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

# ADVANCED TOPICS IN SCIENCE AND TECHNOLOGY IN CHINA

# ADVANCED TOPICS IN SCIENCE AND TECHNOLOGY IN CHINA

Zhejiang University is one of the leading universities in China. In Advanced Topics in Science and Technology in China, Zhejiang University Press and Springer jointly publish monographs by Chinese scholars and professors, as well as invited authors and editors from abroad who are outstanding experts and scholars in their fields. This series will be of interest to researchers, lecturers, and graduate students alike.

Advanced Topics in Science and Technology in China aims to present the latest and most cutting-edge theories, techniques, and methodologies in various research areas in China. It covers all disciplines in the fields of natural science and technology, including but not limited to, computer science, materials science, life sciences, engineering, environmental sciences, mathematics, and physics.

此为试读,需要完整PDF请访问: www.ertongbook.com

# 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

# Contents

**1 Introduction** ..... **1**
1.1 Development of Artificial Intelligence ..... 1
1.2 Characteristics of Artificial Intelligent System ..... 5
1.3 Computational Intelligence ..... 9
1.3.1 Fuzzy Computing ..... 9
1.3.2 Neural Computing ..... 12
1.3.3 Evolutionary Computing ..... 12
1.3.4 Combination of the Three Branches ..... 15
1.4 Process Neural Networks ..... 16
References ..... 17

**2 Artificial Neural Networks** ..... **20**
2.1 Biological Neuron ..... 21
2.2 Mathematical Model of a Neuron ..... 22
2.3 Feedforward/Feedback Neural Networks ..... 23
2.3.1 Feedforward/Feedback Neural Network Model ..... 23
2.3.2 Function Approximation Capability of Feedforward Neural Networks ..... 25
2.3.3 Computing Capability of Feedforward Neural Networks ..... 27
2.3.4 Learning Algorithm for Feedforward Neural Networks ..... 28
2.3.5 Generalization Problem for Feedforward Neural Networks ..... 28
2.3.6 Applications of Feedforward Neural Networks ..... 30
2.4 Fuzzy Neural Networks ..... 32
2.4.1 Fuzzy Neurons ..... 32
2.4.2 Fuzzy Neural Networks ..... 33

2.5 Nonlinear Aggregation Artificial Neural Networks ..... 35
2.5.1 Structural Formula Aggregation Artificial Neural Networks ..... 35
2.5.2 Maximum (or Minimum) Aggregation Artificial Neural Networks ..... 35
2.5.3 Other Nonlinear Aggregation Artificial Neural Networks ..... 36
2.6 Spatio-temporal Aggregation and Process Neural Networks ..... 37
2.7 Classification of Artificial Neural Networks ..... 39
References ..... 40

**3 Process Neurons ..... 43**
3.1 Revelation of Biological Neurons ..... 43
3.2 Definition of Process Neurons ..... 44
3.3 Process Neurons and Functionals ..... 47
3.4 Fuzzy Process Neurons ..... 48
3.4.1 Process Neuron Fuzziness ..... 49
3.4.2 Fuzzy Process Neurons Constructed using Fuzzy Weighted Reasoning Rule ..... 50
3.5 Process Neurons and Compound Functions ..... 51
References ..... 52

**4 Feedforward Process Neural Networks ..... 53**
4.1 Simple Model of a Feedforward Process Neural Network ..... 53
4.2 A General Model of a Feedforward Process Neural Network ..... 55
4.3 A Process Neural Network Model Based on Weight Function Basis Expansion ..... 56
4.4 Basic Theorems of Feedforward Process Neural Networks ..... 58
4.4.1 Existence of Solutions ..... 59
4.4.2 Continuity ..... 62
4.4.3 Functional Approximation Property ..... 64
4.4.4 Computing Capability ..... 67
4.5 Structural Formula Feedforward Process Neural Networks ..... 67
4.5.1 Structural Formula Process Neurons ..... 68
4.5.2 Structural Formula Process Neural Network Model ..... 69
4.6 Process Neural Networks with Time-varying Functions as Inputs and Outputs ..... 71
4.6.1 Network Structure ..... 71

4.6.2 Continuity and Approximation Capability of the Model ..... 73
4.7 Continuous Process Neural Networks ..... 75
4.7.1 Continuous Process Neurons ..... 76
4.7.2 Continuous Process Neural Network Model ..... 77
4.7.3 Continuity, Approximation Capability, and Computing Capability of the Model ..... 78
4.8 Functional Neural Network ..... 83
4.8.1 Functional Neuron ..... 84
4.8.2 Feedforward Functional Neural Network Model ..... 85
4.9 Epilogue ..... 86
References ..... 87

**5 Learning Algorithms for Process Neural Networks ..... 88**
5.1 Learning Algorithms Based on the Gradient Descent Method and Newton Descent Method ..... 89
5.1.1 A General Learning Algorithm Based on Gradient Descent ..... 89
5.1.2 Learning Algorithm Based on Gradient-Newton Combination ..... 91
5.1.3 Learning Algorithm Based on the Newton Descent Method ..... 93
5.2 Learning Algorithm Based on Orthogonal Basis Expansion ..... 93
5.2.1 Orthogonal Basis Expansion of Input Functions ..... 94
5.2.2 Learning Algorithm Derivation ..... 95
5.2.3 Algorithm Description and Complexity Analysis ..... 96
5.3 Learning Algorithm Based on the Fourier Function Transformation ..... 97
5.3.1 Fourier Orthogonal Basis Expansion of the function in $L^2[0,2\pi]$ .... 97
5.3.2 Learning Algorithm Derivation ..... 99
5.4 Learning Algorithm Based on the Walsh Function Transformation ..... 101
5.4.1 Learning Algorithm Based on Discrete Walsh Function Transformation ..... 101
5.4.2 Learning Algorithm Based on Continuous Walsh Function Transformation ..... 105
5.5 Learning Algorithm Based on Spline Function Fitting ..... 108
5.5.1 Spline Function ..... 108
5.5.2 Learning Algorithm Derivation ..... 109
5.5.3 Analysis of the Adaptability and Complexity of a Learning Algorithm ..... 111

5.6 Learning Algorithm Based on Rational Square Approximation and Optimal Piecewise Approximation ..... 112
5.6.1 Learning Algorithm Based on Rational Square Approximation ... 112
5.6.2 Learning Algorithm Based on Optimal Piecewise Approximation ..... 119
5.7 Epilogue ..... 126
References ..... 126

**6 Feedback Process Neural Networks ..... 128**
6.1 A Three-Layer Feedback Process Neural Network ..... 129
6.1.1 Network Structure ..... 129
6.1.2 Learning Algorithm ..... 130
6.1.3 Stability Analysis ..... 132
6.2 Other Feedback Process Neural Networks ..... 135
6.2.1 Feedback Process Neural Network with Time-varying Functions as Inputs and Outputs ..... 135
6.2.2 Feedback Process Neural Network for Pattern Classification ..... 136
6.2.3 Feedback Process Neural Network for Associative Memory Storage ..... 137
6.3 Application Examples ..... 138
References ..... 142

**7 Multi-aggregation Process Neural Networks ..... 143**
7.1 Multi-aggregation Process Neuron ..... 143
7.2 Multi-aggregation Process Neural Network Model ..... 145
7.2.1 A General Model of Multi-aggregation Process Neural Network ..... 145
7.2.2 Multi-aggregation Process Neural Network Model with Multivariate Process Functions as Inputs and Outputs ..... 147
7.3 Learning Algorithm ..... 148
7.3.1 Learning Algorithm of General Models of Multi-aggregation Process Neural Networks ..... 148
7.3.2 Learning Algorithm of Multi-aggregation Process Neural Networks with Multivariate Functions as Inputs and Outputs ..... 152
7.4 Application Examples ..... 155
7.5 Epilogue ..... 159

此为试读,需要完整PDF请访问: www.ertongbook.com

References ..... 160

**8 Design and Construction of Process Neural Networks ..... 161**

8.1 Process Neural Networks with Double Hidden Layers ..... 161

8.1.1 Network Structure ..... 162

8.1.2 Learning Algorithm ..... 163

8.1.3 Application Examples ..... 165

8.2 Discrete Process Neural Network ..... 166

8.2.1 Discrete Process Neuron ..... 167

8.2.2 Discrete Process Neural Network ..... 168

8.2.3 Learning Algorithm ..... 169

8.2.4 Application Examples ..... 170

8.3 Cascade Process Neural Network ..... 172

8.3.1 Network Structure ..... 173

8.3.2 Learning Algorithm ..... 175

8.3.3 Application Examples ..... 176

8.4 Self-organizing Process Neural Network ..... 178

8.4.1 Network Structure ..... 178

8.4.2 Learning Algorithm ..... 179

8.4.3 Application Examples ..... 182

8.5 Counter Propagation Process Neural Network ..... 184

8.5.1 Network Structure ..... 185

8.5.2 Learning Algorithm ..... 185

8.5.3 Determination of the Number of Pattern Classifications ..... 186

8.5.4 Application Examples ..... 187

8.6 Radial-Basis Function Process Neural Network ..... 188

8.6.1 Radial-Basis Process Neuron ..... 188

8.6.2 Network Structure ..... 189

8.6.3 Learning Algorithm ..... 190

8.6.4 Application Examples ..... 192

8.7 Epilogue ..... 193

References ..... 193

**9 Application of Process Neural Networks ..... 195**

9.1 Application in Process Modeling ..... 195

9.2 Application in Nonlinear System Identification ..... 198

9.2.1 The Principle of Nonlinear System Identification ..... 199
9.2.2 The Process Neural Network for System Identification ..... 200
9.2.3 Nonlinear System Identification Process ..... 201
9.3 Application in Process Control ..... 203
9.3.1 Process Control of Nonlinear System ..... 204
9.3.2 Design and Solving of Process Controller ..... 204
9.3.3 Simulation Experiment ..... 208
9.4 Application in Clustering and Classification ..... 210
9.5 Application in Process Optimization ..... 215
9.6 Application in Forecast and Prediction ..... 216
9.7 Application in Evaluation and Decision ..... 224
9.8 Application in Macro Control ..... 226
9.9 Other Applications ..... 227
References ..... 231

**Postscript ..... 233**

**Index ..... 238**

# 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

此为试读,需要完整PDF请访问: www.ertongbook.com