

Foreword

Frank E. Ritter, 30 Dec. 2015

This book describes the background and results from Jun Tani's work of over a decade of building robots that think and learn through interaction with the world. It has numerous useful and deep lessons for modelers developing and using symbolic, sub-symbolic, and hybrid architectures, so I am pleased to see it in the *Oxford Series on Cognitive Models and Architectures*. It is work that is in the spirit of Newell and Simon's (1975) theory of empirical exploration of computer science topics and their work on generation of behavior, and also takes Newell and Simon's and Feynman's motto of understanding through generation of behavior seriously. At the same time, this work extends the physical symbol hypothesis in a very useful way by suggesting by example that the symbols of human cognition need not be discrete symbols manually fed into computers (that we have often done in symbolic cognitive architectures), but can instead be composable neuro-dynamic structures arising through iterative learning of perceptual experience with the physical world.

Tani's work has explored some of the deep issues in embodied cognition, about how interaction with the environment happens, what this means for representation and learning, and how more complex behavior can be created or how it arises through more simple aspects. These lessons include insights about the role of interaction with the environment, consciousness and free will, and lessons about how to build neural-net architectures to drive behavior in robots.

The book starts with a review of the foundations of this work, including some of the philosophical foundations in this area (including the symbol grounding problem, phenomenology, and the role of time in thinking). It argues for a role of hierarchy in modeling cognition, and for modeling and understanding interaction with an external world. The book also notes that state space attractors can be a useful concept in understanding cognition, and, I would add, this could be a useful additional way to measure fit of a model to behavior. This review also reminds us of areas that current symbolic models have been uninformed by—I don't think that these topics have been so much ignored as much as put on a list for later work. These aspects are becoming more timely as Tani's work shows they can be. The review chapters make this book particularly useful as an advanced textbook, which Tani already uses it for.

Perhaps more importantly, in the second half of the book (Chapters 6 to 11) Tani describes lessons from his own work. This work argues that behavior is not always programmed or extant in a system, but that it can or often should arise in systems attempting to achieve homeostasis—that there are positions of stability in a mental representation (including modeling others, imitation), and that differences in knowledge between the levels can give rise to effects that might be seen to be a type of consciousness, a mental trace of what lower levels should do or are doing, or explanations of what they have done based on predictions of the agent's own behavior, a type of self-reflexive mental model. These results suggest that more models should model homeostasis and include more goals and knowledge about how to achieve it.

His work provides another way of representing and generating behavior. This way emphasizes the dynamic behavior of systems rather than the data structures used in more traditional

of insightful advice on this manuscript in regards to its contents, as well as for editing in English examining every page. I would like to commend and thank all members of my former laboratory at RIKEN as well as of the current one in Korean Advanced Institute of Science and Technology (KAIST) who, over the years, have contributed to the research described in this book. I am lucky to have many research friends with whom I can have in-depth discussions about shared interests. Takashi Ikegami has been one of the most inspiring. His stroke of genius and creative insights on the topics of life and the mind are irreplaceable. I admit that many of my research projects described in this book have been inspired by thoughtful discussions with him. Ichiro Tsuda provided me deep thoughts about possible roles of chaos in the brains. The late Joseph Goguen and late Francisco Varela generously offered me much advice about the links between neurodynamics and phenomenology. Karl Friston has provided me thoughtful advice in the research of our shared interests in many occasions. Michael Arbib offered insight into the concept of action primitives and mirror neuron modeling. He kindly read my early draft and sent it to Oxford University Press. I have been inspired by frequent discussions about developmental robotics with Minoru Asada and Yasuo Kuniyoshi. I would like to express my gratitude and appreciation to Masahiro Fujita, Toshitada Doi, and Mario Tokoro of Sony Corporation who kindly provided me with the chance to start my neurorobotics studies more than two decades ago in an elevator hall in a Sony building. I must thank Masao Ito and Shun-ichi Amari at RIKEN Brain Science Institute for their thoughtful advices to my research in general. And, I express my gratitude for Miki Sagara who prepared many figures. I am grateful to Frank Ritter as the Oxford series editor on cognitive models and architectures who kindly provided me advice and suggestions from micro details to macro levels of this manuscript during its development. The book could have not been completed in the present form without his input. I wish to thank my Oxford University Press editor, Joan Bossert for her cordial support and encouragement from the beginning. Finally, my biggest thanks go to my wife, Tomoko, my son Kentaro, and my mother Harumi. I could not have completed this book without their patient and loving support.

This book is dedicated to the memory of my father, Yougo Tani, who ignited my interest in science and engineering before he passed away in my childhood. Some additional resources such as robot videos can be found at <https://sites.google.com/site/tanioupbook/home>. Finally, this work was partially supported by the 2012 KAIST Settlement and Research of New Instructors Fund, titled “Nero-Robotics Experiments with Large Scale Brain Networks”.

Contents

Foreword by Frank E. Ritter

Preface

Part I On the Mind

1. Where do we begin with mind?
2. Cognitivism
 - 2.1 Composition and recursion in symbol systems
 - 2.2 Some cognitive models
 - 2.3 The symbol grounding problem
 - 2.4 Context
 - 2.5 Summary
3. Phenomenology
 - 3.1 Direct experience
 - 3.2 The subjective mind and objective world
 - 3.3 Time perception: How can the flow of subjective experiences be objectified?
 - 3.4 Being-in-the-world
 - 3.5 Embodiment for mind
 - 3.6 Stream of consciousness and freewill
 - 3.7 Summary
4. Introducing the brain and brain science
 - 4.1 Hierarchical brain mechanisms for visual recognition and action generation
 - 4.2 A new understanding of action generation and recognition in the brain
 - 4.3 How can intention arise spontaneously and be aware consciously?
 - 4.4 Deciding between the conflicting evidence
 - 4.5 Summary
5. Dynamical systems approach for embodied cognition
 - 5.1 Dynamical systems
 - 5.2 Gibsonian and Neo-Gibsonian approaches
 - 5.3 Behavior-based robotics
 - 5.4 Modeling the brain at different levels
 - 5.5 Neural network models
 - 5.6 Neurorobotics with the dynamical systems perspectives
 - 5.7 Summary

Part II Emergent minds: Findings from robotics experiments

6. New proposals
 - 6.1 Robots with subjective views
 - 6.2 Engineering subjective views into neurodynamic models
 - 6.3 The subjective mind and the objective world as an inseparable entity
7. Predictive learning about the world from actional consequences
 - 7.1 Development of compositionality: The symbol grounding problem
 - 7.2 Predictive dynamics and self-consciousness
 - 7.3 Summary
8. Mirroring action generation and recognition with articulating sensory-motor flow
 - 8.1 A mirror neuron model: RNNPB

- 8.2 Embedding multiple behaviors in distributed representation
- 8.3 Imitating others by recognizing their mental states
- 8.4 Binding language and actions
- 8.5 Summary
- 9. Development of functional hierarchy for action
 - 9.1 Self-organization of functional hierarchy in multiple timescales
 - 9.2 Robotics experiments on developmental training of complex actions
 - 9.3 Summary
- 10. Freewill for action
 - 10.1 Dynamic account of *spontaneous behaviors*
 - 10.2 Freewill, consciousness, and postdiction
 - 10.3 Summary
- 11. Conclusions
 - 11.1 Compositionality in cognitive mind
 - 11.2 Phenomenology
 - 11.3 Objective science and subjective experience
 - 11.4 Future directions
 - 11.5 Summary of the book

Glossary for abbreviations

References