Neural Control of Sensory Acquisition: The Vestibulo-Ocular Reflex
Michael Paulin, Mark Nelson, James Bower
We present a new hypothesis that the cerebellum plays a key role in ac(cid:173) tively controlling the acquisition of sensory infonnation by the nervous system . In this paper we explore this idea by examining the function of a simple c erebellar-related behavior, the vestibula-ocular reflex or VOR, in which eye movements are generated to minimize image slip on the retina during r apid head movements. Considering this system from the point of view of st atistical estimation theory, our results sug(cid:173) gest that the transfer fu nction of the VOR, often regarded as a static or slowly modifiable feature o f the system, should actually be continu(cid:173) ously and rapidly changed during head movements. We further suggest that these changes are under the dir ect control of the cerebellar cortex and propose experiments to test this hypot hesis.
****

Modeling the Olfactory Bulb - Coupled Nonlinear Oscillators
Zhaoping Li, John J. Hopfield
The olfactory bulb of mammals aids in the discrimination of odors. A ma thematical model based on the bulbar anatomy and electrophysiology is described. Simulations produce a 35-60 Hz modulated activity coherent acro ss the bulb, mimicing the observed field potentials. The decision states (for the odor information) here can be thought of as stable cycles, ra ther than point stable states typical of simpler neuro-computing models. Analysis and simulations show that a group of coupled non-linear oscillator s are responsible for the oscillatory activities determined by the odor in(cid: 173) put, and that the bulb, with appropriate inputs from higher centers, can enhance or suppress the sensitivity to partiCUlar odors. The model provi des a framework in which to understand the transform between odor input and th e bulbar output to olfactory cortex.
****

Efficient Parallel Learning Algorithms for Neural Networks
Alan Kramer, Alberto Sangiovanni-Vincentelli
Parallelizable optimization techniques are applied to the problem of learning i n feedforward neural networks. In addition to having supe(cid:173) rior converge nce properties, optimization techniques such as the Polak(cid:173) Ribiere metho d are also significantly more efficient than the Back(cid:173) propagation algor ithm. These results are based on experiments per(cid:173) formed on small boolea n learning problems and the noisy real-valued learning problem of hand-written character recognition.
****

The Boltzmann Perceptron Network: A Multi-Layered Feed-Forward Network Equivalen t to the Boltzmann Machine
Eyal Yair, Allen Gersho
The concept of the stochastic Boltzmann machine (BM) is auractive for decision making and pattern classification purposes since the probabili ty of attaining the network states is a function of the network energy. Hence, the probability of attaining particular energy minima may be associate d with the probabilities of making certain decisions (or classification s). However, because of its stochastic nature, the complexity of the BM is f airly high and therefore such networks are not very likely to be used in practice. In this paper we suggest a way to alleviate this draw back by converting the sto(cid:173) chastic BM into a deterministic netw ork which we call the Boltzmann Per(cid:173) ceptron Network (BPN). The BPN is functionally equivalent to the BM but has a feed-forward structu re and low complexity. No annealing is required. The conditions under w hich such a convmion is feasible are given. A learning algorithm for the BPN based on the conjugate gradient method is also provided which i s somewhat akin to the backpropagation algorithm.
****

Neural Networks that Learn to Discriminate Similar Kanji Characters

Yoshihiro Mori, Kazuhiko Yokosawa
is
****

Computer Modeling of Associative Learning
Daniel Alkon, Francis Quek, Thomas Vogl
Requests for name changes in the electronic proceedings will be accepted with no questions asked. However name changes may cause bibliographic tracking issues. Authors are asked to consider this carefully and discuss it with their co-authors prior to requesting a name change in the electronic proceedings.
****

Links Between Markov Models and Multilayer Perceptrons
Hervé Bourlard, C. J. Wellekens
Hidden Markov models are widely used for automatic speech recog(cid:173) nition. They inherently incorporate the sequential character of the speech signal and are statistically trained. However, the a-priori choice of the model topology limits their flexibility. Another draw(cid:173) back of these models is their weak discriminating power. Multilayer perceptrons are now promising tools in the connectionist approach for classification problems and have already been successfully tested on speech recognition problems. However, the sequential nature of the speech signal remains difficult to handle in that kind of ma(cid:173) chine. In this paper, a discriminant hidden Markov model is de(cid:173) fined and it is shown how a particular multilayer perceptron with contextual and extra feedback input units can be considered as a general form of such Markov models.
****

Skeletonization: A Technique for Trimming the Fat from a Network via Relevance Assessment
Michael C. Mozer, Paul Smolensky
This paper proposes a means of using the knowledge in a network to determine the functionality or relevance of individual units, both for the purpose of understanding the network's behavior and improving its performance. The basic idea is to iteratively train the network to a cer(cid:173) tain performance criterion, compute a measure of relevance that identi(cid:173) fies which input or hidden units are most critical to performance, and automatically trim the least relevant units. This skeletonization tech(cid:173) nique can be used to simplify networks by eliminating units that con(cid:173) vey redundant information; to improve learning performance by first learning with spare hidden units and then trimming the unnecessary ones away, thereby constraining generalization; and to understand the behavior of networks in terms of minimal "rules."
****

A Passive Shared Element Analog Electrical Cochlea
David Feld, Joe Eisenberg, Edwin Lewis
We present a simplified model of the micromechanics of the human cochlea, realized with electrical elements. Simulation of the model shows that it retains four signal processing features whose importance we argue on the basis of engineering logic and evolutionary evidence. Furthermore, just as the cochlea does, the model achieves massively parallel signal processing in a structurally economic way, by means of shared elements. By extracting what we believe are the five essential features of the cochlea, we hope to design a useful front-end filter to process acoustic images and to obtain a better understanding of the auditory system.
****

Speech Production Using A Neural Network with a Cooperative Learning Mechanism
Mitsuo Komura, Akio Tanaka
We propose a new neural network model and its learning algorithm. The proposed neural network consists of four layers - input, hidden, output and final output layers. The hidden and output layers are multiple. Using the proposed SICL(Spread Pattern Information and Cooperative Learning) algorithm, it is possible to learn analog data accurately and to obtain smooth outputs. Using this neural network, we have developed a speech production system consisting of a phonemic symbol production subsystem and a speech paramete

r production subsystem. We have succeeded in producing natural speech waves with high accuracy.
****

Associative Learning via Inhibitory Search
David Ackley
ALVIS is a reinforcement-based connectionist architecture that learns associative maps in continuous multidimensional environ(cid:173) ments. The discovered locations of positive and negative rein(cid:173) forcements are recorded in "do be" and "don't be" subnetworks, respectively. The outputs of the subnetworks relevant to the cur(cid:173) rent goal are combined and compared with the current location to produce an error vector. This vector is backpropagated through a motor-perceptual mapping network. to produce an action vec(cid:173) tor that leads the system towards do-be locations and away from don 't-be locations. AL VIS is demonstrated with a simulated robot posed a target-seeking task.
****

Performance of a Stochastic Learning Microchip
Joshua Alspector, Bhusan Gupta, Robert Allen
Requests for name changes in the electronic proceedings will be accepted with no questions asked. However name changes may cause bibliographic tracking issues. Authors are asked to consider this carefully and discuss it with their co-authors prior to requesting a name change in the electronic proceedings.
****

Song Learning in Birds
M. Konishi
Requests for name changes in the electronic proceedings will be accepted with no questions asked. However name changes may cause bibliographic tracking issues. Authors are asked to consider this carefully and discuss it with their co-authors prior to requesting a name change in the electronic proceedings.
****

Modeling Small Oscillating Biological Networks in Analog VLSI
Sylvie Ryckebusch, James Bower, Carver Mead
We have used analog VLSI technology to model a class of small os(cid:173) cillating biological neural circuits known as central pattern gener(cid:173) ators (CPG). These circuits generate rhythmic patterns of activity which drive locomotor behaviour in the animal. We have designed, fabricated, and tested a model neuron circuit which relies on many of the same mechanisms as a biological central pattern generator neuron, such as delays and internal feedback. We show that this neuron can be used to build several small circuits based on known biological CPG circuits, and that these circuits produce patterns of output which are very similar to the observed biological patterns.
****

Comparing Biases for Minimal Network Construction with Back-Propagation
Stephen Hanson, Lorien Pratt
learning
****

What Size Net Gives Valid Generalization?
Eric Baum, David Haussler
We address the question of when a network can be expected to generalize from m random training examples chosen from some ar(cid:173) bitrary probability distribution, assuming that future test examples are drawn from the same distribution. Among our results are the following bounds on appropriate sample vs. network size. Assume o < £ $ 1/8. We show that if m > O(~ log~) random exam(cid:173) ples can be loaded on a feedforward network of linear threshold functions with N nodes and W weights, so that at least a fraction 1 - t of the examples are correctly classified, then one has confi(cid:173) dence approaching certainty that the network will correctly classify a fraction 1 - £ of future test examples drawn from the same dis(cid:173) tribution. Conversely, for fully-connected feedforward nets with one hidden layer, any learning algorithm using fewer than O('!') random

training examples will, for some distributions of examples consistent with an appropriate weight choice, fail at least some fixed fraction of the time to find a weight choice that will correctly classify more than a 1 - £ fraction of the future test examples.
****

A Low-Power CMOS Circuit Which Emulates Temporal Electrical Properties of Neuron s
Jack Meador, Clint Cole
This paper describes a CMOS artificial neuron. The circuit is directly derived from the voltage-gated channel model of neural membrane, has low power dissipation, and small layout geometry. The principal motivations behind this work include a desire for high performance, more accurate neuron emulation, and the need for higher density in practical neural network implementations.
****

Linear Learning: Landscapes and Algorithms
Pierre Baldi
Requests for name changes in the electronic proceedings will be accepted with no questions asked. However name changes may cause bibliographic tracking issues. Authors are asked to consider this carefully and discuss it with their co-auth ors prior to requesting a name change in the electronic proceedings.
****

Applications of Error Back-Propagation to Phonetic Classification
Hong Leung, Victor W. Zue
This paper is concerced with the use of error back-propagation in phonetic clas sification. Our objective is to investigate the ba(cid:173) sic characteristics of back-propagation, and study how the frame(cid:173) work of multi-layer percep trons can be exploited in phonetic recog(cid:173) nition. We explore issues such as integration of heterogeneous sources of information, conditioll~ that can a ffect performance of phonetic classification, internal representations, compari sons with traditional pattern classification techniques, comparisons of differ(cid:173) ent error metrics, and initialization of the network. Our investiga(cid :173) tion is performed within a set of experiments that attempts to rec(cid:173) ognize the 16 vowels in American English independent of speaker. Our results are comparable to human performance.
****

An Analog Self-Organizing Neural Network Chip
James Mann, Sheldon Gilbert
A design for a fully analog version of a self-organizing feature map neural net work has been completed. Several parts of this design are in fabrication. The f eature map algorithm was modified to accommodate circuit solutions to the vario us computations required. Performance effects were measured by simulating the d esign as part of a frontend for a speech recognition system. Circuits are inclu ded to implement both activation computations and weight adaption 'or learning. External access to the analog weight values is provided to facilitate weight i nitialization, testing and static storage. This fully analog implementation req uires an order of magnitude less area than a comparable digital/analog hybrid v ersion developed earlier.
****

A Bifurcation Theory Approach to the Programming of Periodic Attractors in Netwo rk Models of Olfactory Cortex
Bill Baird
A new learning algorithm for the storage of static and periodic attra ctors in biologically inspired recurrent analog neural networks is intr oduced. For a network of n nodes, n static or n/2 periodic attractor s may be stored. The algorithm allows programming of the network vect or field indepen(cid:173) dent of the patterns to be stored. Stability of patterns, basin geometry, and rates of convergence may be controll ed. For orthonormal patterns, the l~grning operation reduces to a kind of periodic outer product rule that allows local, additive, commutativ

e, incremental learning. Standing or traveling wave cycles may be stor ed to mimic the kind of oscillating spatial patterns that appear in the neural activity of the olfactory bulb and prepyriform cortex durin g inspiration and suffice, in the bulb, to predict the pattern recogn ition behavior of rabbits in classical conditioning ex(cid:173) periments . These attractors arise, during simulat(cid:173) ed inspiration, through a multiple Hopf bifurca(cid:173) tion, which can act as a critical " decision pOint" for their selection by a very small input pattern.
****

Heterogeneous Neural Networks for Adaptive Behavior in Dynamic Environments
Randall Beer, Hillel Chiel, Leon S. Sterling
Research in artificial neural networks has genera1ly emphasized homogeneous ar chitectures. In contrast, the nervous systems of natural animals exhibit great heterogeneity in both their elements and patterns of interconnection. This hete rogeneity is crucial to the flexible generation of behavior which is essential for survival in a complex, dynamic environment. It may also provide powerful in sights into the design of artificial neural networks. In this paper, we descri be a heterogeneous neural network for controlling the wa1king of a simulated insect. This controller is inspired by the neuroethological It exhibits a and neurobiological literature on insect locomotion. variety of statically stable g aits at different speeds simply by varying the tonic activity of a single cell. It can also adapt to perturbations as a natural consequence of its design.
****

Programmable Analog Pulse-Firing Neural Networks
Alister Hamilton, Alan Murray, Lionel Tarassenko
We describe pulse - stream firing integrated circuits that imple(cid:173) ment asynchronous analog neural networks. Synaptic weights are stored dynam ically, and weighting uses time-division of the neural pulses from a signalling neuron to a receiving neuron. MOS transistors in their "ON" state act as variable resistors to control a capacitive discharge, a nd time-division is thus achieved by a small synapse circuit cell. Th e VLSI chip set design uses 2.5J.1.m CMOS technology.
****

An Analog VLSI Chip for Thin-Plate Surface Interpolation
John Harris
Reconstructing a surface from sparse sensory data is a well-known problem iIi c omputer vision. This paper describes an experimental analog VLSI chip for smoot h surface interpolation from sparse depth data. An eight-node ID network was de signed in 3J.lm CMOS and successfully tested. The network minimizes a second-or der or "thin(cid:173) plate" energy of the surface. The circuit directly impleme nts the cou(cid:173) pled depth/slope model of surface reconstruction (Harris, 1 987). In addition, this chip can provide Gaussian-like smoothing of images.
****

Simulation and Measurement of the Electric Fields Generated by Weakly Electric F ish
Brian Rasnow, Christopher Assad, Mark Nelson, James Bower
The weakly electric fish, Gnathonemus peters;;, explores its environment by gene r(cid:173) ating pulsed elecbic fields and detecting small pertwbations in the fields resulting from nearby objects. Accordingly, the fISh detects and disc riminates objects on the basis of a sequence of elecbic "images" whose tempora l and spatial properties depend on the tim(cid:173) ing of the fish's electric organ discharge and its body position relative to objects in its en(cid:173) vi ronmenl We are interested in investigating how these fish utilize timing and bo dy-po(cid:173) sition during exploration to aid in object discrimination. We ha ve developed a fmite-ele(cid:173) ment simulation of the fish's self-generated e lectric fields so as to reconstruct the elec(cid:173) trosensory consequence s of body position and electric organ discharge timing in the fish. This paper describes this finite-element simulation system and presents preliminary elec(ci d:173) tric field measurements which are being used to tune the simulation.
****

Training Multilayer Perceptrons with the Extended Kalman Algorithm
Sharad Singhal, Lance Wu
trained with
****
An Electronic Photoreceptor Sensitive to Small Changes in Intensity
Tobi Delbrück, C. A. Mead
We describe an electronic photoreceptor circuit that is sensitive to small chan ges in incident light intensity. The sensitivity to change8 in the intensity is achieved by feeding back to the input a filtered version of the output. The fe edback loop includes a hysteretic el(cid:173) ement. The circuit behaves in a ma nner reminiscent of the gain control properties and temporal responses of a var iety of retinal cells, particularly retinal bipolar cells. We compare the thres holds for detection of intensity increments by a human and by the cir(cid:173) cuit. Both obey Weber's law and for both the temporal contrast sensitivities ar e nearly identical.
****
Training a 3-Node Neural Network is NP-Complete
Avrim Blum, Ronald Rivest
We consider a 2-layer, 3-node, n-input neural network whose nodes compute linear threshold functions of their inputs. We show that it is NP-complete to decide whether there exist weights and thresholds for the three nodes of thi s network so that it will produce output con(cid:173) sistent with a given se t of training examples. We extend the result to other simple networks. This result suggests that those looking for perfect training algorithms cannot escape inherent computational difficulties just by considering only simple o r very regular networks. It also suggests the importance, given a training pro blem, of finding an appropriate network and input encoding for that problem. It is left as an open problem to extend our result to nodes with non-linear functions such as sigmoids.
****
Automatic Local Annealing
Jared Leinbach
This research involves a method for finding global maxima in constraint satis faction networks. It is an annealing process butt unlike most otherst requires no annealing schedule. Temperature is instead determined local ly by units at each updatet and thus all processing is done at the unit leve l. There are two major practical benefits to processing this way: 1) processing can continue in 'bad t areas of the networkt while 'good t areas remain stablet and 2) processing continues in the 'bad t areast as long as the constraints remain poorly satisfied (i.e. it does not stop a fter some predetermined number of cycles). As a resultt this method not only avoids the kludge of requiring an externally determined annealing sc hedulet but it also finds global maxima more quickly and consistently tha n externally scheduled systems the to Boltzmann machine (Ackley et alt 198 5) is made). FinallYt implementation of this method is computationally trivial .
****
Adaptive Neural Net Preprocessing for Signal Detection in Non-Gaussian Noise
Richard P. Lippmann, Paul Beckman
A nonlinearity is required before matched filtering in mInimum error r eceivers when additive noise is present which is impulsive and highly non-Gaussian. Experiments were performed to determine whether the corre ct clipping nonlinearity could be provided by a single-input single(cid:173) output multi-layer perceptron trained with back propagation. It was found that a multi-layer perceptron with one input and output node, 20 node s in the first hidden layer, and 5 nodes in the second hidden layer could be trained to provide a clipping nonlinearity with fewer than 5,000 pr esentations of noiseless and corrupted waveform samples. A network train ed at a relatively high signal-to-noise (SIN) ratio and then used as a front end for a linear matched filter detector greatly reduced the

probability of error. The clipping nonlinearity formed by this network was similar to that used in current receivers designed for impulsive noise and provided similar substantial improvements in performance.
****

Learning the Solution to the Aperture Problem for Pattern Motion with a Hebb Rule
Martin Sereno
to
****

GENESIS: A System for Simulating Neural Networks
Matthew Wilson, Upinder Bhalla, John Uhley, James Bower
support simulations at many it is
****

Neural Architecture
Valentino Braitenberg
Requests for name changes in the electronic proceedings will be accepted with no questions asked. However name changes may cause bibliographic tracking issues. Authors are asked to consider this carefully and discuss it with their co-authors prior to requesting a name change in the electronic proceedings.
****

Learning by Choice of Internal Representations
Tal Grossman, Ronny Meir, Eytan Domany
We introduce a learning algorithm for multilayer neural net(cid:173) works composed of binary linear threshold elements. Whereas ex(cid:173)isting algorithms reduce the learning process to minimizing a cost function over the weights, our method treats the internal repre(cid:173)sentations as the fundamental entities to be determined. Once a correct set of internal representations is arrived at, the weights are found by the local aild biologically plausible Perceptron Learning Rule (PLR). We tested our learning algorithm on four problems: adjacency, symmetry, parity and combined symmetry-parity.
****

Adaptive Neural Networks Using MOS Charge Storage
Daniel Schwartz, R. Howard, Wayne Hubbard
MOS charge storage has been demonstrated as an effective method to store the weights in VLSI implementations of neural network models by several workers 2 . However, to achieve the full power of a VLSI implementation of an adaptive algorithm, the learning operation must built into the circuit. We have fabricated and tested a circuit ideal for this purpose by connecting a pair of capacitors with a CCD like structure, allowing for variable size weight changes as well as a weight decay operation. A 2.51-' CMOS version achieves better than 10 bits of dynamic range in a 140 /' X 3501-' area. A 1.25/' chip based upon the same cell has 1104 weights on a 3.5mm x 6.0mm die and is capable of peak learning rates of at least 2 x 109 weight changes per second.
****

Implications of Recursive Distributed Representations
Jordan Pollack
I will describe my recent results on the automatic development of fixed(cid:173)width recursive distributed representations of variable-sized hierarchal data structures. One implication of this wolk is that certain types of AI-style data-structures can now be represented in fixed-width analog vectors. Simple inferences can be perfonned using the type of pattern associations that neural networks excel at Another implication arises from noting that these representations become self-similar in the limit Once this door to chaos is opened. many interesting new questions about the representational basis of intelligence emerge, and can (and will) be discussed.
****

Backpropagation and Its Application to Handwritten Signature Verification
Timothy Wilkinson, Dorothy Mighell, Joseph Goodman

A pool of handwritten signatures is used to train a neural net(cid:173) work for the task of deciding whether or not a given signature is a forgery. The network is a feedforward net, with a binary image as input. There is a hidden layer, with a single unit output layer. The weights are adjusted acco rding to the backpropagation algorithm. The signatures are entered into a C so ftware program through the use of a Datacopy Electronic Digitizing Camera. The binary signa(cid:173) tures are normalized and centered. The performance is examined as a function of the training set and network structure. The bes t scores are on the order of 2% true signature rejection with 2-4% fal se signature acceptance.

****

Theory of Self-Organization of Cortical Maps

Shigeru Tanaka

We have mathematically shown that cortical maps in the primary sensory cortices can be reproduced by using three hypotheses which have physi ological basis and meaning. Here, our main focus is on ocular.dominance colu mn formation in the primary visual cortex. Monte Carlo simulations on the seg regation of ipsilateral and contralateral afferent terminals are carried out. Based on these, we show that almost all the physiological experimental re sults concerning the ocular dominance patterns of cats and monkeys reared un der normal or various abnormal visual conditions can be explained from a viewp oint of the phase transition phenomena.

****

An Adaptive Network That Learns Sequences of Transitions

C. Winter

We describe an adaptive network, TIN2, that learns the transition functi on of a sequential system from observations of its behavior. It integrates tw o subnets, TIN-I (Winter, Ryan and Turner, 1987) and TIN-2. TIN-2 construct s state representations from examples of system behavior, and its dynami cs are the main topics of the paper. TIN-I abstracts transition functions from noisy state representations and environmental data during training, while in operation it produces sequences of transitions in response to variations in inp ut. Dynamics of both nets are based on the Adaptive Resonance Theory of Carpen ter and Grossberg (1987). We give results from an experiment in which TIN2 le arned the behavior of a system that recognizes strings with an even number of l 's .

****

Dynamics of Analog Neural Networks with Time Delay

Charles Marcus, R. Westervelt

A time delay in the response of the neurons in a network can induce sustained o scillation and chaos. We present a stability criterion based on local stability analysis to prevent sustained oscillation in symmetric delay networks, an d show an example of chaotic dynamics in a non-symmetric delay network .

****

On the K-Winners-Take-All Network

E. Majani, Ruth Erlanson, Yaser Abu-Mostafa

We present and rigorously analyze a generalization of the Winner(cid:173) Tak e-All Network: the K-Winners-Take-All Network. This net(cid:173) work ide ntifies the K largest of a set of N real numbers. The network model u sed is the continuous Hopfield model.

****

Does the Neuron "Learn" like the Synapse?

Raoul Tawel

An improved learning paradigm that offers a significant reduction in com(cid:173) putation time during the supervised learning phase is described. It is based on extending the role that the neuron plays in artificial neural systems. Prior wor k

has regarded the neuron as a strictly passive, non-linear processing element, an
d
the synapse on the other hand as the primary source of information processing an
d
knowledge retention. In this work, the role of the neuron is extended insofar as
 allow(cid:173)
ing its parameters to adaptively participate in the learning phase. The temperat
ure
of the sigmoid function is an example of such a parameter. During learning, both
 the
synaptic interconnection weights w[j and the neuronal temperatures Tr are opti(c
id:173)
mized so as to capture the knowledge contained within the training set. The meth
od
allows each neuron to possess and update its own characteristic local temperatur
e.
This algorithm has been applied to logic type of problems such as the XOR or par
ity
problem, resulting in a significant decrease in the required number of training
cycles.
****

Use of Multi-Layered Networks for Coding Speech with Phonetic Features
Yoshua Bengio, Régis Cardin, Renato de Mori, Piero Cosi
Preliminary results on speaker-independant speech recognition are reporte
d. A method that combines expertise on neural networks with expertise o
n speech recognition is used to build the recognition systems. For tr
ansient sounds, event(cid:173) driven property extractors with variable r
esolution in the time and frequency domains are used. For sonorant sp
eech, a model of the human auditory system is preferred to FFT as a
 front-end module.
****

Electronic Receptors for Tactile/Haptic Sensing
Andreas Andreou
We discuss synthetic receptors for haptic sensing. These are based on magnetic
 field sensors (Hall effect structures) fabricated using standard CMOS technolo
gies. These receptors, biased with a small permanent magnet can detect the pre
sence of ferro or ferri-magnetic objects in the vicinity of the sensor. They ca
n also detect the magnitude and direction of the magnetic field.
****

A Programmable Analog Neural Computer and Simulator
Paul Mueller, Jan Van der Spiegel, David Blackman, Timothy Chiu, Thomas Clare, J
oseph Dao, Christopher Donham, Tzu-pu Hsieh, Marc Loinaz
This report describes the design of a programmable general purpose analog
neural computer and simulator. It is intended primarily for real-world real
-time computations such as analysis of visual or acoustical patterns, rob
otics and the development of special purpose neural nets. The machine is sca
lable and composed of interconnected modules containing arrays of neurons, mod
ifiable synapses and switches. It runs entirely in analog mode but connection
 architecture, synaptic gains and time constants as well as neuron parameters
are set digitally. Each neuron has a limited number of inputs and can be conne
cted to any but not all other neurons. For the determination of synaptic gains
and the implementation of learning algorithms the neuron outputs are mu
ltiplexed, AID converted and stored in digital memory. Even at moderate size
 of 1()3 to IDS neurons computational speed is expected to exceed that of any
current digital computer.
****

An Information Theoretic Approach to Rule-Based Connectionist Expert Systems
Rodney Goodman, John Miller, Padhraic Smyth
We discuss in this paper architectures for executing probabilistic rule-bases in
 a par(cid:173) allel manner, using as a theoretical basis recently introduce

d information-theoretic models. We will begin by describing our (non-neural) l earning algorithm and theory of quantitative rule modelling, followed by a d iscussion on the exact nature of two particular models. Finally we work throu gh an example of our approach, going from database to rules to inference networ k, and compare the network's performance with the theoretical limits for speci fic problems.
****

A Connectionist Expert System that Actually Works
Richard Fozzard, Gary Bradshaw, Louis Ceci
Requests for name changes in the electronic proceedings will be accepted with no questions asked. However name changes may cause bibliographic tracking issues. Authors are asked to consider this carefully and discuss it with their co-auth ors prior to requesting a name change in the electronic proceedings.
****

Cricket Wind Detection
John Miller
Requests for name changes in the electronic proceedings will be accepted with no questions asked. However name changes may cause bibliographic tracking issues. Authors are asked to consider this carefully and discuss it with their co-auth ors prior to requesting a name change in the electronic proceedings.
****

ALVINN: An Autonomous Land Vehicle in a Neural Network
Dean A. Pomerleau
ALVINN (Autonomous Land Vehicle In a Neural Network) is a 3-layer back-propagat ion network designed for the task of road following. Cur(cid:173) rently ALVINN takes images from a camera and a laser range finder as input and produces as ou tput the direction the vehicle should travel in order to follow the road. Train ing has been conducted using simulated road images. Successful tests on the Car negie Mellon autonomous navigation test vehicle indicate that the network can e ffectively follow real roads under certain field conditions. The representation developed to perfOIm the task differs dra(cid:173) matically when the networlc is trained under various conditions, suggesting the possibility of a novel adap tive autonomous navigation system capable of tailoring its processing to the co nditions at hand.
****

Fast Learning in Multi-Resolution Hierarchies
John Moody
A class of fast, supervised learning algorithms is presented. They use lo(cid:173) 3)
****

Mapping Classifier Systems Into Neural Networks
Lawrence Davis
Classifier systems are machine learning systems incotporating a genetic al(cid:173)
****

A Network for Image Segmentation Using Color
Anya Hurlbert, Tomaso Poggio
We propose a parallel network of simple processors to find color boundaries irr espective of spatial changes in illumi(cid:173) nation, and to spread uniform co lors within marked re-
****

Training a Limited-Interconnect, Synthetic Neural IC
M. Walker, S. Haghighi, A. Afghan, Larry Akers
Hardware implementation of neuromorphic algorithms is hampered by high degrees of connectivity. Functionally equivalent feedforward networks may be formed b y using limited fan-in nodes and additional layers. but this complicates pr ocedures for determining weight magnitudes. No direct mapping of weights ex ists between fully and limited-interconnect nets. Low-level nonlinearities prevent the formation of internal representations of widely separated spa tial features and the use of gradient descent methods to minimize output error

is hampered by error magnitude dissipation. The judicious use of linear summations or collection units is proposed as a solution.
****

A Model for Resolution Enhancement (Hyperacuity) in Sensory Representation
Jun Zhang, John Miller
Heiligenberg (1987) recently proposed a model to explain how sen(cid:173)
****

Temporal Representations in a Connectionist Speech System
Erich Smythe
SYREN is a connectionist model that uses temporal information in a speech signal for syllable recognition. It classifies the rates and directions of formant center transitions, and uses an adaptive method to associate transition events with each syllable. The system uses explicit spatial temporal representations through de(cid:173) lay lines. SYREN uses implicit parametric temporal representa(cid:173) tions in formant transition classification through node activation onset, decay, and transition delays in sub-networks analogous to visual motion detector cells. SYREN recognizes 79% of six repe(cid:173) titions of 24 consonant-vowel syllables when tested on unseen data, and recognizes 100% of its training syllables.
****

Neural Network Star Pattern Recognition for Spacecraft Attitude Determination and Control
Phillip Alvelda, A. San Martin
computational bottlenecks
****

A Massively Parallel Self-Tuning Context-Free Parser
Eugene Santos
The Parsing and Learning System(PALS) is a massively parallel self-tuning context-free parser. It is capable of parsing sentences of unbounded length mainly due to its parse-tree representation scheme. The system is capable of improving its parsing performance through the presentation of training examples.
****

Neural Net Receivers in Multiple Access-Communications
Bernd-Peter Paris, Geoffrey Orsak, Mahesh Varanasi, Behnaam Aazhang
The application of neural networks to the demodulation of spread-spectrum signals in a multiple-access environment is considered. This study is motivated in large part by the fact that, in a multiuser system, the conventional (matched fil(cid:173) ter) receiver suffers severe performance degradation as the relative powers of the interfering signals become large (the "near-far" problem). Furthermore, the optimum receiver, which alleviates the near-far problem, is too complex to be of practical use. Receivers based on multi-layer perceptrons are considered as a simple and robust alternative to the opti(cid:173) mum solution. The optimum receiver is used to benchmark the performance of the neural net receiver; in particular, it is proven to be instrumental in identifying the decision regions of the neural networks. The back-propagation algorithm and a modified version of it are used to train the neural net. An importance sampling technique is introduced to reduce the number of simulations necessary to evaluate the performance of neural nets. In all examples considered the proposed neu(cid:173) ral ~et receiver significantly outperforms the conventional recelver.
****

A Computationally Robust Anatomical Model for Retinal Directional Selectivity
Norberto Grzywacz, Franklin Amthor
We analyze a mathematical model for retinal directionally selective cells based on recent electrophysiological data, and show that its computation of motion direction is robust against noise and speed.
****

Statistical Prediction with Kanerva's Sparse Distributed Memory
David Rogers
A new viewpoint of the processing performed by Kanerva's sparse distri

buted memory (SDM) is presented. In conditions of near- or over- capaci ty, where the associative-memory behavior of the mod(cid:173) el breaks down, the processing performed by the model can be inter(cid:173) preted as that of a statistical predictor. Mathematical results are presente d which serve as the framework for a new statistical view(cid:173) poi nt of sparse distributed memory and for which the standard for(cid:173) mulation of SDM is a special case. This viewpoint suggests possi(cid: 173) ble enhancements to the SDM model, including a procedure for impr oving the predictiveness of the system based on Holland's work with ' Genetic Algorithms', and a method for improving the capacity of SDM even when used as an associative memory.
****

Learning Sequential Structure in Simple Recurrent Networks
David Servan-Schreiber, Axel Cleeremans, James McClelland
We explore a network architecture introduced by Elman (1988) for predicting suc cessive elements of a sequence. The network uses the pattern of activation over a set of hidden units from time-step t-l, together with element t, to predict element t+ 1. When the network is trained with strings from a particular finite -state grammar, it can learn to be a perfect finite-state recognizer for the gr ammar. Cluster analyses of the hidden-layer patterns of activation showed that they encode prediction-relevant information about the entire path traversed thr ough the network. We illustrate the phases of learning with cluster analyses p erformed at different points during training.
****

A Back-Propagation Algorithm with Optimal Use of Hidden Units
Yves Chauvin
This paper presents a variation of the back-propagation algo(cid:173) ri thm that makes optimal use of a network hidden units by de(cid:173) cr~ asing an "energy" term written as a function of the squared activations of these hidden units. The algorithm can automati(cid:173) cally find o ptimal or nearly optimal architectures necessary to solve known Boolean functions, facilitate the interpretation of the activation of the rem aining hidden units and automatically estimate the complexity of architectu res appropriate for phonetic labeling problems. The general principle of t he algorithm can also be adapted to different tasks: for example, it can b e used to eliminate the [0, 0] local minimum of the [-1. +1] logistic a cti(cid:173) vation function while preserving a much faster convergence and forcing binary activations over the set of hidden units.
****

GEMINI: Gradient Estimation Through Matrix Inversion After Noise Injection
Yann Le Cun, Conrad Galland, Geoffrey E. Hinton
Learning procedures that measure how random perturbations of unit ac(cid:173) ti vities correlate with changes in reinforcement are inefficient but simple to im plement in hardware. Procedures like back-propagation (Rumelhart, Hinton and Wi lliams, 1986) which compute how changes in activities af(cid:173) fect the outpu t error are much more efficient, but require more complex hardware. GEMINI is a hybrid procedure for multilayer networks, which shares many of the implementat ion advantages of correlational reinforce(cid:173) ment procedures but is more e fficient. GEMINI injects noise only at the first hidden layer and measures the resultant effect on the output error. A linear network associated with each hid den layer iteratively inverts the matrix which relates the noise to the error c hange, thereby obtaining the error-derivatives. No back-propagation is involved , thus allowing un(cid:173) known non-linearities in the system. Two simulations demonstrate the effectiveness of GEMINI.
****

A Self-Learning Neural Network
Allan Hartstein, R. Koch
Requests for name changes in the electronic proceedings will be accepted with no questions asked. However name changes may cause bibliographic tracking issues. Authors are asked to consider this carefully and discuss it with their co-auth

ors prior to requesting a name change in the electronic proceedings.
****
Neural Networks for Model Matching and Perceptual Organization
Eric Mjolsness, Gene Gindi, P. Anandan
We introduce an optimization approach for solving problems in com(cid:173)
****
Neuronal Maps for Sensory-Motor Control in the Barn Owl
Clay D. Spence, John Pearson, J. Gelfand, R. Peterson, W. Sullivan
The bam owl has fused visual/auditory/motor representations of space in its mid brain which are used to orient the head so that visu(cid:173) al or auditory sti muli are centered in the visual field of view. We present models and computer s imulations of these structures which address various problems, inclu

****
Performance of Synthetic Neural Network Classification of Noisy Radar Signals
Stanley Ahalt, F. Garber, I. Jouny, Ashok Krishnamurthy
This study evaluates the performance of the multilayer-perceptron and the freq uency-sensitive competitive learning network in iden(cid:173) tifying five commercial aircraft from radar backscatter measure(cid:173) ments. The pe rformance of the neural network classifiers is com(cid:173) pared with tha t of the nearest-neighbor and maximum-likelihood classifiers. Our results indicate that for this problem, the neural network classifiers are re latively insensitive to changes in the net(cid:173) work topology, and to the noise level in the training data. While, for this problem, the t raditional algorithms outperform these sim(cid:173) ple neural classifiers, we feel that neural networks show the poten(cid:173) tial for improved performa nce.
****
Connectionist Learning of Expert Preferences by Comparison Training
Gerald Tesauro
A new training paradigm, caned the "eomparison pa.radigm," is introduce d for tasks in which a. network must learn to choose a prdcrred pattern from a set of n alternatives, based on examplcs of Imma.n expert prderenc es. In this pa.radigm, the inpu t to the network consists of t.wo uf the n alterna tives, and the trained output is the expert's judgement of which pa.ttern is better. This para.digm is applied to the lea,rning of hackg ammon, a difficult board ga.me in wllieh the expert selects a move from a. set, of legal mm·es. \Vith compa.rison training, much higher levels of performance can hc a.chiew~d, with networks that are much smaller, and with coding sehemes t.hat are much simpler and easier to underst and. Furthermorf', it is possible to set up the network so tha.t it always produces consisten t rank-orderings .
****
Winner-Take-All Networks of O(N) Complexity
J. Lazzaro, S. Ryckebusch, M.A. Mahowald, C. A. Mead
We have designed, fabricated, and tested a series of compact CMOS integrated ci rcuits that realize the winner-take-all function. These analog, continuous-time circuits use only O(n) of interconnect to perform this function. We have also modified the winner-take-all circuit, realizing a circuit that computes local n onlinear inhibition.
****
Neural Network Recognizer for Hand-Written Zip Code Digits
John Denker, W. Gardner, Hans Graf, Donnie Henderson, R. Howard, W. Hubbard, L. D. Jackel, Henry Baird, Isabelle Guyon
This paper describes the construction of a system that recognizes hand-printed digits, using a combination of classical techniques and neural-net methods. The system has been trained and tested on real-world data, derived from zip codes s een on actual U.S. Mail. The system rejects a small percentage of the examples as unclassifiable, and achieves a very low error rate on the remaining examples . The system compares favorably with other state-of-the art recognizers. While

some of the methods are specific to this task, it is hoped that many of the tec hniques will be applicable to a wide range of recognition tasks.
****

Analog Implementation of Shunting Neural Networks
Bahram Nabet, Robert Darling, Robert Pinter
An extremely compact, all analog and fully parallel implementa(cid:173) tion of a class of shunting recurrent neural networks that is ap(cid:173) plicable to a wide variety of FET-based integration technologies is propos ed. While the contrast enhancement, data compression, and adaptation to mean i nput intensity capabilities of the network are well suited for processing of sensory information or feature extrac(cid:173) tion for a content addressable m emory (CAM) system, the network also admits a global Liapunov function and can thus achieve stable CAM storage itself. In addition the model can readil y function as a front-end processor to an analog adaptive resonance circuit.
****

Range Image Restoration Using Mean Field Annealing
Griff Bilbro, Wesley Snyder
A new optimization strategy, Mean Field Annealing, is presented. Its appl ication to MAP restoration of noisy range images is derived and experimentally verified.
****

Analyzing the Energy Landscapes of Distributed Winner-Take-All Networks
David Touretzky
DCPS (the Distributed Connectionist Production System) is a neural netwo rk with complex dynamical properties. Visualizing the energy landscapes of some of its component modules leads to a better intuitive understanding of the model, and suggests ways in which its dynamics can be controlled i n order to improve performance on difficult cases.
****

Neural Approach for TV Image Compression Using a Hopfield Type Network
Martine Naillon, Jean-Bernard Theeten
A self-organizing Hopfield network has been developed in the context of Vector Ouantiza(cid:173) -tion, aiming at compression of television images. The metas table states of the spin glass-like network are used as an extra the Minimal Overlap storage resource using and Mezard 1987) to rule (Krauth learning th e organization of the attractors. optimize The sel f-organi zi ng that we hav e scheme devised the generation of an in adaptive codebook for any qiven TV image.
****

Speech Recognition: Statistical and Neural Information Processing Approaches
John Bridle
Requests for name changes in the electronic proceedings will be accepted with no questions asked. However name changes may cause bibliographic tracking issues. Authors are asked to consider this carefully and discuss it with their co-auth ors prior to requesting a name change in the electronic proceedings.
****

Convergence and Pattern-Stabilization in the Boltzmann Machine
Moshe Kam, Roger Cheng
The Boltzmann Machine has been introduced as a means to perform global optimiza tion for multimodal objective functions using the principles of simulated annea ling. In this paper we consider its utility as a spurious-free content-addressa ble memory, and provide bounds on its performance in this context. We show how to exploit the machine's ability to escape local minima, in order to use it, at a constant temperature, for unambiguous associative pattern-retrieval in noisy environments. An association rule, which creates a sphere of influence around each stored pattern, is used along with the Machine's dynamics to match the ma chine's noisy input with one of the pre-stored patterns. Spurious fIxed points, whose regions of attraction are not recognized by the rule, are skipped, due t o the Machine's fInite probability to escape from any state. The results apply to the Boltzmann machine and to the asynchronous net of binary threshold elemen

ts (Hopfield model'). They provide the network designer with worst-case and best-case bounds for the network's performance, and allow polynomial-time tradeoff studies of design parameters.
****

Models of Ocular Dominance Column Formation: Analytical and Computational Results
Kenneth Miller, Joseph Keller, Michael Stryker
We have previously developed a simple mathemati(cid:173)
****

Digital Realisation of Self-Organising Maps
Nigel Allinson, Martin Johnson, Kevin Moon
Kevin J. Moon
****

Further Explorations in Visually-Guided Reaching: Making MURPHY Smarter
Bartlett Mel
MURPHY is a vision-based kinematic controller and path planner based on a connectionist architecture, and implemented with a video camera and Rhino XR-series robot arm. Imitative of the layout of sen(cid:173) sory and motor maps in cerebral cortex, MURPHY'S internal representa(cid:173) tions consist of four coarse-coded populations of simple units represent(cid:173) ing both static and dynamic aspects of the sensory-motor environment. In previously reported work [4], MURPHY first learned a direct kinematic model of his camera-arm system during a period of extended practice, and then used this "mental model" to heuristically guide his hand to unobstructed visual targets. MURPHY has since been extended in two ways: First, he now learns the inverse differential-kinematics of his arm in addition to ordinary direct kinematics, which allows him to push his hand directly towards a visual target without the need for search. Sec(cid:173) ondly, he now deals with the much more difficult problem of reaching in the presence of obstacles.
****

Scaling and Generalization in Neural Networks: A Case Study
Subutai Ahmad, Gerald Tesauro
The issues of scaling and generalization have emerged as key issues in current studies of supervised learning from examples in neural networks. Questions such as how many training patterns and training cycles are needed for a problem of a given size and difficulty, how to represent the inllUh and how to choose useful training exemplars, are of considerable theoretical and practical importance. Several intuitive rules of thumb have been obtained from empirical studies, but as yet there are few rig(cid:173) orous results. In this paper we summarize a study Qf generalization in the simplest possible case-perceptron networks learning linearly separa(cid:173) ble functions. The task chosen was the majority function (i.e. return a 1 if a majority of the input units are on), a predicate with a num(cid:173) ber of useful properties. We find that many aspects of.generalization in multilayer networks learning large, difficult tasks are reproduced in this simple domain, in which concrete numerical results and even some analytic understanding can be achieved.
****

A Model of Neural Oscillator for a Unified Submodule
Alexandr Kirillov, G. N. Borisyuk, R. M. Borisyuk, Ye. Kovalenko, V. Makarenko, V. Chulaevsky, V. Kryukov
Requests for name changes in the electronic proceedings will be accepted with no questions asked. However name changes may cause bibliographic tracking issues. Authors are asked to consider this carefully and discuss it with their co-authors prior to requesting a name change in the electronic proceedings.
****

An Optimality Principle for Unsupervised Learning
Terence Sanger
We propose an optimality principle for training an unsu(cid:173) pervised feed

forward neural network based upon maximal ability to reconstruct the input data from the network out(cid:173) puts. We describe an algorithm which can be used to train either linear or nonlinear networks with certain types of non linearity. Examples of applications to the problems of image coding, featur e detection, and analysis of random(cid:173) dot stereograms are presented.
****

Neural Analog Diffusion-Enhancement Layer and Spatio-Temporal Grouping in Early Vision

Allen Waxman, Michael Seibert, Robert Cunningham, Jian Wu

A new class of neural network aimed at early visual processing is desc ribed; we call it a Neural Analog Diffusion-Enhancement Layer or "NADEL." The network consists of two levels which are coupled through feedfoward and shunted feedback connections. The lower level is a two-dimensional diffusion map which accepts visual features as input, and spreads activity over larg er scales as a function of time. The upper layer is periodically fed the activ ity from the diffusion layer and locates local maxima in it (an extreme form of contrast enhancement) using a network of local comparators. These local maxi ma are fed back to the diffusion layer using an on-center/off-surround s hunting anatomy. The maxima are also available as output of the network. The network dynamics serves to cluster features on multiple scales as a fu nction of time, and can be used in a variety of early visual processing tasks s uch as: extraction of comers and high curvature points along edge contour s, line end detection, gap filling in contours, generation of fixation points, perceptual grouping on multiple scales, correspondence and path impletion in l ong-range apparent motion, and building 2-D shape representations that are invariant to location, orientation, scale, and small deformation on the visual field.
****

Optimization by Mean Field Annealing

Griff Bilbro, Reinhold Mann, Thomas Miller, Wesley Snyder, David van den Bout, M ark White

Nearly optimal solutions to many combinatorial problems can be found using stoc hastic simulated annealing. This paper extends the concept of simulated anneali ng from its original formulation as a Markov process to a new formulation based on mean field theory. Mean field annealing essentially replaces the discrete d e(cid:173) grees of freedom in simulated annealing with their average values as computed by the mean field approximation. The net result is that equilibrium a t a given temperature is achieved 1-2 orders of magnitude faster than with simu lated annealing. A general frame(cid:173) work for the mean field annealing algo rithm is derived, and its re(cid:173) lationship to Hopfield networks is shown. The behavior of MFA is examined both analytically and experimentally for a gene ric combi(cid:173) natorial optimization problem: graph bipartitioning. This ana lysis indicates the presence of critical temperatures which could be im(cid:173) portant in improving the performance of neural networks.
****

An Application of the Principle of Maximum Information Preservation to Linear Sy stems

Ralph Linsker

This paper addresses the problem of determining the weights for a set of line ar filters (model "cells") so as to maximize the ensemble-averaged info rmation that the cells' output values jointly convey about their input values, given the statistical properties of the ensemble of input vectors. The quan tity that is maximized is the Shannon information rate, or equivalently th e average mutual information between input and output. Several models for th e role of processing noise are analyzed, and the biological motivation for con sidering them is described. For simple models in which nearby input signal v alues (in space or time) are correlated, the cells resulting from thi s optimization process include center-surround cells and cells sensitive to temporal variations in input signal.
****

Spreading Activation over Distributed Microfeatures
James Hendler
One att·empt at explaining human inferencing is that of spread(cid:173) ing acti vat,ion, particularly in the st.ructured connectionist para(cid:173) digm. This has resulted in t.he building of systems with semanti(cid:173) cally nameable no des which perform inferencing by examining t.he pat,t.erns of activation spread . In this paper we demonst.rate t.hat simple structured network infert'ncing ca n be p(>rformed by passing art.iva.t.ion over the weights learned by a distribu ted alga(cid:173) rit,hm. Thus , an account, is provided which explains a well(c id:173) behaved rela t ionship bet.ween structured and distri butt'd conn('c(cid :173) t.ionist. a.pproachrs.
****

Consonant Recognition by Modular Construction of Large Phonemic Time-Delay Neura l Networks
Alex Waibel
In this paperl we show that neural networks for speech recognition can be const ructed in a modular fashion by exploiting the hidden structure of previ ously trained phonetic subcategory networks. The performance of resulting la rger phonetic nets was found to be as good as the performance of the subc omponent nets by themselves. This approach avoids the excessive learning t imes that would be necessary to train larger networks and allows for increm ental learning. Large time-delay neural networks constructed incrementall y by applying these modular training techniques achieved a recognition performance of 96.0% for all consonants.
****

Constraints on Adaptive Networks for Modeling Human Generalization
Mark Gluck, M. Pavel, Van Henkle
The potential of adaptive networks to learn categorization rules and to model human performance is studied by comparing how natural and artificial systems respond to new inputs, i.e., how they generalize. Like humans, networ ks can learn a detenninistic categorization task by a variety of alte rnative individual solutions. An analysis of the con(cid:173) straints im posed by using networks with the minimal number of hidden units shows that t his "minimal configuration" constraint is not sufficient to explain and pr edict human performance; only a few solu(cid:173) tions were found to be s hared by both humans and minimal adaptive networks. A further analysis o f human and network generalizations indicates that initial conditions ma y provide important constraints on generalization. A new technique, which we call "reversed learning", is described for finding appropriate initial c onditions.
****

Self Organizing Neural Networks for the Identification Problem
Manoel Tenorio, Wei-Tsih Lee
This work introduces a new method called Self Organizing Neural Network (SONN) algorithm and demonstrates its use in a system identification t ask. The algorithm constructs the network, chooses the neuron functions, a nd adjusts the weights. It is compared to the Back-Propagation algorithm in the identification of the chaotic time series. The results shows that SONN c onstructs a simpler, more accurate model. requiring less training data and e pochs. The algorithm can be applied and generalized to appilications as a class ifier.
****

Learning with Temporal Derivatives in Pulse-Coded Neuronal Systems
David Parker, Mark Gluck, Eric Reifsnider
A number of learning models have recently been proposed which involve calculati ons of temporal differences (or derivatives in continuous-time models). These m odels. like most adaptive network models. are formulated in tenns of frequency (or activation), a useful abstraction of neuronal firing rates. To more precise ly evaluate the implications of a neuronal model. it may be preferable to devel op a model which transmits discrete pulse-coded information. We point out that

many functions and properties of neuronal processing and learning may depend. in subtle ways. on the pulse-coded nature of the informa(cid:173) tion coding and transmission properties of neuron systems. When com(cid:173) pared to formulations in terms of activation. computing with temporal derivatives (or differences) as proposed by Kosko (1986). Klopf (1988). and Sutton (1988). is both more stable and easier when refor(cid:173) mulated for a more neuronally realistic pulse-coded system. In refor(cid:173) mulating these models in terms of pulse-coding. our motivation has been to enable us to draw further parallels and connections between real-time behavioral models of learning and biological circuit models of the substrates underlying learning and memory.

****

Using Backpropagation with Temporal Windows to Learn the Dynamics of the CMU Direct-Drive Arm II

Kenneth Goldberg, Barak Pearlmutter

Computing the inverse dynamics of a robot ann is an active area of research in the control literature. We hope to learn the inverse dynamics by training a neural network on the measured response of a physical ann. The input to the network is a temporal window of measured positions; output is a vector of torques. We train the network on data measured from the first two joints of the CMU Direct-Drive Arm II as it moves through a randomly-generated sample of "pick-and-place" trajectories. We then test generalization with a new trajectory and compare its output with the torque measured at the physical arm. The network is shown to generalize with a root mean square error/standard deviation (RMSS) of 0.10. We interpreted the weights of the network in tenns of the velocity and acceleration filters used in conventional control theory.

****

Storing Covariance by the Associative Long-Term Potentiation and Depression of Synaptic Strengths in the Hippocampus

Patric Stanton, Terrence J. Sejnowski

In modeling studies or memory based on neural networks, both the selective enhancement and depression or synaptic strengths are required ror effident storage or inrormation (Sejnowski, 1977a,b; Kohonen, 1984; Bienenstock et aI, 1982; Sejnowski and Tesauro, 1989). We have tested this assumption in the hippocampus, a cortical structure or the brain that is involved in long-term memory. A brier, high-frequency activation or excitatory synapses in the hippocampus produces an increase in synaptic strength known as long-term potentiation, or L TP (BUss and Lomo, 1973), that can last ror many days. LTP is known to be Hebbian since it requires the simultaneous release or neurotransmitter from presynaptic terminals coupled with postsynaptic depolarization (Kelso et al, 1986; Malinow and Miller, 1986; Gustatrson et al, 1987). However, a mechanism ror the persistent reduction or synaptic strength that could balance LTP has not yet been demonstrated. We stu(cid:173) died the associative interactions between separate inputs onto the same dendritic trees or hippocampal pyramidal cells or field CAl, and round that a low-frequency input which, by itselr, does not persistently change synaptic strength, can either increase (associative L TP) or decrease in strength (associative long-term depression or LTD) depending upon whether it is positively or negatively correlated in time with a second, high-frequency bursting input. LTP or synaptic strength is Hebbian, and LTD is anti-Hebbian since it is elicited by pairing presynaptic firing with post(cid:173) synaptic hyperpolarization sufficient to block postsynaptic activity. Thus, associa(cid:173) tive L TP and associative L TO are capable or storing inrormation contained in the covariance between separate, converging hippocampal inputs •

****

Dynamic, Non-Local Role Bindings and Inferencing in a Localist Network for Natural Language Understanding

Trent Lange, Michael Dyer

This paper introduces a means to handle the critical problem of non(cid:173) local role-bindings in localist spreading-activation networks. Every conceptual node in the network broadcasts a stable, uniquely-identifying activatio

n pattern, called its signature. A dynamic role-binding is cre(cid:173) ated when a role's binding node has an activation that matches the bound co ncept's signature. Most importantly, signatures are propagated across long pat hs of nodes to handle the non-local role-bindings neces(cid:173) sary for infe rencing. Our localist network model, ROBIN (ROle Binding and Inferencin g Network), uses signature activations to ro(cid:173) bustly represent sch emata role-bindings and thus perfonn the inferenc(cid:173) ing, plan/goal analys is, schema instantiation, word-sense disambigua(cid:173) tion, and dynamic re-i nterpretation portions of the natural language un(cid:173) derstanding process.
****

Fixed Point Analysis for Recurrent Networks
Patrice Simard, Mary Ottaway, Dana Ballard
This paper provides a systematic analysis of the recurrent backpropaga(cid:173) tion (RBP) algorithm, introducing a number of new results. The main limitation of the RBP algorithm is that it assumes the convergence of the network to a sta ble fixed point in order to backpropagate the error signals. We show by experim ent and eigenvalue analysis that this condi(cid:173) tion can be violated and th at chaotic behavior can be avoided. Next we examine the advantages of RBP over the standard backpropagation al(cid:173) gorithm. RBP is shown to build stable f ixed points corresponding to the input patterns. This makes it an appropriate t ool for content address(cid:173) able memories, one-to-many function learning, a nd inverse problems.
****