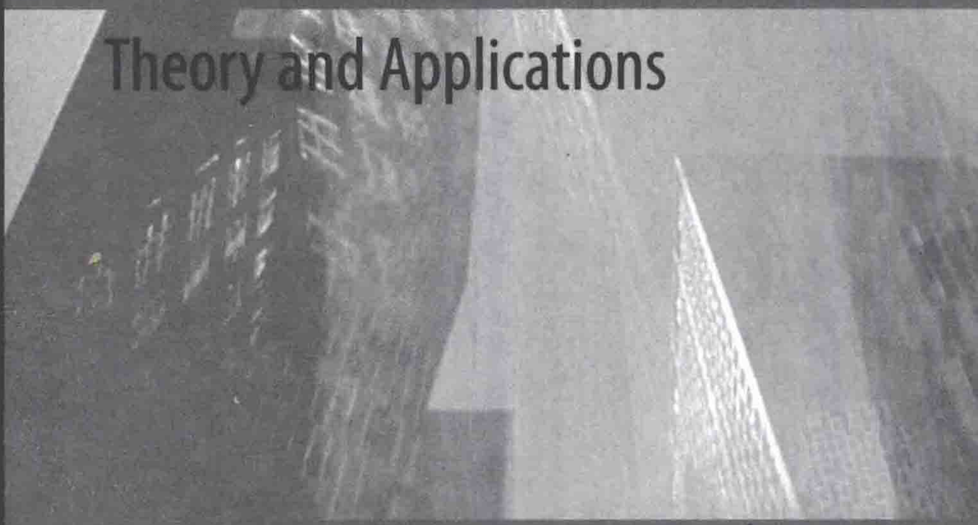
ADVANCED TOPICS IN SCIENCE AND TECHNOLOGY IN CHINA

Xingui He
Shaohua Xu

Process Neural Networks

Theory and Applications



ZHEJIANG UNIVERSITY PRESS
浙江大学出版社
Springer

Xingui He
Shaohua Xu

Process Neural Networks

Theory and Applications

With 78 figures

ZHEJIANG UNIVERSITY PRESS
浙江大学出版社

Springer

图书在版编目（CIP）数据

过程神经元网络＝Process Neural Networks：英文／何新贵，许少华著．—杭州：浙江大学出版社，2009.7
（中国科技进展丛书）
ISBN 978-7-308-05511-6

Ⅰ.过…　Ⅱ.①何…②许…　Ⅲ.神经网络-研究—英文
Ⅳ.TP183

中国版本图书馆 CIP 数据核字（2008）第 137320 号

Not for sale outside Mainland of China
此书仅限中国大陆地区销售

过程神经元网络

何新贵　许少华　著

责任编辑	林昌东
封面设计	Frido Steinen-Broo
出版发行	浙江大学出版社
	网址：http://www.zjupress.com
	Springer-Verlag GmbH
	网址：http://www.springer.com
排　　版	杭州中大图文设计有限公司
印　　刷	杭州富春印务有限公司
开　　本	787mm×960mm　1/16
印　　张	15.75
字　　数	530 千
版 印 次	2009 年 7 月第 1 版　2009 年 7 月第 1 次印刷
书　　号	ISBN 978-7-308-05511-6（浙江大学出版社）
	ISBN 978-3-540-73761-2（Springer-Verlag GmbH）
定　　价	106.00 元

版权所有　翻印必究　印装差错　负责调换
浙江大学出版社发行部邮购电话（0571）88925591

Preface

The original idea for this book came from a conference on applications of agricultural expert systems, which may not seem obvious. During the conference, the ceaseless reports and repetitious content made me think that the problems the attendees discussed so intensely, no matter which kind of crop planting was involved, could be thought of as the same problem, i.e. a "functional problem" from the viewpoint of a mathematical expert. To achieve some planting indexes, e.g. output or quality, whatever the crop grown, different means of control performed by the farmers, e.g. reasonable fertilization, control of illumination, temperature, humidity, concentration of CO_2, etc., all can be seen as diversified time-varying control processes starting from sowing and ending at harvest. They could just as easily be seen as the inputs for the whole crop growth process. The yield or the quality index of the plant can then be considered as a functional dependent on these time-varying processes. Then the pursuit of high quantity and high quality becomes an issue of solving an extremum of the functional.

At that time, my research interest focused on computational intelligence mainly including fuzzy computing, neural computing, and evolutionary computing, so I thought of neural networks immediately. I asked myself why not study neural networks whose inputs and outputs could both be a time-varying processes and why not study some kinds of more general neural networks whose inputs and outputs could be multi-variable functions and even points in some functional space. Traditional neural networks are only used to describe the instantaneous mapping relationship between input values and output values. However, these new neural networks can describe the accumulation or aggregation effect of the outputs on the inputs on the time axis. This new ability is very useful for solving many problems including high-tech applications in agriculture and for elaborate description of the behavior of a biological neuron. The problems that the traditional neural networks solved are function approximation and function optimization, and the problems we need to solve now are functional approximation and functional optimization, which are more complicated. However, as a mathematician my intuition told me that there existed the possibility of resolving these problems with certain definite constraints and that there might be the prospect of broader applications in the future. In research during the following years, I was attracted by these issues. In addition to numerous engineering tasks (e.g. I had assumed responsibility in China for manned airship engineering), almost all the rest of my time was spent on this study. I presented the concept of the "Process Neural Network (PNN)", which would be elaborated in this book. In recent years, we have done some further work on the theories, algorithms, and applications of process neural networks, and we have solved some basic theory issues, including the existence of solutions under certain conditions, continuity of the process neural network models, several approximation theorems (which are the theoretical foundations on which process neural network models can be applied to various practical problems), and we have investigated PNN's computational capability. We have also put forward some useful learning algorithms for process neural networks, and achieved some preliminary applications including process control of chemical reactions, oil recovery, dynamic fault inspection, and communication alert and prediction. It is so gratifying to obtain these results in just a few years. However, the research is arduous and there is a long way to go. Besides summarizing the aforementioned preliminary achievements, this monograph will highlight some issues that need to be solved.

At the time of completing this book, I would like to express my sincere thanks to my many students for their hard work and contributions throughout these studies. Furthermore, I also wish to thank those institutes and persons who generously provided precious data and supported the actual applications.

Xingui He
Peking University
Beijing
April, 2009

此为试读,需要完整PDF请访问: www.ertongbook.com

Contents

此为试读，需要完整PDF请访问：www.ertongbook.com

1

Introduction

As an introduction to this book, we will review the development history of artificial intelligence and neural networks, and then give a brief introduction to and analysis of some important problems in the fields of current artificial intelligence and intelligent information processing. This book will begin with the broad topic of "artificial intelligence", next examine "computational intelligence", then gradually turn to "neural computing", namely, "artificial neural networks", and finally explain "process neural networks", of which the theories and applications will be discussed in detail.

1.1 Development of Artificial Intelligence

The origins of artificial intelligence (AI) date back to the 1930s–1940s. For more than half a century, it can be said that the field of artificial intelligence has made remarkable achievements, but at the same time has experienced many difficulties.

To give a brief description of artificial intelligence development, most events and achievements (except for artificial neural networks) are listed in Table 1.1.

The main purpose of artificial intelligence (AI) research is to use computer models to simulate the intelligent behavior of humans and even animals, to simulate brain structures and their functions, the human thinking process and its methods. Therefore, an AI system generally should be able to accomplish three tasks: (a) to represent and store knowledge; (b) to solve various problems with stored knowledge; (c) to acquire new knowledge when the system is running (that is the system has the capability of learning or knowledge acquisition).

AI has been developing rapidly over the past 50 years. It has been widely and successfully applied in many fields, such as machine learning, natural language comprehension, logic reasoning, theorem proving, expert systems, etc.

Along with the continuous extension of AI application fields and with the problems to be solved becoming more and more complex, traditional AI methods based on a symbol processing mechanism encountered more and more difficulties

Table 1.1 The milestones of artificial intelligence

Date	Leading players	Description and significance of event or production
1930s – 1940s	Frege, Whitehead, and Russell	Established mathematical logic system and gave us new ideas about computation
1936	Turing	Established automata theory, promoted the research of "thinking" machine theory, and proposed the recursive function based on discrete quantities as the basis of intelligent description
1946	Turing	Pointed out the essence of the theory "thinking is computing" and presented formal reasoning in the process of symbolic reasoning
1948	Shannon	Established information theory which held that human psychological activities can be researched in the form of information, and proposed some mathematical models to describe human psychological activities
1956	McCarthy *et al.*	Proposed the terminology "artificial intelligence" (AI) for the first time which marks the birth of AI based on symbol processing mechanism
1960	McCarthy	Developed the list processing language LISP which could deal with symbols conveniently and later was applied widely in many research fields of AI
1964	Rubinson	Proposed the inductive principle which marks the beginning of research into machine proving of theorems in AI
1965	Zadeh	Proposed the fuzzy set, and pointed out that the membership function can describe fuzzy sets, which marked the beginning of fuzzy mathematics research. Binary Boolean logic especially was extended to fuzzy logic
1965	Feigenbaum	Proposed an expert system which used normative logical structure to represent expert knowledge with enlightenment, transparency, and flexibility which was widely applied in many fields
1977	Feigenbaum	Proposed knowledge engineering that used the principles and methods of AI to solve application problems. Established expert systems by developing intelligent software based on knowledge

with artificial intelligence technology when solving problems such as knowledge representation, pattern information processing, the combinatorial explosion, etc. Therefore, it has practical significance to seek a theory and method that have intelligent characteristics such as self-organization, self-adaptation, self-learning, etc., and which is suitable for large-scale parallel computation.

Almost at the same time as the above research activities, some scientists were also seeking methods of representing and processing information and knowledge from different viewpoints and research domains. In 1943, the physiologist

McCulloch and the mathematician Pitts abstracted the first mathematical model of artificial neurons [1] by imitating the information processing mechanism of biological neurons, which marked the beginning of artificial neural networks research based on connectionism. In 1949, the psychologist Hebb proposed the Hebb rule [2], which can achieve learning by modifying the connection intensity among neurons, and make the neuron have the ability to learn from the environment. In 1958, Rosenblatt introduced the concept of the perceptron [3]. From the viewpoint of engineering, this was the first time that an artificial neural network model was applied in information processing. Although the perceptron model is simple, it has characteristics such as distributed storage, parallel processing, learning ability, continuous computation, etc. In 1962, Widrow proposed an adaptive linear element model (Adaline) [4] that was successfully applied to adaptive signal processing. In 1967, Amari implemented adaptive pattern classification [5] by using conferring gradients. The period from 1943 to 1968 can be considered as the first flowering of artificial neural networks research. In this period, there were many more important research achievements, but we have not listed all of them here.

In 1969, Minsky and Papert published *Perceptrons* [6], which indicated the limitation of function and processing ability of the perceptron, that it cannot even solve simple problems such as "Xor". The academic reputation of Minsky and the rigorous discussion in the book, led their viewpoints to be accepted by many people, and this made some scholars who had engaged in artificial neural networks earlier to turn to other research fields. Research in artificial neural networks came into a dormant period that lasted from 1969 to 1982.

Although research in neural networks encountered a cold reception, many scholars still devoted themselves to theoretical research. They proposed lots of significant models and methods, such as Amari's neural network mathematical theory [7] (in 1972), Anderson *et al.*'s BSB (Brain-State-in-Box) model [8] (in 1972), Grossberg's adaptive theory [9] (in 1976), etc.

In the early 1980s, the physical scientist Hopfield proposed a feedback neural network (HNN model) [10] (in 1982) and successfully solved the TSP (Traveling Salesman Problem) by introducing an energy function. Rumelhart *et al.* proposed the BP algorithm in 1986 that preferably solved the adaptive learning problem [11] of feedforward neural networks. From 1987 to 1990, Hinton [12], Hecht-Nielson [13], Funahashi [14] and Hornik *et al.* [15] separately presented the approximation capability theorem of multi-layer BP network which proved that multi-layer feedforward neural networks can approximate any L_2 function. This theorem established the theoretical basis for the practical application of neural networks, and helped the theory and application of neural networks to mature gradually. Artificial neural networks came into a second flowering of research and development.

In 1988, Linsker proposed a new self-organizing theory [16] based on perceptron networks, and formed the maximum mutual information theory based on Shannon's information theory.

In the 1990s, Vapnik and his collaborators proposed a network model called Support Vector Machine (SVM) [17-19] according to the structural risk minimization principle based on learning theory with a limited sample, and it was widely applied to many problems such as pattern recognition, regression, density estimation, etc.

In recent years, many novel artificial neural network models have been established and broadly applied in many areas such as dynamic system modeling [20,21], system identification [22], adaptive control of nonlinear dynamic systems [23,24], time series forecasting [25], fault diagnosis [26], etc.

In 2000, we published process neuron and process neural network (PNN) models after years of intensive study [27,28]. The input signals, connection weights, and activation thresholds of process neurons can be time-varying functions, or even multivariate functions. Based on the spatial weighted aggregation of traditional neurons, an aggregation operator on time (or even more factors) is added to make the process neuron have the ability to process space-time multidimensional information. This expands the input–output mapping relationship of the neural networks from function mapping to functional mapping, and greatly improves the expression capability of neural networks. A series of basic theorems (including existence theorem, approximation theorem, etc.) of process neural networks have been proved and some related theoretical problems have been solved. Practice shows that PNN models have broad applications in many actual signal processing problems relating to process. These will be the core content in this book.

At present, there are thousands of artificial neural network models, of which there are more than 40 primary ones. The application scope of these models covers various fields including scientific computation, system simulation, automatic control, engineering applications, economics, etc., and they show the tremendous potential and development trends of artificial neural networks. However, most present neural networks are traditional neural networks with spatial aggregation and have no relation with time.

Traditional AI methods based on symbol processing mechanisms and neural networks based on connectionism are two aspects of AI research, and each of them has its own advantages and limitations. We assume that the combination of both methods can draw strengths from each other to offset the weaknesses. For example, the setting and connection mode of neural network nodes (neurons) can definitely connect the solving goal with the input variables. We once observed that the specific reasoning rules can be considered as the network nodes (neurons) and "reasoning" can be converted into "computing". At the same time, according to the rules described by knowledge in the practical field, the connection mode and activation threshold among the network nodes can be properly chosen and modified to express more reasonable logical relationships among the described problems, and the corresponding expert system can be designed in terms of a structure of a neural network.

The term AI, as its name suggests, involves making "intelligence" artificially, or even making an intelligent system. Its short-term goal is to implement intelligence simulation in an existing computer and endow the computer with some intelligent behavior, while its long-term goal is to manufacture an intelligent system and endow it with intelligence similar to (or perhaps exceeding in some aspects) that of animals or human beings. Using AI to study autocorrelation problems in the human brain seems to be a paradox in logic and involves complex recursive processes in mathematics, and is high in difficulty. The problem is that how the brain works might never be understood in some sense, because the brain itself is also changing and developing while people are studying it. If some aspects of the brain at some time were studied clearly, the brain function at that time might develop, the former state might change again, and this would not be the same as the original research objective. However, such a spiral research result is still very significant and can be applied in various practical problems. Therefore, we think that, on the one hand, AI should have a long-term research goal and this goal can be gradually approximated; on the other hand, we still need to propose various short-term goals and these goals should not deviate from practical applications to reach for that which is beyond our grasp. The development history of AI in this respect has already given us many lessons, which are worth remembering by AI researchers.

In short, the development of Artificial Intelligence has experienced ups and downs during the past 60 years. Because of the increased demands in science fields and practical applications, we believe that AI will undergo further development, play a more important role in the advancement of science and technology through its role in tackling human and other problems that are difficult to solve with traditional methods at present, and that it will also make great contributions to producing intelligent systems for human beings in the future.

1.2 Characteristics of Artificial Intelligent System

What system can be called an intelligent system? This is a question that we should answer before setting about researching intelligent systems. It can be said that we should set up a research goal. Of course, the understanding of this question changes dynamically, and we cannot answer it clearly in a moment. In fact, we can first find some rough answers from analysis of the intelligent behavior of biological systems.

(1) An intelligent system is a memory system

From the perspective of neurophysiology, what is called memory is the storage capacity and the processing procedure for information obtained from the external or produced from the internal. There is a large amount of information that comes from the outside world, through sense organs, inwards to the brain. The brain does not store all the information that the sensory organ directly receives, but only stores the information obtained through learning or that is of some significance. Therefore, the intelligent system must have memory storage capacity; otherwise, it will lose the object and cannot store processing results, just as a person who has completely lost his memory will no longer have intelligence. In addition, the artificial intelligent system is not completely identical with the human brain; the latter has a powerful memory ability which will decrease gradually, so the former should simulate the latter in aspects of memory and forgetting in some way.

(2) An intelligent system is a computation system

Cognitive science considers that "cognition is computing": it combines intelligence with computation closely and forms a new concept—computational intelligence. What is called computation refers to the process by which we carry out various operations and combinations (including digital or analog) repeatedly using a certain symbol set according to some rules. The acquisition, representation, and processing of knowledge can all come down to the computation process. Therefore, an artificial intelligence system should also have this computing capability to accomplish the corresponding functions. In the Chinese language, there is an alias "electronic cerebra" for the computer, which is of great significance. Carrying out various digital or analog operations fleetingly is the strong point of the computer, so a computer is quite suitable for simulating some intelligent behaviors. However, there are some troubles and problems when we directly use current digital machines or analog machines to handle fuzzy information or qualitative data, and indeed sometimes they cannot handle it at all, so we expect to use a digital machine that has an analog operation component. Such a machine is different from a general digital/analog mixed machine, it should have a uniform digital/analog mixed memory in which to deposit the processing object, and its processor should possess a uniform mixed processing ability for this mixed information. We believe that research on the computing capability of an intelligent system is very important and worth strengthening, and that research and development of the computing capability of an intelligent system (such as fuzzy neural computation) will greatly promote basic research on intelligence, or even the whole development of computer science.

(3) An intelligent system is a logical system

Traditional logic is binary logic and it is adequately utilized in the von Neumann computer, but in fact, the reasoning logic of humans does not strictly abide by binary logic. Especially when the cognition of something is unclear or not completely clear, we only describe this by a qualitative or fuzzy concept, and handle this with a qualitative method or by fuzzy logic. Therefore, an artificial intelligent system should be able not only to carry out routine logical reasoning, but also to represent and process various qualitative and fuzzy concepts that are described by natural language, and then execute the corresponding qualitative or fuzzy reasoning. Consequently, an artificial intelligent system becomes a strong logical processing system. In addition to logical reasoning, the system should also be able to execute complex logical judgments, and adopt appropriate action or reaction according to the judgments. The current computer is competent for binary logic or finite multi-valued logic, but is helpless when it comes to some continuous-valued logic (for example, fuzzy logic) and qualitative logical reasoning. We need the above digital/analog unified processing hybrid computer to meet these demands.

(4) An intelligent system is a perceptive system

An important characteristic of a biological system is that it can perceive the outside

此为试读，需要完整PDF请访问：www.ertongbook.com