

# **Theoretical Neuroscience**

## Computational Neuroscience

Terrence J. Sejnowski and Tomaso Poggio, editors

*Neural Nets in Electric Fish*, Walter Heiligenberg, 1991.

*The Computational Brain*, Patricia S. Churchland and Terrence J. Sejnowski, 1992

*Dynamic Biological Networks: The Stomatogastric Nervous System*, edited by Ronald M. Harris-Warrick, Eve Marder, Allen I. Selverston, and Maurice Moulins, 1992

*The Neurobiology of Neural Networks*, edited by Daniel Gardner, 1993

*Large-Scale Neuronal Theories of the Brain*, edited by Christof Koch and Joel L. Davis, 1994

*The Theoretical Foundations of Dendritic Function: Selected Papers of Wilfrid Rall with Commentaries*, edited by Idan Segev, John Rinzel, and Gordon M. Shepherd, 1995

*Models of Information Processing in the Basal Ganglia*, edited by James C. Houk, Joel L. Davis, and David G. Beiser, 1995

*Spikes: Exploring the Neural Code*, Fred Rieke, David Warland, Rob de Ruyter van Steveninck, and William Bialek, 1997

*Neurons, Networks, and Motor Behavior*, edited by Paul S.G. Stein, Sten Grillner, Allen I. Selverston, and Douglas G. Stuart, 1997

*Methods in Neuronal Modeling: From Ions to Networks*, second edition, edited by Christof Koch and Idan Segev, 1998

*Fundamentals of Neural Network Modeling: Neuropsychology and Cognitive Neuroscience*, edited by Randolph W. Parks, Daniel S. Levine, and Debra L. Long, 1998

*Neural Codes and Distributed Representations: Foundations of Neural Computation*, edited by Laurence Abbott and Terrence J. Sejnowski, 1998

*Unsupervised Learning: Foundations of Neural Computation*, edited by Geoffrey Hinton and Terrence J. Sejnowski, 1998

*Fast Oscillations in Cortical Circuits*, Roger D. Traub, John G.R. Jeffreys, and Miles A. Whittington, 1999

*Computational Vision: Information Processing in Perception and Visual Behavior*, Hanspeter A. Mallot, 2000

*Theoretical Neuroscience: Computational and Mathematical Modeling of Neural Systems*, Peter Dayan and L.F. Abbott, 2001

# **Theoretical Neuroscience**

Computational and Mathematical Modeling of  
Neural Systems

Peter Dayan and L.F. Abbott

The MIT Press  
Cambridge, Massachusetts  
London, England

First MIT Press paperback edition, 2005

© 2001 Massachusetts Institute of Technology

All rights reserved. No part of this book may be reproduced in any form by any electronic or mechanical means (including photocopying, recording, or information storage and retrieval) without permission in writing from the publisher.

MIT Press books may be purchased at special quantity discounts for business or sales promotional use. For information, please email [special\\_sales@mitpress.mit.edu](mailto:special_sales@mitpress.mit.edu) or write to Special Sales Department, The MIT Press, 55 Hayward Street, Cambridge, MA 02142.

Typeset in Palatino by the authors using  $\text{\LaTeX}2_{\epsilon}$ .

Printed and bound in the United States of America.

Library of Congress Cataloging-in-Publication Data

Dayan, Peter.

Theoretical neuroscience : computational and mathematical modeling of neural systems / Peter Dayan and L.F. Abbott.

p. cm. — (Computational neuroscience)

Includes bibliographical references.

ISBN 0-262-04199-5 (hc. : alk. paper) — 0-262-54185-8 (pb.)

1. Neural networks (Neurobiology) — Computer simulation. 2. Human information processing — Computer simulation. 3. Computational neuroscience.

I. Abbott, L.F. II. Title. III. Series

QP363.3 .D39 2001

573.8'01'13--dc21

2001044005

10 9 8 7 6 5 4

*To our families*



# Contents

<b>Preface</b>	<b>xiii</b>
<b>I Neural Encoding and Decoding</b>	<b>1</b>
<b>1 Neural Encoding I: Firing Rates and Spike Statistics</b>	<b>3</b>
1.1 Introduction	3
1.2 Spike Trains and Firing Rates	8
1.3 What Makes a Neuron Fire?	17
1.4 Spike-Train Statistics	24
1.5 The Neural Code	34
1.6 Chapter Summary	39
1.7 Appendices	40
1.8 Annotated Bibliography	43
<b>2 Neural Encoding II: Reverse Correlation and Visual Receptive Fields</b>	<b>45</b>
2.1 Introduction	45
2.2 Estimating Firing Rates	45
2.3 Introduction to the Early Visual System	51
2.4 Reverse-Correlation Methods: Simple Cells	60
2.5 Static Nonlinearities: Complex Cells	74
2.6 Receptive Fields in the Retina and LGN	77
2.7 Constructing V1 Receptive Fields	79
2.8 Chapter Summary	81
2.9 Appendices	81
2.10 Annotated Bibliography	84
<b>3 Neural Decoding</b>	<b>87</b>
3.1 Encoding and Decoding	87
3.2 Discrimination	89
3.3 Population Decoding	97
3.4 Spike-Train Decoding	113
3.5 Chapter Summary	118

3.6	Appendices	119
3.7	Annotated Bibliography	122
<b>4</b>	<b>Information Theory</b>	<b>123</b>
4.1	Entropy and Mutual Information	123
4.2	Information and Entropy Maximization	130
4.3	Entropy and Information for Spike Trains	145
4.4	Chapter Summary	149
4.5	Appendix	150
4.6	Annotated Bibliography	150
<b>II</b>	<b>Neurons and Neural Circuits</b>	<b>151</b>
<b>5</b>	<b>Model Neurons I: Neuroelectronics</b>	<b>153</b>
5.1	Introduction	153
5.2	Electrical Properties of Neurons	153
5.3	Single-Compartment Models	161
5.4	Integrate-and-Fire Models	162
5.5	Voltage-Dependent Conductances	166
5.6	The Hodgkin-Huxley Model	173
5.7	Modeling Channels	175
5.8	Synaptic Conductances	178
5.9	Synapses on Integrate-and-Fire Neurons	188
5.10	Chapter Summary	191
5.11	Appendices	191
5.12	Annotated Bibliography	193
<b>6</b>	<b>Model Neurons II: Conductances and Morphology</b>	<b>195</b>
6.1	Levels of Neuron Modeling	195
6.2	Conductance-Based Models	195
6.3	The Cable Equation	203
6.4	Multi-compartment Models	217
6.5	Chapter Summary	224
6.6	Appendices	224
6.7	Annotated Bibliography	228
<b>7</b>	<b>Network Models</b>	<b>229</b>
7.1	Introduction	229
7.2	Firing-Rate Models	231
7.3	Feedforward Networks	241



7.4	Recurrent Networks	244
7.5	Excitatory-Inhibitory Networks	265
7.6	Stochastic Networks	273
7.7	Chapter Summary	276
7.8	Appendix	276
7.9	Annotated Bibliography	277
<b>III</b>	<b>Adaptation and Learning</b>	<b>279</b>
<b>8</b>	<b>Plasticity and Learning</b>	<b>281</b>
8.1	Introduction	281
8.2	Synaptic Plasticity Rules	284
8.3	Unsupervised Learning	293
8.4	Supervised Learning	313
8.5	Chapter Summary	326
8.6	Appendix	327
8.7	Annotated Bibliography	328
<b>9</b>	<b>Classical Conditioning and Reinforcement Learning</b>	<b>331</b>
9.1	Introduction	331
9.2	Classical Conditioning	332
9.3	Static Action Choice	340
9.4	Sequential Action Choice	346
9.5	Chapter Summary	354
9.6	Appendix	355
9.7	Annotated Bibliography	357
<b>10</b>	<b>Representational Learning</b>	<b>359</b>
10.1	Introduction	359
10.2	Density Estimation	368
10.3	Causal Models for Density Estimation	373
10.4	Discussion	389
10.5	Chapter Summary	394
10.6	Appendix	395
10.7	Annotated Bibliography	396
	<b>Mathematical Appendix</b>	<b>399</b>
A.1	Linear Algebra	399
A.2	Finding Extrema and Lagrange Multipliers	408
A.3	Differential Equations	410

A.4 Electrical Circuits	413
A.5 Probability Theory	415
A.6 Annotated Bibliography	418
<b>References</b>	<b>419</b>
<b>Index</b>	<b>439</b>
<b>Exercises</b>	<a href="http://mitpress.mit.edu/dayan-abbott">http://mitpress.mit.edu/dayan-abbott</a>

## Series Foreword

Computational neuroscience is an approach to understanding the information content of neural signals by modeling the nervous system at many different structural scales, including the biophysical, the circuit, and the systems levels. Computer simulations of neurons and neural networks are complementary to traditional techniques in neuroscience. This book series welcomes contributions that link theoretical studies with experimental approaches to understanding information processing in the nervous system. Areas and topics of particular interest include biophysical mechanisms for computation in neurons, computer simulations of neural circuits, models of learning, representation of sensory information in neural networks, systems models of sensory-motor integration, and computational analysis of problems in biological sensing, motor control, and perception.

Terrence J. Sejnowski

Tomaso Poggio



## Preface

Theoretical analysis and computational modeling are important tools for characterizing what nervous systems do, determining how they function, and understanding why they operate in particular ways. Neuroscience encompasses approaches ranging from molecular and cellular studies to human psychophysics and psychology. Theoretical neuroscience encourages crosstalk among these subdisciplines by constructing compact representations of what has been learned, building bridges between different levels of description, and identifying unifying concepts and principles. In this book, we present the basic methods used for these purposes and discuss examples in which theoretical approaches have yielded insight into nervous system function.

The questions what, how, and why are addressed by descriptive, mechanistic, and interpretive models, each of which we discuss in the following chapters. Descriptive models summarize large amounts of experimental data compactly yet accurately, thereby characterizing what neurons and neural circuits do. These models may be based loosely on biophysical, anatomical, and physiological findings, but their primary purpose is to describe phenomena, not to explain them. Mechanistic models, on the other hand, address the question of how nervous systems operate on the basis of known anatomy, physiology, and circuitry. Such models often form a bridge between descriptive models couched at different levels. Interpretive models use computational and information-theoretic principles to explore the behavioral and cognitive significance of various aspects of nervous system function, addressing the question of why nervous systems operate as they do.

*descriptive models*

*mechanistic models*

*interpretive models*

It is often difficult to identify the appropriate level of modeling for a particular problem. A frequent mistake is to assume that a more detailed model is necessarily superior. Because models act as bridges between levels of understanding, they must be detailed enough to make contact with the lower level yet simple enough to provide clear results at the higher level.

## Organization and Approach

This book is organized into three parts on the basis of general themes. Part I, Neural Encoding and Decoding, (chapters 1–4) is devoted to the coding of information by action potentials and the representation of in-

formation by populations of neurons with selective responses. Modeling of neurons and neural circuits on the basis of cellular and synaptic biophysics is presented in part II, Neurons and Neural Circuits (chapters 5–7). The role of plasticity in development and learning is discussed in part III, Adaptation and Learning (chapters 8–10). With the exception of chapters 5 and 6, which jointly cover neuronal modeling, the chapters are largely independent and can be selected and ordered in a variety of ways for a one- or two-semester course at either the undergraduate or the graduate level.

### *background*

Although we provide some background material, readers without previous exposure to neuroscience should refer to a neuroscience textbook such as Kandel, Schwartz, & Jessell (2000); Nicholls, Martin, & Wallace (1992); Bear, Connors, & Paradiso (1996); Shepherd (1997); Zigmond et al. (1998); or Purves et al. (2000).

Theoretical neuroscience is based on the belief that methods of mathematics, physics, and computer science can provide important insights into nervous system function. Unfortunately, mathematics can sometimes seem more of an obstacle than an aid to understanding. We have not hesitated to employ the level of analysis needed to be precise and rigorous. At times, this may stretch the tolerance of some of our readers. We encourage such readers to consult the Mathematical Appendix, which provides a brief review of most of the mathematical methods used in the text, but also to persevere and attempt to understand the implications and consequences of a difficult derivation even if its steps are unclear.

### *exercises*

Theoretical neuroscience, like any skill, can be mastered only with practice. Exercises are provided for this purpose on the web site for this book, <http://mitpress.mit.edu/dayan-abbott>. We urge the reader to do them. In addition, it will be highly instructive for the reader to construct the models discussed in the text and explore their properties beyond what we have been able to do in the available space.

## Referencing

In order to maintain the flow of the text, we have kept citations within the chapters to a minimum. Each chapter ends with an annotated bibliography containing suggestions for further reading (which are denoted by a bold font), information about works cited within the chapter, and references to related studies. We concentrate on introducing the basic tools of computational neuroscience and discussing applications that we think best help the reader to understand and appreciate them. This means that a number of systems where computational approaches have been applied with significant success are not discussed. References given in the annotated bibliographies lead the reader toward such applications. Many people have contributed significantly to the areas we cover. The books and review articles in the annotated bibliographies provide more comprehensive references to work that we have failed to cite.

## Acknowledgments

We are extremely grateful to a large number of students at Brandeis, the Gatsby Computational Neuroscience Unit, and MIT, and colleagues at many institutions who have painstakingly read, commented on, and criticized numerous versions of all the chapters. We particularly thank Bard Ermentrout, Mark Goldman, John Hertz, Mark Kvale, Zhaoping Li, Eve Marder, and Read Montague for providing extensive discussion and advice on the entire book. A number of people read significant portions of the text and provided valuable comments, criticism, and insight: Bill Bialek, Pat Churchland, Nathaniel Daw, Dawei Dong, Peter Földiák, Fabrizio Gabbiani, Zoubin Ghahramani, Geoff Goodhill, David Heeger, Geoff Hinton, Ken Miller, Phil Nelson, Sacha Nelson, Bruno Olshausen, Mark Plumbley, Alex Pouget, Fred Rieke, John Rinzel, Emilio Salinas, Sebastian Seung, Mike Shadlen, Satinder Singh, Rich Sutton, Nick Swindale, Carl van Vreeswijk, Chris Williams, David Willshaw, Charlie Wilson, Angela Yu, and Rich Zemel.

We received significant additional assistance and advice from Greg DeAngelis, Andy Barto, Matt Beal, Sue Becker, Tony Bell, Paul Bressloff, Emery Brown, Matteo Carandini, Frances Chance, Yang Dan, Kenji Doya, Ed Erwin, John Fitzpatrick, David Foster, Marcus Frean, Ralph Freeman, Enrique Garibay, Frederico Girosi, Charlie Gross, Andreas Herz, Mike Jordan, Sham Kakade, Szabolcs Káli, Christof Koch, Simon Laughlin, John Lisman, Shawn Lockery, Guy Mayraz, Josh McDermott, Markus Meister, Earl Miller, Quaid Morris, Tony Movshon, Yuko Munakata, Randy O'Reilly, Simon Osindero, Tomaso Poggio, Clay Reid, Max Riesenhuber, Dario Ringach, Horacio Rotstein, Sam Roweis, Lana Rutherford, Ken Sugino, Alexei Samsonovich, Bob Shapley, Wolfram Schultz, Idan Segev, Terry Sejnowski, Jesper Sjöström, Haim Sompolinsky, Fiona Stevens, David Tank, Emo Todorov, Alessandro Treves, Gina Turrigiano, David Van Essen, Martin Wainwright, Xiao-Jing Wang, Chris Watkins, Max Welling, Jenny Whiting, Matt Wilson, Laurenz Wiskott, Danny Young, and Kechen Zhang. Thanks also to Quentin Huys, Philip Jonkers, Alexander Lerchner, Shih-Chii Liu, Máté Lengyel, Alex Loebel, Hadi Murr, Iain Murray, Jihwan Myung, John van Opstal, David Simon, Ed Snelson, and Rafael Yuste for pointing out errors in the text.

We thank Maneesh Sahani for advice and for indexing a substantial part of the text, Heidi Cartwright for creating the cover art, and Michael Rutter for his patience and consistent commitment. P.D. acknowledges the support of the Gatsby Charitable Foundation. Karen Abbott provided valuable help with the figures and with proofreading. Finally, we apologize to anyone we have inadvertently omitted from these lists.

