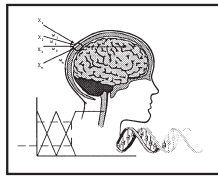


RECURRENT NEURAL NETWORKS

Design and Applications



Edited by

L.R. Medsker

Departments of Physics and Computer
Science and Information Systems
American University
Washington, D.C.

L.C. Jain

Knowledge-Based Intelligent Engineering Systems Centre
Faculty of Information Technology
Director/Founder, KES
University of South Australia, Adelaide
The Mawson Lakes, SA
Australia



CRC Press

Boca Raton London New York Washington, D.C.

PREFACE

Recurrent neural networks have been an interesting and important part of neural network research during the 1990's. They have already been applied to a wide variety of problems involving time sequences of events and ordered data such as characters in words. Novel current uses range from motion detection and music synthesis to financial forecasting. This book is a summary of work on recurrent neural networks and is exemplary of current research ideas and challenges in this subfield of artificial neural network research and development. By sharing these perspectives, we hope to illuminate opportunities and encourage further work in this promising area.

Two broad areas of importance in recurrent neural network research, the architectures and learning techniques, are addressed in every chapter. Architectures range from fully interconnected to partially connected networks, including recurrent multilayer feedforward. Learning is a critical issue and one of the primary advantages of neural networks. The added complexity of learning in recurrent networks has given rise to a variety of techniques and associated research projects. A goal is to design better algorithms that are both computationally efficient and simple to implement.

Another broad division of work in recurrent neural networks, on which this book is structured, is the design perspective and application issues. The first section concentrates on ideas for alternate designs and advances in theoretical aspects of recurrent neural networks. Some authors discuss aspects of improving recurrent neural network performance and connections with Bayesian analysis and knowledge representation, including extended neuro-fuzzy systems. Others address real-time solutions of optimization problems and a unified method for designing optimization neural network models with global convergence.

The second section of this book looks at recent applications of recurrent neural networks. Problems dealing with trajectories, control systems, robotics, and language learning are included, along with an interesting use of recurrent neural networks in chaotic systems. The latter work presents evidence for a computational paradigm that has higher potential for pattern capacity and boundary flexibility than a multilayer static feedforward network. Other chapters examine natural language as a dynamic system appropriate for grammar induction and language learning using recurrent neural networks. Another chapter applies a recurrent neural network technique to problems in controls and signal processing, and other work addresses trajectory problems and robot behavior.

The next decade should produce significant improvements in theory and design of recurrent neural networks, as well as many more applications for the creative solution of important practical problems. The widespread application of recurrent neural networks should foster more interest in research and development and raise further theoretical and design questions.

ACKNOWLEDGMENTS

The editors thank Dr. R. K. Jain, University of South Australia, for his assistance as a reviewer. We are indebted to Samir Unadkat and Mălina Ciocoiu for their excellent work formatting the chapters and to others who assisted: Srinivasan Guruswami and Aravindkumar Ramalingam. Finally, we thank the chapter authors who not only shared their expertise in recurrent neural networks, but also patiently worked with us via the Internet to create this book. One of us (L.M.) thanks Lee Giles, Ashraf Abelbar, and Marty Hagan for their assistance and helpful conversations and Karen Medsker for her patience, support, and technical advice.

THE EDITORS

Larry Medsker is a Professor of Physics and Computer Science at American University. His research involves soft computing and hybrid intelligent systems that combine neural network and AI techniques. Other areas of research are in nuclear physics and data analysis systems. He is the author of two books: *Hybrid Neural Network and Expert Systems* (1994) and *Hybrid Intelligent Systems* (1995). He co-authored with Jay Liebowitz another book on *Expert Systems and Neural Networks* (1994). One of his current projects applies intelligent web-based systems to problems of knowledge management and data mining at the U.S. Department of Labor. His Ph.D. in Physics is from Indiana University, and he has held positions at Bell Laboratories, University of Pennsylvania, and Florida State University. He is a member of the International Neural Network Society, American Physical Society, American Association for Artificial Intelligence, IEEE, and the D.C. Federation of Musicians, Local 161-710.

L.C. Jain is a Director/Founder of the Knowledge-Based Intelligent Engineering Systems (KES) Centre, located in the University of South Australia. He is a fellow of the Institution of Engineers Australia. He has initiated a postgraduate stream by research in the Knowledge-Based Intelligent Engineering Systems area. He has presented a number of keynote addresses at International Conferences on Knowledge-Based Systems, Neural Networks, Fuzzy Systems and Hybrid Systems. He is the Founding Editor-in-Chief of the *International Journal of Knowledge-Based Intelligent Engineering Systems* and served as an Associate Editor of the *IEEE Transactions on Industrial Electronics*. Professor Jain was the Technical chair of the ETD2000 International Conference in 1995, Publications Chair of the Australian and New Zealand Conference on Intelligent Information Systems in 1996 and the Conference Chair of the International Conference on Knowledge-Based Intelligent Electronic Systems in 1997, 1998 and 1999. He served as the Vice President of the Electronics Association of South Australia in 1997. He is the Editor-in-Chief of the International Book Series on Computational Intelligence, CRC Press USA. His interests focus on the applications of novel techniques such as knowledge-based systems, artificial neural networks, fuzzy systems and genetic algorithms and the application of these techniques.

Table of Contents

Chapter 1

Introduction

Samir B. Unadkat, Mălina M. Ciocoiu and Larry R. Medsker

- I. Overview
 - A. Recurrent Neural Net Architectures
 - B. Learning in Recurrent Neural Nets
- II. Design Issues And Theory
 - A. Optimization
 - B. Discrete-Time Systems
 - C. Bayesian Belief Revision
 - D. Knowledge Representation
 - E. Long-Term Dependencies
- III. Applications
 - A. Chaotic Recurrent Networks
 - B. Language Learning
 - C. Sequential Autoassociation
 - D. Trajectory Problems
 - E. Filtering And Control
 - F. Adaptive Robot Behavior
- IV. Future Directions

Chapter 2

Recurrent Neural Networks for Optimization:

The State of the Art

Youshen Xia and Jun Wang

- I. Introduction
- II. Continuous-Time Neural Networks for QP and LCP
 - A. Problems and Design of Neural Networks
 - B. Primal-Dual Neural Networks for LP and QP
 - C. Neural Networks for LCP
- III. Discrete-Time Neural Networks for QP and LCP
 - A. Neural Networks for QP and LCP
 - B. Primal-Dual Neural Network for Linear Assignment
- IV. Simulation Results
- V. Concluding Remarks

Efficient Second-Order Learning Algorithms for Discrete-Time Recurrent Neural Networks

Eurípedes P. dos Santos and Fernando J. Von Zuben

- I. Introduction
- II. Spatial x Spatio-Temporal Processing
- III. Computational Capability
- IV. Recurrent Neural Networks as Nonlinear Dynamic Systems
- V. Recurrent Neural Networks and Second-Order Learning Algorithms
- VI. Recurrent Neural Network Architectures
- VII. State Space Representation for Recurrent Neural Networks
- VIII. Second-Order Information in Optimization-Based Learning Algorithms
- IX. The Conjugate Gradient Algorithm
 - A. The Algorithm
 - B. The Case of Non-Quadratic Functions
 - C. Scaled Conjugate Gradient Algorithm
- X. An Improved SCGM Method
 - A. Hybridization in the Choice of β^j
 - B. Exact Multiplication by the Hessian
- XI. The Learning Algorithm for Recurrent Neural Networks
 - A. Computation of $\nabla E_T(w)$
 - B. Computation of $H(w)v$
- XII. Simulation Results
- XIII. Concluding Remarks

Chapter 4

Designing High Order Recurrent Networks for Bayesian Belief Revision

Ahsraf Abdelbar

- I. Introduction
- II. Belief Revision and Reasoning Under Uncertainty
 - A. Reasoning Under Uncertainty
 - B. Bayesian Belief Networks
 - C. Belief Revision
 - D. Approaches to Finding Map Assignments
- III. Hopfield Networks and Mean Field Annealing
 - A. Optimization and the Hopfield Network
 - B. Boltzmann Machine
 - C. Mean Field Annealing
- IV. High Order Recurrent Networks
- V. Efficient Data Structures for Implementing HORNs
- VI. Designing HORNs for Belief Revision
- VII. Conclusions

Chapter 5

Equivalence in Knowledge Representation: Automata, Recurrent Neural Networks, and Dynamical Fuzzy Systems

C. Lee Giles, Christian W. Omlin, and K. K. Thornber

- I. Introduction
 - A. Motivation
 - B. Background
 - C. Overview
- II. Fuzzy Finite State Automata
- III. Representation of Fuzzy States
 - A. Preliminaries
 - B. DFA Encoding Algorithm
 - C. Recurrent State Neurons with Variable Output Range
 - D. Programming Fuzzy State Transitions
- IV. Automata Transformation
 - A. Preliminaries
 - B. Transformation Algorithm
 - C. Example
 - D. Properties of the Transformation Algorithm
- V. Network Architecture
- VI. Network Stability Analysis
 - A. Preliminaries
 - B. Fixed Point Analysis for Sigmoidal Discriminant Function
 - C. Network Stability
- VII. Simulations
- VIII. Conclusions

Chapter 6

Learning Long-Term Dependencies in NARX Recurrent Neural Networks

Tsungnan Lin, Bill G. Horne, Peter Tino, and C. Lee Giles

- I. Introduction
- II. Vanishing Gradients and Long-Term Dependencies
- III. NARX Networks
- IV. An Intuitive Explanation of NARX Network Behavior
- V. Experimental Results
 - A. The Latching Problem
 - B. An Automaton Problem
- VI. Conclusion

Appendix

Oscillation Responses in a Chaotic Recurrent Network

Judy Dayhoff, Peter J. Palmadesso, and Fred Richards

- I. Introduction
- II. Progression to Chaos
 - A. Activity Measurements
 - B. Different Initial States
- III. External Patterns
 - A. Progression from Chaos to a Fixed Point
 - B. Quick Response
- IV. Dynamic Adjustment of Pattern Strength
- V. Characteristics of the Pattern-to-Oscillation Map
- VI. Discussion

Chapter 8

Lessons From Language Learning

Stefan C. Kremer

- I. Introduction
 - A. Language Learning
 - B. Classical Grammar Induction
 - C. Grammatical Induction
 - D. Grammars in Recurrent Networks
 - E. Outline
- II. Lesson 1: Language Learning Is Hard
- III. Lesson 2: When Possible, Search a Smaller Space
 - A. An Example: Where Did I Leave My Keys?
 - B. Reducing and Ordering in Grammatical Induction
 - C. Restricted Hypothesis Spaces in Connectionist Networks
 - D. Lesson 2.1: Choose an Appropriate Network Topology
 - E. Lesson 2.2: Choose a Limited Number of Hidden Units
 - F. Lesson 2.3: Fix Some Weights
 - G. Lesson 2.4: Set Initial Weights
- IV. Lesson 3: Search the Most Likely Places First
- V. Lesson 4: Order Your Training Data
 - A. Classical Results
 - B. Input Ordering Used in Recurrent Networks
 - C. How Recurrent Networks Pay Attention to Order
- VI. Summary

Chapter 9

Recurrent Autoassociative Networks: Developing Distributed Representations of Structured Sequences by Autoassociation

Ivelin Stoianov

- I. Introduction
- II. Sequences, Hierarchy, and Representations

- III. Neural Networks And Sequential Processing
 - A. Architectures
 - B. Representing Natural Language
- IV. Recurrent Autoassociative Networks
 - A. Training RAN With The Backpropagation Through Time Learning Algorithm
 - B. Experimenting with RANs: Learning Syllables
- V. A Cascade of RANs
 - A. Simulation With a Cascade of RANs: Representing Polysyllabic Words
 - B. A More Realistic Experiment: Looking for Systematicity
- VI. Going Further to a Cognitive Model
- VII. Discussion
- VIII. Conclusions

Chapter 10

Comparison of Recurrent Neural Networks for Trajectory Generation

David G. Hagner, Mohamad H. Hassoun, and Paul B. Watta

- I. Introduction
- II. Architecture
- III. Training Set
- IV. Error Function and Performance Metric
- V. Training Algorithms
 - A. Gradient Descent and Conjugate Gradient Descent
 - B. Recursive Least Squares and the Kalman Filter
- VI. Simulations
 - A. Algorithm Speed
 - B. Circle Results
 - C. Figure-Eight Results
 - D. Algorithm Analysis
 - E. Algorithm Stability
 - F. Convergence Criteria
 - G. Trajectory Stability and Convergence Dynamics
- VII. Conclusions

Chapter 11

Training Algorithms for Recurrent Neural Nets that Eliminate the Need for Computation of Error Gradients with Application to Trajectory Production Problem

Malur K. Sundareshan, Yee Chin Wong, and Thomas Condarcuru

- I. Introduction

- II. Description of the Learning Problem and Some Issues in Spatiotemporal Training
 - A. General Framework and Training Goals
 - B. Recurrent Neural Network Architectures
 - C. Some Issues of Interest in Neural Network Training
- III. Training by Methods of Learning Automata
 - A. Some Basics on Learning Automata
 - B. Application to Training Recurrent Networks
 - C. Trajectory Generation Performance
- IV. Training by Simplex Optimization Method
 - A. Some Basics on Simplex Optimization
 - B. Application to Training Recurrent Networks
 - C. Trajectory Generation Performance
- V. Conclusions

Chapter 12

Training Recurrent Neural Networks for Filtering and Control

Martin T. Hagan, Orlando De Jesús, and Roger Schultz

- I. Introduction
- II. Preliminaries
 - A. Layered Feedforward Network
 - B. Layered Digital Recurrent Network
- III. Principles of Dynamic Learning
- IV. Dynamic Backprop for the LDRN
 - A. Preliminaries
 - B. Explicit Derivatives
 - C. Complete FP Algorithms for the LDRN
- V. Neurocontrol Application
- VI. Recurrent Filter
- VII. Summary

Chapter 13

Remembering How To Behave: Recurrent Neural Networks for Adaptive Robot Behavior

T. Ziemke

- I. Introduction
- II. Background
- III. Recurrent Neural Networks for Adaptive Robot Behavior
 - A. Motivation
 - B. Robot and Simulator
 - C. Robot Control Architectures
 - D. Experiment 1
 - E. Experiment 2
- IV. Summary and Discussion