

N88-17220

NASA JSC NEURAL NETWORK SURVEY RESULTS

DAN GREENWOOD, NETROLOGIC, INC., 4241 Jutland Drive, San Diego, CA 92117

ABSTRACT

VERAC conducted a survey of Artificial Neural Systems in support of NASA's (Johnson Space Center) Automatic Perception for Mission Planning and Flight Control Research Program. Several of the world's leading researchers contributed papers containing their most recent results on Artificial Neural Systems. These papers were broken into categories and descriptive accounts of the results make up a large part of this report. Also included is material on sources of information on Artificial Neural Systems such as books, technical reports, software tools, etc. This paper is an abridged version of the report to NASA.

ACKNOWLEDGEMENT

VERAC gratefully acknowledges the contributions of all the participants of this survey and to Professor **Terence Smith** of the University of California at Santa Barbara and Dr. **Christopher Bowman** of VERAC for their advice concerning the approach to the survey and elucidating many subtle points concerning the subject matter. **Cris Kobryn** played a key role in obtaining survey data and analyzing public domain Artificial Neural Systems software. The foresight of our NASA sponsors, **Robert Savely** and **James Villareal**, in realizing the importance and promise of this exciting new field is greatly appreciated. **Sue Lani Andrassy** and **Susan Foster**, of VERAC, kept everything under control and imparted order where there was chaos.

1.0 INTRODUCTION

Artificial Neural Systems (ANS's) have captured the interest of many computer scientists, robotic engineers, mathematicians, and neurophysiologists as a result of progress made in solving problems which have eluded solution by conventional computer approaches. The aim of this study was to establish a database for NASA which contains the recent results of researchers in this rapidly growing and dynamic field. Since the field is so dynamic and the time to publish current research is often protracted, it was decided to broadcast an appeal for contributions of recent papers/reports or pre-prints to help form NASA's database and this survey. The response was overwhelming both from the standpoint of quantity and quality. So much so that the initial plan to perform a survey approaching the standards set by **Daniel Levine** (LEVINE) was soon abandoned and it was decided to:

- 1) Describe the papers as they appeared to us,
- 2) Not evaluate the papers or try to provide an integrated point of when different authors covered related ANS subject matter.

No effort was made to ascertain the accuracy of data or the validity of mathematics from any of the papers and, many of the papers were submitted as preliminary versions. The topics were broken down into the following areas:

- 1) ANS theory, 2) Computation and optimization, 3) Memory, 4) Learning, 5) Pattern recognition, 6) Speech, 7) Vision, 8) Knowledge processing, 9) Robotics/control, which reflects the format of the International Conference on Neural Networks.

The phrase, Artificial Neural Systems, was selected for this study over: connectionist models, parallel distributed processing, and neural networks, although neural networks seems to be gaining the edge in terms of general acceptance and preference. Perhaps, it is not too late to introduce yet another word to encompass the same meaning associated with the above terms and even a little more. The word "Netrology" seems to be one which includes neural networks and possible expansions which supersede neural networks as they are commonly understood. It would encompass, for example, units which are not neuron-like in their behavior but which, nevertheless, exhibit interesting or useful properties. It seems that a key component belonging to networks falling under the concept of netrology should be that a network be able to learn; thereby circumventing the ordinarily difficult problem of programming on ensemble of parallel processors.

Based on reviewing the papers submitted in support of this study, the following issues are considered to be of importance to future progress in netrology:

- 1) A rigorous definition of "structure" or "regularity" which is often attributed to networks which discover features. Psychophysical measurements and fractal concepts (such as fractal dimension) will probably be necessary to define net "structure" rigorously.
- 2) The construction of netrological experiments and concepts which help to define ANS situational awareness, task management, and planning.
- 3) A rigorous definition of similarity corresponding to the efforts made in numerical taxonomy and classical pattern recognition so that net recall of "distorted" images, taxonomy and classical pattern recognition so that net recall of "distorted" images, etc. really corresponds to the goals of an application. Nets may have to be more or less discriminating per application.
- 4) The integration of sensory data from sensors of the same or different types (e.g., nets with three eyes and four ears) and a priori data concerning the environment and constraints.
- 5) Endowing nets with desirable human-like traits such as artificial modesty, humor, perseverance, honesty, etc. This will be of importance in merging net-workers with human workers in real world industrial, academic and military applications (user friendliness/congeniality is the goal here).
- 6) Establishing bounds of net autonomy. Asimov's robotic laws are anthropocentric. Future neural networks may look more kindly at us early designers if we make an effort to ensure their autonomy and provide the means for gratifying their creative instincts.
- 7) Training humans to be tolerant and accepting of net solutions to problems. Ohm was ridiculed for 30 years, and everyone knows about Galileo, so this issue is not as farfetched as it may seem.
- 8) Establishing a taxonomy of computational devices which shows which problem domains are best suited for systolic arrays, neural networks, symbolic processors, signal processors and conventional processors.
- 9) Establishing neural net design rules which facilitate configuring a neural net per problem application.
- 10) Establishing ANS figures of merit so the value of a particular learning rule or net design can be meaningfully estimated.

Considering issues 4 through 7 is certainly fun — now back to more immediate concerns. This ANS review contains an overview of the small but rapidly growing number of commercial ANS products, public-domain research tools, and some ANS books and educational materials. (LIPPMANN) contains one of the best short introductions as well as a penetrating analysis of ANS's as they are today, and, undoubtedly, the proceedings of the 1987 International Conference on Neural Networks will represent the state-of-the-art for ANS's when it is published.

2.0 GENERAL SOURCES OF INFORMATION ON ANS's

Since the recent re-birth of interest in ANS's, there has been a virtual flood of papers in engineering as well as scientific journals. So many technical papers on ANS's currently exist that are scattered among journals such as **Biological Cybernetics**, **Behavioral and Brain Sciences**, **Psychological Review**, and the **Journal of the**

National Academy of Sciences, that retrieving the papers alone is a time consuming and tediously difficult problem. Once a given paper is retrieved, it is then a major task to decipher the often new definitions, notation and technical style. For an engineer whose sole interest is to understand the potential of an ANS to solve a problem in pattern recognition or image understanding, the neurophysiological as well as the psychophysical flavor of many of the often-cited articles poses a major obstacle. The small number of books that exist for the most part are collections of papers submitted to technical journals or conferences. ANS courses are given at a few universities and there is a video tape of Dr. Robert Hecht-Nielsen's week-long course on ANS's. Training sessions are available both at TRW and HNC with the purchase of TRW's Mark-III Neurocomputer and HNC's ANZA Neurocomputer, respectively.

In view of the widely scattered and varied information on ANS's, this section will be devoted to describing that information on ANS's which is of a general, educational nature. The remaining sections are reserved for presenting the results of the survey based on the papers, reports and discussions generously provided by researchers throughout the world.

2.1 ANS Books

The books listed below are available for purchase and in many university libraries. There are, as yet, no textbooks on ANS's although the books: "Self-Organization and Associative Memory", by T. Kohonen, and "Parallel Distributed Processing", Volumes 1 and 2, by Rumelhart, McClelland and the PDP Research Group are semblances of textbooks.

2.1.1 Parallel Distributed Processing (Volumes 1 and 2), (MIT Press, 1986)

These volumes are the best introductory books to the field, the members of the PDP research group at the Institute for Cognitive Science at the University of California at San Diego, under the leadership of James McClelland and David Rumelhart, combined their talents to write (in various combinations of authors) both tutorial and research-oriented chapters on "Parallel Distributed Processing".

The history of ANS's is traced in a fair amount of detail and a wide range of related topics are covered. Basic mechanisms such as feature discovery by competitive learning, information processing in dynamical systems (Harmony Theory), Boltzmann machines, and back propagation are covered with many excellent examples.

2.1.2 Self-Organization and Associative Memory (T. Kohonen, Springer-Verlag, 1984)

This book was written before the days of back propagation and is mainly concerned with linear transformations. Even with these restrictions, it is a good source on adaptive filters, optimal associative mappings, and self-organizing feature maps. There is a good discussion of, with examples of topology preserving mappings but in general, many of the applications and alternative approaches in ANS's are not considered. The book complements the "Parallel Distributed Processing" book as a result of the extra attention to mathematical rigor and its linear systems perspective.

2.1.3 Parallel Models of Associative Memory (G.D. Hinton, J.A. Anderson, Lawrence Erlbaum Associates, 1981)

This book contains a collection of papers by well-known researchers in ANS such as T. Sejnowski, S. Fahlman, G. Hinton, etc. Topics covered are: models of information processing in the brain, a connectionist model of visual memory, holography, distributed associative memory, representing implicit knowledge, implementing semantic networks in hardware, and many other topics.

2.1.4 Neural Networks for Computing (G.S. Denker, Editor, American Institute of Physics, 1986)

This book contains 64 short papers by leading ANS researchers. The papers encompass applications, mathematical theory, implementations, and biological modeling. A paper by Lapedes and Farabeault presented an interesting method for circumventing the limitations of a Hopfield Network. Another paper by Personnaz, et al. introduces a simple modification to Hebbian learning to give a more biologically plausible selectionist learning scheme.

2.1.5 Brain Theory: Proceedings of the First Trieste Meeting on Brain Theory, 1984 (G. Palm, A. Aertsen, Editors, Springer-Verlag)

"Brain Theory" contains papers by researchers primarily concerned with the workings of the brain itself and, secondarily, with methods for defining and exploiting information processing principles obtained along the way to understanding brain operations.

2.1.6 Competition and Cooperation in Neural Nets (S. Amari, M. Arbib, Editors, Springer-Verlag, 1982)

The proceedings of a 1982 conference on neural nets are presented in this book. Leading neural net theoreticians and brain theorists such as S. Grossberg, M. Arbib, A. Pellionisz, T. Kohonen, S. Amari presented papers at the conference.

2.1.7 The Adaptive Brain (Stephen Grossberg, Editor, North Holland, 1987)

Professor Grossberg and members of the Center for Adaptive Systems at Boston University (which Grossberg leads) wrote the papers for this highly theoretical book. Chapters of the book cover: psychophysiological theory of reinforcement, drive motivation, and attention, psychophysiological and pharmacological correlates of a developmental cognitive and motivational theory, conditioning and attention, memory consolidation, a neural theory of circadian rhythms, and other topics.

2.2 Reports

2.2.1 How the Brain Works: The Next Generation of Scientific Revolution (by David Hestenes, Third Workshop on Maximum Entropy and Bayesian Methods in Applied Statistics, University of Wyoming, Aug 1-4, 1983)

Professor Hestenes, a mathematician from Arizona State University, was persuaded by a former student (now Dr. Robert Hecht-Nielsen of HNC, Inc.) to spend some time and hard work getting familiar with the work of Stephen Grossberg. Hestenes came away from his efforts as a firm believer in Grossberg's approaches and outlined the basis for his beliefs in a tutorial report dedicated exclusively to Grossberg's work.

2.2.2 Neural Network Models of Learning and Adaptation (J.S. Denker, AT&T Bell Laboratories, N.J.)

This Bell Labs technical report provides a good overview of neural network basics. Hopfield's ideas are clearly presented as are discussions of simple Hebbian learning, ADALINE, Geometric and Pseudo-Inverse rules, and the practical effects of clipping. The report ends with an interesting presentation of open questions in neural network theory.

2.2.3 Performance Limits of Optical, Electro-Optical, and Electronic Neurocomputers ("Optical and Hybrid Computing", SPIE, Vol. 634, 1986)

Hecht-Nielsen did an excellent job of summarizing neural network theory and implementations up to 1986. He covered ANS modeling philosophy, technology organization, theory, neurocomputers and their performance limits, the Cohen/Grossberg Adaptive Resonance Network Learning Theorem, Hopfield's and Kohonen's theories and their implications for implementation issues.

2.2.4 Neural Population Modeling and Psychology: A Review (D. Levine, Mathematical Biosciences, 66: 1983)

Professor Levine's excellent review is highly recommended for anyone interested in neural networks including those with either theoretical or applications oriented interest. This well written review addresses all of the major neural assembly models from 1938 to 1983. The works of Grossberg, Barto, Sutton, Klopff, Anderson, Uttley, von der Malsburg, and others are presented in a tutorial fashion and the significance of the respective models in relation to neurophysiological and psychological data is addressed in detail.

2.2.5 Stochastic Interated Genetic Hillclimbing (D. Ackley, March 1987, CMU-CS-87-107, Carnegie-Mellon University)

David Ackley's PhD dissertation contains a new method for performing function optimization in high dimensional binary vector spaces. The method can be compactly implemented in a neural network architecture and provides an effective network training rule which combined genetic search algorithms properties with hillclimbing algorithm properties.

2.2.6 A Survey of Artificial Neural Systems (P. Simpson, Unisys, San Diego)

Patrick Simpson of Unisys reported the results of a survey of ANS's in [SIMPSON, abs]. This survey, completed in early 1987, discusses some of the well known neural models and contains computer codes for different learning rules (Hebbian, Hopfield, Boltzmann) and recall rules. Many of the ANS's such as the Sejnowski/Rosenberg NETtalk, are described. A brief history of ANS's is also included.

2.2.7 Efficient Algorithms with Neural Network Behavior (S. Omohundro, April 1987) Report No. UIUCDCS-R-87-133, Department of Computer Science, University of Illinois at Urbana-Champaign

Although this report is not concerned with ANS's per se, it does discuss alternatives to ANS approaches and, in so doing, sheds light on the capabilities and properties of ANS's. Using hierarchical data structures well known in computer science (arrays, hashing, tries, trees, adaptive grids) Omohundro was able to show very significant implementation advantages in solving problems where ANS's are now being applied. In explaining why his data structure based approach is in many cases much more efficient than corresponding ANS approaches, Omohundro claims:

- 1) ANS's must evaluate each neuron's activity and consider the effect of each weight each time an input is made, while the "algorithm approach" only looks at stored values along a path of logarithm depth.
- 2) ANS learning requires that all weights be updated with each input, but data structures only modify parameters where regions are relevant in determining the output on the given input.

2.3 Other Sources of General ANS Information

The following media provide additional sources of ANS information:

2.3.1 Neuron Digest

Subscribers to ARPANET can avail themselves of neural network information on upcoming lectures, publications, conferences, abstracts, government research grants, opinions and needs. The *Neuron Digest* is distributed each month and is a great way to keep informed about this rapidly growing field.

2.3.2 A Video Tape on Artificial Neural System Design

A video tape of Dr. Robert Hecht-Nielsen's 5-day course on ANS design is available through HNC, Inc. or the University of California at San Diego through the University Extension. The course is an excellent way to get introduced to the state-of-the-art of ANS's by one of the field's leading experts and educators. The full spectrum of ANS's from deep theoretical issues (Grossberg, et al.) to practical ANS implementations are covered.

2.3.3 HNC, Inc. Month-Long Course on ANS's

A hands-on course in ANS is included in the purchase price of an HNC Neurocomputer (the ANZA). The course emphasizes the practical aspects of applying ANS methods to problems in sensor processing, knowledge processing, control, optimization, data base management and statistical analysis. The course is aimed at enabling students to become productive ANS experts in a short period of time.

2.3.4 TRW Mark-III Neurocomputer Training

TRW's ANS Center, headed by Michael Myers, provides a one-week training with the purchase of a TRW Mark-III Neurocomputer. The very powerful machine with a staff experienced at solving problems (much beyond the text-book variety of most applications of ANS's) provides an effective way for ANS's users to augment their skills.

3.0 IMPLEMENTATIONS OF ARTIFICIAL NEURAL SYSTEMS

The resurgence of interest in ANS's and critical evaluations of their potential have resulted in new commercial enterprises whose charters are to bring ANS products to market. There also exist activities in many businesses aimed at developing hardware and software available commercially and under development by members of the commercial sector.

3.1 Currently Available Commercial ANS Products

The following companies sell ANS products and may be contacted directly to obtain literature containing ANS product descriptions.

3.1.1 HNC (Hecht-Nielsen Neurocomputer Corporation)

HNC is a San Diego based company founded by Robert Hecht-Nielsen and Todd Guschow who developed the Mark series Neurocomputers while at TRW. HNC's main products are the ANZA Neurocomputer and the ANZA basic "Netware" (neurocomputer software) package. A month-long course is offered and course participants are expected to have a basic knowledge of college level mathematics, and problems of interest to them will be addressed in the course.

The Netware packages are loaded into the neurocomputer (in combinations of one or more), and their constants and parameters tuned and selected by the user to fit the application problem at hand.

3.1.2 Nestor, Incorporated

Nestor is a publicly traded ANS company which currently sells two products developed by nobel laureat, Leon Cooper, and Brown University physics professor, Charles Elbaum. The company is located in Providence, Rhode Island. ANS methods used in current Nestor products are also being considered for postal sorting, robotic mission systems, fingerprint and voice identification, speech and speaker identification, medical diagnostic systems, check processing and encoding and credit card identification/validation.

3.1.3 Neuraltech, Incorporated

Dr. John Voevodsky founded Neurotech, Incorporated and sells a software product called "PLATO/ARISTOTLE". The software package is a neural-based expert system for the IBM PC-AT and COMPAQ 286/386 personal computers.

3.1.4 Texas Instruments

Andrew Perry and Richard Wiggins of Texas Instruments developed a digital signal processor for accelerating neural network simulations [DENKER]. Using TI's Odyssey boards, they developed a system which was forty times faster than the VAX 8600 for ANS applications.

3.1.5 TRW

TRW sells a neurocomputer called the Mark-III and provides a one-week on-site (San Diego, CA) course in the purchase price. TRW researchers Michael Myers and Bob Kuczewski are defining the state-of-the-art in applications of ANS methods to signal processing, spatial temporal pattern learning, classification of time varying spectrograms, and image analysis.

3.1.6 SAIC's Σ -1 Neurocomputer

SAIC announced a Neurocomputer, called the " Σ -1", which is expected to be available in October of 1987. The machine was developed by a SAIC research team headed by Dr. James Solinsky, a renown computer vision researcher. the Σ -1 software includes shells for most well known learning rules.

3.2 Currently Available Public Domain ANS Products

At the present time software ANS simulation packages can be obtained without charge from Brown University and the University of Rochester for use by ANS researchers.

3.2.1 The Brown University ANS Simulation

Professor James A. Anderson of Brown University released an ANS simulator based on his "Brain State in a Box" neural network model. The software was developed over the last 12 years and continues to be a useful tool for ANS experimentation.

3.2.2 The University of Rochester Simulation

The University of Rochester reported on two ANS software packages. One is intended to be executed in a BBN Butterfly Multi-Processor [FANTY] and another can be executed on either a VAX minicomputer (with UNIX) or the Sun Microcomputer.

3.3 ANS's and Components Being Developed but Not Currently Available

Several high technology companies and laboratories are sponsoring internal research and development programs aimed at producing commercial ANS's or ANS components. The following subsections given an overview of such activities for some of the companies initiating product oriented research programs.

3.3.1 AT&T Bell Laboratories

Bell labs has a neural network working group with H. Graf, L. Jackel, J. Denker, et al. as members. This group has successfully implemented 54 neurons and 54 input channels with 3000 synapses (interconnects) connecting the input channel with each neuron. Standard CMOS was used in the implementation. Design work for a 256 neuron-chip and 512 neuron-chip has been completed. In addition, an associative memory with an analog processor and digital I/O was reported. Details of the chip designs are given in [GRAF, JACKEL, et al.].

Joshua Alspector and Robert Allen of Bell Communications Research described a VLSI implementation of a modified Boltzmann Machine in [ALSPECTOR]. Their paper contains a short review of Pitts formed neuron, ADALINE, Hopfield's model, the Boltzmann Machine and back-propagation.

3.3.2 IBM

Under Dr. Claude Cruz's leadership, IBM's Palo Alto Research facility has developed a "Network Emulation Processor (NEP) with an IBM PC/XT host computer and professional graphics adapter yields a workstation to interactively design, debug, and analyze networks.

3.3.3 California Institute of Technology and the University of Pennsylvania

P. Mueller of the University of Pennsylvania and J. Lazzaro of the California Institute of Technology have assembled an ANS of 400 analog neurons for analyzing and recognizing acoustical patterns (including speech) [MUELLER]. Up to 100,000 interconnects can be made and synaptic gains and time constants are determined by plugging in

resistors and capacitors. The system currently performs adequately for well articulated phonemes and diphones. New energy consonants present problems but it is expected that different coding schemes and better understanding of the invariant clues for speech perception will lead to improvements.

3.3.4 Hughes Research Laboratories (Malibu, CA)

Researchers G. Dunning, E. Maron, Y. Owenchko, and B. Soffer at Hughes Research Laboratories have developed and tested an all optical nonlinear associative memory using a hologram in an optical cavity formed by phase conjugate mirrors [SOFFER]. They were able to store multiple superimposed images and reconstruct a complete image, inputting only a portion of the stored image.

3.3.5 UCSD Analog Back Propagation System

Stan Tomlinson, a graduate student at UCSD, discovered a method for increasing the speed of a back propagation ANS. Tomlinson defined a "continuous time back propagation ANS" where the forward pass, backward pass, and weight modifications are performed simultaneously.

3.3.6 Oregon Graduate Research Center

Dan Hammerstrom and his colleagues at the Oregon Graduate Research Center are in the process of designing a water-scale integrated silicon system that implements a variety of ANS's [HAMMERSTROM]. In their investigations of implementations of the back propagation learning rule, their simulations show that nodes can compute asynchronously and that the rule is robust enough to accommodate incomplete information on each learning cycle.

3.4 ANS Properties

K.L. Babcock and R.M. Westervelt of Harvard University investigated the stability and dynamics of electronic neural networks with added inertia and reported their results in [BABCOCK]. Babcock and Westervelt reviewed the Hopfield ANS model and then expanded it by adding an inertial term to the rate equations.

A. Guez, V. Protopopescu, and J. Barnden of the Oak Ridge National Laboratory Studies the stability, storage capacity and design of nonlinear results in [GUEZ]. They determined sufficient conditions for the existence and asymptotic stability of any ANS's equilibrium. The conditions take the form of a set of piecewise linear inequality constraints solvable by a feedforward binary network or other methods such as Fourier Elimination.

4.0 ANS MODELS AND APPLICATIONS

Despite essentially continuous research and development since the introduction of computers in the late 1940-s, the connection between brain processing and computer processing is still undergoing theoretical development. No attempt is made at forming a complete integrated view of the large amount of material of the different ANS topics covered in the papers and reports from the survey participants. Time restrictions of the study prevent undertaking the very important tasks of interpreting, evaluating, and integrating the many excellent and undoubtedly important contributions by ANS researchers. These crucial tasks remain to be done in the future.

It is difficult to make a clear distinction between brain theory and ANS (neural network or connectionist) theory since ANS researchers typically proceed by abstracting the essence of a theory aimed at explaining results obtained from neurophysiological or psycho-physical experiments and then derive a method for designing processors (electronic or optical) which can obtain the same or nearly the same experimental results. [FRISBY] contains an excellent and well illustrated discussion of Poggio's and Marr's approaches at modeling stereopsis, and his book serves to indicate the basic practice of going from pure brain theoretic modeling to computer based ANS models. It is generally accepted that no completely satisfactory brain theory exists although some models, such as von der Malsburg's model of the visual cortex, exhibit behavior that experiments confirm. In spite of sometimes incomplete theories, ANS researchers often attempt to develop real world applications of any plausible theories in areas such as vision, speech recognition, etc. Many of the current ANS theories basing their roots in brain research are rich enough in information theoretic content to have, so to speak, lives of their own. The current successful ANS model seems to be one reflecting human-like information processing capabilities, implementable in some computational device which provides reasonable results in close to real-time. The present conventional AI impasse reached by attempts at computer vision and artificial intelligence so ably described in [LERNER] and [DREYFUS] leaves no other alternative than the ANS based approaches presented in the papers described below. [von der Malsburg] Contains many modern brain models.

4.1 Brain Theory Applicable to ANS's

The question of how much brain theory is enough to enable the design of processors which can produce acceptable human (or animal)-like processing is very difficult to answer. It is also very difficult to ascertain the level of detail required of a model: is it sufficient to characterize the average behavior of assemblies of neurons as in much of Grossberg's work or are individual cells behavior required? How to choose between and extract the "essential" properties of Grossberg's, Freeman's, Edelman's and Reeke's models, for example is not at all clear. Recent ANS/computer theory history compels modern computer scientists and robotic engineers to achieve at least some familiarity with the different brain theories represented in this subsection.

4.1.1 Nonlinear Dynamics with Chaotic Solutions (Laboratory Measurements and Models)

Walter Freeman, Christine Skarda, and Bill Baird developed models of the formulation and recognition of patterns in the rabbits olfactory bulb. Baird gave an excellent overview of neural modeling which relates the well known ANS models to results from laboratory measurements [BAIRD]. Baird simplified Freeman's ANS model of the rabbit's olfactory bulb while capturing the essence of the pattern formation/recognition behavior. A key definition for Baird's work is "pattern formation": the emergence of macroscopic order from microscopic disorder. Freeman, Skarda, and Baird define dynamical systems which have chaotic behavior (fractal solutions) for ground states as opposed to fixed point attractors, as in Simulated annealing or the Hopfield model. It is speculated that such ground state behavior is essential for real-time continuous perception.

Freeman argues, very congenitally, that neural dynamic system destabilization provides the best description of the essentials of neural functioning, and Baird finds that the mechanism of competing instabilities (nonlinear mode selection) is implicit in dynamical associative memories and provides the key ingredient in pattern recognition. While admitting that their neural dynamical models have many similarities with well known connectionist models, they point out significant dissimilarities essential for recognition and discrimination. Freeman and Baird's models are unique in that they possess dense local feedback between neuron assemblies. Such feedback is necessary to generate chaotic and limit cycle system ground states.

Gail Carpenter and Stephen Grossberg developed a neural network model for mammalian circadian rhythms [CARPENTER] which can have chaotic solutions for some system parameter ranges. Their model explains many phenomena in mammal behavior such as the role of eye closure during sleep and the stability of the circadian period.

4.1.2 Brain Models of Knowledge, Conditioning, Perception, and Learning Processing

Stephen Grossberg and Ennio Mingolla in their paper [GROSSBERG, MINGOLLA] show how computer simulations can be used to guide the development of neural models of visual perception. They use cooperation/competition mathematical models to simulate textual segmentation and perceptual grouping as well as boundary completion.

Daniel Levine, the brain theorist who authored the very extensive and well written neural modeling review [LEVINE] also recently contributed two papers to brain theory. In "A Neural Network Model of Temporal Order Effects in Classical Conditioning", Levine demonstrated in a computer simulation that one of Grossberg's neural networks can reproduce the experimental findings that the strength of a conditioned response is an inverted U-function of the time interval between conditional and unconditioned stimuli. The network also can reproduce blocking of a neutral stimulus by another stimulus that has been previously conditioned. Levine also traces the history of conditioning models in this paper and reviews Grossberg's theories [LEVINE, 1986].

A very novel model of cortical organization was invented and investigated by Gordon Shaw and colleagues [SHAW] at the University of California at Irvine. Their model was motivated by V.B. Mountcastle's organizational principle for neocortical function and M.E. Fisher's model of spinglass systems. Their network is composed of interconnected "trions", units which have three possible states (-1, 0, +1) which represent firing below background, at background, and above background respectively. Trions represent a localized group of neurons and symmetrical interaction between trions exhibit behavior where hundreds of thousands of quasi-stable periodic firing patterns exist and any can be selected out and enhanced, with only small changes in interaction strengths, by using a Hebbian-type of algorithm.

Dana Ballard presented his local representation ideas for modeling the cerebral cortex in a stimulating paper in the Behavioral and Brain Sciences [BALLARD]. A large part of the interest associated with the paper resulted from the peer commentary following the paper (members in the Veer group included Grossberg, Hopfield, Edelman). In the paper Ballard presented his local representation model and value unit concept. He included methods for perceiving shape and motion by exploiting his model.

John Hopfield, whose famous ANS for solving the traveling salesman problem, generated a great amount of interest and enthusiasm for, once again, applying ANS's to real world problems, also performed theoretical and experimental work on the nervous system of the Limax maximum (A terrestrial mollusk). In [HOPFIELD] a description is given of behavioral and neurophysiological attributes of Limax learning. A model which also includes a memory network is related to experimental data and predictions are made.

George Hoffmann presented some very novel ANS concepts in a paper which explored analysis between the brain and the immune system [HOFFMANN]. His model involved a neuron with hysteresis which eliminated the need for learning with modifiable synapses. The Hoffmann ANS learns through interacting with the environment and being driven to regions in phase-space. The system has 2N attractors for N neurons.

In their paper "Selective Neural Networks and Their Implications for Recognition Automation" [REEKE], George Reeke and Gerald Edelman observed that the models of

McCulloch and Pitts, Marr, Hopfield, Hinton and Anderson, and McClelland and Rumelhart all deal with the mechanisms for acquiring and processing information, but not with the ways that the categories of information and the mechanisms to process them come to exist. As a consequence, Reeke and Edelman devised a model which exploits basic biological principles (Darwin's principles of natural selection) to explain the discovery of perceptual categories, the representation of categories without pre-arranged codes, the manipulation of retrieval keys, and the selection of actions on the basis of imperfect and inconsistent information without a program.

A. Pellionisz and **R. Linas** wrote many papers on the last ten years where "Tensor Network Theory" is introduced [PELLIONISZ]. Pellionisz asserts that the conventional representations of McCulloch, Pitts, Kohonen, Hinton and Anderson are expressed in extrinsic, orthogonal systems of coordinates and represented coordinates in Euclidean vector space. He finds that new important insights into the control nervous system are gained by describing the central nervous system in terms of intrinsic coordinates using tensor analysis.

Michael Jordan continued the tradition of excellence in ANS research associated with the UCSD Institute for Cognitive Sciences with the publication of his paper "Serial Order: A parallel Distributed Processing Approach" [JORDAN]. ANS trajectories (attractors) follow desired paths as a result of learning with constraints which generate the required serial order. His report contains an ANS tutorial section, an approach to the coarticulation problem in speech, and examples from various simulations.

David Zipser, also from UCSD's Institute for Cognitive Science, authored two reports dealing with problems of the representation of spatial entities [ZIPSER]. He described a map retrieval mechanism based on an ANS and illustrated mapping tasks such as recognition of previously visited locations, path finding using landmarks, and finding a path between two locations that do not share landmarks. He also points out similarities between ANS features associated with map representation, and known features of the hippocampus.

James McClelland and **David Rumelhart** presented a very thorough exposition of the distributed models of memory and learning and compared their model to other well known models from cognitive science [MCCLELLAND]. They point out that their distributed model is capable of storing many different patterns, determining the central tendency of a number of different patterns, create perceptual categories without using labels, and capturing the structure inherent in a set of patterns with or without prototype characterization.

Jerome Feldman of the University of Rochester performed a theoretical analysis of the behavior of connectionist mode, in [FELDMAN]. He analyzed ANS's using energy concepts such as the Hopfield, Hinton, and Sejnowski, and Smolensky models and pointed out that such approaches are most relevant for problem domains lacking significant structure and questioned the utility of such approaches in highly structured cognitive domains (such as compiling or parsing). He also discussed the relevance of automation theory and control theory to ANS formalization.

In a paper entitled "On Applying Associative Networks", submitted to the IEEE First Annual Conference on Neural Networks, **A.D. Fisher** treated approaches to formulating basic organizational principles for mapping problems onto associative processors. He addressed goal directed learning and structures for knowledge representation, and configuring a simulation environment for evaluating and developing the organizational principles.

4.2 Computation/Optimization

The computational generality of ANS's and the ability of ANS's to solve some interesting optimization problems are now widely known [DENKER]. In the last few years some interesting papers appeared which view ANS's from complexity theory, computational, and mathematical perspectives and which strive to characterize ANS's in more classical systems theoretic manner. This subsection describes these types of papers from various contributors to the survey.

Ian Parberry and **Gero Schnitger**, of the University of Pennsylvania, investigated the relationship between Boltzmann Machines and conventional computers [PARBERRY]. They concentrated on determining the computing power of Boltzmann machines which are resource bounded. They measured machine running time and hardware requirements as functions of problem size. They found that:

- 1) The connection graph can be made to be acyclic (no feedback loops)
- 2) Random behavior can be removed from the machine
- 3) All synapse weights can be made equal to one.

These properties make the machines equivalent to a combinatorial circuit and the resulting machine will have its running time increased by a constant multiple and its hardware requirement increased by a polynomial.

Terence Smith, **Omar Egecioglu**, and **John Moody** analyzed computational complexity issues in ANS's such as the generalized feed-forward networks, perceptrons, and Hopfield devices [EGECIOGLU]. For each ANS type they describe programming the ANS, functions computable by the ANS, complexity issues, and the architecture. They point out the need for the expansion of traditional computer science to include dynamical systems and statistical mechanics.

Pierre Baldi of UCSD (formerly of the California Institute of Technology) established an upper bound for generalized Hopfield type models. That is, models with energy functions or Hamiltonians of degree d :

$$H(x) = \sum_{i_1 \dots i_d} T_{i_1 \dots i_d} x_{i_1} \dots x_{i_d}.$$

Based on counting arguments they established a storage capacity upper bound of $O(N^{d-1})$ for ANS's with dynamics which minimize $H(x)$. Baldi points out that these higher order systems have local updating rules as in the quadratic (Hopfield) case and they recur very naturally in optimizing problems [BALDI].

John Hopfield, of the California Institute of Technology, and **David Tank**, of AT&T Bell Laboratories, introduced a new conceptual framework and a minimization principle which provide increased understanding of computation in neural circuits [HOPFIELD]. They derive a model abstracted from knowledge of biological neurons and discuss how their model dispenses with many known properties of neurons while still capturing those aspects necessary for performing computations essential to organism adaptations and survival. The classical model of the neural dynamics is taken as a point of departure and Hopfield and Tank show that the assumption of interconnect (weight) symmetry is not overly restrictive and that many functions such as edge detection, stereoptics, and motion detection can be cast as optimization problems and solutions result from the convergence of symmetric dynamic neural systems.

Morris Hirsch, a mathematician from the University of California, investigated convergence in neural nets [HIRSCH]. He discussed the Liapunov functions discovered by Cohen and Grossberg and the connection to Hopfield's results. He was able to remove some restrictions on the state equations (neuronic equations) that show that important convergence properties still hold.

Harold Szu, of the Naval Research Laboratory, presented a method for speeding up the conventional simulating annealing algorithm [SZU]. Szu's non-convex optimization method, called a Cauchy machine, is derived from the Boltzmann machine by using the Cauchy/Lorentzian probability density function in place of the Gaussian probability density function. It is shown that the "cooling" schedule varies inversely with the time versus the inverse logarithm of time for conventional simulated annealing.

4.3 Memory in ANS's

A large body of ANS research is devoted solely to the modeling of human or biological memory (see [LEVINE] for a review). This section presents descriptions of recent work on ANS memory and is broken into theory, applications, and implementation categories.

4.3.1 ANS Memory Theory

Despite the extensive amount of research, there is still no complete and universally accepted theory of human or biological memory. However, as in other categories of ANS modeling, many useful models of memory have been identified in the pursuit of devising an accurate theory of biological memory. The work on the theory of memory immediately below is representative of modern theoretical advances in ANS memory theory.

David Rumelhart and **Donald Norman** of the UCSD Institute for Cognitive Science are important contributors to the development of parallel distributed ANS's and they also did important work in Artificial Intelligence. They published an ICS report on memory, "Representation in Memory," which gives an excellent overview on AI approaches in the theory of human memory. Though ANS's are not addressed in this report, it is, nevertheless, recommended for gaining insights into the issues concerning the representation of knowledge. They discuss spreading activation in semantic networks and discuss the ideas of Fahlman and Anderson.

Stephen Grossberg and **Gregory Stone** devised models of the effects of attention switching and temporal-order information in short term memory. In their paper, "Neural Dynamics of Attention Switching and Temporal-Order Information in Short-Term Memory" [GROSSBERG, STONE], they argue that attention-switching influences initial

storage of items in short term memory, but competitive interactions among representations in short term memory control the subsequent dynamics of temporal-order information as new items are processed.

Teuvo Kohonen summarized his current views on the theory of memory in a paper titled, "Self Organization, Memorization, and Associative Recall of Sensory Information by Brain-like Adaptive Networks" [KOHONEN]. Kohonen asserts that the two main functions of memory are: 1) to act as mechanisms which collect sensory information and transform it into various internal models or representations, and 2) to interrelate the signal processes in these representations and store them as collective state changes of the neural network.

In 1984, **Pentti Kanerva** published his PhD thesis entitled "Self-Propagating Search: A Unified Theory of Memory" [KANERVA]. His dissertation introduces the sparsely distributed memory model and the concept of using a neuron as an address decoder for accessing memory. Kanerva's memory model overcomes limitations in the Hopfield memory model such as dependence of storage capacity on the number of neurons ($1.14N$, $N = \#$ of neurons), the inability to store temporal sequences, symmetric interconnects, and a new limited ability to store correlated inputs.

The dissertation includes background material on related ideas by Marr and Kohonen and includes mathematical estimates of convergence rates and memory capacity. Kanerva further elaborated upon his memory model in a paper, "Parallel Structures in Human and Computer Memory," and discussed the application of his ideas to the "frame" problem of Artificial Intelligence and showed that the part of the problem concerned with manipulating vast quantities of data about the real world can be handled with his sparse distributed memory concepts.

James Keeler of UCSD compared Kanerva's model with Hopfield's model in a Research Institute for Cognitive Science Paper: "Comparison between Sparsely Distributed Memory and Hopfield-Type Neural Network Models", [KEELER]. In this very well written and mathematically rigorous paper, Keeler developed a mathematical framework for comparing the two patterns. Keeler extended Kanerva's sparse distributed memory model and showed that Hopfield's model was a special case of this extension. Keeler showed that Kanerva's model corresponds to a three layer network with the middle layer consisting of many more neurons and that Kanerva's formulation allows context to aid in the retrieval of stored information.

Alan Lapedes and **Robert Farber**, of the Theoretical Division of Los Alamos National Laboratory, reported on a new method for designing a content addressable memory which is free of major limitations associated with the Hopfield content addressable memory [LAPEDES]. Their method consists of dividing a network into two groups of neurons. One group, called the Master net function, basically has a Hopfield optimization network while the other group, called the Slave net can have asymmetric connections. Advantages associated with this master/slave formulation are: 1) two bases of attraction may be merged together, 2) weighting of certain components of a fixed point so that it attracts more strongly (sculpting a basis of attraction), and 3) biologically plausible division of neurons into excitatory and inhibitory sub-groups.

Tarig Samad and **Paul Harper** of Honeywell used back-propagation in linear array of fully connected units to construct a content addressable memory [SAMAD]. They cite several advantages in their method over the Hopfield content addressable memory: 1) asymmetric weights, 2) their method is guaranteed to recognize stored patterns, 3) close to perfect recall if a retrieval cue is not very far from any stored memory, 4) up to 2^N patterns can be stored ($N =$ number of neurons and bits in the patterns), 5) 20% perturbations in learned weights and thresholds effected performance by less than 1%, and 6) robustness in degradation of weights, learning rates, and stored pattern cue hamming distance.

Demetri Psaltis and **Cheol Hoon Park** of the California Institute of Technology, designed an associative memory with a quadratic discriminant function [PSALTIS]. Their neural net memory can be shown to have a capacity proportional to N^2 where N is the number of bits in a storage sector. The square law nonlinearity is conducive to an optical implementation. The added capacity can be combined with the shift invariant property of an optical correlator to yield a shift invariant associative memory.

Santosh Venkatesh and **Demetri Psaltis** of the California Institute of Technology, discovered an associative memory which uses the spectrum of a linear operator [VENKATOSH]. They showed that their method has a capacity which is linear in the dimension of the state space while that of the outer-product method has a capacity asymptote of $n/(4 \log n)$. Their method requires full connectivity and, consequently, is more suitable for an optical implementation. The larger storage capacity of the "spectral" method is paid for with increased pre-processing cost.

4.3.2 Applications of ANS Memory

An interesting application of an ANS memory was devised by Michael Mozer of the Institute of Cognitive Science at UCSD [MOZER]. Mozer's ANS performed inductive information retrieval. The ANS retrieval system takes dynamic use of the internal structure of test databases to infer relationships among items in the database.

The inferred relationship helped the system overcome incompleteness and imprecision in request for information as well as in the database. The ANS used neuron-like equations from McClelland and Rumelhart's interactive activation model of word perception. The model handles queries in a document retrieval application by activating a set of descriptor units and seeing which document units become active as a result.

4.3.3 Implementations

Shift Invariant Optical Associative Memory implementations were analyzed by D. Psaltis, J. Hong, and S. Venkatesh in [PSALTIS]. They found that without special encoding techniques associative memories with linear interconnections did not retrieve shifted images well. Two systems, one with a square law interconnection and the other with a novel encoding scheme, were found to be shift invariant and to achieve the performance of the outer product method.

Arthur Fisher, **Robert Fukuda**, and **John Lee** discussed the implementation of parallel processing architectures consisting of multiple optical adaptive associative modules [FISHER]. The modules adaptively learn and store a series of associations in the form of electronic charge distributions in an optical control device termed a micro-channel spatial light modulator. The optical adaptive associative modules have a gated learning capability, where adaptivity is easily switched on or off. The associative modules have an accumulative learning capability where even one exposure to an associated pair of vectors produces a weak association and subsequent exposures improves to the optimum pseudo-inverse solution.

4.4 Learning

It is well known that progress in Artificial Neural Systems came to an abrupt halt shortly after the publication of the Minsky and Papert treatise; "Perceptrons." The ability to program a multilayered network proved to be a very stubborn problem until the work of Kohonen, Rumelhart, Parker, Hopfield, Barto, and Sutton showed how such nets could be programmed. This section describes the recent results in ANS learning.

In this monumental PhD dissertation, **David Ackley** describes a multi-dimensional space search strategy which combines hill climbing methods and search methods based on genetic algorithms [ACKLEY]. Ackley's method called "Stochastic Iterated Genetic Hillclimbing", has a coarse-to-fine search strategy as in the case for simulated annealing and genetic algorithms but the convergence process is reversible. That is, in the implementation it is possible to diverge the search after it has converged and recover coarse-grained information about the space that was suppressed during convergence. Successful optimization typically has a series of converge/diverge cycles.

Ronald Williams, formerly of the Institute for Cognitive Science at UCSD and currently with Northeastern University, investigated a class of algorithms designed to allow the self-organization of feature mappings in ANS's [WILLIAMS]. Williams introduces a measure termed "faithfulness" which is intended to measure how well an input pattern can be constructed from knowledge of an output pattern. The algorithm he devised for feature detection maximizes the faithfulness measure. Williams defines an algorithm called "Symmetric-error correction" and he proves that if a system is completely linear and symmetric of rank R greater than m , the number of output units, then when the algorithm converges the resulting weight vectors are of unit length, orthogonal, and span the space spanned by n eigenvectors having the largest eigenvalues.

Andrew Barto of the University of Massachusetts, discusses an ANS based on the concepts introduced by **Harry Klopf** [BARTO]. Barto's ANS is called A_{R-p} (associative reward-penalty) and has a learning rule which adjusts weights according to four types of information: 1) presynaptic signals, 2) postsynaptic signals, 3) a reinforcement signal from the environment that reflects the consequences of a neuron's activity, and 4) a signal that indicates what a neuron usually does for a given stimulus pattern.

Barto shows how the A_{R-p} ANS solves the "exclusive or" problem as well as other non-trivial problems. He also includes a very informative comparison of "associative reinforcement learning" (to which A_{R-p} corresponds), supervised and unsupervised learning and points out subtle distinctions between these basic learning types. He argues, for example, that unsupervised learning is more accurately regarded as supervised learning with a fixed built in teacher. Theoretical convergence and comparative analyses of A_{R-p} are also included in the report.

Gail Carpenter and **Stephen Grossberg** collaborated on paper on associative learning, adaptive pattern recognition, and cooperative-competitive decision making by neural networks [CARPENTER]. The paper contains a discussion of the "universal theorem" (proved elsewhere) which show how arbitrarily many cells computing arbitrary transfer functions can interact asynchronously through complex nonlinear feedback and experience no learning bias resulting from cross-talk of their feedback signals — a result called "absolute stability". There are also discussions of feature discovery, category learning, adaptive pattern recognition, and Adaptive Resonance Theory (ART). ART is an ANS which self-organizes its recognition code and the environment can also modulate the learning process and, as a result, carry out a teaching role.

Harry Klopf discussed drive reinforcement learning in a paper for the IEEE First Annual International Conference in Neural Networks (June 1987) [KLOPF]. Klopf's learning model uses signal levels and changes in signal levels in such a way to yield an unsupervised learning which predicts classical conditioning phenomena such as delay and track conditioning, stimulus duration and amplitude effects, second-order conditioning, extinction as well as other phenomena. In his model, sequentiality replaces simultaneity and is an extension of the Sutton-Barto model (1981).

D.G. Bounds, of the Royal Signals and Radar Establishment, performed an analysis of Boltzmann Machines and reported the results in [BOUNDS]. Bound's paper contains a discussion of the Boltzmann machine algorithm and the encoder problem for the task of communicating information between components of a parallel network. An extensive simulation was conducted which assessed the effect of temperature on learning rate, and a comparison between the Boltzmann Machine Hamiltonian and the Sherrington-Kirkpatrick Spin-glass Hamiltonian.

W.S. Stornetta and B.A. Huberman of the Xerox Palo Alto Research Center presented an analysis of the back-propagation algorithm in [STORNETTA, ICCN,87]. They capitalized on the fact that if an input is zero there will be no modification to the weights extending out from that unit and, consequently, only half of the weights from the input to the hidden unit layer will be changed. The back-propagation algorithm was modified so that the dynamic range of all units was $(-1/2, 1/2)$ rather than $(0,1)$. The input and output patterns consists of series of $-1/2$'s and $1/2$'s and the squashing function is given by:

$$-1/2 (\exp(-w_{ij}0_j + \text{bias}_i) + 1)^{-1}$$

and the rest of the conventional back-propagation algorithm remains unchanged.

Bart Kosko, of VERAC Corporation, devised a new learning method associated with Bidirectional Associated Memory (BAM) and described its ANS properties in [KOSKO]. This very well written and informative paper contains a review of classical associative memory theory and a tutorial section on BAM. An earlier proof by Kosko that energy matrix is bivalently bidirectionally stable is reviewed and BAM correlation encoding is discussed.

Continuous BAM's, introduced in Kosko's earlier work, are reviewed and a proof that every matrix is continuously bidirectionally stable is given. BAM learning, the main interest of the paper, is achieved by programming the BAM connection matrix to adapt according to a generalized Hebbian learning law where adaptive resonance occurs, that is, neurons and interconnections quickly reach equilibrium. The connection weight learning law is called the Signal Hebb law and is given by:

$$m_{ij} = m_{ij} + (a_i)(b_j)$$

$$A_i, B = \text{activation}$$

$$S(\) = \text{sigmoid function}$$

Adaptive BAM's are characterized as to their classification properties and it is shown that they converge to local energy minima. In addition, the capabilities of the adaptive BAM are compared and contrasted to Grossberg's Adaptive Resonance Theory (ART) ANS.

Kosko extended adaptive BAM's to competitive adaptive BAM's by including lateral-inhibitory interactions [KOSKO]. When field inhibitory connections are taken to be symmetric, stability can be shown, although it was found that in practice, non-symmetric within field connections exhibit stability in many cases.

Variations on the Grossberg adaptive resonance model were explored by **R.W. Ryan and C.L. Winter**, of Science Applications International Corporation and reported in the Proceedings of the ICNN (1987). Ryan and Winter found that the adaptive resonance model may activate a coded recognition node whose top-down pattern bears little resemblance to the corresponding input pattern and there is no guarantee that the adaptive resonance circuit (ARC) will reset and, hence, the ARC will be recoded by the input pattern unless the set of weights is prevented from changing.

4.5 Pattern Recognition

Bill Baird of the Department of Biophysics at the University of California, reported the results of an analyses of pattern recognition in the olfactory bulb of the rabbit [BAIRD]. Baird simplified Freeman's model by neglecting synaptic delay and did not perform a full simulation of the rabbits olfactory system but achieved universality as a result. His mathematical analysis explains how an oscillating system can pattern recognize. Baird conducted experiments with an array of 64 electrodes which yielded EEG patterns which showed the emergence of order from disorder and indicated that specific EEG patterns are correlated with specific recognition responses. A theory which combines the mathematical description of the emergence of order by instability with the mathematics of associative memory is required to model learning and memory in neural networks.

Robert Hecht-Nielsen, of HNC, showed how banks of matched filters can be used as pattern classifiers for complex spatiotemporal pattern environments such as speech, sonar, radar, and communication [HECHT-NIELSEN]. He defined an ANS, called a "simple avalanche matched filter bank" which closely approximates the theoretical classifier to be introduced. He defined a "nearest" match filter, discussed its error rate, and observed that such classifiers can only carry out the first "local in time" stage of pattern recognition and that context must be exploited to achieve high performance pattern recognition.

Paul Gorman, of Bendix Aerospace and **Terrence Sejnowski**, of John Hopkins University, collaborated on a paper which discussed the application of back-propagation to classify sonar targets [GORMAN]. The signal representation used for input to the network was selected as the result of experiments with human listeners. A short term Fourier transform using 60 frequency samples per temporal slice was generated for each signal. By starting at the onset of the signal and increasing the position of the time slice, approximately a linear FM/chirp, essentially formed the diagonal of a 60×60 time/frequency matrix. Normalized values from the matrix served as input to the matrix. Results from experiments indicated that the network is able to discover target classification features from examples of sonar signals with performance comparable to human experts.

Results on applying back-propagation to handwritten numerical recognition and spoken numeral recognition were reported by **D.J. Burr** of Bell Communications Research [BURR]. Burr gave tutorial explanations of related geometric hyperplane analysis and back-propagation learning in addition to thorough treatments of recognition of handwritten and spoken numerals. For the handwritten numerals a two-stage process of normalization followed by feature extraction was used. It was found necessary to subtract a constant representing signal level from all feature vectors. Removing the D.C. component in this was dramatically increased the learning rate of the network. The neural networks were configured with up to 64 hidden units, but it was found that a maximum recognition score around 98% occurred with 6 and 14 hidden units for the written and spoken recognition tasks respectively. Burr's results compared favorably to nearest neighbor pattern recognition applied to the same problem.

Alan Kawamoto and James Anderson extended Anderson's Brain State-in-a-Box (BSB) model to define a new ANS which models multistable perception [KAWAMOTO]. The BSB extension allows the ANS to shift between hyperspace box corners in a way corresponding to multistable perception. After explaining the analytical properties of their new ANS, Kawamoto and Anderson showed how their ANS qualitatively agrees with published psycho-physical results on multi-stable perception regarding bias, adaptation, hysteresis, and dynamics. Most of the results discussed concern visual ambiguities but speculations concerning lexical ambiguity resolution are presented also.

4.6 ANS's Applied to Vision

Maureen Gremillion, Arnold Mandell and Bryan Travis reported on their design and implementation of a neural net model of the mammalian visual system in [TRAVIS]. These researchers from Los Alamos National Laboratory modeled a scaled down version of primary visual cortex, the lateral geniculate nucleus, and a 45,000 neuron retina. Research types are embedded in two-dimensional layers and differentiated cell types are distributed in space. Several different retina models from the visual modeling literature were tested.

Kunihiko Fukushima, of NHK Science and Technical Research Laboratories in Japan, discussed a multilayered hierarchical ANS called the "Neocognitron", in [FUKUSHIMA]. The Neocognitron consists of alternating layers of feature-extracting neurons and neurons which are fed outputs from feature-extracting neurons and fire only if one input is active. Inhibitory cells exist to suppress irrelevant features. The Neocognition is capable of supervised or unsupervised learning and can recognize shifted or deformed variations of a pattern. Excessive distortion results in windowed recognition response. A Neocognition with selective attention is currently being investigated by Fukushima.

Ronald Williams, of Northeastern University, investigated the ability of self-organizing networks to infer spatial relations [WILLIAMS]. Williams was motivated by the dipole pattern neural network experiments conducted by Rumelhart and Zipser where, without a priori knowledge of the spatial layout of the input, an ANS learned aspects of input spatial structure. He mathematically formalized the notion that if two patterns are strongly correlated then they must be nearer in some metric than if their values are weakly correlated. The concept of a spatial relationship or pattern element is formalized in terms of a distance metric on pairs of elements and the notion of an environment with spatial structure is formalized as a random field over a metric space.

Richard Golden, of Brown University, developed an ANS which models the process of visual perception of a letter in the context of a word. Interconnections between neurons represent any spatially or sequentially redundant and transgraphic information in displays of letter strings. Golden's model uses Anderson's BSB model and enhancements derived from commonly accepted principles of information processing in the central neurons system. In Golden's model, a word is represented as a pattern of neural activity over a set of position-specific feature neurons.

Image restoration involves removing degradations from images arising from blur from optical aberrations, atmospheric turbulence, motion, defraction, and noise. The application of ANS methods to image restoration is currently being investigated by Y.T. Zhou, R. Chellapa and B.K. Jenkins of the University of Southern California [ZHOU]. They designed an ANS containing redundant neurons to restore gray level images degraded by shift invariant blur function and noise.

Ralph Linsker, of the IBM Thomas J. Watson Research Center, addressed the origin and organization of spatial-opponent and orientation-selective neurons in ANS's based on biologically plausible roles for development [LINSKER]. Linsker treats the emerging of network structures from spontaneous electrical activity and simple biologically based rules for synaptic modification.

Ralph Siegel, currently of Rockefeller University and formerly of the Salk Institute, analyzed the abilities of Rhesus monkeys and human subjects to detect the change in three-dimensional structure of cylinder using only motion cues [SIEGEL]. Siegel implemented a three layer ANS with 100 neurons in the input layer homologous to the neurons in the middle temporal area of the brain. Each input neuron was tuned for a given velocity at a retinotopic location and corrected to all ten units in the middle layer. The middle ten units were connected to only one neuron in the output layer which indicated structure as constructural motion.

J.M. Oyster, W. Broadwell and F. Vicuna of the IBM Los Angeles Scientific Center, investigated the application of associative networks to robot vision in [OYSTER]. The approach taken to robotic vision by these IBM researchers was to show how the well known methods for robotic vision, such as image acquisition, segmentation, object recognition, etc., can be implemented in an associative network. They rigorously convert a discrete convolution to an associative network and then argue that the standard low-level edge-detection operators, such as the Roberts, Sobel, and Laplacian and Gaussian operators can be implemented as ANS's.

Christof Koch, Jose Marroquin and Alan Yuille demonstrated how Hopfield ANS models can be generalized to solve nonconvex energy functionals corresponding to functions of early vision such as computing depth from two stereoscopic images, reconstructing and smoothing images from sparsely reconstruction can be formulated in terms of minimizing a quadratic energy function and, subsequently, show how quadratic variational principles fail to detect image discontinuities. An energy function containing cubic terms is defined which is shown to handle discontinuities.

G. Cottrell, P. Munro and D. Zipser applied back-propagation to compressing images [COTTRELL]. Current (1987) image compression techniques are briefly treated as in back-propagation. Image compression is considered to be a type of encoder problem in that an identity mapping over some set of inputs must be performed. An ANS is forced to perform the mapping of a narrow channel of the network and thereby causes an efficient encoding. Two noteworthy facts are: 1) the network developed a compact representation of its environment, 2) although the algorithms were developed as supervised learning schemes, the problem really involves learning without a teacher since the input and output are identical, i.e., the ANS self-organizes to encode the environment. Most of their image compression results were obtained with a three layer network: G4 inputs, 16 hidden units and 64 outputs.

Stephen Grossberg developed a theory of vision which offers an explanation of the coherent synthesis of three-dimensional form, color, and brightness percepts [GROSSBERG]. The theory identifies several uncertainty principles that limit the extraction of visual information. Particular modules are resolved by hierarchical parallel interactions between many processing stages. Grossberg asserts that when a neural processing stage removes one type of uncertainty from any input pattern it often generates a new type of uncertainty which is passed to the next processing stage. That is, information is not progressively reduced in a succession of neural processing stages. Based on results of monocular boundary segmentation and feature filling-in and the interaction between these processes, Grossberg suggests that the commonly accepted hypothesis of independent modules in visual perception is wrong and misleading.

Harold Szu and Richard Messner derived multiple-channel novelty filters of associative memory from a retina ANS point-spread function [SZU]. They present a novelty filter as a remainder operator and mathematically drive the multiple-channel model from associative memory formulation. Their ANS is shown to be scale and rotation invariant and simulation results are cited. They also point out a new relationship between adaptive novelty filtering and adaptive associative memory.

Michael Mozer of the Institute for Cognitive Science at UCSD, investigated early parallel processing in reading [MOZER] and developed an ANS capable of recognizing multiple words appearing simultaneously on an artificial retina. The ANS is called "BLIRNET" since it builds location-independent representations of multiple words. BLIRNET is a multilayered hierarchical network which learns via back-propagation.

4.7 ANS's for Speech and Language

Bryan Travis, of the Los Alamos National Laboratory, described a layered ANS model of sensory cortex in [TRAVIS]. The goal of Travis' research program was to construct a model of the human sensory system which reflects what is known structurally and physiologically at several levels from the ear to the midbrain nuclei to the cortex. Travis made simplifications in scale to accommodate current computer technology constraints but claimed that his model contained the following desirable features: 1) more structure (based on neurophysiological data) than previous models, 2) inputs based on biological data, 3) emphasizes dynamics, and 4) provides a means of testing theories about sensory perception.

D.W. Tank and J. Hopfield developed an analog ANS capable of solving a general pattern recognition problem for time-dependent input signals [TANK]. In order to solve such sequence recognition problems, Tank and Hopfield expanded the Hopfield ANS's Energy function to include time dependence. In the case of a time-dependent energy function, a convergent computation can occur if the problem's data produce a channel on the space-time surface which guides the circuit trajectory to a position corresponding to a correct solution. A key concept for applying a time-dependent Hopfield model is the using of a set of delay filters as a sequence detector. An energy function is defined which, when presented with a known sequence builds a deep pit on the space-time energy surface with a wide valley leading to it.

An analysis of the hidden structure of speech was performed by **Jeffrey Elman** and **David Zipser** of the Institute of Cognitive Science at the University of California at San Diego [ELMAN]. Using back-propagation, Elman and Zipser taught an ANS, a series of speech recognition tasks, and after examining the resulting ANS internal representations found that the representations often corresponded to known speech representational units such as: diphones, context-sentence allophones, phonemes, syllables, and morphemes.

Michael Meyers, Robert Kuczewski and William Crawford of TRW's AI Center in San Diego, ran experiments to investigate ANS self-organization and temporal compression methods using English (a form of artificial speech) based on text from a children's book [MEYERS]. They report that the development of a self-organization ANS which concurrently performs dimension reduction, pattern recognition, and new pattern learning. Pre-processing consisted of a linear shift invariant KCM transform and using such a stabilized high dimensional vector time series the ANS learns hierarchical features and their correlations.

Tariq Samad of Honeywell, Incorporated, described an application of back propagation to determine the correct set of features corresponding to words in an input sentence [SAMAD]. Samad states that human cognitive functions such as the acquisition of concepts, tolerance of error and noisy input, graceful degradation have eluded solutions in traditional computer approaches but can be solved as side-effects with ANS's. He reviewed previous related work on parsing, case-role assignment, and word-sense disambiguation and related his ANS to recent work by McClelland and Kawamoto.

- 1) The ANS outputs, an association of features with input words where the ANS learned concepts such as "proper noun", "animate-common-noun", and "inanimate-common-noun".
- 2) The ANS outputs were as in 1) concepts such as "plural" and "read", the grammar was also extended preferred associations.
- 3) More connections were made with central (in a window) words than off-center words to enable preferred associations.

K. Torkkola, H. Rittinen and T. Kohonen reported on the results of a microprocessor-based word recognizer for a large vocabulary (1000 words) in [TORKKOLA]. Their system is capable of phonemic recognition using an ANS in the form of a phonotopic map. The map consists of a two-dimensional array of processing units which constitute matched filters to different phonemes. Each unit is tuned to a particular acoustical spectrum and the spectral templates of the units have a distribution which corresponds to the optimal clustering of the various phonemes. Word recognition is performed by comparison of phonemic transcriptions with reference transcriptions stored in a dictionary.

M. Cohen and S. Grossberg, of the Center for Adaptive Systems at Boston University, presented a computational theory explaining how an observer parses a speech stream into context-sensitive language representations in [COHEN, abs]. Cohen and Grossberg's theory stress the real time dynamical interactions that control the development of languages as well as learning and memory. Properties of the performance of language result from an analysis of the system constraints governing stable language learning. The process whereby internal language representations encode a speech stream in a context-sensitive fashion are analyzed. Cohen and Grossberg also show how organizational principles, important for visual processing, can be applied to language processing and, thus, a similar model can be used for spatial processing as well as temporal processing.

Stephen Grossberg and **Gregory Stone** presented models of the neural dynamics of word recognition and recall in [GROSSBERG]. A major goal of this paper was to synthesize the many experiments and models of human language processing and to show that learning rules and information rules are intimately connected. Grossberg and Stone maintain that to understand word recognition and recall data, it is necessary to analyze the computational units that subserve speech and language and it is also necessary to consider how computational units acquire behavioral memory by reacting to behavioral inputs and generating behavioral outputs. Furthermore, they assert that explanations of hidden processing assumptions that go into a model along with tests of their plausibility and ability to arise through self-organization are required. Models such as the Logogen model, verification model, and the Posner and Snyder model are reviewed using concepts such as automatic activation, limited capacity, attention, serial search, and interactive activation.

4.9 Knowledge Processing

Lokendra Shastri, currently of the University of Pennsylvania, investigated evidential reasoning in ANS's for his PhD thesis at the University of Rochester [SHASTRI]. The following two issues were considered by Shastri to be crucial to making progress in knowledge representation:

- 1) The necessity of identifying and formalizing inference structures that are appropriate for dealing with incompleteness and uncertainty. An agent cannot maintain complete knowledge about any but the most trivial environments, and therefore, he must be capable of reasoning with incomplete and uncertain information.
- 2) The importance of computational tractability. An agent must act in real-time. Human agents take a few hundred milliseconds to perform a broad range of intelligent tasks, and we should expect agents endowed with artificial intelligence to perform similar tasks in comparable time.

Shastri regards these two issues to be intimately related and stressed that computational tractability is not solely concerned with efficiency, optimizing programs, or faster machines. The main issue, he believes, is to establish the existence of a computational account of how an ANS may draw valid conclusions within time constraints which a given environment allows. Shastri contends that the full power of parallelism can be exploited only if it is taken, as an essential premise used to guide searches for interesting problem solutions in the space of possible knowledge representation frameworks, as opposed to finding serial solutions and subsequent parallel implementations.

John Barnden of Indiana University, devised an abstract computational architecture that can embody complex data structures and associated manipulations [BARNDEN]. Barnden addresses many of the issues raised in an NSF sponsored 1986 workshop on ANS's [MCCLELLAND, J., FELDMAN, J., BOWER, G., and MCDERMOTT, D., "Report of a Workshop on Connectionism Instigated by NSF, 1986]. In particular, an opinion expressed at the workshop that "Connectionism has not yet shown its adequacy for dealing with complex — perhaps deeply nested—representations and connectionist problems concerning crosstalk, explosive proliferation, binding and control, are addressed by Barnden in his exposition of his architecture. Barnden's architecture is based on two-dimensional arrays called "configuration matrices" which contain positioned occurrence of basic symbols.

James A. Anderson of Brown University described cognitive capabilities such as concept formation, inference, and guessing associated with the "Brain-State-in-a-Box" (BSB) ANS are nearly identical to those used by Hopfield for continuous valued systems. Information handling capabilities were characterized as follows: 1) Poor handling of precise data, 2) Inefficient use of memory in a traditional serial computer, 3) Several parameters must be tuned, and 4) Outputs may be distorted (or incorrect).

Anderson contends that the above information processing deficiencies are the cost that must be paid for the advantages obtained from ANS's, like BSB when applied to knowledge databases.

James Anderson, Richard Golden and **Gregory Murphy** discussed an ANS with a Hebbian learning rule in [ANDERSON]. They showed that the Brain-State-in-a-Box (BSB) model is similar to a gradient descent algorithm and how the behavior of many ANS's can be viewed in a probabilistic framework. They further showed that Hebbian learning and autoassociative Widrow-Hoff learning can be considered to be ways of estimating the form of an associated probability density function of the form:

$$P(x) = k \exp(-\frac{1}{2} x^T A x)$$

where "A" = a weight matrix and the pdf can adequately represent any arbitrary pdf of binary-valued stimulus vectors that may occur in the model's environment.

Concepts in distributed representations, implications of error correction for concept formation, retrieval, redundancy, and disambiguation by context were also discussed in the paper.

Stephen Grossberg, of Boston University and **William Gutowski**, of Merrimack College, analyzed the neural dynamics of decision making under risk [GROSSBERG]. This paper contains a review of models of human decision making under risk. Utility theory and prospect theory are discussed and model deficiencies are pointed out. Prospect theory, for example, does not account for non-rational decision theory such as preference reversal where in a binary choice situation an individual prefers an alternative which has been judged to be worth less than the nonpreferred alternative. A new theory of decision making under risk, called affective balance theory, is described which is an application of previous Grossberg theories of how cognitive and emotional processes interact. The previous theory was used to model perception, attention, motivation, learning, and memory. Affective balance theory is based on psychophysiological mechanisms and processes derived from analyzing relevant data.

Data Ballard of the University of Rochester, investigated a completely parallel connectionist inference mechanism based on energy minimization [BALLARD]. Ballard used a relaxation algorithm to produce inferences in first order logic derived from a very large knowledge base. He showed that for first order predicate calculus formulas and inferences rules a proof producer (resolution) can be uniquely expressed as a neural network with a very simple form. The implementation of first order logic constraints yield two coupled networks, namely, a clause network that represents clause syntax and a binding network representing relationships between terms in different clauses.

Michael Cohen and **Stephen Grossberg** of the Center for Adaptive Systems at Boston University, discussed a network with very extensive capabilities in hypothesis formation, anticipation, and prediction [COHEN]. The network is called a "masking field" and is a multiple scale, self-similar, cooperative-competitive feedback network with automatic gain control. Cohen and Grossberg discuss context sensitive grouping in recognition processes, cognitive rules arising from network interactions, and how a masking field possesses predictive anticipatory or priming capabilities. Their analytic arguments contain several references to simulation results obtained in their research on masking fields.

Claude Cruz, William Hanson and **Jason Tom** of the IBM Palo Alto Scientific Center, described the research activities in knowledge processing in [CRUZ, abs]. They introduced a conceptual scheme for representing knowledge and for performing inferences called "Knowledge Representation Networks" (KRN). The network consists of a set of knowledge entities (KE's) having varying states of activity. It is postulated that only three basic KE's are required to produce more complex KE's, namely: 1) features, 2) relationships between features, and 3) operations. A KRN inference results from a change in state of a KE which leads to a change in state of another KE. A KRN contains five mechanisms which results in knowledge processing: 1) **Implication**: a forward-chaining evidential reasoning mechanism; 2) **Context**: the KRN's current state affects subsequent sensory processing; 3) **Goal-decomposition**: a forward chaining mechanism reducing the operations to sets of coordinated sub-operations; 4) **Goal-biasing**: a mechanism which enables events to trigger operations in a production-rule-like manner; and 5) **Expectations**: a mechanism which enables operations to be executed in a closed loop position using knowledge of events expected to occur as the operation is executed.

4.10 Robotics/Control

Stephen Grossberg presented an analysis of adaptive sensory-motor control in [GROSSBERG]. The analysis considered the developmental and learning problems that an ANS (also real brain system) must solve to enable accurate performance in a dynamic real-world environment. The main emphasis of Grossberg's sensory-motor control analysis was on visually guided motor behavior and his results are considered to be relevant to issues such as localization, orienting, sensorimotor interfacing, and the design of motor pattern generating circuitry. Grossberg illustrates general organizational principles and mechanisms through analyzing the mammalian saccadic eye movement system. The issue of infinite regress, where changes in serial subsystems can interact to undo changes in other subsystems, was addressed by introducing attentionally mediated interesting sensory cues.

Geoffrey Hinton of Carnegie-Mellon University and **Paul Smolensky** of the Institute for Cognitive Science at the University of California at San Diego, analyzed mass-spring model of motor control using a neural-network [HINTON]. ANS methods which can solve the problems of finding necessary torques and generating internal representations of desired trajectories for reaching movements of arm and body are identified. Hinton and Smolensky comparatively analyzed many conventional methods of robotic control and found, for example, that desired final robot configurations via using length-tension characteristics to set end points is ineffective control. They also discussed Raibert's massive memory control table approach and the Luh-Walker-Paul sequential control algorithm.

5.0 REFERENCES

1. David H. Ackley, Stochastic Iterated Genetic Hillclimbing, CMU-CS-87-107, Department of Computer Science, Carnegie Mellon University, Pittsburgh, PA 15213
2. Joshua Alspector and Robert B. Allen, A Neuromorphic VLSI Learning System, Proc. of the 1987 Stanford Conf. on Adv. Res. in VLSI
3. James A. Anderson, Richard M. Golden, and Gregory L. Murphy, Concepts in Distributed Systems, S.P.I.E. Institute on Hybrid and Optical Computing, Dept. of Psychology, Brown University, Providence RI 02912
4. James A. Anderson and Gregory L. Murphy, Psychological Concepts in a Parallel System, Physica D 22 318-331, North-Holland, Amsterdam, Dept. of Psychology & Center for Cognitive Science, Brown University Providence, RI
5. James A. Anderson, Disordered Systems and Biological Organization, Nato Advanced Research Workshop
6. K.L. Babcock and R.M. Westervelt, Stability and Dynamics of Simple Electronic Neural Networks with Added Inertia, Physica 23D (1986) 464-469 North-Holland, Amsterdam
7. K.L. Babcock and R.M. Westervelt, Complex Dynamics in Simple Neural Circuits, Harvard University
8. Bill Baird, Nonlinear Dynamics of Pattern Formation and Pattern Recognition in the Rabbit Olfactory Bulb, Physica 22D (1986) 150-175, North-Holland, Amsterdam, Department of Biophysics, University of California, Berkeley, CA 94611
9. Pierre Baldi, Neural Networks, Orientations of the Hypercube and Algebraic Threshold Functions, IEEE Transactions on Information Theory, Dept. of Mathematics, University of California, San Diego, La Jolla, CA 92093
10. Pierre Baldi and Santosh S. Venkatesh, Number of Stable Points for Spin-Glasses and Neural Networks of Higher Orders, Physical Review Letters Vol. 58, Number 9, 1 March 1987
11. Dana H. Ballard, Cortical Connections and Parallel Processing: Structure and Function, The Behavioral and Brain Sciences (1986) 9, 67-120, University of Rochester, Rochester, NY 14627
12. Dana H. Ballard, Cortical Connections and Parallel Processing: Structure and Function, TR 133 (revised) January 1985, University of Rochester
13. John Barnden, Complex Cognitive Information-Processing: A Computational Architecture with a Connectionist Implementation. TR 211, December 1986, Indiana University, Bloomington, IN 47405-4101
14. Andrew G. Barto, Learning by Statistical Cooperation of Self-Interested Neuron-like Computing Elements, COINS Technical Report 85-11 April 1985, University of Massachusetts, Amherst, MA 01003
15. D. G. Bounds, Numerical Simulations of Boltzmann Machines, Proc Neural Network for Computing, Snowbird, Utah 1986
16. D. J. Burr, Matching Elastic Templates, Proceedings of the Royal Society Symposium on Physical and Biological Processing of Images, London, September 27-29, 1982
17. D. J. Burr, Experiments with a Connectionist Text Reader, Bell Comm. Res., 1987
18. D. J. Burr, A Neural Network Digit Recognizer, Bell Comm. Res.
19. David J. Burr and Stephen Jose Hanson, Knowledge Representation in Connectionist Networks, Bell Comm. Res.
20. Gail A. Carpenter and Stephen Grossberg, Associative Learning, Adaptive Pattern Recognition, and Cooperative-Competitive Decision Making by Neural Networks Source: Hybrid and Optical Computing, SPIE 1986
21. Gail A. Carpenter and Stephen Grossberg, Northeastern University, Boston, MA 02115 and Boston University, A Massively Parallel Architecture for a Self-Organizing Neural Pattern Recognition Machine, "Computer Vision and Image Processing", 37 (1987)
22. Gail A. Carpenter and Stephen Grossberg, Mammalian Circadian Rhythms: A Neural Network Model, Lectures on Mathematics in the Life Sciences, Vol. 19, 1987
23. Patrick Castelaz, John Angus, and Joan Mahoney, Application of Neural Networks to Expert System and Command & Control Systems, Hughes Aircraft, Fullerton, CA
24. Michael A. Cohen and Stephen Grossberg, Masking Fields: A Massively Parallel Neural Architecture for Learning, Recognizing, and Predicting Multiple Groupings of Patterned Data, Applied Optics, in press, May, 1987, Center for Adaptive Systems, Boston University, Boston, MA 02215
25. Garrison W. Cottrell, Paul Munro, David Zipser, Image Compression by Back Propagation: An Example of Extensional Programming, UCSD Institute for Cognitive Sciences
26. William Crawford, Michael Myers, Robert Kuczewski, Application of New Artificial Neural System Information Processing Principles to Pattern Classification, TRW (Rancho Carmel, San Diego) Technical Memo
27. Claude A. Cruz, William A. Hanson, Jason Y. Tam, Knowledge Processing Through Flow-Of-Activation, To Appear in Proceedings of 1987 IEEE First Annual International Conference on Neural Networks IBM Palo Alto Scientific Center, Palo Alto, CA 94304
28. Claude A. Cruz, William A. Hanson, Jason Y. Tam, Neural Network Emulation Hardware Design Considerations, To appear in Proceedings of 1987 IEEE First Annual International Conference on Neural Networks IBM Palo Alto Scientific Center, Palo Alto, CA 94304

29. M. Cohen and S. Grossberg, Neural Dynamics of Speech and Language Coding: Developmental Programs, Perceptual Grouping, and Competition for Short-Term Memory, *Human Neurobiology* (1986)5:1-22
30. John S. Denker, Neural Network Models of Learning and Adaptation, AT&T Bell Laboratories, Holmdel, NJ 07733
31. G. Ed. Denker, Neural Networks for Computing, American Institute of Physics, 1986, Proceedings 151.
32. H. Dreyfus, From Socrates to Expert Systems: The Limits of Calculative Rationality, *Bulletin of the Academy of Arts and Sciences*, January 1987, No. 4, Vol XL.
33. Omer Egecioglu, Terence R. Smith and John Moody, Computable Functions and Complexity in Neural Networks, Computer Science Dept, U. of Cal. Santa Barbara, Santa Barbara, CA
34. Jeffrey L. Elman, James L. McClelland, Exploiting Low Variability in the Speech Wave, For MIT Conference on Invariance and Variability in Speech Processes October 8-10, 1983
35. Jeffrey L. Elman, Connectionist Approaches to Acoustic/Phonetic Processing, Dept. of Linguistics, University of California, San Diego
36. Mark Fanty, A Connectionist Simulator for the BBN Butterfly Multiprocessor, TR 164 January 1986, U. of Rochester
37. Jerome A. Feldman, Connectionist Models and Parallelism in High Level Vision, *Computer Vision, Graphics, and Image Processing* 31, 178-200 (1985)
38. Jerome A. Feldman, Energy and the Behavior of Connectionist Models, TR 155, November 1985, U. of Rochester
39. Jerome A. Feldman, Neural Representation of Conceptual Knowledge, TR 189 June 1986, U. of Rochester
40. A. D. Fisher, On Applying Associative Networks, Submitted to: IEEE First Annual International Conference on Neural Networks, San Diego, CA June 1987.
41. Arthur D. Fisher, Robert C. Fukuda, John N. Lee, Implementations of Adaptive Associative Optical Computing Elements, SPIE, Vol. 625 Optical Computing (1986)
42. W. J. Freeman, Simulation of Chaotic EEG Patterns with a Dynamic Model of the Olfactory, *System Biol. Cybern.* 55 (1987)
43. Walter J. Freeman, M.D., Hardware Simulation of Brain Dynamics in Learning: The SPOCK, Submitted for the IEEE 1st Annual Intl. Conf. on Neural Networks, San Diego, CA 21-24 January 1987
44. Walter J. Freeman, Semi-Autonomous Control of Input by Nerve Cell Assemblies Through Genesis of Images, CH1935-6/83, 1983 IEEE, U. of Calif., Berkeley
45. Frisby, Seeings, Oxford University Press.
46. Kuniyiko Fukushima, Neocognitron: A Brain-Like System for Pattern Recognition, *Denshi Tokyo No. 25*(1986) NHK Science & Technical Research Laboratories 1-10-1, Kinuta, Setagaya, Tokyo 157, Japan
47. Kuniyiko Fukushima, Sei Miyake and Takayuki Ito, Neocognitron: A Biocybernetic Approach to Visual Pattern Recognition, *NHK No. 336* September 1986 (ISSN 0027-657X), NHK Science & Technical Research Laboratories 1-10-11, Kinuta, Setagaya, Tokyo 157, Japan
48. Kuniyiko Fukushima, A Neural Network Model for Selective Attention in Visual Pattern Recognition, *Biol. Cybern.* 55, 5-15(1986)
49. Kuniyiko Fukushima and Takayuki Ito, A Neural Network Model Extracting Features from Speech Signals, *The Transactions of the Institute of Electronics, Information and Communication Engineers, Japan*, Vol. J70-D, No. 2, pp. 451-462
50. Kuniyiko Fukushima, Neocognitron: A Brain-Like System for Pattern Recognition, *Denshi Tokyo* (published by the Tokyo Section, IEEE) No. 25 (1986)
51. Richard M. Golden, A Developmental Neural Model of Visual Word Perception, *Cognitive Science* 10, 241-276(1986)
52. Richard M. Golden, The "Brain-State-in-a-Box" Neural Model Is a Gradient Descent Algorithm, Reprinted from *Journal of Mathematical Psychology* Vol. 30, No. 1, March 1986
53. Leslie M. Goldschlager, A Computational Theory of Higher Brain Function, Stanford Computer Science Report
54. R. Paul Gorman, Terrence J. Sejnowski, Learned Classification of Sonar Targets Using a Massively-Parallel Network, *Proc. of the Digital Signal Processing Workshop* (1986) Sponsored by: IEEE Acoustics, Speech and Signal Processing Society Rev. 4/26/87
55. H.P. Graf, L.D. Jackel, R.E. Howard, B. Straughn, J.S. Denker, W. Hubbard, D.M. Tennat, and D. Schwartz, VLSI Implementation of a Neural Network Memory with Several Hundreds of Neurons, AT&T Bell Laboratories, Holmdel, NJ 07733
56. Stephen Grossberg, Cortical Dynamics of Three-Dimensional Form, Color, and Brightness Perception: I. Monocular Theory, *Perception & Psychophysics* 1987, 41 (2), 87-116
57. Stephen Grossberg and Ennio Mingolla, Computer Simulation of Neural Networks for Perceptual Psychology, *Behavior Research Methods, Instruments, & Computers* 1986, 18 (6), 601-607
58. Stephen Grossberg, Cortical Dynamics of Three-Dimensional Form, Color, and Brightness Perception: II. Binocular Theory, *Perception & Psychophysics* 1987, 41 (2), 117-158

59. Stephen Grossberg and Ennio Mingolla, Neural Dynamics of Surface Perception: Boundary Webs, Illuminants, and Shape-From-Shading, Source: from "Computer Vision and Image Processing", 37, (1987)
60. Stephen Grossberg and William E. Gutowski Neural Dynamics of Decision Making Under Risk: Affective Balance and Cognitive-Emotional Interactions Source: Psychological Review, in press, 1986
61. Stephen Grossberg Cooperative Self-Organization of Multiple Neural Systems During Adaptive Sensory-Motor Control Source: Center for Adaptive Systems, Boston University, Boston, MA 02215
62. Stephen Grossberg and Gregory Stone Neural Dynamics of Attention Switching and Temporal-Order Information in Short-Term Memory Memory & Cognition 1986, 14 (6), 451-468
63. A. Guez, V. Protopopescu, J. Barhen, On the Stability, Storage Capacity and Design of Nonlinear Continuous Neural Networks, Prepared by the Oak Ridge Natl. Laboratory, Oak Ridge, TN 37831
64. Dan Hammerstrom, Casey Bahr, Jim Bailey, Gary Beaver, Kevin Jagla, Norman May, A Development Environment for Wafer-scale Integrated Silicon Neuron-computers, Oregon Graduate Research Center, 1987
65. Dan Hammerstrom, Jim Bailey, and Mike Rudnick, Interconnect Architectures for WSI Neurocomputers, Oregon Graduate Research Center, 1987
66. Robert Hecht-Nielsen, Performance Limits of Optical, Electro-Optical, and Electronic Neurocomputers, HNC 5893 Oberlin Dr., San Diego, CA 92121, SPIE Optical and Hybrid Computing, Vol. 634, 1986
67. Robert Hecht-Nielsen, Nearest Matched Filter Classification of Spatiotemporal Patterns, HNC Hecht-Nielsen Neurocomputer Corporation
68. David Hestenes, How the Brain Works: the Next Great Scientific Revolution, Presented at the Third Workshop on Maximum Entropy and Bayesian Methods in Applied Statistics (U. of Wyoming, Aug. 1-4, 1983)
69. Geoffrey Hinton and Paul Smolensky, Parallel Computation and the Mass-Spring Model of Motor Control, Institute for Cognitive Science, U. of Calif., San Diego
70. Morris W. Hirsch, Convergence in Neural Nets, Dept. of Mathematics, U. of California, Berkeley, CA 94720
71. Geoffrey W. Hoffmann, A Neural Network Model Based on the Analogy with the Immune System, Dept. of Physics & Microbiology, University of British Columbia, Vancouver, B.C., Canada V6T 2A6
72. Geoffrey W. Hoffmann, Maurice W. Benson, Geoffrey M. Bree and Paul E. Kinahan, A Teachable Neural Network Based on an Unorthodox Neuron, Physica 22D (1986) 233-246, North-Holland, Amsterdam
73. J. J. Hopfield, A. Gelperin and D. W. Tank, The Logic of Limax Learning, AT&T Bell Laboratories, Murray Hill, NJ 07974 and (AG) Dept. Biology, Princeton Univ., Princeton, NJ 08544, (JJH) Div. of Chemistry and Biology, Calif. Institute of Tech., Pasadena, CA 91125
74. John J. Hopfield and David W. Tank, Computing with Neural Circuits: A Model, Articles 8 August 1986
75. W. Hubbard, D. Schwartz, J. Denker, H.P. Graf, R. Howard, L. Jackel, B. Straughn, D. Tennant, Electronic Neural Networks, AT&T Bell Laboratories, Holmdel, NJ 07733
76. B.A. Huberman, W. Scott Stornetta, An Improved Three-Layer, Back Propagation Algorithm, ICNN 1987, pre-print
77. B.A. Huberman, T. Hogg, Adaptation and Self-Repair in Parallel Computing Structures, Vol. 52, Number 12, Physical Review Letters, 19 March 1984
78. James M. Hutchinson and Christof Koch, Simple Analog and Hybrid Networks for Surface Interpolation, "Neural Networks for Computing", ed. T.S. Denber, pp. 235-239, American Institute of Physics, New York, 1986
79. Michael I. Jordan, Serial Order: A Parallel Distributed Processing Approach, ICS Report 8604
80. Pentti Kanerva, Sparse, Distributed Memory for Patterns and Sequences, RIACS, NASA AMES
81. Pentti Kanerva, Parallel Structures in Human and Computer Memory, RIACS Technical Report TR-86.2, January 1986
82. Alan H. Kawamoto and James A. Anderson, A Neural Network Model of Multistable Perception, Acta Psychologica 59 (1985)35-65 North Holland
83. James D. Keeler, Comparison Between Sparsely Distributed Memory and Hopfield-Type Neural Network Models, submitted to J. Cog Sci also RIAES Tech. Report 86.31
84. James D. Keeler, Information Capacity of Hebbian Neural Networks, submitted to Phys. Rev. Letters PACS Numbers: 87.30, 89.70
85. A. Harry Klopff, Drive-Reinforcement Learning: A Real-Time Learning Mechanism for Unsupervised Learning, Submitted to the IEEE First Annual International Conference on Neural Networks, San Diego, California 21-24 June 1987
86. Christof Koch, Jose Marroquin, and Alan Yuille, Analog "Neuronal" Networks in Early Vision, Proc. Natl. Acad. Sci. Vol. 83, pp. 4263-4267, June 1986
87. Teuvo Kohonen, Adaptive, Associative, and Self-Organizing Functions in Neural Computing, #135, Helsinki University of Technology, Dept. of Technical Physics Rakentajanaukio 2 C, SF-02150 Espoo, Finland

88. Teuvo Kohonen, Representation of Sensory Information in Self-Organizing Feature Maps, and Relation of These Maps to Distributed Memory Networks, #129, Helsinki University of Technology, Dept. of Technical Physics Rakentajanaukio 2 C, SF-02150 Espoo, Finland
89. Bart Kosko, Adaptive Bidirectional Associative Memories, To appear in "Applied Optics", Nov. 1987
90. Bart Kosko, Competitive Adaptive Bidirectional Associative Memories, ICNN, 1987
91. S. Y. Kung, and H. K. Liu, An Optical Inner-product Array Processor For Associative Retrieval, A reprint from O-E LASE '86
92. Alan Lapedes and Robert Farber, A Self-Optimizing, Nonsymmetrical Neural Net for Content Addressable Memory and Pattern Recognition, Physica 22D (1986) 247-259
93. Alan Lapedes and Robert Farber, Programming a Massively Parallel, Computation Universal System: Static Behavior, Los Alamos National Laboratory
94. Lerner, Computer Vision Research Looks to the Brain, High Technology, May 1980.
95. Daniel S. Levine, A Neural Network Model of Temporal Order Effects in Classical Conditioning Modelling of Biomedical Systems IMACS, 1986, Modeling of Biomedical Systems, Elsevier Science Pub., 1986
96. Daniel S. Levine, A Neural Network Theory of Frontal Lobe Function, Proc. of the Eighth Annual Conf. of the Cognitive Science Soc. (Amherst, MA, 1986), Erlbaum
97. Daniel S. Levine, Neural Population Modeling and Psychology: A Review, Mathematical Biosciences 6:1983.
98. Ralph Linsker, From Basic Network Principles to Neural Architecture: Emergence of Spatial-Opponent Cells, Proc. Natl. Acad. Sci. USA Vol. 83, pp. 7508-7512, October 1986
99. R. Lippmann, An Introduction to Computing with Neural Nets, IEEE, ASSP Magazine, April 198.
100. Jeffrey Locke, David Zipser, Learning the Hidden Structure of Speech, Report Number ICS 8701, UCSD
101. James L. McClelland, Jeffrey L. Elman, The TRACE Model of Speech Perception, ICS, UCSD
102. Robert J. McEliece, Edward C. Posner, Eugene R. Rodemich and Santosh S. Venkatesh, The Capacity of the Hopfield Associative Memory, IEEE Trans. Inform Theory, 1987 (To Appear)
103. Michael C. Mozer, RAMBOT: A Connectionist Expert System That Learns by Example, ICS, UCSD
104. C. Mozer, Early Parallel Processing in Reading: A Connectionist Approach, ICS, UCSD
P. Mueller, J. Lazzaro, A Machine for Neural Computation of Acoustical Patterns With Application to Real Time Speech Recognition, U. of Penn., Phila., PA 19104-6059, Calif. Inst. of Technology, Pasadena, CA
105. Michael H. Myers, Some Speculations on Artificial Neural System Technology, NAECON 1986 Proc.
106. Stephen M. Omohundro, Efficient Algorithms with Neural Network Behavior, Dept. of Comp. Sci, U. of Illinois at Urbana-Champaign
107. J. Michael Oyster, Walter Broadwell, Fernando Vicuna, Associative Network Applications to Robot Vision, Report No. 320-2777, Los Angeles Scientific Center
Ian Parberry, Georg Schnitger, Relating Boltzmann Machines to Conventional Models of Computation, Dept. of Computer Science, Pennsylvania Univ.
108. David B. Parker, A Comparison of Algorithms for Neuron-Like Cells, 1986 AIP Proc.: Neural Networks for Computing
109. Barak A. Pearlmutter, Geoffrey E. Hinton, G-Maximization: an Unsupervised Learning Procedure for Discovering Regularities, Proceedings of the conference on "Neural Networks for Computing" American Institute of Physics, 1986
110. Andras J. Pellionisz, Sensorimotor Operations: A Ground for the Co-Evolution of Brain Theory with Neurobotics and Neurocomputers, Proc. IEEE 1st Ann. Internatl. Conf. on Neural Networks, San Diego, CA 1987 June
111. P. Andrew Penz, The Closeness Code: An Integer to Binary Vector Transformation Suitable for Neural Network Algorithms, Texas Instruments Inc., Dallas, TX
Tomaso Poggio, Vincent Torre and Christof Koch Computational vision and Regularization Theory Source: Reprinted from Nature, Vol. 317, No. 6035, pp. 314-319, 26 September 1985
112. D. Psaltis, J. Hong, and S. Venkatesh, Shift Invariance in Optical Associative Memories, Presented at Conf. on "Optical Computing", LA, 1986
113. Demetri Psaltis and Cheol Hoon Park, Nonlinear Discriminant Functions and Associative Memories, Calif. Inst. of Tech., Dept. of Electrical Eng., Pasadena, CA 91125
114. George N. Reeke, Jr. and Gerald M. Edelman, Selective Neural Networks and Their Implications for Recognition Automata, The Rockefeller University
115. Kathleen J. Roney and Gordon L. Shaw, Analytic Study of Assemblies of Neurons in Memory Science, Physics Dept., Univ. of Calif., Irvine, CA 92717 Mathematical Biosciences 51:25-41 (1980)

116. David E. Rumelhart, Geoffrey E. Hinton and Ronald J. Williams, Learning Representations by Back-Propagating Errors, *Nature* Vol. 323 9 October 1986 David E. Rumelhart, Donald A. Norman, Representation in Memory, Center for Human Information Processing, U. of Calif., San Diego, La Jolla, CA
117. David E. Rumelhart, et al., *Parallel Distributed Processing*, (Vol 1 & 2) (MIT Press 1986).
118. J.P. Sage, K. Thompson, and R.S. Withers, An Artificial Neural Network Integrated Circuit Based on MNOS/CCD Principles, MIT Lincoln Lab.
119. Tariq Samad, Refining and Redefining the Back-Propagation Learning Rule for Connectionist Networks, Submitted to IEEE Systems, Man & Cybernetics Conf.
120. Tariq Samad, A Connectionist Network That Learns to Process Some (Very) Simple Sentences, Honeywell, Inc., 1000 Boone Ave.N, Golden Valley, MN 55427
121. Richard J. Sasiela, Forgetting As A Way To Improve Neural-Net Behavior, American Inst. of Physics, *Neural Networks for Computing* 1986
122. Lokendra Shastri, Evidential Reasoning in Semantic Networks: A Formal Theory and its Parallel Implementation, TR 166, University of Rochester
123. Gordon L. Shaw and Kathleen J. Roney, Analytic Solution of a Neural Network Theory Based on an Ising Spin System Analogy, *Physics Letters*, Vol. 74A, number 1,2 29 Oct 1979
124. Gordon L. Shaw, Physics Dept., Univ. of California, Irvine, CA 92717, Space-Time Correlation of Neuronal Firing Related to Memory Storage Capacity, *Brain Research Bulletin*, Vol 3, pp. 107-113, 1978
125. Gordon L. Shaw, Dennis J. Silverman, and John C. Pearson, Model of Cortical Organization Embodying a Basis for a Theory of Information Processing and Memory Recall, (Hebb synapse/selective adaptive network/axial next-nearest-neighbor Ising model/fluctuations/synchronous time steps) *Proc. Natl. Acad. Sci. USA* Vol. 82, pp. 2364-2368, April 1985
126. R. M. Siegel and R. A. Andersen, Representation of Head-Centered Visual Space in the Inferior Parietal Lobule of Macaque Monkey, The Salk Inst., San Diego
127. Christine A. Skarda and Walter J. Freeman, Brains Make Chaos to Make Sense of the World, *Behavioral and Brain Sciences*, June 1987
128. S.L. Small, L. Shastri, M.L. Brucks, S.G. Kaufman, G. W. Cottrell, and S. Addanki, ISCON: A Network Construction Aid and Simulator for Connectionist Models, TR 109, U. of Rochester
129. B.H. Soffer, G.J. Dunning, Y. Owechko, and E. Marom, Associative Holographic Memory with Feedback Using Phase-Conjugate Mirrors, Reprinted from *Optics Letters*, Vol. 11, page 118, February 1986
130. H. Sompolinsky, *Neural Networks with Nonlinear Synapses and a Static Noise*, *Physical Review*, Vol 34, No. 3, Sept. 1986
131. Harold H. Szu, Non-Convex Optimization, *SPIE Vol. 698 Real Time Signal Processing IX* (1986)
131. Harold H. Szu and Richard A. Messner, Adaptive Invariant Novelty Filters, *Proc. of the IEEE*, Vol. 74, No. 3, March 1986
132. D. W. Tank and J. J. Hopfield, Neural Computation by Concentrating Information In Time, ATT Bell Labs, 1987
133. J. Ticknor, Harrison H. Barrett, Optical Implementations in Boltzmann Machines, *Optical Engineering* 26(1), 016-021 (Jan. 1987)
134. Max Stanford Tomlinson Jr., Thesis Proposal: Analog Back Propagation Systems, UCSD, Institute for Cognitive Sciences 1987
135. Santosh S. Venkatesh and Demetri Psaltis, Linear and Logarithmic Capacities in Associative Neural Networks, Preprint submitted Mar 1985 to *IEEE Transactions on Information Theory*, Revised Nov. 1986
136. John Voevodsky, A Neural-Based Knowledge Processor, Neuraltech, Mountain View, CA
137. [von der Malsburg, C.], *Disordered Systems and Biological Organization*, Springer (1986)
138. Paul J. Werbos, Building and Understanding Adaptive Systems: A Statistical/Numerical Approach to Factory Automation and Brain Research, *IEEE Transactions on Systems, Man, & Cybernetics*, Vol. SMC-17, No. 1, January/February 1987
139. Ronald J. Williams, Reinforcement Learning in Connectionist Networks: A Mathematical Analysis, ICS, UCSD
140. Ronald J Williams, Inference of Spatial Relations, Institute for Cognitive Science UCSD C-015 La Jolla, CA 92093
141. Y.T. Zhou, R. Chellappa and B.K. Jenkins, A Novel Approach to Image Restoration, Based on a Neural Network To be presented at First Intl. Conf. on Neural Nets San Diego, June 1987
142. David Zipser, Programming Neural Nets To Do Spatial Computations, ICS Report 8608, UCSD