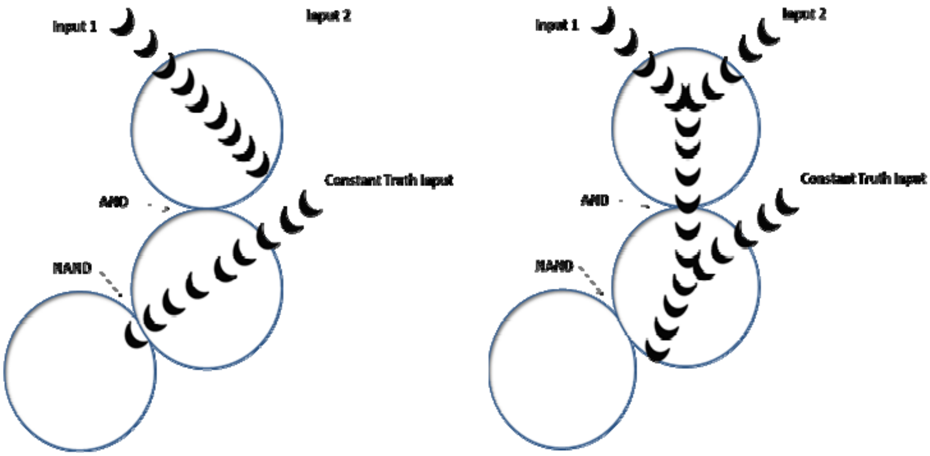


Fig. 4. Showing the construction of a two-input NAND gate under a collision-based scheme using three BZ vesicles. The cases of inputs True-False (left) and True-True (right) are shown. Note the existence of an AND gate also.

4 Imitation Programming

The basic principle of imitation programming is that individuals alter themselves based upon another individual(s), typically with some error in the process. Individuals are not replaced with the descendants of other individuals as in evolutionary search; individuals persist through time, altering their solutions via imitation. Thus imitation may be cast as a directed stochastic search process, thereby combining aspects of both recombination and mutation used in evolutionary computation:

```
BEGIN
INITIALISE population with random candidate solutions
EVALUATE each candidate
REPEAT UNTIL (TERMINATION CONDITION) DO
   FOR each candidate solution DO
         SELECT candidate(s) to imitate
         CHOOSE component(s) to imitate
         COPY the chosen component(s) with ERROR
         EVALUATE new solution
         REPLACE IF (UPDATE CONDITION) candidate with new solution
   OD
OD
END
```

For A-type design, IP utilizes a variable-length representation of pairs of integers defining node inputs, each with an accompanying single bit defining the node's start state. There are three imitation operators - copy a node connection, copy a node start state, and change size through copying. In this paper, each operator can occur with or without error, with equal probability, such that an individual performs one of the six during the imitation process as follows:

To copy a node connection, a randomly chosen node has one of its randomly chosen connections set to the same value as the corresponding node and its same connection in the individual it is imitating. When an error occurs, the connection is set to the next or previous node (equal probability, bounded by solution size). Imitation can also copy the start state for a randomly chosen node from the corresponding node, or do it with error (bit flip here). Size is altered by adding or deleting nodes and depends upon whether the two individuals are the same size. If the individual being imitated is larger than the copier, the connections and node start state of the first extra node are copied to the imitator, a randomly chosen node being connected to it. If the individual being imitated is smaller than the copied, the last added node is cut from the imitator and all connections to it re-assigned. If the two individuals are the same size, either event can occur (with equal probability). Node addition adds a randomly chosen node from the individual being imitated onto the end of the copier and it is randomly connected into the network. The operation can also occur with errors such that copied connections are either incremented or decremented. For a problem with a given number of binary inputs I and a given number of binary outputs O, the node deletion operator has no effect if the parent consists of only $O + I + 2$ nodes. The extra two inputs are constant True and False lines. Similarly, there is a maximum size (100) defined beyond which the growth operator has no effect.

In this paper, each individual in the population P creates one variant of itself and it is adopted if better per iteration. In the case of ties, the solution with the fewest number of nodes is kept to reduce size, otherwise the decision is random. The individual to imitate is chosen using a roulette-wheel scheme based on proportional solution utility, i.e., the traditional reproduction selection scheme used in GAs. Other forms of updating, imitation processes, and imitation selection are, of course, possible [8]. In this form IP may be seen as combining ideas from memetics [12] with Evolutionary Programming [14]. It can be noted GAs have previously been used to design chemical computing systems in various ways (e.g., [9][37][38]).

5 A-Type Experimentation

In the following, three well-known logic problems are used to begin to explore the characteristics and capabilities of the general approach. The multiplexer task is used since they can be used to build many other logic circuits, including larger multiplexers. These Boolean functions are defined for binary strings of length $l = k + 2^k$ under which the k bits index into the remaining 2^k bits, returning the value of the indexed bit. Hence the multiplexer has multiple inputs and a single output. The demultiplexer and adders have multiple inputs and multiple outputs. As such, simple examples of each are also used here. A simple sequential logic task is also used here - the JK latch. In all cases, the correct response to a given input results in a quality increment of 1, with all possible binary inputs being presented per solution evaluation. Upon each presentation of an input, each node in an unorganised machine has its state set to its specified start state. The input is applied to the first connection of each corresponding I input node. The A-type is then executed for 15 cycles. The value on the output node(s) is then taken as the response. All results presented are the average of 20 runs, with P=20. Experience found giving initial random solutions N=O+I+2+30 nodes was useful across all the problems explored here, i.e., with the other parameter/algorithmic settings.

Figure 5 shows the performance of IP to design A-type unorganised machines on k=2 versions of the four tasks: the 6-bit multiplexer (opt. 64), 2-bit adder (opt. 16), 6-bit demultiplexer (opt. 8) and 2-input JK latch (opt. 4). As can be seen, optimal performance is reached in all cases, well within the allowed time, and that the solution sizes are adjusted to the given task. That is, discrete dynamical circuits capable of the given logic functions have been designed. As discussed elsewhere [8], the relative robustness of such circuits to faults, their energy usage, etc. remains to be explored.

However, to begin to consider implementing such designs within BZ vesicles, the time taken for signal propagation between NAND gate nodes needs to included. That is, in Figure 5, as in all previous work with such dynamic representations, any changes in node state are immediately conveyed to any other connected nodes since a traditional computational substrate is assumed. Within the vesicles, changes in NAND gate node state will propagate through chains and hence there will be a time delay proportional to the distance between nodes.

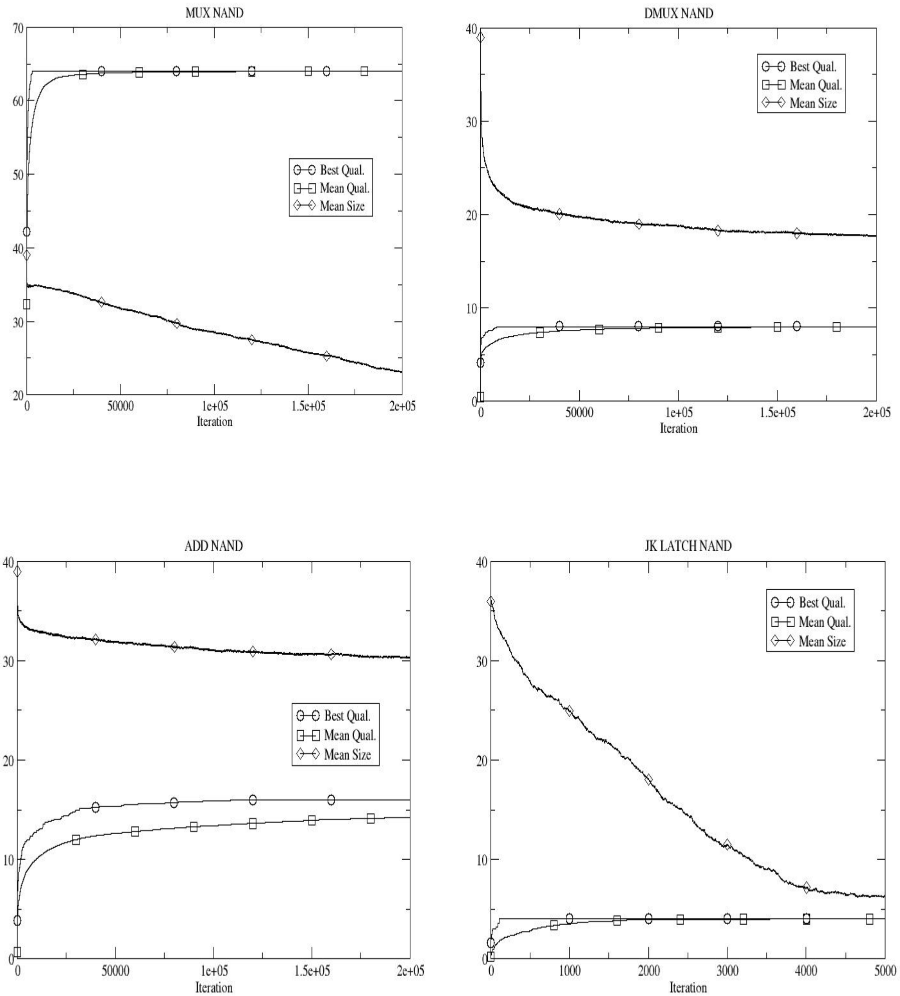


Fig. 5. Showing the performance of IP in designing A-type unorganised machines for the three combinatorial and single sequential logic tasks

Figure 6 shows results for the same experiments and parameters as before but with a form of time delay added to begin to consider the physical implementation in an elementary way. Here NAND gate node states take the same number of update cycles to propagate between nodes as the absolute difference in node number. For example, the state of node 11 at time *t* would take 8 update cycles to reach node 3. Hence at update cycle *t*+8, node 3 would use the state of node 11 as at time *t* as one of its inputs. The number of overall update cycles for the A-types was increased to 50 to help facilitate signal passing across the network.

As Figure 6 shows, it takes longer to reach optimal solutions (T-test, $p<0.05$) and they are perhaps surprisingly smaller (T-test, $p<0.05$, except JK Latch) than before, but suitable dynamic designs are again found in the allotted time, except for the adder which takes longer to reach optimality.

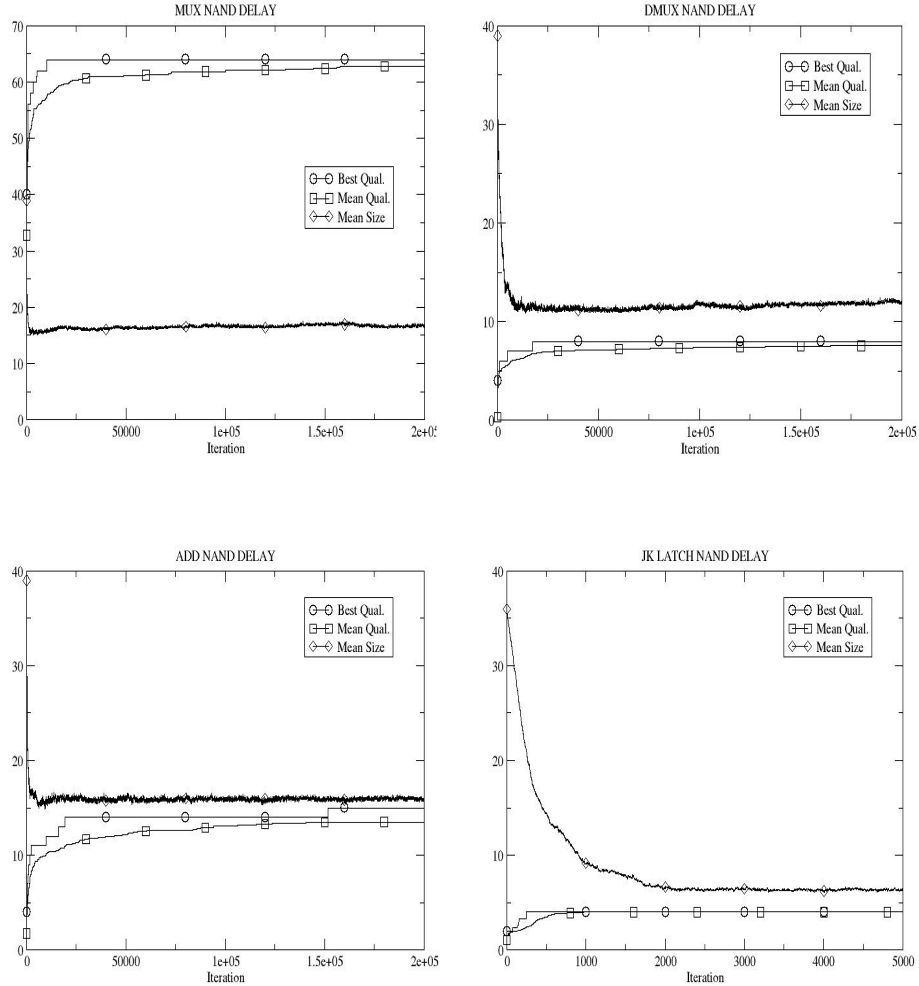


Fig. 6. Showing the performance of IP in designing A-type unorganised machines for the four logic tasks with signal propagation times added

6 RBN Experimentation

Random Boolean networks were originally introduced to explore aspects of biological genetic regulatory networks. Since then they have been used as a tool in a wide range of areas such as self-organisation (e.g., [23]) and computation (e.g., [30]). An RBN typically consists of a network of *N* nodes, each performing a Boolean function with *K* inputs from other nodes in the network, all updating synchronously. As such, RBN may be viewed as a generalization of Turing's A-type unorganised machines which used only the NAND Boolean function with *K*=2. As noted above, Turing's paper was not published until 1968 so it is perhaps not too surprising that Kauffman did not originally discuss his work - although no connection has been made subsequently either, except in [36].

It is well-established that the value of *K* affects the emergent behaviour of RBN wherein attractors typically contain an increasing number of states with increasing *K*. Three phases of behaviour are suggested: ordered when *K*=1, with attractors consisting of one or a few states; chaotic when *K*>2, with a very large numbers of states per attractor; and, a critical regime when *K*=2, where similar states lie on trajectories that tend to neither diverge nor converge and 5-15% of nodes change state

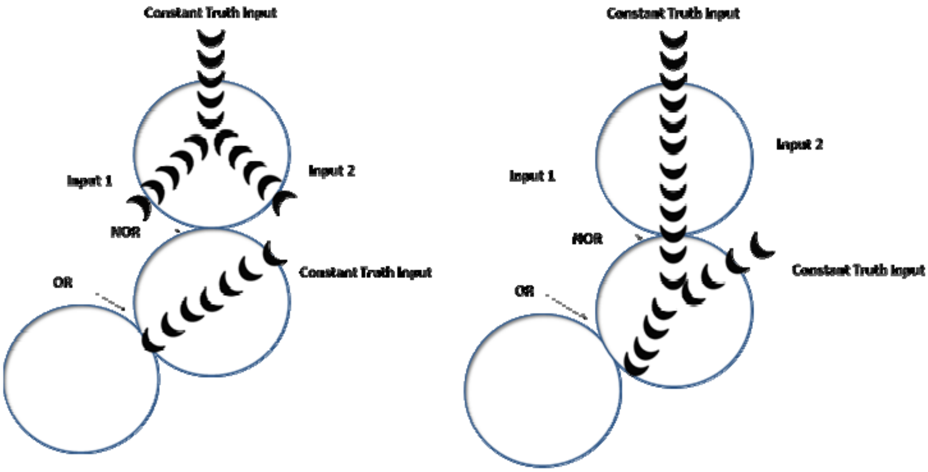


Fig. 7. Showing the construction of a two-input OR and NOR gates under a collision-based scheme using three BZ vesicles. The cases of inputs False-False (left) and True-True (right) are shown.

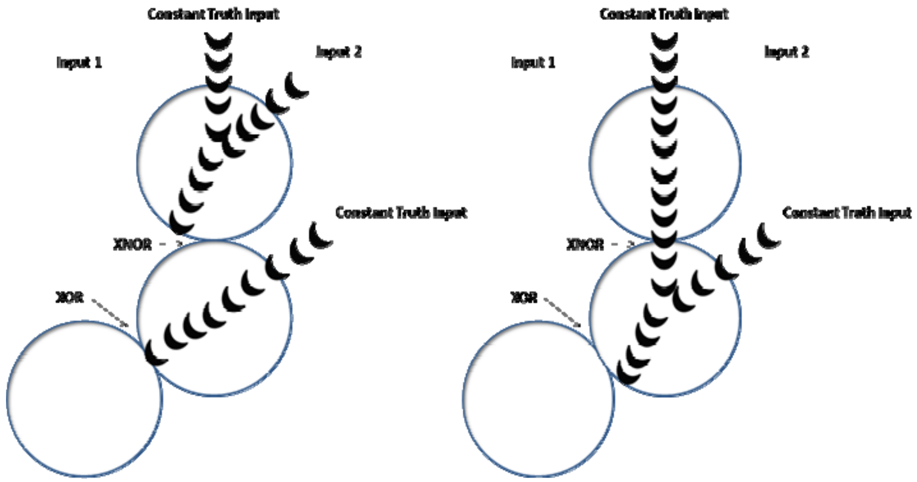


Fig. 8. Showing the construction of a two-input XOR and XNOR gates under a collision-based scheme using three BZ vesicles. The cases of inputs False-False (left) and False-True (right) are shown.

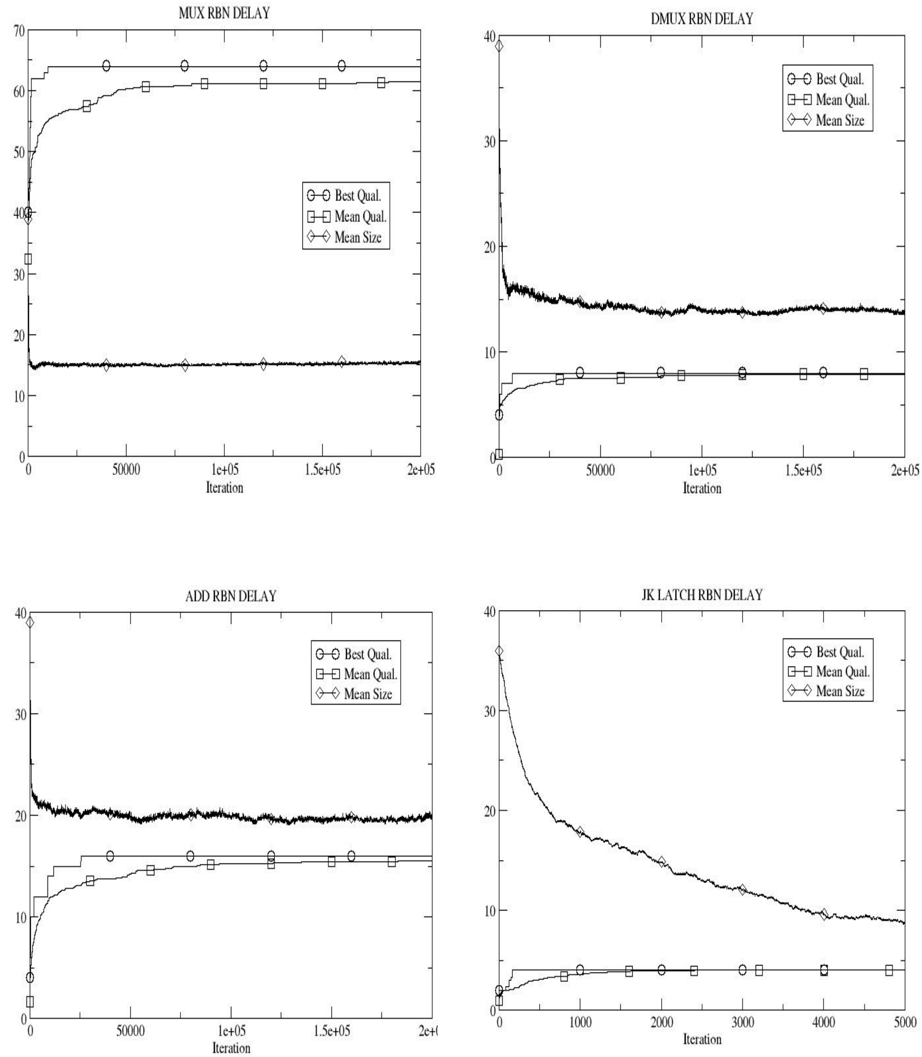


Fig. 9. Showing the performance of IP in designing RBN for the four logic tasks with signal propagation times

per attractor cycle (see [23] for discussions of this critical regime, e.g., with respect to perturbations). Analytical methods have been presented by which to determine the typical time taken to reach a basin of attraction and the number of states within such basins for a given degree of connectivity *K* (again, see [23]).

The previous A-type scenario has been extended to include other well-known Boolean functions through collision-based schemes. Figures 7 and 8 show how two-input OR, NOR, XOR and XNOR can all be achieved using vesicles.

Figure 9 shows how performance is not typically improved in any case considered (T-test, $p<0.05$) with the extra Boolean functionality added - AND, NAND, OR, NOR, XOR, XNOR – and hence Turing's simpler scheme appears to represent a potentially useful approach for implementation with the vesicles.

7 Conclusions

Over sixty years ago, Alan Turing presented a simple representation scheme for machine intelligence – a discrete dynamical system network of two-input NAND gates. Since then only a few other explorations of these unorganized machines are known. As noted above, it has long been argued that dynamic representations provide numerous useful features, such as an inherent robustness to faults and memory capabilities by exploiting the structure of their basins of attraction. For example, unique attractors can be assigned to individual system states/outputs and the map of internal states to those attractors can be constructed such that multiple paths of similar states lead to the same attractor. In this way, some variance in the actual path taken through states can be varied, e.g., due to errors, with the system still responding appropriately. Turing appears to have been thinking along these lines also.

Given the relative simplicity of A-types but their potential for complex behaviour, this paper suggests they may provide a useful representation scheme for unconventional computing substrates. Unconventional computing aims to go beyond traditional architectures and formalisms, much of which is based upon Turing's work on computability, by exploiting the inherent properties of systems to perform computation. A number of experimental systems have been presented in biological, chemical and physical media. Where NAND gate function can be realised, whilst also leaving open the potential utilisation of other aspects of the chosen medium, A-types could be explored. In particular, a substrate of BZ vesicles recently presented as a step towards molecular information processing, e.g., for future smart drugs, was considered and a form of two-input NAND gate designed for it through collision-based computing.

It was then shown how a number of well-known benchmark logic circuits can be designed from A-type unorganised machines using an approach inspired by a comment from Turing on cultural search. Further consideration of the physical implementation within networks of BZ vesicles meant that signal propagation times were also included into the A-types. Results indicate that the design process was slowed relatively but still effective. Extending the NAND gate functionality to include other well-known Boolean logic within the networks showed no improved performance in the more realistic case. Current work is increasing the level of detail of the simulated chemical system both in terms of the vesicle structure and of the BZ reaction therein.

Acknowledgement. The research was supported by the NEUNEU project sponsored by the European Community within FP7-ICT-2009-4 ICT-4-8.3 - FET Proactive 3: Bio-chemistry-based Information Technology (CHEM-IT) program.

References

1. Adamatzky, A. (ed.): Collision-based Computing. Springer, London (2002)
2. Adamatzky, A., De Lacy Costello, B., Asai, T.: Reaction-Diffusion Computers. Elsevier (2005)
3. Adamatzky, A., Holley, J., Bull, L., De Lacy Costello, B.: On Computing in Fine-grained Compartmentalised Belousov–Zhabotinsky Medium. Chaos, Solitons & Fractals 44(10), 779–790 (2011)
4. Adamatzky, A., De Lacy Costello, B., Holley, J., Gorecki, J., Bull, L.: Vesicle computers: Approximating a Voronoi diagram using Voronoi automata. Chaos Solitons and Fractals 44, 480–489 (2011)
5. Agladze, K., Aliev, R.R., Yamaguhi, T., Yoshikawa, K.: Chemical diode. Journal of Physical Chemistry 100, 13895–13897 (1996)
6. Ashby, W.R.: Design for a Brain. Wiley, New York (1954)
7. Bar-Eli, K., Reuveni, S.: Stable stationary-states of coupled chemical oscillators: Experimental evidence. Journal of Physical Chemistry 89, 1329–1330 (1985)
8. Bull, L.: Using Genetical and Cultural Search to Design Unorganised Machines. Evolutionary Intelligence 5(1) (2012)
9. Bull, L., Budd, A., Stone, C., Uroukov, I., De Lacy Costello, B., Adamatzky, A.: Towards Unconventional Computing Through Simulated Evolution: Learning Classifier System Control of Non-Linear Media. Artificial Life 14(2), 203–222 (2008)
10. Copeland, J., Proudfoot, D.: On Alan Turing's Anticipation of Connectionism. Synthese 108, 361–377 (1996)
11. Crowley, M.F., Field, R.J.: Electrically coupled Belousov-Zhabotinskii oscillators 1: Experiments and simulations. Journal of Physical Chemistry 90, 1907–1915 (1986)
12. Dawkins, R.: The Selfish Gene. Oxford Press, Oxford (1976)
13. Dolnik, M., Epstein, I.R.: Coupled chaotic oscillators. Physical Review E 54, 3361–3368 (1996)
14. Fogel, L.J., Owens, A.J., Walsh, M.J.: Artificial Intelligence Through A Simulation of Evolution. In: Maxfield, M., et al. (eds.) Biophysics and Cybernetic Systems: Proceedings of the 2nd Cybernetic Sciences Symposium, pp. 131–155. Spartan Books (1965)
15. Gorecki, J., Yoshikawa, K., Igarashi, Y.: On chemical reactors that can count. Journal of Physical Chemistry A 107, 1664–1669 (2003)
16. Hjelmfelt, A., Ross, J.: Mass-coupled chemical systems with computational properties. Journal of Physical Chemistry 97, 7988–7992 (1993)
17. Hjelmfelt, A., Weinberger, E.D., Ross, J.: Chemical implementation of neural networks and Turing machines. PNAS 88, 10983–10987 (1991)
18. Holland, J.H.: Adaptation in Natural and Artificial Systems. Univ. of Mich. Press (1975)
19. Holley, J., Adamatzky, A., Bull, L., De Lacy Costello, B., Jahan, I.: Computational modalities of Belousov-Zhabotinsky encapsulated vesicles. Nano Communication Networks 2, 50–61 (2011)
20. Holley, J., Jahan, I., De Lacy Costello, B., Bull, L., Adamatzky, A.: Logical and Arithmetic Circuits in Belousov Zhabotinsky Encapsulated Discs. Physical Review E 84, 56110 (2011)
21. Holz, R., Schneider, F.W.: Control of dynamic states with time-delay between 2 mutually flow-rate coupled reactors. Journal of Physical Chemistry 97, 12239 (1993)
22. Kauffman, S.A.: Metabolic Stability and Epigenesis in Randomly Constructed Genetic Nets. Journal of Theoretical Biology 22, 437–467 (1969)
23. Kauffman, S.A.: The Origins of Order. Oxford Press, Oxford (1993)

24. Kawato, M., Suzuki, R.: Two coupled neural oscillators as a model of the circadian pacemaker. Journal of Theoretical Biology 86, 547–575 (1980)
25. King, P.H., Corsi, J.C., Pan, B.-H., Morgan, H., de Planque, M.R., Zauner, K.-P.: Towards molecular computing: Co-development of microfluidic devices and chemical reaction media. Biosystems (2012)
26. Kuhnert, L., Agladze, K.I., Krinsky, V.I.: Image processing using light sensitive chemical waves. Nature 337, 244–247 (1989)
27. Laplante, J.P., Pemberton, M., Hjelmfelt, A., Ross, J.: Experiments on pattern recognition by chemical kinetics. Journal of Physical Chemistry 99, 10063–10065 (1995)
28. Lebender, D., Schneider, F.W.: Logical gates using a nonlinear chemical reaction. Journal of Physical Chemistry 98, 7533–7537 (1994)
29. McCulloch, W.S., Pitts, W.: A Logical Calculus of the Ideas Immanent in Nervous Activity. Bulletin of Mathematical Biophysics 5, 115–133 (1943)
30. Mesot, B., Teuscher, C.: Deducing Local Rules for Solving Global Tasks with Random Boolean Networks. Physica D 211(1-2), 88–106 (2005)
31. Motoike, I.N., Yoshikawa, K., Iguchi, Y., Nakata, S.: Real time memory on an excitable field. Physical Review E 63, 1–4 (2001)
32. Sielewiesiuk, J., Gorecki, J.: Passive barrier as a transformer of chemical frequency. Journal of Physical Chemistry A 106, 4068–4076 (2002)
33. Steinbock, O., Toth, A., Showalter, K.: Navigating complex labyrinths: Optimal paths from chemical waves. Science 267, 868–871 (1995)
34. Steinbock, O., Kettunen, P., Showalter, K.: Chemical wave logic gates. Journal of Physical Chemistry 100, 18970–18975 (1996)
35. Stuchl, I., Marek, M.: Dissipative structures in coupled cells: Experiments. Journal of Physical Chemistry 77, 2956–2963 (1982)
36. Teuscher, C.: Turing's Connectionism. Springer, London (2002)
37. Toth, R., Stone, C., De Lacy Costello, B., Adamatzky, A., Bull, L.: Dynamic Control and Information Processing in the Belousov-Zhabotinsky Reaction using a Co-evolutionary Algorithm. Journal of Chemical Physics 129, 184708 (2008)
38. Toth, R., Stone, C., De Lacy Costello, B., Adamatzky, A., Bull, L.: Simple Collision-based Chemical Logic Gates with Adaptive Computing. Journal of Nanotechnology and Molecular Computation 1(3), 1–16 (2009)
39. Turing, A.: Intelligent Machinery. In: Evans, C.R., Robertson, A. (eds.) Key Papers: Cybernetics, Butterworths, pp. 91–102 (1968)
40. Turing, A.: The Chemical Basis of Morphogenesis. Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences 237(641), 37–72 (1952)
41. Wolfram, S.: A New Kind of Science. Wolfram Media (2002)
42. Zaikin, A.N., Zhabotinsky, A.M.: Concentration wave propagation in two-dimensional liquid-phase self-oscillating system. Nature 225, 535–537 (1970)

Learning to Hypercompute? An Analysis of Siegelmann Networks

Keith Douglas

philosopher.animal@gmail.com

Abstract. This paper consists of a further analysis (continuing that of [11]) of the hypercomputing neural network model of Hava Siegelmann ([21]).

1 Introduction

This paper consists of a further analysis (continuing that of Douglas [11][1]) of the hypercomputing neural network model of Hava Siegelmann ([21]). It consists of three large sections. In the first section, a brief description of Siegelmann's model is presented. This section will be followed by a discussion of the merits of taking this model as a factual model (pace the "abstract" approach of Sieg [20]). Third, a discussion of one of Siegelmann's key, heretofore poorly explored, assumptions (the "linear precision suffices" claim) will be addressed and is the primary focus of the paper. This discussion will proceed along the following three subsections of analysis: it will discuss (1) a not-fully "noticed" suggestion of Arlo-Costa ([1]) and Douglas ([11])[2] that the Siegelmann network model actually requires a supertask to perform; (2) the merits of treating Siegelmann's proposal as one involving an idealization in the sense of Norton ([15]). The latter two will also allow a brief discussion of a factual interpretation of the arithmetic, analytic and similar familiar hierarchies; (3) that pace Davis and Scott ([9]) "non-recursive black boxes" are not exactly untestable, making use of the work of Kelly and Schulte ([16]). Subsections (2) and (3) are not independent and yield similar findings.

I end with summary conclusions. The conclusions will largely be more negative and "skeptical" about the merits of the Siegelmann network than those of herself or some of those (e.g. Copeland) who have defended it, but hope to provide more details on the areas of the model where interested parties could work on improving its plausibility and (nomological) possibility.

---

[1] This paper is dedicated to the memory of Horacio Arlò-Costa and, of course, to that of Alan Turing, who I would like to think would be astonished by the amount of work building on his we have today and by which the results are so omnipresent.

[2] I am not claiming such should have been noticed, per se, but I find it strange that these considerations have not made it into the discussion of our topic by critics of hypercomputing (or considered as a "devil's advocate" objection by proponents).

G. Dodig-Crnkovic and R. Giovagnoli (Eds.): *Computing Nature*, SAPERE 7, pp. 201–211.
DOI: 10.1007/978-3-642-37225-4_12

2 Siegelmann Neural Networks

Hava Siegelmann's monograph ([21]) is the principle source of detailed discussion of her model (hereafter SN), and includes much information about the computational strengths of various forms of neural networks, their properties from the perspective of computational complexity and much off topic for our present purpose. Subsequently here, we only need to focus on aspects that are unique or unusual in her presentation with regards to the question of computability. I assume the reader is familiar enough with usual artificial neural network models (hereafter, ANN) to follow the discussion involving such matters as nodes and weights (see, e.g.,[8]).

These unique/unusual aspects are: (1) the necessity of using an "output protocol", (2) her claims about the (real-valued) weights in the network and (3) the "sensitivity" of the network, i.e., a matter of interpreting the activation function.

This section of the paper will simply remind or inform the audience of these features as they do play a role later on, and not critically discuss them completely at this point. I bring these up to show that there are additional problems with SNs not discussed as much as the problem of weights already familiar in the literature and because they will play a crucial role in the discussions of idealizations later on.

In the case of "output protocol", what is meant is the convention adopted by Siegelmann (see in particular, [21], pp. 23-24) to indicate when a SN is finished its computation and is returning the value so calculated to its users. A state flag called the "output validation line" is set to 1 and is held in this value for the duration of the significant output and is set and held at 0 at all other times. One can then, during the time this line is 1, read off the value of the output from another, more conventional, output line. The "hypercomputing" part of this proposal is somewhat significant and yet hidden in her presentation.

In particular, how does this flag get set? Take the case of a recursively undecidable problem, for which these networks are supposedly useful at solving, like the halting problem for Turing machines (hereafter, TMs). In this case the output is a single encoded output, so in this case, the flag will be set to 1 at one "tick"[3] sometime in the future while (say) 1 comes out of the output line if the encoded TM halts and 0 otherwise. How does one know how to set this flag as a "programmer" of one of the networks? This depends on how the function is calculated, presumably. One has to know that the calculation is finished, that whatever state the network is in is the correct one. But this itself is a hyper-computational task; and so a regress seems to threaten.

Let us move on then to the case of the real valued weights of the network. This feature is the root of the hypercomputational power of the SN. Siegelmann does not tell us how these are to be obtained; merely calculates approximately how many digits of precision are needed after a given amount of run time. Since (by hypothesis) the network does not use registers, it is unclear what gaining digits of precision could refer

---

[3] This assumes the network is somehow equipped with a clock (which at least some ANNs do not have), but in the interest of manageability of this paper, I'll simply grant this part of the SN ex hypothesi.

to. An unspecified learning procedure is appealed to for the source of this extra precision, but without details this is simply an unknown as well. Notice that there are two concerns here - both the learning procedure and how its use gets “recorded”, “stored”, etc. are at stake. As for the activation functions, their “embodiment” or “implementation” also raises unanswered questions. For example, a threshold function (as the name suggests) is typically understood to be some representation of a node’s sensitivity to its inputs. In the case of SNs, these must be infinitely sensitive. Take a threshold function of the form (all of them discussed have the same problem; but since Siegelmann places special emphasis on the truncated linear one I use it here):

f(x) = 0 if x < 0
= x if 0 <= x <=1
= 1 if x > 1

To see the potential concern, consider a value of x = 0 + e, where e is some small value approximately (but not exactly) equal to zero. Represented in the usual notation, this is then some value 0.00000000000000000...1, say. The network has to be able to be able to “recognize” that value, no matter how small its difference is from 0, because the value of the output depends on it[4]. Siegelmann emphasizes truncation or rounding reduces the value represented at a node to a rational value and hence renders the computational properties of the network nonhypercomputational. I call the property of the nodes in question “sensitivity”, and as we have now seen, this is infinite in a real valued network (which allows literally any real value as a weight). Previous critics have pointed out the implausibility of finding (or knowing) that one had a hypercomputable weight in a SN (e.g., [9]); it is hopefully now clear that the problem is at least twice that, since one also needs a way for the network to make use of it, and that requires a “hypersensitive sensor” or something of the kind - subsystems that respond in infinitely precise ways to embody the activation functions. I might add in passing that this mistake or oversight is nothing new. Bunge ([5]) argues that a human brain is not suitably modeled by a TM because even a single electron can be in a continuum of states. Ignoring that this might prove too much, Bunge, like Siegelmann, has to argue that there can be an infinite number of (hyper)computationally (or, in Bunge’s case[5], cognitively) relevant states and events (state transitions: [10]).

---

[4] Consider the required difference in output from two nodes that differ in value by 2e (e.g., one 0+e and the other 0-e). One of these will have activation 0 and the other e. It is also interesting to reflect that a relatively informal presentation of ANNs like in [8] the weights are also described as being real-valued, but nothing in their presentation hinges on it. Presumably it makes explaining the mathematics easy and ensures that a digression about computable PDEs is irrelevant. Presumably also Churchland and Sejnowski regard the plausibility of any real number as a weight to be not worth considering. Note also that the learning algorithms they discuss (pp. 96 ff.) are computable as well.

[5] I will not press the point here, but from my experience in conversation with Bunge (in the late 1990s); he does not think human brains are hypercomputers: rather, he thinks that computational notions are inapplicable to them altogether. The view, although he would be horrified by the comparison, seems to be similar to that of Wittgenstein. But this is all for another time.

3 Sieg (Indirectly) on Siegelmann

Sieg ([20]) has argued (in the context of a discussion of the Church-Turing thesis) that one can dispense with said thesis and instead:

"My strategy, when arguing for the adequacy of a notion, is to bypass theses altogether, and avoid the fruitless discussion of their (un-)provability. This can be done by conceptual analysis, i.e., by sharpening the informal notion, formulating its general features axiomatically, and investigating the axiomatic framework."

This viewpoint dispenses with the need to analyze the Siegelmann network in detail, at least for the present purposes - were it correct. It would make it clear that hypercomputation is doomed to failure as a subject as the axiomatic framework in question makes it perfectly clear that broadly computational devices (including potential hypercomputers[6]) do not include anything like the SN[7].

However, as has been pointed out by Stewart Shapiro ([19]), it does not appear that Sieg successfully dispenses with theses here. In other words, there is the question of whether or not the axioms are correct[8] of the real (or non-abstract, non-Platonic, etc.: replace as necessary according to your favourite philosophy of mathematics) systems under consideration. How do we (hopefully) resolve this impasse? For if Sieg is right, there is nothing to investigate; we simply see that the SNs do not fall under the axioms of computing machines he has usefully provided and that would be the end of it. This seems too hasty for the present purpose, so the concern is pressing.

Here is where Shapiro is mistaken; he thinks that (following Quine [18] and others) one is dealing with some matter which is both mathematical and empirical. For some (perhaps for Sieg) this is impossible or unlikely; instead it is like investigating axioms for (say) groups[9]. If Sieg were right it would be a matter of getting (as he borrows a

---

6 Nothing in the Sieg-Gandy presentation actually rules out accelerated Turing machines ([2]) for example. However, it is unlikely at best that either Sieg or Gandy would approve; the advantage to the SN over many models of computation is that it explicitly includes a clock (a feature it admittedly shares with some ANN models) and thus can be used to more precisely make claims for or against "tricks with time" like the accelerated Turing machine requires. I'd hazard a conjecture that such a machine also requires no lower bound on the size of its parts if described as a Sieg-Gandy machine, and hence runs afoul of the finiteness requirements that way, but such an argument would require delicate physical hypotheses I do not wish to address in the present work.

7 Since I disagree with Sieg that this approach is suitable, I shall not investigate precisely (in the present paper) where Sieg-Gandy machines rule out SNs, however it seems likely they run afoul of the "finite parts" conditions. Sieg and Gandy represent parts by the hereditary finite sets, so, presumably, a similar approach to the SN would need to use hereditary countable sets. This seems to suggest either or both of an infinite number of parts or an infinite magnitude of a property of one.

8 I suspect that Sieg would claim that there is no thesis involved here; one simply investigates whether or not the axioms are fruitful, lead to desired consequences, etc. However useful that approach is for his and many other very important purposes, it amounts to begging the question against hypercomputation without further ado.

9 Using an algebraic analogy here, as opposed to (say) using a geometric one is important. By contrast, say, analysis would lead to questions immediately about the "real" continuum and whether spacetime is or could be discrete; geometry raises similar questions about dimensionality, curvature, etc.

phrase from Hilbert in saying) the "Tieferlegung der Fundamente" right; Shapiro claims instead one has to look to the world too. I claim both are mistaken because they have overlooked the possibility that the matter is not about mathematics at all.

I argue that the debate should be construed as one about doing mathematics (or at least doing calculations or computations). Turing, as Sieg has rightly emphasized, analyzed a human "computor" (in the sense of Gandy [13]). Similarly, Gandy, him, and others have analyzed calculations by machine as well. Using Bunge's ([3]) theory of reference and Sieg's admirable presentation ([20]) of "Gandy machines", one sees that the theory of Gandy machines is, indeed, about computing machines. This makes the subject matter a factual[10] one in Bunge's sense[11]; see also Douglas [12]. In other words, it is not a matter of mathematics - one can (and should) use mathematics as a tool to describe the characteristics of the computers and computors, but this does not make the field mathematics anymore than using differential equations in the theory of reaction mechanisms makes chemistry a branch of mathematics.

Hence Shapiro is right in his claim: Sieg does not dispense with theses - or, if preferred, Church's thesis is in need of "empirical" confirmation and hence SNs' "usefulness" as a model of computing cannot be dismissed so hastily. Also hence in particular, we must address the question of whether SNs are empirically plausible. It is here that we run quickly into previous criticisms of her proposals from the likes of Martin Davis and Dana Scott.

Davis' ([9]) paper quotes Scott concerning how we would recognize a "nonrecursive black box". I feel this quotation is also slightly mistaken: it proves too much. I agree that no finite amount of interaction with a black box could show that it performs hypercomputational processes. However, no finite amount of observation could tell you that a black box contains a Turing machine. Any finite experimentation with input and output is consistent with the black box being a (perhaps very large) finite state automaton[12]. This is not to say Scott and Davis are mistaken concerning the difficulty of determining that one has a hypercomputer of some kind, but instead that it is important not to overstate this difficulty. He emphasizes how hard it would be to tell that one had a non- recursive "transparent box" (i.e. a black box with much of its workings well known). It seems to me that Scott and Davis have adopted almost an instrumentalist attitude towards (what Bunge would call factual) scientific theories here. Since instrumentalism is controversial amongst philosophers of science, we

---

[10] Bunge ([3], [5]) is a mathematical fictionalist and contrasts factual to formal sciences; once again one can translate into one's appropriate philosophy of mathematics idiom. The important matter is that group theory is not the correct analogy; instead, a theory in (say) chemistry - like (say) a theory of solutions - is a better comparison. Using physics would raise questions about "rational mechanics" that might prolong the debate unnecessarily and other sciences would raise equally irrelevant questions for our present purpose. I will use "factual" in his sense throughout.

[11] Bunge (see, e.g., [4]) would claim that this use of "empirical" (traditional in most philosophy of science) is wrong, however, I shall use it here to emphasize what I intend.

[12] Matters are actually not quite this simple. See below about [15]).

should be wary of this approach[13]. After all, how does (say) Newtonian dynamics (ND) get verified? This presumably factual theory uses continuous functions and such; whereas any measurement is only of finite precision and hence renders direct confirmation impossible. Davis and Scott thus "prove too much" with this approach. They might rejoin that one could state ND in terms of computable analysis. However, assuming it could be done in this case does not show it could be done in general. Also, since the theory is then different, how does one decide between the computable version and its (usual) noncomputable counterpart? It would seem one would have to apply more general principles about the nature of theories in (factual) science. Since these are arguably under debate, we are now back where we started.

Nevertheless, Davis and Scott have correctly (in my view) treated SNs as to what sort of proposal they are - namely a family of factual hypotheses. I have mentioned earlier (section 1) that there are what one might call "nomological" areas of discussion (problems, counterproposals, etc.) with the SN approach. I now turn to three of these.

4 Nomological Considerations about "Linear Precision Suffices"

The first of these stems from Arlò-Costa ([1]) and adapts prior remarks of Douglas ([11]) to that end. He asks whether or not the SN require a supertask to implement and hence "inherit" the implausibility of the accelerated Turing machine (see, e.g., Boolos and Jeffrey [2]) which most would agree is a purely "notional" device. In particular, note the difficulty even in computing a constant function with a SN. Since the weights of each node in a SN are of infinite precision, outputting their value directly is impossible by the protocol described. This arises because such a constant is still an infinite precision number, and so outputting its value requires an infinite amount of time[14], followed by a signal to indicate that the output is finished. At best this would require a supertask. A suitable re-encoding[15] would have to be found, and that is not suggested anywhere by Siegelmann. Moreover, such would have to handle rational

[13] Disclosure: As may be noticed, I am a scientific realist (of a somewhat unnamed sort), so I have (what I take to be) good reasons against instrumentalisms. But to be charitable to such esteemed scholars as Scott and Davis, I have tried to avoid dismissing their seemingly instrumentalist views out of hand and tried to find a way to allow both them and Siegelmann the benefit of the doubt about the plausibility of certain hypotheses.

[14] That is, unless one could in every case "program" the Siegelmann network to tell when it had an irrational number and flag rationals appropriately. This ability itself seems to be hypercomputational.

[15] Siegelmann's book spends a lot of time talking about Cantor sets and changes in number bases, etc. As far as I can tell, qua engineering proposal (and one does take SNs as such when one takes them factually, as we are doing) this is largely irrelevant without knowing what physical properties does the representation in the engineering sense. Obviously no registers are involved, and so re-encoding is not well defined at present. This problem is (needless to say) another instance of the same one that we keep encountering: how *do* the weights work?

values as well as surds, transcendental values, and even non-Turing computable numbers, like Chaitin's constant. Of course, giving a finite representation of the latter sort of value cannot in general be done. My earlier remarks about the "output protocol" loom large here.

Similarly, if the precision of an infinitely precise real number is not available at the beginning of the run of a SN, and the precision increases uniformly in time (see further my discussion of "learning networks") it will take an infinite amount of time for the network to become infinitely precise. This entails immediately that the networks are actually Turing-equivalent for any finite period of time. Here, let me note further that this puts Siegelmann's model in an unfortunate dilemma much as the precision consideration proper above provokes. Once again, either the network is infinitely precise in finite time (throwing away the "linear precision suffices" result), in which case the Siegelmann network is implausible from the sensitivity considerations I have canvassed and from related concerns, or it is only infinitely sensitive in infinite time, in which case using it to perform super-Turing computations would again require a supertask. Thus, it seems that Arlò-Costa is in fact correct, though a proof would be nice to have - but in the interests of time I have omitted such. I thus turn to a question which stems also from [11]).

This concerns the nature of idealizations and approximations. It might be argued that the critics of SNs are taking the model too literally. Instead, it should be treated as one involving either an idealization or an approximation. For example, a SN-fond opponent of the critics of hypercomputing may well point out the critic will ask: why should we not grant relevant idealizations to the Siegelmann network? after all, the TM itself (or, equally, a "Gandy-Sieg machine") makes idealizations concerning computing agents and their resources. For example, these are held to have an unbounded amount of memory, do not break down ever, can calculate without running out of energy no matter how long they run, etc. So, the opponent asks, why not grant idealizations to the SN?

One could attempt to respond to this opponent by (1) counting idealizations or (2) intuitively trying to evaluate their merits and plausibilities. In the case of (1) it is likely correct to conclude that the SN includes all those of the TM and then some, it does not seem fruitful to simply claim that the SN has more and hence is more implausible, for what if the TM was already regarded as sufficiently implausible to not merit adoption? Or, in other words, does this objection prove too much? Also, how does one know what is "too many" idealizations and approximations anyhow? Better to look at (2), instead.

To focus attention, let us discuss one particular family of idealizations, that of the weights of the nodes[16] in the SN. Norton ([15]) has circulated a manuscript on idealizations which is useful to apply to the present purpose (cf. also the brief remarks in [14] and [7]). Norton's paper centers around what he has called "the problem of limits", distinguishing between the case when the limit property and the limit system agree on the one hand, and when there is no limit system on the other. I shall argue,

[16] There is an "informal duality" between the weights and what I have called "sensitivity". All the arguments I raise, as far as I can tell, apply to it as well.

based on considerations based on the arithmetic hierarchy (e.g., [16], pp. 362 ff.) that Siegelmann's proposal falls into the latter category. This is because at any finite time, a Siegelmann network is equivalent to a TM in power; only "at" the limit of infinite time is the network super-Turing in power.

To adopt Norton's analysis, one thus has to identify the limit property of a SN and what the limit system could be (if there is one). As just suggested, the limit property is the node weight (it does not matter if we take all the nodes or one, as by interleaving their expansion one can see there is effectively only one "weight" anyway[17]). Let us also assume for definiteness that this weight is Chaitin's constant, ϕ[18]. Then the question becomes how the weight gets used in an idealization. Here I do not know what to do to proceed. The weight is still not explained: in order to evaluate the idealization we have to know what this property actually is. For the sake of definiteness again, assume that we are dealing with a length. Then the idealization involves that of limiting lengths, the idealized value of which is ϕ. This is a case where the limit property is not a property of the limit system. This is because the increases in length precision do not correspond at all the steps on the arithmetic hierarchy: if the idealization here were plausible then each finite increase in precision would correspond to some finite n in the usual labeling of the hierarchy[19]. One simple "imaginable possibility" would be to think that a weight of precision n corresponded to a "strength" of $\Delta^0_{f(n)}$ where f(n) is some finite (i.e., non-divergent) function of n so that an infinite precision would deliver Δ^0_ω as required[20]. But this does not work, no matter how slowly the function f grows, as Δ^0_0 is already the recursive functions. In that sense, since each

[17] For example, taking the digits from each weight value: 0.a1a2a3a4... and 0.b1b2b3b4... becomes 0.a1b1a2b2a3b3a4b4... in the case of a two node network. A similar procedure can be done for any SN, as they all have a finite number of nodes.

[18] A referee asked if we know that this particular constant is calculable. Since it is representable as a function from natural numbers to natural numbers, it is. SNs can calculate all such functions ([21], pp. 59 ff.).

[19] It is vital not to get the order of this (very impressionistic) proposal wrong. Since SNs grow in computational power by increasing precision, and in the conventional theory of computation, more computational power means "climbing" the arithmetic hierarchy, all I am suggesting is how one would have to reconcile these aspects.

[20] There is some potential confusion here, so a clarification is true. It is quite correct to point out that this level of the hierarchy is "infinitely more" than is needed to do super-Turing computations. But precisely what is wrong (in one way) with the SN proposal is that it skips that entire part of the hierarchy. Why? Because that's precisely what a jump from finite precision real numbers (i.e., rational weights) to the full infinitely precise weights of the SN does. On the one hand one needs the infinite precision; on the other hand the jump is an inappropriate idealization for that reason. There is no way an SN, as described, can climb through the hierarchy over finite time and wind up, in the infinite limit, at the ability to calculate all functions from natural numbers to natural numbers. Any finite increase in precision adds at best a finite ability. It is not as if some magic bit, say at the 31337th decimal place, suddenly allows all the sigma 0 1 functions to be computed, etc.

finite n is still sub-recursive, the limit in question of the SNs is then the recursive functions and not the arithmetic ones the SN need, never mind the analytic ones (i.e. all the functions from natural numbers to natural numbers). Hence, the property of the limiting system and the limit property do not agree. Hence further it looks like an inappropriate idealization in Norton's sense. We thus have a clearer way of stating the difference over the idealizations of the SN versus those of the TM.

These brief appeals to the arithmetic hierarchy also allow an answer to Scott and Davis (above) and avoid protracted debates over operationalism. A non-recursive "oracle" is indeed hard to investigate; however, Kelly and Schulte ([15]) draw important connections between the arithmetic hierarchy and the learning of theories with uncomputable predictions. While I will not prove any results here, I suggest that rather than an "operationalist" response to Siegelmann, one can in principle give a learning-theoretic answer at least for some possible uses of the network. The goal in this section is to answer residual worries about operationalism and merely gesture at an area of future investigation, particularly connecting the properties of Siegelmann network to other hypercomputing proposals as has been suggested by the Wikipedia contributors ([22]).

Let us turn to specifics. Kelly and Schulte ([15]) classify (following Gold and Putnam) hypercomputational problems into learning theoretic classes. For example, a hypothesis of the form Π_1^0 is one which is "refutable with certainty". However, what is interesting from the perspectives of this paper (and symposium) is that a Δ_3^0 sentence is sufficiently "complex" that there is no way to investigate it in a computable way[21]. This is "infinitely far" away from the level $\Delta_\omega^0 \ (=\Delta_0^1)$ that characterizes the complete SNs. However, a brief look at learning again might prove useful. If the increase in precision of real valued weights increased through the learning-theoretic hierarchy in a useful way - say, some fixed bound moved the strength of the system from Δ_1^0 to Π_1^0, that would be a useful finding. Unfortunately, it seems to be nowhere in offering, once again for the same reason. To reiterate: any finite bound in increase of precision of the networks preserves their behaviour vis-a-vis the arithmetic hierarchy. This makes it implausible, to say the least, that SNs could increase their precision in a relevant way by "learning" as she proposes (without a supertask). This is not to say that real valued weights in a network could not increase precision by external influence (learning) but rather that they could not do so in a relevant way that makes hypercomputation plausible (or nomologically possible).

The lesson for this subsection is then: Scott and Davis are right to be skeptical of our abilities to investigate purported capabilities of a supposed non-recursive black box. However, they are wrong to say that it is impossible in principle, but SNs pro-

[21] Claiming to investigate it in a hypercomputable way would of course beg the question against the critic of the SNs and also be useless for the proponent of them. After all, if one has a known "Δ_ω^0 device" or procedure already, why use a SN?

vide no way for this investigation to proceed. Partisans of hypercomputation wanting to answer Scott and Davis must look elsewhere (including refining their models).

5 Conclusions

Investigations into the arithmetic hierarchy-related properties of Siegelmann style networks show how they are implausible relative to a Turing machine model of computation for they invoke various versions of the same inappropriate idealization. In future work, I hope to discuss whether any models of hypercomputation meet these requirements. I also hope to provide more details in these areas of specific criticism and also more rigorously analyze the Turing machine model from the perspective of idealizations.

References

[1] Arlò-Costa, H.: Unpublished comment in discussion with the author (2003)
[2] Boolos, G., Jeffrey, R.: Computability and Logic (2e). Cambridge University Press, New York (1980)
[3] Bunge, M.: Sense and Reference. Treatise on Basic Philosophy, vol. 1. Reidel, Dordrecht (1974)
[4] Bunge, M.: Epistemology and Methodology I: Exploring the World. Treatise on Basic Philosophy, vol. 5. Reidel, Dordrecht (1983)
[5] Bunge, M.: Epistemology and Methodology III: philosophy of science and technology: Part I. Formal and Physical Sciences. Treatise on Basic Philosophy, vol. 7. Reidel, Dordrecht (1985)
[6] Bunge, M.: Epistemology and Methodology III: philosophy of science and technology: Part II: Life Science, Social Science and Technology. Treatise on Basic Philosophy, vol. 7. Reidel, Dordrecht (1985)
[7] Bunge, M.: Philosophy of Science (2 vols.). Transaction, New Brunswick (1998)
[8] Churchland, P., Sejnowski, T., Terrence: The Computational Brain. MIT Press, Cambridge (1996)
[9] Davis, M., (adopting remarks of D. Scott). The Myth of Hypercomputation. In: Teuscher, C. (ed.) Alan Turing: Life and Legacy of a Great Thinker. Springer, Berlin (2006)
[10] Douglas, K.: A Special Davidsonian Theory of Events. Unpublished MA Thesis, Department of Philosophy, University of British Columbia (2001), `http://www.philosopher-animal.com/papers/thesis.pdf`
[11] Douglas, K.: Super-Turing Computation: A Case Study Analysis. Unpublished M.S. Thesis, Carnegie Mellon University (2003), `http://www.philosopher-animal.com/papers/take6c.PDF`
[12] Douglas, K.: What Does a Computer Simulation Have to Reproduce? The Case of VM-Ware. Unpublished paper presented at IACAP 2010
[13] Gandy, R.: Church's Thesis and Principles for Mechanisms. In: Barwise, J., Keisler, H.J., Kunen, K. (eds.) The Kleene Symposium. North-Holland, Amsterdam (1980)
[14] Kelly, K., Schulte, O.: The Computable Testability of Theories Making Uncomputable Predictions. Erkenntnis 43, 29–66 (1995)
[15] Niiniluoto, I.: Critical Scientific Realism. Oxford University Press, Oxford (1999, 2004)

[16] Norton, J.: Approximation and Idealization: Why the Difference Matters. (unpublished manuscript) (2011), http://www.pitt.edu/~jdnorton/
[17] Odifreddi, P.: Classical Recursion Theory: The Theory of Functions and Sets of Natural Numbers. Elsevier Science, Amsterdam (1989)
[18] Olszewski, A., Woleński, J., Janusz, R. (eds.): Church's Thesis after 70 years. Ontos Verlag, Heusenstamm (2006)
[19] Quine, W.V.O.: Two Dogmas of Empiricism. The Philosophical Review 60, 20–43 (1951)
[20] Shapiro, S.: Computability, Proof, and Open-Texture. In: [17]
[21] Sieg, W.: Calculation by Man and Machine: Mathematical Presentation. Technical Report No. CMU-PHIL-105 (2000)
[22] Siegelmann, H.: Neural Networks and Analog Computation: Beyond the Turing Limit. Birkhäuser, Boston (1999)
[23] Wikipedia contributors. "Hypercomputation" in Wikipedia, The Free Encyclopedia, http://en.wikipedia.org/wiki/Hypercomputation (accessed January 29, 2012)

Oracle Hypermachines Faced with the Verification Problem

Florent Franchette

IHPST, University of Paris 1 Panthéon-Sorbonne
Florent.Franchette@malix.univ-paris1.fr

Abstract. One of the main current issues about hypercomputation concerns the claim of the possibility of building a physical device that hypercomputes. In order to prove this claim, one possible strategy could be to physically build an oracle hypermachine, namely a device which is be able to use some extern information from nature to go beyond Turing machines limits. However, there is an epistemological problem affecting this strategy, which may be called "verification problem". This problem raises in presence of an oracle hypermachine and it may be set out as follows: even if we were able to build such a hypermachine we would not be able to claim that it hypercomputes because it would be impossible to verify that the machine can compute a non Turing-computable function. In this paper, I propose an analysis of the verification problem in order to know whether it is a genuine problem for oracle hypermachines.

Keywords: computability theory, hypercomputation, oracle hypermachine, Turing machine, verification problem, randomness.

1 Introduction

Hypercomputation is a term that denotes the computation of functions, which are not computable by a Turing machine (non Turing-computable functions) [8]. Current researches about this notion primarily concern its possibility, either conceptual or physical.

On the one hand, the conceptual possibility of hypercomputation is not really contested because several machines able to hypercompute - called "hypermachines" - have been formalized [8], [10], [16], [27], [29]. This conceptual possibility means that effective procedures are not the only methods to compute mathematical functions and that hypermachines can be devised to compute more functions than the Turing machine[1].

On the other hand, the physical possibility of hypercomputation is not fully accepted within scientific and philosophical communities [12], [21], [24]. Although

[1] By "the Turing machine" we want to say "a universal Turing machine", namely a Turing machine that is able, according to the Church-Turing thesis, to compute all functions $f : \mathbb{N}^k \to \mathbb{N}$ computable by effective procedures. It is worth noting that hypercomputation and the Church-Turing thesis are not in opposition with each other, because the latter only concerns effective methods by contrast to the former.

G. Dodig-Crnkovic and R. Giovagnoli (Eds.): *Computing Nature*, SAPERE 7, pp. 213–223.
DOI: 10.1007/978-3-642-37225-4_13

numerous hypermachines have been devised from a logical point of view, current issues are more about the physical domain. In particular, one of these issues directly concerns the physical construction of a hypermachine and the truthfulness of the claim that I will call the "physical hypercomputation thesis": it is physically possible to build a device that hypercomputes.

In order to prove the physical hypercomputation thesis, two strategies have been proposed [21]:

Definition 1 (Strategies P and O)

- According to strategy P (where P stands for programmable), the device should be able to compute a non Turing-computable function from its internal program [19], [25], [26].
- According to strategy O (where O stands for oracle), the device should be able to compute a non Turing-computable function with the help of an extern information coming from nature [2], [10], [28].

Propositions based on the strategy P come from the works of Pitowski and Hogarth [17], [25]. More precisely, they showed that the computation of an infinite number of steps is consistent within a particular relativistic space-time called "Malament-Hogarth space-time". Later, Shagrir and Pitowski devised a hypermachine based on the Malament-Hogarth space-time, which is able to compute the halting function for Turing machines [26]. However, even if it appears that some space-times can be regarded as realistic, it is unclear whether kind of devices could ever be constructed to take practical advantage of their properties [11], [14].

On the contrary, the strategy O is more promising because devices based on the use of an oracle seem to be closer to be built [2], [28]. Nevertheless, there is still an epistemological problem affecting the strategy O, which may be called "verification problem" [8], [20], [28]. This problem raises in presence of an oracle hypermachine and it may be set out as follows: even if we were able to physically build such a hypermachine we would not be able to prove the physical hypercomputation thesis because it would be impossible to verify that the machine can compute a non Turing-computable function.

In this paper, I propose an analysis of the verification problem in order to know whether it is a genuine problem for oracle hypermachines. In section 2 I will start by illustrating two propositions concerning the physical construction of an oracle hypermachine; then, in section 3, I will deal with the verification problem in itself.

2 How to Build an Oracle Hypermachine?

Before explaining how to physically build an oracle hypermachine, it is first necessary to understand how it works in theory.

The idea of computing with the aid of an oracle comes from Turing's work [30]. Turing is indeed behind a kind of machine called "oracle Turing machine" or "O-machine", which is a Turing machine equipped with an oracle, namely a black box whose behaviour is not specified. The particularity of the oracle lies in its capacity to provide some non computable functions results to the machine:

> "Let suppose that we are supplied with some unspecified means of solving number-theoretic problems; a kind of oracle as it were. We shall not go any further into the nature of this oracle apart form saying that it cannot be a machine" [30, p. 167].

Thus, according to its architecture and computational power, the O-machine is not a standard model of computability but a hypermachine.

However, the O-machine is not detailed enough to perfectly understand how it is able to hypercompute. It is why Copeland and Proudfoot have proposed a further definition of that hypermachine [10]. From their point of view an O-machine is a Turing machine which has two more elements: first, a device - the oracle - able to make measurements with an infinite precision; secondly, a memory space containing a real number called "τ". τ is an infinite binary string, which represents the results of a non Turing-computable function. Thus, if such non Turing-computable function is denoted by d, the nth symbol of τ represents $d(n)$, namely 0 or 1. For example, if we want to have acces to $d(239208)$ the device measures the symbol number 239208 and provides the corresponding value.

Now, the first practical step in order to build an O-machine is to find a physical theory in which a device will be able to use some extern information from nature. Such an information would be regarded as an oracle that provides an additional element to go beyond Turing machines limits. Nowadays, there are two proposals that go in this direction: one is based on randomness from radioactive decay [28], while the other on quantum randomness [2]. The interesting fact is that both of these propositions regard randomness as a source of non computable information, which could be used to hypercompute.

This very idea is not really a new one; already Turing pointed out that a machine equipped with a random element could do more than the Turing machine [7]. More precisely, there are two types of random processes: true-random processes and pseudo-random processes. Pseudo-random processes generate strings of numbers from pseudo-random methods (for example the linear congruence method), which numbers "appear" random but that are actually provided by algorithms. An immediate consequence is that machines using pseudo-random processes are equivalent to Turing machines [13]. By contrast, if a machine is able to generate an infinite true-random string of numbers[2], it would provide the results of a non Turing-computable function. Indeed, a simple cardinality consideration shows that, with probability one, the sequence produced is not Turing-computable [3]. There are uncountably many infinite strings of digits and even more strongly, there are uncountably many infnite strings of digits with any given limiting frequency of 0's and 1's. But there are only countably many Turing-computable strings. Therefore, assuming that each infinite string (or each infinite string with a certain limiting frequency) has the same probability of occuring as a result of a random process,

[2] The definition of an infinite true-random string of numbers is both technically and philosophically complex. For further details see [4], [5], [22].

the probability that a random process would generate a Turing-computable string of digits is zero, whereas the probability that the string is not Turing-computable is one.

On the one hand, Stannett described a hypermachine that uses a radioactive sample to generate such an infinite true-random string of numbers [28]. More specifically, this sample decays by "α-radiation", a process in which Helium nuclei are ejected from the nuclei of atoms. The Helium nuclei combine with electrons in their environment to form atoms of Helium gas that can be pumped out of the experimental apparatus, and as a result the sample becomes progressively lighter. In order to generate a random integer, it is necessary to examine the weighting scales to determine how much mass has to be lost by the sample for the reduction to be detectable. Then, large amount of sample material is collected to ensure that when all of it has decayed, the weight loss will be easily noticeable. One chooses a "threshold" value somewhere between the minimum mass loss detectable by the scales, and the total possible mass loss of the sample, and then set a clock running. The number that is generated by the system is defined to be the number of complete seconds that pass before the system generates the chosen amount of mass loss.

Nevertheless, underpinning the system's behaviour is the standard assumption (considered by physicists to be valid for all radioactive materials) that there eventually comes a time when half of the sample can be expected to have decayed (the average time required is called the "half-life" of the sample in question). Because the threshold mass is strictly less than the total possible mass, it should eventually be reached after only finitely many half-lives, but because decay is assumed in the standard model to occur randomly, the number of seconds that pass before this happens must also be random. In other words, this system implements a true random number generator.

On the other hand, Calude's proposition is simpler to explain and consists of fixing on a computer a device able to generate a string of random numbers from a quantum process [2]. The idea of finding this information from quantum randomness comes from the standard model of quantum physics. Precisely, quantum randomness is postulated from the Born postulate, which is the idea that a measurement of a particle will yield a result which follows probability distribution $|\psi|^2$, where ψ is the particle's wave function. In particular, the ID quantique company[3] has created a device whose name is "Quantis", which generates a string of random numbers from an elementary quantum optics process [18]. More specifically, photons are sent one by one onto a semi-transparent mirror and detected. The exclusive events (reflection - transmission) are associated to "0", "1" bit values and each of them have a probability at 50% to occur. The operation of Quantis is continuously monitored to ensure immediate detection of a failure and disabling of the random bit stream.

In theory, a device equipped with Quantis might provide an arbitrarily long string of quantum random strings. However, this device should be considered as a hypermachine only if the quantum random string cannot be generated by

[3] `http://www.idquantique.com/`

a Turing machine, that is to say only if the string includes an infinite number of bits. In that case Quantis would be seen as an oracle able to provide non computable information from nature.

Although the physical construction of these two devices based on randomness seems to be sufficient to prove the physical hypercomputation thesis, an epistemological problem nevertheless remains. This problem is raised when we have an oracle hypermachine: even if we build such a hypermachine we will not be able to prove the physical hypercomputation thesis because it would be impossible to verify that the device is able to compute a non Turing-computable function. I am going to analyze this problem in further details in the next section.

3 The Verification Problem

Suppose to have an oracle hypermachine, how can we verify the results provided by the machine? In practice, this problem does not seem to be different from the verification of results provided by standard computers. Indeed, take a particular function as an example (the argument works no matter which function is considered). Let p the function defined by $p(n) =$ the nth decimal of the expansion of π. It is easy to check that we cannot verify in practice (due to a lack of resources) whether the 10^{12}th decimal of π recently computed by a computer is 5. So why verification would be a real problem in the case of hypercomputation?

According to Copeland, the difference lies in principle and not in practice, namely in the case where we disregard physical computational resources:

> “ There is an epistemological problem with the idea of hypercomputation. Suppose Laplace’s genius says ’Here is a black box for solving the Turing-machine halting problem’ (The problem arises no matter which non Turing-machine-computable function is considered.) Type in any integer x and the box will deliver the corresponding value of the halting function $H(x)$ or so Laplace’s genius assures you. Since there is no systematic method for calculating the values of the halting function, you have no means of checking whether or not the machine is producing correct answers. Even simulating the Turing machine in question will not in general help you, because no matter how long you watch the simulation, you cannot infer that the machine will not halt from the fact that it has not yet halted” [8, p. 471].

More specifically, from a point of view in principle, and since standard computers can be studied from its theoretical model, namely the Turing machine, it is possible to verify that a standard computer provides a correct result. Indeed, according to the definition of an effective procedure we can have access to results in principle by checking step by step the computation from the input to the output. On the contrary, we could not proceed in the same way with an oracle hypermachine because we would not be able to check each computational step because of the absence of an effective procedure. Hence we would not be able to prove in principle the hypercomputational power of the device.

Although Copeland's thought experiment could call into question that devices using oracles really hypercompute, it is nevertheless the case in practice that we need to consider to answer whether the physical hypercomputation thesis could be challenged. According to a verification in practice, namely a verification which takes into account physical computational resources, one solution has been suggested to overcome the verification problem.

This solution was brought both by Shagrir & Pitowski and Cleland and it is to claim that computation does not presuppose verification [6], [26]. More precisely, they claim that since we regard a function as computed by a standard computer, even if we are not able to verify in practice the provided results, the same should be true for hypercomputation. As we mentioned at the beginning of this section, it is indeed impossible in practice (because of a lack of resources) to verify that a computer correctly computes a given function. However, we would be inclined to say that computers do compute functions and, moreover, we develop a real confidence in computers; otherwise why would they be the grounds for the construction of our transport networks, economy and energy systems? But where does such a trust come from? On the one hand, Cleland explains this confidence arises from the fact that it is possible to use some empirical methods such as probabilistic causal relations or parallel computations and theoretical methods such as program verification [6, p.224]. Such tests are used to claim in a plausible way that a computer computes a function even if no perfect verification is possible. In the other hand, Shagrir and Pitowski claim that what computers do, namely computations, can be explained in terms of a formal or physical background theory :

> "We can use our physical theories to explain what function is being computed. Our physical theories mandate that a device with such-and-such a physical structure and initial conditions will behave according to a given set of equations, whose solutions it computes" [26, p. 91].

According to this passage, we can use in the case of computers both program verification to check their computational behavior and physical theories to ensure that programs are correctly physically implemented. Therefore, computation does not presuppose a perfect verification.

In summary, the verification problem should not to be considered as a thought experiment as Copeland did but as an empirical hypothesis. In other words, the problem could not be solved as long as one will assume the physical construction of an oracle hypermachine; on the contrary we must dispose of an oracle hypermachine physically build to achieve theoretical and physical tests, and to claim with a high confidence that it computes a non Turing-computable function. In this way, the verification problem should not be a real problem for hypercomputation.

I do not entirely agree with this conclusion. In my opinion, the solution proposed by Shagrir, Pitowski and Cleland only works for a particular kind of hypermachine, namely programmable hypermachines, but fails to dissolve the verification problem in the case of oracle hypermachines.

As their name suggests, programmable hypermachines have a remarkable property: they are programmable. This property means first that we can study their computational behavior from their internal program. In particular, take the relastivistic hypermachine, which is able to compute an infinite number of steps from Malament-Hogarth space-times properties. One possible physical realisation of this hypermachine consists of two modern computers T_A and T_B that can communicate with each other. From the properties of Malament-Hogarth space-times, when T_A will have computed a finite number of steps, T_B will have computed an infinite number of steps. Now here is the procedure that allows to compute the halting function h:

1. We begin by providing the input n to T_A and transferring it to T_B. T_B is an universal Turing machine, that is to say its program allows to simulate the computation of the nth Turing machine on input n.
2. During the computation, if T_B halts then it immediately sends a signal to T_A. Otherwise, it sends nothing to T_A.
3. Finally, when T_B has achieved an infinite number of steps in a finite time, T_A will print 1 if it has recieved a signal from T_B and will print 0 otherwise.

We can notice then that program verification for this kind of hypermachines is easy because it reduces to the verification of the fact that we have correctly implemented the program of an universal Turing machine on a modern computer.

Secondly, the program of programmable hypermachines can be used to increase the plausibility they are computing a non Turing-computable function. Assume we want to decide whether or not a Diophantine equation has a solution[4]. Even though there is no effective procedure that decides this problem [23], we can nevertheless implement on a modern computer a simple program using the brute-force search method (to test each integer and to find the solution if it exists). Therefore since T_B (a modern computer that belongs to the relastivistic hypermachine) is able to compute an infinite number of steps in a finite time, it will be able to cover all integers in a finite time in order to find an integer satisfying a given equation. Now, if we feed the device with an input, which is the code of a Diophantine equation for which we still do not have a solution and the device gives as output that the equation has no solution, then the plausibility that the device hypercomputes will be increased.

However, even if the solution proposed by Shagrir, Pitowski and Cleland seems to reasonably work for programmable hypermachines, it is a different issue in the case of oracle hypermachines. Here are four arguments why the verification problem is a real problem for this kind of hypermachines.

First, we cannot use the method which has been set out above in order to increase the plausibility that a programmable hypermachine computes a given function. The reason is that we do not know the function that is computed by an oracle hypermachine based of randomness. Indeed, since the hypermachine provides a random string of numbers, there is no way to define the function

[4] A Diophantine equation is an equation whose cœfficients and solutions are integers numbers.

$f : \mathbb{N}^k \to \{0, 1\}$ whose values are being generated by the random process [24]. Of course, f exists in the set-theoretic sense of a set of pairs whose output values process happens to emit. But in this context, defining f means actually specifying the relationship that obtains between the arguments and values of f, as we do when we define, for instance, the halting function for Turing machines. If the process is genuinely random, there is no way to specify f without generating the values of f by running the process. Therefore the computation of an oracle hypermachine gives no hint about the plausibility of having computed a particular non Turing-computable function.

Even worse, it would be impossible to know whether the oracle hypermachine is hypercomputing or not because we would not be able to claim whether the computed function is a non Turing-computable function or a computable one. On the one hand, as a consequence of the impossibility for identifying the non Turing-computable function that is computed by the hypermachine, oracle hypermachines have to be regarded as black boxes, which internal behavior is not specified. On the other hand, it is impossible - exclusively from intput-output black box behavior - to identify the function that is computed by the Turing machine [1], [15]. Intuitively, this is due to the fact that we only have at our disposal a finite number of results, which could every time correspond to other functions. For exemple, if we think that the computed function multiplies a number by 2 and the first 3000 results agree with f, it could be still possible that the next number will be multiplied by 3. Hence we would not be able to verify the device is computing a non Turing-computable function or a computable one, and we would not be able to claim whether it is a hypermachine or a Turing machine.

However, according to Shagrir and Pitowski however, we are not in a total incapacity to claim that the device is hypercomputing because we can use our physical theories to explain that the computed function is non Turing-computable. In particular, from the theory of radioactive decay and the standard interpretation of quantum mechanics, we can infer that processes based on these theories are genuisely random and thus that oracle hypermachines are able to hypercompute. Nevertheless, Stannett has proved in the case of the oracle hypermachine based on the theory of radioactive decay that either the theory is correct, and we can build the true infinite random-number generator using a radiocative sample, or the assumption was flawed, but in this case, we can show that hypercomputation cannot be refuted by any experiment conducted according to the rules of the theory [28]. To sum up, this means that any experimental refutation of the hypercomputational power of the oracle hypermachine is a refutation of the theory of radioactive decay itself. In other words, even if the theory tells us why the device should hypercompute, we cannot have any experimental proof of this claim. It seems therefore that physical theories do not bring any robust answer concerning the verification problem applied to oracle hypermachines. Yet since Stannett's argument only works in the case of oracle hypermachines based on the theory of radioactive decay, further research will need to be carried out about oracle hypermachines based on quantum randomness.

Finally, even if we cannot prove hypercomputation from experiments according to the rules of the theory, it could be still possible to solve the verification problem in an empirical way in increasing the plausibility of the claim that the device computes a non Turing-computable function [6]. In that case, tests on random strings should be achieved in order to show with a high probability that they are not pseudo-random. If such tests were achieved and we concluded that a string is true-random, then we could claim that this string represents the results of a non Turing-computable function. However, the disadvantage is that all current pseudo-random number generators provide strings, which are in practice impossible to distinguish from true-random number strings. Nevertheless, it is true that we cannot dismiss the possibility to have some day reasonable grounds to believe that a string is true-random.

4 Conclusion

In this paper, we have tried to explain that the verification problem, namely the problem of verifying the computed results provided by a machine, could be a threat to a possible proof of the physical hypercomputation thesis. Indeed, even though some researchers have proposed solutions in order to solve this problem, we claimed these solutions fit for programmable hypermachines but not for oracle hypermachines. Our arguments were that in the case of oracle hypermachines (1) we cannot increase the plausibility that the hypermachine computes a given function because we cannot know what is the computed function; (2) it is impossible to know whether the hypermachine is hypercomputing because we would not be able to claim whether the computed function is a non Turing-computable function; (3) From physical theories, we cannot have any experimental proof of the claim that an oracle hypermachine is hypercomputing; (4) no current methods are able to prove the true randomness of the string provided by the hypermachine, which is the property on which is based its hypercomputational power. Thus, as long as the verification problem will not be solved, the physical construction of an oracle hypermachine could not be considered as a proof of the physical hypercomputation thesis.

Acknowledgements. I would like to thank the editors and referees for very helpful comments during the preparation of this paper. In particular, many thanks to Cristian Calude and Jack Copeland for their very interesting advices. Finally, I am very grateful to the university paris 1 Doctoral school for supporting this work.

References

1. Adleman, L.M., Blum, M.: Inductive Inference and Unsolvability. The Journal of Symbolic Logic 56, 891–900 (1991)
2. Calude, C.S.: Algorithmic randomness, quantum physics, and incompleteness. In: Margenstern, M. (ed.) MCU 2004. LNCS, vol. 3354, pp. 1–17. Springer, Heidelberg (2005)

3. Calude, C.S., Svozil, K.: Quantum Randomness and Value Indefiniteness. Advanced Science Letters 1, 165–168 (2008)
4. Chaitin, G.: On the Length of Programs for Computing Finite Binary Sequences. Journal of the Association for Computing Machinery 13, 547–569 (1966)
5. Church, A.: On the Concept of a Random Sequence. Bulletin of the American Mathematical Society 46, 130–135 (1940)
6. Cleland, C.E.: The Concept of Computability. Theoretical Computer Science 317, 209–225 (2004)
7. Copeland, J.: Narrow Versus Wide Mechanism: Including a Re-Examination of Turing's Views of the Mind-Machine Issue. The Journal of Philosophy 1, 5–32 (2000)
8. Copeland, J.: Hypercomputation. Minds and Machines 12, 461–502 (2002)
9. Copeland, J.: Accelerating Turing Machine. Minds and Machines 12, 281–301 (2002)
10. Copeland, J., Proudfoot, D.: Alan Turing's Forgotten Ideas in Computer Science. Scientific American 208, 76–81 (1999)
11. Davies, B.: Building Infinite Machines. British Journal for the Philosophy of Science 52, 671–682 (2001)
12. Davis, M.: The Myth of Hypercomputation. In: Teuscher, C. (ed.) Alan Turing: The Life and Legacy of a Great Thinker, pp. 195–212. Springer, Berlin (2004)
13. De Leeuw, K., Moore, E.F., Shannon, C.E., Shapiro, N.: Computability by Probabilistic Machines. Automata Studies. Princeton University Press (1956)
14. Earman, J.: Bangs, Crunchs, Whimpers and Shrieks - Singularities and Acausalities in Relativistic Spacetimes. Oxford University Press, Oxford (1995)
15. Gold, M.: Limiting Recursion. The Journal of Symbolic Logic 30, 28–48 (1965)
16. Hamkins, J.D.: Infinite Time Turing Machines. Minds and Machines 12, 567–604 (2002)
17. Hogarth, M.: Does General Relativity Allow an Observer to View an Eternity in a Finite Time? Foundations of Physics Letters 5, 173–181 (1992)
18. Jennewein, T., Achleitner, U., Weihs, G., Weinfurter, H., Zeilinger, A.: A Fast and Compact Quantum Random Number Generator. Review of Scientific Instruments 71, 1675–1680 (2000)
19. Kieu, T.: Quantum Hypercomputation. Minds and Machines 12, 541–561 (2002)
20. Leitsch, A., Schachner, G., Svozil, K.: How to Acknowledge Hypercomputation? Complex Systems 18, 131–143 (2008)
21. Loff, B., Costa, J.F.: Five Views of Hypercomputation. International Journal of Unconventional Computing 5, 193–207 (2009)
22. Martin-Löf, P.: The Definition of a Random Sequence. Information and Control 9, 602–619 (1966)
23. Matiiassevitch, Y.: Le Dixième Problème de Hilbert: son Indécidabilité. Masson, Paris (1995)
24. Piccinini, G.: The Physical Church-Turing Thesis: Modest or Bold? The British Journal For The Philosophy of Science 62, 773–769 (2011)
25. Pitowski, I.: The Pysical Church Thesis and Physical Computational Complexity. Iyyun 39, 81–99 (1990)
26. Shagrir, O., Pitowski, I.: Physical Hypercomputation and the Church-Turing Thesis. Minds and Machines 13, 87–101 (2003)

27. Siegelmann, H.T.: Neural and Super-Turing Computing. Minds and Machines 13, 103–114 (2003)
28. Stannett, M.: Computation and Hypercomputation. Minds and Machines 13, 115–153 (2003)
29. Stannett, M.: Hypercomputational Models. In: Teuscher, C. (ed.) Alan Turing: The Life and Legacy of a Great Thinker, pp. 135–157. Springer, Berlin (2004)
30. Turing, A.M.: Systems of Logic Based on the Ordinals. In: Davis, M. (ed.) The Undecidable, pp. 154–222. Dover, New York (1939, 1965)

Does the Principle of Computational Equivalence Overcome the Objections against Computationalism?

Alberto Hernández-Espinosa and Francisco Hernández-Quiroz

Universidad Nacional Autónoma de México,
Departamento de Matemáticas, Facultad de Ciencias,
Ciudad Universitaria, D.F. 04510
Mexico
fhq@ciencias.unam.mx, albertohernandezespinosa@gmail.com

Abstract. Computationalism has been variously defined as the idea that the human mind can be modelled by means of mechanisms broadly equivalent to Turing Machines. Computationalism's claims have been hotly debated and arguments against and for have drawn extensively from mathematics, cognitive sciences and philosophy, although the debate is hardly settled. On the other hand, in his 2002 book New Kind of Science, Stephen Wolfram advanced what he called the Principle of Computational Equivalence (PCE), whose main contention is that fairly simple systems can easily reach very complex behaviour and become as powerful as any possible system based on rules (that is, they are computationally equivalent). He also claimed that any natural (and even human) phenomenon can be explained as the interaction of very simple rules. Of course, given the universality of Turing Machine-like mechanisms, PCE could be considered simply a particular brand of computationalism, subject to the same objections as previous attempts. In this paper we analyse in depth if this view of PCE is justified or not and hence if PCE can overcome some criticisms and be a different and better model of the human mind.

Keywords: Computationalism, Computational Theory of Mind, Representationalism, Principle of Computational Equivalence.

1 Introduction

Computational Theory of the Mind or computationalism is usually attributed to Alan Turing (for instance [15]). In fact Turing compared the human brain to a digital computing machine [34], but also to an analogue type machine [36], but we should point out that Turing never developed a formal theory of thought, despite his foundational work on computability. In contrast, McCulloch and Pitts [21] did talked about mental processes as computations, as Piccinini reminds us [23].

G. Dodig-Crnkovic and R. Giovagnoli (Eds.): *Computing Nature*, SAPERE 7, pp. 225–233.
DOI: 10.1007/978-3-642-37225-4_14

As we said before, computationalism is not a single thesis, but it has been formulated differently by many people. According to Piccinini [23], computationalism claims that cognitive activity is achieved by means of computations carried out by specific components of the mind whose functioning is akin to that of a Turing Machine (TM) or an equivalent mechanism. The fact that cognition happens in the brain (and the brain is based on neural networks and not on TM) can be incorporated into computationalism by considering that neural computations are Turing-computable at least as they are actually realized in the human brain. This wider thesis would make some types of connectionism mere variations of computationalism.

Piccinini [24] also distinguishes two variations: 1) traditional or classic computationalism, which claims that thought can be reduced to computations made over linguistic structures and 2) connectionist computationalism, which claims that thought can be reduced to "computations" carried out by Neural Network Systems.

There are some other theses that frequently have been grouped together under the label of computationalism, for instance, the so called Strong Artificial Intelligence (SAI), which, according to Searle [30], claims that artificial intelligence can eventually reach the ability of becoming self-aware and exhibit human-like thought processes.

We will not dwell on this specific variety of computationalism (if it really can be found beyond Searle's analysis) as we are interested only in the explanatory power of computationalism for understanding the human mind and not in the question of whether computers can really think, and we consider Piccinini's classification perfectly adequate for this purpose.

The debate between supporters of varieties of computationalism and their detractors has raged for decades and both sides have drawn arguments from mathematics, cognitive science and philosophy. The point is hardly settled and we do not intend to review it here even superficially. New arguments and theories keep appearing which can (or cannot) be considered variations of computationalism and claim to deal better with objections against computational explanations of the mind. Our purpose in this paper is to analyse one of this theories, namely, Stephen Wolfram's Principle of Computational Equivalence (PCE), introduced as one of the key elements of his extremely ambitious New Kind of Science program. In his book of the same name, Wolfram contends that PCE can explain the complexity in any natural or artificial phenomenon, including of course the complexity of human mind.

The outline of the paper is as follows: in the second section we review some arguments against computationalism. In the third section, we summarize what we consider some of the essential claims of Wolfram's PCE as a tool for explaining the complexity of the human mind. In the fourth we ponder the ability of PCE for dealing against the counterarguments of computationalism presented in the second section, while at the same time evaluating if PCE is or not just plain computationalism under a new disguise (although we do not offer a definite answer yet). In the final section, we point out to the challenges that PCE should deal with if it has any hope of offering a better alternative to past theories.

2 Four Types of Arguments against Computationalism

Cordeshi [4], Dreyfus [9,10] and Horst [17] have brought forward diverse arguments against computationalism. We have classified them in four types for convenience.

Computationalism contends that is the only scientific explanation in offer. Their supporters argue that computational explanations of cognitive abilities like language and learning are the only viable approach to the mind. Examples of this view can be found in Fodor [13], Pinker [25] and Winograd [37]. Even if they take for granted that the mind "resides" in the brain and the brain is a gigantic neural network, they also claim that electrical signals in neural networks codify symbols and representations which are manipulated according to logical rules [30]. One consequence of this view is that the mind deals basically with representational systems [17]. A first and clear line of attack against computationalism (at least its representational version) is to challenge the contention that it is the only serious candidate for modelling or explaining the mind. As Horst has pointed out [19], in the search for alternatives philosophers and cognitive scientists are reconsidering if models like neural networks can and should be based on rules and representations or if they work in a radically different way.

On the other hand, Dreyfus [9,10] and even Winograd and Flores [37] have argued that a significant part of what we call thought and behaviour cannot be reduced to explicit rules and therefore cannot be formalized (and translated into a computer program). In other words, a sizeable portion of mental phenomena are beyond the reach of techniques dearest to computationalists.

A third line of criticism rejects the use of symbols as the foundation of the semantics of thoughts. Symbolic semantics imply intentionality in thought either through causality [16,27,28] or concepts [18]. But trying to explain intentionality by symbols is a vicious circle. Searle [29] and Horst [18] go further and state that computer "representations" are not even symbolic on their own right as its symbolic nature rests on the intentions and conventions held by their human users.

Supporters of externalist theories of meaning have raised a fourth set of criticisms. Many computationalists were fond of what can be called "methodological solipsism" [12] or individualism: the view that mental states' characterization is insensitive to and independent from any external features of the cognitive subject, as the underlying computational processes only have access to mental representations. But at the same time, computationalism would have this characterization reflecting semantic properties. This is clearly difficult to reconcile with an externalist stand on meaning [27], which would require that the meaning of terms be at least partially determined by factors external to the cognitive subject, for instance, its physical [24] and linguistic [1,2] environment. Of course, the argument can be turned around to reject externalism as Fodor did [11].

3 NKS and the Principle of Computational Equivalence (PCE)

Stephen Wolfram wrote his book A New Kind of Science (NKS) [38] after twenty years of experimentation with Cellular Automata (CA) as tools for solving problems in a very wide range of domains. One of the main guidance of his proposal is the Principle of Computational Equivalence (PCE), which can be summarized by the two following theses:

1. All processes, whether they are produced by human effort or occur spontaneously in nature, can be viewed as computations.
2. In computational terms there is a fundamental equivalence between many different kinds of processes. In particular, almost all processes that are not obviously simple can be viewed as computations of equivalent sophistication. [38]

In very general terms, Wolfram contends that PCE means that there is a maximal (“universal”) level of complexity in computations and this level is easily attainable by most non-trivial systems (even artificial ones). Natural systems can in principle have the same computational power as computers and vice versa. Wolfram claims that, provided a proper translation for inputs and outputs of different systems, all of them are computationally equivalent.[1]

Wolfram states that his NKS has four basic advantages over classical science:

1. An alternative view of randomness: over time, simple rules can produce very complex behaviour which becomes almost impossible to predict. Randomness is then just unpredictability arising from lack of information about deterministic phenomena. But this type of “randomness” can be approximated by means of programs based on very simple rules.
2. The assumption of a maximal degree of simplicity in phenomena, which implies also their irreducibility to even simpler rules (wolfram calls it “computational irreducibility”). When confronted with irreducible cases, direct observation and experimentation are the only way forward for Science. The seemingly paradoxical nature of free will may contain this sort of irreducibility.
3. Scientific insight should be guided by the search of these very simple rules in all natural and human phenomena. Of course, this idea goes counter the “prejudice” that computing simulations of natural phenomena should be based in very complex software. The key, according to Wolfram, is the opposite: look for simple rules.
4. Given that all systems are based on simple rules, individual sciences can proceed to analyse their disparate subjects by means of a uniform methodology which can help to extract more general and abstract explanations.

[1] Sutner claims that Wolfram's view can be taken to mean that there are really only two levels of complexity in natural phenomena: a lower one of systems whose behaviour is decidable and the higher one of systems whose behaviour reaches universal complexity in computational terms [32,33]. This thesis has been called a 0/1 law of computational degrees.

Stephen Wolfram states explicitly that the complexity of the human mind is also covered by PCE. For instance, he claims that perception can be reduced to a process of pattern recognition and information processing [38]. At first sight, PCE seems to be just another version of classical computationalism. But it may not be so simple. For instance: does PCE imply representationalism? Other similar questions can be easily asked and their answers are not straightforward, which makes us think worthwhile to consider in depth if Wolfram's proposal can really offer a valuable alternative to classical computationalism.

4 NKS vs Objections against Computationalism

Following Dodig-Crnkovic's analysis of what she calls info-computationalism (the strong thesis that the universe can be better understood as a series of computational processes operating on informational structures) [5,6,7], we may be inclined to regard PCE as a variety of info-computationalism. Nonetheless, there are at least two reasons why Wolfram's proposal may be considered a different and probably better brew of computationalism which may be able to avoid some criticisms directed against other traditional computationalist views: 1) if he is right (and this a big "if") that there is an upper limit in complexity for all systems and this limit can be reached by some simple rules, then of course computer programs can simulate any degree of complexity; 2) again, if his main thesis is right, the complexity of the mind also falls in the scope of what can be explained by computations based on simple rules.

While it is far from clear that all systems in nature have a complexity limit within the reach of the computable, computable universality is reachable by means of the simple rules advocated by Wolfram [31,3]. The general question of a universal limit is still open and seems bound to remain so for the foreseeable future. On the other hand, even if Wolfram were right about the existence of an upper limit in complexity, he offers no practical clues for the solution of the many problems any theory of mind (let alone a computational one) should face. His optimism becomes evident when he regards a possible explanation of free will as computationally complex decision procedures whose inner details are hidden from consciousness [38].

NKS and the PCE are then just a (sketch of a) proposal for a research program and before embracing it any prospective theoretician of mind should at least make a quick assessment of its potential:

1. A first obvious question is if we are not dealing with a mere variety of computationalism.
2. A second and more interesting one is —if PCE is not simple computationalism (or even despite being computationalism)—, how it can prove its worth as a serious challenge to representationalism's claim of being the only viable explanation in offer.[13].
3. Next it is to be seen if PCE can answer the objection that human thought and behaviour cannot be reduced to explicit rules and therefore cannot be formalized or reduced to computer programs [9,10,37].

4. PCE should also offer a theory of the meaning of thought without the troubles faced by computationalism's symbolic semantics [28,16,27].
5. Finally, PCE should present an alternative explanation of how mental states can be characterized independently of features external to the cognitive subject [1,2].

Many other issues could be raised [4,24], but we consider these some of the most relevant because they touch the core of the theory and we will dwell on them in the next section.

5 Problems to Solve

What are the chances of PEC dealing rightly with the previous questions? It is not our intention to give a definitive answer, but just to offer a very initial assessment and to outline how a NKS practitioner should carry on.

To begin with, the charge of being just computationalism under a different guise. Mathematically speaking, the simple rules on which NKS is based are computationally equivalent to Turing Machines and other Turing-complete models. Claiming that any system (natural or artificial) is of equivalent complexity is highly reminiscent of (a strong form of) Church-Turing's thesis, on its turn one of the pillars of computationalism. Wolfram himself seems to support this view: "But it was not until the 1980s —perhaps particularly following some of my work— that it began to be more widely realized that Church's Thesis should best be considered a statement about nature and about the kinds of computation that can be done in our universe. The validity of Church's Thesis has long been taken more or less for granted by computer scientist, but among physicists there are still nagging doubts, mostly revolving around the perfect continua assumed in space and quantum mechanism in the traditional formalism of theoretical physics" [38]. Wolfram calls Turing's and other scientists' attempts "close approaches", acknowledging their similarity, but he also claims to have a distinctive proposal which is also based on "experimentation" on computers. Of course, these short and sometimes puzzling comments do not settle the point, as (a sort of) mathematical equivalence between Church's thesis and PCE does not imply that PCE has to assume all the baggage of classical computationalism (which in turn is not a consequence of Church's thesis).

Regarding the second question, PCE should be able to attain at least the same degree of success as connectionism, an important rival of classical computationalism. According to some researchers [16] connectionism has been able to explain some intellectual abilities without resorting to syntactical representations and manipulations (let us put aside the issue that Artificial Neural Networks as they exist in this moment are mathematically equivalent to Turing Machines), performing better than actual or potential systems based on techniques dear to computationalists. Can PCE equal these supposed achievements? Again, for the time being Wolfram's NKS can only provide more optimism: "So on the basis of traditional intuition; one might then assume that the way to solve this problem must be to use systems with more complicated underlying rules, perhaps more closely based on details of human psychology or neurophysiology. But from discoveries in this book we know that this is not the case, and that in fact very simple rules are quite sufficient to produce highly complex behaviour" [38].

Searle [27,28] and Horst [18] have provided a powerful argument against the idea that thought can be reduced to the application of simple rules in the style of a computer program, as meaning cannot be derived from rules for manipulating symbols (so covering the core of questions 4 and 5): "The problem of semantics is: How these sentences in the head get their meaning? But that question can be discussed independently of the question: How does the brain work in processing these sentences?" [29]. About this last issue Wolfram says: "One might have imagined that human thinking must involve fundamentally special processes, utterly different from all other processes that we have discussed [here Wolfram talks about thinking and perception as processes]. But just as it has become clear over the past few centuries that the basic physical constituents of human beings are not particularly special, so also —especially after the discoveries in this book (NKS)— I am quite certain that in the end there will turn out to be nothing particularly special about the basic processes that are involved in human thinking. And indeed, my strong suspicion is that despite the apparent sophistication of human thinking most of the most important processes that underlie it are very simple" [38]. To be fair (and therefore not so pessimistic), Wolfram's phrasing of the problem does not imply that the solution should be attached to rules for manipulating symbols.

Finally, there is the issue of defining mental states (which are internal representations according to computationalism) and their complex relation with features external to the cognitive subject [23]. Can Wolfram‘s idea of intelligence being based at least partially on pattern recognition point to a different definition about what a mental state is and how it relates to the external world (the ultimate source from which the pattern is recognized)? We consider that, right now, this idea is too vague to give rise to any serious attempt to formulate the problem of mental states, let alone to lead to its solution.

To conclude: PCE hopes for being a better alternative than classical computationalism are dependent on many "if", namely: if Wolfram is right that all natural and artificial phenomena are under the scope of the kind of simple computational rules he advocates, if these rules can lead to practical ways of explaining what previous models have been unable to explain, if complex behaviour such as meaning and mental states (and their relation with the external world) can be accounted for by the same rules, then NKS can offer a way out of computationalism troubles. On a more positive note, we want to stress an implicit conclusion of our previous analysis: it is not obvious that PCE should fail where classical computationalism has already failed.

But optimism cannot be the only foundation for a scientific account of the human mind. More philosophical and empirical research is needed to see if optimism can be turned into results or, at least, concrete lines of research.

References

1. Burge, T.: Individualism and the mental. In: French, P., Euhling, T., Wettstein, H. (eds.) Studies in Epistemology, Midwest Studies in Philosophy, vol. 4. University of Minnesota Press, Minneapolis (1979)
2. Burge, T.: Individualism and psychology. Philosophical Review 95(1), 3–45 (1986)
3. Cook, M.: Universality in Elementary Cellular Automata. Complex Systems 15, 1–40 (2004)
4. Cordeschi, R.: Computacionalism under attack. Cartographies of the Mind, Chapter 3, 37–49 (2007)
5. Dodig-Crnkovic, G., Mueller, V.: A dialogue concerning two world systems: info-computational vs. mechanistic. In: Dodig Crnkovic, G., Burgin., M. (eds.) Information and Computation, pp. 149–184. World Scientific Publishing Co., Inc., Singapore (2009)
6. Dodig-Crnkovic, G.: Biological Information and Natural Computation. In: Vallverdú, J. (ed.) Thinking Machines and the Philosophy of Computer Science: Concepts and Principles, Information Science Reference. IGI Global, Hershey (2010)
7. Dodig-Crnkovik, G.: Dynamics of Information as Natural Computation. Information 2(3), 460–477 (2011), doi:10.3390/info2030460
8. Dodig-Crnkovic, G.: Physical Computation as Dynamics of Form that Glues Everything Together. Information 3(2), 204–218 (2012)
9. Dreyfus, H.: What Computers Can't Do. Harper and Row, Nueva York (1972)
10. Dreyfus, H.: What Computers Still Can't Do. MIT Press, Cambridge (1992)
11. Fodor, J.: Metodological silopsism considered as a research strategy in cognitive science. Behavioral and Brain Sciences 3, 63–73
12. Fodor, J.A.: Representations: Philosophical Essays on the Foundations of Cognitive Science. MIT Press, Cambridge (1981)
13. Fodor, J.A.: Theory of Content and Other Essays. MIT Press, Cambridge (1990)
14. Fodor, J.A.: The Elm and the Expert. MIT Press, Cambridge (1993)
15. Fodor, J.A.: Concepts. Clarendon Press, Oxford (1998)
16. Garrido, A.: Connectionism vs. Computationalism Theory of Mind. BRAIN (Broad Research in Artificial Intelligence and Neuroscience) 1(1) (2010)
17. Horst, S.: Mind Computationalism Theory. University of California Press, Berkeley (1986)
18. Horst, S.: Symbols, Computation and Intentionality: A Critique of the Computational Theory of Mind. University of California Press, Berkeley (1996)
19. The Stanford Encyclopedia of Philosophy, `http://plato.stanford.edu/`
20. Kelemen, J., Alica, K.: The New Computationalism: a Lesson from Embodied Agents. Silesram University, Institute of Computer Science Press, Opava, Slovakia (2008)
21. McCulloch, W.S., Pitts, W.H.: A Logical Calculus of the Ideas Immanent in Nervous Activity. Bulletin of Mathematical Biophysics 7, 115–133 (1943)
22. Piccinini, G.: The first Computational theory of mind and brain: a close look at McCulloch and Pitts's Logical calculus of ideas immanent in nervous activity. Synthese 141, 175–215 (2004)
23. Piccinini, G.: Computationalism, the Church-Turing thesis, and the Church-Turing falacy. Synthese 154(1), 97–120 (2007)
24. Piccinini, G.: Computatonalism in the Philosophy of Mind. Philosophy Compass, Journal Compilation 4 (2009)

25. Pinker, S.: The Language Instinct, the New Science of Language and Mind. Penguin, London (1995)
26. Pinker, S.: How the Mind Works. Penguin, London (1997)
27. Putnam, H.: The Meaning of "Meaning". In: Gunderson, E.K. (ed.) Language, Mind and Knowledge. Minnesota Studies in the Philosophy of Science, vol. 7, University of Minnesota Press, Minneapolis (1975)
28. Sayre, K.: Intentionality and information processing: An alternative model for cognitive science. Behavioral and Brain Sciences 9(1), 121–138 (1986)
29. Searle, J.: Minds, brains and programs. Behavioral and Brain Sciences 3, 417–424 (1980)
30. Searle, J.: Is the Brain a digital computer? Proceedings and Addresses of American Philosophical Association 64, 21–37 (1990)
31. Smith, A.: Universality of Wolfram's 2,3 Turing Machine (2007)
32. Sutner, K.: Cellular automata and intermediate degrees. Theoretical Computer Science 296, 365–375 (2003)
33. Sutner, K.: Universality and Cellular Automata. MCU, 50–59 (2004)
34. Turing, A.M.: Lecture to the London Mathematical Society on 20 February 1947. In: Ince, D. (ed.) Mechanical Intelligence, pp. 87–105. North-Holland, Amsterdam (1947)
35. Turing, A.M.: Intelligent Machinery. In: Ince, D. (ed.) Mechanical Intelligence, pp. 87–106. North-Holland, Amsterdam (1948)
36. Turing, A.M.: Computing machinery and intelligence. Mind 59, 433–460 (1950)
37. Winograd, T., Flores, F.: Understanding Computers and Cognition. Ablex, Norwood (1986)
38. Wolfram, S.: The New Kind of Science (2002) ISBN 1-57955-008-8

Some Constraints on the Physical Realizability of a Mathematical Construction

Francisco Hernández-Quiroz[1] and Pablo Padilla[2]

[1] School of Science, UNAM
fhq@ciencias.unam.mx
[2] Institute for Applied Mathematics, UNAM
pablo@mym.iimas.unam.mx

Abstract. Mathematical constructions of abstract entities are normally done disregarding their actual physical realizability. The definition and limits of the physical realizability of these constructions are controversial issues at the moment and the subject of intense debate.

In this paper, we consider a simple and particular case, namely, the physical realizability of the enumeration of rational numbers by Cantor's diagonalization by means of an Ising system.

We contend that uncertainty in determining a particular state in an Ising system renders impossible to have a reliable implementation of Cantor's diagonal method and therefore a stronger physical system is required. We also point out what are the particular limitations of this system from the perspective of physical realizability.

Keywords: Diagonalization, Ising systems, Physical Implementation.

1 Introduction

"There is no quantum world. There is only an abstract quantum description. It is wrong to think physics' task is to discover how Nature is. Physics deals with what is possible to say about Nature."

This quote is attributed to Niels Bohr, when he was asked whether the quantum formalism reflected the underlying physical reality. Bohr's, other philosophers' and scientists' opinions aside, a good deal of paper has been used to analyse the possibility of describing and understanding reality by means of formal mathematical tools. Barrow, Chaitin, Hawking and Penrose (among others) have advanced some ideas with varying degrees of formality.

Here we address a reciprocal question: given a mathematical construction and a particular physical system, is the latter adequate to "implement" the former? By implementation we mean an actual physical device that (a) has structural properties that correspond to components of the mathematical entity (some have talked about an isomorphism between physical and mathematical structures [3], but a weaker notion may also do); (b) a physical procedure that can produce experimental results which reflect accurately corresponding properties of the mathematical construction.

G. Dodig-Crnkovic and R. Giovagnoli (Eds.): *Computing Nature*, SAPERE 7, pp. 235–240.
DOI: 10.1007/978-3-642-37225-4_15

These are very intricate and hard questions to be answered definitely in a general case. Our aim is more modest, namely to explore a specific instance of this problem: we take the classical Cantor's diagonalization for the enumeration of the rational numbers [2] and how it can be implemented by an Ising system. We provide a specific implementation and show its limitations deriving from properties of the physical system itself.

This leads us to think that some clearly defined mathematical questions cannot always be posed and answered within the context of a particular physical system. Of course, the more general question of the existence of a physical system realizing a particular mathematical construction is beyond the limits of this work but we hope our example helps to stimulate discussions on this line of thought. The standard interpretation of quantum mechanics regarding physically meaningful questions is that it should be possible to pose them in such a way that they can be answered experimentally.

The reciprocal question is also interesting: to what extent mathematical constructions should be considered valid? One possible approach, would imply that only those mathematical constructions that can actually be implemented by means of a physical system can in fact be used, at least in terms of computation.

In the next section we present—as a reminder—Cantor's diagonalization method for enumerating the rational numbers. The third section deals with Ising systems and its properties.The fourth section presents our implementation of Cantor's method and how to find a specific rational number. In the final section, which is the central part of this paper, we show how our system is unable to perform the task for which it was designed due to intrinsic limitations of Ising systems and other physical principles, and we also discuss some implications.

2 Cantor's Diagonalization

In 1878 Cantor defined rigorously when two sets have the same cardinality. Let A and B be two sets. They have the same number of elements if and only if there exists a bijection between them, i.e., a function $f : A \to B$ which is both injective and surjective.

He also proved that the set of natural numbers and the set of rational numbers are equinumerous, even though the former is a proper subset of the latter. His argument introduced an ingenious device to construct a one-to-one correspondence between the two sets. The idea is that rational numbers are not arranged according to the traditional $<$ relation, but rather, by taking advantage of the fact that a rational number (in accordance with the etimology of the name) can be regarded as the ratio of two integers. For example, the number 0.5 is also represented by the fraction 1/2.

The fractional representation of a number, let us say m/n, can be transformed into the convention that the pair *(m,n)* represents this very number. Now consider the list

$$(1,1),(1,2),(2,1),(1,3),(2,2),(3,1),(1,4),\ldots$$

where pairs are arranged so that the sum of the two components is increasing; pairs whose sum produces the same value are ordered by the traditional $<$ order applied to the first coordinate of the pairs. By omitting pairs representing the same number (which can always be calculated in a finite number of steps as the list is being produced), this is a bijection between natural and rational numbers, and thus both sets have the same cardinality.

If we set aside the traditional objections posed by mathematical constructivists to the idea of actual infinite sets, Cantor's argument seems very straightforward and has been regarded as such ever since. However we could take a mathematical constructive perspective and reject Cantor's device (and his whole set theory, for that matter).

But we can also take a different constructive perspective, which we may name *physical constructivism*: What requirements should a particular physical system meet in order to serve as a basis for implementing Cantor's device? At first sight there must be physical systems on which this may not be possible (although the symmetrical question does not seem easy to answer). Specifically, we will analyse the feasibility of Ising models for this task in the next section.

3 Ising Models

In the last decades, some models in physics have played a central role in understanding specific connections between mathematical aspects of the theory and experiments. One of such is precisely the Ising model. We use it here for different purposes. We suggest that it can be taken as a real system in which Cantor's diagonal procedure could be implemented and therefore as a starting point from which conclusions can be drawn regarding the limitations that mathematical constructions could have in the physical world. This is due to the fact that, in principle, the physical configurations of the system can be put in correspondence with rational numbers. Moreover, for the Ising model a direct relationship between the physical entropy and the informational entropy can be established, allowing a quantitative comparison .

We briefly recall what the Ising model is about and later on we make a few remarks on the entropy of a discrete physical system. What follows is basically adapted from [5].

We consider a magnetic material in which the electrons determining the magnetic behaviour are localized near the atoms of a lattice and can have only two magnetization states (spin up or down). The spin for a given site in this lattice will be identified with the of the 0's or 1's used in the mathematical construction of the previous section to write down the binary expansion of the rational numbers. Notice that we need only a finite number or 0's or 1's since these expansions will be either finite or periodic. For instance, we might put in a row all numbers (m, n) of a fixed height one after the other with a conventional sequence to denote beginning and end of a number. As mentioned before, the magnetization S_i can take only two values ± 1 that we identify with 0 and 1 respectively.

There is a Hamiltonian associated in the presence of an external magnetic force depending on the site, h_i which is given by:

$$H = -J \sum_{i,k} S_i S_j - \sum_i h_i S_i,$$

where the sum over i and k runs over all possible nearest-neighbour pairs of the lattice and J is the so called exchange constant.

The fact that is important to stress is that a possible enumeration of the rationals correspond to a particular physical configuration. Notice that we are disregarding the obvious limitation of size. That is, in Cantor's procedure we need an infinite number of rows and columns, that is an ideal lattice, whereas a physical material will necessarily have finite size. Nevertheless, we will see that even then, there are physical constraints that are imposed by the quantum nature of the system to the entropy, which can be interpreted as informational restrictions on the physical realizability of the mathematical construction.

For a continuous system whose configuration is denoted by C, where the configuration space is assumed to be endowed with a measure μ (for simplicity one may think of $\mathbb{R}^d$), the entropy associated with a specific probability distribution P is given by

$$S[P] = -\int d\mu(C)P(C)lnP(C),$$

that is, the expected value of $-lnP(C)$ with respect to μ.

By dividing the space into cells of size ε^d the entropy of the continuous system can be well approximated by the entropy of the discrete system resulting from the partition:

$$S_{disc} = S_{cont} - dln(\varepsilon).$$

As a matter of fact, the ε can be taken to be the Planck constant for a quantum system. This observation will be important later on.

4 Implementing Cantor's Method

As we mentioned before, we can in principle use the Ising system to physically array and enumerate the rational numbers and locate any of them in this array. In fact the question: "How to find a rational number in the list?" is well defined and would only need a finite number of steps.

In the section devoted to the Ising model, we recalled equation 3 for the entropy of a quantum system. Notice that the second term is positive and independent of the details of the system, only due to the quantum nature of the same. This has an important implication in terms of the possibility of actually determining the state in which the Ising model is. If we relate the information content with the entropy of the system we see that, in order for the state of the system to be completely determined, we would need zero entropy [4]. This is physically impossible. Moreover, a lower bound for the entropy is related not only to the discrete (quantum) nature of the system, but it also depends on the temperature and other parameters. The conclusion is that even when the counting and locating procedure is well defined, there is always an intrinsic error. Of course one might argue that this is probably due to the chosen system, but the reasoning is general enough as to suggest that no matter what physical implementation we choose, there will always exist this limitation.

5 Conclusion: Uncertainty Comes in the Way or How Real Is Reality?

We have argued that uncertainty in determining a particular state in an Ising system renders impossible to have a reliable implementation of Cantor's diagonal method. There are also other related mathematical constructions that could be analysed in a similar way. For instance, Cantor's proof of the uncountability of the real numbers relies on similar ideas. As a matter of fact, in the usual argument, a contradiction is obtained by producing a real number that cannot be included in a proposed enumeration. This is done by considering the diagonal sequence and taking its negation. Once this is done, it can be shown that if t is the truth value of the element of this sequence intersecting the diagonal, then it would have to satisfy the relation

$$t = 1 - t,$$

which leads to a contradiction *if one assumes the only possible truth values are 0 or 1* (see for instance chapter 2 on diagonalization in[1]). However, this equation does not pose any problem if t is interpreted in a probabilistic way and assigned a value of 1/2. This opens up a series of even subtler questions such as whether we can actually have a physical model of the real numbers and many others, that from our perspective, are worth addressing.

Many other people have previously addressed this questions either in general terms or for particular mathematical concepts. A pioneering work is [6], which posed the question of realizing an abstract mapping process within the constraints of a physical version of Church's thesis. A very recent case study in the field of control and quantum systems can be found in [7].

References

1. Boolos, G.S., Burgess, J.P., Jeffrey, R.C.: Computability and Logic. Cambridge Univ. Press, New York (2007)
2. Cantor, G.: Contributions to the Founding of the Theory of Transfinite Number. Dover, New York (1915)
3. Chalmers, D.: On Implementing a Computation. Mind and Machines 4, 391–402 (1994)
4. Kintchin, A.: Mathematical Foundations of Information Theory. Dover, London (1963)

5. Parisi, G.: Statistical Field Theory. In: Frontiers in Physics. Addison-Wesley Publishing Company (1988)
6. Rosen, R.: Church's thesis and its relation to the concept of realizability in biology and physics. Bulletin of Mathematical Biology 24, 375–393 (1962)
7. Petersen, I., Shaiju, A.: A Frequency Domain Condition for the Physical Realizability of Linear Quantum Systems. IEEE Transactions on Automatic Control (9) (2012)

From the Closed Classical Algorithmic Universe to an Open World of Algorithmic Constellations

Mark Burgin[1] and Gordana Dodig-Crnkovic[2]

[1] Dept. of Mathematics, UCLA, Los Angeles, USA
mburgin@math.ucla.edu
[2] Mälardalen University, Department of Computer Science and Networks,
School of Innovation, Design and Engineering, Västerås, Sweden
gordana.dodig-crnkovic@mdh.se

Abstract. In this paper we analyze methodological and philosophical implications of algorithmic aspects of unconventional computation. At first, we describe how the classical algorithmic universe developed and analyze why it became closed in the conventional approach to computation. Then we explain how new models of algorithms turned the classical closed algorithmic universe into the open world of algorithmic constellations, allowing higher flexibility and expressive power, supporting constructivism and creativity in mathematical modeling. As Gödel's undecidability theorems demonstrate, the closed algorithmic universe restricts essential forms of mathematical cognition. In contrast, the open algorithmic universe, and even more the open world of algorithmic constellations, remove such restrictions and enable new, richer understanding of computation.

Keywords: Unconventional algorithms, unconventional computing, algorithmic constellations, Computing beyond Turing machine model.

1 Introduction

Te development of various systems is characterized by a tension between forces of conservation (tradition) and change (innovation). Tradition sustains system and its parts, while innovation moves it forward advancing some segments and weakening the others. Efficient functioning of a system depends on the equilibrium between tradition and innovation. When there is no equilibrium, system declines; too much tradition brings stagnation and often collapse under the pressure of inner or/and outer forces, while too much innovation leads to instability and frequently in rupture.

The same is true for the development of different areas and aspects of social systems, such as science and technology. In this article, we are interested in computation, which has become increasingly important for society as the basic aspect of information technology. Tradition in computation is represented by conventional computation and classical algorithms, while unconventional computation stands for the far-reaching innovation.

G. Dodig-Crnkovic and R. Giovagnoli (Eds.): *Computing Nature*, SAPERE 7, pp. 241–253.
DOI: 10.1007/978-3-642-37225-4_16

It is possible to distinguish three areas in which computation can be unconventional:

1. *Novel hardware* (e.g. quantum systems) provides material realization for unconventional computation.
2. *Novel algorithms* (e.g. super-recursive algorithms) provide operational realization for unconventional computation.
3. *Novel organization* (e.g. evolutionary computation or self-optimizing computation) provides structural realization for unconventional computation.

Here we focus on algorithmic aspects of unconventional computation and analyze methodological and philosophical problems related to it, making a distinction between three classes of algorithms: *recursive*, *subrecursive*, and *super-recursive algorithms*.

Each type of *recursive algorithms* form a class in which it is possible to compute exactly the same functions that are computable by Turing machines. Examples of recursive algorithms are partial recursive functions, RAM, von Neumann automata, Kolmogorov algorithms, and Minsky machines.

Each type of *subrecursive algorithms* forms a class that has less computational power than the class of all Turing machines. Examples of subrecursive algorithms are finite automata, primitive recursive functions and recursive functions.

Each type of *super-recursive algorithms* forms a class that has more computational power than the class of all Turing machines. Examples of super-recursive algorithms are inductive and limit Turing machines, limit partial recursive functions and limit recursive functions.

The main problem is that conventional types and models of algorithms make the algorithmic universe, i.e., the world of all existing and possible algorithms, closed because there is a rigid boundary in this universe formed by recursive algorithms, such as Turing machines, and described by the Church-Turing Thesis. This closed system has been overtly dominated by discouraging incompleteness results, such as Gödel incompleteness theorems.

Contrary to this, super-recursive algorithms controlling and directing unconventional computations break this boundary leading to an open algorithmic multiverse – world of unrestricted creativity.

The paper is organized as follows. First, we summarize how the *closed algorithmic universe* was created and what are advantages and disadvantages of living inside such a closed universe. Next, we describe the breakthrough brought about by the creation of super-recursive algorithms. In Section 4, we analyze super-recursive algorithms as cognitive tools. The main effect is the immense growth of cognitive possibilities and computational power that enables corresponding growth of information processing devices.

2 The Closed Universe of Turing Machines and other Recursive Algorithms

Historically, after having an extensive experience of problem solving, mathematicians understood that problem solutions were based on various algorithms. Construction algorithms and deduction algorithms have been the main tools of mathematical research. When they repeatedly encountered problems they were not able to solve, mathematicians, and especially experts in mathematical logic, came to the conclusion that it was necessary to develop a rigorous mathematical concept of algorithm and to prove that some problems are indeed unsolvable. Consequently, a diversity of exact mathematical models of algorithm as a general concept was proposed. The first models were λ-*calculus* developed by Church in 1931 – 1933, *general recursive functions* introduced by Gödel in 1934, ordinary *Turing machines* constructed by Turing in 1936 and in a less explicit form by Post in 1936, and *partial recursive functions* built by Kleene in 1936. Creating λ-calculus, Church was developing a logical theory of functions and suggested a formalization of the notion of computability by means of λ-definability. In 1936, Kleene demonstrated that λ-definability is computationally equivalent to general recursive functions. In 1937, Turing showed that λ-definability is computationally equivalent to Turing machines. Church was so impressed by these results that he suggested what was later called the Church-Turing thesis. Turing formulated a similar conjecture in the Ph.D. thesis that he wrote under Church's supervision.

It is interesting to know that the theory of Frege [1] actually contains λ-calculus. So, there were chances to develop a theory of algorithms and computability in the 19^{th} century. However, at that time, the mathematical community did not feel a need of such a theory and probably, would not accept it if somebody created it.

The Church-Turing thesis explicitly marked out a rigid boundary for the algorithmic universe, making this universe closed by Turing machines. Any algorithm from this universe was inside that boundary.

After the first breakthrough, other mathematical models of algorithms were suggested. They include a variety of Turing machines: *multihead, multitape Turing machines, Turing machines with n-dimensional tapes, nondeterministic, probabilistic*, *alternating* and *reflexive Turing machines*, *Turing machines with oracles*, *Las Vegas Turing machines*, etc.; *neural networks* of various types – *fixed-weights, unsupervised, supervised, feedforward,* and *recurrent neural networks*; *von Neumann automata* and general *cellular automata*; *Kolmogorov algorithms finite automata* of different forms – *automata without memory, autonomous automata, automata without output or accepting automata, deterministic, nondeterministic*, *probabilistic automata*, etc.; *Minsky machines*; *Storage Modification Machines* or simply, *Shönhage machines*; *Random Access Machines* (RAM) and their modifications - *Random Access Machines with the Stored Program* (RASP), *Parallel Random Access Machines* (PRAM); *Petri nets* of various types – *ordinary* and ordinary *with restrictions*, *regular, free, colored,* and *self-modifying Petri nets*, etc.; *vector machines*; *array machines*; *multidimensional structured model of computation and computing systems*; *systolic arrays*; *hardware modification machines*; *Post productions*; *normal Markov algorithms*; *formal grammars* of many forms – *regular, context-free, context-sensitive, phrase-structure*, etc.; and so on. As a result, the theory of algorithms, automata and computation has become one of the foundations of computer science.

In spite of all differences between these models and diversity of algorithms, there is a unity in the system of algorithms. While new models of algorithm appeared, it was proved that no one of them could compute more functions than the simplest Turing machine with a one-dimensional tape. All this give more and more evidence to validity of the Church-Turing Thesis.

Even more, all attempts to find mathematical models of algorithms that were stronger than Turing machines were fruitless. Equivalence with Turing machines has been proved for many models of algorithms. That is why the majority of mathematicians and computer scientists have believed that the Church-Turing Thesis was true. Many logicians assume that the Thesis is an axiom that does not need any proof. Few believe that it is possible to prove this Thesis utilizing some evident axioms. More accurate researchers consider this conjecture as a law of the theory of algorithms, which is similar to the laws of nature that might be supported by more and more evidence or refuted by a counter-example but cannot be proved.

Besides, the Church-Turing Thesis is extensively utilized in the theory of algorithms, as well as in the methodological context of computer science. It has become almost an axiom. Some researchers even consider this Thesis as a unique absolute law of computer science.

Thus, we can see that the initial aim of mathematicians was to build a closed algorithmic universe, in which a universal model of algorithm provided a firm foundation and as it was found later, a rigid boundary for this universe supposed to contain all of mathematics.

It is possible to see the following advantages and disadvantages of the closed algorithmic universe.

Advantages:

1. Turing machines and partial recursive functions are feasible mathematical models.
2. These and other recursive models of algorithms provide an efficient possibility to apply mathematical techniques.
3. The closed algorithmic universe allowed mathematicians to build beautiful theories of Turing machines, partial recursive functions and some other recursive and subrecursive algorithms.
4. The closed algorithmic universe provides sufficiently exact boundaries for knowing what is possible to achieve with algorithms and what is impossible.
5. The closed algorithmic universe provides a common formal language for researchers.
6. For computer science and its applications, the closed algorithmic universe provides a diversity of mathematical models with the same computing power.

Disadvantages:

1. The main disadvantage of this universe is that its main principle - the Church-Turing Thesis - is not true.
2. The closed algorithmic universe restricts applications and in particular, mathematical models of cognition.
3. The closed algorithmic universe does not correctly reflect the existing computing practice.

3 The Open World of Super-Recursive Algorithms and Algorithmic Constellations

Contrary to the general opinion, some researchers expressed their concern for the Church-Turing Thesis. As Nelson writes [2], "*Although Church-Turing Thesis has been central to the theory of effective decidability for fifty years, the question of its epistemological status is still an open one.*" There were also researchers who directly suggested arguments against validity of the Church-Turing Thesis. For instance, Kalmar [3] raised intuitionistic objections, while Lucas and Benacerraf discussed objections to mechanism based on theorems of Gödel that indirectly threaten the Church-Turing Thesis. In 1972, Gödel's observation entitled "A philosophical error in Turing's work" was published where he declared that: *"Turing in his 1937, p. 250 (1965, p. 136), gives an argument which is supposed to show that mental procedures cannot go beyond mechanical procedures. However, this argument is inconclusive. What Turing disregards completely is the fact that mind, in its use, is not static, but constantly developing, i.e., that we understand abstract terms more and more precisely as we go on using them, and that more and more abstract terms enter the sphere of our understanding. There may exist systematic methods of actualizing this development, which could form part of the procedure. Therefore, although at each stage the number and precision of the abstract terms at our disposal may be finite, both (and, therefore, also Turing's number of distinguishable states of mind) may converge toward infinity in the course of the application of the procedure."* [4]

Thus, pointing that Turing disregarded completely the fact that mind, in its use, is not static, but constantly developing, Gödel predicted necessity for super-recursive algorithms that realize inductive and topological computations [5]. Recently, Sloman [6] explained why recursive models of algorithms, such as Turing machines, are irrelevant for artificial intelligence.

Even if we abandon theoretical considerations and ask the practical question whether recursive algorithms provide an adequate model of modern computers, we will find that people do not see correctly how computers are functioning. An analysis demonstrates that while recursive algorithms gave a correct theoretical representation for computers at the beginning of the "computer era", super-recursive algorithms are more adequate for modern computers. Indeed, at the beginning, when computers appeared and were utilized for some time, it was necessary to print out data produced by computer to get a result. After printing, the computer stopped functioning or began to solve another problem. Now people are working with displays and computers produce their results mostly on the screen of a monitor. These results on the screen exist there only if the computer functions. If this computer halts, then the result on its screen disappears. This is opposite to the basic condition on ordinary (recursive) algorithms that implies halting for giving a result.

Such big networks as Internet give another important example of a situation in which conventional algorithms are not adequate. Algorithms embodied in a multiplicity of different programs organize network functions. It is generally assumed that any computer program is a conventional, that is, recursive algorithm. However, a recursive algorithm has to stop to give a result, but if a network shuts down, then something is wrong and it gives no results. Consequently, recursive algorithms turn out to be too weak for the network representation, modeling and study.

Even more, no computer works without an operating system. Any operating system is a program and any computer program is an algorithm according to the general understanding. While a recursive algorithm has to halt to give a result, we cannot say that a result of functioning of operating system is obtained when computer stops functioning. To the contrary, when the operating system does not work, it does not give an expected result.

Looking at the history of unconventional computations and super-recursive algorithms we see that Turing was the first who went beyond the "Turing" computation that is bounded by the Church-Turing Thesis. In his 1938 doctoral dissertation, Turing introduced the concept of a *Turing machine with an oracle*. This, work was subsequently published in 1939. Another approach that went beyond the Turing-Church Thesis was developed by Shannon [7], who introduced the *differential analyzer*, a device that was able to perform continuous operations with real numbers such as operation of differentiation. However, mathematical community did not accept operations with real numbers as tractable because irrational numbers do not have finite numerical representations.

In 1957, Grzegorczyk introduced a number of equivalent definitions of computable real functions. Three of Grzegorczyk's constructions have been extended and elaborated independently to super-recursive methodologies: the *domain approach* [8,9], *type* 2 *theory of effectivity* or *type* 2 *recursion theory* [10,11], and the *polynomial approximation approach* [12]. In 1963, Scarpellini introduced the class $\mathbf{M}_1$ of functions that are built with the help of five operations. The first three are elementary: substitutions, sums and products of functions. The two remaining operations are performed with real numbers: integration over finite intervals and taking solutions of Fredholm integral equations of the second kind.

Yet another type of super-recursive algorithms was introduced in 1965 by Gold and Putnam, who brought in concepts of *limiting recursive function* and *limiting partial recursive function*. In 1967, Gold produced a new version of limiting recursion, also called *inductive inference*, and applied it to problems of learning. Now inductive inference is a fruitful direction in machine learning and artificial intelligence.

One more direction in the theory of super-recursive algorithms emerged in 1967 when Zadeh introduced *fuzzy algorithms*. It is interesting that limiting recursive function and limiting partial recursive function were not considered as valid models of algorithms even by their authors. A proof that fuzzy algorithms are more powerful than Turing machines was obtained much later (Wiedermann, 2004). Thus, in spite of the existence of super-recursive algorithms, researchers continued to believe in the Church-Turing Thesis as an absolute law of computer science.

After the first types of super-recursive models had been studied, a lot of other super-recursive algorithmic models have been created: *inductive Turing machines*, *limit Turing machines*, *infinite time Turing machines*, *general Turing machines*, *accelerating Turing machines*, *type* 2 *Turing machines*, *mathematical machines*, δ-Q-*machines*, *general dynamical systems*, *hybrid systems*, *finite dimensional machines* over real numbers, ***R***-*recursive functions* and so on.

To organize the diverse variety of algorithmic models, we introduce the concept of an algorithmic constellation. Namely, an *algorithmic constellation* is a system of algorithmic models that have the same type. Some algorithmic constellations are disjoint, while other algorithmic constellations intersect. There are algorithmic constellations that are parts of other algorithmic constellations.

Below some of algorithmic constellations are described.

The *sequential algorithmic constellation* consists of models of sequential algorithms. This constellation includes such models as deterministic finite automata, deterministic pushdown automata with one stack, evolutionary finite automata, Turing machines with one head and one tape, Post productions, partial recursive functions, normal Markov algorithms, formal grammars, inductive Turing machines with one head and one tape, limit Turing machines with one head and one tape, reflexive Turing machines with one head and one tape, infinite time Turing machines, general Turing machines with one head and one tape, evolutionary Turing machines with one head and one tape, accelerating Turing machines with one head and one tape, type 2 Turing machines with one head and one tape, Turing machines with oracles.

The *concurrent algorithmic constellation* consists of models of concurrent algorithms. This constellation includes such models as Petri nets, artificial neural networks, nondeterministic Turing machines, probabilistic Turing machines, alternating Turing machines, Communicating Sequential Processes (CSP) of Hoare, Actor model, Calculus of Communicating Systems (CCS) of Milner, Kahn process networks, dataflow process networks, discrete event simulators, View-Centric Reasoning (VCR) model of Smith, event-signal-process (ESP) model of Lee and Sangiovanni-Vincentelli, extended view-centric reasoning (EVCR) model of Burgin and Smith, labeled transition systems, Algebra of Communicating Processes (ACP) of Bergstra and Klop, event-action-process (EAP) model of Burgin and Smith, synchronization trees, and grid automata.

The *parallel algorithmic constellation* consists of models of parallel algorithms and is a part of the concurrent algorithmic constellation. This constellation includes such models as pushdown automata with several stacks, Turing machines with several heads and one or several tapes, Parallel Random Access Machines, Kolmogorov algorithms, formal grammars with prohibition, inductive Turing machines with several heads and one or several tapes, limit Turing machines with several heads and one or several tapes, reflexive Turing machines with several heads and one or several tapes, general Turing machines with several heads and one or several tapes, accelerating Turing machines with several heads and one or several tapes, type 2 Turing machines with several heads and one or several tapes.

The *discrete algorithmic constellation* consists of models of algorithms that work with discrete data, such as words of formal language. This constellation includes such models as finite automata, Turing machines, partial recursive functions, formal grammars, inductive Turing machines and Turing machines with oracles.

The *topological algorithmic constellation* consists of models of algorithms that work with data that belong to a topological space, such as real numbers. This constellation includes such models as the differential analyzer of Shannon, limit Turing machines, finite dimensional and general machines of Blum, Shub, and Smale, fixed point models, topological algorithms, neural networks with real number parameters.

Although several models of super-recursive algorithms already existed in 1980s, the first publication where it was explicitly stated and proved that there are algorithms more powerful than Turing machines was [13]. In this work, among others, relations between Gödel's incompleteness results and super-recursive algorithms were discussed.

Super-recursive algorithms have different computing and accepting power. The closest to conventional algorithms are inductive Turing machines of the first order because they work with constructive objects, all steps of their computation are the same as the steps of conventional Turing machines and the result is obtained in a finite time. In spite of these similarities, inductive Turing machines of the first order can compute much more than conventional Turing machines [14, 5].

Inductive Turing machines of the first order form only the lowest level of super-recursive algorithms. There are infinitely more levels and as a result, the algorithmic universe grows into the algorithmic multiverse becoming open and amenable. Taking into consideration algorithmic schemas, which go beyond super-recursive algorithms, we come to an open world of information processing, which includes the algorithmic multiverse with its algorithmic constellations. Openness of this world has many implications for human cognition in general and mathematical cognition in particular. For instance, it is possible to demonstrate that not only computers but also the brain can work not only in the recursive mode but also in the inductive mode, which is essentially more powerful and efficient. Some of the examples are considered in the next section.

4 Absolute Prohibition in the Closed Universe and Infinite Opportunities in the Open World

To provide sound and secure foundations for mathematics, David Hilbert proposed an ambitious and wide-ranging program in the philosophy and foundations of mathematics. His approach formulated in 1921 stipulated two stages. At first, it was necessary to formalize classical mathematics as an axiomatic system. Then, using only restricted, "finitary" means, it was necessary to give proofs of the consistency of this axiomatic system.

Achieving a definite progress in this direction, Hilbert became very optimistic. As a response to the Latin dictum: "*Ignoramus et ignorabimus*" or "*We do not know, we cannot know*", in his speech in Königsberg in 1930, he made a famous statement:

Wir müssen wissen. Wir werden wissen.
(We must know. We will know.)

Next year the Gödel undecidability theorems were published [15]. They undermined Hilbert's statement and his whole program. Indeed, the first Gödel undecidability theorem states that it is impossible to validate truth for all true statements about objects in an axiomatic theory that includes formal arithmetic. This is a consequence of the fact that it is impossible to build all sets from the arithmetical hierarchy by Turing machines. In such a way, the closed Algorithmic Universe imposed restriction on the mathematical exploration. Indeed, rigorous mathematical proofs are done in formal mathematical systems. As it is demonstrated (cf., for example, [16]), such systems are equivalent to Turing machines as they are built by means of Post productions. Thus, as Turing machines can model proofs in formal systems, it is possible to assume that proofs are performed by Turing machines.

The second Gödel undecidability theorem states that for an effectively generated consistent axiomatic theory T that includes formal arithmetic and has means for formal deduction, it is impossible to prove consistency of T using these means.

From the very beginning, Gödel undecidability theorems have been comprehended as absolute restrictions for scientific cognition. That is why Gödel undecidability theorems were so discouraging that many mathematicians consciously or unconsciously disregarded them. For instance, the influential group of mostly French mathematicians who wrote under the name Bourbaki completely ignored results of Gödel [17].

However, later researchers came to the conclusion that these theorems have such drastic implications only for formalized cognition based on rigorous mathematical tools. For instance, in the 1964 postscript, Gödel wrote that undecidability theorems "do not establish any bounds for the powers of human reason, but rather for the potentialities of pure formalism in mathematics."

Discovery of super-recursive algorithms and acquisition of the knowledge of their abilities drastically changed understanding of the Gödel's results. Being a consequence of the closed nature of the closed algorithmic universe, these undecidability results lose their fatality in the open algorithmic universe. They become relativistic being dependent on the tools used for cognition. For instance, the first undecidability theorem is equivalent to the statement that it is impossible to compute by Turing machines or other recursive algorithms all levels of the Arithmetical Hierarchy [18]. However, as it is demonstrated in [19], there is a hierarchy of inductive Turing machines so that all levels of the Arithmetical Hierarchy are computable and even decidable by these inductive Turing machines. Complete proofs of these results were published only in 2003 due to the active opposition of the proponents of the Church-Turing Thesis [14]. In spite of the fast development of computer technology and computer science, the research community in these areas is rather conservative although more and more researchers understand that the Church-Turing Thesis is not correct.

The possibility to use inductive proofs makes the Gödel's results relative to the means used for proving mathematical statements because decidability of the Arithmetical Hierarchy implies decidability of the formal arithmetic. For instance, the first Gödel undecidability theorem is true when recursive algorithms are used for proofs but it becomes false when inductive algorithms, such as inductive Turing machines, are utilized. History of mathematics also gives supportive evidence for this conclusion. For instance, in 1936 by Gentzen, who in contrast to the second Gödel undecidability theorem, proved consistency of the formal arithmetic using ordinal induction.

The hierarchy of inductive Turing machines also explains why the brain of people is more powerful than Turing machines, supporting the conjecture of Roger Penrose [20]. Besides, this hierarchy allows researchers to eliminate restrictions of recursive models of algorithms in artificial intelligence described by Sloman [6].

It is important to remark that limit Turing machines and other topological algorithms [21] open even broader perspectives for information processing technology and artificial intelligence than inductive Turing machines.

5 The Open World of Knowledge and the Internet

The *open world*, or more exactly, the *open world of knowledge*, is an important concept for the knowledge society and its knowledge economy. According to Rossini [12], it emerges from a world of pre-Internet political systems, but it has come to encompass an entire worldview based on the transformative potential of open, shared, and connected technological systems. The idea of an open world synthesizes much of the social and political discourse around modern education and scientific endeavor and is at the core of the *Open Access* (OA) and *Open Educational Resources* (OER) movements. While the term *open society* comes from international relations, where it was developed to describe the transition from political oppression into a more democratic society, it is now being appropriated into a broader concept of an open world connected via technology [22]. The idea of openness in access to knowledge and education is a reaction to the potential afforded by the global networks, but is inspired by the sociopolitical concept of the open society.

Open Access (OA) is a knowledge-distribution model by which scholarly, peer-reviewed journal articles and other scientific publications are made freely available to anyone, anywhere over the Internet. It is the foundation for the open world of scientific knowledge, and thus, a principal component of the open world of knowledge as a whole. In the era of print, open access was economically and physically impossible. Indeed, the lack of physical access implied the lack of knowledge access - if one did not have physical access to a well-stocked library, knowledge access was impossible. The Internet has changed all of that, and OA is a movement that recognizes the full potential of an open world metaphor for the network.

In OA, the old tradition of publishing for the sake of inquiry, knowledge, and peer acclaim and the new technology of the Internet have converged to make possible an unprecedented public good: "the world-wide electronic distribution of the peer-reviewed journal literature" [23].

The open world of knowledge is based on the Internet, while the Internet is based on computations that go beyond Turing machines. One of the basic principles of the Internet is that it is always on, always available. Without these features, the Internet cannot provide the necessary support for the open world of knowledge because ubiquitous availability of knowledge resources demands non-stopping work of the Internet. At the same time, classical models of algorithms, such as Turing machines, stop after giving that result. This contradicts the main principles of the Internet. In contrast to classical models of computation, as it is demonstrated in [5], if an automatic system, e.g., a computer or computer network, works without halting, gives results in this mode and can simulate any operation of a universal Turing machine, then this automatic (computer) system is more powerful than any Turing machine. This means that this automatic (computer) system, in particular, the Internet, performs unconventional computations and is controlled by super-recursive algorithms. As it is explained in [5], attempts to reduce some of these systems, e.g., the Internet, to the recursive mode, which allows modeling by Turing machines, make these systems irrelevant.

6 Conclusions

This paper shows how the universe (the world) of algorithms became open with the discovery of super-recursive algorithms, providing more powerful tools for computational cognition and artificial intelligence.

Here we considered only some of the consequences of the open world environment of unconventional algorithms and algorithmic constellations for mathematical (computation-theoretical) cognition. It would be interesting to study other consequences of current break through into an open world of unconventional algorithms and computation.

It is known that not all quantum mechanical events are Turing-computable. So, it would be interesting to find a class of super-recursive algorithms that compute all such events or to prove that such a class does not exist.

It might be methodologically and philosophically interesting to contemplate relations between the Open World of Algorithmic Constellations and the Open Science in the sense of Nielsen [24]. For instance, one of the pivotal features of the Open Science is accessibility of research results on the Internet. At the same time, as it is demonstrated in [5], the Internet and other big networks of computers are always working in the inductive mode or some other super-recursive mode. Moreover, actual accessibility depends on such modes of functioning.

One more interesting problem is to explore relations between the Open World of Algorithmic Constellations with the theoretical framework of Info-computationalism, a synthesis of Pancomputationalism (Naturalist Computationalism) with Informational Structural Realism – the model of a universe as a network of computational processes on informational structures. Info-computationalism connects algorithms with interactive computing in natural (physical) systems [25,26][28]. Connecting new unconventional models of super-recursive algorithms and Algorithmic Constellations with unconventional computations performed by natural systems opens new possibilities for the development of innovative models of physical computation with "Trans-Turing" algorithms and "Non-Von" computing architectures. [27].

Acknowledgements. The authors would like to thank Andree Ehresmann, Hector Zenil and Marcin Schroeder for useful and constructive comments on the previous version of this work.

References

1. Frege, G.: Grundgesetze der Arithmetik, Begriffschriftlich Abgeleitet, Viena (1893/1903)
2. Nelson, R.J.: Church's thesis and cognitive science. Notre Dame J. of Formal Logic 28(4), 581–614 (1987)
3. Kalmar, L.: An argument against the plausibility of Church's thesis. In: Constructivity in Mathematics, pp. 72–80. North-Holland Publishing Co., Amsterdam (1959)
4. Gödel, K.: Some Remarks on the Undecidability Results (1986–1995); Gödel, K.: Collected Works, vol. II, pp. 305–306. Oxford University Press, Oxford (1972)
5. Burgin, M.: Super-recursive Algorithms. Springer, Heidelberg (2005)
6. Sloman, A.: The Irrelevance of Turing machines to AI Aaron Sloman. In: Scheutz, M. (ed.) Computationalism: New Directions. MIT Press (2002), `http://www.cs.bham.ac.uk/~axs/`
7. Shannon, C.: Mathematical Theory of the Differential Analyzer. J. Math. Physics 20, 337–354 (1941)
8. Abramsky, S., Jung, A.: Domain theory. In: Abramsky, S., Gabbay, D.M., Maibaum, T.S.E. (eds.) Handbook of Logic in Computer Science, vol. III. Oxford University Press (1994)
9. Edalat, A.: Domains for computation in mathematics, physics and exact real arithmetic. Bulletin of Symbolic Logic 3, 401–452 (1997)
10. Ko, K.: Computational Complexity of Real Functions. Birkhauser, Boston (1991)
11. Weihrauch, K.: Computable Analysis. An Introduction. Springer, Heidelberg (2000)
12. Pour-El, M.B., Richards, J.I.: Computability in Analysis and Physics. In: Perspectives in Mathematical Logic, vol. 1, Springer, Berlin (1989)
13. Burgin, M.: The Notion of Algorithm and the Turing-Church Thesis. In: Proceedings of the VIII International Congress on Logic, Methodology and Philosophy of Science, Moscow, vol. 5, part 1, pp. 138–140 (1987)
14. Burgin, M.: Nonlinear Phenomena in Spaces of Algorithms. International Journal of Computer Mathematics 80(12), 1449–1476 (2003)
15. Gödel, K.: Über formal unentscheidbare Sätze der Principia Mathematics und ver-wandter System I. Monatshefte für Mathematik und Physik 38, 173–198 (1931)
16. Smullian, R.M.: Theory of Formal Systems. Princeton University Press (1962)
17. Mathias, A.R.D.: The Ignorance of Bourbaki. Physis Riv. Internaz. Storia Sci (N.S.) 28, 887–904 (1991)
18. Rogers, H.: Theory of Recursive Functions and Effective Computability. MIT Press, Cambridge Massachusetts (1987)
19. Burgin, M.: Arithmetic Hierarchy and Inductive Turing Machines. Notices of the Russian Academy of Sciences 299(3), 390–393 (1988)
20. Penrose, R.: The Emperor's New Mind. Oxford University Press, Oxford (1989)
21. Burgin, M.: Topological Algorithms. In: Proceedings of the ISCA 16th International Conference "Computers and their Applications", pp. 61–64. ISCA, Seattle (2001)
22. Rossini, C.: Access to Knowledge as a Foundation for an Open World. EDUCAUSE Review 45(4), 60–68 (2010)
23. Budapest Open Access Initiative, `http://www.soros.org/openaccess/read.shtml`
24. Nielsen, M.: Reinventing Discovery: The New Era of Networked Science. Princeton University Press, Princeton and Oxford (2012)

25. Dodig-Crnkovic, G., Muller, V.C.: A Dialogue Concerning Two World Systems. In: Information and Computation, pp. 107–148. World Scientific, New York (2011)
26. Rozenberg, G., Bäck, T.H.W., Kok, J.N. (eds.): Handbook of Natural Computing. Springer, Heidelberg (2012)
27. Dodig Crnkovic, G., Burgin, M.: Unconventional Algorithms: Complementarity of Axiomatics and Construction. Entropy, Special issue "Selected Papers from the Symposium on Natural/Unconventional Computing and its Philosophical Significance" (forthcoming) (2012), http://www.mdpi.com/journal/entropy/special_issues/unconvent_computing
28. Dodig-Crnkovic, G.: Significance of Models of Computation from Turing Model to Natural Computation. Minds and Machines 21(2), 301 (2011)

What Makes a Computation Unconventional?

S. Barry Cooper

University of Leeds, UK

Turing's standard model of computation, and its physical counterpart, has given rise to a powerful paradigm. There are assumptions underlying the paradigm which constrain our thinking about the realities of computing, not least when we doubt the paradigm's adequacy.

There are assumptions concerning the logical structure of computation, and the character of its reliance on the data it feeds on. There is a corresponding complacency spanning theoretical – but not experimental – thinking about the complexity of information, and its mathematics. We point to ways in which classical computability can clarify the nature of apparently unconventional computation. At the same time, we seek to expose the devices used in both theory and practice to try and extend the scope of the standard model. This involves a drawing together of different approaches, in a way that validates the intuitions of those who question the standard model, while providing them with a unifying vision of diverse routes "beyond the Turing barrier".

The results of such an analysis are radical in their consequences, and break the mould in a way that has not been possible previously. The aim is not to question, invalidate or supplant the richness of contemporary thinking about computation. A modern computer is not *just* a universal Turing machine. But the understanding the model brought us was basic to the building of today's digital age. It gave us *computability*, an empowering insight, and computing with *consciousness*. What is there fundamental that unconventional computation directs us to? What *is it* makes a computation unconventional? And having fixed on a plausible answer to this question, we ask: To what extent can the explanatory power of the mathematics clarify key issues relating to emergence, basic physics, and the supervenience of mentality on its material host?

1 Method over Matter

There is a huge literature concerning computability. It has grown beyond what anyone might have anticipated back in the 1930s. The subject has taken on a life of its own, the context has spread across disciplinary boundaries in a startling way and ideas are increasingly hard to categorise and evaluate within traditional structures. During the 2012 centenary of Alan Turing's birth, we were reminded that some of the most important early contributions to the computing revolution came from people who thought deeply about the way the world computes, while gaining strength from their independence of what other people had done. Thank

G. Dodig-Crnkovic and R. Giovagnoli (Eds.): *Computing Nature*, SAPERE 7, pp. 255–269.
DOI: 10.1007/978-3-642-37225-4_17

you, Turing. Relieved of the burden of the taxonomy of who said what and when, let us start by mapping out some of the main features of the standard model of computation. Or rather, the main features of how a starting student in computability might unwrap the model.

The most basic feature, a feature not just of Turing's 1936 Turing machine but of all the equivalent models, is its *disembodiment*. We might have been told at some point that it was devised as a disembodying model of machine computation. Not so, of course. 2012 has made everyone aware of the very specific physicality of the computing situation that Turing was modelling, the predominantly women 'computers' following instructions. One of the special strengths of the Turing model is its close relationship with physical computation, via a very specific deconstruction of a *typical* computational context. We have but a confused idea of what a machine might be. We have a firmer grip on what a 'computer' following instructions might be doing, using well-defined workspace, tools and conventions. The underlying physicality may be highly complex. But such things as the human computer's aches and pains, her feelings of hunger or boredom, are factored out of the process. We extract emergent features of the material context which are far from disembodiments of the computation, but which give us a model which we may re-embody in quite different contexts, and whose mathematical properties can be investigated with a realistic hope of relevance to a wide spectrum of controlled situations.

This provides a green light to those who would turn such a superficial take on physical computation back on its host. In the absence of a corresponding deconstruction of more complex physicality – and ignoring the fact that even more ad-hoc descriptions of particle physics, life, cosmology, human mentality etc. are incomplete – the temptation is to turn the Turing model into metaphor, and then into extended model, in ways we can only argue about. The disembodiment implicit in the standard model is not so simple. It has a character which we should not ignore. Surfing the computational world is fun, but the underlying complexity may still surprise.

We have dwelled on the basic particularity and oddness of the standard model. It is relevant to what we expect of other aspects of the dominant paradigm. Before Turing and logicians like Emil Post, Kurt Gödel, Alonzo Church, and Stephen Kleene came on the scene, an important input to a computation was the computing machine itself. Physically the machine embodied a weighty piece of data. The logician's overview provided an extraction of its essentials in the form of a simple code – a natural number, a finite binary string, or other similar mathematical object. This provided two hugely important features of the standard model, and the modern computational world.

Firstly, however we structure our machines, their descriptions can be converted into data used by machines based on a different logical analysis, enabling the construction of algorithms for converting 'programs' within one framework into one fitting with another. Much activity of the early investigators involved devising such algorithms (say, for converting a description of a λ-computable function into one for an equivalent Turing computable function). The natural conclusion was that Turing machines could 'compute anything'.

Secondly, having trivialised the hardware, the power of the computing paradigm lay in the programming. And with a model which turned machines into data, it was a short step to building a machine which could mine machines for different purposes out of whatever data you gave it – and having decoded the program from the data, could implement it. This was the origin of the so-called 'universal Turing machine'. By a simple mathematical sleight of hand, one had moved machine from physical world to the realm of pure thought. Well, not quite, as Charles Babbage found out in the process of getting his Analytical Engine built prior to 1871. Of course, Babbage's machine, like others pre-dating the 1948 Manchester Small-Scale Experimental Machine (SSEM), or 'Manchester Baby', was not universal machine in the sense of a modern stored-program computer.

Today's computers are a true embodiment of Turing's universal machine in that they enable programs to be combined and edited in increasingly creative ways, without the need for any rebuilding of the programmer-computer physical interface. Nowadays, the program, once input, becomes part of the computer, to be stored, adapted or discarded by the programmer without any grappling with punched cards or realigning of wired connections or switches. Early designs of Babbage, Konrad Zuse and others are 'Turing complete', but lack the vital stored-program feature. This important ingredient of Turing's 1936 logical analysis was incorporated by John von Neumann in his June 1945 EDVAC report, and also features, much less influentially, in Turing's report on the ACE of later that same year.

Universality; the transposability of computational activity from one computing platform to another; the supplanting of the physical by the logical; the redundancy of information beyond the type 1 or type 2 mathematical level — these are familiar aspects of an overarching computational paradigm. The underlying assumptions have served us well, and moulded our thinking about the wider context. One can recognise it in early approaches to artificial intelligence. In the philosophy of mind one has various functionalist viewpoints, with Hilary Putnam explicitly drawing on the universal machine metaphor in his seminal 1960s writings on the topic. Again, in computer science one has the allied notion of 'virtual machine' quite validly useful in both practical and theoretical contexts. One observes the paradigm in the drive to reduce social interaction and development to the algorithmic, setting complex interactive processes within simple rule-based game structures. The feedback between the emergent and the algorithmic, to which we return below, does not fit well with 'corporate thinking', with its drive for strategic certainty.

In computer science and mathematics the paradigm can be detected in sophisticated approaches to the logic of computation, focused on the value of frameworks transferable not just between specific contexts, but between different disciplines. Categorical methods have been productive in the computational context, where according to Samson Abramsky "in the work on concurrent processes, the behaviour *is* the object of interest.

2 Process and Embodiment

History brings its own reminders that computers are not 'just' universal Turing machines. Moving the model from the human 'computers' platform to a more efficient and cost-effective electronic platform was highly non-trivial. Apart from this, the re-embodiment of computing brought us closer to the main point of Turing's 1936 paper — a proof that there are interesting questions beyond the reach of algorithms. In retrospect, one can rephrase this as "computers are stupid", and go on to ask if a 14 billion year old universe is subject to similar limitations. Is the mathematics of Turing's simple diagonalisation of the computable reals unembodiable?

The sort of problems Turing and Alonzo Church showed to be unsolvable by a computer were very natural in an everyday sense. From Turing we know that if **U** is a universal Turing machine then there is no computer that can tell us , for an arbitrary input x, whether **U** will ever produce an output from x. This is the 'unsolvability of the Halting Problem' for **U**, with the set of numbers x on which **U** halts called the Halting Set for **U**. Remembering that x can code a program, this gives us an indication of why computer program checking is such a tough problem. The process tends to be experimental, with a new piece of software requiring a sequence of updates to fix various bugs.

Even closer to home is 'Church's Theorem' — actually Turing proved it too, it is just the negative solution to Hilbert's *Entscheidungsproblem.* This says that if you have a sentence in everyday language (as formalised in first-order logic), there is no computer that can tell us of any such sentence whether it is logically valid or not. To many this is quite informative and counter-intuitive.

One can extract from each of these problems a binary expression for a real number r. Say $r = 0 \cdot r_0 r_1 \ldots$, with each $r_i = 0$ or 1, where $r_i = 1$ exactly when **U** successfully computes on input $x = i$. **U** can be thought to 'compute' r in the sense that this number is uniquely decided by the actions of **U** in computing on $0, 1, 2, \ldots$ successively. r is a very real *feature* of the real world in which **U** lives and operates. However, the level of abstraction of r means that even though we can 'see' **U** computing, we cannot 'see' r at all well. If we could see r we would be moved to allow that it it is 'computed' in some sense. Of course, r does not fit into the classical paradigm, since r is not available as an input to further computation by **U**. Not only can *we* not see r, nor can **U**. And this is not just due to the incomputability. r as a mathematical object is of higher *type* than the natural numbers which **U** usually accepts as an input. Anyway, in the absence of an embodied presence, r is not considered a computed outcome.

Mathematically, r is *definable* from **U**, but the existential quantifier needed to define it puts r on the other side of an unembodied chasm. The question arises: Can this chasm be crossed given the right material conditions?

There is an obvious counterpart of this elevation of type, and Turing's proof of a resulting incomputability. The non-locality of the view amounts to a logical interactivity between computations. The observing process provides the connectivity, with us a player in the physical environment. We are no longer in the presence of an individual computation, it is an interactive *process* at work, with what we will subsequently recognise as an *emergent* incomputable real r. We will come to regard *emergent* as real world analogue of 'definable'. Emergence plays an important role in many dynamical systems, such as weather systems, large scale social interactions, the internet, biology, creative thinking, and turbulent environments of many kinds.

Definability is commonly ignored, or regarded as a logician's playground, with important instantiations in the wider mathematical context. A useful 'missing link' is the fractal, with both precise mathematical description and a visual presence, often enhanced via computer simulation.

In a formal sense, the Halting Problem is in the same world as the Mandelbrot set, for instance. We have gone up another level of the type structure, but there is an in-principle connection. We have a simple computable rule hosted by the complex numbers. Based on this there is a two-quantifier definition of the members of the Mandelbrot set, which with a little manipulation can be reduced to a one-quantifier expression for the purpose of the well-known computer simulation. The computability or otherwise of the Mandelbrot set is still an open problem. But unlike our incomputable Halting Set, the Mandelbrot set comes beautifully and interestingly embodied, with quite visual counterparts to the suspected incomputability. There it is on our computer screen! And we can delve as deeply as we like into this fascinatingly surprising type-2 object. The reason for this is that we are *sampling* this set of complex numbers. The computer screen image involves a trick reduction of type. Turing himself was familiar with the usefulness of statistical sampling for reducing complex information to something computationally approachable. It is not a purely ad hoc methodology. It is a way of recognising the higher type computability enabled by a definition, or by some real world process to whose computational underpinnings we are not privy.

Moving beyond our mathematical comfort zone, we can observe many everyday examples of emergence as instances of objects definable in a real context. We see apparently chaotic environments involving generation of informational complexity via simple rules with a computational character. And we further observe the accompanying, often surprising, emergence of new regularities — such as those of Robert Shaw's dripping taps — entropic resting points, often at most observable via the sort of selective sampling which made visible the embodied Mandelbrot set.

The embodied computation of higher type objects is not in itself a challenge to the classical model. But its character does mell with intuitions concerning unconventionality of computation. And the parallel with the established incomputable r and its mathematical context certainly rings the alarm bells.

3 Emergence and Definability

All around us we see a world exhibiting algorithmic content accompanied by hierarchical structure not easily explainable in terms of the familiar underlying rules. Our everyday lives are built around what appears to be a computable environment, but nature continually surprises us. Much of that surprise is attached to natural form which does appear to be part of a universe which 'knows what it is doing', and it is this we think of as 'emergent'.

The importance of getting a mathematical grip on this omnipresent phenomenon — in evidence from 'strange attractors' to human creativity, and from the origins of life to large-scale cosmic structure — is illustrated by the history of 'British Emergentism', and its heyday in the 1920s. One of the leaders of the movement was the Cambridge philosopher C. D. Broad. Here he is[1] in 1925, attempting an explanation of what emergence is, while pointing to illustrative examples:

> ... the characteristic behaviour of the whole ... could not, even in theory, be deduced from the most complete knowledge of the behaviour of its components ... This .... is what I understand by the 'Theory of Emergence'. I cannot give a conclusive example of it, since it is a matter of controversy whether it actually applies to anything ... I will merely remark that, so far as I know at present, the characteristic behaviour of Common Salt cannot be deduced from the most complete knowledge of the properties of Sodium in isolation; or of Chlorine in isolation; or of other compounds of Sodium, ...

Dramatic scientific developments were in progress around this time. 1925 saw the key elements of the new quantum mechanics put in place by Werner Heisenberg and Erwin Schrödinger, and by the 5th Solvay conference in 1927 quantum theory was revolutionising the foundations of chemistry. The mystery was stripped from the examples from chemistry of Broad and others.[2]

For Stuart Kauffman[3] emergence is not just an example of unconventional computation, it calls into question basic assumptions about the computational content of causality and the deterministic character of the universe:

> We are beyond reductionism: life, agency, meaning, value, and even consciousness and morality almost certainly arose naturally, and the evolution of the biosphere, economy, and human culture are stunningly creative often in ways that cannot be foretold, indeed in ways that appear to be partially lawless. The latter challenge to current science is radical. It runs starkly counter to almost four hundred years of belief that natural

[1] C.D. Broad, *The Mind and Its Place In Nature*, Kegan-Paul, London, 1925, p.59.

[2] See Brian McLaughlin's article "The Rise and Fall of British Emergentism", in *Emergence or Reduction? – Essays on the Prospects of Nonreductive Physicalism* (A. Beckermann, H. Flohr, J. Kim, eds.), de Gruyter, Berlin, 1992, pp.49–93.

[3] Stuart Kauffman, *Reinventing the Sacred: A New View of Science, Reason and Religion*, Basic Books, 2008, p.281.

> laws will be sufficient to explain what is real anywhere in the universe, a view I have called the Galilean spell. The new view of emergence and ceaseless creativity partially beyond natural law is a truly new scientific worldview in which science itself has limits.

Such claims are counterbalanced by words of caution from Ronald Arkin[4]:

> Emergence is often invoked in an almost mystical sense regarding the capabilities of behavior-based systems. Emergent behavior implies a holistic capability where the sum is considerably greater than its parts. It is true that what occurs in a behavior-based system is often a surprise to the system's designer, but does the surprise come because of a shortcoming of the analysis of the constituent behavioral building blocks and their coordination, or because of something else?

In the face of historic confusions, and radical contemporary speculations, the clarifying role of mathematics is urgently needed. This is not to brush aside the more detailed proposals of Kauffman and others. The aim is to place them in a more foundational framework.

To this end, one needs more than the codifying of current 'best observational practice' represented by the Test of Emergence of Ronald, Sipper and Capcarrère in *Design, observation, surprise! A test of emergence* (Artificial Life, **5** (1999), 225–239). Here is a summary of their qualifying criteria:

1) **Design:** The system has been constructed by the designer, by describing local elementary interactions between components (e.g., artificial creatures and elements of the environment) in a language $\mathfrak{L}_1$.
2) **Observation:** The observer is fully aware of the design, but describes global behaviors and properties of the running system, over a period of time, using a language $\mathfrak{L}_2$.
3) **Surprise:** The language of design $\mathfrak{L}_1$ and the language of observation $\mathfrak{L}_2$ are distinct, and the causal link between the elementary interactions programmed in $\mathfrak{L}_1$ and the behaviors observed in $\mathfrak{L}_2$ is non-obvious to the observer – who therefore experiences surprise. In other words, there is a cognitive dissonance between the observer's mental image of the system's design stated in $\mathfrak{L}_1$ and his contemporaneous observation of the system's behavior stated in $\mathfrak{L}_2$.

Might this serve as a test for unconventional computation? Unconventionality certainly requires some obstacle to reduction to basic algorithmic structure. And it is hard to *design* a computational device which has no underpinning of classical ingredients. On the other hand, there are potentially incomputable processes in nature for which 1) or 2) fail. Can a foundational approach make computational sense of the outcome of a quantum measurement leading to a collapse of the wave function?

A nice aspect of the above test is its differentiation between 'designer' and 'observer' languages. This is a feature of the fragmentary nature of science, where

[4] Ronald C. Arkin, *Behaviour-Based Robotics*, MIT Press, 1998, p.105.

it is common to view, say, biology as emergent from an underlying quantum mechanical base, with its own emergent rules and language, non-reducible to the quantum level on which it depends. In the case of the Halting Set for a universal Turing machine, $\mathfrak{L}_2$ is distinguished by the addition of quantification.

Alan Turing recognised something computationally interesting in emergence when he investigated the mathematics of morphogenesis. In the early 1950s Turing wrote his groundbreaking paper on *The chemical basis of Morphogenesis*, in which he proposed a simple reaction-diffusion system describing chemical reactions and diffusion to account for morphogenesis in a range of cases. He even ran computer programs on the early Manchester Mark 1 computer (a more powerful successor to the 'Baby') with the aim of verifying his reaction-diffusion 'design' underlying such emergent patterns as the familiar black and white dappling on a Holstein dairy cow.

What is specially interesting about this work is how it related the powerful descriptive framework of differential equations to emergent form in nature, so exhibiting a connection between the mathematics of higher type objects and apparent emergence. It is hard to claim computational unconventionality on this basis – the solutions to Turing's equations tended to be computable – but then mathematics provides us with little means of identifying real world incomputability. Reducing the Halting Problem to an elusive solution to a non-linear differential equation is not very likely. On the other hand, Marian Pour-el and Ian Richards had some success designing 'A computable ordinary differential equation which possesses no computable solution'[5]

To summarise: Turing provided examples of descriptions of emergent phenomena, whereby one might characterise the emergence as an expression of a higher type computation. And this fits well with the Ronald-Sipper-Capcarrère test for emergence, via the provision of each of design, observation and surprise. With the latter mathematically traceable back to the type-climbing and concomitant potential incomputability of the emergent form.

Is it pure serendipity the discovery that some phenomena can be described in terms of material context? There is a strong intuition that form in the universe arises for a reason. Scientifically this intuition takes the form of an expectation of finding descriptions of phenomena in terms of basic laws of nature. An echo of such an expectation be traced back to Gottfried Leibniz's 1714 description[6] of his 'principle of sufficient reason':

> ...there can be found no fact that is true or existent, or any true proposition, without there being a <u>sufficient reason</u> for its being so and not otherwise, although we cannot know these reasons in most cases.

The intuition that natural phenomena not only generate descriptions, but arise and derive from them, connects with a useful abstraction associated with Alfred Tarski, and growing out of his 1930s work on the notion of truth for formal languages. *Mathematical definability*, or more generally *invariance* under

[5] In: Annals of Mathematical Logic Volume 17, November 1979, Pages 6190.

[6] See *The Monadology*, sections 31, 32.

automorphisms of an appropriate structure, provides an effective organiser of the relative ontology of relations over a structure.

Definability is a basic notion which deserves to be better known in the mathematical world, and in the wider scientific community. It's relevance to physics has been long recognised. Hans Reichenbach worked to axiomatise Einstein's relativity in the 1920s, a project carried forward in relation to general relativity today by the Budapest group of István Németi and Hajnal Andréka. This extension of the fundamental mathematics enables us to deal with a wider range of phenomena, taking us beyond the classical computational model. It gives precision to our experience of emergence as a potentially trans-algorithmic determinant of phenomena.

The overarching aim now is to describe global relations in range of contexts in terms of local structure, so capturing the emergence of large scale formations. And mathematically to formalise such descriptions as definability, or as invariance over basic computational structure. Although Stephen Kleene provided formal content to the notion of higher type computation via a series of papers spanning over 30 years (1959–1991), the physical relevance of his take on the topic needs to be clarified. A forthcoming book on "Computability At Higher Types" by John Longley and Dag Normann is eagerly anticipated. The intuition is that computational unconventionality certainly entails higher type computation, with a correspondingly enhanced respect for embodied information. There is some understanding of the algorithmic content of descriptions. But so far we have merely scratched the surface.

4 Physics and Definability

When a Nobel Prize winner in Physics is quoted as saying[7]:

> The state of physics today is like it was when we were mystified by radioactivity ... They were missing something absolutely fundamental. We are missing perhaps something as profound as they were back then.

people take notice. And this from 2004 winner David Gross did cause something of a stir.

This section is in the nature of a road test for the conceptual framework we have been building up around the notion of unconventionality of a computation. We briefly outline various gaps in the 'standard model' of physics and point to the how a more basic viewpoint can help. The discussion will consist of a brief commentary centred around some revealing quotations from physicists themselves.

We start with Einstein himself complaining about the resort to ad hoc elements of physical theories:

[7] David Gross, quoted in New Scientist, Dec. 10 2005, "Nobel Laureate Admits String Theory Is In Trouble".

> . . . I would like to state a theorem which at present can not be based upon anything more than upon a faith in the simplicity, i.e. intelligibility, of nature . . . nature is so constituted that it is possible logically to lay down such strongly determined laws that within these laws only rationally completely determined constants occur (not constants, therefore, whose numerical value could be changed without destroying the theory) . . .

Notice that this is not just an exhortation to physicists to look for a better theory. It is an expression of faith in the fact that a theory which successfully defines the observable universe should itself be determined by the universe. That is, what we observe is there because the universe is 'self organising' itself, as one would expect of an emergent system with sufficient invariance of its structure to exhibit a high degree of mathematical rigidity. An interesting question is the extent to which constants of the model which make it work, but which are not measurable, are actually defined. In general, one can interpret the necessity of certain values of the constants to make the model work as a sort of invariance. What we suspect of invariance is an elusiveness of algorithmic infrastructure to the relationship between the local and global which makes it possible that aspects of reality are dependent on basic information in a way that is impossible for us to theoretically unravel. We identify below a mathematical model within which to host basic computable causality. Characterising the automorphisms of this model promises to be a key task.

Here is a more recent questioning of progress towards a more comprehensive model of physics, from Peter Woit, author of the book *Not Even Wrong – The Failure of String Theory and the Continuing Challenge to Unify the Laws of Physics* (Jonathan Cape, 2006):

> By 1973, physicists had in place what was to become a fantastically successful theory of fundamental particles and their interactions, a theory that was soon to acquire the name of the standard model. Since that time, the overwhelming triumph of the standard model has been matched by a similarly overwhelming failure to find any way to make further progress on fundamental questions.

And one of Peter Woit's concerns is those undefined constants:

> One way of thinking about what is unsatisfactory about the standard model is that it leaves seventeen non-trivial numbers still to be explained, . . .

Alan Guth, originator of the inflationary hypothesis, would like to see the laws of physics defined:

> If the creation of the universe can be described as a quantum process, we would be left with one deep mystery of existence: What is it that determined the laws of physics?

If we think we are observing the universe defining its own laws, we can but hope to have access to the defining process in the course of time.

We are talking here about hugely unconventional computation. It may be so unconventional that for us it is hardly computation at all. But its existence can be framed as a something feasibly approachable, at least in principle. Roger Penrose[8] calls it 'Strong Determinism':

> [According to Strong Determinism] ... all the complication, variety and apparent randomness that we see all about us, as well as the precise physical laws, are all exact and unambiguous consequences of one single coherent mathematical structure.

In our final section, we fill in the missing ingredient — namely, the fundamental mathematical host for all this embodied information, definability and higher order computation. Before that, a remark regarding mathematical structures: Mathematical structures commonly consist of objects connected by operations or relations. Sometimes the difference between these classes is blurred, but in an interesting structure there are objects which accumulate *information* expressive of their context in the structure. Sometimes this information can be 'read' by the relations on the structure, which express a formal 'causality', whereby the distribution of information itself has a structure. This appears to be a feature of our own universe.

5 Modelling Basic Causality

Another quotation, this time from Lee Smolin's 2006 book on *The Trouble with Physics*, p.241:

> ... causality itself is fundamental ...

The 'early champions' of the role of causality mentioned by Smolin – Roger Penrose, Rafael Sorkin, Fay Dowker, Fotini Markopoulou – make a doughty bunch, formidable protagonists in contemporary turf wars around quantum gravity, causal sets and a hydra-headed superstring theory. The aim, as outlined by Smolin, is a more comprehensively immanent universe[9]:

> It is not only the case that the spacetime geometry determines what the causal relations are. This can be turned around: Causal relations can determine the spacetime geometry ... Its easy to talk about space or spacetime emerging from something more fundamental, but those who have tried to develop the idea have found it difficult to realize in practice. ... We now believe they failed because they ignored the role that causality plays in spacetime. These days, many of us working on quantum gravity believe that causality itself is fundamental – and is thus meaningful even at a level where the notion of space has disappeared.

---

[8] Roger Penrose: Quantum physics and conscious thought, in *Quantum Implications: Essays in honour of David Bohm* (B.J. Hiley and F.D. Peat, eds.), pp.106-107.

[9] Lee Smolin, *The Trouble With Physics*, p.241.

Note that we are talking about very specific observed and computationally well-served *causality* here, which largely frees us from the strictures of John Earman[10] regarding the more wide-ranging use of the term:

> ...the most venerable of all the philosophical definitions [of determinism] holds that the world is deterministic just in case every event has a cause. The most immediate objection to this approach is that it seeks to explain a vague concept - determinism - in terms of a truly obscure one - causation.

In fact, a primary objective of this modelling of basic computable causality is the clarity that the mathematics of the model brings to less easily described causality, and to issues regarding over causation, downward causation and non-locality.

The question is, what kind of causality fails to engage with the informational content of the reality it structures? The relevance of the question derives from the fact that it is causal structure from which information derives and whereby it is stored. Computation is about information, and potentially equipped to model the way in which the basic causality of our universe respects and transports information. Once again it was Turing gave us a precise formulation corresponding to the fundamentality of the intuition.

Alan Turing's 1939 paper is a neglected masterpiece — less cited than the more famous trio of papers that gave us the universal Turing machine, the Turing test for intelligence, and the mathematics of morphogenesis — but crackling with ideas and perceptive intuitions. The *oracle Turing machine* as it came to be called appears on just one page of this densely argued article. Essentially, it equips the computer – in the form of a Turing machine – to roam the scientific universe of real numbers, accepting type 1 inputs, and outputting, if we are lucky, type 1 outputs. This sometimes described as *relative* computation.

An oracle Turing machine exactly expresses the character of basic causality in the world, progressively sampling information and transferring it comprehensively across time and space.

The mathematician or computer scientist — and maybe Turing himself at the time of its invention — regards the oracle machine as a model of how we might compute using data given to us from an unknown source. This viewpoint, together with observation of the apparent actuality of incomputability in the natural universe, provides the basis for Jack Copeland's notion of *hyper computation* (beyond the Turing barrier).

But the physicist is presented with a model — the *Turing universe* — within which the computable content of Newtonian dynamics comfortably fits; at a basic level of course. As Poincaré speculated, and researchers from Kreisel in 1970, to Beggs, Costa and Tucker today observed, more broadly interactive contexts based on Newton's laws can generate infinitary mathematics with attendant incomputabilities.

[10] In *A Primer On Determinism*, D. Reidel/Kluwer, 1986, p.5.

Mathematically, the type-2 computable functions Φ over the reals are termed *Turing* – or *partial computable* (p.c.) – *functionals*. Turing, despite his longterm interest in interactive computation (mainly between humans and machines), seems to have never mentioned his oracle machines again. It was left to another highly creative but under-appreciated mathematician, Emil Post, 15 years older than Turing, to set in motion the mathematical development of Turing's model. In 1944 Post defined the *degrees of unsolvability* – later called the *Turing degrees* – as a classification of reals in terms of their *relative* computability.

Strangely, the subsequent investigation of the mathematical character of the Turing degree structure was a process entirely detached from reality. There was no sense at all of relevance to the real world. The fact that the Turing universe underpinned a wide range of dynamical contexts in which the 'design' was understood meant nothing. The possibility that all sorts of higher structure might be better understood via an analysis of definability or invariance in the basic underlying model was never entertained. I was there through a golden age of technical development. Turing was gone, taking with him his broadly questioning brilliance, leaving behind a universally adopted computational paradigm. We recursion theorists were busy doing our sums while the natural world around us computed in mysterious and wondrous ways. There was no such thing as unconventional computation.

There were mathematical events beneath which one can retrospectively detect a sort of subliminal prescience. In 1965 Hartley Rogers gave a fascinating talk (judging by the 1967 paper[11] that came out of it) at the Tenth Logic Colloquium in Leicester, England. What was remarkable was the focus[12] on the large scale structure of the Turing universe, via the notions of invariance and definability that we have identified as relevant to the emergence of form in wide range of different environments. There was in evidence a 'Hartley Rogers Agenda', built around a number of deep and difficult questions about the global character of the Turing degrees. Over the years, there has been a growing intuition that Rogers' questions are key to pinning down how higher order relations on the real world can appear to be computed. Much of the progress with these questions rests on the richness of Turing structure discovered so far. Mathematically structural pathology is a disappointment. Out in the real world pathology is super-abundant, both generator of and avatar of a richness of real world definability. In the Turing universe. the pathology takes on a parallel role, becoming the raw material for a multitude of definable relations, counterparts to visibly 'computable' structure out in the real world.

[11] H. Rogers, Jr., Some problems of definability in recursive function theory, in *Sets, Models and Recursion Theory* (J. N. Crossley, ed.), Proceedings of the Summer School in Mathematical Logic and Tenth Logic Colloquium, Leicester, August–September, 1965, North Holland, Amsterdam, pp. 183–201.

[12] Hartley Rogers' 1965 Agenda, in 'Logic Colloquium '98' (S.R. Buss, P. Hajek and P. Pudlak, eds.), Proceedings of the Annual European Summer Meeting of the ASL, held in Prague, Czech Republic, August 9-15, 1998, Lecture Notes in Logic 13, Association for Symbolic Logic/A.K. Peters, Natick, Massachusetts, pp. 154-172.

6 Undefinable Relations

We might have developed the view of unconventional computation as higher type computation in various guises – emergence, definability etc – in other contexts. Another fruitful workspace would have been that of artificial intelligence/neuroscience. Each has it's special strengths. That of the physics is the clear way in which it displays the fragility of definability, and the consequences of its failure.

Failures of definability are not necessarily negative in their impact. Many friendly features of our universe depend on them. In physics, there is a wide range of special symmetries underpinning aspects of our observed world. A symmetry is of course an instance of an automorphism at work, maybe small-scale or very selective in its scope. More important broad impact symmetries include the relationship of the SU(3) symmetry group to the quark model underlying hadrons, for which Murray Gell-Mann got a nobel prize in 1969. One of the interests of such particular examples is that they point to the possibility of a rich automorphism structure underlying the basic causal structure, and hence to the identification on new relations defined/computed within the physics with potentially far-reaching explanatory power.

Back in the underlying Turing model, there is some disagreement about the potential character of the automorphism group of both the local and the global structures. An interest in the so-called 'Bi-interpretability Conjecture' originating with Leo Harrington goes back around 30 years, during which time various people have managed to prove partial versions of the conjecture, with interesting consequences for the automorphism groups. Essentially, what the conjecture says is that there is a close enough correspondence between the structure of the Turing degrees and that of second order arithmetic for the two structures to share a number of characteristics, particularly related to automorphisms and definability. A full verification of bi-interpretability would impose rigidity on the Turing universe, and invalidate it as a model for the real universe, which appears to be far from rigid. There is no consensus of informed guesses concerning rigidity.

Failure of rigidity would have potentially dramatic consequences for the long-standing search for a 'realistic' interpretation of quantum non locality and the collapse of the wave function in conjunction with a measurement. What we commonly have is the deterministic continuous evolution of the wave equation describing a physical system via Schrödinger's equation, involving the superposition of basis states. We may then have a probabilistic non-local discontinuous change due to a measurement – and observe a jump to a single basis state. There are various interpretations of this. The simplest is that what we are encountering is a level of failure of definability at the ontological level of the quantum world — there is just not enough connectivity and information down there to uniquely identify basis states. While the intervention entailed by the measurement changes the situation. If a higher type relationship of definability gets unconventionally computed, it is allowed to operate non-locally, without any of the usual problems for the physics.

There are wider ramifications to such a 'realistic' and immanent deciding of physical transitions. The Many Worlds interpretation and its Multiverse derivatives begin to look pleasingly redundant. It is not good that this becomes yet another misfortune to impact on Hugh Everett III and his family – nowadays represented by his son Mark, the talented lead singer of the EELS. No doubt, even with such a powerfully persuasive replacement for Many Worlds and the various Multiverses, via unconventional computation, it will be hard to divert David Deutsch from his view[13] that:

> ... understanding the multiverse is a precondition for understanding reality as best we can. Nor is this said in a spirit of grim determination to seek the truth no matter how unpalatable it may be ... It is, on the contrary, because the resulting world-view is so much more integrated, and makes more sense in so many ways, than any previous world-view, and certainly more than the cynical pragmatism which too often nowadays serves as surrogate for a world-view amongst scientists.

But there are many others, like presumably George Ellis,[14] would breath a sigh of relief:

> The issue of what is to be regarded as an ensemble of 'all possible universes' is unclear, it can be manipulated to produce any result you want ... The argument that this infinite ensemble actually exists can be claimed to have a certain explanatory economy (Tegmark 1993), although others would claim that Occam's razor has been completely abandoned in favour of a profligate excess of existential multiplicity, extravagantly hypothesized in order to explain the one universe that we do know exists.

There are many other ways in which the admission of an extended computational repertoire can bolster the integrity of the observed 'one universe that we do know exists'. Physics is just one area that can benefit from the mathematics of definability, invariance, emergence and higher type computation. Alan Turing would have been fascinated.

---

[13] David Deutsch, *The Fabric of Reality*, Allen Lane, 1997, p.48.

[14] The Unique Nature of Cosmology, in *Revisiting the Foundations of Relativistic Physics* (eds. Abhay Ashtekar et al), Kluwer, 1996, p.198.

Author Index

Computing Nature Cover Bios

Gordana Dodig-Crnkovic is Associate Professor of Computer Science at Mlardalen University, Sweden. Her research interests include: Computing paradigms, Natural computing, Social Computing and Social Cognition, Info-computational models, Foundations of Information, Theory of Science/Philosophy of Science, Computing and Philosophy and Ethics. She is a co-organizer, PC member and speaker at numerous conferences. She is on scientific board of several journals, research institutes and research networks.

Dodig Crnkovic published the authored books *Investigations into Information Semantics and Ethics of Computing* (2006) and *Information and Computation Nets* (2009) and edited volumes *Computation, Information, Cognition* (with Susan Stuart) 2007 and *Information and Computation* (with Mark Burgin) 2011

Raffaela Giovagnoli is Adjunct Professor at the Pontifical Lateran University in Rome, where she teaches the course "Issues of Social Ontology". She has been nominated Contract Professor in Political Philosophy at the University of Salerno (2006) and Contract Professor in Philosophy of Language at the University of Rome "Tor Vergata" (2008). She has been carrying her research work at the Universities of Frankfurt, Pittsburgh and Berkeley. She is on the scientific board of the journals "Etica & politica/Ethics & Politics" and "Philosophy Study". She is also a member of the Berkeley Group on Social Ontology.

Giovagnoli is author of the following books: *Habermas: agire comunicativo e Lebenwelt* (Roma 2000), *Razionalità espressiva. Scorekeeping: inferenzialismo, pratiche sociali e autonomia* (Milano 2004) and *Autonomy. A Matter of Content* (Firenze 2007). Moreover, she published numerous scholarly essays in international and national journals.