Studies in Applied Philosophy, Epistemology and Rational Ethics

Volume 35

About this Series

Studies in Applied Philosophy, Epistemology and Rational Ethics (SAPERE) publishes new developments and advances in all the fields of philosophy, epistemology, and ethics, bringing them together with a cluster of scientific disciplines and technological outcomes: from computer science to life sciences, from economics, law, and education to engineering, logic, and mathematics, from medicine to physics, human sciences, and politics. It aims at covering all the challenging philosophical and ethical themes of contemporary society, making them appropriately applicable to contemporary theoretical, methodological, and practical problems, impasses, controversies, and conflicts. The series includes monographs, lecture notes, selected contributions from specialized conferences and workshops as well as selected Ph.D. theses.

Advisory Board

More information about this series at http://www.springer.com/series/10087

Simone Pinna

Extended Cognition and the Dynamics of Algorithmic Skills

Springer

Simone Pinna
Dipartimento di Pedagogia, Psicologia,
Filosofia. ALOPHIS—Applied LOgic,
Philosophy, and HIstory of Science
Università di Cagliari
Cagliari
Italy

ISSN 2192-6255 ISSN 2192-6263 (electronic)
Studies in Applied Philosophy, Epistemology and Rational Ethics
ISBN 978-3-319-51840-4 ISBN 978-3-319-51841-1 (eBook)
DOI 10.1007/978-3-319-51841-1

Library of Congress Control Number: 2016963180

Printed on acid-free paper

This Springer imprint is published by Springer Nature
The registered company is Springer International Publishing AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

For Marta

Foreword

From Turing Machines to the Dynamical Explanation of Algorithmic Skills

What is the basis of our ability to transform symbols by deliberately applying appropriate computational rules? What is an algorithmic skill and how does it develop? Are such skills essentially internal and purely mental, or do they depend on the dynamical interaction between internal mental factors and external (environmental, bodily, etc.) ones? Simone Pinna's book starts from these questions, and it gradually constructs precise and very well-argued answers to them.

To dispel possible misunderstandings, it is important to make clear from the start that the very delimitation of the research subject and the consequent formulation of the related problems is no trivial matter. The subject of this book is a particular *kind* of cognitive phenomenon that, *as such*, has never undergone before a systematic investigation of similar scope or depth. The problem at issue is an inquiry of *all* human cognitive activities that consist in the deliberate and controlled execution of any *set of calculation rules* (*algorithm*) capable of producing a determinate transformation of symbols.

Most familiar examples of these cognitive phenomena can be found in the field of mathematics: the execution of determinate rules for arithmetical or algebraic calculation, such as the well-known right-to-left column algorithm for adding two or more addends, or the quadratic formula for solving a second-degree equation. However, the domain of the cognitive phenomena under investigation is not limited to mathematics, but it includes all the high-level cognitive activities specified above, which elsewhere have been called *phenomena of human computation* (Giunti 2009, Sect. 4; Giunti and Pinna 2016, Sect. 5).

Given the wide extension and variety of these phenomena, which can be found in almost any field of symbolic thinking, one might be tempted to identify them with thought itself. Such an identification might even seem obvious if one accepts the Computational Hypothesis of Mind (CHM), according to which any thought is ultimately reducible to a computational activity.

Nevertheless, such an identification is untenable. In the first place, Pinna makes clear (see Sect. 3.2.2, pp. 66–67) that any investigation of the phenomena of human computation is independent of CHM. For it is conceivable that cognitive phenomena are stratified in different levels, whose distinctive characters are emergent with respect to lower levels and thus irreducible to them. According to this view, the computational character would only belong to the phenomena of some higher level, which would not exhaust the much wider field of all thought activities.

In the second place, even if we take CHM to be true, it still does not follow that the phenomena of human computation are identifiable with thought itself. CHM in fact asserts a generic computational, or algorithmic, nature of any thought activity. But this is not sufficient for being a phenomenon of human computation, in the precise sense specified above. The distinctive feature of these phenomena is the *deliberate and controlled execution of a well-specified algorithm*, and not just the fact that a thought activity can be described, analyzed, or explained as a computation, or even identified with it.

For example, subcognitive activities, or even unconscious ones, such as the unexpected emergence of the solution of a complex problem, can be analyzed and explained from a computational point of view, but this is not sufficient for them to be a phenomenon of human computation. For, even if we take for granted that they result from the execution of some algorithm, this execution is neither deliberate nor controlled, nor does it presuppose any specification of the algorithm itself.

It is quite surprising to realize that, to date, phenomena of human computation have not been the subject of *specific* studies addressed to highlight their basic structure and their relations to the development of corresponding algorithmic skills. In fact, cognitive science has not been able so far to devise a well-defined *theoretical framework* for studying phenomena of *this kind* from a systematic and unifying point of view.

In this respect, in the first chapter of this book (see Sect. 1.4), Simone Pinna draws attention to a fact whose implications have been overlooked so far. Alan Turing, in the first part (pp. 249–252) of Sect. 9 of his fundamental article "On computable numbers, with an application to the *Entscheidungsproblem*" (1936), very clearly states that the computing machines there defined—today known as Turing machines (TMs)—are to be thought as adequate abstract models of a human being that, in a deliberate and controlled way, executes algorithms of an *exceedingly simple kind*, which nonetheless are capable to produce all that a human being is able to compute. Let us agree to call this statement "Turing Thesis" (TT).

TT is usually interpreted as the intuitive ground on which we can erect a very strong justification for Church Thesis (CT), according to which all that a human being is able to compute is recursive. For, if TT is true, the truth of CT immediately follows from it and the fact that whatever is computable by a Turing machine is recursive. This argument, sometimes called the *analogical argument*, is often considered as the strongest argument in favor of CT. For this reason, CT *itself* is jointly attributed to Church and Turing, and sometimes, CT is also called Human version of the Church-Turing Thesis (HCTT).

Nonetheless, this is not the only possible interpretation of TT and, perhaps, not even the most obvious one. It would seem quite natural to also interpret TT as a methodological statement, which suggests the most adequate *kind* of model to describe *all* phenomena of human computation. The first part of TT asserts that Turing machines are adequate models of a *proper subclass* of human computation phenomena (namely, all those phenomena whose algorithm is exceedingly simple); its second part asserts that these phenomena, even though not exhausting the totality of human computation phenomena, are nonetheless very much representative of it, for their simple algorithms are sufficient for also producing whatever is computable by means of the presumably more complex algorithms of all other phenomena. But then, as TMs are adequate models of a very much representative subclass of human computation phenomena, it is natural to suppose that the *kind* of model adequate to describe *all* these phenomena must be an appropriate generalization of the TMs that preserve their structure and basic design.

To sum up, this second interpretation of TT asserts that TMs supply us with the theoretical background, or the conceptual horizon, within which to define the *kind* of model adequate to describe *all* phenomena of human computation. Nevertheless, as just mentioned earlier, the methodological suggestion provided by the second interpretation of TT has not received so far much attention or credit in cognitive science research. It is natural to ask why this has happened.

To explain this fact, we need to clearly understand the kind of computation, exceedingly simple, of which a TM is an adequate abstract model. According to Turing (1936, pp. 249–251), this kind of computation is based on two fundamental elements: a finite number of *internal states*, which correspond to the possible states of the working memory of the human being, and a tape divided into cells, whose number is finite but can be increased without limits as needed. The tape is to be thought as the *external support* where the human being *writes* the partial result of the computation in a single cell and then *shifts* his/her attention to right or left, within a finite number of cells, as prescribed by the particular instruction executed. At each computation step, the instruction to be executed is chosen according to the internal state and the symbol *read* in the cell where the human being concentrates his/her attention.

The crucial point of the simple kind of computation modeled by a TM is that all symbolic transformations are not performed *internally*, in the working memory of the human being, but externally, for all symbols that are subsequently read or modified are written on the tape, which is, according to Turing (1936, p. 249), the simplest *external support* that enables the human being to perform any calculation. To put it in slightly different terms, in a phenomenon of human computation of the simple kind modeled by a TM, *all symbolic transformations take place externally*.

Simone Pinna points out very clearly that this kind of computation, whose symbolic transformations are all external, does not belong to the domain of the *cognitive operations* admitted by classic computationalism (CC). In fact, for this philosophical and scientific movement of the second half of last century,

> […] cognitive operations consist essentially of internal symbolic transformations based on purely syntactic rules. This central idea, that lies behind many important cognitive theories, like Newells Physical Symbol System (Newell 1980) and Fodors Computational Theory of Mind (Fodor 1975), has represented the standard view in cognitive science since the outset of this discipline, and has been the theoretical background on which the first Artificial Intelligence (AI) programs were constructed (Chap. 1, p. 10).

Let us now sum up what has been said so far. For CC, all cognitive operations are *internal* symbolic transformations based on purely syntactic rules, but the symbolic transformations of the simple phenomena of human computation modeled by TMs are all external. It thus follows that, for CC, these symbolic transformations are not cognitive operations and, consequently, the simple phenomena modeled by TMs are not within the domain of inquiry of cognitive science.

Given the strong influence that CC exerted on the development of cognitive science, it is now clear why the methodological suggestion provided by the second interpretation of TT could not receive much attention or credit. For the simple phenomena of human computation on which this suggestion is based were widely perceived to be outside the proper domain of inquiry of cognitive science.

According to Pinna (see Sect. 1.3), the fact that the simple kind of phenomenon modeled by a TM was taken to be out of the domain of cognitive science resulted into two effects. On the one hand, Turing's psychological interpretation of a TM, which implied the interaction of a human calculator with an external support through the psychophysical operations of reading, writing, and attention shift, was reduced to a mere metaphor of the real symbolic computation that would take place internally to the subject. On the other one, the basic structure of a TM was only seen as the abstract precursor of the architecture of a modern digital computer, but it was not considered as an adequate source of inspiration for the construction of detailed cognitive models of phenomena of human computation, as the obvious methodological interpretation of TT would instead have suggested.

However, starting at the end of last century, the emergence a number of new elements has begun to modify this scenario. According to Pinna (see Introduction and Sects. 1.4, 2.4, 3.1.2, 3.1.3), the first element to be considered is the revival and rediscovery, by Wells (1998, 2005), of the original psychological interpretation of a TM (Turing 1936, Sect. 9.I). Such a rediscovery is the foundation on which Wells grounds his proposal of a new computational paradigm for cognitive science, the so-called Ecological Functionalism (EF).

Like many others, Wells maintains that cognition is ultimately due to computational processes, but it is the very concept of *cognitive computation* that must be thoroughly rethought in light of the original psychological interpretation of a TM. Contrary to CC, cognitive computations are not wholly internal symbolic transformations but, analogously to the simple kind of computation modeled by a TM, they are the result of a definite interaction between an *agent* (the human calculator) and his/her *environment* (the external calculation support), which supplies the agent with the most adequate means for the particular cognitive task to be carried out.

The second new element is that, at the end of last century, in the philosophy of mind and cognitive science the idea has begun to spread that cognitive operations are not purely internal, but depend on the interaction of internal and external factors of an embodied agent situated in a definite environment. In this respect, in Chap. 2, Pinna reviews the main tenets of the Extended Mind Hypothesis (EMH) (Clark and Chalmers 1998), and he highlights its strong similarities and connections with Wells' EF.

Lastly, the third element to be taken into account is the uprising and diffusion of the so-called *dynamical approach to cognition*. Analogously to EF and EMH, the dynamical approach places the interaction between the agent's internal dynamics and the one of his/her body and environment at the heart of the study of cognition. This dynamics produces a definite trajectory of the observed cognitive parameters.

In Chap. 3, Pinna clearly explains how Wells' Ecological Functionalism can be seen as a particular form of dynamical approach, in which both time and state space are taken to be discrete. This may seem somehow surprising because, according to a quite widespread interpretation (van Gelder and Port 1995; van Gelder 1998, 1999), the dynamical approach would only admit continuous systems, with respect to both time and state space. However, this interpretation is surely too restrictive, for according to several authors (Giunti 1992, 1995, 1997; Beer 1998; Wells 1998) there is no principled reason to exclude discrete dynamical systems from the domain of possible cognitive models.

We have seen why, while classic computationalism was the dominant view in cognitive science, the methodological suggestion contained in the second interpretation of Turing Thesis did not stand many chances of being entertained or developed. For this to happen, it was necessary to accept first the idea that the simple kind of phenomenon modeled by a TM—such that all symbolic transformations take place externally—could be the keystone for the explanation of the much wider class of all phenomena of human computation. The change in the very concept of cognition that was then promoted by Ecological Functionalism, the Extended Mind Hypothesis, and the Dynamical Approach created the favorable context for such an idea to become acceptable.

In Chap. 4, Pinna illustrates the Bidimensional Turing Machine Theory (BTM-T). I set forth this theory (Giunti 2009) as a first attempt to carry through on the methodological suggestion provided by the second interpretation of Turing Thesis, according to which the *kind* of model adequate to describe *all* phenomena of human computation is to be searched as an appropriate generalization of the TMs, which preserve their structure and basic design.

I identified such a basic structure with (i) the three components of the complete state of a TM (internal state, content of the tape, and position of the read-write-shift head on the tape), each interpreted according to Turing's original psychological interpretation (respectively, mental state, content of the external calculation support, region of the external support on which attention is focused), and (ii) the interaction mechanism that transforms the present complete state into the next one; such a mechanism is based on the three fundamental operations of reading, writing, and (attention) shifting.

The generalizations that I proposed concern instead (i) the external support, which, in a BTM, is not just a tape but a bidimensional grid; (ii) the more complex structure of the internal state, which is composed of different registers, whose contents are allowed to be strings of symbols or natural numbers; and (iii) the introduction of auxiliary functions and relations, which provide for more sophisticated modifications of the internal state and the scanned symbol, as well as for more complex shifting than a traditional TM.

The intended application domain of BTM-T is the class of all phenomena of human computation that, at most, employ an external support analogous to a squared sheet of paper[1] or, as a limit case, no external support. However, in Giunti (2009), I only proposed some examples of application of the theory to the description of well-known algorithms, but I did not make sufficiently clear how BTM-T could foster a systematic and unified study of the phenomena of human computation that belong to its intended domain. This study should be able to highlight the basic structural features of all these phenomena, as well as their relevant differences, and relate them to the development of corresponding algorithmic skills.

In Chap. 5, Simone Pinna confronts this very problem head-on, by employing a quite definite method. The first step of this method consists in selecting a small number of phenomena of human computation, which nonetheless be representative of different types of algorithm that a human being can execute with different degrees of difficulty. Pinna takes into account three distinct arithmetical tasks concerning additions of natural numbers expressed in decimal notation: (a) adding two one-digit addends; (b) adding an arbitrary number of addends of any number of digits; and (c) mentally adding a three-digit addend and a two-digit addend.

As regards task (a), Pinna considers three different ways to carry it out, which consist in the execution of three different algorithms: (1) *direct sum with written result*: Mentally add the two one-digit addends and write the result on a sheet of paper; (2) *counting-on with written result*: Memorize the two one-digit addends and mentally add to the first addend as many units as those of the second addend. Employ your fingers to keep track of the number of units already added and stop when the number represented by your fingers equals the second addend. Finally, write the result on a sheet of paper; (3) *counting-on without written result*: This procedure is identical to the previous one, but the final result is just kept in memory, not written on paper. As regards task (b), Pinna considers just one computation strategy that consists in (4) the execution of the well-known *right-to-left column rule* for addition, by using *paper and pencil*. Finally, as regards task (c), Pinna takes into account two different algorithms: (5) *mental decomposition in tens and units*:

[1]This and other formal limitations of BTM-T have been superseded by the Algorithmically enhanced Turing Machine Theory (ATM-T), which has been set forth in a recent paper (Giunti and Pinna 2016). ATM-T is a natural generalization of BTM-T, which applies to *all* phenomena of human computation. They include the deliberate execution of algorithms whose external support has an arbitrary dimension and, as a limit case, also the deliberate execution of algorithms that employ no external support—so that the whole computation takes place internally.

Decompose both addends in tens and units, add tens first and then units, and finally add the two results (every step is purely mental); (6) *mental right-to-left column rule*: Add the two addends by mentally applying the right-to-left column rule for addition, but without writing anything at all.

Pinna then shows how it is possible to construct, for each of the phenomena (1)–(6), the instruction table of a BTM that formalizes the corresponding algorithm. In other terms, Pinna employs the informal description of the algorithm to pinpoint the BTM that formally expresses, as well as possible, the algorithm itself. The BTM thus selected can then be used as a *dynamical model* of the phenomenon under investigation.

Having built a dynamical model for each of the phenomena at stake, Pinna goes on to analyze and compare these models. The fact that all models are BTMs and thus share the same basic structure allows to capture the common elements of all phenomena and, at the same time, it facilitates the comparison and individuation of their specific characters.

On the basis of these analyses, Pinna is finally able to advance a series of general hypotheses, empirically testable in principle, which explain (i) how more sophisticated algorithmic skills can gradually develop (Sect. 5.3.4, Hypothesis 1b); (ii) why, and in what precise sense, computations that use an external support are advantageous (Sect. 5.4.3, Hypothesis 2b); and (iii) the specific role of this kind of computation for the development of more complex algorithmic skills (Sect. 5.5.3, Hypothesis 3).

We can thus safely conclude that Pinna's book is a groundbreaking work, which for the first time shows how it is possible to study the phenomena of human computation from a systematic and unifying point of view, by inserting them within the well-defined theoretical framework of BTM-T.[2]

Florence, Italy | Marco Giunti
September 2016

References

Beer R (1998) Framing the debate between computational and dynamical approaches to cognitive science. Behav Brain Sci 21:630

Clark A, Chalmers D (1998) The extended mind. Analysis 58:10–23

Fodor J (1975) The language of thought. Harvard University Press, Cambridge, MA

Giunti M (1992) Computers, dynamical systems, phenomena, and the mind. Ph.D. dissertation, Indiana University, Bloomington

Giunti M (1995) Dynamical models of cognition. In: Port R, van Gelder T (eds) Mind as motion. The MIT Press, Cambridge, MA

Giunti M (1997) Computation, dynamics, and cognition. Oxford University Press, New York

[2]Or the more general theoretical framework of ATM-T (see Footnote 1).

Giunti M (2009) Bidimensional Turing machines as Galilean models of human computation. In: Minati G, Abram M, Pessa E (eds) Processes of emergence of systems and systemic properties. World Scientific, Cambridge, MA

Giunti M, Pinna S (2016) For a dynamical approach to human computation. Logic J IGPL 15 (2):557–569

Newell A (1980) Physical symbol systems. Cogn Sci 4:135–183

Turing A (1936) On computable numbers, with an application to the *Entscheidungsproblem*. In: Proceedings of the London mathematical society. Oxford Journals, London, pp 230–265

van Gelder T (1998) The dynamical hypothesis in cognitive science. Behav Brain Sci 21:615–665

van Gelder T (1999) Dynamic approaches to cognition. In: Wilson RA, Keil FC (eds) The MIT encyclopedia of the cognitive sciences. MIT Press, Cambridge, MA, pp 244–246

van Gelder T, Port R (1995) Its about time: An overview of the dynamical approach to cognition. In: Port R, van Gelder T (eds) Mind as motion. MIT Press, Cambridge, MA

Wells A (1998) Turing's analysis of computation and theories of cognitive architecture. Cogn Sci 22:269–294

Wells A (2005) Rethinking cognitive computation: Turing and the science of the mind. Palgrave Macmillan, Basingstoke, UK

Acknowledgements

The research for this book was supported by the Regione Autonoma della Sardegna (RAS), L.R. 7/2007, within the project “Investigation on the foundations of computation.” I gratefully acknowledge Roberto Giuntini for supervising this project.

This book has been developed from my doctoral dissertation. First of all, I would like to express all my gratitude to my former Ph.D. supervisor Marco Giunti for providing me illuminating ideas, suggestions, and support during the whole writing process.

Many people from the University of Cagliari deserve to be mentioned here. I am especially thankful to Francesco Paoli, who was a source of interesting ideas and suggestion about the psychology of mathematics. Giorgio Fumera has given me full technical support for the neural-net simulation I present in Chap. 5. I am also thankful to my colleagues Simonluca Pinna, Luisa Peruzzi, and Stefano Bonzio for all the help and support I received from them, and to everyone in the ALOPHIS group.

I am very thankful to Lorenzo Magnani, editor in chief of the SAPERE series, not only for his efforts and helpfulness during the review process, but also for his friendliness and warmth during a visit in Cagliari. I am also grateful to Leontina DiCecco, my editor at Springer, for her patience and for giving me many helpful suggestions.

I would especially like to thank Giuseppe Longo and an anonymous reviewer for their precious comments and advices.

Last, I would like to thank my family for all the support and encouragement that I always received from them.

Contents

Acronyms

AI	Artificial Intelligence
ANS	Approximate Number System
BTM	Bidimensional Turing Machine
CC	Classic Computationalism
CF	Computational Functionalism
CHM	Computational Hypothesis of Mind
CTM	Computational Theory of Mind
CTT	Church-Turing Thesis
DD	Developmental Dyscalculia
DH	Dynamical Hypothesis
EF	Ecological Functionalism
EMH	Extended Mind Hypothesis
HCTT	Human version of the Church-Turing Thesis
HOC	Hypothesis of Organism-Centered Cognition
IPS	Intra-Parietal Sulcus
MCTT	Mechanical Church-Turing Thesis
MNL	Mental Number Line
OTS	Object Tracking System
SDA	Single-Digit Addition
SNARC	Spatial Numerical Association of Response Codes
TM	Turing Machine
UTM	Universal Turing Machine
WM	Working Memory

Introduction

The fields of philosophy of mind and cognitive science have been characterized, in the last few decades, by a growing interest for explanations of mind's activity in terms of *interaction* between brains, bodies, and the world (Clark 1997; Varela and Thompson 1991). *Embodiment*, *embeddedness*, and *situatededness* are keywords that most often can be found in contemporary cognitive studies. This fact cannot be simply intended as a matter of fashion or a philosophical mannerism, for these concepts have arisen from a number of influential and somehow *revolutionary* studies which collaborated toward a change of shared philosophical and scientific views of the mind. Contributions have come from many different fields, such as robotics (e.g., Brooks 1991), neuropsychology (e.g., Damasio et al. 1991; Edelman 1987), linguistics (e.g., Lakoff and Johnson 1980), and developmental psychology (e.g., Thelen and Smith 1994).

However, some cognitive activities seem recalcitrant to this kind of treatment. Mathematical thinking is one of them. Explanations of human computational competencies, indeed, focus typically on *representational* issues, while underestimating or, at least, giving less importance to the role of mind/body/environment interaction in the development of *algorithmic skills*, namely those capacities which are essential in order to operate with numbers and carry out symbolic transformation.

The significance of these skills for a general understanding of computational activities is explicitly recognized in Alan Turing's theory of computation (Turing 1936), which is focused on the construction of idealized models of the mechanisms at work in a *real* cognitive system, namely the one consisting of *a human being performing calculations with paper and pencil*.

This kind of cognitive activity has then been recognized as a true example of *extended cognition*. Rumelhart et al. (1986), e.g., referred to the kinds of operations needed to carry out a long multiplication with the aid of paper and pencil as an example of *online symbolic transformations*, while Andy Clark (2008) proposed that a human being which performs such an activity instantiates a "transient extended cognitive system" (TECS):

> TECSs are soft-assembled (i.e., temporary and easily dissoluble) wholes that mesh the problem-solving contribution of human brain and central nervous system with those of the rest of the body and various elements of local "cognitive scaffolding" (Clark 2008, p. 158).

Turing's description of the basic operations at stake in the execution of an algorithm seems indeed in the same line of thought:

> Let us imagine the operations performed by the computer to be split up into "simple operations" which are so elementary that it is not easy to imagine them further divided. Every such operation consists of some change of the *physical system consisting of the computer and his tape*. We know the state of the system if we know the sequence of symbols on the tape, which of these are observed by the computer (possibly with a special order), and the state of mind of the computer (Turing 1936, p. 250) [emphasis added].

Here, Turing refers clearly to a *physical system* which not only consists of a *computer*—namely a human being involved in a computation—but comprises also *external* features that in the case of his computing machines are represented by symbols on the tape. The behavior of such a kind of physical system results from a strict interaction between internal (*mental*) and external (*environmental*) features, so that it is impossible to explain this behavior without referring to what happens to both kinds of features.

Despite these considerations, the *cognitive* importance of Turing's theory of computation has been so far underestimated.[3] Turing's work, indeed, has been primarily appreciated for its purely mathematical content—the formalization of the notion of *effective procedure*. As regards its specific cognitive content, it is instead widely held that the way computations are performed by a Turing Machine (TM) makes it a *psychologically implausible* model of a real cognitive system.

In this work, I endorse Andrew Wells' opinion, according to which a reinterpretation of the TM's architecture is needed, so as to restore Turing's original view and finally eliminate the misinterpretation originated with *classic computationalism*. By this term, I mean the view that computations are performed by a cognitive system *through internal symbolic transformations based on purely syntactic rules*. This idea lies behind many important cognitive theories, like Newell's Physical

---

[3]An important exception is represented by Andrew Hodges, who maintains that Turing's primary interest has ever been, since 1936, the study of the mind:

> The problem of mind is the key to 'Computable Numbers'. Somehow, it would appear, Turing sensed in the questions about definite, mechanical methods an opportunity to abstract and refine the notion of being determined, and apply this newly refined concept to the old question of mind. Somehow he perceived a link between what to anyone else would have appeared the quite unrelated questions of the foundations of mathematics, and the physical description of mind. The link was a scientific, rather than philosophical view; what he arrived at was a new materialism, a new level of description based on the idea of discrete states, and an argument that this level (rather than that of atoms and electrons, or indeed that of the physiology of brain tissue) was the correct one in which to couch the description of mental phenomena. It was to promoting and exploring this idea that he gave much of his subsequent life (Hodges 1995, p. 6).

Symbol Systems hypothesis (Newell 1980) and Fodor's Computational Theory of Mind (Fodor 1975).

Wells claims that the main mistake of classic computationalism in the interpretation of TM's architecture is treating the tape as an *internal memory*, rather than the *external environment*. By contrast, Wells considers the behavior of a TM as strictly dependent on the interaction between its internal and external parts (Wells 1998, 2005). In his view, the TM cannot be viewed as carrying out totally *internal* symbolic transformations, for the symbols written on the tape should be considered as objects in the environment.[4] Wells' alternative interpretation thus connects the long established notion of a TM to the aforementioned recent ideas in the philosophy of mind and cognitive science.

According to Wells, it would be possible to construct computational models of a wide range of cognitive functions, drawing inspiration from a TM design. However, Wells does not set any boundary to the cognitive functions which could be modeled in this way, and this is likely to be a major weakness of his proposal. An adequate strategy to make Wells' view more viable is trying to use his interpretation of the TM to explain cognitive phenomena of a specific type, and then extending the same kind of model to a wider range of cognitive tasks. This kind of strategy has been employed in the present book, in which I take seriously Marco Giunti's proposal to use a TM-based computational architecture, namely the Bidimensional Turing Machine (BTM) (Giunti 2009), in order to study human algorithmic skills.

This work consists of two main parts. The first part, philosophically oriented, deals with Wells' *ecological* interpretation of the TM's architecture and its relations with a set of philosophical and psychological positions such as classic computationalism, the Extended Mind Model and the Dynamical Approach to cognition (Chaps. 1–3); the second, more technical part, sets up a theoretical and methodological framework for the development and justification of BTM-based models of human algorithmic skills (Chaps. 4 and 5).

Outline of the Work

In Chap. 1, I describe the architecture and functioning of a TM taking a neutral stance about its interpretation. Then, I show how both historical and philosophical reasons have contributed to the widely accepted *classic computational* interpretation of the TM's architecture, which comes together with the claim of *psychological implausibility* deriving from this (mis)interpretation. In the last part of this chapter,

[4]A similar position has been defended by Wilson (1994). In this author's view, the cognitive system to which mental states and processes belong may be part of a *wide computational system*, namely "the corresponding computational system could transcend the boundary of the individual and include parts of that individual's environment" (Wilson 1994, p. 352).

I introduce the interpretation of Turing's theory of computation that I hold in this book, namely Andrew Wells' Ecological Functionalism (Wells 2005).

Chapter 2 focuses on the theoretical bases of Ecological Functionalism. First, I briefly discuss the main philosophical grounds of the Extended Mind Model, with special regard to the *parity principle* and to the concept of *active externalism*. Second, I explain the reasons why, on the one side, we should not worry (at least for the moment) about the ontological issues concerning the Extended Mind Model. On the other side, we should focus our attention on the pursuit of explaining the functioning and role of any kind of relevant features for a cognitive system, internal or external to the organism. Third, I discuss Well's arguments against the Computational Theory of Mind and the connectionist approach. Lastly, I present Well's original view of Turing machines as formal models for Gibson's concept of *affordance*.

In Chap. 3, I draw an overview of the main lines of research included in the so-called dynamical approach to cognition. Then, I elaborate on the original aspects of this approach, with a particular focus on the theoretical differences with classic computationalism and the analogies with Wells' ecological functionalism. Lastly, I present a dynamical interpretation of Turing machines' architecture. This interpretation is the starting point for the new kind of analysis of algorithmic skills that I present in this book.

In Chap. 4, first, I introduce a special kind of TM-based models of human computational skills, namely BTMs. Then, I define the concept of Galilean Model (Giunti 2010a, b), i.e. a special kind of *empirically adequate* cognitive model, and show how and why BTM-based models can be thought as possible Galilean Models of algorithmic skills.

Lastly, in Chap. 5, I employ an original method in order to verify (i) if some specific BTM-based models may be considered genuine Galilean Models of definite algorithmic skills and (ii) if it is possible to extrapolate from the analysis of such models of algorithmic skills some typical features of the performance and development of human computational skills. This last point is elaborated through the formulation of some specific hypotheses which may receive, at least in principle, empirical confirmation.

References

Brooks R (1991) Intelligence without representation. Artif Intel 47:139–159

Clark A (1997) Being there. Putting mind, brain and the world together again. MIT Press, Cambridge, MA

Clark A (2008) Supersizing the mind. Oxford University Press, New York

Damasio AR, Tranel D, Damasio H (1991) Somatic markers and the guidance of behavior: theory and preliminary testing. In: Levin HS, Eisenberg HM, Benton AL (eds) Frontal lobe function and dysfunction. Oxford University Press, New York, pp 217–229

Edelman GM (1987) Neural Darwinism: the theory of neuronal group selection. Basic Books, New York

Fodor J (1975) The language of thought. Harvard University Press, Cambridge, MA
Giunti M (2009) Bidimensional Turing machines as Galilean models of human computation. In: Minati G, Abram M, Pessa E (eds) Processes of emergence of systems and systemic properties. World Scientific, Cambridge, MA
Giunti M (2010a) Panorama e prospettive dell'approccio dinamico in scienza cognitiva. Logic Philos Sci (Online journal http://www2.unitsit/episteme/) 8:101–118
Giunti M (2010b) Reduction in dynamical systems. In: D'Agostino M, Giorello G, Laudisa F, Pievani T, Sinigaglia C (eds) SILFS New essays in logic and philosophy of science. College Publications, London
Hodges A (1995) Alan Turing and the Turing machine. In: Herken R (ed) The Universal Turing machine: a half-century survey, 2nd edn. Springer, pp 3–14
Lakoff G, Johnson M (1980) Metaphors we live by. Chicago University Press, Chicago
Newell A (1980) Physical symbol systems. Cogn Sci 4:135–183
Rumelhart DE, Smolensky P, McClelland JL, Hinton GE (1986) Schemata and sequential thought processes in PDP models. In: McClelland JL, Rumelhart DE and PDP Research Group (eds) Parallel distributed processing. Volume 2: psychological and biological models. MIT Press, Cambridge, MA, pp 7–57
Thelen E, Smith L (eds) (1994) A dynamic systems approach to the development of cognition and action. MIT Press, Cambridge, MA
Turing A (1936) On computable numbers, with an application to the *Entscheidungsproblem*. In: Proceedings of the London mathematical society. Oxford Journals, London, pp 230–265
Varela F, Thompson E (1991) the embodied mind: cognitive science and human experience. MIT Press, Cambridge, MA
Wells A (1998) Turing's analysis of computation and theories of cognitive architecture. Cogn Sci 22:269–294
Wells A (2005) Rethinking cognitive computation: Turing and the science of the mind. Palgrave Macmillan, Basingstoke, UK
Wilson RA (1994) Wide computationalism. Mind 103(411):351–372