

AD-A212 697

# ABSTRACTS OF PAPERS

2

1987



IEEE

DTIC  
ELECTE  
AUG 11 1989  
S D  
CB

## IEEE CONFERENCE ON "NEURAL INFORMATION PROCESSING SYSTEMS — NATURAL AND SYNTHETIC"

November 8-12, 1987

Sheraton Denver Tech Center Hotel,  
Denver, Colorado

DISTRIBUTION STATEMENT A  
Approved for public release  
Distribution Unlimited

IEEE Catalog Number 87 CH 2386-1

REPORT DOCUMENTATION PAGE				Form Approved OMB No. 0704-0188	
1a. REPORT SECURITY CLASSIFICATION <b>UNCLASSIFIED</b>		1b. RESTRICTIVE MARKINGS -			
2a. SECURITY CLASSIFICATION AUTHORITY -		3. DISTRIBUTION/AVAILABILITY OF REPORT Approved for public use; distribution unlimited.			
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE -					
4. PERFORMING ORGANIZATION REPORT NUMBER(S) 87 CH 2386-1		5. MONITORING ORGANIZATION REPORT NUMBER(S) <b>AFOSR-TR-89-1061</b>			
6a. NAME OF PERFORMING ORGANIZATION IEEE (Information Theory Society)		6b. OFFICE SYMBOL (if applicable) -	7a. NAME OF MONITORING ORGANIZATION AFOSR/NE		
6c. ADDRESS (City, State, and ZIP Code) c/o E.C. Posner Caltech 116-81 Pasadena, CA 91125		7b. ADDRESS (City, State, and ZIP Code) AFOSR (AFSC) Bolling AFB, DC 20332-6448			
8a. NAME OF FUNDING/SPONSORING ORGANIZATION AFOSR/NE		8b. OFFICE SYMBOL (if applicable) -	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER AFOSR-88-0007		
8c. ADDRESS (City, State, and ZIP Code) AFOSR (AFSC) Electronic & Material Sciences-NE Bolling AFB, DC 20332-6448		10. SOURCE OF FUNDING NUMBERS			
		PROGRAM ELEMENT NO 61102F	PROJECT NO. 2305	TASK NO. B4	WORK UNIT ACCESSION NO.
11. TITLE (Include Security Classification) IEEE Conf. on "Neural Information Processing Systems - Natural and Synthetic"					
12. PERSONAL AUTHOR(S) E.C. Posner, General Chair					
13a. TYPE OF REPORT Abstract of Papers		13b. TIME COVERED FROM 11/1/87 TO 10/31/88	14. DATE OF REPORT (Year, Month, Day) 6/30/89		15. PAGE COUNT 55
16. SUPPLEMENTARY NOTATION Conference held Nov. 8-12, 1987					
17. COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number) Neural Networks, Connectionism, Parallel Distributed Processing		
FIELD	GROUP	SUB-GROUP			
19. ABSTRACT (Continue on reverse if necessary and identify by block number) The first IEEE Conference on "Neural Information Processing Systems - Natural and Synthetic" was held in Denver, Colorado, November 1987, with partial travel grant support from AFOSR and other agencies plus the private sector. Six hundred were in attendance. The travel of approximately 60 students was supported by grants as well as travel of about 10 young investigators. This meeting was widely publicized and was held again in 1988 and is being held in 1989. It brought together neural network researchers from various disciplines. A book was published based mostly on papers presented. Both theory and applications were covered.					
20. DISTRIBUTION/AVAILABILITY OF ABSTRACT <input checked="" type="checkbox"/> UNCLASSIFIED/UNLIMITED <input type="checkbox"/> SAME AS RPT <input type="checkbox"/> DTIC USERS			21. ABSTRACT SECURITY CLASSIFICATION		
22a. NAME OF RESPONSIBLE INDIVIDUAL DR. C. LEE GILES		22b. TELEPHONE (Include Area Code) 202-767-4932		22c. OFFICE SYMBOL NE	

89 8 10 066

(2)

**1987 IEEE CONFERENCE ON  
"NEURAL INFORMATION PROCESSING  
SYSTEMS -  
NATURAL AND SYNTHETIC"**

**NOVEMBER 8-12, 1987**

**SHERATON DENVER TECH CENTER HOTEL, DENVER,  
COLORADO**

**WITH A POST-MEETING WORKSHOP, NOVEMBER 12-15,  
COPPER MOUNTAIN RESORT, COLORADO**

*Sponsored by:*

The Institute of Electrical and Electronic Engineers, Co-sponsorship with the American Physical Society. A satellite meeting of the Society for Neuroscience. IEEE Group and Societies involved: Information Theory; Acoustics Speech and Signal Processing; Circuits and Systems. Financial support from the National Science Foundation, Bell Communications Research, US WEST Advanced Technologies, Office of Naval Research, Air Force Office of Scientific Research, Innovative Science and Technology Center (SDIO), and others.

**ORGANIZING COMMITTEE:**

Edward C. Posner	General Chairman
Yaser Abu-Mostafa	Program Chairman
Jawad Salehi	Treasurer
Clifford Lau	Publicity Chairman
Howard Wachtel	Local Arrangements Chairman
Dana Anderson	Publications Chairman
Larry Jackel	Physics Liaison
James Bower	Neurobiology Liaison

**S DTIC ELECTE D**  
**AUG 11 1989**  
**B**

**DISTRIBUTION STATEMENT A**

Approved for public release;  
Distribution Unlimited

IEEE Catalog Number 87CH2386-1  
Library of Congress Number 86-82857

#### INTERDISCIPLINARY ADVISORY COMMITTEE:

Shun-ich Amari	Tokyo
James A. Anderson	Brown
Leon N. Cooper	Brown
G. Dreyfuss	Paris
Scott Fraser	Irvine
Kunihiko Fukushima	Tokyo
Stephen Grossberg	Boston University
Robert Hecht-Nielsen	HNC
Geoffrey Hinton	Carnegie Mellon
John Hopfield	Caltech
Scott Kirkpatrick	IBM
Carver Mead	Caltech
John P. Miller	Berkeley
G. Palm	Max-Planck
Tomaso Poggio	MIT
Terry Sejnowski	Johns Hopkins
David W. Tank	Bell Labs
Roger D. Traub	IBM

#### PROGRAM COMMITTEE:

Y.S. Abu-Mostafa, Caltech  
C.H. Anderson, JPL  
D.Z. Anderson, U of Colorado  
P. Baldi, UC San Diego  
D.H. Ballard, U of Rochester  
J.M. Bower, Caltech  
G. Carpenter, Northeastern  
C. Elbaum, Brown  
W.J. Freeman, UC Berkeley  
A. Gelperin, AT&T Bell Labs  
C.L. Giles, AFOSR  
W.F. Heiligenberg, UC San Diego  
R. Linsker, IBM Watson  
R. Llinas, NYU  
J.L. McClelland, Carnegie Mellon  
D. Sabbah, IBM Watson  
J. Sage, MIT Lincoln Labs  
A.I. Selverston, UC San Diego  
R.F. Thompson, Stanford  
D.S. Touretzky, Carnegie Mellon  
D.C. Van Essen, Caltech  
W.B. Veldkamp, MIT Lincoln Labs  
S.S. Venkatesh, U of Pennsylvania

PROGRAM OUTLINE

Sunday, November 8, 1987

6:00 PM: Registration and Reception

Monday, November 9, 1987

8:45-9:00 AM: Opening Remarks  
 9:00-11:00 AM: Plenary Session - *Carver Mead and Terry Sejnowski*  
 11:00 AM-12:15 PM: Poster Session P1  
  
 1:30-3:10 PM: Oral Session O1 - *Biological Models*  
 3:40-5:20 PM: Oral Session O2 - *Mathematical Models*  
  
 5:30 PM: University of Colorado Reception

Tuesday, November 10, 1987

8:30-11:00 AM: Oral Session O3 - *Learning Theory*  
 11:00 AM-12:15 PM: Poster Session P2  
  
 1:30-3:10 PM: Oral Session O4 - *Learning Techniques*  
 3:40-5:20 PM: Oral Session O5 - *Applications*  
  
 7:00 PM: Banquet - *Banquet address by Leon Cooper*

Wednesday, November 11, 1987

8:30-11:00 AM: Oral Session O6 - *Computer Simulations in Neurobiology*  
  
 11:00 AM-12:15 PM: Poster Session P3  
  
 1:30-3:10 PM: Oral Session O7 - *Simulation and Implementation*  
 3:40-5:20 PM: Oral Session O8 - *Technology*

Thursday, November 12, 1987

8:30-11:00 AM: Oral Session O9 - *Analysis of Neural Networks*  
 11:00 AM-12:15 PM: Poster Session P4

Accession For	
NTIS GRA&I	<input checked="" type="checkbox"/>
DTIC TAB	<input type="checkbox"/>
Unannounced	<input type="checkbox"/>
Justification	
By _____	
Distribution/	
Availability Codes	
Dist	Avail and/or Special
A-1	



## **POST-MEETING WORKSHOP**

November 12 - 15

Copper Mountain Resort

The outline of the workshop schedule is as follows:

### **THURSDAY**

7:00 PM Opening remarks and introduction to discussion topics by moderators.

9:00 PM Forming of first-day groups.

### **FRIDAY**

8:30 Parallel group discussions (morning session).

10:30 Informal discussions and activities.

4:30 Parallel group discussions (afternoon session).

6:30 Dinner

8:00 First-day moderators report on discussion results.

9:30 Forming of second-day groups.

### **SATURDAY**

8:30 Parallel group discussions (morning session).

10:30 Informal discussions and activities.

4:30 Parallel group discussions (afternoon session).

6:30 Dinner

8:00 Second-day moderators report on discussion results.

### **SUNDAY**

Informal discussions and activities.

**ONLY THOSE PAPERS WHICH HAD THEIR ABSTRACTS READY AT THE TIME OF PREPARING THIS BOOK APPEAR HERE. REFER TO THE CONFERENCE PROGRAM FOR THE COMPLETE LIST OF PAPERS.**

## TABLE OF CONTENTS

### POSTER SESSION P1

THE PERFORMANCE OF CONVEX SET PROJECTION BASED NEURAL NETWORKS, Robert J. Marks II, Les E. Atlas and James A. Ritcey	1
A DYNAMICAL APPROACH TO TEMPORAL PATTERN PROCESSING, W. Scott Stornetta, Tad Hogg and B.A. Huberman	1
A NEURAL MODEL FOR PERCEPTUAL STABILITY, Bartlett W. Mel	1
MECHANISMS OF SELF-ORGANIZATION FOR INFORMATION PROCESSING BY DEVICES MODELED ON CEREBRAL CORTEX, Walter J. Freeman	2
SHIFTERS AND ADAPTIVE FILTERS-POSSIBLE FUNCTIONS OF CORTICAL LAYER 4, Charles H. Anderson and David C. Van Essen	2
CENTRIC MODELS OF THE ORIENTATION MAP IN PRIMATE VISUAL CORTEX, Bruce Dow and William Baxter	3
QUANTITATIVE RECONSTRUCTIONS OF COMMAND NEURONS AND MOTONEURONS IN THE TAILFLIP ESCAPE CIRCUIT OF CRAYFISH, Donald H. Edwards, Jr.	3
HIGH DENSITY ASSOCIATIVE MEMORY, A. Dembo and O. Zeitouni	3
DISTRIBUTED NEURAL PROCESSING IN THE VESTIBULO-OCULAR SYSTEM, Clifford Lau and Vicente Honrubia	4
THE HOPFIELD MODEL WITH MULTILEVEL NEURONS, M. Fleisher and E. Levin	4
SIMPLIFIED NEURAL NET ARCHITECTURES, John Moody	4
ASSOCIATIVE MEMORY DESIGN WITH HIGH-ORDER NEURAL NETWORKS, L. Personnaz, I. Guyon and G. Dreyfus	5
REFLEXIVE ASSOCIATIVE MEMORIES, Hendricus G. Loos	5
PHASOR NEURAL NETWORKS, Andre J. Noest	5
A MEAN FIELD THEORY OF LAYER IV, OF THE VISUAL CORTEX AND ITS APPLICATION TO ARTIFICIAL NEURAL NETWORKS, Christopher L. Scofield	6
EXTEND CONNECTIONIST NETWORK MODELS TO DO SEQUENTIAL REASONING, Y.C. Lee, H.H. Chen, G.Z. Sun, H.Y. Lee, Tom Maxwell, and C. Lee Giles	6

SPATIAL ORGANIZATION OF NEURAL NETWORKS: A PROBABILISTIC MODELING APPROACH, Andreas Stafylopatis, Marios Dikaiakos and Dimitris Kontoravdis 7

## ORAL SESSION 01 - BIOLOGICAL MODELS

BIFURCATION ANALYSIS OF A NETWORK MODEL OF RABBIT OLFACTORY BULB WITH PERIODIC ATTRACTORS STORED BY A SEQUENCE LEARNING ALGORITHM, Bill Baird 9

MODEL OF THE CEREBELLUM AS AN ARRAY OF ADJUSTABLE PATTERN GENERATORS, James C. Houk 9

SPONTANEOUS AND INFORMATION-TRIGGERED SEGMENTS OF SERIES OF HUMAN BRAIN ELECTRIC FIELD MAPS, D. Lehmann, D. Brandeis, H. Ozaki, and I. Pal 9

TOWARDS AN ORGANIZING PRINCIPLE FOR PERCEPTION: THE ROLE OF HEBBIAN SYNAPSES IN THE EMERGENCE OF FEATURE-ANALYZING FUNCTION, Ralph Linsker 10

## ORAL SESSION 02 - MATHEMATICAL MODELS

SYNCHRONIZATION IN NEURAL NETS, Jacques J. Vidal and John Haggerty 12

NEW 'NEURAL' ALGORITHMS FOR ASSOCIATIVE MEMORY, Eric B. Baum, John Moody and Frank Wilczek 12

A NEURAL NETWORK CLASSIFIER BASED ON CODING THEORY, Tzi-Dar Chiuch and Rodney Goodman 13

A TRELIS-STRUCTURED NEURAL NETWORK, Thomas Petsche and Bradley W. Dickinson 13

CONSTRAINED NETWORK REPRESENTATIONS, John Platt and Alan Barr 13

## ORAL SESSION 03 - LEARNING THEORY

MINKOWSKI-R BACK-PROPAGATION: LEARNING IN CONNECTIONIST MODELS WITH NON-EUCLIDIAN ERROR SIGNALS, Stephen Jose Hanson and David J. Burr 15

GENERALIZATION OF BACKPROPAGATION TO RECURRENT AND HIGH-ORDER NETWORKS, Fernando J. Pineda 15

PARTITIONING OF SENSORY DATA BY A CORTICAL NETWORK, R. Granger, H. Henry and G. Lynch 16

THE GENERATION OF EFFICIENT REPRESENTATIONS IN NEURAL NET ARCHITECTURES USING HIGH ORDER CORRELATIONS, Tom Maxwell, C. Lee Giles and Y.C. Lee 16

## POSTER SESSION P2

LEARNING STEREOPSIS IN A HIGH-ORDER NEURAL NETWORK, H.H. Chen, G.Z. Sun and Y.C. Lee	19
LEARNING BY STATE RECURRENCE, Bruce E. Rosen, James M. Goodwin and Jacques J. Vidal	19
MANUAL TRAINING TECHNIQUES OF AUTONOMOUS SYSTEMS BASED ON ARTIFICIAL NEURAL NETWORKS, J. Shepanski and S. Macy	19
TIME-SEQUENTIAL SELF-ORGANIZATION OF HIERARCHICAL NEURAL NETWORKS, A.S. Noetzel and R.H. Silverman	20
SUPERVISED LEARNING OF PROBABILITY DISTRIBUTIONS BY NEURAL NETWORKS, Eric B. Baum and Frank Wilczek	20
STATIC AND DYNAMIC ERROR PROPAGATION NETWORKS WITH APPLICATION TO SPEECH CODING, A.J. Robinson and F. Fallside	20
DISCOVERING INPUT STRUCTURE USING FEED-FORWARD NETWORKS, K. Prazdny	20
AN ARTIFICIAL NEURAL NETWORK FOR SPATIO-TEMPORAL BINARY PATTERNS: APPLICATION TO PHONEME CLASSIFICATION, Toshiteru Homma, Les E. Atlas and Robert J. Marks II	21
HIERARCHICAL LEARNING CONTROL - AN APPROACH WITH NEURON- LIKE ASSOCIATIVE MEMORY, E. Ersu, St. Gehlen and H. Tolle	21
LEARNING ON A GENERAL NETWORK, Amir F. Atiya	22
A FRAMEWORK FOR SPEECH RECOGNITION USING NEURAL NETWORKS, Richard P. Lippmann and William Y. Huang	22
LEARNING IN NETWORKS OF NONDETERMINISTIC ADAPTIVE LOGIC ELEMENTS, Richard C. Windecker	22
NEUROMORPHIC NETWORKS FOR OPTICAL COMMUNICATIONS AND SIGNAL PROCESSING, M.P. Vecchi and J.A. Salehi	23
ROTATION AND SCALE INVARIANT OBJECT RECOGNITION USING A DISTRIBUTED ASSOCIATIVE MEMORY, Harry Wechsler and George Lee Zimmerman	23
AN IMPROVED NEURAL-NETWORK ANALOG-TO-DIGITAL CONVERTER, R.A. Games, E.L. Key and D. Moulin	23

## **ORAL SESSION 04 - LEARNING TECHNIQUES**

- LEARNING REPRESENTATIONS BY RECIRCULATION, Geoffrey E. Hinton, James L. McClelland and Geoffrey J. Goodhill 25
- STRATEGIES FOR TEACHING LAYERED NETWORKS CLASSIFICATION TASKS, Ben S. Wittner, John S. Denker, L.D. Jackel, and R.E. Howard 25
- STOCHASTIC LEARNING NETWORKS AND THEIR ELECTRONIC IMPLEMENTATION, Joshua Alspector, Robert B. Allen, Victor Hu, and Srinagesh Satyanarayana 25
- SPEECH STABILIZATION AND ROBUST FRONT ENDS FOR NEURAL NETWORKS, D.J. Burr 26

## **ORAL SESSION 05 - APPLICATIONS**

- A 'NEURAL' NETWORK THAT LEARNS TO PLAY BACKGAMMON, G. Tesauro and T.J. Sejnowski 28
- GENETIC DATABASE ANALYSIS WITH NEURAL NETS, Alan Lapedes and Robert Farber 28
- LEARNING A COLOR ALGORITHM FROM EXAMPLES, A. Hurlbert and T. Poggio 28
- COMPUTING MOTION USING NEURAL NETWORKS, Christof Koch and James Hutchinson 29
- DYNAMICAL NEURAL NETWORKS AND THEIR APPLICATION TO ROBOT CONTROL, Athanasios Sideris, Alan Yamamura and Demetri Psaltis 29

## **ORAL SESSION 06 - COMPUTER SIMULATIONS IN NEUROBIOLOGY**

- SIMULATIONS SUGGEST INFORMATION PROCESSING ROLES FOR THE DIVERSE CURRENTS FOUND IN HIPPOCAMPAL NEURONS, Lyle J. Borg-Graham 31
- A COMPUTER SIMULATION OF CEREBRAL NEOCORTEX: COMPUTATIONAL CAPABILITIES OF NONLINEAR NETWORKS, A. Singer and J.P. Donoghue 31
- AN ADAPTIVE AND HETERODYNE FILTERING PROCEDURE FOR THE IMAGING OF MOVING OBJECTS, H.A.K. Mastebroek, F.H. Schuling and W.H. Zaagman 31

## POSTER SESSION P3

- ON TROPICSTIC PROCESSING AND ITS APPLICATIONS,  
Manuel F. Fernandez 34
- NEW MULTI-NEURON RECORDING TECHNIQUES, M.E. Nelson,  
B. Rasnow, J.J. Banik, Y.F. Wong, A. Atiya, and  
J.M. Bower 34
- COMPUTER SIMULATIONS OF DYNAMIC FUNCTIONAL CHANGES  
IN THE TOPOGRAPHIC MAPS OF ADULT ANIMALS,  
Leif H. Finkel and Gerald M. Edelman 34
- DISCOVERING STRUCTURE FROM VISUAL MOTION IN MONKEY,  
MAN AND MACHINE, M. Siegel 35
- COMPETITIVE LEARNING, UNLEARNING AND FORGETTING IN  
OPTICAL RESONATORS, Dana Z. Anderson and Michael J.  
O'Callaghan 25
- INTRODUCTION TO A SYSTEM FOR IMPLEMENTING NEURAL  
NETWORKS ON SIMD ARCHITECTURES, Sherryl Tombouliau 35
- THE SIGMOID NONLINEARITY IN OLFACTORY SYSTEM OF RAT,  
F.H. Eeckman 36
- PULSED NEURAL NETWORKS: HARDWARE, SOFTWARE, AND THE  
HOPFIELD A/D CONVERTER EXAMPLE, Martin E. Gaspar 36
- USING NEURAL NETWORKS TO IMPROVE COCHLEAR IMPLANT  
SPEECH PERCEPTION, Manoel F. Tenorio 36
- PRESYNAPTIC NEURAL INFORMATION PROCESSING, L. Richard  
Carley 37
- ADAPTATIONS OF THE ERROR PROPAGATION LEARNING  
ALGORITHM FOR VLSI IMPLEMENTATION, Tom Baker, Dan  
Hammerstrom and Hal McCartor 37
- OPTIMAL NEURAL SPIKE CLASSIFICATION, Amir Atiya and  
James Bower 37
- A NEURAL-NETWORK SOLUTION TO A CONCENTRATOR ASSIGNMENT  
PROBLEM, Gene A. Tagliarini and Edward W. Page 38
- HOW THE CATFISH TRACKS ITS PREY: GUSTO-MODULATED  
RETICULOSPINAL NEURONS MAY GUIDE NAVIGATIONAL  
BEHAVIOR IN THE CHANNEL CATFISH, *ICTALURUS PUNCTATUS*,  
Jagmeet S. Kanwal 38
- CYCLES: A SIMULATION TOOL FOR STUDYING CYCLIC NEURAL  
NETWORKS, Michael T. Gately 38

SYNAPTIC INTERACTIONS BETWEEN NEURONS: FUNCTIONAL SIGNIFICANCE OF CORRELATIONAL LINKAGES, E.E. Fetz 39

## ORAL SESSION 07 - SIMULATION AND IMPLEMENTATION

BASINS OF ATTRACTION FOR ELECTRONIC NEURAL NETWORKS, C.M. Marcus and R.M. Westervelt 41

TEMPORAL PATTERNS OF ACTIVITY IN NEURAL NETWORKS, Paolo Gaudiano 41

OPTIMAL MAPPING OF THE CONFIGURATION SPACE OF A NETWORK OF NEURONS, G.J. Carman, B. Rasnow and J.M. Bower 42

NEURAL NET AND CONVENTIONAL CLASSIFIERS, William Y. Huang and Richard P. Lippmann 42

NEURAL NETWORK IMPLEMENTATION APPROACHES FOR THE CONNECTION MACHINE, Nathan H. Brown, Jr. 42

## ORAL SESSION 08 - TECHNOLOGY

MICROELECTRONIC IMPLEMENTATION OF CONNECTIONIST NEURAL NETWORK MODELS, H.P. Graf, W.S. Mackie and J.S. Denker 44

A PROGRAMMABLE BINARY SYNAPTIC MATRIX CHIP FOR ELECTRONIC NEURAL NETWORKS, A. Moopenn, H. Langenbacher, A.P. Thakoor, and S.K. Khanna 44

BIT - SERIAL VLSI NEURAL NETWORKS, Alan F. Murray, Anthony V.W. Smith and Zoe F. Butler 44

NEW HARDWARE FOR MASSIVE NEURAL NETWORKS, A.G.U. Perera and D.D. Coon 45

## ORAL SESSION 09 - ANALYSIS OF NEURAL NETWORKS

ON PROPERTIES OF NETWORKS OF NEURON-LIKE ELEMENTS, Santosh S. Venkatesh and Pierre Baldi 47

PHASE TRANSITIONS IN NEURAL NETWORKS, Joshua Chover 47

ANALYSIS OF DISTRIBUTED REPRESENTATION OF CONSTITUENT STRUCTURE IN CONNECTIONIST SYSTEMS, Paul Smolensky 47

PATTERN CLASS DEGENERACY IN AN UNRESTRICTED STORAGE DENSITY MEMORY, Christopher L. Scofield, Douglas L. Reilly, Charles Baum, and Leon N. Cooper 48

ON THE POWER OF NEURAL NETWORKS FOR SOLVING HARD PROBLEMS, Jehosua Bruck and Joseph W. Goodman 48

## POSTER SESSION P4

SCALING PROPERTIES OF COARSE-CODED SYMBOL MEMORIES, Ronald Rosenfeld and David S. Touretzky	50
SELF-ORGANIZATION OF ASSOCIATIVE DATA-BASE AND ITS APPLICATIONS, Hisashi Suzuki and Suguru Arimoto	50
MEMORY CAPACITY IN SYMMETRIC NEURAL NETWORKS: RIGOROUS BOUNDS, Charles M. Newman	50
STABILITY RESULTS FOR NEURAL NETWORKS, A.N. Michel, W. Porod and J.A. Farrell	51
ANALYSIS AND COMPARISON OF DIFFERENT LEARNING ALGORITHMS FOR PATTERN ASSOCIATION PROBLEMS, J. Bernasconi	51
THE CONNECTIVITY REQUIREMENTS OF SIMPLE ASSOCIATION -OR- HOW MANY CONNECTIONS DO YOU NEED?, Dan Hammerstrom	51
MAPPING OPTIMIZATION PROBLEMS ONTO ARTIFICIAL NEURAL NETWORK SYSTEMS: A GENERAL PRINCIPLE AND A COMPARISON TO GRADIENT DESCENT (ASCENT), Harrison MonFook Leong	52
CONNECTING TO THE PAST, Bruce A. MacDonald	52
THE CAPACITY OF THE KANERVA ASSOCIATIVE MEMORY, P.A. Chou and R.M. Gray	52
CONNECTIVITY VERSUS ENTROPY, Yaser S. Abu-Mostafa	53
PROBABILISTIC CHARACTERIZATION OF NEURAL MODEL COMPUTATIONS, Richard Golden	53
PERFORMANCE MEASURES FOR ASSOCIATIVE MEMORIES THAT LEARN AND FORGET, Anthony Kuh	53
INFORMATION STORAGE CAPACITY OF CONNECTIONIST SYSTEMS: THE LINEAR ASSOCIATOR, Dean C. Mumme and Walter Schnelder	54
'ENSEMBLE' BOLTZMANN UNITS HAVE COLLECTIVE COMPUTATIONAL PROPERTIES LIKE THOSE OF GRADED HOPFIELD NEURONS, Mark A. Derthick	54
MATHEMATICAL ANALYSIS OF LEARNING BEHAVIOR OF NEURONAL MODELS, Massoud Omidvar and John Y. Cheung	54
AN OPTIMIZATION NETWORK FOR MATRIX INVERSION, Ju-Seog Jang, Soo-Young Lee and Sang-Yung Shin	55
INFORMATION CAPACITY OF HEBBIAN NEURAL NETWORKS, James D. Keeler	55

## POSTER SESSION P1

### THE PERFORMANCE OF CONVEX SET PROJECTION BASED NEURAL NETWORKS,

ROBERT J. MARKS II, LES E. ATLAS AND JAMES A. RITCEY,  
Interactive Systems Design Laboratory, University of Washington, FT-10, Seattle,  
WA 98195, USA.

Convex set projection theory has been shown to be a potentially rich source of material for formulating neural networks. The performance of continuous level associative memory neural networks based on convex set projection theory can be predicted analytically. In either the extrapolation or association mode, the network performs exactly under relatively loose conditions. When asked to retrieve memory data not previously stored, the network can respond negatively and, if appropriate, go into a learning mode. Learning can be done using a modified Gram-Schmit procedure with possible nonlinearity modifications made at the neural level. Application of similar neural networks to signal classification is also considered.

### A DYNAMICAL APPROACH TO TEMPORAL PATTERN PROCESSING,

W. SCOTT STORNETTA, Physics Department, Stanford University, Stanford, CA 94305, USA, and TAD HOGG and B.A. HUBERMAN, Xerox Palo Alto Research Center, Palo Alto, CA 94304, USA.

We discuss those issues involved in massively parallel network processing of patterns with temporal context which are distinct from static pattern processing. These include the internal representation of temporal relations between patterns, training at intermediate stages and interpretation of output. In general, the paradigm of relaxation to equilibrium is shown to be inappropriate to temporal pattern processing.

An approach to the problem is outlined which differs from work done to date. A central distinction of the approach is the idea of storing the history of input patterns in nodes, rather than replicating the architecture for each time step. This recasts the problem as the evolution of a dynamical system which can be analyzed in terms of attractors in discrete spaces.

As a simple example, a system is described that embodies some of these ideas. The system learns general temporal patterns, and can solve a temporal equivalent of the XOR problem.

### A NEURAL MODEL FOR PERCEPTUAL STABILITY,

BARTLETT W. MEL, Center for Complex Systems Research, University of Illinois, Champaign, IL 61820, USA.

A sequential neural-network architecture has been applied to the problem of maintaining stable perception in a moving observer. The basis for stable perception lies in the network's ability to predict the next sensory state on the

basis of its current sensory state and internal motor command. A high degree of biological plausibility was sought in the design of this model. As such, the network is only locally interconnected, makes use of the topography of sensory and motor representations, and requires only a simple Hebb-like learning rule. Furthermore, the network has no explicit teacher, but rather trains itself by simply moving through its environment and observing the associated changes in the sensory array. A slightly enhanced single-unit model is also used, in which simple non-linear spatio-temporal interactions in a unit's 'dendritic field' influence its next state of activation.

#### **MECHANISMS OF SELF-ORGANIZATION FOR INFORMATION PROCESSING BY DEVICES MODELED ON CEREBRAL CORTEX,**

WALTER J. FREEMAN, Department of Physiology-Anatomy, University of California, Berkeley, CA 94720, USA.

The outstanding characteristic of biological intelligence is its capacity to set and pursue its own goals, including the subordinate task of pattern recognition. The key property shared by self-organizing systems is nonlinear feedback among large arrays of interconnected elements, each having simple dynamics, leading to the existence of a collection of ordered states reached by continuous trajectories through multiple and repeated bifurcations. A simple but instructive example of a goal-seeking device modeled on the vertebrate olfactory system has been devised through a combination of hardware and software simulation. The system consists of an array of nonlinear oscillators coupled globally by time-dependent multiplexing under the control of a digital computer. Connection weights are stored digitally and are internally modifiable under learning algorithms for habituation and attentiveness. The system does pattern classification by multiple Hopf bifurcation to learned limit cycle attractors emergent from a basal chaotic attractor.

#### **SHIFTERS AND ADAPTIVE FILTERS-POSSIBLE FUNCTIONS OF CORTICAL LAYER 4,**

CHARLES H. ANDERSON, Jet Propulsion Laboratory, Pasadena, CA 91109, USA and DAVID C. VAN ESSEN, Biology Division, California Institute of Technology, Pasadena, CA 91125, USA.

Layer 6 cells of the cortex have a massive feedback projection into layer 4 which, through local interneurons, provides a strong inhibitory influence on the ascending excitatory inputs to the cortex. This suggests that layer 6 may serve an important regulatory role, by controlling flow of input data into each cortical area. In this regard, we have proposed that the enlarged layer 4 that is characteristic of primary visual cortex in primates provides a shifting action which stabilizes the image data against retinal slip. A specific mechanism for producing these shifts will be described, as well as a number of other adaptive spatial-temporal filters that could also be implemented in this region.

## **CENTRIC MODELS OF THE ORIENTATION MAP IN PRIMATE VISUAL CORTEX,**

**BRUCE DOW AND WILLIAM BAXTER,** School of Medicine, State University of New York, Buffalo, NY 14226, USA.

In the visual cortex of the monkey the horizontal organization of the preferred orientations of orientation-selective cells follows two opposing rules: 1) neighbors tend to have similar orientation preferences, and 2) many different orientations are observed in a local region. Several orientation models which satisfy these constraints are found to differ in the spacing and the topological index of their singularities. The models are compared with respect to rates of orientation change, reversals, and discontinuities as obtained from published experimental results.

## **QUANTITATIVE RECONSTRUCTIONS OF COMMAND NEURONS AND MOTONEURONS IN THE TAILFLIP ESCAPE CIRCUIT OF CRAYFISH,**

**DONALD H. EDWARDS, JR.,** Department of Biology, Georgia State University, Atlanta, GA 30303, USA.

The integrative properties of the command neuron and two motoneurons of the crayfish lateral giant escape circuit were studied to determine the effects of those properties on the cells' responses of the excitatory and inhibitory motoneurons to common inputs were observed and compared with the corresponding responses of multicompart ment electrical models of the cells to simulated inputs. Agreement between recorded and computed responses indicates that the passive integrative properties of these cells have a major effect on their roles in the escape circuit.

## **HIGH DENSITY ASSOCIATIVE MEMORY,**

**A. DEMBO and O. ZEITOUNI,** Division of Applied Mathematics, Brown University, Providence, RI 02912, USA.

A novel design of neural networks operating as associative memories is presented. Unlike the schemes suggested by J.J. Hopfield and others, the network is not modelled by a spin-glass like system. An axiomatic approach leads to the characterization of energy functions with no spurious local minima. Combining these energy functions in a descent-type relaxation algorithm leads to an associative memory with only the desired memories being attractors. The continuous-time version of this model has unlimited capacity, whereas the capacity of the discrete-time/finite-state-space version is optimal in the sense of error-correcting capability. Carefully chosen energy functions enable the restoration of the closest (to a given probe input) memory, at least when the distance between the input and his memory is less than half the distances between any two memories. The proposed associative memory is powerful enough to approximate any (predetermined) pattern classification rule with arbitrarily good accuracy.

## **DISTRIBUTED NEURAL PROCESSING IN THE VESTIBULO-OCULAR SYSTEM,**

CLIFFORD LAU, Office of Naval Research Detachment, Pasadena, CA 91106-2485, USA and VICENTE HONRUBIA, Division of Head and Neck Surgery, University of California, Los Angeles, CA 90024, USA.

A new distributed neural processing model is proposed to explain the response characteristics of the vestibulo-ocular system and to more accurately reflect the latest neuroanatomical and neurophysiological data on the vestibular afferent fibers and vestibular nuclei. In this model, the head motion is sensed topographically by haircells in the semicircular canals with a continuum of varying dynamics. The head motion information is then processed by multiple synapses both at the haircell level and at the vestibular nuclei level. The reflexive vestibular eye movement is the result of summation of the distributed responses of this network of neurons.

## **THE HOPFIELD MODEL WITH MULTILEVEL NEURONS,**

M. FLEISHER AND E. LEVIN, Department of Electrical Engineering, Technion - Israel Institute of Technology, 32000 Haifa, Israel.

The Hopfield neural network model for associative memory is generalized. The generalization involves the replacement of two-state neurons by neurons taking on a larger set of values. A class of certain allowed quantization rules for the neurons is developed, guaranteeing convergence to stable states.

The dynamic information capacity of the network is defined as the  $\text{LOG}_2$  (# of different possible state evolutions of network) and is found to be of order  $N^3$  bits for a network with  $N$  neurons.

Next, two learning algorithms for this network are proposed. The first is a generalization of the well-known "Sum of outer products" construction of the connection matrix. Asymptotic bound on the number of stable states is derived for this construction and is found to be of order  $N/\log N$ .

The second, is a perceptron-based learning algorithm, and conditions for its convergence are developed.

## **SIMPLIFIED NEURAL NET ARCHITECTURES,**

JOHN MOODY, Department of Computer Science, Yale University, New Haven, CT 06520, USA.

Novel architectures are introduced which dramatically reduce the hardware and programming complexities of many of the best known neural networks. The Little/Hopfield associative memory model and the  $p^{\text{th}}$ -order correlation models of YC Lee, et al. can be faithfully implemented using only  $2MN$  binary connections ( $M$  is the number of memories). Since the memory bit patterns are stored directly, no advance calculation of outer products is required. Dramatic reductions in complexity also occur for the Traveling Salesman and Hitchcock networks, the quadratic network, the ill-posed inversion network, and the simple  $p$ -flop.

## **ASSOCIATIVE MEMORY DESIGN WITH HIGH-ORDER NEURAL NETWORKS,**

L. PERSONNAZ, I.GUYON, G. DREYFUS, Ecole Supérieure de Physique et Chimie  
Laboratoire d'Electronique, 75005 Paris, France.

The introduction of high-order interactions in fully connected networks of formal neurons is shown to be an efficient way of storing and retrieving correlated informations reliably. A new simple learning rule is derived, giving a large storage capacity and allowing to solve a class of problems which could not be handled by standard networks. It is shown that the choice of the number of high-order interactions is an important parameter in the design of an associative memory, in order to prevent an excessive growth of the number of synaptic coefficients. Local approximations of the learning rule are presented, and some typical applications are investigated.

## **REFLEXIVE ASSOCIATIVE MEMORIES,**

HENDRICUS G. LOOS, Laguna Research Laboratory, Fallbrook, CA 92028-9765,  
USA.

No spurious memories occur for reflexive neural nets which utilize orthonormal bipolar labels together with a nonlinear operation which selects the dominant label. A fault-tolerant associative memory may be obtained by choosing the non-linear operator as a conventional thresholder or a winner-take-all net, together with an orthogonal transformation, such as a Hadamard transform or a Fourier transform. Using  $m$ -dimensional orthonormal bipolar labels,  $m$  analog data vectors of any dimension can be stored with perfect recall.

## **PHASOR NEURAL NETWORKS,**

ANDRE J. NOEST, NIBR, NL-1105, AZ Amsterdam, The Netherlands.

I analyse the behavior of coupled networks of unit-length 2-vectors (phasors). The information is represented in their relative phases, and the couplings and dynamics are chosen so that learned patterns have a low energy to which the network relaxes. Results on accuracy of recall, capacity and spurious stable states in Content Addressable Memory applications are reported. These illustrate by a simple example the extent to which one can go beyond sigmoidal scalar signals in collective processing of information in neural networks without the need for awkward recoding of signals with rotational symmetry.

## **A MEAN FIELD THEORY OF LAYER IV, OF THE VISUAL CORTEX AND ITS APPLICATION TO ARTIFICIAL NEURAL NETWORKS,**

**CHRISTOPHER L. SCOFIELD, Nestor Incorporated, Providence, RI 02906, USA.**

A single cell theory for the development of selectivity and ocular dominance in visual cortex has been generalized to incorporate more realistic neural networks that approximate the actual anatomy of Layer IV of area 17 cortex. In particular we have analyzed, with a mean field approximation, a network consisting of excitatory and inhibitory cells, both of which may receive input information from the LGN. These two cortical cell types then interact through intralayer connections that are either excitatory or inhibitory. Our investigation of the evolution of a cell in this mean field network indicates that many of the results on existence and stability of fixed points that have been obtained previously in the single cell theory can be successfully generalized here. We can, in addition, make explicit further statements concerning the independent effects of excitatory and inhibitory neurons on selectivity and ocular dominance. In particular, shutting off inhibitory cells lessens selectivity and alters ocular dominance (masked synapses). The inhibitory cells may be selective but there is no theoretical necessity that they be so. Further, the intra-layer inhibitory synapses do not have to be very responsive to visual experience. Most of the learning process can occur among the excitatory LGN-cortical synapses. Finally, it can be shown that the N2 connectivity often proposed for artificial neural networks can be reduced to an N+1 connectivity. Some of these theoretical ideas will be compared with experimental results. These results are of significance to the design and implementation of artificial neural networks since they predict that much of the important behavior that is observed within natural networks is a result of inter-layer connectivity and not intra-layer connectivity. Note that this allows the construction of high density memories with networks that are strictly feedforward in architecture and which have optimal recall accuracies. We will discuss several such artificial networks and their application to information processing.

## **EXTEND CONNECTIONIST NETWORK MODELS TO DO SEQUENTIAL REASONING,**

**Y.C. LEE, H.H. CHEN, G.Z. SUN, H.Y. LEE, Department of Physics and Astronomy, University of Maryland, College Park, MD 20742, USA and TOM MAXWELL, Sachs Freeman Associates/Naval Research Laboratory, Landover, MD 20785 USA, and C. LEE GILES, AFOSR/Bolling AFB, DC 20332, USA.**

A neural 'blackboard' architecture is proposed in which a high order associative memory network serves as a global shared memory for one or more feedforward-type neural "inference engines" to allow the system to reason in a time-delayed, or sequential, manner. New learning algorithms are proposed to speed up "slow learning" and to partially alleviate the tough "credit apportionment" problem inherent in any multistep decision.

**SPATIAL ORGANIZATION OF NEURAL NETWORKS: A PROBABILISTIC MODELING APPROACH,**

**ANDREAS STAFYLOPATIS, MARIOS DIKAIAKOS, DIMITRIS KONTORAVDIS,**  
Department of Electrical Engineering, National Technical University of Athens,  
157 73 Zographos - Athens, Greece.

We develop in this paper a probabilistic model of neural networks based upon the theory of general product-form queueing networks. A neural network is modeled as an open network of nodes, in which customers moving from node to node represent stimulation and connections between nodes are expressed in terms of suitably selected routing probabilities. We obtain the solution of the model under different disciplines affecting the time spent by a stimulation at each node visited. Results concerning the distribution of excitation as a function of network topology are compared with measures obtained by simulating the behaviour of conventional neural networks.

## NOTES

## ORAL SESSION 01 BIOLOGICAL MODELS

### **BIFURCATION ANALYSIS OF A NETWORK MODEL OF RABBIT OLFACTORY BULB WITH PERIODIC ATTRACTORS STORED BY A SEQUENCE LEARNING ALGORITHM,**

**BILL BAIRD**, Department of Biophysics, University of California, Berkeley, CA 94720, USA.

The spatio-temporal dynamics of pattern recognition are investigated in a network model of the rabbit olfactory bulb where time varying spatial patterns have been stored by an error correction algorithm. During inspiration such patterns appear in the EEG through a bifurcation from a homogeneous steady state and suffice to predict the animal's response in conditioning experiments. In an effort to obtain mathematical insight into the intrinsic mechanisms and capabilities of this type of system, application is made of the numerical and analytic tools of bifurcation theory - both center manifold theory and the newly emerging singularity theory of bifurcations with symmetry.

### **MODEL OF THE CEREBELLUM AS AN ARRAY OF ADJUSTABLE PATTERN GENERATORS,**

**JAMES C. HOUK**, Department of Physiology, Northwestern University Medical School, Chicago, IL 60611, USA.

Microelectrode studies of signal transmission through the intermediate cerebellum in alert animals will be reviewed. These data provide important constraints on the information processing which must go on within the cerebellum. While some sources of input (mossy fibers) signal detailed parametric information about the state of the body, other inputs (climbing fibers) signal only the occurrences of sensory events. Signals recorded on the output side appear to represent motor commands; they are movement-related with only weak sensory components of response. On the basis of these results and the knowledge that cerebellar lesions interfere with sensorimotor learning, a model is proposed which treats the cerebellum as an array of adjustable motor pattern generators.

### **SPONTANEOUS AND INFORMATION-TRIGGERED SEGMENTS OF SERIES OF HUMAN BRAIN ELECTRIC FIELD MAPS,**

**D. LEHMANN, D. BRANDEIS, H. OZAKI, and I. PAL**, Neurology Department, University Hospital, 8091 Zurich, Switzerland.

Spontaneous and information-triggered brain electric activity is a series of momentary field maps (=functional states). Adaptive segmentation of spontaneous series into spatially stable epochs exhibited 210 msec mean segments, discontinuous changes. Different maps imply activity of different neural populations, hence expectedly different effects on information processing: Reaction time differed between map classes at stimulation (ANOVA p.02). Segments might be units of

brain information processing (content/mode/step), possibly operationalizing consciousness time. Related units (e.g. in activity triggered by stimuli during figure perception and during voluntary attention) might specify brain sub-mechanisms of information treatment.

**TOWARDS AN ORGANIZING PRINCIPLE FOR PERCEPTION: THE ROLE OF HEBBIAN SYNAPSES IN THE EMERGENCE OF FEATURE-ANALYZING FUNCTION,**

RALPH LINSKER, IBM Thomas J. Watson Research Center, Yorktown Heights, NY 10598, USA.

I show that a Hebb-type synaptic modification rule induces the emergence, in a layered network, of feature-analyzing functions similar to those observed in the mammalian visual pathway. A suitable Hebbian algorithm is shown to implement a design optimization principle that tends to maximize (subject to constraints) the accuracy with which input activity information can be inferred, given only the output activity of a cell. I illustrate how this principle leads to the emergence of orientation selectivity in a model visual system, and suggest that this principle of "optimal neural encoding" may serve as a generic organizing principle for the development of biological and synthetic perceptual networks.

## NOTES

## ORAL SESSION 02

### MATHEMATICAL MODELS

#### SYNCHRONIZATION IN NEURAL NETS,

JACQUES J. VIDAL and JOHN HAGGERTY, University of California, Los Angeles, CA 90024, USA.

The paper presents an artificial neural network model where nodes are synchronizable oscillators. The relative timing of neural firing and spatio-temporal coherence are determinant dimensions of information coding. It is shown that synchronizable oscillator arrays are capable of energy minimization computations similar to those achieved with more traditional analog representations, and that stochastic concepts and simulated annealing apply as well.

#### NEW 'NEURAL' ALGORITHMS FOR ASSOCIATIVE MEMORY,

ERIC B. BAUM, Jet Propulsion Laboratory, Pasadena, CA 91109, USA, and JOHN MOODY, Yale University, New Haven, CT 06520, USA, and FRANK WILCZEK, Institute for Theoretical Physics, University of California, Santa Barbara, CA 93106, USA.

We describe a class of models for associative content addressable memory. The models utilize layered neural networks analogous to those of the brain, but possess analogies to standard digital computer memories as well. When a cue is presented at the input layer, an intermediate processing layer retrieves a label associated with a stored word and the output layer displays the stored word. All computation is massively parallel. Our algorithms have been constrained by limitations such as fan out so as to be implementable using analog VLSI technology currently undergoing intensive and promising development for other connectionist algorithms. Excellent performance can be achieved even for synaptic values with large imprecision and no grey scale. Our algorithms use the identical technology, but have substantial advantages in storage capacity, recall reliability, and flexibility when compared to the Hopfield model.

The choice of internal label representations is integral to the capabilities of the algorithm. We describe two classes of internal labels: the unary or 'grandmother cell' representation and various sparse distributed representations. A new, fixed weight code is present. Generalized inverse methods are discussed. We consider storage of sparse data as well as postprocessing layers and preprocessing layers which employ 'feature detectors' to sparsify general data in order to achieve high capacity. The models saturate fundamental information theoretic bounds on capacity and retrieval. We believe our results will be of interest both for biological modelling and for VLSI implementation.

### **A NEURAL NETWORK CLASSIFIER BASED ON CODING THEORY,**

**TZI-DAR CHIUEH** and **RODNEY GOODMAN**, Department of Electrical Engineering, California Institute of Technology, Pasadena, CA 91125, USA.

The new neural network classifier we propose transforms the classification problem into the coding theory problem of decoding a noisy codeword. Input vectors in the feature space are transformed to vectors in a code space, and then error correction decoded in this space to classify the input feature vector to its class. We show that the number of classes stored in an N-neuron system is linear in N and significantly more than that obtainable by using the Hopfield model, at the expense of some extra complexity. For example, with N=63 the former model has a reliable classification capacity of 5, while our new models stores 48, for less than twice the complexity. Finally, we give a practical application of the classifier, showing how it can be used as an efficient edge detector in digital image processing.

### **A TRELLIS-STRUCTURED NEURAL NETWORK,**

**THOMAS PETSCHÉ**, Siemens Corporate Research and Support Inc., Princeton, NJ 08540, USA, and **BRADLEY W. DICKINSON**, Department of Electrical Engineering, Princeton University, Princeton, NJ 08544, USA.

We have developed a neural network which consists of cooperatively interconnected Grossberg on-center off-surround subnets and which can be used to optimize a function related to the log likelihood function for decoding convolutional codes or more general FIR signal deconvolution problems. Connections in the network are confined to neighboring subnets, and it is representative of the types of networks which lend themselves to VLSI implementation. Analytical and experimental results for convergence and stability of the network have been found. The structure of the network can be used for distributed representation of data items while allowing for fault tolerance and replacement of faulty units.

### **CONSTRAINED NETWORK REPRESENTATIONS,**

**JOHN PLATT** and **ALAN BARR**, Department of Computer Science, California Institute of Technology, Pasadena, CA 91125, USA.

Many mathematical models that are represented as neural networks need constraints to restrict the space of possible outputs or to accept input. For example, to solve the travelling salesman problem or analog decoding, a matrix of neurons needs to represent a permutation matrix. Another example is that of an elastic model which represents shape. The model modifies its shape based on information supplied by an image.

Quadratic energy constraints are not guaranteed to continuously satisfy the constraint conditions as a function of time, particularly when there are other forces on the neural models and when there are multiple constraints. We will create forces which satisfy the constraints over time, using inverse dynamics, Lagrange multipliers, and projection. These extra forces correspond to modifying the inputs of the neurons.

## NOTES

## ORAL SESSION 03

### LEARNING THEORY

#### **MINKOWSKI-R BACK-PROPAGATION: LEARNING IN CONNECTIONIST MODELS WITH NON-EUCLIDIAN ERROR SIGNALS,**

STEPHEN JOSE HANSON and DAVID J. BURR, Bell Communications Research, Morristown, NJ 07960 USA.

Learning in many connectionist models has been implemented using gradient descent in a least squares error function of the network output and the teacher signal. The present model generalizes, in particular, Back-propagation by using Minkowski- $r$  power metrics. For small  $r$ 's a 'city block' error metric is approximated and for large  $r$ 's the 'maximum' or 'supremum' metric is approached, while for  $r=2$  the standard back-propagation model results. An implementation of Minkowski- $r$  Back-Propagation is described, and several experiments showing that different values of  $r$  may be desirable for various purposes, including reduction of the effects of outliers (noise), modeling the input space with more compact clusters, or modeling the statistics of a particular domain more naturally or in a way that may be more perceptually or psychologically meaningful (e.g. in speech or vision).

#### **GENERALIZATION OF BACKPROPAGATION TO RECURRENT AND HIGH-ORDER NETWORKS,**

FERNANDO J. PINEDA, Applied Physics Laboratory, Johns Hopkins University, Laurel, MD 20707, USA.

The backpropagation algorithm of Rumelhart et al. is shown to be a special case of a general class of algorithms which apply to networks with feedback and feedforward connections. The propagation of activation in such networks is determined by dissipative differential equations.

A general method for deriving backpropagation algorithms for networks with feedback and higher order connectivity is introduced. As an example, this method is used to derive an algorithm for adaptively modifying the weights of a recurrent generalization of the network introduced by Rumelhart et al. The backpropagation involves the integration of an associated differential equation. The weight matrix is updated only after the forward-propagation and backward-propagation differential equations have reached steady state.

This approach may be significant for several reasons, first, there are no constraints on the connectivity of the network. Preliminary experiments show that learning is robust even in sparse randomly connected networks. Second, learning algorithms for quite arbitrary networks can be derived, e.g. the Hopfield network and higher order networks. Third, the recurrent algorithms overcome the synchronization difficulties which occur in implementing the feedforward algorithm in VLSI. Fourth, preliminary numerical experiments show that, in some cases, recurrent networks learn faster than networks with feedforward connections only.

## **PARTITIONING OF SENSORY DATA BY A CORTICAL NETWORK,**

R. GRANGER, H. HENRY, and G. LYNCH, Center for the Neurobiology of Learning and Memory, University of California, Irvine, CA 92717, USA.

Processing sensory data requires preservation of information about both similarities and differences among learned cues, without which either acuity would be lost or degraded versions of a cue would be erroneously thought to be distinct cues, and would not be recognized. We have constructed a model of piriform cortex in which fidelity to a large number of biophysical, anatomical and physiological parameters yields a marked ability to learn multiple, similarity- and difference-preserving encodings of cues.

In particular, probabilistic quantal transmitter-release properties of piriform synapses give rise to probabilistic postsynaptic voltage levels which, in combination with the activity of local patches of inhibitory interneurons in layer II, differentially select bursting vs. single-pulsing layer-II cells. Time-locked firing to the theta rhythm (Larson and Lynch, 1986) enables distinct spatial patterns to be read out against a relatively quiescent background firing rate. Training trials simulate the conditions for induction of long-term potentiation (LTP), stabilizing layer-II-cell spatial firing patterns for learned cues, such that initial responses to strongly-overlapping simulated inputs preserve the input overlap (e.g., input cues with 80% overlap result in early output responses with equal or greater overlap). Enhanced synapses cause stronger cell firing, yielding strong, cell-specific afterhyperpolarizing (AHP) currents. Alternate cells, selected by the same inhibitory interneuron mechanisms described above, then activate their caudally-flowing recurrent collaterals, activating distinct populations of synapses in caudal layer Ib, which when potentiated selectively stabilize the response of caudal cells that tend to enhance the differences among even very-similar cues on late response firing patterns (e.g., inputs with 90% overlap result in late responses with less than 10% overlap). The difference-enhancing response can be measured with respect to its acuity, whereas the similarity-enhancing response is properly viewed as a *partitioning* mechanism, mapping distinct input cues onto nearly-identical response patterns (category indicators). Using a statistical metric for the information value of categorizations enables measurement of the value of partitionings produced by the piriform simulation network.

## **THE GENERATION OF EFFICIENT REPRESENTATIONS IN NEURAL NET ARCHITECTURES USING HIGH ORDER CORRELATIONS,**

TOM MAXWELL, Sachs-Freeman Associates, Landover, MD 20785, USA, and C. LEE GILES, AFOSR/NE, Bolling Air Force Base, DC 20332, USA, and Y. C. LEE, Department of Physics, University of Maryland, College Park, MD 20742, USA.

In this paper the learning and generalization capabilities of single and multiple slab high order architectures will be measured on test problems such as contiguity, symmetry detection, and parity. We have found that for certain sets of problems a single slab of high order units will generalize nearly perfectly in cases in which back propagation shows very little generalization capacity. High order cascaded slab architectures are capable of learning very high order problems (such as ninth

order parity) which are not handled efficiently by single slab architectures. These architectures require only integer arithmetic and converge much faster than back propagation.

## NOTES

## POSTER SESSION P2

### LEARNING STEREOPSIS IN A HIGH-ORDER NEURAL NETWORK,

H.H. CHEN, G.Z. SUN and Y.C. LEE, Department of Physics, University of Maryland, College Park, MD 20742, USA.

A high order recursive network is constructed to learn the task of stereopsis from random dot stereograms. Hebbian learning and translational invariance are exploited to analytically study the connection weights and the performance of the network. It is demonstrated that the network learned the continuity and the uniqueness constraints automatically.

### LEARNING BY STATE RECURRENCE,

BRUCE E. ROSEN, JAMES M. GOODWIN and JACQUES J. VIDAL, Department of Computer Science, University of California, Los Angeles, CA 90024, USA.

This research investigates Recurrence learning, a new technique for nonsupervised learning of nonlinear control problems. When applied to Barto and Sutton's associative search element (ASE), it significantly improves the convergence rate of stochastically based learning automata.

Recurrence learning exploits information found during learning trials to reward a decision which produces a return to nonfailing states. Whereas in the standard BOXES or ASE algorithms, negative reinforcement is only applied only on failure, Recurrence learning applies positive reinforcement during the exploration of the search space.

Other applications, where detection of recurring oscillations are punished instead of rewarded, should benefit from this approach.

### MANUAL TRAINING TECHNIQUES OF AUTONOMOUS SYSTEMS BASED ON ARTIFICIAL NEURAL NETWORKS,

J. SHEPANSKI and S. MACY, TRW Defense Systems Group, Redondo Beach, CA 90278, USA.

We are developing autonomous control systems based on artificial neural networks (ANN), and methods for training these systems manually. We find that properly constructed networks internally generate rules of behavior that allow them to function autonomously if they have been trained on a spanning set of possible data combinations. This training can be provided manually, either under the direct supervision of a system trainer, or indirectly using a background mode where the network assimilates its training data by correlating input data with the expert's responses as he performs his day-to-day tasks.

## **TIME-SEQUENTIAL SELF-ORGANIZATION OF HIERARCHICAL NEURAL NETWORKS,**

A.S. NOETZEL, Polytechnic University, Brooklyn, NY 11201, USA, and R.H. SILVERMAN, Cornell University Medical College, New York, NY 10021, USA.

Self organization of multi-layered hierarchical networks can be realized by time-sequential organization of successive neural layers. Lateral inhibition operating within clusters of cells in each layer provides for capture of excitation patterns presented by the previous layer. By presenting patterns grouped into sets of increasing complexity, in co-ordination with network self-organization, higher levels of the hierarchy capture concepts implicit in the pattern sets.

## **SUPERVISED LEARNING OF PROBABILITY DISTRIBUTIONS BY NEURAL NETWORKS,**

ERIC B. BAUM, Jet Propulsion Laboratory, Pasadena, CA 91109, USA, and FRANK WILCZEK, Department of Physics, Harvard University, Cambridge, MA 02138, USA.

We propose that the back propagation algorithm for supervised learning can be generalized, put on a satisfactory conceptual footing, and very likely made more efficient by defining the values of the output and input neurons as probabilities and varying the synaptic weights in the gradient direction of the log likelihood, rather than the 'error.'

## **STATIC AND DYNAMIC ERROR PROPAGATION NETWORKS WITH APPLICATION TO SPEECH CODING,**

A.J. ROBINSON and F. FALLSIDE, Department of Engineering, Cambridge University, Cambridge, England.

Error propagation nets have been shown to be able to learn a variety of tasks in which a static input pattern is mapped onto a static output pattern. This paper presents a generalisation of these nets to deal with time varying, or dynamic patterns, and three possible architectures are explored. As an example, dynamic sets are applied to the problem of speech coding, in which a time sequence of speech data are coded by one net and decoded by another. The use of dynamic nets gives a better signal to noise ratio than that achieved using static nets.

## **DISCOVERING INPUT STRUCTURE USING FEED-FORWARD NETWORKS,**

K. PRAZDNY, Artificial Intelligence Center, FMC Corporation, Santa Clara, CA 95052, USA.

One of the objections against supervised learning procedures is the necessity of an external teacher that specifies, for each input vector, the desired output vector. There is, however, at least one situation in which the teacher is explicitly and inherently available to guide the learning process: the detection or discovery of structure in the input vector itself. The goal of such (perceptual) learning is to determine which environmental features are predictive of others and distinguish

useful cues from context and background noise. We address the problem of determining which input variables are predictive of which other variables without knowing anything about the actual relationship.

Fully connected, feed-forward networks (where each unit on a given layer is connected to each unit on *every* layer above it) and the backpropagation credit/blame assignment can be used to solve this problem for a wide range of input/output dependencies. When such network converges for a given set of input/output examples the weight matrix contains information about irrelevant input variables. 'Ablation' of such units has minimal consequences for the accuracy of the output produced by the network. Thus, an automatic procedure can be developed that is capable of automatically discovering the dependencies between the input variables and checking on its conclusion by 'simulation.'

#### **AN ARTIFICIAL NEURAL NETWORK FOR SPATIO-TEMPORAL BINARY PATTERNS: APPLICATION TO PHONEME CLASSIFICATION,**

TOSHITERU HOMMA, LES E. ATLAS and ROBERT J. MARKS II, Department of Electrical Engineering, University of Washington, Seattle, WA 98195, USA.

An artificial neural network is developed to recognize spatio-temporal binary patterns associatively. the function of a formal neuron is generalized by replacing multiplication with convolution, weights with transfer functions, and thresholding with sigmoidal transform following adaptation. The Hebbian learning rule and the delta learning rule are generalized accordingly, resulting in the learning of delays as well as of weights. The neural network which was first developed for spatial patterns was thus generalized for spatio-temporal patterns. It was tested using sets of input patterns modelled on speech signals, showing robust classification of phonemes where 100 neurons and 30 model phonemes were used.

#### **HIERARCHICAL LEARNING CONTROL - AN APPROACH WITH NEURON-LIKE ASSOCIATIVE MEMORY,**

E. ERSU, ISRA Systemtechnik GmbH, 6100 Darmstadt, FRG, and ST. GEHLEN and H. TOLLE, Control Systems Theory Laboratory, Technische Hochschule Darmstadt, 6100 Darmstadt, FRG.

Based on the learning control concept LERNAS with neuron-like associative memory systems, the paper discusses a multi-level hierarchical control where each level is composed of LERNAS-loops with local tasks operating on a specific subprocess.

The LERNAS-control scheme represents a new type of self-organizing control system motivated by neurobiological and psychological research on the brain. The main novelty is that the concept uses a control action optimization strategy via a k-step ahead output predictive algorithm (similar to human problem solving) with both the predictive model of the unknown environment and the control strategy being represented by general mathematical (i.e. non-linear) mappings. Both mappings are carried out by special neuron-like associative memory system similar to information storage in neuronal networks.

The contribution will discuss the multilevel composition of LERNAS-loops, where in each level the lower levels are considered as a pseudo-sub-process. Thus the inter-level communication is fully directed by the upper levels which leads to a pyramidal control structure.

#### **LEARNING ON A GENERAL NETWORK,**

AMIR F. ATIYA, Department of Electrical Engineering, California Institute of Technology, Pasadena, CA 91125, USA.

This paper generalizes the backpropagation method to a general network containing feedback connections. The network model considered consists of interconnected groups of neurons, where each group could be fully interconnected (it could have feedback connections, with the weights not necessarily being symmetric), but no loops between the groups. A stochastic descent algorithm is applied, under a certain inequality constraint on each intra-group weight matrix, which ensures for the network to possess a unique stable state for every input.

#### **A FRAMEWORK FOR SPEECH RECOGNITION USING NEURAL NETWORKS,**

RICHARD P. LIPPMANN and WILLIAM Y. HUANG, Lincoln Laboratory, Massachusetts Institute of Technology, Lexington, MA 02173-0073, USA.

The preliminary design of a neural net speech recognizer and an automatic training algorithm for this recognizer is presented. The design uses many simple computational elements which operate in parallel to shift the focus of processing attention over time based on expectations of serial order. The training algorithm specifies both parameters (weights) and the structure of the net. The structure of sub-word and word segments is built up over time from training data instead of being pre-specified by hand as in most current recognizers.

#### **LEARNING IN NETWORKS OF NONDETERMINISTIC ADAPTIVE LOGIC ELEMENTS,**

RICHARD C. WINDECKER, AT&T Bell Laboratories, Middletown, NJ 07748, USA.

This work presents a model of nondeterministic adaptive automata that are constructed as networks of simpler nondeterministic adaptive information processing elements. In this model, all adaptive elements in a network receive the same positive or negative reinforcement simultaneously. The model was developed partly to support the hypothesis that animal nervous systems are nondeterministic adaptive automata constructed of simpler nondeterministic adaptive parts. It may also facilitate new approaches to building machines that can learn to perform (artificially) intelligent things. Computer simulations show that multiple-input, multiple-output, multiple-layered, combinational and sequential networks built within this model have reasonable learning characteristics.

## **NEUROMORPHIC NETWORKS FOR OPTICAL COMMUNICATIONS AND SIGNAL PROCESSING,**

M.P. VECCHI and J.A. SALEHI, Bell Communications Research, Morristown, NJ 07960-1961, USA.

Neuromorphic networks specifically designed for communications and optical signal processing applications are presented. Sparse Optical Orthogonal Codes sequences are used to encode the information, on the basis of unipolar, binary (0,1) signals. The connectivity matrix is also unipolar, and clipped to binary (0,1) values. In addition to associative memory, the resulting neural networks provide a variety of more general functions, such as code filtering, code mapping, code joining, code shifting, and code projecting.

## **ROTATION AND SCALE INVARIANT OBJECT RECOGNITION USING A DISTRIBUTED ASSOCIATIVE MEMORY,**

HARRY WECHSLER and GEORGE LEE ZIMMERMAN, Department of Electrical Engineering, University of Minnesota, Minneapolis, MN 55455, USA.

We describe an approach to 2-dimensional object recognition. Complex-log conformal mapping is combined with a distributed associative memory to create a system which recognizes objects regardless of changes in rotation or scale. Recalled information from the memorized database is used to classify an object, reconstruct the memorized version of the object, and estimate the magnitude of changes in scale or rotation. The system response is resistant to moderate amounts of noise and occlusion. Several experiments, using real, gray scale images, are presented to show the feasibility of our approach.

## **AN IMPROVED NEURAL-NETWORK ANALOG-TO-DIGITAL CONVERTER,**

R.A. GAMES, E.L. KEY and D. MOULIN, The MITRE Corporation, Bedford, MA 01730, USA.

This paper presents a new neural-network analog-to-digital (A/D) converter. Given an input analog voltage, the network is guaranteed to converge to the correct digital state, regardless of the state used to initialize the network or the gain of the analog neurons, and so overcomes two difficulties associated with the A/D converter derived from the Hopfield minimization technique. The paper develops a new neural-network A/D converter formulation by considering the partitions of the input parameter space produced by a neural network. Simulation results are presented that compare the performance of the old and new formulations. A 4-bit neural-network A/D converter was assembled with off-the-shelf electronic hardware.

## NOTES

## ORAL SESSION O4 LEARNING TECHNIQUES

### LEARNING REPRESENTATIONS BY RECIRCULATION,

GEOFFREY E. HINTON, Computer Science Department, University of Toronto, Toronto M5S 1A4, Canada, and JAMES L. MCCLELLAND and GEOFFREY J. GOODHILL, Computer Science Department and Psychology Department, Carnegie Mellon University, Pittsburgh, PA 15213, USA.

We describe a new learning procedure for networks that contain several groups of non-linear units arranged in a closed loop. The procedure modifies the weights on the connections between groups so that the training patterns over the input group return unaltered after passing around the loop. The learning rule amounts to changing each weight by an amount proportional to the product of the pre-synaptic activity and the *rate of change* of the post-synaptic activity. It is much simpler to implement in hardware than methods like back-propagation. Simulations show that it usually converges rapidly, and analysis shows that in certain restricted cases it performs gradient descent in a measure of how much the training patterns are altered by passing around the loop.

### STRATEGIES FOR TEACHING LAYERED NETWORKS CLASSIFICATION TASKS,

BEN S. WITTNER, JOHN S. DENKER, L.D. JACKEL, and R.E. HOWARD, AT&T Bell Laboratories, Holmdel, NJ 07733, USA.

It is widely believed that the generalized delta method, applied to a one-layer network, is guaranteed to be equivalent to the perceptron learning rule. This is definitely not the case. We show, however, that with certain modifications and in certain limits, the generalized delta rule *does* reduce to the perceptron learning rule, and naturally extends it to multiple layers.

We illustrate this with a simple classification problem which can be solved by a one layer perceptron using the standard rule, but cannot be solved by the standard version of the generalized delta rule. We also present an example of how our improved version works better on data taken from a handwritten digit recognition experiment.

### STOCHASTIC LEARNING NETWORKS AND THEIR ELECTRONIC IMPLEMENTATION,

JOSHUA ALSPECTOR, ROBERT B. ALLEN, VICTOR HU, and SRINAGESH SATYANARAYANA, Bell Communications Research, Morristown, NJ 07960, USA.

We describe a variety of learning algorithms that operate on a recurrent, symmetrically connected, neuromorphic, Hopfield-style network that, like the Boltzmann machine, settles to a global minimum in its Liapunov function in the presence of noise. These networks learn by modifying synaptic connections

strengths on the basis of correlations seen locally by each synapse. We describe a version of the supervised learning algorithm for a network with analog activation functions. We also demonstrate unsupervised competitive learning with this approach and describe preliminary experiments in reinforcement learning.

These algorithms were chosen for ease of implementation in VLSI. We have designed a CMOS test chip in 2 micron rules that can speed up the learning about a millionfold over an equivalent simulation on a VAX 11/780. The speedup is due to parallel analog computation for summing and multiplying weights and activations, and the use of physical processes for generating random noise. The components of the test chip are a noise amplifier, a neuron amplifier, and a 300 transistor adaptive synapse, each of which is separately testable. These components are also integrated into a 6 neuron and 15 synapse network. Finally, we point out techniques for reducing the area of the electronic correlational synapse both in technology and design and show how the algorithms we study can be implemented naturally in electronic systems.

#### **SPEECH STABILIZATION AND ROBUST FRONT ENDS FOR NEURAL NETWORKS,**

D.J. BURR, Bell Communications Research, Morristown, NJ 07960, USA.

Experiments on speech and handwriting recognition with neural networks indicate that fast and accurate learning is possible by careful attention to front end feature analysis. Irregularities associated with elastic deformations of space/time/level, if treated early, can reduce substantially the Kolmogorov complexity, network size, and training requirements. Networks using robust front ends learn to recognize spoken and printed materials in seconds and generalize to better than 98% recognition accuracy. A neural-based word recognizer reads words with only two or three errors per thousand words (98.8%) when aided by an English dictionary.

The current speech front end system uses weighted cepstral analysis to stabilize spectral information and a simplified energy model to time stabilize a monosyllabic word. The model is extended to allow time stabilization of polysyllabic inputs and hence efficient recognition of arbitrary words. Recent successes on hidden Markov modeling of speech imply the existence of quasi-stable acoustic events and provide encouragement for the success of this approach.

## NOTES

## ORAL SESSION 05

### APPLICATIONS

#### A 'NEURAL' NETWORK THAT LEARNS TO PLAY BACKGAMMON,

G. TESAURO, Center for Complex Systems Research, University of Illinois at Urbana-Champaign, Champaign, IL 61820, USA, and T.J. SEJNOWSKI, Biophysics Department, Johns Hopkins University, Baltimore, MD 21218, USA.

We describe a class of connectionist networks that have learned to play backgammon at an intermediate-to-advanced level. The networks were trained by a supervised learning procedure on a large set of sample positions evaluated by a human expert. In actual match play against humans and conventional computer programs, the networks demonstrate substantial ability to generalize on the basis of expert knowledge. Our study touches on some of the most important issues in network learning theory, including the development of efficient coding schemes and training procedures, scaling, generalization, the use of real-valued inputs and outputs, and techniques for escaping from local minima. Practical applications in games and other domains are also discussed.

#### GENETIC DATABASE ANALYSIS WITH NEURAL NETS,

ALAN LAPEDES and ROBERT FARBER, Theoretical Division, Los Alamos National Laboratory, Los Alamos, NM 87545, USA.

Genetic databases such as the DNA sequence database at Los Alamos or the Brookhaven protein sequence database contain real world information that is encoded in a relatively few number of symbols. We have considered analyzing DNA and amino acid data using backpropagation. The neural net can make an 80% accurate prediction of whether a fragment of DNA codes for a protein or not given only 30 bases, whereas conventional statistical methods to perform the same task degenerate to chance at the 30 base level. In attempting to train a neural net to assign a correct secondary structure characteristic to an amino acid we are currently competing with conventional methods and are optimistic about exceeding conventional prediction capability.

#### LEARNING A COLOR ALGORITHM FROM EXAMPLES,

A. HURLBERT and T. POGGIO, Artificial Intelligence Laboratory, Massachusetts Institute of Technology, Cambridge, MA 02139, USA.

We show that a color algorithm capable of separating illumination from reflection in a Mondrian world can be learned from a set of examples. The learned algorithm resembles the latest retinex color algorithm proposed by Land. This result is a specific instance of our earlier result that a standard regularization algorithm can be learned from examples. It illustrates that the constraints needed to solve a problem in inverse optics can be extracted directly from a sufficient set of associated input and output data.

## COMPUTING MOTION USING NEURAL NETWORKS,

CHRISTOF KOCH, Divisions of Biology, Engineering and Applied Science, California Institute of Technology, Pasadena, CA 91125, USA, and JAMES HUTCHINSON, Jet Propulsion Laboratory, Pasadena, CA 91109, USA.

Computing the velocity field from an image sequence is an ill-posed problem that can be made well-posed by imposing a smoothness constraint (Poggio, Torre and Koch, 1985). Optimization of the corresponding quadratic energy functional has been successfully used to recover smooth motion (area-based approach: Horn and Schunck, 1981; contour-based approach: Hildreth, 1984). Linear resistive networks with local connectivity can implement the appropriate Euler-Lagrange equations (Poggio and Koch, 1985). However, standard regularization theory does not take account of any discontinuities in the velocity field. Thus, we introduce "motion line processes" to segment the computed velocity field and incorporate discontinuities using a Markov Random Field model (Geman and Geman, 1984). This approach allows us to include a *a priori* knowledge about discontinuities (e.g., they occur along extended contours). Simple analog or mixed analog/digital networks with switches between resistive connections can implement the resulting non-quadratic functionals in an efficient manner. Finally, following Gamble and Poggio (1987) co-localization of depth-discontinuities and edges, we introduce the additional constraint that motion discontinuities can only occur if an intensity discontinuity (e.g., edge or zero-crossing) co-exists at the same location. Using image sequences from video cameras, we will show that this approach successfully recovers the piece-wise smooth velocity field in addition to explicitly labeling the motion discontinuities.

## DYNAMICAL NEURAL NETWORKS AND THEIR APPLICATION TO ROBOT CONTROL,

ATHANASIOS SIDERIS, ALAN YAMAMURA and DEMETRI PSALTIS, Department of Electrical Engineering, California Institute of Technology, Pasadena, CA 91125, USA.

In this paper we investigate the ability of neural networks to 'learn' simple physical laws by means of storing the interdependence of the physical quantities involved in the form of connection strengths in a multilayered network. We also investigate the properties of such trained neural networks for the control of dynamical systems.

Specifically, we consider a problem that humans solve repeatedly and without hesitation: positioning their arm at a desired location. We study how well simulated neural networks can accomplish the task of translating the desired end-positions, or perhaps trajectories, of the arm to the necessary torques that effect the motion of the arm.

## NOTES

## ORAL SESSION 06

### COMPUTER SIMULATIONS IN NEUROBIOLOGY

#### **SIMULATIONS SUGGEST INFORMATION PROCESSING ROLES FOR THE DIVERSE CURRENTS FOUND IN HIPPOCAMPAL NEURONS,**

LYLE J. BORG-GRAHAM, Center for Biological Information Processing, Massachusetts Institute of Technology, Cambridge, MA 02139, USA.

A computer model of the hippocampal pyramidal cell (HPC) is described which integrates data from a variety of sources in order to develop a consistent description for this cell type. The model presently includes descriptions of eleven non-linear somatic currents of the HPC. The electronic structure of the HPC is modelled with a soma/short-cable approximation. Model simulations qualitatively or quantitatively reproduce a wide range of somatic electrical behavior in HPCs, and explicitly demonstrate possible roles for the various currents in information processing.

#### **A COMPUTER SIMULATION OF CEREBRAL NEOCORTEX: COMPUTATIONAL CAPABILITIES OF NONLINEAR NETWORKS,**

A. SINGER AND J.P. DONOGHUE, Center for Neural Science, Brown University, Providence, RI 02912, USA.

A neural network simulation of cerebral neocortex was developed based on detailed anatomy and physiology. Processing elements possess temporal nonlinearities and connection patterns similar to those of cortical neurons. The network was able to replicate spatial and temporal integration properties found experimentally in neocortex. A certain level of asynchrony was found to be crucial for the robustness of at least some of the network's computational capabilities.

#### **AN ADAPTIVE AND HETERODYNE FILTERING PROCEDURE FOR THE IMAGING OF MOVING OBJECTS,**

H.A.K. MASTEBROEK, F.H. SCHULING and W.H. ZAAGMAN, Biophysics Department, Groningen State University, 9718 CM Groningen, The Netherlands.

Recent experimental work on the temporal behaviour of a wide-field movement sensitive interneuron in the highest order optical ganglion of the blowfly *Calliphora erythrocephala* (*M*), revealed the amazing phenomenon that at this high level of the fly visual system, the timeconstants which are involved in the processing of neural activity evoked by moving objects, are inverse proportional to the velocity of those objects over an extremely wide range.

In this contribution we discuss the implementation of a two-dimensional heterodyne adaptive filter construction into a computer simulation model. This model includes the ability to account for the observed stimulus-tuned adaptive

temporal behaviour of timeconstants in the fly visual system. The simulation results obtained, clearly show that the application of such an adaptive processing procedure delivers an improved imaging technique of moving patterns in the high velocity range.

## NOTES

## POSTER SESSION P3

### ON TROPISTIC PROCESSING AND ITS APPLICATIONS,

MANUEL F. FERNANDEZ, General Electric Advanced Technology Laboratories, Syracuse, NY 13221, USA.

The interaction of a set of tropisms is sufficient in many cases to explain the seemingly complex behavioral responses exhibited by varied classes of biological systems to combinations of stimuli. It can be shown that a straightforward generalization of the tropism phenomenon allows the efficient implementation of effective algorithms which appear to respond "intelligently" to changing environmental conditions. Examples of the utilization of tropistic processing techniques will be presented in this paper in applications entailing simulated behavior synthesis, path-planning, pattern analysis (clustering), and engineering design optimization.

### NEW MULTI-NEURON RECORDING TECHNIQUES,

M.E. NELSON, B. RASNOW, J.J. BANIK, Y.F. WONG, A. ATIYA, AND J.M. BOWER, Divisions of Biology, Engineering and Physics, California Institute of Technology, Pasadena, CA 91125, USA.

The experimental neurobiologist is faced with a variety of technical challenges when exploring the functional properties of biological neural networks. One such challenge is that of developing techniques to simultaneously record and analyze signals from a large number of individual neurons. We describe the latest methods being developed in our laboratory to address this problem and present multi-neuron data recorded from the rat cerebellar cortex using these techniques. The recording methods involve the fabrication of multi-channel, silicon-based microelectrodes with on-chip preamplification and multiplexing, the precise placement of the probes in the brain and accurate reconstruction of the recording site positions in three dimensions, the amplification of up to 100 neural signals with programmable gain control, analog filtering, thresholding and monitoring for error conditions, and finally, the possibility of analyzing and classifying neural waveforms in real-time using Hopfield-type networks.

### COMPUTER SIMULATIONS OF DYNAMIC FUNCTIONAL CHANGES IN THE TOPOGRAPHIC MAPS OF ADULT ANIMALS,

LEIF H. FINKEL and GERALD M. EDELMAN, Neurosciences Institute, The Rockefeller University, New York, NY 10021, USA.

A computer model of plasticity in topographic maps has been constructed that accounts for several recent experimental findings in monkey somatosensory cortex. The model incorporates realistic neural features and is based on competition between functional groups of neurons. The mechanisms underlying receptive field organization are studied at the synaptic, neuronal, and network levels.

## **DISCOVERING STRUCTURE FROM VISUAL MOTION IN MONKEY, MAN AND MACHINE,**

M. SIEGEL, Laboratory of Neurobiology, The Rockefeller University, New York, NY 10021, USA.

The ability to obtain three-dimensional structure from two dimensional visual motion is important for survival of both humans in the city and monkeys in the jungle. Stimuli have been developed to study this visual function under controlled conditions (Siegel and Andersen, submitted); the psychophysical evidence indicates that ability to perceive three-dimensional structure is similar for both species. Studies in human infants (Spitz, Stiles-Davis and Siegel, submitted) and in normal and brain damaged adults are allowing us to trace the development and physiology of this visual function in man. Furthermore this work provides new computational constraints for any model for this visual function.

Studies in the monkey in which the cortical motion areas MT or MST are removed by a chemical agent indicate that each is behaviorally necessary for the performance of this task. The time course of the effect supports the anatomical and physiological data that indicates that MT is lower in the visual hierarchy than area MST.

These results have lead to the development of a parallel distributed computational model for the analysis performed between areas MT and MST. Psychophysical analysis of the model using the same stimuli as in the primate studies suggest that it satisfies a number of the biological constraints.

## **COMPETITIVE LEARNING, UNLEARNING AND FORGETTING IN OPTICAL RESONATORS,**

DANA Z. ANDERSON and MICHAEL J. O'CALLAGHAN, Joint Institute for Laboratory Astrophysics, University of Colorado, Boulder, CO 80309-0440, USA.

We consider the processes of learning, unlearning and forgetting in two types of optical resonators having real-time volume holographic elements and gain. The learning and recall that takes place in a resonator is through a competitive process in the gain medium. In one type of resonator, information is stored in the form of the resonator's eigenmodes. We call this an eigenstate memory resonator. This type of resonator also goes through an unlearning and forgetting process. The other type of resonator makes associations between its eigenmodes and input information. We call this an encoding resonator. The encoding resonator will forget its associations in time.

## **INTRODUCTION TO A SYSTEM FOR IMPLEMENTING NEURAL NETWORKS ON SIMD ARCHITECTURES,**

SHERRYL TOMBOULIAN, Institute for Computer Applications in Science and Engineering, NASA Langley Research Center, Hampton, VA 23665, USA.

A new algorithm is proposed which allows a realistic implementation of neural networks on massively parallel architectures. The algorithm allows the formation of arbitrary connections between the "neurons." A feature is the ability to add

new connections quickly. It also has error recovery ability and is robust over a variety of network topologies. Simulations of the system, and its actual implementation on the Connection Machine, indicate that the time and space requirements are proportional to the product of the average number of connections per neuron and the diameter of interconnection network.

#### **THE SIGMOID NONLINEARITY IN OLFACTORY SYSTEM OF RAT,**

F.H. EECKMAN, Department of Physiology-Anatomy, University of California, Berkeley, CA 94720, USA.

We report a study on the relationship between EEG amplitude values and unit spike output in the olfactory cortex of awake rats. This relationship takes the form of a sigmoid curve, that describes normalized pulse-output for normalized wave-input. The curve is fitted using nonlinear regression and described by a parameter  $Q_{max}$ , that relates to slope and maximum height. The data are compared to previous sets obtained in the olfactory bulb of awake rats. The differences in slope between the structures are striking and may be related to the different functions of both parts of the system. Analogies with existing neural nets are discussed.

#### **PULSED NEURAL NETWORKS: HARDWARE, SOFTWARE, AND THE HOPFIELD A/D CONVERTER EXAMPLE,**

MARTIN E. GASPAR, Department of Electrical Engineering, University of Utah, Salt Lake City, Utah 84112, USA.

Neuron models are presented that generate and accept pulses, and the Hopfield A/D converter simulation is redone with these pulsed models and shown to function correctly. A description is given of how the model was extracted from the biology, and a tested circuit implementation is presented. Development of the simulation is explained, and the significant segments of code are listed. This work shows that while pulsed models can reduce to existing models under special conditions, in the general case they may introduce a new temporal summation aspect to neural nets.

#### **USING NEURAL NETWORKS TO IMPROVE COCHLEAR IMPLANT SPEECH PERCEPTION,**

MANOEL F. TENORIO, School of Electrical Engineering, Purdue University, West Lafayette, IN 47907, USA.

An increasing number of profoundly deaf patients suffering from sensorineural deafness are using cochlear implants as prostheses. After the implant, sound can be detected through the electrical stimulation of the remaining peripheral auditory nervous system. Although great progress has been achieved in this area, no useful speech recognition has been attained with either single or multiple channel cochlear implants. Since cochlear implants are replacing natural receptors, the signal provided by the implants must be coded to contain the same attributes of the natural receptors. An artificial neural computing model of the receptor functions is presented. The model computes some known responses of the

sensorineural system using an artificial neural network. The model, its implementation, and speech recognition results are presented. The objective is to create a portable artificial cochlear for use as prostheses for patients who have sensorineural deafness caused by hair cell disease or damage, which would permit them to perform useful speech recognition.

#### **PRESYNAPTIC NEURAL INFORMATION PROCESSING,**

L. RICHARD CARLEY, Department of Electrical and Computer Engineering, Carnegie Mellon University, Pittsburgh, PA 15213, USA.

A possible mechanism for information processing within the arbor of a single axon (presynaptic neural information processing) based on measurements of the activity dependence of the firing threshold, the conditions required for conduction failure, and the similarity of nodes along a single axon will be presented in this paper. Recent studies report that the primary effect of several common anesthetics is to abolish the activity dependence of the firing threshold without interfering with impulse conduction, suggesting that presynaptic processing may play an important role in human consciousness. Possible electronic analogs of presynaptic information processing will be discussed.

#### **ADAPTATIONS OF THE ERROR PROPAGATION LEARNING ALGORITHM FOR VLSI IMPLEMENTATION,**

TOM BAKER, DAN HAMMERSTROM and HAL MCCARTOR, Department of Computer Science and Engineering, Oregon Graduate Center, Beaverton, OR 97006-1999, USA.

Connectionist network learning algorithms influence the physical architecture of VLSI or wafer scale implementations. One promising learning technique for connectionist networks is the Error Propagation method proposed by Rumelhart, Hinton, and Williams. This paper addresses some difficulties in a silicon implementation of this learning theorem, and presents possible solutions. Our research has shown:

1. Strictly sequenced node evaluations are not required.
2. Floating point computations can be replaced with integer calculations.
3. Unique error messages for each connection are unnecessary.

#### **OPTIMAL NEURAL SPIKE CLASSIFICATION,**

AMIR ATIYA, Department of Electrical Engineering, and JAMES BOWER, Division of Biology, California Institute of Technology, Pasadena, CA 91125, USA.

Extracellular microelectrode recording of neural signals is a useful way for studying small populations of nerve cells. The recorded signal contains usually spikes from several neurons adjacent to the electrode. This paper considers the problem of assigning the spikes, based upon their shapes, to the different neurons.

A method is proposed, which results in minimum probability of error, even in case of the occurrence of overlapping spikes. Various degrees of a priori knowledge could be taken into account by this method.

#### **A NEURAL-NETWORK SOLUTION TO A CONCENTRATOR ASSIGNMENT PROBLEM,**

GENE A. TAGLIARINI and EDWARD W. PAGE, Department of Computer Science, Clemson University, Clemson, SC 29634-1906, USA.

A network of simple analog processors having neuron-like properties have been employed to compute good solutions to a variety of optimization problems. This paper presents a neural-net solution to a resource allocation problem that arises in providing local access to the backbone of a wide-area communication network. The problem is described in terms of an energy function that can be mapped onto an analog computational network. Simulation results characterizing the performance of the algorithm are also presented.

#### **HOW THE CATFISH TRACKS ITS PREY: GUSTO-MODULATED RETICULOSPINAL NEURONS MAY GUIDE NAVIGATIONAL BEHAVIOR IN THE CHANNEL CATFISH, *ICTALURUS PUNCTATUS*,**

JAGMEET S. KANWAL, Department of Cellular and Structural Biology, University of Colorado, Denver, CO 80262, USA.

Ictalurid catfish use a highly developed gustatory system to localize food in the aquatic environment. In the present study, neuronal tracers enabled visualization of specific connections between the gustatory (sensory) inputs and the medullary (premotor) outputs to the spinal cord. The neuroanatomical and electrophysiological data indicate the presence of separate neural circuits which, while having a common input system, may be selectively activated during separate phases of the feeding behavior. The diverging-converging nature of the circuits and their selective information content suggest a unique 'gustation-seeking' mechanism for food-oriented navigation.

#### **CYCLES: A SIMULATION TOOL FOR STUDYING CYCLIC NEURAL NETWORKS,**

MICHAEL T. GATELY, Texas Instruments Incorporated, Dallas, TX 75243, USA.

A computer simulation program has been designed and implemented for the Texas Instruments Explorer/Odyssey computer system. This program allows a researcher to analyze the oscillatory behavior of neural networks with cyclic connectivity. The computer program, as well as the results of numerous experiments, will be discussed. The program, CYCLES, allows users to construct, operate, and inspect cyclic neural networks with the aid of a powerful graphics-based interface. Numerous cycles have been studied, including cycles with one or more activation points, non-interruptible cycles, cycles with variable path lengths, and interacting cycles. The final class, interacting cycles, are important due to their ability to implement goal processing in neural networks.

**SYNAPTIC INTERACTIONS BETWEEN NEURONS: FUNCTIONAL SIGNIFICANCE OF CORRELATIONAL LINKAGES,**

E.E. FETZ, Department of Physiology and Biophysics, University of Washington, Seattle, WA 98195, USA.

The strength and significance of synaptic interactions between neurons of the vertebrate motor system were investigated in physiological experiments and computer modelling.

Recordings in cat motoneurons revealed that excitatory postsynaptic potentials produce an increase in motoneuron firing rate in the cross-correlogram that is largely proportional to the EPSP derivative. A computer model was used to investigate how correlogram peaks are related to the EPSP shape and noise parameters. The model was validated by showing that simulations using intracellularly recorded EPSP's replicated the corresponding correlogram peaks obtained in cat motoneurons.

The cross-correlograms of neurons may be affected by polysynaptic as well as monosynaptic connections. The correlations produced by polysynaptic linkages are convolutions of the component monosynaptic linkages.

Recordings of motor cortex cells in behaving monkeys indicate that the response patterns of task-related cells are often remarkably independent of their correlational linkages. These data indicate that the activity propagated through neural networks of the vertebrate motor system is not strongly constrained by any particular synaptic linkages.

# NOTES

## ORAL SESSION 07

### SIMULATION AND IMPLEMENTATION

#### **BASINS OF ATTRACTION FOR ELECTRONIC NEURAL NETWORKS,**

C.M. MARCUS and R.M. WESTERVELT, Division of Applied Sciences and Department of Physics, Harvard University, Cambridge, MA 02138, USA.

We have used an analog electronic network and computer models to study the basins of attraction and the trajectories through phase space of continuous-time neural networks. The analog network consists of 8 neurons with non-linear gains and controllable time delay. By raster-scanning initial conditions we can measure the basins of attraction in 2 dimensional slices through phase space. We find that as the idealizations characteristic of well understood neural network models are modified to more realistically describe hardware implimentable networks, the dynamics and the shapes of the basins of attraction can become very complicated. We present examples of how memory overloading, time delay, clipped interconnection matrices and random faults can result in irregular basins of attraction, including basins for spurious memories and collective oscillatory modes. We also discuss some techniques for relating the dynamics of small networks like ours to the global dynamics of large networks.

#### **TEMPORAL PATTERNS OF ACTIVITY IN NEURAL NETWORKS,**

PAOLO GAUDIANO, Department of Aerospace Engineering, University of Colorado, Boulder, CO 80309, USA.

Patterns of activity over real neural structures are known to exhibit time-dependent behavior. Most current models of neural nets not only ignore the temporal structure of activity patterns, but actually attempt to eliminate temporal irregularities in the patterns by means of delay lines and other devices that will synchronize activity in the nets.

Using computer programs that simulate highly-realistic populations of spatially-organized neural nets, it was found that certain types of small nets can be shown to exhibit highly structured temporal responses to external stimuli. This is characterized by the net's tendency to fall into a limit cycle some time after presentation of input. Furthermore, a net that learns a specific spatial pattern is able to recall the same spatial pattern in a variety of distinct temporal modes.

These results seem to indicate that nets can use purely temporal patterns to discriminate between input stimuli. Such nets may be useful in performing tasks such as continuous speech recognition, which are not easily handled by other models because of temporal variability in the input.

## **OPTIMAL MAPPING OF THE CONFIGURATION SPACE OF A NETWORK OF NEURONS,**

G.J. CARMAN, B. RASNOW and J.M. BOWER, Divisions of Biology and Physics, California Institute of Technology, Pasadena, CA 91125, USA.

We have used simulated annealing to map the  $N$  dimensional configuration space of a network of  $N$  binary neurons into observation spaces of three or four dimensions. Our mapping minimizes the difference between distances over a neighborhood in the configuration space and the corresponding distances in the observation space for a choice of a metric in each space. In one such mapping, we chose to represent the Hamming distance in the configuration space by Euclidean distance in the observation space. The distribution of points on the resultant maps approximates a spherical surface, and shows a hierarchical organization which is a consequence of the approximate ultrametricity of the configuration space. Application of this mapping to the analysis of simulated and real neural networks will be discussed.

## **NEURAL NET AND CONVENTIONAL CLASSIFIERS,**

WILLIAM Y. HUANG and RICHARD P. LIPPMANN, Lincoln Laboratory, Massachusetts Institute of Technology, Lexington, MA 02173, USA.

Multi-layer perceptron classifiers with continuous-valued inputs can form complex decision regions. Previous work led to generative procedures to construct convex decision regions with two-layer perceptrons (one hidden layer) and arbitrary decision regions with three-layer perceptrons. We demonstrate that two-layer perceptron classifiers trained with back propagation can form both convex and disjoint decision regions. Such training, however, often takes excessively long and error rates are often higher than those provided by more traditional  $k$ -nearest neighbor classifiers. Three neural net classifiers are presented that provide improved performance. Two use fixed or random weights in the first one or two layers. A third can perform  $k$ -nearest neighbor classification.

## **NEURAL NETWORK IMPLEMENTATION APPROACHES FOR THE CONNECTION MACHINE,**

NATHAN H. BROWN, JR., MJR/Perkin-Elmer, Oakton, VA 22124, USA.

The SIMD parallelism of the Connection Machine (CM) allows the construction of neural network simulations by the use of simple data and control structures. Two approaches are described which allow parallel computation of a model's nonlinear functions, parallel modification of a model's weights, and parallel propagation of a model's activation and error. Each approach also allows a model's interconnect structure to be physically dynamic. A Hopfield model is implemented with each approach at three sizes over the same number of CM processors to provide a performance comparison.

## NOTES

## ORAL SESSION 08

### TECHNOLOGY

#### MICROELECTRONIC IMPLEMENTATION OF CONNECTIONIST NEURAL NETWORK MODELS,

H.P. GRAF, W.S. MACKIE and J.S. DENKER, AT&T Bell Laboratories, Holmdel, NJ 07733, USA.

Experiments on a chip containing 54 amplifiers with programmable interconnections led to two new chip designs tailored for the associated memory task. One is a streamlined version of the previous chip, designed to store 32 100-bit vectors and find the best match to an input vector in around 100ns. The other design is an all digital pipeline chip with a bit-serial processing unit attached to each stored vector. It remains to be seen which approach will be the more useful in applications, and how closely, if at all, circuits should resemble the models they represent.

#### A PROGRAMMABLE BINARY SYNAPTIC MATRIX CHIP FOR ELECTRONIC NEURAL NETWORKS,

A. MOOPENN, H. LANGENBACHER, A.P. THAKOOR, and S.K. KHANNA, Jet Propulsion Laboratory, Pasadena, California 91109, USA.

A binary synaptic matrix chip for electronic neural networks has been designed and fabricated. The matrix chip, implemented in 3-micron bulk CMOS, contains a programmable 32x32 array of NMOSFET's each functioning as a binary resistive connection. The matrix chip is cascadable for expansion to a binary resistive connection matrix up to 512x512 in size. A 32-neuron breadboard system has been built to test and evaluate the chip operation. The performance of the breadboard system in several test applications is discussed. A smaller version of the matrix chip consisting of an array of thin film resistors for the ON connections has also been designed and fabricated. The thin film resistors were deposited in an additional special process step to the normal CMOS fabrication run. The performance of various candidate materials for the resistive elements is also discussed.

#### BIT - SERIAL VLSI NEURAL NETWORKS,

ALAN F. MURRAY, ANTHONY V.W. SMITH and ZOE F. BUTLER, Department of Electrical Engineering, University of Edinburgh, EH9 3JL Edinburgh, Scotland.

We describe the implementation of a class of neural networks as VLSI devices, using synchronous, bit - serial pipelined arithmetic. The neural activation function is an approximation to a sigmoid that is piecewise - linear in the form of a staircase function with threshold and width control. Synaptic weights ( $T_{ij}$ ) are fully programmable. The psuedo - sigmoid activation function protracts but improves learning over simple threshold activation, and weight saturation during learning is best left for the learning algorithm to deal with naturally.

## NEW HARDWARE FOR MASSIVE NEURAL NETWORKS,

A.G.U. PERERA and D.D. COON, Department of Physics, University of Pittsburgh, Pittsburgh, PA 15260, USA.

Transient phenomena associated with silicon  $p^+n-n^+$  structures at 4.2K show remarkable similarities with biological neurons. The devices play a role similar to the two-terminal switching elements in Hodgkin-Huxley equivalent circuit diagrams. The devices provide simpler and more realistic neuron emulation than transistors or op-amps. They have such low power and current requirements that they could be used in massive neural networks. Some observed properties of simple circuits containing the devices include action potentials, refractory periods, threshold behavior, excitation, inhibition, summation over synaptic inputs, synaptic weights, temporal integration and memory.

## NOTES

## ORAL SESSION 09

### ANALYSIS OF NEURAL NETWORKS

#### ON PROPERTIES OF NETWORKS OF NEURON-LIKE ELEMENTS,

SANTOSH S. VENKATESH, Moore School of Electrical Engineering, University of Pennsylvania, Philadelphia, PA 19104, USA, and PIERRE BALDI, Department of Mathematics, University of California (San Diego), La Jolla, CA 92093, USA.

Quadratic Hamiltonians describing the dynamics of single-layered neural networks and spin glasses have a natural extension to algebraic Hamiltonians of higher degree. Systems subscribing to dynamics regulated by such algebraic Hamiltonians include random structures such as higher-order spin glasses, as well as dynamically programmed structures such as multi-layered neural nets. In this paper fundamental limits are posited for such higher-order structures. The static behaviour of such networks is examined using ideas from geometric probability to obtain rigorous estimates of their *intrinsic* information storage capacities. Network dynamics are characterised using landscapes of algebraic Hamiltonians. Formal results on learning rates and generalisation, and algorithms and storage rules approaching fundamental limits are presented. The complexity of feedforward computational circuits is examined, and classes of functions that can be implemented in shallow circuits characterised.

#### PHASE TRANSITIONS IN NEURAL NETWORKS,

JOSHUA CHOVER, Department of Mathematics, University of Wisconsin, Madison, WI 53706, USA.

Various simulations of cortical subnetworks have evidenced something like phase transitions with respect to key parameters. We demonstrate that such transitions must indeed exist in analogous infinite array models. For related finite array models classical phase transitions (which describe steady-state behavior) may not exist, but there can be distinct qualitative changes in ('metastable') transient behavior as key system parameters pass through critical values.

#### ANALYSIS OF DISTRIBUTED REPRESENTATION OF CONSTITUENT STRUCTURE IN CONNECTIONIST SYSTEMS,

PAUL SMOLENSKY, Department of Computer Science, University of Colorado, Boulder, CO 80309-0430, USA.

Any model of complex information processing in networks of simple processors must solve the problem of representing complex structures over network elements. This representational problem is not well understood, and this has been a principal criticism of the connectionist approach to AI. This paper reports a general analysis of the connectionist representation of complex symbolic structures (e.g., strings, stacks, trees). A general scheme is formulated, the *tensor product representation*,

which generalizes a number of the representational approaches found in existing connectionist systems. The tensor product representation can be proved to possess a number of the properties that are required of a computationally adequate connectionist representation.

#### **PATTERN CLASS DEGENERACY IN AN UNRESTRICTED STORAGE DENSITY MEMORY,**

CHRISTOPHER L. SCOFIELD, Nestor Inc., Providence, RI 02906, USA, and DOUGLAS L. REILLY, CHARLES ELBAUM and LEON N. COOPER, Center for Neural Science and Department of Physics, Brown University, Providence, RI 02912, USA.

We discuss an extension of the method of Reilly et al. as the square well limit of the N-dimensional Coulomb potential of Bachmann et al., and introduce a general method of state space tunnelling when square well potentials are close or overlap. We define the degeneracy problem as the existence of points in a feature or representation space which may have multiple class labels associated with them. We show that the state space tunnelling gives the statistically correct classification of new patterns even under conditions of complete class degeneracy.

#### **ON THE POWER OF NEURAL NETWORKS FOR SOLVING HARD PROBLEMS,**

JEHOSHUA BRUCK and JOSEPH W. GOODMAN, Department of Electrical Engineering, Stanford University, Stanford, CA 94305, USA.

This paper deals with a neural network model in which each neuron performs a threshold logic function. An important property of the model is that it always converges to a stable state when operating in a serial mode. This property is the basis of the potential applications of the model such as associative memory devices and combinational optimization.

One of the motivations for use of the model for solving hard combinational problems is the fact that it can be implemented by optical devices and thus operate at a higher speed than conventional electronics.

The main theme in this work is to investigate the power of the model for solving NP-hard problems, and to understand the relation between speed of operation and the size of a neural network. In particular, it will be shown that:

1. A network with polynomial (in the size of the input) number of neurons can not solve an NP-hard problem even if it operates for an exponential length of time (unless NP = co-NP).
2. A network with polynomial (in the size of the input) number of neurons which always gets to an epsilon-approximate solution for the Travelling Salesman Problem (TSP) does not exist unless P = NP.

The above results are of great practical interest, because right now it is possible to build neural networks which will operate fast but are limited in the number of neurons.

## NOTES

## POSTER SESSION P4

### SCALING PROPERTIES OF COARSE-CODED SYMBOL MEMORIES,

RONALD ROSENFELD and DAVID S. TOURETZKY, Computer Science Department, Carnegie Mellon University, Pittsburgh, PA 15213, USA.

Coarse coded memories have appeared in several neural network symbol processing models, such as Touretzky and Hinton's distributed connectionist production system DCPS, Touretzky's distributed implementation of Lisp S-expressions on a Boltzmann machine, BoltzCONS, and St. John and McClelland's PDP model of case role defaults. In order to determine how these models would scale, one must first have some understanding of the mathematics of coarse coded representations. We characterize a random coarse coded memory by its alphabet size, number of units, memory capacity, and probability of false memories, and derive mathematical relationships between these parameters. We then compare this approach with other, non-random approaches.

### SELF-ORGANIZATION OF ASSOCIATIVE DATA-BASE AND ITS APPLICATIONS,

HISASHI SUZUKI and SUGURU ARIMOTO, Mechanical Engineering, Engineering Science, Osaka University, 560 Osaka, Japan.

A simple and efficient method of organizing associative data-bases is proposed with some applications to learning systems. The data-bases associate input data with some related information in memory. In the first half part of discussion, an algorithm of self-organization is constructed, and the complexity of computation and other factors of performance is evaluated mathematically. In the latter half part, an applicability to typographic and handwritten letter recognition and that to an autonomous mobile robot system are demonstrated.

### MEMORY CAPACITY IN SYMMETRIC NEURAL NETWORKS: RIGOROUS BOUNDS,

CHARLES M. NEWMAN, Department of Mathematics, University of Arizona, Tucson, AZ 85721, USA.

We consider neural networks with  $N$  binary neurons and symmetric  $l^{\text{th}}$  order synaptic connections, in which  $m$  randomly chosen  $N$ -bit patterns are stored and retrieved with a small fraction  $\delta$  of bit errors allowed. We prove rigorous versions of the following:

1.  $M$  can grow as fast as  $\alpha N^{l-1}$ .
2.  $\delta = 0(\exp(-B_l/\alpha))$  for small  $\alpha$ .

Our analysis yields rigorous lower bounds for the maximum possible value of  $(M/\alpha)$ , the number of stored bits per distinct synapse: 0.11 for  $l = 2$  (compared to 0.29 as estimated by Hopfield and Amit et al.), 0.22 for  $l = 3$  and 0.16 for  $l = 4$ .

## **STABILITY RESULTS FOR NEURAL NETWORKS,**

A.N. MICHEL, W. POROD and J.A. FARRELL, Department of Electrical and Computer Engineering, University of Notre Dame, Notre Dame, IN 46556, USA.

We present new results for the asymptotic stability, exponential stability, instability, and complete instability of equilibrium points in neural networks. Also, we present estimates of the domain of attraction of an asymptotically stable equilibrium in such networks, as well as estimates of trajectory bounds. In our approach, we view neural networks as an interconnection of free subsystems and we make use of Lyapunov functions not previously used. Our results are phrased in terms of the qualitative properties of the free subsystems and in terms of the qualitative properties of the system interconnecting structure. Furthermore, our results do not require a symmetric interconnecting structure of the neural networks and some of our results apply to systems with structural perturbations. The methodology advanced herein is applicable to several models of neural networks. However, we will primarily consider the Hopfield model.

## **ANALYSIS AND COMPARISON OF DIFFERENT LEARNING ALGORITHMS FOR PATTERN ASSOCIATION PROBLEMS,**

J. BERNASCONI, Brown Boveri Research Center, CH-5405 Baden, Switzerland.

We investigate and compare the properties of different learning algorithms for networks on neuron-like units. As test cases we use simple pattern association problems, such as the XOR-problem and symmetry detection problems, and the algorithms considered are either versions of the Boltzmann machine learning rule or based on the backpropagation of errors. In particular, we propose and analyze a generalized delta rule for linear threshold units. We find that the learning speed and the sensitivity with respect to parameter variations not only depends on the learning algorithm and on the network structure, but also on the type of units used.

## **THE CONNECTIVITY REQUIREMENTS OF SIMPLE ASSOCIATION -OR- HOW MANY CONNECTIONS DO YOU NEED?,**

DAN HAMMERSTROM, Department of Computer Science and Engineering, Oregon Graduate Center, Beaverton, Oregon 97006-1999, USA.

The efficient realization, using current silicon technology, of neural networks with more than a billion connections requires the networks exhibit a high degree of communication locality. A network exhibits *locality of communication* if most of its processing elements connect to physically adjacent nodes in any reasonable mapping onto a planar surface. Real neural networks exhibit significant locality, yet most connectionist/neural network models have little. In this paper a communication channel model of simple associative processing in a massively parallel network is presented. Using this model, the effects of a node's receptive field size and predicate order on the associative capacity of the network can be analyzed.

## MAPPING OPTIMIZATION PROBLEMS ONTO ARTIFICIAL NEURAL NETWORK SYSTEMS: A GENERAL PRINCIPLE AND A COMPARISON TO GRADIENT DESCENT (ASCENT),

HARRISON MONFOOK LEONG, Research Institute for Advanced Computer Science, NASA Ames Research Center, Moffett Field, CA 94035, USA.

General formulae for mapping optimization problems into systems of ordinary differential equations arising from artificial neural network research are presented. A comparison is made to optimization using dynamical systems that represent gradient-descent (ascent) methods. The performance measure used is the time required for the systems to move from an initial state to a target state. Two analytical examples illustrate situations where dynamical systems representing artificial neural network methods would settle faster than those representing gradient-descent (ascent). Computer simulations used to estimate settling times for finding minima of a multidimensional quadratic surface having many local minima suggest that gradient-descent systems can settle 1 to 2 orders of magnitude faster than systems based on current neural network optimization methods.

## CONNECTING TO THE PAST,

BRUCE A. MACDONALD, Department of Computer Science, The University of Calgary, T2N 1N4 Calgary, Alberta, Canada.

Recently there has been renewed interest in neural-like processing systems, evidenced for example in the two volumes *Parallel Distributed Processing* (MIT Press, 1986) edited by Rumelhart and McClelland, and discussed as parallel distributed systems, connectionist models, neural nets, value passing systems and multiple context learning systems. Dissatisfaction with symbolic manipulation paradigms for artificial intelligence seems partly responsible for this attention, encouraged by the promise of massively parallel systems implemented in hardware. This paper relates simple neural-like systems to some other well-known notions - namely production systems, k-length sequence prediction, finite-state machines and Turing machines - and presents earlier sequence prediction results in a new light.

## THE CAPACITY OF THE KANERVA ASSOCIATIVE MEMORY,

P.A. CHOU and R.M. GRAY, Information Systems Laboratory, Stanford University, Stanford, CA 94305, USA.

We define the capacity of the Kanerva associative memory to be the logarithm of the number of  $\rho$ -stable words that can be stored at random, divided by the wordlength  $n$ , in the limit as  $n \rightarrow \infty$ . A word (record) is  $\rho$ -stable if, when retrieved by a key with fewer than  $n\rho$  bit errors, the probability of a word error is arbitrarily small when the wordlength is large. Under these definitions, the capacity of the Kanerva associative memory is equal to  $1 - h_2(\rho)$ , where  $h_2(\cdot)$  is the binary entropy function. This is in sharp contrast to the capacity of the Hopfield neural net, which is zero under these definitions.

## CONNECTIVITY VERSUS ENTROPY,

YASER S. ABU-MOSTAFA, Department of Electrical Engineering, California Institute of Technology, Pasadena, CA 91125, USA.

How does the connectivity of a neural network (number of synapses per neuron) relate to the complexity of the problems it is capable of solving? Switching theory would suggest no relation at all, since any Boolean function can in principle be implemented using a circuit with very low connectivity (e.g., NAND gates). However under the assumption that the circuit itself has to learn the function it is to implement without global supervision, the entropy of the function becomes a lower bound for the connectivity of the circuit.

## PROBABILISTIC CHARACTERIZATION OF NEURAL MODEL COMPUTATIONS,

RICHARD GOLDEN, Learning Research and Development Center, University of Pittsburgh, Pittsburgh, PA 15260, USA.

Information retrieval in a neural network is a procedure in which the network computes a MAP estimate of the unknown information with respect to an a posteriori density function,  $P$ . With a minimal number of assumptions, the 'energy' function that a neural network model minimizes during information retrieval is shown to *uniquely* specify the form of  $P$ . Inspection of the form of  $P$  reveals the class of probabilistic environments that can be learned. Optimal learning algorithms can be designed by using maximum likelihood estimation techniques to estimate the parameters of  $P$ . Assumed PDFs are constructed for the nonlinear auto-associative network models of Cohen and Grossberg (1983), Hopfield (1982, 1984), and Anderson (1977), nonlinear associative networks created using the back-propagation algorithm (Rumelhart et al., 1986), and certain classes of nonlinear multi-stage networks.

## PERFORMANCE MEASURES FOR ASSOCIATIVE MEMORIES THAT LEARN AND FORGET,

ANTHONY KUH, Department of Electrical Engineering, University of Hawaii at Manoa, Honolulu, HI 96822, USA.

Recently, many modifications to the Hopfield Associative Memory Network have been proposed where both learning and forgetting occur. Given that the network never saturates (ceases to function effectively due to an overload of information), the learning updates can continue indefinitely. For these networks, we need to introduce performance measures in addition to the information capacity to evaluate the different networks. We mathematically define quantities such as the plasticity of a network, the efficacy of an information vector, and the probability of network saturation. From these quantities we compare the different networks both analytically and through simulations.

## **INFORMATION STORAGE CAPACITY OF CONNECTIONIST SYSTEMS: THE LINEAR ASSOCIATOR,**

DEAN C. MUMME and WALTER SCHNELDER, Learning Research and Development Center, University of Pittsburgh, Pittsburgh, PA 15260, USA.

The information-storage capacity of hetero-associative memory systems is addressed. The associator can be treated as an M-ary symmetric channel when M associations are stored. The maximum number of associations storable is bounded asymptotically by  $N/2$  where N is the number of connection weights. Storage efficiency is bounded by  $M/N$  so that it never exceeds  $1/2$ . Information capacity degrades as inter-vector correlations increase and also when classification tasks are performed. The correlation effect is most pronounced in high-dimensional systems storing a large number of associations. Specifically, when vectors of  $\pm 1$ 's are stored, the probability that a vector component is 1 must be within the range  $1/2 \pm 2/\sqrt{3} \cdot M^{-1/4}$  for optimal capacity.

## **'ENSEMBLE' BOLTZMANN UNITS HAVE COLLECTIVE COMPUTATIONAL PROPERTIES LIKE THOSE OF GRADED HOPFIELD NEURONS,**

MARK A. DERTHICK, Department of Computer Science, Carnegie Mellon University, Pittsburgh, PA 15213, USA.

[Hopfield 84] shows that a deterministic neural network model with continuous outputs on  $[0,1]$  serves as a mean field approximation to his earlier stochastic model in which the outputs were binary valued. This paper presents an analogous result for Boltzmann Machines. For any Boltzmann Network there is also another Boltzmann Machine which can be interpreted as its mean field approximation. This network has essentially the same deterministic, continuous behavior as the Hopfield and Tank model. There is a single parameter which determines whether the behavior is completely stochastic, deterministic, or intermediate in character. Changing this parameter allows the experimenter to do his debugging on a deterministic system, while allowing the more robust behavior to be had with a stochastic system at other times.

## **MATHEMATICAL ANALYSIS OF LEARNING BEHAVIOR OF NEURONAL MODELS,**

MASSOUD OMIDVAR and JOHN Y. CHEUNG, School of Electrical Engineering and Computer Science, University of Oklahoma, Norman, OK 73019, USA.

Neural adaptation, or learning is an important aspect in the analysis of neuronal behavior. In the present study, we have developed mathematical tools to analyze some of these models to predict the behavior and outputs of neuronal models with respect to conditioned and unconditioned stimuli. Specifically, two analytical approaches are used: the first is the Z-transform approach and the second is based on nonlinear analysis approach. As an example, the Hebb model is used for illustration. The synaptic strength and neuronal output based on the Hebb model can be completely predicted using these two approaches. The developed techniques can be used to study the learning rate and convergence criteria of the various models.

### **AN OPTIMIZATION NETWORK FOR MATRIX INVERSION,**

**JU-SEOG JANG, SOO-YOUNG LEE and SANG-YUNG SHIN,** Department of Electrical Engineering, Korea Advanced Institute of Science and Technology, Seoul, Korea.

Inverse matrix calculation can be considered as an optimization problem. We have demonstrated that this problem can be rapidly solved by highly interconnected simple neuron-like analog processors. A network for matrix inversion was designed based on the concept of Hopfield's neural network, and implemented with electronic hardwares. With slight modifications, the network is readily applicable to solving a matrix equation  $Ax=b$  efficiently. Notable features of this circuit are potential speed due to parallel processing, and robustness against variations of device parameters.

### **INFORMATION CAPACITY OF HEBBIAN NEURAL NETWORKS,**

**JAMES D. KEELER,** Institute for Nonlinear Science, University of California (San Diego), La Jolla, CA 92093, USA.

The information capacity is investigated for two classes of neural network models using the Hebbian (outer-product) storage rule. A generalization of Hopfield-type models using higher-order interactions is analyzed, as well as a similar generalization of a three layer network that uses Hebbian learning among the second and third layer. It is shown that the total information stored in these systems is a constant times the number connections in the network, independent of the particular model, the order of the model, or whether clipped weights are used or not.

## NOTES