Neural Computers

Edited by

Rolf Eckmiller   Christoph v. d. Malsburg

NATO ASI Series

Series F: Computer and Systems Sciences, Vol. 41

Neural Computers

Edited by

Rolf Eckmiller

Department of Biophysics
University of Düsseldorf
Universitätsstraße 1
D-4000 Düsseldorf, FRG

Christoph v. d. Malsburg

Department of Neurobiology
Max-Planck-Institute of
Biophysical Chemistry
Nikolausberg
D-3400 Göttingen, FRG

Springer-Verlag
Berlin Heidelberg New York London Paris Tokyo
Published in cooperation with NATO Scientific Affairs Division

Proceedings of the NATO Advanced Research Workshop on Neural Computers, held in Neuss, Federal Republic of Germany, September 28 – October 2, 1987.

ISBN 3-540-18724-3 Springer-Verlag Berlin Heidelberg New York
ISBN 0-387-18724-3 Springer-Verlag New York Berlin Heidelberg

Printed in Germany

Printing: Druckhaus Beltz, Hemsbach; Binding: J. Schäffer GmbH & Co. KG, Grünstadt
2145/3140-543210

PREFACE

This book is the outcome of a NATO Advanced Research Workshop (ARW) held in Neuss (near Düsseldorf), Federal Republic of Germany from 28 September to 2 October, 1987.

The workshop assembled some 50 invited experts from Europe, America, and Japan representing the fields of Neuroscience, Computational Neuroscience, Cellular Automata, Artificial Intelligence, and Computer Design; more than 20 additional scientists from various countries attended as observers.

The 50 contributions in this book cover a wide range of topics, including: Neural Network Architecture, Learning and Memory, Fault Tolerance, Pattern Recognition, and Motor Control in Brains versus Neural Computers. Twelve of these contributions are review papers.

The readability of this book was enhanced by a number of measures:

* The contributions are arranged in seven chapters.
* A separate List of General References helps newcomers to this rapidly growing field to find introductory books.
* The Collection of References from all Contributions provides an alphabetical list of all references quoted in the individual contributions.
* Separate Reference Author and Subject Indices facilitate access to various details.

Group Reports (following the seven chapters) summarize the discussions regarding four specific topics relevant for the 'state of the art' in Neural Computers.

It is hoped that this book will prove useful as a reference book for future research in the field of Neural Computers and as a catalyzer for more international cooperation in the endeavor to TRANSFER CONCEPTS OF BRAIN FUNCTION AND BRAIN ARCHITECTURE TO THE DESIGN OF SELF-ORGANIZING COMPUTERS WITH NEURAL NET ARCHITECTURE.

The editors wish to thank the other two members of the organizing committee for the workshop, Elie Bienenstock (Paris) and Jerome Feldman (Rochester) for their engagement.

The steady and efficient managerial and secretarial assistance of Sabine Canditt, Barbara Lohmann, Dietmar Ott, and Anneliese Thelen, which made it possible to prepare the book manuscript for publication within only three weeks after the workshop, is gratefully acknowledged.

Neuss, October 1987

R. Eckmiller
C. v.d. Malsburg

ACKNOWLEDGEMENT OF SPONSORSHIP

The generous sponsorship of the following institutions, which made this Advanced Research Workshop (ARW) possible and provided support for observers, is gratefully acknowledged:

NATO Scientific Affairs Division/Brussels, Belgium

Minister for Science and Research of the State of Northrhine-Westfalia/Düsseldorf, Fed. Rep. of Germany

International Computer Science Institute/Berkeley, USA

Siemens Corporation/München, Fed. Rep. of Germany

TABLE OF CONTENTS

THE ROLE OF ADAPTIVE AND ASSOCIATIVE CIRCUITS IN FUTURE COMPUTER DESIGNS

Teuvo Kohonen
Helsinki University of Technology
Laboratory of Computer and Information Science
Rakentajanaukio 2 C, SF-02150 Espoo, Finland

ABSTRACT

The fifth-generation computers have essentially been designed around knowledge data bases. Originally they were supposed to accept natural data such as speech and images; these functions, however, have now been abandoned. Some vistas to sixth- and later-generation computers have already been presented, too. The latter are supposed to be even more "natural". Certainly the fifth-generation computers have applied very traditional logic programming, but what could the alternative be? What new functions can evolve from the "neural computer" principles? This presentation expounds a few detailed questions of this kind. Also some more general problem areas and motivations for their handling are described.

"All forecasting is difficult, especially forecasting of the future."
Old Chinese proverb

1. PARALLEL ASSOCIATIVE SEARCH AND INFERENCE IN KNOWLEDGE DATA BASES

It is an old notion that the biological brain operates according to *associative* principles, and that the recall of items from memory is *content-addressable* (Aristotle, 384-322 B.C.). Naturally, associative mechanisms are characteristic of thinking processes, whereas nobody may want to claim that all the problem-solving procedures applied, say, in scientific thinking are directly derivable from them. It seems necessary to emphasize that cultural development has always been based on *instrumental means*: language, writing, logical and mathematical notations and systems, etc. which have developed during a very long time, and which obey their own mechanisms. People using them must only *learn* them as some kind of *internal*

NATO ASI Series, Vol. F41
Neural Computers
Edited by R. Eckmiller and Ch. v. d. Malsburg

representations in their minds. On the other hand, the *procedures* in problem-solving can be activated without all the time referring to *neural* processes. The great number of rules usually involved in these procedures may easily lead one into concluding that our brain were full of detailed circuits to implement these rules; consider, however, that we apply route learning even for the simplest formal symbolisms such as addition and multiplication tables, thereby occupying substantial portions of the brain, like the visual, auditory, and motor areas. It seems more plausible that only a few central items or rules are recalled from the brain, whereas the majority of rules is supported by various artificial tools and aids.

Everybody is expert in thinking, but when we are asked to specify what kind of abilities we want from "intelligent" machines, the views tend to be very concrete and straightforward. Examples of this are the "expert systems". It seems highly desirable to possess computer systems to which various kinds of queries, concerning facts as well as advices, can be presented. The function to be implemented by the computer and its memories is then a rather straightforward *searching* on a number of conditions. As long as the stored items are discrete-valued, such as symbol strings, their storage is most effectively made in separate memory locations. To find an item quickly on the basis of its content, there exist both software solutions (*hash-coding*) as well hardware mechanisms (*content-addressable memory, CAM*) both of which have been known since 1955. "Neural" principles are not needed unless the items or their relations are defined in a continuous scale, as further pointed out below.

Let us now consider the act of searching in somewhat more detail. *Information* usually consists of data items and their links (relations, "associations") to subsets of other items. The *knowledge* acquired in such a data base is realized and can be managed through long chains of such links, which are realized when the associations are made to overlap.

It is perhaps illustrative to compare a data base and the searching process with a system of mathematical equations and their solution. When we present a query, we in fact set up one or more "equations" which contain unknown variables: for instance, we may specify a number of partial relations in which some of the members are unknown, and the system has to find from memory all the relations which match with the "equations" in their specified parts, whereby a number of solutions for the unknown variables become known.

Above I was in fact talking about the so-called *relational data bases* which are widely used for business data. There also exist other types of data structures, (e.g., inverted lists) but in principle they are not very much different. All of these systems seem to call for the following elementary operations: 1. Parallel or otherwise very fast matching of a number of search arguments (such as known members in the above relations) with all the elementary items stored in memory, and their provisional marking. 2. Analysis of the markings and sequential readout of the results that satisfy all conditions.

In the simplest case I was also assuming that the representations of all items are unique, whereby for matching, full identity of the stored item with the search argument is required. Under these conditions, the data base machines can also be realized by software in standard computer systems. If parallel searching memories, such as the CAMs are available (they may be 50 times more costly than usual memories), one may match arbitrary *parts* of the items, or compare them on the basis of *numerical magnitude relations*. Natural signals, such as images, speech, and physical measurements which comprise a significant bulk of all available information, however, are seldom unique or even numerical; for one thing they may contain a significant amount of noise. Naturally, their numerical values are easily stored, but parallel searching on the basis of *approximate* matching is a task which calls for quite special hardware. In principle, every memory location could be provided with arithmetic circuits which can analyze various approximate magnitude relations. This kind of solution is hardly practicable for more than, say, 10,000 memory locations, which is a rather small figure for data bases.

For most effective approximate searching, a new generation of "associative" networks, namely, the *artificial neural networks* has been suggested, and among other things they have aroused expectations of new kinds of data bases and even "neural expert systems". The searching arguments are imposed as an initial condition to the network, and solution for the "answer" results when the state of the network relaxes to some kind of "energetic minimum". One has to note the following facts characteristic of these devices. 1. Their network elements are *analog devices*, whereby representation of numerical variables, and their matching can only be defined with a relatively low accuracy. This, however, may be sufficient for prescreening purposes which is most time-consuming. 2. A vast number of relations in memory which only approximately match with the search argument can be activated. On the other hand, since the "conflicts" then cannot be totally resolved but only minimized, the state of the neural network to which it converges in the process represents some kind of *optimal* answer (usually, however, only in the sense of Euclidean metric). 3. The "answer", or the asymptotic state which represents the searching result has *no alternatives*. Accordingly, it is not possible, except in some rather weird constructs, to find the complete set of solutions, or even a number of the best candidates for them. It is neither sure that the system will converge to the *global* optimum; it is more usual that the answer corresponds to one of the local optima which, however, may be an acceptable solution in practice.

With regard to the possible role of the "neural networks" as *memory units* in future computers, one has to consider whether the above three points are acceptable or not. It is self-evident that theoretical mathematical, physical, and engineering problems cannot be handled that way, whereas there may exist applications, say, in writing reports and planning strategies, where even incomplete answers from a large data base will be useful. There may also exist on-line tasks, e.g. with "intelligent" robots, or various *ad hoc* tasks, say, relating to space missions, where a unique although suboptimal answer which is obtained promptly is better than a set of alternative answers obtained in an exhaustive search.

2. PATTERN RECOGNITION PROBLEMS (INPUT OF IMAGES, SPEECH, ETC.)

The first "adaptive" and "associative" machines which were devised around 1960 were in fact already intended for statistical interpretation of patterned variables, such as speech signals, weather maps, and handwritten symbols. The task was not so much to *search* for stored items but to *categorize* or *classify* input data. This research area became later known under the name *Pattern Recognition*, although a more proper name for it might have been *Artificial Perception*, and its methods (based on general-purpose computers) became more heuristic. As a matter of fact, however, the objective always seems to have been to implement sensory and cognitive functions to interprete the observations at different levels of abstraction, like the biological beings are doing.

The most important *application areas* for "neural pattern recognition" could be the same as those for which conventional, heuristic methods have been developed during the past thirty years:

- remote sensing
- medical image analysis
- industrial computer vision (especially for robotics)
- input devices for computers, etc.

More concrete tasks for which special computer equipment has already been developed are:

- segmentation and classification of regions from images
- recognition of handwritten characters and text
- recognition of speech
- processing, especially restoration of noisy pictures, etc.

On a more ambitious level, one may wish to achieve the capabilities of:

- image analysis (referring to different thematic levels of abstraction, such as monitoring of land use on the basis of satellite pictures)
- image understanding (interpretation of scenes)
- speech understanding (parsing and interpretation of spoken sentences), etc.

To implement these tasks, certain basic problems still call for better understanding, for instance, those concerning the *intrinsic properties* (features) of input information, such as

- the most natural pattern primitives (lines, their curvatures and end points, edges, statistics of point groups)
- visual information which describes the surface curvature and cusps
- texture
- phonological invariants in speech, etc.

According to an estimate which I have based on computerized document search, the number of papers published on Pattern Recognition nowadays amounts to some 30,000. It is highly improbable that any easy heuristic method had escaped our attention; but still the performance of artificial methods falls far short from that of the biological sensory systems. My personal explanation to this is that in biological sensory systems, a very thorough multilevel optimization of the information processing functions and resources has been achieved during evolution. It is very difficult to describe all the input data to which this ability is related. It would similarly be hopeless to try any analytical definition of these mappings. If artificial machines are expected to have comparable abilities, the only possibility seems to be to provide them with adaptive properties which optimize their network structures as well as their tuning. The former type of development seems to need some new kind of *natural choice* from many alternative structures, whereas adjustment of the "matched filters" (similar to Perceptron, Adaline, etc.) has already many solutions. There also still remain the fundamental problems of how to make the recognition operation *invariant* with respect to various transformation groups of observations, and how to take *context* into account in the most general way.

It seems that the mainstreams of Pattern Recognition research are nowadays directed to Picture Processing and Computer Vision, whereby the input data appear as large, regular arrays of topologically related variables like the picture elements (pixels) in digital images. Although many syntactic and other structural methods have been developed for the analysis of the objects and their relations, processing and interpretation of the picture data is usually made in a highly parallel but regular fashion such as in the SIMD (single instruction stream, multiple data stream) computers. This principle, however, is not met in biological sensory