A Connectionist/Neural Network Bibliography

Dan Hammerstrom
Oregon Graduate Center
19600 N.W. von Neumann Dr.
Beaverton, Oregon 97006

Technical Report No. CS/E-86-010
August 1986

Abstract

This technical report contains a bibliography of connectionist and neural modelling papers and books as well as related references in the areas of computer architecture and brain science. Abstracts are included for many of the references.

BIBLIOGRAPHY

Competition and Cooperation in Neural Nets, Lecture Notes in Biomathematics, 45, Springer-Verlag, Berlin, 1982.

Agrawal, Dharma P., Virendra K. Janakiram, and Girish C. Pathak, "Evaluating the Performance of Multicomputer Configurations," IEEE Transactions on Computers, pp. 23-37, May 1986.

Allman, J., F. Miezin, and E. McGuinness, "Stimulus Specific Responses From Beyond The Classical Receptive Field: Neurophysiological Mechanisms for Local-Global Comparisons in Visual Neurons," in Ann. Rev. Neurosci., vol. 8, pp. 407-30, Annual Reviews Inc., 1985.

We perceive the visual world as a unitary whole, yet one of the guiding principles of nearly a half century of neurophysiological research since the early recordings by Hartline (1938) has been that the visual system consists of neurons that are driven by stimulation within small discrete portions of the total visual field. These classical receptive fields (CRFs) have been mapped with the excitatory responses evoked by a flashed or moving stimulus, usually a spot or bar of light. Most of the visual neurons, in turn, are organized in a series of maps of the visual field, at least 10 of which exist in the visual cortex in primates as well as additional topographic representations in the lateral geniculate body, pulvinar and optic tectum (Allman 1977, Newsome & Allman 1980, Allman & Kaas 1984). It has been widely assumed that perceptual functions that require the integration of inputs over large portions of the visual field must be either collective properties of arrays of neurons representing the visual field, or features of those neurons at the highest processing levels in the visual system, such as the cells in inferotemporal or posterior parietal cortex that typically possess very large receptive fields and do not appear to be organized in visuotopic maps. These assumptions have been based on the results of the many studies in which receptive fields were mapped with conventional stimuli, presented one at a time, against a featureless background. However, unlike the neurophysiologist's tangent screen, the natural visual scene is rich in features, and there is a growing body of evidence that in many visual neurons stimuli presented outside the CRF strongly and selectively influence neural responses to stimuli presented within the CRF. These results suggest obvious mechanisms for local-global comparisons within visuotopically organized structures. Such broad and specific surround mechanisms could participate in many functions that require the integration of inputs over wide regions of the visual space such as the perceptual constancies, the segregation of figure from ground, and depth perception through motion parallax. In the first section of this paper, we trace the historical development of the evidence of response selectivity for visual stimuli presented beyond the CRF; in the second, examine the anatomical pathways that subserve these far-reaching surround mechanisms; and in the third, explore the possible relationships between these mechanisms and perception.

Amari, S., "Field Theory of Self-Organizing Neural Nets," IEEE Transactions on Systems, Man, and Cybernetics, vol. SMC-13, no. 5, pp. 741-748, Sept/Oct 1983.

A field theory is proposed as a mathematical method for analyzing learning and self-organizing nerve nets and systems in a unified manner. It is shown by the use of the theory that a nerve net has an ability of automatically forming categorizers or signal detecting cells for the signals which the net receives from its environment. Moreover, when the set of signals has a topological structure, the detectors of these signals are arranged in the nerve system (or field) to preserve the topology, so that the topographical structure is introduced in the nerve system by self-organization.

Andersen, R. A., G. K. Essick, and R. M. Siegel, "Encoding of Spatial Location by Posterior Parietal Neurons," Science, vol. 230, pp. 456-458, Oct. 25, 1985.

The cortex of the inferior parietal lobule in primates is important for spatial perception and spatially oriented behavior. Recordings of single neurons in this area in behaving monkeys showed that the visual sensitivity of the retinotopic receptive fields changes systematically with the angle of gaze. The activity of many of the neurons can be largely described by the product of a gain factor that is a function of the eye position and the response profile of the visual receptive field. This operation produces an eye position-dependent tuning for locations in head-centered coordinated space.

Anderson, J.A. and G.E. Hinton, "Models of Information Processing in the Brain," in Parallel Models of Associative Memory, ed. G.E. Hinton, pp. 9-48, Lawrence Erlbaum Assoc., Hillsdale, NJ, 1981.

This chapter introduces some models of how information may be represented and processed in a highly parallel computer like the brain. Despite the staggering amount of information available about the physiology and anatomy of the brain, very little is really known about the nature of the higher-level processing performed by the nervous system. There is no established theory about the kinds of neural activity that occur when we hear a sentence, perceive an object, or form a plan, though data on many fascinating and significant bits and pieces are now available.

Anderson, J.A. and M.C. Mozer, "Categorization and Selective Neurons," in Parallel Models of Associative Memory, ed. J. Anderson, pp. 213-236, Erlbaum Associates, Hilsdale, N.J., 1981.

Anderson, J.A., "Cognitive and Psychological Computation with Neural Models," IEEE Trans. on Systems, Man, and Cybernetics, vol. SMC-13, no. 5, pp. 799-815, Sept./Oct. 1983.

Biological support exists for the idea that large-scale models of the brain should be parallel, distributed, and associative. Some of this neurobiology is reviewed. It is then assumed that state vectors, large patterns of activity of groups of individual somewhat selective neurons, are the appropriate elementary entities to use for cognitive computation. Simple neural models using this approach are presented that will associate and will respond to prototypes of sets of related inputs. Some experimental evidence supporting the latter model is discussed. A model for catagorization is then discussed. Educating the resulting systems and the use of error correcting techniques are discussed, and an example is presented of the behavior of the system when diffuse damage occurs to the memory, with and without compensatory learning. Finally a simulation is presented which can learn partial information, integrate it with other material, and use that information to reconstruct missing information.

Arbib, M. A., Brains, Machines and Mathematics, McGraw-Hill, New York, NY, 1964.

Arden, Bruce W. and Hikyu Lee, "A Regular Network for Multicomputer Systems," IEEE Trans. on Computers, vol. C-31, pp. 60-69, January 82.

Bailey, J. and D. Hammerstrom, "How to Make a Billion Connections," Tech. Report CS/E-86-007, Dept. of Computer Science/Engineering, Oregon Graduate Center, Beaverton, Oregon, July 1986.

This paper addresses some of the problems involved in implementing a connection network consisting of a million nodes with a billion connections between them. It is organized into three general parts. The first explains what connection networks are, how they work and what some of the proposed implementations are. The second portion is a preliminary comparison of some of the interconnect strategies that have been suggested for parallel systems and how the connectionist model maps onto them. The final contribution of this paper is a suggestion of future directions for research. For the purposes of this discussion, a connection network is a set of simple computing elements interconnected in some specified manner. Each element has a single output that is a small integral value. This value is the

result of the computational function for that node. The function is assumed to be a "Sigma-Pi" function, i.e. the sum of a set of products. The inputs to this function are normally the prior state of the node together with the outputs of all nodes connected to it. A set of metrics for analyzing different interconnection topologies are defined. They include speed of communication, interconnect cost, locality of communication and degree of fault tolerance. Several potential interconnect architectures are compared using the given metrics. The candidates studied fall into one of five general classifications: direct implementation of the connection network in silicon, a simplistic model such as a grid, a message routing model such as the hypercube, a shared memory model and a new proposed solution - the broadcast hierarchy. Preliminary results show that some interconnection structures are incompatible with most existing connection network models.

Ballard, D.H., "Cortical Connections and Parallel Processing: Structure and Function," TR 133, Computer Science Department, Rochester, NY, January 1985.

The cerebral cortex is a rich and diverse structure that is the basis of intelligent behavior. One of the deepest mysteries of the function of cortex is that neural processing times are only about one hundred time faster than the fastest response times for complex behavior. At the very least, this would seem to indicate that the cortex does massive amounts of parallel computation. This paper explores the hypothesis that an important part of the cortex can be modeled as a connectionsist computer that is especially suited for parallel problem solving. The connectionist computer uses a special representation, termed value unit encoding, that represents small subsets of parameters in a way that allows parallel access to many different parameter values. This computer can be thought of as computing hierarchies of sensory- motor invariants. The neural substrate can be interpreted as a commitment to data structures and algorithms that compute invariants fast enough to explain the behavioral response times. A detailed consideration of this model has several implications for the underlying anatomy and phsiology.

Barnden, J.A., "Diagrammatic Short-Term Information-Processing by Neural Mechanisms," in Cognition and Brain Theory, vol. 3-4, pp. 285-328, Lawrence Erlbaum Associates, Inc., 1984.

There are problems in trying to deploy prevailing ideas about representation in neural networks to account for complex short-term information-processing. An apparently unrelated set of difficulties arise in research on visual imagery. The theory presented in this paper solves or sidesteps both sets of problems, and it does so in an integrated way by means of a simple set of basic postulates.

The most important postulate is that array-like neural areas as seats for short-term data structures. This hypothese allows the use of physical adjacency as the fundamental means for encoding associations between items in data structures. It also naturally suggests the idea of using spatial relationships in the domain of discourse. One benefit of the proposal is the consequent smooth unification of "propositional" and "imagistic" representation and processing. The theory is sufficient well-developed for computer simulation of the proposed information-processing system to be imitated.

Barnden, J.A., "On Short-Term Information-Processing in Connectionist Theories," in Cognition and Brain Theory, vol. 7, pp. 25-59, Lawrence Erlbaum Associates, Inc., 1984.

Connectionist theories of the neural basis of cognition have concentrated on long-term memory and certain specialized short-term processing operations. They have not been much concerned with general issues of short-term information-processing. Several varieties of connectionist theory are investigated here. It is pointed out that short-term information-processing must be pervaded by the manipulation of ancillary information structures. For instance, neural assemblies must frequently be marked to indicate that they are temporarily playing some special role. Also, lists of assemblies must be rapidly created and updated. It is shown that most varieties of connectionist theory face problems of efficiency in trying to account for the manipulation of ancillary information structures.

Barto, Andrew G., "Learning by Statistical Cooperation of Self-Interested Neuron-Like Computing Elements," COINS Technical Report 85-11, Dept. of Computer and Information Science, University of Massachusetts, April 1985.

Since the usual approaches to cooperative computation in networks of neuron-like computing elements do not assume that network components have any "preferences," they do not make substantive contact with game theoretic concepts, despite their use of some of the same terminology. In the approach presented here, however, each network component, or adaptive element, is a self-interested agent that prefers some inputs over others and "works" toward obtaining the most highly preferred inputs. Here we describe an adaptive element that is robust enough to learn to cooperate with other elements like itself in order to further its self-interests. It is argued that some of the long-standing problems concerning adaptation and learning by networks might be solvable by this form of cooperativity, and computer simulation experiments are described that show how networks of self-interested components that are sufficiently robust can solve rather difficult learning problems. We then place the approach in its proper historical and

theorectical perspective through coparison with a number of related algorithms. A secondary aim of this article is to suggest that beyond what is explicitly illustrated here, there is a wealth of ideas from game theory and allied disciplines such as mathematical economics that can be of use in thinking about cooperative computation in both nervous system and man-made systems.

Blasdel, G.G. and G. Salama, "Voltage-sensitive dyes reveal a modular organization in monkey striate cortex," Nature, vol. 321, pp. 579-585, 5 June 1986.

Voltage-sensitive dyes allow neuronal activity to be studied by non-invasive optical techniques. They provide an attractive means of investigating striate cortex, where important response properties are organized in two dimensions. In the present study, patterns ocular dominance and orientation selectivity were obtained repeatedly from the same patch of cortex using the dye merocyanine oxazoline, together with current image-processing techniques. The patterns observed agree with most established features of monkey stiate cortex and suggest a new unit of cortical organization; one that is modular in structure and which appears to link the organization selectivity with that of ocular dominance.

Bobrow, D.G. and P.J. Hayes, "Artificial Intelligence - Where are we?," Artificial Intelligence, vol. 25, pp. 375-415, Elsevier Science Publishers (North-Holland), 1985.

Brachman, R. J. and J. G. Schmolze, "An Overview of the KL-ONE Knowledge Representation System," Cognitive Science, vol. 9, pp. 171-216, 1985.

KL-ONE is a system for representing knowledge in Artificial Intelligence programs. It ahs been developed and refined over a long period and has been used in both basic research and implemented knowledge-based systems in a number of places in the AI community. Here we present the kernal ideas of KL-ONE, emphasizing its ability to form complex structured descriptions. In addition to detailing all of KL-ONE's description-forming structures, we discuss a bit of the philosophy underlying the system, highlight notions of taxonomy and classification that are central to it, and include an extended example of the use of KL-ONE and its classifier in a recognition task.

Braitenberg, V., "Thoughts on the Cerebral Cortex," J. Theor. Biol., vol. 46, pp. 421-447, 1974.

The cortex is often described as a network processing information in the direction from sensory to motor areas. However, the structure of the cortex is asymmetrical only in the vertical direction, suggesting an input-output transformation between layers rather between areas. This operation must be a very generally applicable one, since the plan of the cortex is basically the same everywhere. In an attempt to understand it, a skeleton cortex of only pyramidal cells is considered. They are characterized by a double dendritic expansion, an apical one in the first layer, which is considered as the input layer, and a basal one which receives excitation from the axon collaterals of other pyramidal cells. If pyramidal cells learn (perhaps by growing dendritic spines) to respond to frequent constellations of activity in their afferents, each will learn a property of the input (through its apical dendrites) provided that it was preceded by other properties sensed by neighbouring pyramidal cells (which influences it through its basal dendrites). Thus the pyramidal cells will code the input in terms of properties which have a tendency to follow each other. This will be a coding which reflects the causal structure of the world. Various uses of a network embodying the conditional probabilities of events in the input are described, including recognition of familiar sequences and prediction. The local variation of fiber patterns in the cerebral cortex of man, described as myeloarchitectonics, is interpreted as a macroscopical expression of the different statistics of the set of conditional probabilities linking the events represented by individual pyramidal cells in different areas (in different functional contexts).

Braitenberg, V., "Cortical Architectonics: General and Areal," in _Architectonics of the Cerebral Cortex_, ed. H. Petsche, pp. 443-465, Raven Press, New York, 1978.

The most significant development since the time of von Economo is the shift of emphasis from the brain considered as a collection of cells to the brain considered as a network of fibers, from the neuropathologist's view of the brain to that of the computer engineer; in short, from Nissl pictures to wiring diagrams. This paper is intended as a contribution to speculations on the nature of the generalized cortical machinery, with some anatomical data from the mouse cortex to support those ideas which are at variance with current views.

Brown, Chappell, "Chips Designed to Mimic Nervous System," _Electronic Engineering Times_, pp. 61-62, March 24, 1986.

Carpenter, G. and S. Grossberg, "Neural Dynamics of Category Learning and Recognition: Structural Invariants, Reinforcement, and Evoked Potentials," in _Pattern Recognition and Concepts in Animals, People, and Machines_, ed. R.J. Herrnstein, Erlbaum Associates, Hillsdale, NJ, 1986.

Carpenter, G. and S. Grossberg, "A Massively Parallel Architecture for a Self-Organizing Neural Pattern Recognition Machine," in _Computer Vision, Graphics, and Image Processing_, 1986. In press

A neural network architecture for the learning of recognition categories is derived. Real-time network dynamics are completely characterized through mathematical analysis and computer simulations. The architecture self-organizes and self-stabilizes its recognition codes in response to arbitrary orderings of arbitrarily many and arbitrarily complex self-stabilizing the code learning process. The architecture embodies a parallel search scheme which updates itself adaptively as the learning process unfolds. After learning self-stabilizes, the search process is automatically disengaged. Thereafter input patterns directly access their recognition codes without any search. Thus recognition time does not grow as a function of code complexity. A novel input pattern can directly access a category if it shares invariant properties with the set of familiar exemplars of that category. These invariant properties emerge in the form of learned critical feature patterns, or prototypes. The architecture possesses a context-sensitive self-scaling property which enables its emergent critical feature patterns to form. They detect and remember statistically predictive configurations of featural elements which are derived from the set of all input patterns that are ever experienced. Four types of attentional process-- priming, gain control, vigilance, and intermodal competition-- are mechanistically characterized. Top down priming and gain control are needed for code matching and self-stabilization. Attentional vigilance determines how fine the learned categories will be. If vigilance increases due to an environmental disconfirmation, then the system automatically searches for and learned finer recognition categories. A new nonlinear matching law (the 2/3 Rule) and new nonlinear associative laws (the Weber Law Rule, the Associative Decay Rule, and the Template Learning Rule) are needed to achieve these properties. All the rules describe emergent properties of parallel network interactions. The architecture circumvents the noise, saturation, capacity, orthogonality, and linear predictability constraints that limit the codes which can be stably learned by alternative recognition models.

Churchland, P.S., Neurophilosophy, MIT Press, Cambridge, MA, 1986.

Cohen, M. and S. Grossberg, "Neural dynamics of speech and language coding: development programs, perceptual grouping, and competition for short-term memory," Human Neurobiology, vol. 5, pp. 1-22, 1986.

A computational theory of how an observer parses a speech stream into context-sensitive language representations is described. It is shown how temporal lists of events can be chunked into unitized representations, how perceptual groupins of past item sublists can be reorganized due to information carried by newly occurring items, and how item information and temporal order information are bound together into context-sensitive codes. These language units are emergent properties due to intercellular interactions among large numbers of nerve cells. The controlling neural networks can arise through simple rules of neuronal devlopment: random growth of connections along spatial gradients, activity-dependent self-similar sites. Within these networks, a spatial frequency analysis of temporally evolving activity patterns leads to competitive masking of unappropriate list encodings in short term memory. The neurons obey membrane equations undergoing shunting recurrent on-center off-surround interactions. Several design principles are embodied by the networks, such as the sequence masking principle, the long-term memory principle, and the principle of self-similar growth.

Cohen, M.A. and S. Grossberg, "Absolute Stability of Global Pattern Formation and Parallel Memory Storage by Competitive Neural Networks," IEEE Trans. on Systems, Man, and Cybernetics, vol. SMC-13, no. 5, pp. 815-826, Sept/Oct 1983.

The process whereby input patterns are transformed and stored by competitive cellular networks is considered. This process arises in such diverse subjects as the short-term storage of visual or language patterns by neural networks, pattern formation due to the firing of morphogenetic gradients in developmental biology, control of choice behavior during macromolecular evolution, and the design of stable context-sensitive parallel processors. In addition to systems capable of approaching one of perhaps infinitely many equilibium points in response to arbitrary input patterns and initial data, one finds in these subjects a wide variety of other behaviors, notably traveling waves, standing waves, resonance, and chaos. The question of what general dynamical constraints cause global approach to equilibria rather than large amplitude waves is therefore of considerable interest. In other terminology, this is the question of whether global pattern formation occurs. A related question is whether the global pattern formation property persists when system parameters slowly change in an unpredictable fashion due to self-organization (development, learning).

Cooper, L.N., F. Liberman, and E. Oja, "A Theory For The Acquisition And Loss Of Neuron Specificity In Visual Cortex," Biological Cybernetics, pp. 9-28, Springer-Verlag, 1979.

We assume that between lateral geniculate and visual cortical cells there exist labile synapses that modify themselves in a new fashion called threshold passive modification and in addition, non-labile synapses that contain permanent information. In the theory which results there is an increase in the specificity of response of a cortical cell when it is exposed to stimuli due to normal patterned visual experience. Non-patterned input, such as might be expected when an animal is dark-reared or raised with eyelids sutured, results in a loss of specificity, with details depending on whether noise to labile and non-labile junctions is correlated. Specificity can sometimes be regained, however, with a return of input due to patterned vision. We propose that this provides a possible explanation of experimental results obtained by Imbert and Buisseret (1975); Blakemore and Van Sluyters (1975); Buisseret and Imbert (1976); and Fregnac and Imbert (1977, 1978)

Cottrell, G., "Connectionist Parsing," Proceedings of the 7th Annual Cognitive Society Conference, pp. 201-211, 1985.

We have proposed a neural network style model of language processing in an effort to build a cognitive model which would simultaneously satisfy constraints from psychology and neurophysiology. This model was successful in disambiguating word senses in semantically determined sentences, but was unable to distinguish Agent from Object in semantically reversible sentences such as "John loves Mary." In this paper we rectify the matter by specifying the syntactic portion of the model, which is a massively parallel, completely distributed connectionist parser. We also describe the results of a simulation of the model.

Cottrell, G.W. and S.L. Small, "Viewing Parsing as Word Sense Discrimination: A Connectionist Approach," in Computational Models of Natural Language Processing, ed. G. Guida, pp. 91-119, Elsevier Science Publishers (North-Holland), 1984.

This paper advocates the interdisciplinary development of a computational theory of human language comprehension and proposes a collection of initial constraints from which to start on such an enterprise. In order to satisfy these constraints, our modelling effort employs an architecture significantly different from the typical computer and closer to that of the human brain. We use a particular spreading activation or active semantic network scheme, called connections, which entails a massive number of appropriately connected computing units that communicate through weighted levels of exclusion and inhibition. While such an architecture does not solve any problems per se, we believe that a number of questions become easier to set forth and more straightforward to solve. This paper surveys a number of fundamental language comprehension issues from the new perspective, and presents some simulation results of a parsing model based on these considerations.

Creutzfeldt, O.D., "The Neocortical Link: Thoughts on the Generality of Structure and Function of the Neocortex," in Architectonics of the Cerebral Cortex, ed. H. Petsche, pp. 357-383, Raven Press, New York, 1978.

The thesis is put forward that there are fundamental similarities between various neocortical areas. The evidence in support of this thesis is shortly reviewed in pointing out the common afferent, intrinsic, and efferent organizational principles of the various cortical areas with special emphasis on the generality of thalamocortical circuits.

Crick, F., "Function of the thalamic reticular complex: The searchlight hypothesis," Proc. Natl. Acad. Sci. USA, vol. 81, pp. 4586-4590, July 1984.

It is suggested that in the brain the internal attentional searchlight, proposed by Treisman and others, is controlled by the reticular complex of the thalamus (including the closely related perigeniculate nucleus) and that the expression of the searchlight is the production of rapid bursts of firing in a subset of thalamic neurons. It is also suggested that the concjuntions produced by the attentional searchlight are mediated by rapidly modifiable synapses - here called Malsburg synapses - and especially by rapid bursts acting on them. The activation of Malsburg synapses is envisaged as producing transient cell assemblies, including "vertical" ones that temporarily unit neurons at different levels in the neural network.

Davis, A.L. and S.V. Robinson, "The FAIM-1 Symbolic Multiprocessing System," Spring Compcon 1985, pp. 370-375, Febuary 1985.

Davis, B.R., "An Associative Hierarchical Self-Organizing System," IEEE Trans. on Systems, Man, and Cybernetics, vol. SMC-15, no. 4, pp. 570-579, July/Aug 1985.

A system that learns to predict events in various environments is described. The system is associative and distributed wherein a hierarchical self-organization of low-level units into high-level units takes place based upon experience in a particular domain. Its design is inspired by widely held principles of brain organization and by some newly developed techniques in nonparametric statistical inference. The system can be regarded as a realization of a nonparametric statistical algorithm. This is demonstrated by a discussion of system architecture and a presentation of an application in a "number theory" environment.

Davis, L.S. and A. Rosenfeld, "Cooperating Processes for Low-level Vision: A Survey," Artificial Inelligence, vol. 17, pp. 245-263, 1981.

Cooperating local parallel processes can be used as aids in assigning numerical or symbolic labels to image or scene parts. Various approaches to using such processes in low-level vision are reviewed, and thier advantages are discussed. Methods of designing and controlling such processes are also considered.

Derthick, M., "Learning in Boltzmann Machines and Why it's Slow," CMU-CS-84-120, Computer Science Department, Carnegie-Mellon University, 1982.

Boltzmann Machines learn to model the structure of an environment by modifying internal weights. The algorithm used for changing a weight depends on collecting statistics about the behavior of the two units that the weight connects. The success and speed of the algorithm depends on the accuracy of the statistics, the size of the weight changes, and the way in which the accuracy of the machine's model varies as the weights are changed. This paper presents theoretical analysis and empirical results that can be used to select more effective parameters for the learning algorithm.

Dwork, C., P.C. Kanellakis, and J.C. Mitchell, "On the Sequential Nature of Unification," MIT/LCS/TM-257, Lab. for Computer Science, MIT, Cambridge, MA. 02139, March 1984.

The problem of unification is log-space complete for P. In deriving this lower bound no use is made of the potentially concise representation of terms by directed acyclic graphs. In addition, the problem remains complete even if infinite substitution are allowed. A consequence of this result is that parallelism cannot significantly improve on the best sequential solutions for unification. The "dual" problem of computing the congruence closure of an equivalence relation is also log-space complete for P. However, we show that for the problem of term matching, an important subcase of unification, there is a good parallel algorithm using O(log**2 n) time and n**O(1) processors on a PRAM. For the O(log**2 n) parallel time upper bound we assume that the terms are represented by directed acyclic graphs; if the longer string representation is used we obtain an O(log n) parallel time bound.

Eccles, J.C., "The Modular Operation of the Cerebral Neocortex Considered as the Material Basis of Mental Events," Neuroscience, vol. 6, no. 10, pp. 1839-1856, 1981.

The module of the neocortex forms a unit that is central to the various conceptual developments here formulated. Its structure and its modes of operation and communication are described in some detail. The attempt is made to see if the known synaptic excitatory and inhibitory mechanisms of the module would give it unique properties as a unitary component of that part of the brain specially related to mental events. Particular reference is made to the important new discoveries, axo-axonic synapses inhibiting pyramidal cell axons and the disinhibitory action of the cellule a double bouquet of Ramon y Cajal. There seems to be no internal system of reverberatory communication within the module. However, the complexities of reverberatory operation are given with the module as a unit in three communication systems of the brain: the cortico-cortical, the thalamo-cortical and the horizontal fibre systems. It is agreed that all mental experiences are embedded in memories. An hypothesis is developed in respect of the generation of modular patterns in relation to memory and its retrieval that is based on the conjunction hypothesis of Marr. The horizontal fibres of lamina I play a key role. These patterns would also be the modular patterns correlated with self-consciousness in all of its manifestations. Diagrams of modular patterns illustrate the hypothesis and could give rise to experimental testing. It is pointed out that this hypothesis leads to concepts of modular patterning that are not at variance with the conjectures of Mountcastle, Sperry and Szentagothai with respect to the neural events in correlation with mental events, though here the dualist-interactionist hypothesis is favoured.

Edelman, G.M., "Group Selection and Phasic Reentrant Signaling: A Theory of Higher Brain Function," in The Mindful Brain, MIT Press, Cambridge, MA, 1977.

The remarkable diversity of nervous systems in various animal species and their exquisite capacity for adaptive function are both intriguing and confounding to neurobiologists. Despite their complexity, however, all nervous systems appear to obey similar general principles at the level of morphological expression of neuronal structures and in their mechanisms of signal transmission. The recognition of these general principles and their application to the study of simple nervous systems as well as to subsystems in more complex brains have been among the greatest triumphs of neurobiology in this century.

Edelman, G.M., "Group Selection as the Basis for Higher Brain Function," in The Organization of the Cerebral Cortex, ed. Stephen G. Dennis, pp. 535-563, The MIT Press, Cambridge, Mass., 1981.

The purpose of this chapter is to consider and extend several aspects of the group-selection theory of higher brain function. This theory is Darwinian in outlook: its basic premise is that preexisting structural variations in multiply connected and distributed cellular groups of the brain provide a basis for selection of those groups whose function proves to be sufficient for adaptive behavior. It is meant to apply to backboned animals, mainly mammals, and specifically to man. At present, it has no direct reference to other nervous systems such as those of insects.

This theory is posed in phenomenological terms; although it does not depend upon the exact details of localization, neural coding, or synaptic plasticity, it nonetheless must directly confront these major unsolved problems of neurobiology.

Fahlman, S.E., _NETL: A System for Representing and Using Real-World Knowledge_, MIT Press, Cambridge, MA., 1979.

Fahlman, S. E., _Design Sketch for a Million-Element NETL Machine_, Carnegie-Mellon University Department of Computer Science, Pittsburgh, PA, 1980.

Fahlman, S.E., "The Hashnet Interconnection Scheme," CMU-CS-80-125, Dept. of Computer Science, Carnegie-Mellon Univ., Pittsburgh, PA, June 2, 1980.

This paper describes a type of switching network which can simultaneously connect many inputs to many outputs. Such networks are useful in communications, in dynamically reconfigurable compter systems, and in building semantic network memories for artificial intelligence and data base applications. This paper will explore the issues that arise in the design of practical hashnets, and will describe how to design a near-optimal hashnet for any desired combination of size and blocking probability. It will also present a scheme for time-sharing such networks, which greatly multiplies the number of (virtual) connections available with a given amount of hardware.

Fahlman, S.E., G.E. Hinton, and T.J. Sejnowski, _Massively Parallel Architectures for AI: NETL, Thistle, and Boltzmann Machines_, Proc. of the AAAI-83, Washington, D.C., August 1983.

It is becomming increasingly apparent that some aspects of intelligent behavior require enormous computational power and that some sort of massively parallel computing architecture is the most plausible way to deliver such power. Parallelism, rather than raw speed of the computing elements, seems to be the way that the brain gets such jobs done. But even if the need for massive parallelism is admitted, there is still the question of what kind of parallel architecture best firts the needs of various AI tasks. In this paper we will attempt to isolate a number of basic computational tasks that an intelligent system must perform. We will describe several families of massively parallel computing architectures, and we will see which of these computational tasks can be handled by each of these families. In particular, we will describe a new architecture, which we call the Boltzmann machine, whose abilitiese to include a number of tasks that are inefficient or impossible on the other architectures.

Fanty, Mark, "A Connectionist Simulator for the BBN Butterfly Multiprocessor," TR 164, Dept. of Computer Science, University of Rochester, Rochester NY, January, 1986.

Feldman, J. A., "A Connectionist Model of Visual Memory," in _Parallel Models of Associative Memory_, ed. J. Anderson, pp.

49-81, Erlbaum Associates, Hilsdale, N.J., 1981.

Feldman, J.A., "Memory and Change in Connection Networks," TR 96, Dept. of Computer Science, Univ. of Rochester, Rochester, NY, December 1981.

Artificial intelligence and articulating cognitive sciences have made great progress by employing models based on conventional digital computers as theories of intelligent behavior, but a number of crucial phenomena have not yielded to this treatment. The Rochester group (among others) has been exploring the alternative of employing massively parallel connectionist models (CM). A series of papers has developed the foundations for detailed modeling and applied these ideas to a number of difficult problems in vision with considerable success. But all of the detailed technical work was concerned with the structure and behavior of fixed networks. The purpose of this paper is to extend the methodology to cover several aspects of change and memory. The emphasis here is on general mechanisms that appear to be applicable in a wide range of modeling tasks; subsequent papers will treat specific problems with adequate care.

Feldman, J.A. and D.H. Ballard, "Connectionist Models and Their Properties," Cognitive Science, vol. 6, pp. 205-254, 1982.

Much of the progress in the fields constituting cognitive science has been based upon the use of explicit information processing models, almost exclusively patterned after conventional serial computers. An extension of these ideas to massively parallel, connectionist models appears to offer a number of advantages. After a preliminary discussion, this paper introduces a general connectionist model and considers how it might be used in cognitive science. Among the issues addressed are: stability and noise-sensitivity, distributed decision-making, time and sequence problems, and the representation of complex concepts.

Feldman, J.A., "Energy and the Behavior of Connectionist Models," TR 155, Dept. of Computer Science, Univ. of Rochester, Rochester, NY, November 1985.

Massively parallel (connectionist) compuational models are playing an increasingly important role in cognitive science. Establishing the behavioral correctness of a connectionist model is exceedingly difficult, as it is with any complex system. For a restricted class of models, one can define an analog to the energy function of physics and this can be used to help prove properties of a network. This paper explores energy and other techniques for establishing that a network meets its specifications. The treatment is elementary, computational, and focuses on specific examples. No free lunch is offered.

Feng, Tse-yun, "A Survey of Interconnection Networks," IEEE Computer, pp. 12-27, December 1981.

Freeman, W.J., Mass Action in the Nervous System, Academic Press, New York, 1975.

Examination of the Neurophysiological basis of adaptive behavior through the EEG.

Freeman, W.J., "A Physiological Hypothesis of Perception," in Perspectives in Biology and Medicine, pp. 561-592, Summer 1981.

The brain is a particular kind of physical object that has the property of making within itself representations of the outside world. These representations exist in transient bursts of energy that constitute the objective aspect of mental images. The representations are shaped by sensory input; following on sequential transformations as they serve to shape motor activity and to predict future sensory input. The task of the physiologist is to describe the material substrates of these representations, their physical locations and patterns, and the operations by which they are constructed and transformed in the brain. In short, what are the physical forms of mental images, and how does the brain make them? This essay reviews the conceptual bases of the task, and outlines an experimental analysis of the operations that underlie perception in the simplest mammalian sensory system, that for olfaction. It is shown that the operation of constructing a perceptual representation from sensory input cannot be described by the classical stimulus-response paradigm. The operation consists of a state change in the brain that leads to self-ordering of neural activity. It can be modeled with mathematics used to describe self-organization in nonequilibruim systems and not with deterministic equations.

Freeman, W.J. and W. Schneider, "Changes in Spatial Patterns of Rabbit Olfactory EEG with Conditioning to Oders," Psychophysiology, vol. 19, no. 1, pp. 44-56, 1982.

Electrode arrays were implanted epidurally on the olfactory bulbs of rabbits for EEG recording. The rabbits were trained to give a conditional response to a warning oder paired with an electric shock. EEGs were recorded and edited, and representative EEG bursts with oder and preceding the oder were selected for measurement. Each burst was displayed in a contour map of amplitude. The contour maps revealed active EEG foci in the bulb with size, shape and location unique to each rabbit. Changes in shape and location took place only during familiarization and during training, when a warning oder was paired with the aversive stimulus. The EEG spatial patterns did not change when visual or auditory stimuli were used as CS. EEG spatial

patterns did not reflect conformal mapping of oder stimulus to neural activity response, but were determined by state variables of the animal related to olfactory conditioning history. The implications for human EEG are briefly discussed.

Fukushima, K., S. Miyake, and T. Ito, "Neocognitron: A Neural Network Model for a Mechanism of Visual Pattern Recognition," IEEE Trans. on Systems, Man, and Cybernetics, vol. SMC-13, no. 5, pp. 826-834, Sept/Oct 1983.

A neural network model, called a "neocognitron," for a mechanism of visual pattern recognition was proposed earlier, and the result of computer for a small-scale network was shown. A neocognitron with a larger-scale network is now simulated on a digital computer and is shown to have a great capability for visual pattern recognition. The neocognitron is a multilayer network. A learning-with-a-teacher process is used to reinforce these modifiable synapses in the new model, instead of the learning without a teacher process which was applied to the previous small scale model.

Fukushima, K., "A Hierarchical Neural Network Model for Associative Memory," Biological Cybernetics, vol. 50, pp. 105-113, 1984.

A hierarchical neural network model with feedback interconnections, which has the function of associative mmeory and the ability to recognize patterns, is proposed. The model consists of a hierarchical multi-layered network to which efferent connections are added, so as to make positive feedback loops in pairs with afferent connections. The cell-layer at the initial stage of the network is the input layer which receives the stimulus input and at the same time works as an output layer for associative recall. The deepest layer is the output layer for pattern-recognition. Pattern recognition is performed hierarchically by integrating information by convergivng afferent paths in the network. For the purpose of associative recall, the integrated information is again distributed to lower-order cells by diverging efferent paths. These two operations progress simultaneously in the network. If a fragment of a training pattern is presented to the network which has completed its self-organization, the entire pattern will gradually be recalled in the initial layer. If a stimulus consisting of a number of training patterns becomes predominant in the recalled output after competition between the patterns, and the others disappear. At about the same time when the recalled pattern reaches a steady state in the initial layer, in the deepest layer of the network, a response is elicited from the cell corresponding to the category of the finally-recalled pattern. Once a steady state has been reached, the response of the network is automatically extinguished by in-

hibitory signals from a steadiness-detecting cell. If the same stimulus is still presented after inhibition, a response for another pattern, formerly suppressed, will now appear, because the cells of the network have adaptation characteristics which makes the same response unlikely to recur. Since inhibition occurs repeatedly, the superposed input patterns are recalled one by one in turn.

Geman, S., "Notes on a Self-Organizing Machine," in Parallel Models of Associative Memory, ed. J. Anderson, pp. 237-263, Erlbaum Associates, Hilsdale, N.J., 1981.

This chapter sets forth a design for a system whose purpose is to discover temporal and spatial regularities in a high-dimensional environment. Whereas the goal of this research is the realization of an "intelligent system," the model is based on principles of organization and self-modification widely believed to be in force in the nervous system. The design is of a parallel-processing machine composed of non-linear and highly interconnected devices. As it is presented here, the model is a general one in the sense that it is not dedicated to any particular environment or task. A specific implementation of the model requires the specification of two sets of parameters: the "input parameters" and the "direction primitives." The input primitives represent the information about the environment that is available to the system. The direction primitives define what is "good" and what is "bad" relative to the system.

Goldman, Patricia S. and Walle J. H. Nauta, "Columnar Distribution of Cortico-Cortical Fibers in the Frontal Association, Limbic, and Motor Cortex of the Developing Rhesus Monkey," Brain Research, vol. 122, pp. 393-413, Elsevier/North-Holland Biomedical Press, Amsterdam, The Netherlands, 1977.

Goldman-Rakic, Patricia S., "Development and Plasticity of Primate Frontal Association Cortex," in The Organization of the Cerebral Cortex, ed. Stephen G. Dennis, pp. 69-97, The MIT Press, Cambridge, Mass., 1981.

The neuronal organization and connections of the prefrontal association cortex in primates are beginning to be understood as these features are brought increasingly to light by the use of advanced morphological techniques. In the present report, the corticocortical and corticostiatal fiber systems of the prefrontal cortex have been selected to illustrate the timing, sequence, mode, and mechanisms of development and plasticity of the prefrontal cortex. At early fetal ages, axons of prefrontal corticocortical and corticostriatal systems are distributed in a diffuse manner in their target structures. The remarkable neuronal plasticity exhibited by the output neurons of the developing primate neocortex may be an important biological mechanism for recovery of function following injury to the telencephalon at various stages of prenatal and early postnatal development.

Goldschlager, L. M., "A Computational Theory of Higher Brain Function," Report No. STAN-CS-84-1004, Department of Computer Science, Stanford, CA, April 1984.

The higher functions of the brain are believed to occur in the cortex. This region of the brain is modelled as a memory surface which performs both storage and computation. Concepts are modelled as patterns of activity on the memory surface, and the model explains how these patterns interact with one another to give the computations which the brain performs. The method of interaction can explain the formation of abstract concepts, association of ideas and train of thought. It is shown that creativity, self, consciousness and free will are explainable within the same framework. A theory of sleep is presented which is consistent with the model.

Grossberg, S., "How Does a Brain Build a Cognitive Code?," Psychological Review, vol. 87, no. 1, pp. 1-51, January 1980.

This article indicates how competition between afferent data and learned feedback expectations can stabilize a developing code by buffering committed populations of detectors against continual erosion by new environmental demands. The gating phenomena that result lead to dynamically maintained critical periods, and to attentional phenomena such as overshadowing in the adult. The functional unit of cognitive coding is suggested to be an adaptive resonance, or amplification and prolongation of neural activity, that occurs when afferent data and efferent expectancies reach consensus through a matching process. The resonant state embodies the perceptual event, or attentional focus, and its amplified and sustained activities are capable of driving slow changes of long-term memory. Mismatch between afferent data and efferent expectancies yields a global suppression of activity and triggers a reset of short-term memory, as well as rapid parallel search and hypothesis testing for uncommitted cells. These mechanisms help to explain and predict, as manisfestations of the unified theme of stable code development, positive and negative aftereffects, the McCollough effect, spatial frequency adaptation, monocular rivalry and hysteresis, pattern completion, and Gestalt switching; analgesia, partial reinforcement acquisition effect, conditioned reinforcers, underaroused versus overaroused depression; the contingent negative variation, P300, and ponto-geniculo-occipital waves; olfactory coding, corticogeniculate feedback, matching of proprieceptive and terminal motor maps, and cerebral dominance. The psychophysiological mechanisms that unify these effects are inherently nonlinear and parallel and are inequivalent to the computer, probabilistic, and linear models currently in use.

Grossberg, S., "The Temporal Unfolding and Stability of STM and LTM Patterns," in Competition and Cooperation in Neural Nets, ed. M. Arbib, pp. 295-341, Springer-Verlag, Berlin, 1982.

This article reviews some principles, mechanism, and theorems from my work over the past twenty-five years. I review these results here to illustrate their interconnectedness from a recent perspective, to indicate directions for future work, and to reaffirm an approach to theorizing on problems of mind and brain that is still not fashionable despite growing signs that it needs to become so soon.

Grossberg, S., Studies of Mind and Brain, D. Reidel Publ. Co., Dordrecht, 1982.

Grossberg, S. and G. Stone, "Neural Dynamics of Word Recognition and Recall: Attentional Priming, Learning, and Resonance," Psychological Review, vol. 93, no. 1, pp. 46-74, 1986.

Data and models about recognition and recall of words and nonwords are unified using a real-time network processing theory. Lexical decision and word frequency effect data re analyzed in terms of theoretical concepts that have unified data about development of circular reactions, imitation of novel sounds, the matching of phonetic to articulatory requirements, serial and paired associate verbal learning, free recall, unitization, categorical perception, selective adaptation, auditory contrast, and word superiority effects. The theory, called adaptive resonance theory, arose from an analysis of how a language system self-organizes in real time in response to its complex input environment. Such an approach emphasizes the moment-by-moment dynamical interactions that control language development, learning, and stability. Properties of language performance emerge from an analysis of the system constraints that govern stable language learning. Concepts such as logogens, verification, automatic activation, interactive activation, limited-capacity processing, conscious attention, serial search, processing stages, speed-accuracy trade-off, situational frequency, familiarity, and encoding specificity are revised and developed using this analysis. Concepts such as adaptive resonance, resonant equilibriation of short-term memory, bottom-up adaptive filtering, top-down adaptive template matching, competitive masking field, unitized list representation, temporal order information over item representations, attentional priming, attentional gain control, and list-item error trade-off are applied.

Haken, H., Synergetics, Springer-Verlag, Berlin, 1978.

Hammerstrom, D., "A Connectivity Analysis of Recursive, Auto-Associative Connection Networks," Tech. Report CS/E-06-009, Dept. of Computer Science/Engineering, Oregon Graduate Center, Beaverton, Oregon, August 1986.

The efficient realization using current technology of Very Large Connection Networks (VLCN) with more than a billion connections requires that networks exhibit a high degree of locality of communication. A network exhibits locality of communication, if most of its processing elements connect to other physically adjacent processing elements in any reasonable mapping of the elements onto a planar surface. Real neural networks exhibit significant locality, yet most neural network models have little locality. In this report, the connectivity of a simple class of recursive, auto-associative networks similar to those of Hopfield is analyzed, and several techniques are presented that improve the robustness of the network in the face of sparse, local interconnect structures.

Hammerstrom, D.W., "Dynamic, Decentralized Load Leveling," _EuroMicro_, London, ENGLAND, 1980.

This paper presents a study of methods for the distribution of tasks around a loosely coupled network of distributed processors. This load leveling is performed in a completely decentralized manner by the individual processors. A method of distributing these tasks around the network to maximize total network headway was developed and then simulated using a simplified performance estimation model. In a static environment, where the task charcteristics were fixed in time, the load leveling algorithm performed well. For non-trivial configurations under random strating process assignments, the optimal allocation was found between 50-100% of the time, dpending on the configuration, and the final cost was on the average between 0-20% from optimal. The algorithm is shown to be deadlock free and finds a stable state quickly due to the inherent parallelism in decentralized task movement. It is then shown that with a restricted class of networks, the algorithm is stable in a dynamic system.

Hammerstrom, D. W., D. Maier, and S. Thakkar, "The Cognitive Architecture Project," _Computer Architecture News_, 1986.

Hebb, D.O., _The Organization of Behavior_, John Wiley, New York, 1949.

Hecht-Nielsen, R. and T. Gutschow, _Sensor Processing using Artificial Neural Systems_.

The advent of theoretical neurobiological research has led to the creation of a number of new information processing principles that have been shown to be applicable to the problem of pattern recognition. It is anticipated that

these principles will also provide enhanced or completely new capabilities in other areas of sensor processing as well (such as preprocessing). The current bottleneck in the development of operation systems utilizing these principles is the lack of real-time processing hardware to implement them. To remove this bottleneck, the DARPA ADAPT Program is devleoping hardware designs that will provide real-time processing capabilities for systems as large as 10^6 processing elements with 10^8 interconnects. Such systems will be able to carry out initial operational demonstrations. Development of advanced hardware in subsequent phases of the ADAPT Program will lead to deployable systems.

Hecht-Nielsen, R., "Neural Analog Processing," Proc. of SPIE, vol. 360, pp. 180-189, San Diego, CA, Aug. 24-27, 1982.

This paper presents a bionic approach to pattern classification entitled Neural Analog Processing (NAP). NAP systems are based upon information processing principles of discovered by neural modelers, but are not themselves neural models. To set the stage for a discussion of how NAP systems work, the theory of a particular type of local-in-time template- matching classifier -- the Generalized Nearest Neighbor (GNN) classifier -- for general time-varying patterns (imagery, spectra, tactile signals, etc.) is reviewed. The definition and function of the fundamental NAP structure -- the slab -- is then presented and it is shown that a GNN classifier can, in principle, be implemented using slabs. The embellishments necessary to allow NAP systems to be realized in hardware are then described. Finally, a summary of NAP system characteristics is presented.

Hecht-Nielsen, R., DARPA ADAPT Processor, Mark IV ADAPT Processor, TRW Military Electronics Division, San Diego, CA, 23 Oct 1985.

The DARPA/TRW ADAPT (Adaptive Distributive Analog Processor Technology) Program is exploring architectures and hardware for artificial neural systems. These systems are expected to solve critical processing and data storage problems in the Defense and intelligence communities in areas such as sensor processing, knowledge representation and retrieval, and autonomous system control.

During Phase II of the ADAPT Program (Aug 85 - Nov 86), an experimental processor (the Mark IV ADAPT system) is being built and used to carry out selected demonstrations of ADAPT system capabilities and using real world data. The Mark IV is based upon the design developed during Phase I (Sept 84- Aug 85). Earlier ADAPT systems at TRW (such as the Mark I and Mark II VAX software packages and the smaller-scale Mark III hardware processor) have established the utility of the Mark IV design by means of exact emulation. The CAD tools developed for these earlier systems will be used with the Mark IV.

The purpose of this document is to provide a precise functional specification of the Mark IV experimental ADAPT processor. The primary goal is to allow groups outside of TRW to evaluate the Mark IV's capabilities for use in their experimental programs. If sufficient interest in using the Mark IV exists, it may be possible to configure the processor for convenient remote use through the ARPANET. Another possibility is to establish an on-site center at Rancho Carmel. Naturally, both of these concepts are temporary measures designed to allow early (late 1986/1987) access to the Mark IV. It is anticipated that by the 1988 ADAPT systems of various sizes will be available at most major universities and all relevant Research Centers.

Hecht-Nielsen, R., "Performance Limits of Optical, Electro-Optical, and Electronic Neurocomputers," Artificial Neural Networks, UC San Diego Extension, August 1986.

The performance limits of optical, electro-optical, and electronic artificial neural systems (ANS) processors (also known as neurocomputers) are discussed. After a brief introduction, an overview is provided of the recently revived field of ANS. Next, ANS performance measures are defined and a neurocomputer taxonomy is presented. Finally, the designs and performance limits of the various types of neurocomputers are discussed.

Hecht-Nielsen, R., Artificial Neural Systems Technology, TRW Ranco Carmel AI Center, San Diego, CA, March 10, 1986.

An Artificial Neural System (ANS) is a Hamiltonian dynamical system (or a couple collection of such systems) with an adaptive or selectable energy function that can carry out useful information processing by means of its state response to initial or continuous input. The purpose of this document is to elaborate this terse definition and to familiarize the reader with recent progress in ANS architecture engineering and in the electronic, electro-optical, and optical implementations of such systems.

Hendler, J. A., "Integrating Marker-Passing and Problem Solving," The 7th Annual Conf. of The Cognitive Science Society, pp. 130-139, Irvine, CA., 1985.

In this paper we describe how an efficient underlying mechanism, a parallel spreading activation algorithm, can be used during problem solving. We present this mechanism, chosen due to its demonstrated usefulness for several other cognitive tasks, and show how it can be used to guide plans as they are generated. Examples of how this technique is used are given, and an implementation of such as system, integrating a marker-passer with a problem solver, is described. The paper discusses some of the desiderata in designing such systems and some of the issues that arise.

Some future directions for the work are also described.

Hestenes, D., "How the Brain Works: the next great scientific revolution," Proc. 3rd Workshop on Max. Entropy and Bayesian Methods in Applied Statistics, Aug. 1-4, 1983.

In spite of the enormous complexity of the human brain, there are good reasons to believe that only a few basic principles will be needed to understand how it processes sensory input and controls motor output. In fact, the most important principles may be known already! These principles provide the basis for a definite mathematical theory of learning, memory, and behavior.

Hillis, W. Daniel, "The Connection Machine (Computer Architecture for the New Wave)," A.I. Memo No. 646, MIT Artificial Intelligence Laboratory, Sept 1981.

This paper describes the Connection Machine, a machine for concurrently manipulating knowledge stored in semantic networks. We need the connection memory because conventional serial computers cannot move through such networks fast enough. The connection memory sidesteps the problem by providing processing power proportional to the size of the network. Each node and link in the network has its own simple processor.

Hinton, G., "Shape Representation in Parallel Systems," Proc. IJCAI-81, pp. 1088-1096, Vancouver, B.C., 24-28 August 1981.

There has been a recent revival of interest in parallel systems in which computation is performed by excitatory and inhibitory interactions within a network of relatively simple, neuronlike units. At the early stages of visual processing, individual units can represent hypotheses about how small local fragments of the visual input should be interpreted, and interactions between units can encode knowledge about the constraints between local interpretations. Higher up in the visual system, the representational issues are more complex. This paper considers the difficulties involved in representing shapes in parallel systems, and suggests ways of overcoming them. In doing so, it provides a mechanism for shape perception and visual attention which allows a novel interpretation of the Gestalt slogan that the whole is more than the sum of its parts.

Hinton, Geoffrey E., "Distributed Representations," TR CMU-CS-84-157, Computer Science Dept., Carnegie-Mellon University, Pittsburgh, PA 15213, 1984.

Given a network of simple computing elements and some entities to be represented, the most straightforward scheme is to use one computing element for each entity. This is called a local representation. It is easy to understand and easy to implement because the structure of the physical network mirrors the structure of the knowledge it contains. This report describes a different type of representation that is less familiar and harder to think about than local representations. Each entity is represented by a pattern of activity distributed over many computing elements, and each computing element is involved in representing many different entities. The strength of this more complicated kind of representation does not lie in its notational convenience or its ease of implementation in a conventional computer, but rather in the efficiency with which it makes use of the processing abilities of networks of simple, neuron-like computing elements. Every representational scheme has its good and bad points. Distributed representations are no exception. Some desirable properties like content-addressable memory and automatic generalization arise very naturally from the use of patterns of activity as representations. Other properties, like the ability to temporarily store a large set of arbitrary associations, are much harder to acheive. The best psychological evidence for distributed representation is the degree to which their strengths and weaknesses match those of the human mind.

Hinton, Geoffrey E., Terrence J. Sejnowski, and David H. Ackley, "Boltzmann Machines: Constraint Satisfaction Networks that Learn," Technical Report CMU-CS-84-119, Computer Science Dept., Carnegie-Mellon University, Pittsburgh, PA 15213, May 1984.

The computational power of massively parallel networks of simple processing elements resides in the communication bandwidth provided by the hardware connections between elements. These connections can allow a significant fraction of the knowledge of the system to be applied to an instance of a problem in a very short time. One kind of computation for which massively parallel networks appear to be well suited is large constraint satisfaction searches, but to use the connections efficiently two conditions must be met First, a search technique that is suitable for parallel networks must be found. Second, there must be some way of choosing internal representations which allow the pre-existing hardware connections to be used efficiently for encoding the constraints in the domain being searched. We describe a general parallel search method, based on statistical mechanics, and we show how it leads to a general learning rule for modifying the connection strengths so as to incorporate knowledge about a task domain in an efficient way. We describe some simple examples in which the learning algorithm creates internal representations that are demonstrably the most efficient way of using the pre-existing connectivity structure.

Hogg, T. and B.A. Huberman, "Understanding biological computation: Reliable learning and recognition," Proc. Natl. Acad.

Sci. USA, vol. 81, pp. 6871-6875, November 1984.

We experimentally examine the consequences of the hypothesis that the brain operates reliably, even though individual components may intermittently fail, by computing with dynamical attractors. Specifically, such a mechanism exploits dynamic collective behavior of a system with attractive fixed proints in its phase space. In contrast to the usual methods of reliable computation involving a large number of redundant elements, this technique of self-repair only requires collective computation with a few units, and it is amenable to quantitative investigation. Experiments on parallel computing arrays show that this mechanism leads naturally to rapid self-repair, adaptation to the environment, recognition and discrimination of fuzzy inputs, and conditional learning, properties that are commonly associated with biological computation.

Hopfield, J.J., "Neural networks and physical systems with emergant collective computational abilities," Proc. Nat. Acad. Sci. USA, vol. 79, pp. 2554-2558, April 1982.

Computational properties of use to biological organisms or to the construction of computers can emerge as collective properties of systems having a large number of simple equivalent components (or neurons). The physical meaning of content-addressable memory is described by an appropriate phase space flow of the state of the system. A model of such a system is given, based on aspects of neurobiology but readily adapted to integrated circuits. The collective properties of this model produce a content-addressable memory which correctly yields an entire memory from any given subpart of sufficient size. The algorithm for the time evolution of the state of the system is based on asynchronous parallel processing. Additional emergent collective properties include some capacity for generalization, familiarity recognition, categorization, error correction, and time sequence retention. The collective properties are only weakly sensitive to details of the modeling or the failure of individual devices.

Hopfield, J.J., "Neurons with graded response have collective computational properties like those of two-state neurons," Proc. Natl. Acad. Sci., vol. 81, pp. 3088-3092, May, 1984.

A model for a large network of "neurons" with a graded response (or sigmoid input-output) relation is studied. This deterministic system has collective properties in very close correspondance with the earlier stochastic model based on McCulloch-Pitts neurons. The content-addressable memory and other emergant collective properties of the original model also are present in the graded response model. The ideas that such collective properties are used in biological systems is given credence by the continued presence of such properties for more nearly biological "neurons." Collective analog electrical circuits of the kind described will certainly function. The collective states of the two models have a simple correspondence. The original model will continue to be useful for simulations, because its connection to graded response systems is established. Equations that include the effect of action potentials in the graded response system are also developed.

Hopfield, J.J. and D.W. Tank, "Neural Computation of Decisions in Optimizing Problems," Bio. Cybernetics, vol. 52, pp. 141-152, July 1985.

Hopfield, J.J. and D.W. Tank, "Computing with Neural Circuits: A Model," Science, vol. 233, pp. 625-633, 8 August 1986.

A new conceptual framework and a minimization priniciple together provide an understanding of computation in model neural circuits. The circuits consist of nonlinear graded-response model neurons organized inot networks with effectively symmetric synaptic connections. The neurons represent an approximation to biological neurons in which a simplified set of important computational properties is retained. Complex circuits solving problems similar to those essential in biology can be analyzed and understood without the need to follow the circuit dynamics in detail. Implementation of the model with electronic devices will provide a class of electronic circuits of novel form and function.

Irvine, D.R.F. and D.P. Phillips, "Polysensory "Association" Areas of the Cerebral Cortex," in Cortical Sensory Organization, ed. C.N. Woolsey, vol. 3, pp. 111-156, Humana Press, Clifton, NJ, 1982.

Jones, E.G., "Anatomy of Cerebral Cortex: Columnar Input-Output Organization," in The Organization of the Cerebral Cortex, ed. Stephen G. Dennis, pp. 199-235, The MIT Press, Cambridge, Mass., 1981.

In the cerebral cortex, the search for manageable circuit elements comparable, say, to the mossy fiber-granule cell-Purkinje cell circuit of the cerebellum has been hampered by lack of knowledge regarding the meaning of cortical lamination and a lack of agreement regarding the classes and distribution of interneurons. Recent work indicates that in higher primates, each lamina of the cortex to a alrge extent represents an aggregation of pyramidal-cell somata whose axons project to the same cortical or subcortical site. There is even evidence for a sublaminar organization along these lines in layers III and V.

Kandel, E. R., Cellular Basis of Behavior, W.H. Freeman and Co., 1976.

Kanerva, P., "Self-Propagating Search," CSLI-84-7, Cntr for the Study of Language and Information, Stanford Univ., Palo Alto, CA, March 1984.

At issue is the ability of humans to retrieve information from memory according to content (recalling and recognizing previously encountered objects) and temporal sequence (performing a learned sequence of actions). Retrieval times indicate the direct retrieval of stored information. A mathematical theory of memory is developed. Memory items are represented by n-bit binary words (points of the space {0,1}^n). The unifying principle is that the address space and the datum space are the same. As in the conventional random-access memory of a computer, any stored item can be accessed directly, and sequential retrieval is achieved by storing the memory record as a pointer chain. Accessing of many locations at once accounts for recognition. Three main results are obtained: (1) The properties of neurons allow their use as address decoders for a generalized random-access memory; (2) distributing the storage of an item in a set of locations makes very large address spaces (2^1,000) practical; and (3) structures similar to those suggested by the theory are found in the cerebellum.

Keller, J. M. and D. J. Hunt, "Incorporating Fuzzy Membership Functions into the Perceptron Algorithm," IEEE Trans. on Pattern Analysis and Machine Intelligence, vol. PAMI-7, no. 6, pp. 693-699, Nov. 1985.

The perceptron algorithm, one of the class of gradient descent techniques, has been widely used in pattern recognition to determine linear decision boundaries. While this algorithm is guaranteed to converge to a separating hyperplane if the data are linearly separable, it exhibits erratic behavior if the data are not linearly separable. Fuzzy set theory is introduced into the perceptron algorithm to produce a "fuzzy algorithm" which ameliorates the convergence problem in the nonseparable case. It is shown that the fuzzy perceptron, like its crisp counterpart, converges in the separable case. A method of generating membership functions is developed, and experimental results comparing the crisp to the fuzzy perceptron are presented.

Keller, Robert M. and Frank C.H. Lin, "Simulated Performance of a Reduction-based Multiprocessor," IEEE Computer, vol. 17, no. 7, pp. 71-81, July 1984.

Multiprocessor systems present unique concurrency problems. Rediflow combines disciplined von Neumann processes with a hybrid reduction and dataflow model in an effective packet-switching network.

Kirkpatrick, S., C.D. Gelatt, and M.P. Vecchi, "Optimization by Simulated Annealing," Science, pp. 671-680, 1983.

In this articl we briefly review the central constructs in combinatorial optimization and in statistical mechanics and then develop the similarities between the two fields. We show how the Metropolis algorithm for approximate numerical simulation of behavior of a many-body system at a finite temperature provides a natural tool for bringing the techniques of statistical mechanics to bear on optimization.

Klopf, A. Harry, The Hedonistic Neuron, Hemisphere Publishing Corporation, Washington, 1982.

Kohonen, T., Associative Memory: A System-Theoretical Approach, Springer-Verlag, Berlin, 1978.

Kohonen, T., Self-Organization and Associative Memory, Springer-Verlag, Berlin, 1984.

Konishi, M. and E.I. Knudsen, "A Theory of Neural Auditory Space," in Cortical Sensory Organization, ed. C.N. Woolsey, vol. 3, pp. 219-229, Humana Press, Clifton, NJ, 1982.

Kuffler, S.W. and J.G. Nicholls, From Neuron to Brain, Sinauer Associates, Inc., Sunderland, Massachusetts, 1977.

Lansner, A., "Associative Processing in Brain Theory and Artificial Intelligence," TRITA-NA-8505, Dept. of Numerical Analysis and Computing Science, The Royal Institute of Technology, Stockholm, SWEDEN, 1985.

There is an obvious difference in terms of design and operation between contemporary computers and the brain. Computers are good at arithmetics and logics, biological information processing is vastly superior in tasks involving elements of advanced pattern recognition. The brain is an extremely parallel machine whereas most computers today are not. In this paper some theories of associative brain functions are reviewed and elaborated further on. A distributed and self-organizing (learning) comutational model, an associative net, is proposed. It is intended both as a brain model, albeit abstract and simplified, and as a sketch for a technical device capable of powerful pattern processing. One mechanism behind learning and pattern completion, recognition, and association are described. Issues related to the representation in time and storage capacity of an associative net together with the reliability and computational complexity of associative recall are treated. The relevance of this kind of model for brain theory and the technical potential of associative processing is also discussed.

Lansner, A., "Prototype Extraction and Matching in Associative Nets.," TRITA-NA-8516, Dept. of Numerical Analysis and Computing Science, The Royal Institute of Technology, Stockholm, SWEDEN, 1985.

In this paper, an associative network model is described and learning rules, threshold control strategies and relaxation procedures are discussed. The function properties of the model are examined by means of computer simulation with respect to prototype extraction (clustering) and Hamming distance matching. It is demonstrated that this model is capable of performing prototype extraction from synthetic input (randomly generated training patterns) containing clusters. The storage capacity in terms of the number of prototypes possible to store while maintaining a reliable performance increases with the size of the net as O((N/logN)^2). In the associative recall, a new input pattern activates the prototype which is closest is terms of Hamming distance. The representation of information and the need to incorporate mechanisms for feature extraction into this model is also discussed. Furthermore, we give an example of how a conventional Euclidean clustering problem can be mapped into and solved by an associative net.

Lansner, A. and O. Ekeberg, "Reliability and Speed of Recall in an Associative Network," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. PAMI-7, no. 4, pp. 490-498, July 1985.

Previous investigations of the storage capacity of associative nets have not explicitly considered quantitative aspects of the tradeoff between storage capacity and reconstructive power in these systems. Furthermore, few comparisons have been made between theoretical estimates and experimental results (simulations). In this correspondance, we describe some results recently obtained and relevant to these issues. It is shown that a high storage capacity is possible, without sacrificing reliability in the recall process. Furthermore, an efficient algorithm for retrieval of the information stored is presented, and the speed of recall employing various degrees of parallelism is discussed.

Lansner, Anders and Oerjan Ekeberg, "A Program Package Implementing a System of Interacting Adaptive Associative Networks," TRITA-NA-8302, Dept. of Numerical Analysis, Royal Inst. of Technology, Stockholm, SWEDEN, February 1983.

Recently a new and promising theory of nervous system function has emerged. It is often referred to as the theory of cell assemblies and its origin partly dates back to early work on learning automata. This new theory deserves special interest, since it manages to combine consistency with fundamental facts from experimental neuroscience with a general, constructive and so far largely unexplored approach to information processing. The underlying hypothetical structures are sometimes called adaptive associative networks, or for short associative networks. Although having much in common with neuronal networks, they only retain those "biological" properties judged to be of primary importance for the information processing. An associative network is capable of performing several interesting operations, e.g. learning new patterns that subsequently can be recognized, filtered from noise, completed and associated with each other. A set of interacting associative nets should also be able to allow different types of information (text, image, sound) to interact according to the same principles.

This new theory has also triggered a new interest in applying similar approaches to artificial intelligence and related areas. So far, most work within this framework has been theoretically oriented. However, it seems as if such an approach has to be complemented by experimentally oreiented efforts. To support such work, we have developed a program package suitable for the simulation of a system of interacting associative networks. It is designed with an application independent kernal that might be interfaced to different application dependent parts. The preliminary results using this program are promising, but to learn more we have to attack more realistic, large-scale problems. Such work has recently been initiated.

Legendy, C.R., "On the Scheme by Which the Human Brain Stores Information," Math. Biosciences, vol. 1, pp. 555-597, 1967.

The assumption is explored that the effect of information on the brain is to cause the neurons to form groups. Each group can "ignite" to reverberatory firing in response to a certain kind of stimulation, similar to that which orginally formed it. The groups have to be larer than a certain critical size in order to be able to ignite. The critical size is expressible in terms of the number of branches per neuron and the number of neurons in the brain; it is found to be around 1M neurons. It is shown that about 100M groups can be fitted into the brain before they begin to interact significantly. (One neuron can can participate in many groups.) We assume that the latter number of group is something like the maximum number of elementary "things" that the brain can "know." Attention is devoted to such questions as reliability, interaction, and the mechanism of ignition, and to possible indications that the brain is actually organized in the scheme described.

Leiserson, C.E., "Fat-Trees: Universal Networks for Hardware-Efficient Supercomputing," IEEE Trans. on Computers, vol. C-34, no. 10, pp. 892-901, October 1985.

This paper presents a new class of universal routing networks called fat-trees, which might be used to interconnect the processors of a general-purpose parallel supercomputer. A fat-tree routing network is parameterized not only in the number of processors, but also in the amount of simultaneous communication it can support. Since communication can be scaled independently from number of processors, substantial hardware can be saved over, for example, hypercube-based networks, for such parallel processing applications as finite-element analysis, but without resorting to a special-purpose architecture.

Longuet-Higgens, H.C., D.J. Willshaw, and O.P. Buneman, "Theories of Associative Recall," Quart. Rvw. of Biophysics, vol. 3, no. 2, pp. 223-244, 1970.

The problem of how the brain stores and retrieves information is ultimately an experimental one, and its solution will doubtless call for the combined resources of psychology, physiology and molecular biology. But it is also a problem of great theoretical sophistication; and one of the major tasks confronting brain scientists is the construction of theoretical models which are worthy of, and open to, experimental test. In this review we shall be concerned with the latter aspect of the problem of memory, which has attracted quite a lot of attention in the past few years.

Marr, D., "A Theory of Cerebellar Cortex," Journal Physiology, vol. 202, pp. 437-470, 1969.

A detailed theory of cerebellar cortex is proposed whose consequence is that the cerebellum learns to perform motor skills. Two forms of input-output relation are described, both consistent with the cortical theory. One is suitable for learning movements (actions), and the other for learning to maintain posture and balance (maintenance reflexes).

Marr, D., "A Theory for Cerebral Neocortex," Proc. Roy. Soc. London, vol. 176, pp. 161-234, 1970.

It is proposed that the learning of many tasks by the cerebrum is based on using a very few fundamental techniques for organizing information. It is argued that this is made possible by the prevalence in the world of a particular kind of redundancy, which is characterized by a "Fundamental Hypothesis."

This hypothesis is used to found a theory of the basic operations which, it is proposed, are carried out by the cerebral neocortex. They involve the use of past experience to form so-called "classificatory units" with which to interpret subsequent experience. Such classificatory units are imagined to be created whenever either something occurs frequently in the brain's experience, or enough redundancy appears in the form of clusters of slightly differing inputs.

Marr, D., Vision, W. H. Freeman and Co., 1982.

McClelland, J.L. and D.E. Rumelhart, "An Interactive Activation Model of the Effect of Context in Perception," Psychological Review, vol. 1, 1981.

McDonald, J., E. Rodgers, K. Rose, and A. Steckl, "The Trials of Wafer-Scale Integration," IEEE Spectrum, pp. 32-39, October 84.

Minsky, M. and S. Papert, Perceptrons, The MIT Press, Cambridge, MA, 1969.

Mountcastle, V.B., "An Organizing Principle for Cerebral Function: The Unit Module and the Distributed System," in The Mindful Brain, MIT Press, Cambridge, MA, 1977.

There can be little doubt of the dominating influence of the Darwinian revolution of the mid-nineteenth century upon concepts of the structure and function of the nervous system. The ideas of Spencer and Jackson and Sherrington and the many who followed them were rooted in the evolutionary theory that the brain develops in phylogeny by the successive addition of more cphalad parts. On this theory each new addition or enlargement was accompanied by the elaboration of more complex behavior and, at the same time, imposed a regulation upon more caudal and primitive parts and the presumably more primitive behavior they control. Dissolution of this hierarchy is thought to be revealed by disease or lesions of the brain in humans and by lesions or truncation of the neuraxis in experimental animals. The importance of these ideas can scarecly be exaggerated; for nearly a cnetury they dominated the theory and practice of brain research.

Mountcastle, V.B., R.A. Anderson, and B.C. Motter, "The Influence of Attentive Fixation Upon the Excitability of the Light-Sensitive Neurons of the Posterior Parietal Cortex," The Journal of Neurosciences, vol. 1, no. 11, pp. 1218-1235, November, 1981.

We describe the effect of behavioral state upon the excitability of light-sensitive (LS) neurons of the inferior parietal lobule, area 7a, studied in waking monkeys. The responses of parietal LS neurons to visual stimuli are facilitated during the state of attentive fixation of a target light as compared to their responses to physically and retinotopically identical test stimuli delivered during the eye pauses of alert wakefulness. Seventy percent of the neurons tested (n = 55) showed significant increments in responses in the state of attentive fixation; the median value of the increments was 3.5 times. Only 4 of the 55 cells examined completely showed the reverse relation. Three sets of control experiments were done. The facilitation occurred when the responses evoked during the trials of a reaction task with attentive fixation of a target were compared with those evoked by identical stimuli delivered to the same retinotopic locations at the end of each intertrial interval: the facilitation of attentive fixation is not due to a shift in the general level of arousal. The facilita-

tion occurred when the animal maintained attentive fixation of a spot on the tangent screen without a target light or when an additional light mimicking the target light was presented along with testing stimuli in the state of alert wakefulness without attentive fixation: the facilitation is not produced by a sensory-sensory interaction between target and testing lights. Finally, the facilitation was observed whether or not the test stimuli were behaviorally relevant.
We conclude that the act of attentive fixation exerts a specific and powerful effect upon the excitability of the neural systems linking the retinae and the inferior parietal lobule and that the facilitation plays an important role in visually guided behavior.

Murdock, B.B. Jr., "A Distributed Memory Model For Serial-Order Information," Psychological Review, vol. 90, no. 4, pp. 316-338, 1983.

A theory for the storage and retrieval of item and serial-order information is presented. In the theory, items or events are represented as random vectors. Convolution is used as the storage operation, and correlation is used as the retrieval operation. A distributed memory system is assumed; all information is stored in a common memory vector. In principle, the theory can apply to item recognition, order recognition, probe recall, ordered recall, judgements of recency and frequency, lexical decision, and storage of higher order units such as chunks or propositions. It applies to both accuracy and latency. Performance is predicted from the moments (expectation and variance) of the similarity distributions, and these can be derived from the theory. A canonical version of discrete memory models is outlined as a basis for comparison, and it does not fit some serial-order recognition data as well as the distributed memory model does.

Murdock, B.B., "A Distributed Memory Model for Serial-Order Information," Psychological Review, vol. 90, no. 4, pp. 316-338, 1983.

A theory for the storage and retrieval of item and serial-order information is presented. In the theory, items or events are represented as random vectors. Convolution is used as the storage operation, and correlation is used as the retrieval operation. A distributed memory system is assumed; all information is stored in a common memory vector. In principle, the theory can apply to item recognition, order recognition, probe recall, ordered recall, judgments of recency and frequency, lexical decision, and storage of higher order units such as chunks or propositions. It applies to both accuracy and latency. Performance is predicted from the moments (expectation and variance) of the similarity distributions, and these can be derived from the theory. A canonical version of discrete memory is outlined

as a basis for comparison, and it does not firt some serial-order recognition data as well as the distributed memory model does.

Narendra, K.S. and M.A.L. Thathachar, "Learning Automata - A Survey," IEEE Trans. on Systems, Man and Cybernetics, vol. SMC-4, no. 4, pp. 323-345, July 1974.

Stochastic automata operating in an unknown random environment have been proposed ealier as models of learning. These automata update their action probabilities in accordance with the inputs received from the environment and can improve their own performance during operation. In this context they are referred to as learning automata. A survey of the available results in the area of learning automata has been attempted in this paper. Attention has been focused on the norms of behavior of learning automata, issues in the design of updating schemes, convergance of the action probabilities, and interaction of several automata. Utilization of learning automata in parameter optimization and hypothesis testing is discussed, and potential areas of application are suggested.

Negro, Alberto, The Perfect k-Shuffle: a broad purpose architecture for parallel computation, Oregon State University Department of Computer Science, Corvallis, OR, February 1986.

Norman, D.A., "Copycat Science or Does the mind really work by table look-up?," in Perception and Production of Fluent Speech, ed. R.E. Cole, pp. 381-395, 1980.

Palm, G., Neural Assemblies. An Alternative Approach to Artificial Intelligence, Springer-Verlag, Berlin, 1982.

Palm, G., Local Synaptic Modification Can Lead to Organized Connectivity Patterns in Associative Memory, pp. 229-242, ???.

Without their cerebral cortex people seem to be unable to perform the more interesting types of behavior. On the other hand the cortex is anatomically and electrophysically surprisingly uniform. So all the different important capabilities that have been attributed to different cortical capabilities that have been attributed to different cortical areas seem to be achieved by invoking almost the same machinery. How is this possible? We propose that the cortex is merely a large memory. The organizing priniciple of this memory is simple and local: local correlations in cortical activity are stored by enhancing the local connectivity between the active elements. This prinicple (called Hebb's law) leads to the long-term storage of "preferred" global activity patterns in the cortex (called cell assemblies). Each of these patterns can be activated by any sufficiently large part of it. Viewed as a retrieval procedure in a

memory, this process is known as self-association or as autoassociation in the context of associative memories. As a data storage technique, Hebb's local rule or the corresponding global mechanism of autoassociation turn out to be indeed efficient, even for the cerebral cortex, there is now experimental evidence for variable synaptic connectivities obeying Hebb's law.

Pearlmutter, B.A. and G.E. Hinton, "G-Maximization: an Unsupervised Learning Procedure for Discovering Regularities," in Proc. Neural Networks for Computing, American Institute of Physics, 1986.

Hill climbing is used to maximize an information theoretic measure of the difference between the actual behavior of a unit and the behavior that would be predicted by a statistician who knew the first order statistics of the inputs but believed them to be independent. This causes the unit to detect higher order correlations among its inputs. Initial simulations are presented, and seem encouraging. We describe an extension of the basic ideas which makes it resemble competitive learning and which causes members of a population of these units to differentiate, each extracting different structure from the input.

Pellionisz, A. and R. Llinas, "Space-Time Representation in the Brain. The Cerebellum as a Predictor Space-Time Metric Tensor," Neuroscience, vol. 7, pp. 2949-2970, 1982.

The concept of space-time representation in the brain is redefined using tensor network theory. We make the following suggestions. (a) In order to deal with the external world, the brain embeds the external space-time continuum into a high dimensional internal space. External space-time events are represented within the CNS in overcomplete, inherently oblique, reference frames where space and time information is detected as a continuum over each coordinate axis. (b) The CNS may be seen as imposing a geometry on this internal hyperspace in such manner that neuronal networks transform inputs in a metric tensor-like manner. (c) In order to coordinate movements the cerebellum acts as a predictive motor space-time metric which allows the establishment of coincidences of goal-directed movements of limbs in space-time with external targets.

Perkel, D.H., "Functional role of dendritic spines," J. Physio., Paris, vol. 78, pp. 695-699, 1983.

A number of speculations have been made on the functional role of dendritic spines. Some emphasize that a spine modulates the effect of a chemical synapse on the spine head. Others propose that spines isolate neighboring synapses form each others' effects. Still others suggest that spines play a role in short- or long-term plasticity, while others deny

any functional role for spines at all. This paper brings some quantitative calculations to bear on these questions.

Pollack, J.B. and D.L. Waltz, "Parallel Interpretation of Natural Language," Proc. Intl. Conf. on Fifth Generation Computer Systems, pp. 686-691, Tokyo, 1984.

This is a description of research in developing a natural language processing system with modular knowledge sources but strongly interactive processing. The system offers insights into a variety of linguistic phenomena and allows easy testing of a variety of hypotheses. Language interpretation takes place on an activation network which is dynamically created from input, recent context, and long-term knowledge. Initially ambiguous and unstable, the network settles on a single interpretation, using a parallel, analog relaxation process. We also describe a parallel model for the representation of context and of the priming of concepts. Examples illustrating contextual influence on meaning interpretation and "semantic garden path" sentence processing are included. Furthermore, our model has been designed with the constraints of New Generation Computing in mind, and we describe our first pass at the architectural design of a message-merging communications network which implements the relaxation process.

Posner, M.I., Chronometric Explorations of Mind, L.E. Erlbaum Associates, Hillsdale, NJ, 1978.

Preparata, Franco P. and Jean Vuillemin, "The Cube-Connected Cycles: A Versatile Nework for Parallel Computation," CACM, pp. 300-309, May 1981.

Rakic, Pasco, "Development Events Leading to Laminar and Areal Organization of the Neocortex," in The Organization of the Cerebral Cortex, ed. Stephen G. Dennis, pp. 7-28, MIT Press, Cambridge, Mass., 1981.

Evidence from a series of studies on the development of the primate neocortex indicates that the tangential (horizontal) coordinates of cortical neuron are determined by the relative position of its precursor cell in the proliferative zones lining the cerebral ventricle, while the radial (vertical) position is determined by the time of its genesis and rate of migration. Thus, the development of topography and/or modality are determined by spatial parameters, while the hierarchical organization of neurons within each radial columnar unit is determined by temporal factors. The development of thalamocortical connections can be divided into three phases. In the first phase, thalamic fibers grow into the intermediate zone of the cerebral wall and accumulate below the developing cortical plate; in the second phase, fibers enter the appropriate cortical region but distribute in a relatively diffuse manner; finally, in the third phase, fibers sort out into specific terminal fields. The three phase are separated in time and are probably regulated by different mechanisms.

Reggia, J.A., "Virtual Lateral Inhibition in Parallel Activation Models of Associative Memory," Proc. 9th IJCAI, 1985.

This paper describes a new theory of how spreading activation may occur in associative memory models formulated as parallel activation networks. The theory postulates that competition for activation by nodes/concepts in a network is a fundamental principle of memory retrieval. Using only excitatory connections between concepts, a specific implementation of this model is able to demonstrate "virtual lateral inhibition" between competitors and other interesting behaviors that have required use of explicit inhibitory connections in the past.

Reilly, D. L., L. N. Cooper, and C. Elbaum, "A Neural Model for Category Learning," Biol. Cybern., vol. 45, pp. 35-41, 1982.

We present a general neural model for supervised learning of pattern categories which can resolve pattern classes separated by nonlinear, essentially arbitrary boundaries. The concept of a pattern class develops from storing in memory a limited number of class elements (prototypes). Associated with each prototype is a modifiable scalar weighting factor which effectively defines the threshold for categorization of an input with the class of the given prototype. Learning involves (1) commitment of prototypes to memory and (2) adjustment of the various lambda factors to eliminate classification errors. In tests, the model ably defined classsification boundaries that largely separated complicated pattern regions. We discuss the role which divisive inhibition might play in a possible implementation of the model by a network of neurons.

Roberts, Charles S., "Partial-Match Retrieval via the Method of Superimposed Codes," Proc. of the IEEE, vol. 67, no. 12, pp. 1624-1642, IEEE, December 1979.

Rosenblatt, F., Principles of Perceptrons, Spartan, Washington, D.C., 1962.

Rumelhart, D.E. and D. Zipser, "Feature Discovery by Competitive Learning," Cognitive Science, vol. 9, pp. 75-112, 1985.

This paper reports the results of our studies with an unsupervised learning paradigm which we have called "Cooperative Learning." We have examined competitive learning using both computer simulation and formal analysis and have found that when it is applied to parallel networks of neuron-like elements, many potentially useful learning tasks can be accomplished. We were attracted to competitive learning because it seems to provide a way to discover the salient, general features which can be used to classify a set of patterns. We show how a very simply competitive mechanism can discover a set of feature detectors which capture important aspects of the set of stimulus input patterns. We also show how these feature detectors can form the basis of a multilayer system that can serve to learn categorizations of stimulus sets which are not linearly separable. We show how the use of correlated stimuli can serve as a kind of "teaching" input to the system to allow the development of feature detectors which would not develop otherwise. Although we find the competitive learning mechanism a very interesting and powerful learning priniciple, we do not, of course, imagine that it is the only learning principle. Competitive learning is an essentially nonassociative statistical learning scheme. We certainly imagine that other kinds of learning mechanisms will be involved in the building of associations among patterns of activation in a more complete neural network. We offer this analysis of these competitive learning mechanisms to further our understanding of how simple adaptive networks can discover features important in the description of the stimulus environment in which the system finds itself.

Rumelhart, D.E. and J.L. McClelland, Parallel Distributed Processing: Explorations in the Microstructure of Cognition, 1 and 2, Bradford Books/MIT Press, Cambridge, MA, 1985.

Rumelhart, D.E., G.E. Hinton, and R.J. Williams, "Learning Internal Representations by Error Propagation," ICS Report 8506, Institute for Cognitive Science, La Jolla, CA, September 1985.

Seitz, C. L., "Concurrent VLSI Architectures," IEEE Transactions on Computers, vol. C33, pp. 1247-1265, December 1984.

Seitz, C. L., "The Cosmic Cube," Communications of the ACM, vol. 28, pp. 23-33, January 1985.

Sejnowski, T.J. and C.R. Rosenberg, "NETtalk: A Parallel Network that Learns to Read Aloud," JHU/EECS-86/01, The Johns Hopkins Univ. Elec. Eng. and Comp. Sci. Tech. Rpt, 1986.

Unrestricted English text can be converted to speech by applying phonological rules and handling exceptions with a look-up table. However, this approach is highly labor intensive since each entry and rule must be hand-crafted. NETtalk is an alternative approach that is based on an automated learning procedure for a parallel network of deterministic processing units. After training on a corpus of informal continuous speech, it achieves good performance and generalizes to novel words. The distributed internal representations of the phonological regularities discovered by the network are damage resistant.

Selman, Bart, "Rule-Based Processing in a Connectionist System for Natural Language Understanding," CSRI Technical Report No. 168, Computer Systems Research Institute, University of Toronto, Toronto, Ont., Canada, 1985.

In this thesis we give a connectionist model for natural language processing. In contrast with previously proposed schemes, this scheme handles traditionally sequential rule-based processing in a general manner in the network. Another difference is the use of a computational scheme similar to the one used in the Boltzmann machine. This allows us to formulate general rules for setting of weights and thresholds. We give a detailed description of a parsing system based on context-free grammar rules. Using simulated annealing, we show that at low temperatures the time average of the visited states at thermal equilibrium represents the correct parse of the input sequence. The system is built from a small set of connectionist primitives that represent the grammar rules. These primitives are linked together using pairs of computing units that behave like discrete switches. These units are used as binders between concepts. They can be linked in such a way that individual rules can be selected from a collection of rules, and are very useful in the construction of connectionist schemes for any form of rule-based processing. We also consider two variations on the formalism of the Boltzmann machine. First we show how the use of +1 and -1 as output values for the computing units facilitates the stting of thresholds. Secondly we introduce an alternative energy function. Using this function we are able to choose a set of weights, such that only the state of the network that corresponds to the correct parse of the input sentence has the lowest possible energy.

Sequin, C. H, "Doubly Twisted Torus Networks for VLSI Processor Array," ACM - SIGARCH Newsletter, pp. 471-480, June 81.

Shastri, L. and J.A. Feldman, "Semantic Networks and Neural Nets," TR 133, Dept. of Computer Science, Univ. of Rochester, Rochester, NY, June, 1984.

Connected networks of nodes representing conceptual knowledge are widely employed in aritificial intelligence and cognitive science. This report describes a direct way of realizing these semantic networks with neuron-like computing units. The proposed framework appears to offer several advantages over previous work. It obviates the need for a centralized knowledge base interpreter, thereby partially solving the problem of computational information. The model employs a class of inference that may be characterized as working with a set of competing hypotheses, gathering evidence for each hypothesis and selecting the best among these. The resulting system has been simulated and is capable of supporting existing semantic network applications dealing with problems of recognition and recall in a uniform manner.

Shastri, L. and J. Feldman, "Evidential Reasoning in Semantic Networks: A Formal Theory," Proc. 9th IJCAI, pp. 467-474, 1985.

This paper presents an evidential approach to knowledge representation and inference wherein the principle of maximum entropy is applied to deal with uncertainty and incompleteness. It focuses ona restricted representation language - similar in expressive power to semantic network formalisms, and develops a formal theory of evidential inheritance within this language. The theory applies to a limited, but we think interesting, class of inheritance problems including those that involve exceptions and multiple inheritance hierarchies. The language and the accompanying evidential inference structure provide a natural treatment of defaults and conflicting information. The evidence combination rule proposed in this paper is incremental, commutative and features of the Dempster-Shafer evidence combination rule in the context of the problems addressed in this paper. The resulting theory can be implemented as a highly parallel (connectionist) network made up of active elements that can solove iheritance problems in time proportional to the depth of the conceptual hierarchy.

Shephard, G., The Synaptic Organization of the Brain, Oxford University Press, New York, 1979.

Shepherd, G.M., R.K. Brayton, J.P. Miller, I. Segev, J. Rinzel, and W. Rall, "Signal Enhancement In Distal Cortical Dendrites By Means Of Interactions Between Active Dendritic Spines," Proc. Natl. Acad. Sci. USA, vol. 82, pp. 2192-2195, April 1985.

Pyramidal neurons in the cerebral cortex characteristically give rise to an apical dendrite, whose distal dendritic branches in layer I are covered with spines. These spines are known to be sites of synaptic connections, but the physiological properties of the spines and the functional significance of their responses are still largely unknown. The main function attributed thus far to these synaptic responses, situated at a great distance from the neuronal cell body, is slow background modulation of impulse output in the axon. In pursuing computer simulation analysis of electrical properties of dendrites, we have obtained results suggesting interactions between distal dendritic spines. If the heads of dendritic spines have excitable membrane properties, the spread of current from one or several spines could bring adjacent spines to their thresholds for impulse generation. This could give rise to a sequence of spine head action potentials, representing a saltatory propagation, from one or more excitable spine heads to nearby excitable spine heads, in the distal dendritic branches. Both the amplification due to several spine action potentials and the possibility of propagation into more proximal branches would increase the efficacy of distal synaptic inputs. Because of nonlinear dependence upon several modifiable parameters (such as spine stem resistance and membrane excitability) and upon the spatio-temporal pattern of synaptic input, such contingent synaptic enhancement would be particularly relevant to cortical functions underlying information processing and to plasticity underlying learning and memory.

Sivilotti, M., M. Emerling, and C. Mead, A Novel Associative Memory Implemented Using Collective Computation, pp. 329-342, 1985 Chapel Hill Conference on VLSI, May 1985.

A radically new type of associative memory, the ASSOCMEM, has been implemented in VLSI and tested. Analog circuit techniques are used to construct a network that evolves towards fully restored (digital) fixed-points that are the memories of the system. Association occurs on the whole source word, each bit of which may assume a continuous analog value. The network does not require the distinction of a search key from a data field in either the source or target words. A key may be dynamically defined by differentially weighting any subset of the source word. The key need not be eact; the system will evolve to the closest memory. In the case when the key is the whole input word, the system may be thought of as performing error correction.

Small, S.L., L. Shastri, M.L. Brucks, S.G. Kaufman, G.W. Cottrell, and S. Addanki, "ISCON: A Network Construction Aid and Simulation for Connectionist Models," TR 109, Dept. of Computer Science, University of Rochester, Rochester, New York 14627, April 1983.

This paper describes the organization and use of an interactive computer program called ISCON, which aids in the development of connectionist models of cognitive processes and neural systems. ISCON was developed during a six week period in Spring 1982 as a prototypic implementation both to help determine the requirements for a useful simulation system and to use as a short-term measure to facilitate research into the construction of connectionist models in various domains, especially language comprehension, motor control, and general control structures. Such models can be explicitly neurobiological in nature, in which connectionist computing units represent neurons, or cognitive in nature, in which these units constitute an architectural substrate for higher-level processes. ISCON development has stopped with the implementation of the initial system, and new features are being added as they are suggested by ongoing experimentation with existing versions. The design aid and simulation systems are implemented in Franz Lisp and the graphic subsystem is implemented in C.

Snyder, Lawrence, "Introduction to the Configurable, Highly Parallel Computer," IEEE Computer, vol. 15, no. 1, pp. 47-56, Jan. 1982.

Architectures for this computer are built around a lattice of programmable switches and data paths that allows arbitrary connection patterns - an approach that preserves locality.

Squire, L. R., "Mechanisms of Memory," Science, vol. 232, pp. 1612-1619, June 27, 1986.

Recent studies of animals with complex nervous systems, including humans and other primates, have improved our understanding of how the brain accomplishes learning and memory. Major themes of recent work include the locus of memory storage, the taxonomy of memory, the distinction between declarative and procedural knowledge, and the question of how memory changes with time, that is, the concepts of forgetting and consolidation. An important recent advance is the development of an animal model of human amnesia in the monkey. The animal model, together with newly available neuropathological information from a well-studied human patient, has permitted the identification of brain structures and connections involved in memory functions.

Stone, H. S., "Parallel Processing with the Perfect Shuffle," IEEE Transactions on Computers, vol. C-20, No 2, pp. 153-161, Febuary 81.

Suga, N., "Functional Organization of the Auditory Cortex," in Cortical Sensory Organization, ed. C.N. Woolsey, vol. 3, pp. 157-218, Humana Press, Clifton, NJ, 1982.

Szentagothai, J. and M. Arbib, Conceptual Models of Neural Organization, MIT Press, Cambridge, MA., 1975.

Tanenbaum, Andrew S., Computer Networks, PRHALL, 1981.

A general text covering computer networks via the ISO model, shows alternates for each level and discusses some topology problems.

Touretzky, D. S. and G. E. Hinton, "Symbols Among the Neurons: Details of a Connectionist Inference Architecture," Proc. 9th IJCAI, 1985.

Pattern matching and variable binding are easily implemented in conventional computer architectures, but not necessarily in all architectures. In a distributed neural network architecture each symbol is represented by activity in many units and each unit contributes to the representation of many symbols. Manipulating symbols using this type of distributed representation is not as easy as with a local representation where each unit denotes one symbol, but there is evidence that the distributed approach is the one chosen by nature. We describe a working implementation of a production system interpreter in a neural network using distributed representations for both symbols and rules. The research provides a detailed account of two important symbolic reasoning operations, pattern matching and variable binding, as emergent properties of collections of neuron-like elements. The success of our production system implementation goes some way towards answering a common criticism of connectionist theories: that they aren't powerful enough to do symbolic reasoning.

Trehub, A., "Neuronal Models for Cognitive Processes: Networks for Learning, Perception, and Imagination," Journal of Theoretical Biology, vol. 65, pp. 141-169, 1977.

Model neuronal mechanisms are described in detail and organized into a functional system which can perform a wide variety of high level visualization tasks. A conservative estimate of the number of nerve cells required to realize the complete system in the human brain indicates that it can be accommodated within the visual areas of a single cortical hemisphere leaving a "surplus" capacity of 6.62x10^8 visual neurons in the occupied hemisphere.

Ullman, J.D., "Flux, Sorting, and Supercomputer Organization for AI Applications," Journal of Parallel and Distributed Computing, vol. 1, pp. 133-151, 1984.

A central issue for the design of massively parallel computers is their ability to sort. We consider organizations that are suitable for fast sorting, both those that use point-to-point connections and those that connect processors with multipoint nets. We show that for fast sorting and minimal area, nets must connect at least sqrt(n) nodes each, if the network has n nodes. We then discuss some of the ways that fast-sorting networks can be used to speed up other processes, such as combinatorial search.

Valiant, L.G., "Learning Disjunctions of Conjunctions," Proc. 9th IJCAI, pp. 560-566, 1985.

The question of whether concepts expressible as disjunctions of conjunctions can be learned from examples in polynomial time is investigated. Positive results are shown for significant subclasses that allow not only propositional predicates but also some relations. The algorithms are extended so as to be provably tolerant to a certain quantifiable error rate in the examples data. It is further shown that under certain restrictions on these subclasses the learning algorithms are well suited to implementation on neural networks of threshold elements. The possible importance of disjunctions of conjunctions as a knowledge representation stems from the observations that on the one hand humans appear to like using it, and, on the other, that there is circumstantial evidence that significantly larger classes may not be learnable in polynomial time. An NP-completeness result corroborating the latter is also presented.

Wickelgren, W.A., "Chunking and Consolidation: A Theoretical Synthesis of Semantic Networks, Configuring in Conditioning, S-R Versus Cognitive Learning, Normal Forgetting, the Amnesic Syndrome, and the Hippocampal Arousal System," Psychological Review, vol. 86, no. 1, pp. 44-60, 1979.

Horizontal versus vertical associative memory concepts are defined. Vertical associative memory involves chunking: the specification of new (previously free) nodes to represent combinations of old (bound) nodes. Chunking is the basis of semantic memory, configuring in conditioning, and cognitive (as opposed to stimulus-response) learning. The cortex has the capacity for chunking, but the hippocampal (limbic) arousal system plays a critical role in this chunking process by differentially priming (paritally activating) free, as opposed to bound, neurons. Binding a neuron produces negatively accelerated repression of its connections to the hippocampal input and thus retarding forgetting. Disruption of the hippocampal arousal system produces the amnesic syndrome of an inability to do new chunking (cognitive learning) - anterograde amnesia - and an inability to retrieve recently specified chunks - retrograde amnesia.

Wickelgren, Wayne A., Cognitive Psychology, Prentice-Hall, Englewood Cliffs, New Jersey, 1979.

Yasuura, H., "On Parallel Computational Complexity of Unification," Proc. Intl. Conf. on Fifth Generation Computer Systems, pp. 235-243, Tokyo, 1984.

The computational complexity of unification in the first-order logic under a parallel computation scheme is discussed. A parallel unification algorithm is presented for a combinational logic circuit model and its complexity is analyzed. In order to discuss the lower bound of parallel time comlexity of unification, we show that unification is one of the hardest problems in the class of problems resolved by polynomial size circuits. Namely, it seems to be difficult to design very efficient parallel algorithms for unification.