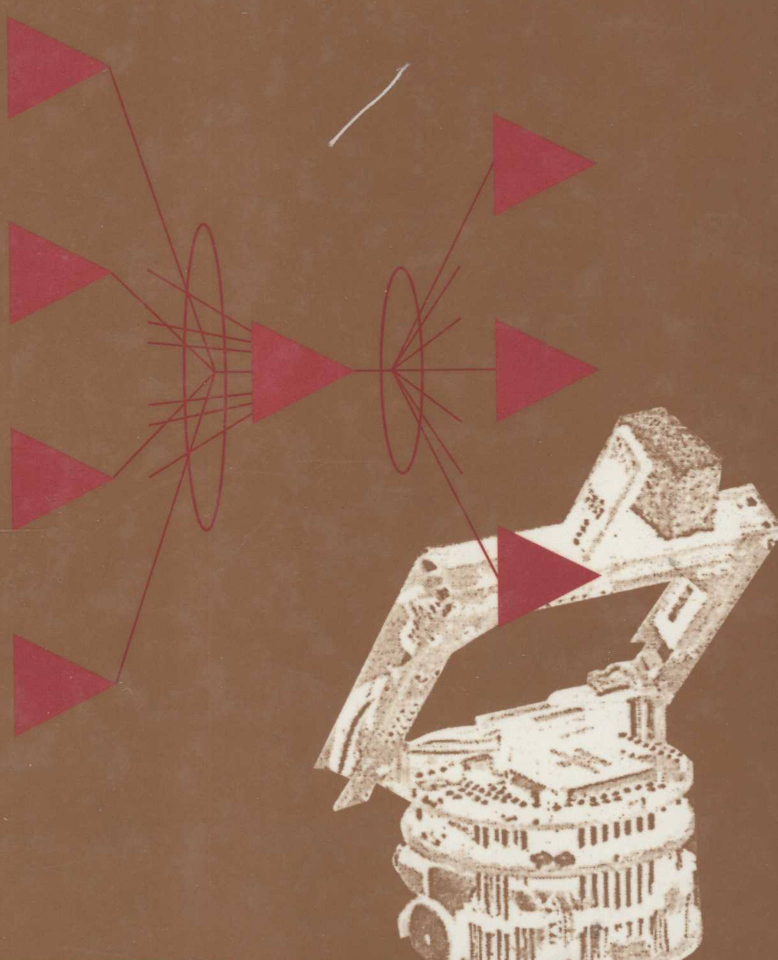
NEURAL NETWORK PERSPECTIVES ON COGNITION AND ADAPTIVE ROBOTICS

EDITED BY ANTONY BROWNE



IoP

Neural Network Perspectives on Cognition and Adaptive Robotics

Edited by

Antony Browne

Nene College, UK

Institute of Physics Publishing
Bristol and Philadelphia

British Library Cataloguing-in-Publication Data

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

ISBN 0 7503 0455 3

Library of Congress Cataloging-in-Publication Data are available

Published by Institute of Physics Publishing, wholly owned by The Institute of Physics, London

Institute of Physics Publishing, Dirac House, Temple Back, Bristol BS1 6BE, UK

US Editorial Office: Institute of Physics Publishing, The Public Ledger Building, Suite 1035, 150 South Independence Mall West, Philadelphia, PA 19106, USA

Typeset in TeX using the IOP Bookmaker Macros
Printed in the UK by J W Arrowsmith Ltd, Bristol

Neural Network Perspectives on Cognition and Adaptive Robotics

Related titles published by
Institute of Physics Publishing

Neural Computing: An Introduction
R Beale and T Jackson

Handbook of Neural Computation
Edited by E Fiesler and R Beale

Neural Network Analysis, Architectures and Applications
Edited by A Browne

To Jayne and Liam

Preface

Neural networks have been used extensively to model human cognition. Human reasoning and language are two areas where, so far, no computational system has reached the level of human performance. There is much interest in neural network models as they provide an approach to the modelling of cognition different from that of traditional symbolic artificial intelligence (AI). Perhaps because of the success of neural network systems in modelling some aspects of cognition and human language processing, many questions have been raised by the symbolic AI community as to the representational power of connectionist systems. Some of these questions are discussed in depth in this book.

Most of us are familiar with images of robots, and often these robots can be seen to be performing useful tasks. However, most of the robots we are familiar with have an important limitation. They must be pre-programmed for a fixed environment and a fixed set of tasks. If their environment changes they fail ungracefully. Adaptive robots using neural computing techniques are able to change their behaviour as their environment changes.

This book is designed to give an overview of some of the many perspectives that neural computing gives on cognition and autonomous robotics. It is not written as an introductory textbook; it is assumed that the reader has some previous knowledge of neural networks and an understanding of their basic mechanisms. The book is divided into several parts:

- *Representation.* This part delves into questions on the representational power of neural networks and discusses the use of information theory in neural computing.
- *Cognitive modelling.* The use of neural networks for cognitive modelling is discussed, including both the modelling of human reasoning and the implications of neurophysiological data.
- *Adaptive robotics.* Robots which can adapt to their environment are described, together with a discussion on biologically realistic learning mechanisms.

Many chapters in the book are cross-referenced to a companion volume *Neural Network Analysis, Architectures and Algorithms* (1997, Institute of

Physics Publishing), which can be seen as complementary to this volume as it contains chapters on the following areas:

- *Understanding and simplifying networks.* Methods for extracting information about what a trained neural network has learned are outlined, together with a method for simplifying network architectures based on information theory.
- *Novel architectures and algorithms.* Two novel hardware implementations of neural networks are described, together with a discussion of fast training algorithms for feed-forward network architectures.
- *Applications.* Some applications of neural networks in the diverse fields of control (including neuro-fuzzy control), data compression and target identification are discussed.

Hopefully the following chapters will clarify some of the many fields in which neural computing is expanding in its attempts to model human cognition more accurately and produce flexible and adaptive robotic systems. Perhaps in the (possibly very distant) future these two areas will meet, and we will have robotic systems with human-like capabilities of thought and speech.

No textbook can ever hope to give a comprehensive review of the myriad directions in which current research is headed. However, this volume attempts to give a flavour of some of the most promising areas.

A Browne
July 1997

Contents

PART 1

REPRESENTATION

Now that the hysteria has died down about neural computing being a magical solution, both to the problems faced by cognitive scientists in modelling human cognition and to the problems faced by engineers in developing intelligent solutions to solve practical problems, it is time to take stock of the situation. Several questions can be asked:

How powerful are neural networks? What sort of things can they represent and what operations can they perform? Are there limits to the sort of operations they can perform and the types of structures they can represent? How do they compare to symbolic computational systems?

Because of the recent popularity of neural computing and the perhaps extravagant claims put forward by some of its proponents, the field has come under fire from many directions. This part of the book concentrates on aspects of the representational power of neural networks and how some of the challenges regarding this representational power have been answered. A formal analysis of what is happening inside neural network systems, using information theory, is also discussed.

Chapter 1

Challenges for Neural Computing

Antony Browne
School of Information Systems, Nene College, UK
antony.browne@nene.ac.uk

1.1 Introduction

In recent years neural networks have been touted as the miracle solution, both to the problems that cognitive scientists face in trying to build models of cognition and to the practical problems that engineers face in trying to develop intelligent solutions for commerce and industry. However, neural networks have also been subject to much criticism. Some of these criticisms, together with the attempts that scientists and engineers working with neural networks have made to answer them, are given below.

1.2 Two schools of thought for intelligent systems

A substantial proportion of artificial intelligence (AI) research, and the application of the technology produced from this research, is based on the assumption that all the important aspects of human cognition, and all the useful tasks which humans require their machines to do, may (at some time) be described or carried out by a computational model. This view (computationalism) follows on from the belief that any aspect of cognition or practical task can be modelled computationally (the Church–Turing thesis). This can be contrasted with the approach of some theorists who hold that certain properties of the mind (and hence some properties that we may wish to capture in practical intelligent systems) cannot be captured algorithmically [228]. Many researchers in cognitive science believe that implementation details are irrelevant for cognition, as do many engineers who implement the practical intelligent systems currently used in commerce and industry. As a result of this some of those involved with these systems have not tried to obtain inspiration from the physical structure of the brain, but have instead concentrated on modelling the abstract structure of the mind and the practical application of the associated technology. In recent years there has been much disagreement as to how to produce systems with the desired properties and research in intelligent systems can be seen to have split into two schools. These are the classical (in so far as they have been researching in the field of AI for the longest time) symbolic AI school and the connectionist school. It is a matter of some debate as to which school actually has the most powerful model and members of the connectionist school implementing their models with neural networks are currently attempting to face up to the criticisms levelled at their models by the symbolic AI camp. One could argue that since both neural networks and classical symbolic systems are universal Turing machines [111], at one level of abstraction there is no distinction between them. However, in the nature of the representations they use, in their observed performance given different tasks, and in their adequacy for the modelling of intelligent systems, there are many differences between these two schools of thought.

1.2.1 The symbolic AI school

The classical symbolic school states that the correct level at which to model the mind is that of the symbol, an entity in a computer program that is taken to refer to an entity in the real world. The main assumptions on which the symbolic AI paradigm rests were first explicitly stated under the 'physical symbol system hypothesis' [210] which states that 'a physical symbol system has the necessary and sufficient means for general intelligent action'. In this definition:

- Necessary means that any physical system that exhibits general intelligence will be an instance of a physical symbol system (PSS).
- Sufficient means that any physical symbol system can be organized further to exhibit general intelligent action.
- General intelligent action means the same scope of intelligence seen in humans.

This implies that in real situations behaviour can occur that is both appropriate to the needs of the system and adaptive to the demands of the external environment (within some physical limits imposed by processing speed and memory requirements). This hypothesis states that an entity can only be intelligent if it instantiates a PSS. Symbols correspond to unitary concepts and are taken to represent objects, events, relations between objects and relations between events. In this way the word symbol comes to represent a concept or entity we can put a meaningful label on such as '**apple**' or '**dog**'. At any one time a symbol represents a single entity or concept. Symbols are atomic, they may combine to form symbol structures, but individual symbols may not be broken down. An obvious example of such symbols occurs in programming languages such as Prolog, where atoms such as '**man**', '**fred**', or