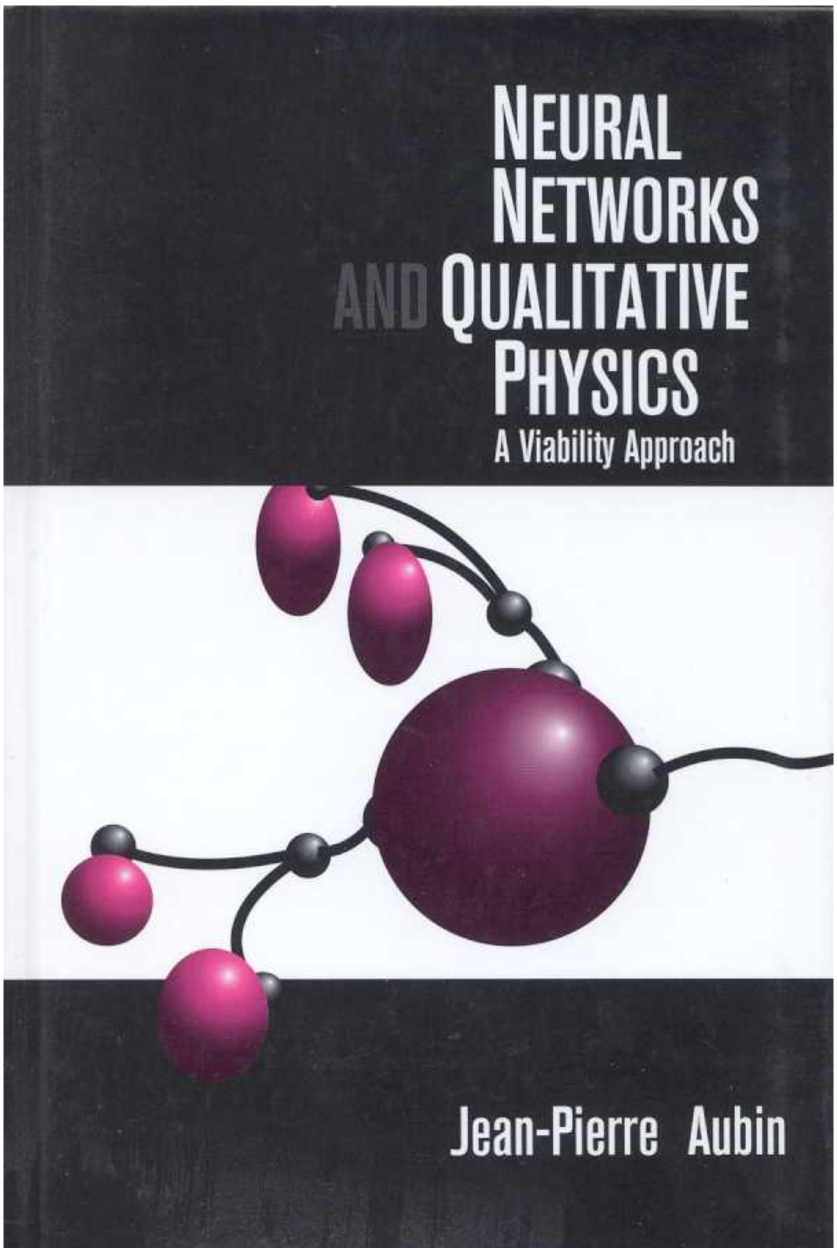
NEURAL
NETWORKS
AND QUALITATIVE
PHYSICS
A Viability Approach
Jean-Pierre Aubin

This book is dedicated to
Philippe Boutry

Introduction

This book is devoted to some mathematical methods arising in two domains of Artificial Intelligence: Neural Networks and Qualitative Physics[1].

These two topics are treated independently.

Rapid advances in these two areas left unanswered many mathematical questions which should motivate and challenge mathematicians.

The mathematical techniques we chose to present in this book are

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 class of problems, data and solutions. Set-Valued analysis underlies also mathematical morphology[2] which provides useful techniques for images recognition.

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

These mathematical techniques may also be efficient to contribute to the various attempts to devise mathematical metaphors of cognitive processes. We present here a very speculative one 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 to construct learning rules which describe how conceptual controls evolve in terms of sensory-motor states to adapt to viability constraints.

[1] that we shall call here Qualitative Analysis.

[2] See the forthcoming book Morphological Image Analysis at Cambridge University Press by Michel Schmitt & Luc Vincent. The links between mathematical morphology and set-valued analysis and viability theory shall be exposed in a forthcoming book.

There is always a combination of two basic motivations for dealing with formal models of cognisciences, Neural Networks being content with implementation of "neural-like" systems on computers, Cognitive Systems attempting to model actual biological nervous systems. Every model lies between these two requirements, the first one allowing more freedom in the choice of a particular representation, computing efficiency being the main criterion, the second one constraining the modeling to be closer to biological reality[3].

Symbolic processing capabilities that Neural Networks try to achieve are unexpected technological consequences of digital computers that were not designed for such purpose at their inception. In the same way, the skills of logical reasoning and solving mathematical problems are also some kind of unexpected "technological fallouts" of the human brain, since they certainly did not belong to the relevant advantages necessary to the survival of the human species when they appeared.

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

1. require less precision in the results (for instance, signs of the components of vectors instead of their numerical values),

2. take into account a broader universality or robustness of these results with respect to uncertainty, disturbances and lack of precisions.

As in numerical analysis, which deals both with approximation of problems in infinite dimensional spaces by problems in finite-dimensional spaces and of the algorithms for solving such approximated problems, 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[4]. In particular, Kuipers' QSIM algorithm for tracking the

[3]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 member's perceptions of their physical, biological, social and cultural environments. This concept of reality is thus relative to a social group and subject to evolution.

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

monotonicity properties of solutions to differential equations is revisited by placing it in a rigorous mathematical framework. This allows 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 behavior of the solutions is the same. Once these qualitative cells computed, the Dordan QSIM algorithm provides the transition laws from one qualitative cell to the others.

Ten chapters are presented. The seven first ones deal with neural networks and some mathematical background needed to treat them (pseudo-inverses, tensor products, gradient methods for convex potentials), the eighth one to cognitive systems and the two last ones deal with mathematical questions raised by Qualitative Physics, in the static and dynamical respectively.

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

Chapter 2 deals with some mathematical tools, pseudo-inverses 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 defined on spaces of synaptic matrices are tensor products and also, to the fact that the pseudo-inverse of a tensor product of linear operators is the tensor product of their pseudo-inverses.

Chapter 3 is devoted to the case of linear neural networks, also called associative memories. We begin to show that the Heavy Learning Algorithm to neural networks which are affine with respect to the synaptic matrices (but nonlinear with respect to the signals) has an Hebbian character. We proceed with purely linear networks, with a single layer, or with a finite number or a continuum of layers. This 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 convex criterion with or without constraints. We provide an application to the Minover Algorithm of Mzard which replaces the perceptron algorithm. Much more features of convex analysis could be used in the study of a class of neural networks, but these results go beyond the scope of this book and the common knowledge of its expected audience.

Chapter 5 adapt these results to the case of nonlinear networks and present 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 made of learning rules based on the Newton method.

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

In Chapter 7, internal learning rule 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 bench-mark 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 sensory-motor 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 sensori-motor states are then obtained by using viability theorems, and among them, the ones which obey the inertia principle: change the conceptual controls only when the viability of the cognitive system is at stakes. This chapter is more oriented towards mathematical metaphors motivated by cognisciences, of which we present few relevant facts.

Chapter 9 treats the qualitative resolution of static problems described both in the form of 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' QSIM algorithm to track the monotonicity properties of solutions to differential equations. We present 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 has been designed to study the qualitative behavior of a class of differential systems, the replicator systems, which play an important role in several domains of biology and biochemistry. It presents several examples obtained by using a software designed by O. Dordan.

Two appendices are presented to conclude this monograph.

The first one provides a survey of convex optimization and set-valued analysis which goes beyond the minimal survey of Chapter 4.

The second exposes the applications of Nicolas Seube's algorithms presented in Chapters 6 and 7 to the control of Autonomous Underwater Vehicles (AUV) tracking the trajectory of an exosystem.

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.

Jean-Pierre Aubin

September 18, 1992