INTRODUCTION

JOSEF SKRZYPEK and WALTER KARPLUS

Machine Perception Laboratory
Computer Science Department
University of California at Los Angeles
Los Angeles, CA 90024, USA

1. NEURAL NETWORKS IN VISION AND PATTERN RECOGNITION

Neural Networks in Vision and Pattern Recognition explores the application of the neural network paradigm to research focused on understanding both human and machine perception of patterns. This collection of papers, which is also published as a special issue of the *International Journal of Pattern Recognition and Artificial Intelligence* (Vol. 6, No. 1, 1992), is an attempt to bridge the gap between two well-established fields, Pattern Recognition and Artificial Intelligence. While the former often deals with signal processing, the latter has been dominated by symbolic, top-down approaches to problems of machine ''cognition''. This dichotomy reflects the basic difficulties in understanding perceptual and cognitive activities; the stimuli from the environment must go through signal processing stages before the symbolism of its content can be communicated. Vision is a prime example of the problems that we face where the early stages of visual processing, best understood in terms of concepts developed in signal processing, must interact with expectations and models of the environment that are best manipulated in the symbolic domain. Solutions are difficult because in some tasks the problem could be solved in the signal processing domain while other tasks require interactions between signal and symbolic representations. It is not always clear at what level of depth should signals and symbols interact and the transition from one to the other is generally difficult. For this reason the neural network paradigm might be the best bridge between AI and PR that has come into existence.

2. NEURAL NETWORK PARADIGM

The neural network paradigm has been extensively detailed; many good reviews of the field are available in the literature. A common view of the neural and/or connectionist networks are collections of simple, interconnected computational elements. The network's behavior results or ''emerges'' from the *interconnection* of these simple computational elements. Popular aliases for these computational elements include *neurons*, *units*, and *nodes*.

Most of the connectionist and neural models proposed in the past several years are based on the simplified model of a ''neuron'' where each neuron: (1) calculates a weighted sum of its inputs, (2) performs a threshold operation on the sum, (3) optionally passes the thresholded value through a linear or non-linear transfer function, (4) modifies its behavior (via its connection weights) based on its sensory experience (past inputs). In general, neural models can be decomposed into three major components: (1) an interconnected network of nodes, (2) a node activation rule (function) which dictates how the node's inputs are to be processed, and (3) a node learning rule (function) which indicates how to modify the node's connectivity weights.

3. ADVANTAGES OF THE NEURAL NETWORK PARADIGM

In the neural network paradigm, the environment or data can be used as an active element which influences the answer. In a biological system, the function is represented by the structure of the network, which can be dynamically reconfigured in order to extract the desired answer from the input data. In other words, the network could "self-organize" pathways for signal flow so that after a certain period of relaxation, the answer can emerge from the network's structure. In this case, it may be impossible or perhaps meaningless to know the exact pattern of signal flow. This differs from traditional computing, where the solution procedure has to be known before the program can be written. An interesting side effect of this concept is that neural networks can adapt to input data by comparing obtained results with the desired ones. This discrepancy can be used to modify the behavior of the network in order to improve the algorithm. Other forms of learning are also possible.

Neural networks appear to be better suited for the realization of very complex systems. The information, at least in the analog domain, could be represented by a signal's amplitude varying in time. This limited dimensionality makes interfaces between modules easier than multidimensional symbols such as words. Nevertheless, a formal methodology does not exist at the present time to develop separate modules and to have an *a priori* specification of the link weights needed to properly connect the modules into more complex systems. Supervised and/or unsupervised learning offer some solutions to these problems and represent a very significant advantage of synthetic neural systems over the more traditional programming based approaches to system building.

The massive parallelism exhibited by neural networks makes it easy to create systems that are fault tolerant. A correct decision made at the output of a neural network depends on the pattern of activity of a subset of neurons. Thus, knowledge representation does not depend on the exact connectivity between neurons, but rather is based on an approximate pattern of activity distributed across the network's topology. Does parallelism imply merely an increased speed of computation? Since we can implement any parallel computer as a universal Turing machine, parallelism does not offer any theoretical advantages. However, cognitive and perceptual tasks are physical functions underlying survival, and they must be completed in an acceptable amount of time; interactions with a natural environment impose "real-time" computing constraints. This suggests that the success of a specific function is determined by the underlying computing structure and vice versa, so that a desired function will dictate the optimal computing architecture. Furthermore, perceptual functions seem dependent on the task and the environment in which the behavior is to take place. Thus, it may be difficult or even impossible to algorithmically solve a problem that is based on interacting constraints between the network's structure, the desired function and the environment. Consequently, it seems plausible that neural networks offer not only speed improvements but also a qualitatively different, non-procedural computational approach, necessary for implementation of cognitive and perceptual functions.

Whereas traditional computer science approaches have separated hardware and software for engineering simplicity, nature has integrated them for performance and efficiency; neural networks embody coupling between structure and function of their "hardware" to maximize performance. Conceivably, future solutions to perception could benefit from the close match between function and the underlying architecture. In other words, the desired function must dictate the computing architecture. In fact, this paradigm of distributed representation and computation may provide not only improvements in computing speed, but also the conceptual tools necessary to realize perception in machines.

Neural systems capable of distributed knowledge representation display plasticity. This guarantees graceful degradation in case of erroneous decisions or novel stimuli. Consequently, the network is always capable of some response which eventually converges to the desired output. In other words, it is often very difficult to enumerate *a priori* all of the possible rules that may be relevant to computing the answer to a given problem. An equally impossible task is to completely specify all of the contextual conditions that might affect such rules. Whenever the precise calculation/analysis of all of the numerical decisions necessary to solve a problem is infeasible, an approximate response provided by the neural network is better than no response at all.

4. COMPUTATIONAL NEUROSCIENCE AND ARTIFICIAL NEURAL SYSTEMS

Almost all of the contributions to Neural Networks can be categorized as attempts at modeling various processes. This is simply because all of the current implementations of "neural" tasks have been so far only in software; without the special "neural" hardware, successes will be limited. In this respect, the current euphoria (neuronia) over the "neural" paradigm seems to span the spectrum of modeling approaches from synthesis to analysis of the neural networks; from engineering of synthetic "neural" systems to computational neuroscience. Of course, the usefulness of a model is eventually determined by the goal behind its creation.

Many engineering models, presented as remote abstractions of the "brain", are of little use in understanding cognitive processes. This is in contrast to many studies where models are employed as a means to understanding various physical aspects of a nervous system. In such studies, conclusions inferred from specific facts are supported by experimental data. These conclusions advance the science of the brain. The "engineering" models are based on deductive conclusions reasoned out from general principles; frequently engineering models become an end in themselves rather than a mechanism for understanding the physical process of a cognitive function. Nevertheless, some of these synthesized "neural" structures might advance the engineering aspects of computation.

A complete enumeration and description of all activities that are within the bounds of computational neuroscience is beyond the scope of this editorial and it would be difficult to do justice to the enormous effort in this area that has been expended over the last 45 years. In one sense, modeling efforts in neurophysiology such as the Hodgkin-Huxley model of spiking membranes, Rall's theory of signal propagation in dendritic trees, Marr's model of the cerebellum, and Robinson's theory of the occulomotory system are prime examples of excellence in computational neuroscience. Many good texts are available where physiological processes are modeled with mathematical physics. One common characteristic of all of these theories is that they are driven primarily by the need to explain existing physiological data. No attempt has ever been made to develop an *a priori* plan that would guide the development of such theories and the subsequent collection of data; despite the enormous amount of collected data little of it can be used to develop a coherent model of the brain. Consequently, the potential usefulness of the modeling effort in brain research has not been fully appreciated or utilized within the neurosciences.

5. HIGHLIGHTS

This volume contains ten papers which are highlighted below. The first paper by J. Skrzypek and D. Gungner addresses the problem of lightness constancy, first defined by Hering in 1878. The motivation is to determine to what extent the luminance contrast captured by a simplified neural network of the retinal Outer Plexiform Layer explains lightness constancy. They report that while some aspects of this phenomenon can be explained with such a simplified neural network, the information contained within the ON channel response is not sufficient to reconstruct the surface lightness. The second paper by S. Edelman and T. Poggio is also in Computational Neuroscience. The authors describe a prototype-based learning scheme for recognition of wire-frame objects. The approach is based on network approximations of multivariate functions. In essence a network synthesizes by learning a standard representation of an object and by "observing" its perspective views.

The third and fourth papers are focused on Pattern Recognition problems. D. Michaels addresses the more theoretical question of the internal organization of classifier networks trained by the backpropagation algorithm. The empirical results suggest that in a feedforward net, trained on separable and nonseparable problems, "hidden" units act as difference operators that reject features common to all input patterns. The fourth paper by R. Tisdale and W. Karplus addresses the problem of system identification, viewed as a pattern recognition, with artificial neural networks. The classifier called General Vector Quantizer is proposed where discriminants evolve automatically when learning their function thus facilitating correct classification of system response within the response space.

The fifth, sixth, and seventh papers, represent the area of computational vision. David Suter looks at the problem of visual reconstruction using analog networks. Since reconstruction is an ill-formed problem, the regularization of it is first sought by finding a function that minimizes a certain functional. This paper proposes a solution where various orders of image function derivatives are explicitly reconstructed at the same time as the function itself. The approach is computationally advantageous as it allows handling of constraints on derivatives and leads to natural analog network implementation. The sixth, short paper by V. Atalay, E. Gelenbe, and N. Yalabik describes the use of the random neural network model to generate various visual textures. The model starts with a randomly generated gray-level image and produces textures with controllable features, as for example granularity. Furthermore the model is computationally less complex than, for example, Markov random fields. The seventh paper by Lai-Wan Chan introduces a novel method for the translational invariant recognition of objects in scenes with multiple and overlapping objects.

The eighth paper by J. Goodwin, B. Rosen and J. Vidal presents a technique for image recognition and reconstruction using a Boltzmann machine type neural network based on the use of magnetic thin films and opto-magnetic control. Images imprinted on magnetic film in an external magnetic field can be learned and recalled later when the new input is a corrupted version of the stored prototype. The ninth paper by B. Kämmerer proposes a method to filter the known uncertainty of input data that influences weight values during training. The motivation is that during the recognition stage these variations in weight values could mask out the real data. The final paper by P. Martin and C. Bellissant describes a neural network for the recognition of images representing musical scores. A multilayer perceptron is used to perform image segmentation. Classical image analysis techniques in conjunction with decision tree neural networks are used to perform the final classification.

ACKNOWLEDGEMENTS

We are grateful to many authors who submitted their manuscripts in response to our initial call for papers. Out of 53 submissions ten were selected based on multiple reviews by 38 referees whose unselfish contribution is greatly appreciated.

We would like to express our gratitude to all the contributors who made this book possible. Thanks also to Prof. Dr. Horst Bunke, Editor-in-charge of the *International Journal of Pattern Recognition and Artificial Intelligence* for facilitating the realization of this book, and to Ms. Jennifer Gan of World Scientific. Special thanks to the students in the Machine Perception Laboratory at UCLA for helping out with countless details necessary to integrate individual manuscripts into a book. Finally, thanks to Deborah Estrin and Joshua E. Skrzypek for granting JS the time necessary to complete this project.

Support for the UCLA Machine Perception Laboratory environment is provided in part by generous grants from IBM and Hewlett Packard. We sincerely acknowledge support by ARCO-UCLA Grant #1, ONR grant #N00014-86-K-0395, ARO grant DAAL03-88-K-0052 and PMTC-ATI grant #N00123-87-D-0364.