The 1991 "Neural Information Processing Systems-Natural and Synthetic" (NIPS) was held in Denver, Colorado, from 2-5 December 1991. Since its inception in 1987, the NIPS conference has attracted researchers from many disciplines who are applying their expertise to problems in the field of neural networks. The conference and the following two-day workshop have become a forum for presenting the latest research results and for leading researchers to gather and exchange ideas. The 1991 conference maintained the high level of excitement of its predecessors. Important new theoretical results were presented concerning the capability and generalization performance of networks.
FINAL TECHNICAL REPORT
FOR
YALE UNIVERSITY
AFOSR-91-0438
30 SEP 91 - 29 SEP 92
Capt. Steven Suddarth, Ph.D.
AFOSR/NE, Bldg. 410
Bolling Air Force Base
Washington, DC 20332

Dear Dr. Suddarth:

This letter and the attached materials constitute the final report for AFOSR Grant 91-0438 which provided $5,000 for student travel grants for the 1991 Neural Information Processing Systems Conference. The money was used to help 20 students as indicated in the attached list.

Also attached is a copy of the front matter of the proceedings which resulted from NIPS '91. As is evident, many of the students we were able to help made substantial contributions to the conference program. We are very grateful for your generous support and hope that you will be able to continue to support NIPS conferences in the future.

Sincerely,

John Moody
Associate Professor
NIPS'91 General Chairman
<table>
<thead>
<tr>
<th>Rank</th>
<th>$250</th>
<th>Name</th>
<th>Email</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td>Quanfeng Wu</td>
<td><a href="mailto:qw0w@andrew.cmu.edu">qw0w@andrew.cmu.edu</a></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>David Plaut</td>
<td><a href="mailto:dcp@cs.cmu.edu">dcp@cs.cmu.edu</a></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>John Hampshire</td>
<td><a href="mailto:hamps@speech2.cs.cmu.edu">hamps@speech2.cs.cmu.edu</a></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td>Frederick Waugh</td>
<td><a href="mailto:waugh@curly.harvard.edu">waugh@curly.harvard.edu</a></td>
</tr>
<tr>
<td>6</td>
<td></td>
<td>Lori Pratt</td>
<td><a href="mailto:pratt@paul.rutgers.edu">pratt@paul.rutgers.edu</a></td>
</tr>
<tr>
<td>7</td>
<td></td>
<td>Barak Pearlmutter</td>
<td><a href="mailto:barak@james.psych.yale">barak@james.psych.yale</a></td>
</tr>
<tr>
<td>8</td>
<td></td>
<td>Alexandre Pouget</td>
<td><a href="mailto:alex@cajal.edec.edu">alex@cajal.edec.edu</a></td>
</tr>
<tr>
<td>9</td>
<td></td>
<td>Steve Fisher</td>
<td><a href="mailto:fisher@helmholtz.sdsc.edu">fisher@helmholtz.sdsc.edu</a></td>
</tr>
<tr>
<td>10</td>
<td></td>
<td>Sowmya Ramachandran</td>
<td><a href="mailto:sowmya@cs.utexas.edu">sowmya@cs.utexas.edu</a></td>
</tr>
<tr>
<td>11</td>
<td></td>
<td>Andrew Moore</td>
<td><a href="mailto:awm@ai.mit.edu">awm@ai.mit.edu</a></td>
</tr>
<tr>
<td>12</td>
<td></td>
<td>Zoubin Ghahramani</td>
<td><a href="mailto:zoubin@psych.ee.mit.edu">zoubin@psych.ee.mit.edu</a></td>
</tr>
<tr>
<td>13</td>
<td></td>
<td>Ying Zhao</td>
<td><a href="mailto:yzhao@ai.mit.edu">yzhao@ai.mit.edu</a></td>
</tr>
<tr>
<td>14</td>
<td></td>
<td>David Cohn</td>
<td><a href="mailto:pablo@cs.washington.edu">pablo@cs.washington.edu</a></td>
</tr>
<tr>
<td>15</td>
<td></td>
<td>Steve Nowlan</td>
<td><a href="mailto:nowlan@helmholtz.sdsc.edu">nowlan@helmholtz.sdsc.edu</a></td>
</tr>
<tr>
<td>16</td>
<td></td>
<td>Sherif Botros</td>
<td><a href="mailto:smb@ai.mit.edu">smb@ai.mit.edu</a></td>
</tr>
<tr>
<td>17</td>
<td></td>
<td>Gary Scott</td>
<td><a href="mailto:scott@chewi.che.wisc.edu">scott@chewi.che.wisc.edu</a></td>
</tr>
<tr>
<td>18</td>
<td></td>
<td>Tony Bell</td>
<td><a href="mailto:tony@helmholtz.sdsc.edu">tony@helmholtz.sdsc.edu</a></td>
</tr>
<tr>
<td>20</td>
<td></td>
<td>Elizabeth Thomas</td>
<td><a href="mailto:cmjv465@hermes.chpc.utexas.edu">cmjv465@hermes.chpc.utexas.edu</a></td>
</tr>
<tr>
<td>22</td>
<td></td>
<td>Tony Zador</td>
<td><a href="mailto:zador@cs.yale.edu">zador@cs.yale.edu</a></td>
</tr>
<tr>
<td>23</td>
<td></td>
<td>Antonette Logar</td>
<td><a href="mailto:N0801@ttacs1.ttu.edu">N0801@ttacs1.ttu.edu</a></td>
</tr>
</tbody>
</table>
ADVANCES IN

NEURAL

INFORMATION

PROCESSING

SYSTEMS 4
OTHER TITLES OF INTEREST
FROM MORGAN KAUFMANN PUBLISHERS

NIPS-3-
Advances in Neural Information Processing Systems
Proceedings of the 1990 Conference
Edited by Richard P. Lippmann, John E. Moody, and David S. Touretzky

NIPS-2-
Advances in Neural Information Processing Systems
Proceedings of the 1989 Conference
Edited by David S. Touretzky

NIPS-1-
Advances in Neural Information Processing Systems
Proceedings of the 1988 Conference
Edited by David S. Touretzky

Computer Systems That Learn: Classification and Prediction Methods from Statistics, Neural Nets, Machine Learning, and Expert Systems
By Sholom M. Weiss and Casimir A. Kulikowski

Connectionist Models Summer School Proceedings
1990 Edited by David S. Touretzky, Jeffrey L. Elman, Terrence J. Sejnowski, and Geoffrey E. Hinton
1988 Edited by David S. Touretzky, Geoffrey E. Hinton, and Terrence J. Sejnowski

Learning Machines
By Nils J. Nilsson, with an Introduction by Terrence J. Sejnowski and Halbert White

Foundations of Genetic Algorithms
Edited by Gregory J.E. Rawlins

Genetic Algorithms: Proceedings of the Fourth International Conference
Edited by Rick Belew and Lashon Booker

Genetic Algorithms: Proceedings of the Third International Conference
Edited by David Schaffer

COLT—Proceedings of the Annual Workshops on Computational Learning Theory
1990 Edited by Mark Fulk and John Case
1989 Edited by Ron Rivest, Manfred Warmuth, and David Haussler
1988 Edited by David Haussler and Leonard Pitt
CONTENTS

Preface xv

Part I NEUROBIOLOGY

Models Wanted: Must Fit Dimensions of Sleep and Dreaming ................. 3
J. Allan Hobson, Adam N. Mamelak, and Jeffrey P. Sutton

Stationarity of Synaptic Coupling Strength Between Neurons with Nonstationary
Discharge Properties ............................................ 11
Mark Sydorenko and Eric D. Young

Perturbing Hebbian Rules ........................................ 19
Peter Dayan and Geoffrey Goodhill

Statistical Reliability of a Blowfly Movement-Sensitive Neuron ............ 27
Rob de Ruyter van Steveninck and William Bialek

The Clusteron: Toward a Simple Abstraction for a Complex Neuron ........ 35
Bartlet W. Mel

Network activity determines spatio-temporal integration in single cells .... 43
Oivind Bernander, Christof Koch, and Rodney J. Douglas

Nonlinear Pattern Separation in Single Hippocampal Neurons
with Active Dendritic Membrane .................................. 51
Anthony M. Zador, Brenda J. Claiborne, and Thomas H. Brown

Self-organisation in real neurons: Anti-Hebb in 'Channel Space'? .......... 59
Anthony J. Bell

Single Neuron Model: Response to Weak Modulation in the Presence of Noise ... 67
A.R. Bulsara, E.W. Jacobs, and F. Moss

Dual Inhibitory Mechanisms for Definition of Receptive Field Characteristics
in a Cat Striate Cortex .............................................. 75
A.B. Bonds
A comparison between a neural network model for the formation of brain maps and experimental data

K. Obermayer, K. Schulten, and G. G. Blasdel

Retinogeniculate Development: The Role of Competition and Correlated Retinal Activity

Ron Keesing, David G. Stork, and Carla J. Shatz

Part II NEURO-DYNAMICS

Locomotion in a Lower Vertebrate: Studies of the Cellular Basis of Rhythmogenesis and Oscillator Coupling

James T. Buchanan

Adaptive Synchronization of Neural and Physical Oscillators

Kenji Doya and Shuji Yoshizawa

Sustained Synchronization without Frequency Locking in a Completely Solvable Network Model

Heinz Schuster and Christof Koch

Oscillatory Model of Short Term Memory

David Horn and Marius Uther

Part III SPEECH

Multi-State Time Delay Neural Networks for Continuous Speech Recognition

Patrick Haffner and Alex Waibel

Modeling Applications with the Focused Gamma Net

Jose C. Principe, Bert de Vries, Jyh Ming Kuo, Pedro Guedes de Oliveira

Time-Warping Network: A Hybrid Framework for Speech Recognition

Esther Levin, Roberto Pieraccini, and Enrico Bocchieri

Improved Hidden Markov Model Speech Recognition Using Radial Basis Function Networks

Elliot Singer and Richard P. Lippmann

Connectionist Optimisation of Tied Mixture Hidden Markov Models

Steve Renals, Nelson Morgan, Hervé Bourlard, Horacio Franco, and Michael Cohen

Neural Network—Gaussian Mixture Hybrid for Speech Recognition or Density Estimation

Yoshua Bengio, Renato De Mori, Giovanni Flammia, Ralf Kompe

JANUS: Speech-to-Speech Translation Using Connectionist and Non-Connectionist Techniques

Alex Waibel, Ajay N. Jain, Arthur McNair, Joe Tebelskis, Louise Osterholz, Hiroaki Saito, Otto Schmidbauer, Tilo Sloboda, and Monika Woszczyna
Forward Dynamics Modeling of Speech Motor Control
Using Physiological Data
Makoto Hirayama, Eric Vatskios-Bateson,
Mitsu Kawato, and Michael I. Jordan

English Alphabet Recognition with Telephone Speech
Mark Fanty, Ronald A. Cole, and Krut Rogniski

Part IV LANGUAGE

Generalization Performance in PARSEC—A Structured Connectionist Parsing Architecture
Ajay N. Jain

Constructing Proofs in Symmetric Networks
Gadi Pinkas

A Connectionist Learning Approach to Analyzing Linguistic Stress
Prabhlad Gupta and David S. Touretzky

Propagation Filters in PDS Networks for Sequencing
and Ambiguity Resolution
Ronald A. Sumida and Michael G. Dyer

A Segment-based Automatic Language Identification System
ieshwant K. Muthusamy and Ronald A. Cole

Part V TEMPORAL SEQUENCES

The Efficient Learning of Multiple Task Sequences
Satinder P. Singh

Practical Issues in Temporal Difference Learning
Gerald Tesauru

HARMONET: A Neural Net for Harmonizing Chorales
in the Style of J.S. Bach
Hermann Hild, Johannes Feulner, and Wolfram Menzel

Induction of Multiscale Temporal Structure
Michael C. Mozer

Network Model of State-Dependent Sequencing
Jeffrey P. Sutton, Adam N. Mamelak, and J. Allan Hobson

Learning Unambiguous Reduced Sequence Descriptions
Jürgen Schmidhuber

Part VI RECURRENT NETWORKS

Recurrent Networks and NARMA Modeling
Jerome Connor, Les E. Atlas, and Douglas R. Martin
Induction of Finite-State Automata Using Second-Order Recurrent Networks
Raymond L. Watrous, and Gary M. Kuhn

309

Extracting and Learning an Unknown Grammar
with Recurrent Neural Networks

317

Operators and curried functions: Training and analysis
of simple recurrent networks
Janet Wiles and Anthony Bloesch

325

Green's Function Method for Fast On-line Learning Algorithm
of Recurrent Neural Networks
Guo-Zheng Sun, Hsing-Hen Chen, and Yee-Chun Lee

333

Dynamically-Adaptive Winner-Take-All Networks
Trent E. Lange

341

PartVII VISION

Information Processing to Create Eye Movements
David A. Robinson

351

Decoding of Neuronal Signals in Visual Pattern Recognition
Emad N. Eskandar, Barry J. Richmond, John A. Hertz,
Lance M. Optican, and Troels Kjaer

356

Learning How to Teach or Selecting Minimal Surface Data
David Geiger and Ricardo A. Marques Pereira

364

Learning to Make Coherent Predictions in Domains with Discontinuities
Suzanna Becker and Geoffrey E. Hinton

372

Recurrent Eye Tracking Network Using a Distributed Representation
of Image Motion
P.A. Viola, S.G. Lisberger, and T.J. Sejnowski

380

Against Edges: Function Approximation with Multiple Support Maps
Trevor Darrell and Alex Pentland

388

Markov Random Fields Can Bridge Levels of Abstraction
Paul R. Cooper and Peter N. Prokopowicz

396

Illumination and View Position in 3D Visual Recognition
Amnon Shashua

404

Hierarchical Transformation of Space in the Visual System
Alexandre Pouget, Stephen A. Fisher, and Terrence J. Sejnowski

412

VISIT: A Neural Model of Covert Visual Attention
Subutai Ahmad

420

Visual Grammars and their Neural Nets
Eric Mjolness

428
Learning to Segment Images Using Dynamic Feature Binding .......................... 436
  Michael C. Motzer, Richard S. Zemel, and Marlene Behrmann

Combined Neural Network and Rule-Based Framework
for Probabilistic Pattern Recognition and Discovery .................. 444
  Hayit K. Greenspan, Rodney Goodman, and Rama Chellappa

Linear Operator for Object Recognition ........................................... 452
  Ronen Basri and Shimon Ullman

3D Object Recognition Using Unsupervised Feature Extraction .......... 460
  Nathan Intrator, Josh I. Gold, Heinrich H. Bulthoff, and Shimon Edelman

Part VIII  OPTICAL CHARACTER RECOGNITION

Structural Risk Minimization for Character Recognition ................. 471
  I. Guyon, V. Vapnik, B. Boser, L. Bottou, and S.A. Solla

Image Segmentation with Networks of Variable Scales ................... 480
  Hans P. Graf, Craig R. Nohl, and Jan Ben

Multi-Digit Recognition Using a Space Displacement Neural Network .... 488
  Ofer Manas, Christopher J.C. Burges, Yann Le Cun, and John S. Denker

A Self-Organizing Integrated Segmentation and Recognition Neural Net ... 496
  Jim Keeler and David E. Rumelhart

Recognizing Overlapping Hand-Printed Characters
by Centered-Object Integrated Segmentation and Recognition .......... 504
  Gale L. Martin and Mosfeg Rashid

Adaptive Elastic Model: for Hand-Printed Character Recognition .......... 512
  Geoffrey E. Hinton, Christophe K.I. Williams, and Michael D. Reevow

Part IX  CONTROL AND PLANNING

Obstacle Avoidance through Reinforcement Learning ..................... 523
  Tony J. Prescott and John E.W. Mayhew

Active Exploration in Dynamic Environments ................................. 531
  Sebastian B. Thrun and Knut Möller

Oscillatory Neural Fields for Globally Optimal Path Planning ........... 539
  Michael Lemmon

Recognition of Manipulated Objects by Motor Learning ................... 547
  Hiroaki Gomi and Mitsuo Kawato

Refining PID Controllers using Neural Networks .......................... 555
  Gary M. Scott, Jude W. Shavlik, and W. Harmon Ray

Fast Learning with Predictive Forward Models ............................. 563
  Carlos Brody
<table>
<thead>
<tr>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fast, Robust Adaptive Control by Learning only Forward Models</td>
<td>571</td>
</tr>
<tr>
<td>Andrew W. Moore</td>
<td></td>
</tr>
<tr>
<td>Reverse TDNN: An Architecture for Trajectory Generation</td>
<td>579</td>
</tr>
<tr>
<td>Patrice Simard and Yann Le Cun</td>
<td></td>
</tr>
<tr>
<td>Learning Global Direct Inverse Kinematics</td>
<td>589</td>
</tr>
<tr>
<td>David DeMers and Kenneth Kreutz-Delgado</td>
<td></td>
</tr>
<tr>
<td>A Neural Net Model for Adaptive Control of Saccadic Accuracy</td>
<td>595</td>
</tr>
<tr>
<td>by Primate Cerebellum and Brainstem</td>
<td></td>
</tr>
<tr>
<td>Paul Dean, John E. W. Mayhew, and Pat Langdon</td>
<td></td>
</tr>
<tr>
<td>Learning in the Vestibular System: Simulations of Vestibular</td>
<td>603</td>
</tr>
<tr>
<td>Compensation</td>
<td></td>
</tr>
<tr>
<td>Thomas J. Anastasio</td>
<td></td>
</tr>
<tr>
<td>A Cortico-Cerebellar Model that Learns to Generate Distributed</td>
<td>611</td>
</tr>
<tr>
<td>Commands to Control a Kinematic Arm</td>
<td></td>
</tr>
<tr>
<td>N.E. Berthier, S.P. Singh, A.G. Barto, and J.C. Houk</td>
<td></td>
</tr>
<tr>
<td>A Computational Mechanism to Account for Averaged Modified</td>
<td>619</td>
</tr>
<tr>
<td>Hand Trajectories</td>
<td></td>
</tr>
<tr>
<td>Eitan A. Henis and Tamar Flash</td>
<td></td>
</tr>
<tr>
<td>Simulation of Optimal Movements Using the</td>
<td>627</td>
</tr>
<tr>
<td>Minimum-Muscle-Tension-Change Model</td>
<td></td>
</tr>
<tr>
<td>Menashe Dornay, Yoji Uno, Misuo Kawato, and Ryoji Suzuki</td>
<td></td>
</tr>
<tr>
<td>Part X APPLICATIONS</td>
<td></td>
</tr>
<tr>
<td>ANN Based Classification for Heart Defibrillators</td>
<td>637</td>
</tr>
<tr>
<td>M. Jabri, S. Pickard, P. Leong, Z. Chi, B. Flower, and Y. Xie</td>
<td></td>
</tr>
<tr>
<td>Neural Network Diagnosis of Avascular Necrosis</td>
<td>645</td>
</tr>
<tr>
<td>from Magnetic Resonance Images</td>
<td></td>
</tr>
<tr>
<td>Armando Manduca, Paul Christy, and Richard Ehman</td>
<td></td>
</tr>
<tr>
<td>Neural Network Analysis of Event: Related Potentials</td>
<td>651</td>
</tr>
<tr>
<td>and Electroencephalogram Predicts Vigilance</td>
<td></td>
</tr>
<tr>
<td>Rita Venterini, William W. Lytton, and Terrence J. Sejnowski</td>
<td></td>
</tr>
<tr>
<td>Neural Control for Rolling Mills: Incorporating Domain Theories</td>
<td>659</td>
</tr>
<tr>
<td>to Overcome Data Deficiency</td>
<td></td>
</tr>
<tr>
<td>Martin Röschisen, Reimar Hofmann, and Volker Tresp</td>
<td></td>
</tr>
<tr>
<td>Fault Diagnosis of Antenna Pointing Systems Using Hybrid Neural</td>
<td>667</td>
</tr>
<tr>
<td>Network and Signal Processing Models</td>
<td></td>
</tr>
<tr>
<td>Padhraic Smyth and Jeff Mellstrom</td>
<td></td>
</tr>
<tr>
<td>Multimodular Architecture for Remote Sensing Options</td>
<td>675</td>
</tr>
<tr>
<td>Sylvie Thiria, Carlos Mejia, Fouad Badran, Michel Crépon</td>
<td></td>
</tr>
</tbody>
</table>
Principled Architecture Selection for Neural Networks: Application to Corporate Bond Rating Prediction

John Moody and Joachim Ulans

683

Adaptive Development of Connectionist Decoders for Complex Error-Correcting Codes

Sheri L. Gish and Mario Blaum

691

Application of Neural Network Methodology to the Modelling of the Yield Strength in a Steel Rolling Plate Mill

Ah Chung Tsoi

698

Computer Recognition of Wave Location in Graphical Data by a Neural Network

Donald T. Freeman

706

A Neural Network for Motion Detection of Drift-Balanced Stimuli

Hilary Tunley

714

Neural Network Routing for Random Multistage Interconnection Networks

Mark W. Goudreau and C. Lee Giles

722

Networks for the Separation of Sources that are Superimposed and Delayed

John C. Platt and Federico Faggin

730

Part XI IMPLEMENTATION

CCD Neural Network Processors for Pattern Recognition

Alice M. Chiang, Michael L. Chuang, and Jeffrey R. LaFranchie

741

A Parallel Analog CCD/CMOS Signal Processor

Charles F. Neugebauer and Amnon Yariv

748

Direction Selective Silicon Retina that uses Null Inhibition

Ronald G. Benson and Tobi Delbruck

756

A Contrast Sensitive Silicon Retina with Reciprocal Synapses

Kwabena A. Boahen and Andreas G. Andreou

764

A Neurocomputer Board Based on the ANNA Neural Network Chip

Edward Sückinger, Bernhard E. Boser, and Lawrence D. Jackel

773

Software for ANN training on a Ring Array Processor

Phil Kohn, Jeff Bilmes, Nelson Morgan, and James Beck

781

Constrained Optimization Applied to the Parameter Setting Problem for Analog Circuits

David Kirk, Kurt Flescher, Lloyd Watts, and Alan Barr

789

Segmentation Circuits Using Constrained Optimization

John G. Harris

797
Analog LSI Implementation of an Auto-Adaptive Network for Real-Time Separation of Independent Signals
Marc H. Cohen, Philippe O. Pouliquen, and Andreas G. Andreou

Temporal Adaptation in a Silicon Auditory Nerve
John Lazzaro

Optical Implementation of a Self-Organizing Feature Extractor
Dana Z. Anderson, Claus Benkert, Verena Hebler, Ju-Seog Jang, Don Montgomery, and Mark Saffman

Part XII LEARNING AND GENERALIZATION

Principles of Risk Minimization for Learning Theory
V. Vapnik

Bayesian Model Comparison and Backprop Nets
David J.C. MacKay

The Effective Number of Parameters: An Analysis of Generalization and Regularization in Nonlinear Learning Systems
John E. Moody

Estimating Average-Case Learning Curves Using Bayesian, Statistical Physics and VC Dimension Methods
David Haussler, Michael Kearns, Manfred Opper, and Robert Schapire

Constant-Time Loading of Shallow 1-Dimensional Networks
Stephen Judd

Experimental Evaluation of Learning in a Neural Microsystem
Joshua Al恻ctor, Anthony Jayakumar, and Stephan Luna

Threshold Network Learning in the Presence of Equivalences
John Shawe-Taylor

Gradient Descent: Second Order Momentum and Saturating Error
Barak Pearlmutter

Tangent Prop--A formalism for specifying selected invariances in an adaptive network
Patrice Simard, Bernard Victorri, Yann Le Cun, and John Denker

Polynomial Uniform Convergence of Relative Frequencies to Probabilities
Alberto Bortoni, Paola Campadelli, Anna Morpurgo, and Sandra Ponzetta

Unsupervised learning of distributions on binary vectors using 2-layer networks
Yoav Freund and David Haussler

Incrementally Learning Time-varying Half-planes
Anthony Kuh, Thomas Petsche, and Ron L. Rivest
The VC-Dimension versus the Statistical Capacity of Multilayer Networks
Chuanyi Ji and Demetri Palioura

Some Approximation Properties of Projection Pursuit Learning Networks
Ying Zhao and Christopher G. Atkinson

Neural Computing with Small Weights
Kai-Yeung Siu and Jehoshua Bruck

A Simple Weight Decay Can Improve Generalization
Anders Krogh and John A. Hertz

Best-First Model Merging for Dynamic Learning and Recognition
Stephen M. Omohundro

Part XIII  ARCHITECTURES AND ALGORITHMS

Rule Induction through Integrated Symbolic and Subsymbolic Processing
Clayton McMillan, Michael C. Mozer, and Paul Smolensky

Interpretation of Artificial Neural Networks:
Mapping Knowledge-Based Neural Networks into Rules
Geoffrey Towell and Jude W. Shavlik

Hierarchies of adaptive experts
Michael I. Jordan and Robert A. Jacobs

Adaptive Soft Weight Tying using Gaussian Mixtures
Steven J. Nowlan and Geoffrey E. Hinton

Repeat Until Bored: A Pattern Selection Strategy
Paul W. Munro

Towards Faster Stochastic Gradient Search
Christian Darken and John Moody

Competitive Anti-Hebbian Learning of Invariants
Nicol N. Schraudolph and Terrence J. Sejnowski

Merging Constrained Optimisation with Deterministic Annealing
to "Solve" Combinatorially Hard Problems
Paul Stolorz

Kernel Regression and Backpropagation Training with Noise
Patri Koiranen and Lasse Holmstrom

Splines, Rational Functions and Neural Networks
Robert C. Williamson and Peter L. Bartlett

Networks with Learned Unit Response Functions
John Moody and Norman Yarvin

Learning in Feedforward Networks with Nonsmooth Functions
Nicholas J. Redding and T. Downs
PREFACE

This volume contains 144 papers summarizing the talks and posters presented at the fifth NIPS conference (short for "Neural Information Processing Systems—Natural and Synthetic"), held in Denver, Colorado, from 2–5 December 1991. Since its inception in 1987, the NIPS conference has attracted researchers from many disciplines who are applying their expertise to problems in the field of neural networks. The conference and the following two-day workshop have become a forum for presenting the latest research results and for leading researchers to gather and exchange ideas.

The 1991 conference maintained the high level of excitement of its predecessors. Important new theoretical results were presented concerning the capability and generalization performance of networks. Of particular interest are papers included in this volume by Vapnik, MacKay, Haussler, and others, which describe how to relate the complexity of networks to generalization performance on unseen test data. Many new network architectures were described. Some integrate expert system rules with networks, build hierarchies of networks, use radial basis function hidden nodes, and impose pre-specified invariance on the final solution. Neurobiological papers analyzed and modeled neurons in the hippocampus, in cat striate cortex, and in the blowfly. They also modeled biological networks that control eye movement, form topological maps, and compensate for head movement. Successful applications of neural networks were described in the areas of speech, vision, language, control, medical monitoring, and system diagnostics. Of particular interest was a paper by Tesauro, which demonstrated how a network could be trained to play backgammon at an expert level; papers by Jain, Warroux, and Giles, which described approaches to learning grammars; hybrid hidden-Markov-model/neural-network speech recognizers described by Haffner, Levin, Singer, Renals, and Bengio; papers on optical character recognition; a paper by Jabri, which describes a network to control a wearable heart defibrillator; a paper by Smyth for diagnosis of large-dish antenna pointing systems; and a paper by Roscheisen concerning control of force on rollers in steel rolling mills. Papers also described new analog and digital VLSI chips, systems for neural network implementation, and compared neural network and statistical approaches to pattern classification.
An historical milestone was reached this year, NIPS-91 was the fifth NIPS conference since the first conference was held in 1987. To mark this anniversary, we decided to review the history of events that led to the foundation of the NIPS conference and to discuss the evolution of the conference since its foundation. The following history is based in part on the recollections of Jim Bower, Larry Jackel, and Ed Posner. Some of this history was presented by Larry Jackel at the opening banquet.

While the first NIPS conference met in 1987, its origins can be traced back to the “Hopfest” meetings named in honor of John Hopsfield, held at Caltech. The first few, 1984-1986, were organized by Ed Posner of Caltech. These meetings met in the fall and included researches mainly from the Caltech campus and JPL. In 1985, Larry Jackel of Bell Labs and Demetri Psaltis of Caltech organized the first of what were to become the “Snowbird” meetings. The meetings were intended to be small informal workshops and convened in Santa Barbara. Twenty people were invited, but news of the meeting spread by word of mouth, so that attendance ended up growing to 60. In 1986, the meeting reconvened at Snowbird, which offered better snow conditions. Jackel, Psaltis, and the other organizers intended to keep the attendance down to 100 people, but the interest was so great that many people were turned away even after the attendance was capped at 160. The first Snowbird proceedings was edited by John Denker of Bell Labs and published by the American Institute of Physics (AIP) press.

In 1986, the Snowbird meeting was the only neural network conference, and it clearly could not accommodate the exploding numbers of researchers becoming interested in the field and still maintain the character of a small workshop. To respond to demand, the organizers decided to make Snowbird a more closed meeting, but to set in motion organization of a large meeting that would be open to all interested. The goal was to have a non-commercial meeting, dedicated to scholarship, which would capture some of the flavor of the workshop. The Snowbird organizers nominated a committee with Ed Posner as General Chairman and Yaser Abu Mostafa as Program Chairman (both of Caltech), to organize and run the 1987 NIPS conference, which was officially sponsored by the IEEE Information Theory Society. Denver was chosen as the site due to its central geographical location, ease of access by air, and close proximity to the mountains and the University of Colorado at Boulder.

The 1987 organizers designed the NIPS conference to have many of the advantages of a workshop, while still accommodating a large audience. To maximize scientific interchange, they decided to limit the oral presentations to a single stream, have posters be the majority of presentations, and include poster preview as well as formal poster sessions. Furthermore, a set of post-conference workshops was organized at the Copper Mountain ski resort after the main conference to enable small groups to discuss specific topics. The 1987 conference proved to be a great success, with about 450 attendees and 91 papers making it into the proceedings. Dana Z. Anderson of CU Boulder edited the proceedings, which were published by the AIP press and are now informally known as NIPS Volume 0.

Since 1987, some changes and refinements have been made, but the basic structure of the conference has remained the same. The NIPS 1988 proceedings (NIPS Volume 1, edited by David Toureszky of Carnegie Mellon) were the first published by Morgan Kaufmann. Also in 1988, the post-conference workshops were moved to Keystone, CO. The refinement processes (three reviewers instead of two), a more cross-disciplinary grouping of presenta-
tions, finer presentation categories, and the addition of five-minute oral poster spotlight presentations. A major and very successful addition to the 1991 conference was the introduction of a day of tutorials preceding the main conference. The 1991 workshops were held at Vail, which proved to be a popular move.

Finally, 1991 marked the drafting of articles of incorporation for the Neural Information Processing Systems Foundation, which will be responsible for the continuity of the NIPS conference in future years. The initial board of directors of the foundation consists of the 1987 to 1992 NIPS General Chairs (Ed Posner of Cal Tech, Terry Sejnowski of the Salk Institute and UCSD, Scott Kirkpatrick of IBM, Richard Lippmann of MIT Lincoln Labs, John Moody of Yale, and Stephen Hanson of Siemens), a member of the IEEE Information Theory Society (Terry Fine of Cornell), and our legal counsel (Philip Sotel).

The NIPS conference continues to be an exciting, successful meeting due to the efforts of a large group of people. We would first like to thank all the other members of the 1991 program and organizing committees who helped make this conference possible. In particular, we would like to thank Renate Crowley of Siemens for her extensive work throughout the year as the conference secretary and both Renate and Kate Fuqua of CU Boulder for running the conference desk so smoothly. Student contributions are an important part of the NIPS program, and we gratefully thank Tom McKenna of ONR and Steve Suddarth of AFOSR for the student travel funding provided by their agencies. Finally, we thank everyone who attended and submitted papers and the 105 referees who carefully read and reviewed 20 papers each.

John Moody
Stephen Hanson
Richard Lippmann
NIPS-91 ORGANIZING COMMITTEE

General Chair                      John Moody, Yale University
Program Chair                     Stephen Hanson, Siemens
Workshop CoChair                  Gerry Tesauro, IBM
Publicity Chair                   Scott Kirkpatrick, IBM
Publications Chair                John Pearson, David Sarnoff Research Lab
Treasurer                         Richard Lippmann, MIT Lincoln Laboratory
Government/Corporate Liaison       Bob Allen, Bellcore
Local Arrangements Chair          Lee Giles, NEC Research Institute
IEEE Liaisons                     Mike Mozer, University of Colorado
Tutorials Chair                   Rodney Goodman and Ed Posner, Caltech
APS Liaisons                      John Moody, Yale University
Neurobiology Liaisons             Eric Baum, NEC Research Institute
IEEE Liaisons                     Larry Jackel, AT&T Bell Labs
Overseas Liaison (Japan)          James Bower, Caltech
Overseas Liaison (Australia)      Tom Brown, Yale University
Overseas Liaison (United Kingdom) Rodney Goodman and Ed Posner, Caltech
Overseas Liaison (South America)  Maran Jabri, University of Sydney

NIPS-91 PROGRAM COMMITTEE

Program Chair                     Stephen Hanson, Siemens
Program CoChairs                  David Ackley, Bellcore
                                  Pierre Baldi, JPL and Caltech
                                  William Bialek, NEC Research Institute
                                  Lee Giles, NEC Research Institute
                                  Steve Omohundro, ICSI
                                  Mike Jordan, MIT
                                  John Platt, Synaptics
                                  Terry Sejnowski, Salk Institute
                                  David Stork, Ricoh and Stanford
                                  Alex Waibel, CMU

xviii
NIPS-91 Reviewers

Martin Brady, Penn. State Univ.
Gary Correll, UCSD
John Denker, AT&T Lab
John Herz, NORDITA
James Keeler, Advanced Computing Technology
Alan Lapedes, Los Alamos
David Rogers, RAACS
David G. Stork, Ricoh
Peter Todd, Stanford Univ.
Andreas Weigend, Stanford Univ.
Sue Becker, Univ. of Toronto
David A. Cohn, Univ. of Washington
Michael Littman, Bellcore
David Plaut, CMU
Lori Pratt, Rutgers Univ.
Jürgen Schmidhuber, Technische Univ. Muenchen
Rich Zemel, Univ. of Toronto
Paul Munro, Siemens
Ben Yuhas, Bellcore
Steve Gallant, Northeastern Univ.
Marwen Jabri, Sydney, Univ.
Gary M. Kuhn, CCRP-IDA
Y.C. Lee, Univ. of Maryland
Ken Marko, Ford Motor Company
Michiel Noordewier, Rutgers Univ.
Kevin Lang, NEC
Steven C. Saddarth, AFOSR
Robert Allen, Bellcore
Jonathan Bachrach
Gary Dell, Beckman Institute
Michael G. Dyer, UCLA
Jeff Elman, UCSD
Michael Gasser, UCSD
Lee Giles, NEC
Robert Jacobs, MIT
Stephen Hanson, Siemens
Christopher Atkeson, MIT
Andrew Baro, Univ. of Massachusetts
Judy Franklin, GTE
Vijaykumar Gullipalli, Univ. of Massachusetts
Thomas Martinez, Beckman Institute
Kumpati Narendra, Yale Univ.
Richard Sutton, GTE
Manoel Tenorio, Purdue Univ.
Lyle Ungar, Univ. of Pennsylvania
Josh Alspector, Bellcore
Jim Burr, Stanford Univ.
Federico Faggin, Synaptics
Hans Peter Graf, AT&T
Kristina Johnson, Univ. of Colorado
John Lazzaro, Univ. of Colorado
Dick Lyon, Apple Computer
Dan Schwartz, GTE
Larry Abbott, Brandeis Univ.
Thomas Anastasio, Univ. of Southern California
Joseph Artick, Princeton
William Bialek, Univ. of California at Berkeley
A.B. Bonds, Vanderbilt Univ.
James Bower, California Inst. of Tech.
Thomas Brown, Yale Univ.
Jack Cowan, Univ. of Chicago
Shawn Lockery, Salk Institute
Bartlett Mel, California Inst. of Tech.
Kenneth Miller, California Inst. of Tech.
Eric Schwartz, New York Univ. Medical Center
Terry Sejnowski, Salk Institute
Gordon Shepherd, New Haven
David van Essen, California Inst. of Tech.
Frank Fallside, Univ. of Cambridge
Hervé Bourlard, Phillips
Patrick Haafner, Centre National d'Etudes des Telecommunications