This book is devoted to some mathematical methods that arise in two domains of artificial intelligence: neural networks and qualitative physics. The rapid advances in these two areas have left unanswered several mathematical questions that should motivate and challenge mathematicians.

Professor Aubin makes use of control and viability theory in neural networks and cognitive systems, regarded as dynamical systems controlled by synaptic matrices, and set-valued analysis that plays a natural and crucial role in qualitative analysis and simulation. This allows many examples of neural networks to be presented in a unified way. In addition, several results on the control of linear and nonlinear systems are used to obtain a "learning algorithm" of pattern classification problems, such as the back-propagation formula, as well as learning algorithms of feedback regulation laws of solutions to control systems subject to state constraints.

Mathematical models involve many features of a problem that may not be relevant to its solution. Qualitative physics, however, deals with an imperfect knowledge of the problem model. It is therefore more suited to the study of expert systems, which are shallow models and do not require structural knowledge of the problem.

This book should be a valuable introduction to the field for researchers in neural networks and cognitive systems, and should help to expand the range of study for viability theorists.

Cambridge University Press
978-0-521-44532-0 - Neural Networks and Qualitative Physics
Jean-Pierre Aubin
Frontmatter
More information

Neural Networks and Qualitative Physics

Cambridge University Press
978-0-521-44532-0 - Neural Networks and Qualitative Physics
Jean-Pierre Aubin
Frontmatter
More information

Neural Networks and Qualitative Physics

JEAN-PIERRE AUBIN

Ecole Doctorale de Mathématiques de la Décision
Université de Paris - Dauphine

CAMBRIDGE
UNIVERSITY PRESS

Published by the Press Syndicate of the University of Cambridge
The Pitt Building, Trumpington Street, Cambridge CB2 1RP
40 West 20th Street, New York, NY 10011–4211, USA
10 Stamford Road, Oakleigh, Melbourne 3166, Australia

First Published 1996

Printed in the United States of America

Library of Congress Cataloging-in-Publication Data available

A catalog record for this book is available from the British Library.

ISBN 0–521–44532–9 Hardback

Contents

Cambridge University Press
978-0-521-44532-0 - Neural Networks and Qualitative Physics
Jean-Pierre Aubin
Frontmatter
More information

Preface

This book is devoted to some **mathematical methods** that arise in two domains of artificial intelligence: neural networks and qualitative physics (which here we shall call "qualitative analysis"). These two topics are treated independently. Rapid advances in these two areas have left unanswered many mathematical questions that should motivate and challenge a wide range of mathematicians. The mathematical techniques that I choose to present in this book are as follows:

control and viability theory in neural networks and cognitive systems, regarded as dynamical systems controlled by synaptic matrices.

set-valued analysis, which plays a natural and crucial role in qualitative analysis and simulation by emphasizing properties common to a class of problems, data, and solutions. Set-valued analysis also underlies *mathematical morphology*,[1] which provides useful techniques for image recognition.

This allows us to present in a unified way many examples of neural networks and to use several results on the control of linear and nonlinear systems to obtain a *learning algorithm* of pattern-classification problems (including time series in forecasting), such as the *back-propagation formula*, in addition to learning algorithms concerning feedback-regulation laws for solutions to control systems subject to state constraints (inverse dynamics).

These mathematical techniques may also serve to contribute to the various attempts to devise mathematical metaphors for cognitive pro-

[1]See the forthcoming book by Michel Schmitt and Luc Vincent, *Morphological Image Analysis* (Cambridge University Press). The links between mathematical morphology and set-valued analysis and viability theory will be explored in a subsequent book, *Mutational and Morphological Analysis: tools for shape regulation and optimization.*

cesses. I present here a very speculative metaphor under the name of *cognitive systems* based on these mathematical techniques. They go beyond neural networks in the sense that they involve the problem of adaptation to viability constraints. They can *recognize* the state of the environment and *act* on the environment to *adapt* to given viability constraints. Instead of encoding knowledge in synaptic matrices as neural networks do, the knowledge is stored in *conceptual controls*. Given the mechanism of recognition of the state of the environment by conceptual controls, perception and action laws, and viability constraints, the viability theorems allow one to construct *learning rules* that describe how conceptual controls evolve in terms of sensorimotor states to adapt to viability constraints.

There is always a combination of two basic motivations for dealing with formal models of cognition - *neural networks* being content with implementation of "neural-like" systems on computers, and *cognitive systems* attempting to model actual biological nervous systems. Every model lies between these two requirements - the first allowing more freedom in the choice of a particular representation (computing efficiency being the main criterion), and the second constraining the modeling to be closer to biological reality.[2]

The symbolic processing capabilities that neural networks try to achieve are unexpected technological consequences of the use of digital computers, which were not designed for such a purpose at their inception. In the same way, the skills of reasoning logically and solving mathematical problems also represent a kind of unexpected "technological fallout" of the human brain, because they certainly were not among the advantages necessary for the survival of the human species when they appeared.

Expert systems are *shallow models* that do not require any formal and structural knowledge of the problem, whereas a mathematical model might involve too many features that would not be relevant for solving the problem at hand. For many problems, we have only imperfect knowledge of the model, and we may be interested in only a few features (often of a qualitative nature) of the solution, and so we see at once that the concept of *partial knowledge* involves two types of ideas:

[2] Actually, we should say "degree of reality for a social group at a given time," which is understood here in terms of the consensus interpretations of the group members' perceptions of their physical, biological, social, and cultural environments. This concept of reality is thus relative to a social group and is subject to evolution.

1. We require less precision in the results (e.g., signs of the components of vectors instead of their numerical values),
2. We take into account a broader universality or robustness of these results with respect to uncertainty, disturbances, and lack of precision.

As in numerical analysis, which deals both with approximation of problems in infinite-dimensional spaces by problems in finite-dimensional spaces and with the algorithms for solving such approximated problems, the problems of qualitative analysis arise at two levels: the passage from quantitative analysis to qualitative analysis (which deals with the association of discrete problems with continuous problems) and the algorithms to solve discrete problems.[3] In particular, Kuipers's QSIM algorithm for tracking the monotonicity properties of solutions to differential equations is revisited by placing it in a rigorous mathematical framework. This allows us to determine a priori the *landmarks* (i.e., the states at which the monotonicity properties change) instead of discovering them a posteriori by tracking the qualitative evolution of the solutions to the differential equation. These landmarks delineate *qualitative cells*, in which the monotonicity behaviors of the solutions are the same. Once these qualitative cells are computed, the Dordan QSIM algorithm provides the transition laws from one qualitative cell to the others.

This book is divided into 10 chapters. Chapters 1-7 deal with neural networks and some mathematical background needed to treat them (pseudoinverses, tensor products, gradient methods for convex potentials), Chapter 8 deals with cognitive systems, and Chapters 9 and 10 deal with some mathematical questions raised by qualitative physics, in the static and dymanic cases, respectively.

Chapter 1 provides the definitions of neural networks and learning processes (including the perceptron algorithm) and the *heavy learning algorithm*, which allows learning without forgetting.

Chapter 2 deals with some mathematical tools: pseudoinverses of linear operators and tensor products. Indeed, we have to use the specific structure of the space of synaptic matrices as a tensor product to justify mathematically the *connectionist features* of neural networks. Tensor products *explain* the Hebbian nature of many learning algorithms. This is due to the fact that derivatives of a wide class of nonlinear maps de-

[3]The first aspect has been quite neglected, and it is the one we shall emphasize in this book.

fined on spaces of synaptic matrices are tensor products and also to the fact that the pseudoinverse of a tensor product of linear operators is the tensor product of their pseudoinverses.

Chapter 3 is devoted to the case of linear neural networks, also called *associative memories.* We begin by showing that the *heavy learning algorithm* for neural networks that are affine with respect to the synaptic matrices (but nonlinear with respect to the signals) has a Hebbian character. We proceed with purely linear networks with a single layer or with a finite number or a continuum of layers. The chapter ends with an introduction to associative memories with gates, which are well adapted to compute Boolean and fuzzy Boolean functions.

Chapter 4 is devoted to the proof of the convergence of the gradient method for minimization problems involving a convex criterion with or without constraints. We discuss an application to the Minover algorithm of Mézard that replaces the perceptron algorithm. Many more features of convex analysis could be used in the study of a class of neural networks, but such results would go beyond the scope of this book and the common knowledge of its expected audience.

Chapter 5 adapts these results to the case of nonlinear networks and presents two main types of learning rules. The first class consists of algorithms derived from the gradient method and includes in particular the back-propagation rule. The second class is composed of learning rules based on the Newton method.

Chapter 6 is devoted to the use of neural networks for finding viable solutions to control systems, that is, solutions to control systems that will satisfy given viability (or state) constraints. The purpose of this chapter is to derive learning processes for *regulation feedback* for control problems through neural networks. Two classes of learning rules are presented. The first, called the class of *external learning rules*, is based on the gradient method (of optimization problems involving nonsmooth functions). The second deals with *uniform algorithms*.

In Chapter 7, the *internal-learning algorithm* provides learning rules based on viability theory. Two sections are devoted to a short presentation of the main results of viability theory and its application to the regulation of viable solutions to control systems. Applications to the control of cart-pole problems and other benchmark problems have been designed by N. Seube. This algorithm is applied to stabilization problems.

Chapter 8 goes beyond neural networks as they are usually defined. It proposes a very speculative mathematical model of what is called a *cognitive system*. A cognitive system is a dynamical system describing the evolution of sensorimotor states, recognized and controlled by *conceptual controls*, according to perception and action laws, and required to obey some viability constraints. *Adaptive learning processes* associating conceptual controls with sensorimotor states are then obtained by using viability theorems, including the ones that obey the *inertia principle: Change the conceptual controls only when the viability of the cognitive system is at stake.* This chapter is oriented toward mathematical metaphors motivated by cognisciences, of which we present a few relevant facts.

Chapter 9 treats the qualitative resolution of static problems described in the form of both equations and inclusions. It proposes a general framework (*confluence frames*) to link quantitative problems with qualitative ones. In particular, sign confluences are thoroughly investigated.

Chapter 10 is devoted to qualitative simulation of differential equations and to a mathematical treatment of Kuipers's QSIM algorithm to track the monotonicity properties of solutions to differential equations. We also consider Dordan's QSIM algorithm, which provides the qualitative cells delineated by the landmarks, and then the transition map associating with each qualitative cell its successor(s). Dordan's QSIM algorithm was designed to study the qualitative behaviors of a class of differential systems, the *replicator systems*, which play important roles in several domains of biology and biochemistry. We consider several examples obtained by using software designed by O. Dordan.

Two appendixes conclude this book. Appendix A provides a survey of convex optimization and set-valued analysis that goes beyond the minimal survey of Chapter 4. Appendix B describes applications of Nicolas Seube's algorithms, presented in Chapters 6 and 7, to the control of autonomous underwater vehicles (AUVs) tracking the trajectory of an exosystem.

Acknowledgements

I thank Alain Bensoussan for asking me to report to INRIA about neural networks and qualitative physics. That triggered a deeper interest in this topic.

This book owes much to Olivier Dordan and Nicolas Seube, who developed most of the material of this book and the computer applications. I have the pleasure to express here my gratitude. I also thank Luc Doyen and Juliette Mattioli for correcting parts of the text and for developing other applications of set-valued analysis and viability theory to visual control and mathematical morphology.