ATLANTIS THINKING MACHINES

VOLUME 1

SERIES EDITOR: KAI-UWE KÜHNBERGER

Atlantis Thinking Machines

Series Editor:

Kai-Uwe Kühnberger

Institute of Cognitive Science
University of Osnabrück, Germany

(ISSN: 1877-3273)

Aims and scope of the series

This series publishes books resulting from theoretical research on and reproductions of general Artificial Intelligence (AI). The book series focusses on the establishment of new theories and paradigms in AI. At the same time, the series aims at exploring multiple scientific angles and methodologies, including results from research in cognitive science, neuroscience, theoretical and experimental AI, biology and from innovative interdisciplinary methodologies.

For more information on this series and our other book series, please visit our website at:

www.atlantis-press.com/publications/books

AMSTERDAM – PARIS

© ATLANTIS PRESS / WORLD SCIENTIFIC

Enaction, Embodiment, Evolutionary Robotics

Simulation Models for a Post-Cognitivist Science of Mind

Marieke Rohde

Multisensory Perception and Action Group
Max Planck Institute for Biological Cybernetics
Spemannstrasse 41
72076 Tübingen, Germany

With a preface by

Dr. Ezequiel A. Di Paolo

University of Sussex
Brighton, UK

**ATLANTIS
PRESS**

AMSTERDAM – PARIS

World Scientific

Atlantis Press

29, avenue Laumière
75019 Paris, France

For information on all Atlantis Press publications, visit our website at: www.atlantis-press.com

Copyright

This book is published under the Creative Commons Attribution-Non-commercial license, meaning that copying, distribution, transmitting and adapting the book is permitted, provided that this is done for non-commercial purposes and that the book is attributed.

In case of commercial purposes, this book, or any parts thereof may not be reproduced in any form or by any means, electronic or mechanical, including photocopying, recording or any information storage and retrieval system known or to be invented, without prior permission from the Publisher.

Atlantis Thinking Machines

ISBN: 978-90-78677-23-9
ISSN: 1877-3273

e-ISBN: 978-94-91216-34-3

For the people of Brighton and the CCNR

Preface

This is an unusual book. It launches a new style of research into the nature of the mind, a style that proficiently uncovers, explores and exploits the synergies between complex systems thinking, sophisticated theoretical critique, synthetic modeling technologies and experimental work. Rather than adopting a grandiose programmatic approach, Marieke Rohde presents us with a pragmatic conjunction of elements, each of them strongly feeding off the others and making it impossible to shelf her work strictly under any one rubric such as psychology, robotics, artificial intelligence or philosophy of mind. Perhaps the least unjust choice is to call this a work of *new cognitive science*.

It is yesterday's news to remark on how our conceptual framework for understanding complex systems is changing. There is a recognized need to supplement the scientific categories of mechanistic, XIX century thought for new ways of thinking about non-linear forms of interaction and inter-relation between events and processes at multiple scales. Since the times of cybernetics and in parallel to the development of the computer as a scientific tool, we have witnessed several proposals for "revolutionary" ways of dealing with complexity: catastrophe theory, general systems theory, chaos, self-organized criticality, complex networks, etc. Despite not always fulfilling their stated potential, these ideas have helped us increase our capability to understand complex systems and have in general left us with new concepts, new tools and new ways of formulating questions.

This conceptual change, however, has not been homogeneous. Only very recently have some of these ideas begun to make some way into mainstream theoretical biology (even if they were contained *in nuce* in the work of many pioneers) for example, in models of protocells and minimal metabolic systems, genetic regulatory networks, embryogenesis, immune networks, to evolutionary and ecosystems dynamics.

It is often the case that the sophisticated theoretical developments necessary to tackle a specific problem (or to reformulate it in a workable manner) have already existed for some

time, but only become acceptable once they are seen to work in the form of conceptual models, synthetic machines, or novel data.

If there is a contemporary domain of enquiry where complexity reigns supreme and where mono-disciplinary linear thinking is bound to fail, this is the realm of cognition. The future history of science will write that today we still know close to nothing about the mind. It will point out that essential categories such as autonomy, agency, values, meaning, intentionality, and many others remain poorly defined at the start of the XXIst century, and that we are, in ways still too preliminary, only beginning to grasp the complexity not only of brains, but of bodily physiologies and mechanisms, of experience, of structured and structuring environments and of social interactions. A proper study of the mind, or for that matter cognition, requires us to get a handle on biological (evolutionary, psycho-physical, neuroscientific), psychological, technological, socio-cultural, linguistic and experiential constraints. Mind is the realm of the über-complex. How to begin to think about it?

According to Rohde, this is how:

We need new concepts

The enactive framework that serves as the basis for Rohde's investigations attempts to explore the relations between life and mind. This approach, with roots in the work of Francisco Varela, has recently become a wellspring of novel conceptual developments, many of which are described and put to work in this book. As examples one could mention a deeper sense of embodiment as rooted in the autonomous organisation of a cognitive being and offering a route towards a naturalisation of normativity and meaning as well as workable novel concepts such as adaptivity and sense-making. Workable, and improvable, as there is no pretension that the last word on the subject has been spoken.

These ideas in combination with the tools of synthetic modelling and dynamical systems theory are the ingredients of the results presented here. The emerging picture is by no means a simple one. It reveals, on the contrary, the subtle complexities and interweaving of factors that are not always easy to isolate. This is perhaps just as it should be – to study, as Rohde does, real cognition. We should suspect any story that renders the complexity of the mind too easily graspable, because that is likely a sign that we have not really made the effort to change our ways of thinking. This is not a self-defeating position – we can indeed understand the mind and develop scientific methods, like the one presented in this book, to this end, but chances are we shall have to radically change our own minds in the process. If this change has not taken place yet, then any understanding will be a façade

that works simply by covering the difficulties with its own blind-spots. Any semblance of an explanation is suspect if all it does is to sweep complexity under the carpet of concepts designed to stop us from asking questions. Such is the case with traditional distinctions between content and vehicles, use of representational language, and widespread functionalist assumptions about the mind – all of which are explicitly or implicitly rejected in this work.

We need new tools

The combined use of dynamical systems ideas and synthetic approaches such as evolutionary robotics allows Rohde to create a micro-loop of scientific enquiry running within her laptop. Pre-conceptions are put to the test by a process capable of generating (under certain constraints) exemplars of the behaviours of interest without having to be too specific about the underlying mechanisms. The result is precisely a process of exploration of mechanisms, which is rendered possible by the use of artificial evolution, a method that is often less biased than our own engineering-laden approach towards designing systems. Thus, evolutionary robotics can be put to the service of generating novel proofs of concept, to question intuitions and overall to exercise our scientific mind and train it in understanding complex embodied and embedded systems. Such models are often deliberately simple thus seeking maximum conceptual impact. In the phrase of Randall Beer, one of the pioneers of this methodology, working with these models is a form of mental gymnastics.

The versatility of this modeling methodology is clearly demonstrated in this book, where it is put to work in conceptual models (querying the logical consistency of the idea of value systems), empirically inspired problems (studying the role of linear synergies in human pointing) and in a direct dialogue with empirical studies where evolutionary robotics models serve the role of hypotheses-generators as well as providing guidance in the dynamical analysis of data. The subtlety with which the present work demonstrates this methodological versatility should not be missed and is likely to set the standard for important forthcoming developments.

We need not shy away from the grand challenges

A common remark about the enactive, dynamical approach to cognition and also about evolutionary robotics is that these ideas and tools sit well at the lower levels of intelligence, where sensorimotor constraints dominate. In other words, they serve us well to understand the intelligence of insects, but we shall find these concepts and methods lacking when dealing with the complexities of the human mind.

This is a real challenge, although at the same time, one could wonder how well traditional approaches have really helped us understand the complexities of the human mind (presumably if they had been successful we should by now have been able to build an artificial person?). Instead, this book takes a practical approach to this question and in doing so it undermines the commonly held assumption that, in order to be useful, a scientific model must match the complexity of the target phenomena that needs to be explained. Rohde provides several clear demonstrations that this is not the case. That it is possible to learn about important factors affecting human cognitive performance (bodily synergies, social perception and time perception) with models that do not nearly match humans in complexity and still can provide us with clear insights about the problem. These models allow us to enter into a concrete dialogue with empirical and theoretical efforts.

We need interdisciplinary dialogue and cross-fertilisation

Rohde does a remarkable job at doing justice to the different disciplines involved in her research moving easily between domains like a Renaissance mind. At no point are concepts trivialised or assumed to map into each other unproblematically. At no point does she render any of the disciplines or methods redundant or secondary to others. A sense of integration, not of unification, and dialogue comes through and hopefully this book will be followed by similar efforts and similar collaborations. Some of the methodological connections she draws, for instance between evolutionary robotics modeling and psychological experiments, are uncommon and probably presented here for the first time.

These are Rohde's proposals for the new sciences of cognition. If they sound idealistic in this preface, this impression will be corrected when the reader is confronted with the practical thrust of her work. This is what gives her book a genuine chance of changing the way we study the mind.

It would be useful to put Rohde's proposal in context. The computational-representational view of the mind, with traditional AI as its theoretical core, has been the target of multiple criticisms during its 6-decade long history. Some of these criticisms have been very insightful and sometimes apparently devastating. On the table of conceptual disputes one wonders how the traditional perspective managed to outlive such attacks. Probably for different reasons (the strong support of funding bodies since the 1960s playing no small part). But importantly due to a genuine, *bona fide* scientific advantage of this perspective: it has been able to provide enough friction to drive a scientific process. Thanks to the computational-representational view of the mind it is possible for psychologists to for-

ulate clear hypotheses about human and animal perception, decision-making, learning, memory and problem-solving and perform clean elegant experiments to probe these hypotheses. Thanks to the computational-representational view of the mind it is possible for neuroscientists to spell out neural function as information processing and turn this idea into the guiding heuristic for experiments and models both at the level of the cellular and biochemical processes as well as on the level of brain anatomical organisation. Thanks to the computational-representational view of the mind cognitive science can define itself as the scientific programme of specifying the functional architecture of intelligence, building new formalisms and deriving their specific implications to be ultimately tested against empirical evidence and by construction in the form of artificially intelligent machinery. In short, thanks to the computational-representational view of the mind, it has been possible for science to move on.

This is what marks the computational-representational view of the mind as a fruitful scientific paradigm, not its theoretical soundness or its in/ability to deal with problems that cognitive scientists choose to ignore (when they arguably should not). Those of us who sustain that this perspective is flawed must still recognise its fruitfulness and be realistic about this oft-neglected fact: the success of a scientific paradigm is not solely judged at the court of logical consistency and empirical evidence. Its maturity, by definition, also lies in the fact that it provides the right set of ideas and tools to tackle the problems that it sets itself, never mind the critical stance that points to its blind-spots as also being genuine and relevant problems.

Consequently, it should not surprise us that the biggest shocks to the system have come less from well argued conceptual attacks than from real practical evidence of its limitations. The connectionist revolution-come-reform clearly demonstrated the benefits of breaking with a rigid understanding of functionality as expressed only in logical rules. It expanded the paradigm without overturning its central tenets. And this was achieved with the help of workable tools leading to better models of brain function, better match with empirical data at least in some domains, and novel technologies for AI and robotics as well as for wider applications.

The Brooksonian revolution in autonomous robotics ushered in a more radical break. It explicitly questioned the view of intelligence as complex information processing in the shape of world-modelling strategies. It showed how simple, loosely coupled systems not necessarily organised hierarchically but working in parallel, could achieve real world performance in ways that had been eluding traditional AI engineers. Conceptually, the field

of situated robotics also questioned the understanding of intelligence promoted by traditional cognitive science as human-level, cold reasoning. This achievement, accordingly, is the icing in the evolutionary cake that represents the history of natural intelligence, much of which is opportunistic, affective, embodied and dynamic. Natural intelligence (including much of human everyday intelligence) is fond of “cheap tricks” of local applicability and less prone to general problem solving. In a large part, the questioning of computational-representational views, which has emerged during the last decade, stemmed from this Brooksonian revolution. It grew from the ensuing questioning of concepts such as computation and representation and the probing of otherwise abstract ideas like embodiment and situatedness, which were now beginning to yield concrete and measurable results in the fields of autonomous and evolutionary robotics.

Clearly, it is factual, observable changes that fuel conceptual shifts, and as a benefit these changes re-signify existing conceptual criticisms, giving them new tools and techniques to drive a research programme, in other words, to provide an alternative, non-computational framework with the chance to become a new paradigm in its own right.

This is the process that we are witnessing today and which is likely to keep on developing over the next decade. The present book is an example of this process. Recently at a talk at the University of Sussex (in June, 2009), the philosopher of cognitive science Andy Clark was asked for his opinion about the future of cognitive science and philosophy of mind over the next 10 years. He said it would be something dominated very much like what goes on today under the name of enactivism, but “without the silly bits”. What better way to get rid of the “silly bits” (assuming there are any!) – and so enter into a new phase of scientific development – than to put these ideas to work and see what they can do? And what better way to start this process than with interdisciplinary research spanning conceptual critique, experimental work, target oriented and abstract modeling in an exemplary methodological dialogue? This is what this book is about – a foray into what’s to come in the sciences and technologies of the mind. This is what makes it an unusual book.

Ezequiel A. Di Paolo
University of Sussex
Brighton, UK

Acknowledgments

I would like to thank Miriam Kyselo and Kai-Uwe Kühnberger for comments on a draft of this book. Also, Ezequiel Di Paolo has been a great mentor and friend. Thanks to all my friends, family and colleagues for their support.

Marieke Rohde
Tübingen, October 2009

Contents

Preface	vii
Acknowledgments	xiii
List of Figures	xix
1. Introduction	1
2. Enactive Cognitive Science	9
2.1 The Rise and Fall of Traditional Cognitive Science	10
2.2 Alternative Paradigms	13
2.2.1 Connectionism	14
2.2.2 Dynamicism	15
2.2.3 Cybernetics, ALife, Behaviour Based Robotics	16
2.2.4 Minimal Representationalism and Extended Mind	17
2.2.5 Methodological Overlap, Ideology Worlds Apart	18
2.3 The Enactive Approach	19
2.3.1 Autonomy	20
2.3.2 Sense-Making	21
2.3.3 Emergence	22
2.3.4 Embodiment	23
2.3.5 Experience	23
2.3.6 The Roots	25
2.4 Challenges, Criticisms and Simulation Models	25
3. Methods and Methodology	29
3.1 The Scientist as Observing Subject	30
3.2 Dynamical Systems Theory	35
3.2.1 Definition	35
3.2.2 The Explanatory Role of DST	38
3.3 Simulation Models, Evolutionary Robotics and CTRNN Controllers	39
3.3.1 Evolutionary Robotics Simulations	39
3.3.2 Simulation Models as Scientific Tools	44
3.4 Sensory Substitution and Sensorimotor Recalibration	47
3.5 The Study of Experience	52
3.5.1 First and Second Person Methods to Study Experience	53

3.5.2	Perceptual Judgements as Second Person Method?	58
3.6	Combining Experimental, Experiential and Modelling Approaches	62
4.	Linear Synergies as a Principle in Motor Control	67
4.1	Motor Synergies	68
4.1.1	The Degree-of-Freedom Problem and Motor Synergies	68
4.1.2	Directional Pointing	70
4.2	Model	72
4.3	Results	75
4.3.1	Number of Degrees of Freedom	76
4.3.2	Forcing Linear Synergy	78
4.3.3	Evolved Synergies	79
4.4	Discussion	81
5.	An Exploration of Value System Architectures	85
5.1	Value Systems	86
5.1.1	Reductionist Approaches and Value System Architectures	86
5.1.2	Value System Simulation	88
5.2	Model	91
5.3	Results	94
5.3.1	Co-evolution of Light-Seeking and Fitness Estimation	94
5.3.2	A Caricature of ‘Value-Guided Learning’	97
5.4	Discussion	99
5.5	Enactive Sense Making, Value Generation, Meaning Construction	102
5.6	Conclusion	106
6.	Perceptual Crossing in One Dimension	109
6.1	Perceptual Crossing in a One-Dimensional Environment	110
6.2	Model	112
6.3	Results	113
6.4	Discussion	117
7.	Perceptual Crossing in Two Dimensions	123
7.1	Perceptual Crossing in a Two-Dimensional Environment	123
7.2	Model	125
7.3	Results	128
7.3.1	Evolvability	128
7.3.2	Behavioural Strategies Evolved	129
7.3.3	Two-Wheeled Agent	132
7.3.4	‘Euclidean’ Agent	134
7.3.5	Arm Agent	135
7.4	Discussion	137
8.	The Embodiment of Time	143
8.1	Newton Meets Descartes: The Classical Approach	144
8.2	Time and its Many Dimensions in our Mind	148
8.3	Phenomenology	149
8.4	The Construction of Time	152
8.5	Findings on Cognitive Concepts of Time	156

8.6	The Brain, the World and Time Perception	162
8.7	Time Experience	172
9.	An Experiment on Adaptation to Tactile Delays	175
9.1	Adaptation to Sensory Delays and the Experience of Simultaneity	175
9.2	Methods	179
9.3	Results	183
10.	Simulating the Experiment on Tactile Delays	187
10.1	Model	187
10.2	Results	189
10.2.1	Systematic Displacements	191
10.2.2	Stereotyped Trajectories	193
10.2.3	Velocity	195
10.3	Summary	196
10.4	Revisiting the Human Data	198
10.4.1	Systematic Displacements	199
10.4.2	Stereotyped Trajectories	201
10.4.3	Velocity	202
10.5	Discussion	203
11.	Perceived Simultaneity and Sensorimotor Latencies	207
11.1	Summary of the Results	207
11.2	The Sensorimotor Basis of Present-Time	209
12.	Outlook	217
12.1	Summary	217
12.2	Evolutionary Robotics Simulations for a Post-cognitivist Science of Mind	220
12.2.1	Reception in the Scientific Community	221
12.2.2	Representationalist Strongholds	222
12.2.3	Simulating Human Perceptual Behaviour	224
12.3	Conclusion	226
Appendix A	List of Abbreviations and Symbols	229
Bibliography		231
Author Index		241

List of Figures

3.1	Illustration of ascriptional judgements of autonomy based on naïve observation and scientific study of the generative mechanisms.	34
3.2	Illustration of the social dimension of scientific knowledge construction.	35
3.3	Illustration of the evolutionary cycle in ER.	40
3.4	Illustration of brain-body-environment interaction.	43
3.5	Illustration of interplay between disciplines in computationalism, neuropsychology and the approach proposed in this book.	63
4.1	Visualisation of the simulated arm and schematic diagram of the task.	72
4.2	Network diagrams for the unconstrained, modularised and forced synergy condition.	74
4.3	Number of parameters evolved.	75
4.4	Average number of starting positions reached in incremental evolution.	76
4.5	Squared difference in normalised performance as individual joints are paralysed or blocked in two- and three-dimensional conditions.	77
4.6	An example evolved RBFN for a forced synergy network for the three-dimensional condition.	79
4.7	Sum of squared deviation from linear synergy in CTRNN controllers and example strategies for forced synergy and CTRNN controllers.	80
5.1	An illustration of different views on values.	90
5.2	The controller of the agent that seeks light and estimates its distance from the light.	94
5.3	Successful light seeking behaviour (trajectory and fitness/sensorimotor values).	95
5.4	Light-avoiding behaviour of an agent after 50 generations of ‘value-guided learning’, trajectory and performance.	99

5.5	Life-cognition continuity and the scale of increasing mediacy.	100
6.1	Schematic diagram of the one-dimensional task environment.	112
6.2	Example behaviour evolved.	115
6.3	Trajectories and sensorimotor values of interaction with a fixed object and with the other (details).	116
7.1	Schematic diagram of the simulation environment and control network.	126
7.2	Schematic diagram of the different types of agents evolved.	127
7.3	Population fitness average $\bar{F}$ (mean and maximum from 10 evolutionary runs with and without delay) and fitness of best individual from best run.	129
7.4	Example evolution profiles for different agents and parameters.	130
7.5	Average of populations in which rhythmic behaviour was evolved and correlated fitness.	131
7.6	Example trajectory and sensorimotor diagram for the best wheeled agent evolved.	133
7.7	Example trajectory and sensorimotor diagram for the best Euclidean agent evolved.	134
7.8	Example trajectory and sensorimotor diagram for an arm agent that evolved a neural oscillator as central pattern generator.	136
7.9	Example trajectory and sensorimotor diagram for the best arm agent evolved.	137
7.10	Example trajectory and sensorimotor diagram for an arm agent evolved with proprioceptive feedback.	138
9.1	The Tactos tactile feedback platform.	180
9.2	The repetitive lateral displacement of rows of objects.	182
9.3	Participants' performance with and without the delay before and after training.	183
9.4	Example trajectories from one participant over the course of the experiment.	185
10.1	Evolutionary Robotics simulation model of the experiment on adaptation to delays.	188
10.2	Performance profile averaged over 9 evolutionary runs in an unperturbed condition as opposed to perturbation through scaling the velocity.	191
10.3	Trajectories for different agent starting positions across time, presentation of a single object.	192
10.4	Performance profile with the modified fitness function F'	193
10.5	Steady state velocities v^* for different I_1 for the analysed evolved agents.	194

10.6	Trajectories for different agent starting positions across time, presentation of a single object (reactive agent).	195
10.7	Change in systematic displacements from the object centre across the phases of the experiment.	200
10.8	Logarithm of the <i>MSE</i> from mean trajectories throughout the phases of the experiment.	202
10.9	Average velocity before making contact with the object to be caught across the phases of the experiment.	203
11.1	Illustration of ideas on the relation between temporal experience and sensorimotor loops from the observer's perspective.	213
11.2	Illustration of how adaptation to increased sensorimotor latencies may change experienced simultaneity.	214
12.1	Illustration of the interdisciplinary enactive framework proposed.	218
12.2	Illustration of the hermeneutic circle of understanding.	225