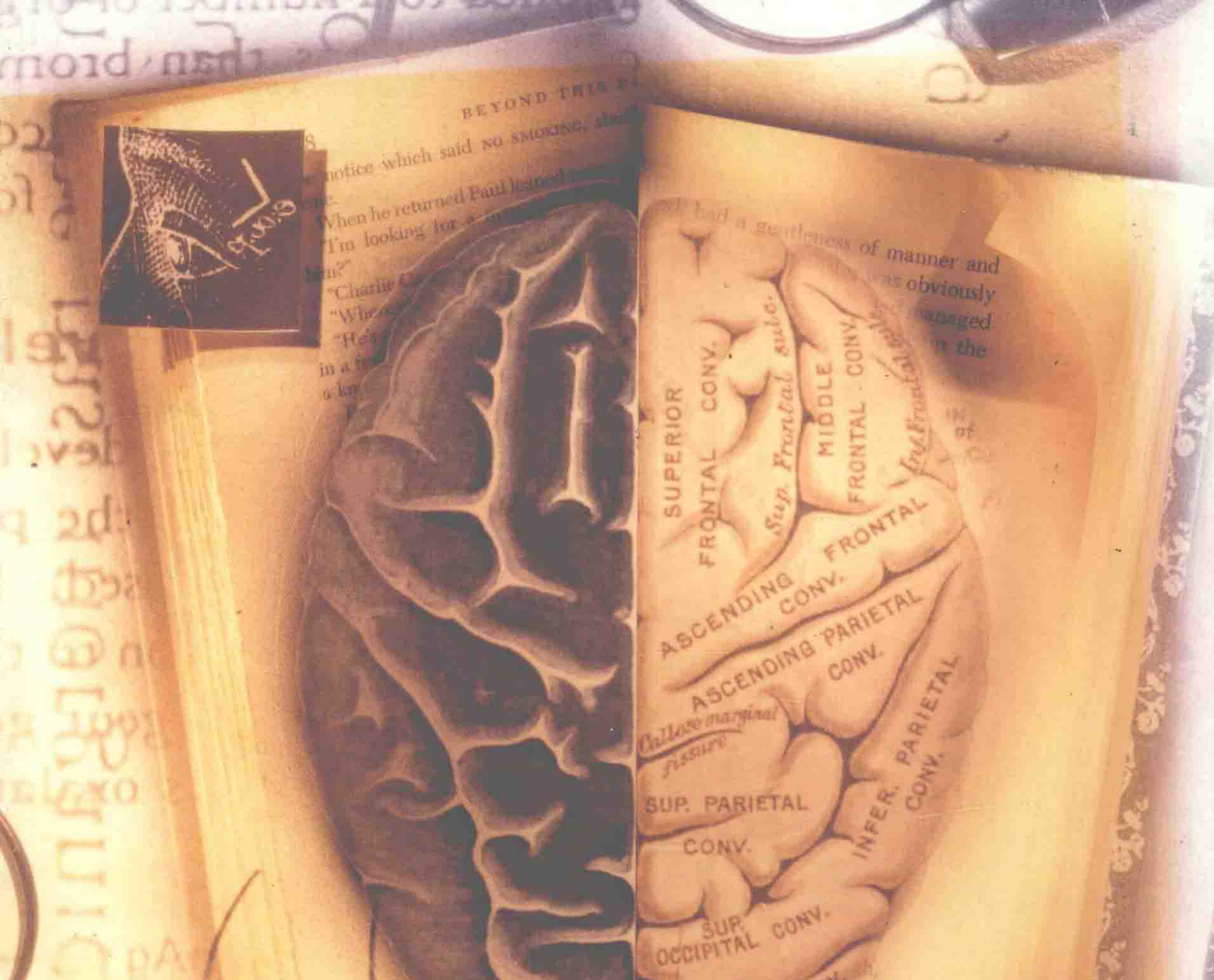

Fundamentals *of* Computational Neuroscience

Thomas P. Trappenberg

Foreword by John G. Taylor

Fundamentals of Computational Neuroscience

Thomas P. Trappenberg
Dalhousie University

OXFORD
UNIVERSITY PRESS

OXFORD
UNIVERSITY PRESS

Great Clarendon Street, Oxford OX2 6DP

Oxford University Press is a department of the University of Oxford.
It furthers the University's objective of excellence in research, scholarship,
and education by publishing worldwide in

Oxford New York

Auckland Bangkok Buenos Aires Cape Town Chennai
Dar es Salaam Delhi Hong Kong Istanbul Karachi Kolkata
Kuala Lumpur Madrid Melbourne Mexico City Mumbai Nairobi
São Paulo Shanghai Taipei Tokyo Toronto

and an associated company in Berlin

Oxford is a registered trade mark of Oxford University Press
in the UK and in certain other countries

Published in the United States
by Oxford University Press Inc., New York

First published 2002

A catalogue record for this book is available from the British Library

Library of Congress Cataloging in Publication Data

(Data available)

ISBN 0 19 851582 0 (Hbk)
ISBN 0 19 851583 9 (Pbk)

10 9 8 7 6 5 4 3 2 1

Typeset by the author
Printed in Great Britain
on acid-free paper by
T.J. International, Padstow, Cornwall

Fundamentals of Computational Neuroscience

In memory of my father
Rüdiger
dedicated to my children
Kai and **Nami**

Foreword

J.G. Taylor, King's College London

One of the main scientific adventures of this newly started century is that of understanding the brain, both of ourselves and of the animals we care for. New experimental results are pouring in on all sides, with most exciting data coming from the new brain imaging machines, as well as from advances in single cell recordings and from new measurement techniques only now being developed. All these new results have, however, to be understood in terms of some framework strong enough to bear their increasing weight. Without such an apparatus at our disposal we are in danger of not being able to see the wood for the trees, and even not seeing any wood at all. What should such a framework be constructed from in order to bear this increasing weight of experimental data?

Experiment can only progress effectively hand in hand with a developing theory. The two-way dialogue between them is strongly supported by Popper's 'falsification' approach to science: scientific understanding progresses by proposing an explanation of a range of phenomena which itself can be tested and is then modified so as to take account of the difficulties it meets when confronted by new experimental data. It is thus necessary to include a suitably strong theoretical component to this developing framework in order to understand the brain. It is also necessary to allow this theoretical component to develop by strong interaction with experiment.

It is this crucial need of providing a theoretical framework into which the burgeoning study of the brain can be fitted that is provided by the subject of Neural Networks. This discipline, now properly being recognized as such, allows the requisite theoretical ideas to be formulated and developed in order to make predictions testable against experiment in a falsifiable manner.

Neural networks have been developed for over 50 years, since their early description by McCulloch and Pitts to give a remarkable result as to the logical powers of the simplest form of neural network, composed of elegantly simple neurons. This result helped spark off a revolution in artificial intelligence, both in leading to the development of the digital computer and in the creation of an array of artificial neural networks for industrial applications. The subject of neural networks has now amply justified its continued existence in these areas. However, it now comes back to its roots when we turn back to use it, not in the important but still limited industrial applications I just mentioned. It is now necessary to attempt to create really intelligent machines, even up to seemingly conscious ones (if not actually so) in order to be able to build a theoretical framework which can talk in a useful dialogue to the data of neuroscience.

Here is the nub: how do we move neural networks towards the greater complexity of the brain? At the same time how can we develop them so they include both a non-conscious as well as a conscious component. The latter has so far been in the domain of artificial intelligence, working on a symbol system. The problem faced by neural networks is to create through them a powerful enough architecture to encompass both the non-conscious and conscious processing levels. This is not yet achievable, but progress is being made.

This book lays the foundations for such an approach. Its author has had considerable experience in applying neural networks to neuro-scientific modelling tasks. Yet he writes to emphasize the basic concepts involved in applying neural networks to the brain. As such this book is an essential component in the learning process for any who wish to join those trying to go from models of unconscious brain activity to those of a higher order, and who wish to complete their education on how lower level models can be constructed and understood.

Preface

Computational neuroscience is still a very young and dynamically developing discipline. Many different facets are explored by a growing number of scientists. It is beyond the scope of this book to review the research literature extensively, and I was not able to avoid some personal bias in the choice of topics included in this book. Some colleagues might disagree on the emphasis I have given specific topics, such as the detailed discussion on the information processing in networks of relatively simple processing devices. However, my aim in writing this book is to offer a trail through some of the exciting discoveries in this area, and to introduce some of the basic thoughts that guide the current thinking of many researchers about information processing in the brain. I believe that a basic understanding of many of the concepts outlined in this book is becoming mandatory for all neuroscientists, and I hope that I can help to build some foundations in theoretical techniques that are becoming important in neuroscience. I believe that these attempts will be useful for students and researchers who are interested in finding out 'how the brain computes'.

In writing this book I challenged myself in particular with the task of highlighting some reasons for the assumptions we make commonly in recent models; a discussion that is often omitted in the research literature. Furthermore, I intended to be brief in my description so that a reading of the entire book is possible in reasonable time (note that most of the formulas can be skipped in a first reading). At the same time I did not want to give the impression that the content of this book is the entire working knowledge of computational neuroscience. I did, therefore, not hesitate to mention some very specific terms (usually highlighted in italic), even without sufficient discussions, in order to provide the reader with starting points to search the common literature if further studies of such topics are desired. In line with the intent to provide a teaching book, I concentrated in the main text on describing ideas (rather than outlining a research trail). I have therefore avoided literature references in the text and have instead provided a short and commented 'Further reading' list at the end of each chapter. This list consists mainly of books that can provide further overviews and references for further detailed studies. A few references to some original research literature have been included when we followed such work closely.

I am aware of the fact that many readers interested in this subject may not be mathematically oriented, but I included mathematical formulas for several reasons. The most important is that formulas allow compact yet detailed specifications of models that can be translated directly into computer programs. It is, however, important to see beyond the formulas and to understand their meaning. There is no reason to be afraid of formulas; they only represent a shorthand notation that allows precision while avoiding lengthy verbal descriptions. Not all formulas have to be derived personally and can even be skipped in a first reading. Some formulas are included mainly for completeness and for reference if they should become important in your research. Please ask your instructor or mathematically inclined colleague to discuss with you any formulas that are not directly transparent to you. You don't have to be or become a mathematician to be able to understand the formulas in this book.

All chapters on different network architectures begin with an outline of some neuroscientific issues or some examples of information-processing principles that are aimed to motivate the study of the specific models in the corresponding chapter. The models in the book are, however, relatively general and are intended to illustrate basic information processing mechanisms in the brain. More specific models in the literature are often composed of such basic elements, and a study of the basic models in this book will enable the reader to follow some of the recent literature in this area. I have included only a few sample discussions of original research papers that highlight some of the points I wanted to make. I apologize for not mentioning many other research papers that have had major impacts in the field as well as papers that have guided much of my thinking.

This book is about neuroscience, and several issues mentioned in this book can be found in basic textbooks of neuroscience, in particular some basic descriptions of neurons and brain organizations. I want to encourage every reader who is not familiar with these subjects to consult these books, some examples of which I have included in the 'Further reading' list. In contrast to widely spread neuroscience textbooks, we concentrate in this book on the computational aspects, which are often not the major focus of other books. This book focuses on the modelling aspects, in particular the teaching of modelling, how to make useful abstractions, and how to use tools to analyse models. In contrast to some books on the technological aspects of neural networks I wanted to outline the relations of theoretical concepts or particular technological solutions to brain functions.

In Chapters 2 and 3 we focus our attention on single neurons. These chapters are very brief in comparison to the large amount of knowledge we have gathered in this area over the last decades. We start with a very short review of basic neuronal components and functionalities, mainly to remind us of some terms and basic facts that are well known and treated in the literature. We mention only briefly compartmental models of neurons as these are the subject of several excellent dedicated publications. We concentrate instead on simpler neuron models that are frequently used in network models. The reader familiar with the basic anatomy and functionality of neurons can skip most of this outline and instead concentrate on the models of single neurons in Chapter 3 that are more specific to studies in computational neuroscience.

In Chapter 4 we begin to explore networks of neurons. It includes the formal introduction of node models, which are often viewed as abstractions of single neurons, because I strongly believe that these models are much better motivated on a network level as outlined in the text. These node models will be essential to much of the discussion in the rest of the book. In Chapter 5 we discuss how information is represented and communicated within the nervous system, which motivates many aspects of the subsequent models.

Chapters 6–9 introduce in turn each one of the fundamental network models that we think are essential for information processing in the brain. These include mapping networks such as perceptrons, associators, auto-associators, and continuous attractor and competitive networks. We commonly discuss these models as examples in conjunction with specific brain-processing issues, although such models are found at the heart of many models in computational neuroscience. In Chapter 10 we follow up on some more specific issues on motor control, supervised learning, and reward learning, which are fundamental issues in computational neuroscience. Chapter 11 then discusses combinations of the fundamental network models more formally, which we have to consider when we want to understand information processing in the brain on a system level. Such modular networks include the mixture of experts, product of experts, and coupled auto-associator networks. In addition, we include in this chapter some examples of more specific system level issues on working memory and translation-invariant object recognition.

Most of the models discussed in this book can be implemented and simulated with small computer programs of only a few lines. We have included an introduction to programming such models within the MATLAB®[1] programming environment in Chapter 12. MATLAB is very convenient for neural network simulations and scientific visualization, and this tool can be mastered within a very short time. The hands-on guide to some of the models in this book is also aimed to further the understanding of such models and the formulas provided in the main text. Many of the simulation details are encapsulated in this chapter to keep the main text focused on the general issue. MATLAB provides a Neural Network Toolbox with ready-to-use routines of basic and advanced techniques in neural computation. We are not using this toolbox in our treatment as most of the basic algorithms, and even advanced research projects, can be programmed easily within the basic MATLAB programming language itself. This gives the greatest flexibility in creating our own solutions and in implementing novel ideas. Employing only the basic MATLAB programming environment will also make the working of neural networks most transparent and will teach us, at the same time, a useful tool.

[1] MATLAB is a registered trademark of The MathWorks, Inc.

Acknowledgements

Many colleagues and friends have contributed to this book through personal conversations, their enlightening research, and general guidance, too many indeed to be able to mention them all. Keeping with my theme to be brief I will thus only mention some colleagues who particularly have encouraged me to study computational neuroscience. On the theory side these were Shun-ichi Amari, Edmund Rolls, and John Taylor, and I have benefited much from conversations with Andrew Back, Gustavo Deco, Shiro Ikeda, Simon Stringer, and Masami Tatsuno. On the physiology front I have to thank in particular Doug Munoz and Okahide Hikosaka who gave me the opportunity to learn at first hand from brilliant researchers, as well as the members (and past members) of their active teams including Mike Dorris, Stefan Everling, Johan Lauwereyns, and Masamichi Sakagami. My interest in the brain would not have occurred without the cognitive scientists Raymond Klein and Patricia McMullen; Ray in particular has guided me for most of my way through neuroscience, and I want to thank him for his continuous support, guidance, and collaboration. I am very grateful to Justus Verhagen, Sue Becker, Masami Tatsuno, and Theodore Chiasson for reading through very rough drafts and pointing out many mistakes and unclear sections. Needless to say that I am responsible for all remaining mistakes. Last, but certainly not least, I want to thank my wife Kanayo for her support, without which the writing of this book would not have been possible.

Contents

1 Introduction

We will outline in this introductory chapter the major focus, specific tools, and the strategies of computational neuroscience. The complexity of the brain makes it necessary to clarify how we attempt to describe it and what we expect from explanatory models. We will argue that theoretical and computational studies are important in understanding experimental measurements. We will discuss the specific role of models in computational neuroscience and argue that they have to be chosen carefully if they are to be useful for advancing our knowledge of the information-processing principles in the brain. Models must not only be able to summarize known experimental data, but must also be able to make predictions that can be verified experimentally. We also include a short guide to the book to show the path we are going to take to explore this fascinating scientific subject.

1.1 What is computational neuroscience?

In the scientific area now commonly called computational neuroscience we utilize distinct techniques and ask specific questions aimed at advancing our understanding of the nervous system. This specific scientific area might be defined as:

> *Computational neuroscience is the theoretical study of the brain to uncover the principles and mechanisms that guide the development, organization, information processing, and mental abilities of the nervous system.*

Thus, computational neuroscience is a specialization within neuroscience. Neuroscience itself is a scientific area with many different aspects. Its aim is to understand the nervous system, in particular the central nervous system that we call the brain. The brain is studied by researchers who belong to diverse disciplines such as physiology, psychology, medicine, computer science, and mathematics, to name but a few. Neuroscience emerged from the realization that interdisciplinary studies are vital to further our understanding of the brain. The brain is one of the most complex systems ever encountered in nature, and there are still many questions that we can only attack through combined efforts. How does the brain work? What are the biological mechanisms involved? How is it organized? What are the information-processing principles used to solve complex tasks such as perception? How did it evolve? How does it change during the lifetime of the organisms? What is the effect of damage to particular areas and the possibilities of rehabilitation? What are the origins of degenerative diseases and possible treatments? All these basic questions are asked by neuroscientists.

1.1.1 The tools and specializations in neuroscience

Many techniques are employed within neuroscience to answer these questions. These include genetic manipulation, *in vivo* and *in vitro* recording of cell activities, optical imaging, functional magnetoresonance scanning, psychophysical measurement, and computer simulation. Each of these techniques is complicated and laborious enough to justify a specialization within neuroscience. We therefore speak of neurophysiologists, cognitive scientists, and anatomists. It is, however, vital for any neuroscientist to develop a basic knowledge of all major techniques in order to understand and use the contributions made by them. The significance of any technique has to be evaluated with a view to its specific problems and limitations as well as its specific aim of the technique. Computational neuroscience is an increasingly important area of neuroscience and a basic understanding of this field has become essential for all neuroscientists.

1.1.2 The focus of computational neuroscience

Computational neuroscience attempts to develop and test hypotheses about the functional mechanisms of the brain. A major focus is therefore the development and evaluation of *models*. The scientific area that is the subject of this book is also known as 'theoretical neuroscience'. Computational neuroscience can be viewed as a specialization within theoretical neuroscience, which employs computers to simulate models. The major reason for using computers is the complexity of many models, which are often analytically intractable. For such models we have to employ carefully designed numerical experiments in order to be able to compare these models to experimental data. However, we do not want to restrict our studies to this tool because analytical studies can often give us a deeper and more controlled insight into the features of models and the reasons behind numerical findings. Whenever possible, we try to include analytical techniques, and we will use the term 'computational neuroscience' synonymously with theoretical neuroscience. The word 'computational' also emphasizes that we are interested in particular in the computational aspects of brain functions.

We have included some examples of analytical techniques in order to give you a taste of some of those powerful techniques. Not every neuroscientist has to perform such calculations, but it is necessary to comprehend the general ideas if you are to get support from specialists in these techniques when required in your own research. However, it is very instructive to perform some numerical experiments yourself. We therefore included an introduction to a modern programming environment that is very well suited to many models in neuroscience. Writing programs and creating advanced graphics can be easily learned within a short time even without extensive computer knowledge.

Although computational neuroscience is theoretical by nature, it is important to bear in mind that the models have to be measured against experimental data; they are otherwise useless for understanding the brain. Only experimental measurements on the real brain can verify 'what' the brain does. In contrast to the experimental domain, computational neuroscience tries to speculate 'how' the brain does it. The speculations are developed into hypotheses, realized into models, evaluated analytically or numerically, and tested against experimental data. We will discuss specific examples of models on several levels (and related evaluation techniques) throughout this book.