DESIGN AND PERFORMANCE OF A PROTOTYPE GENERAL PURPOSE ANALOG NEURAL COMPUTER

PAUL MUELLER, JAN VAN DER SPIEGEL, VINCENT AGAMI, PERVES AZIZ, DAVID BLACKMAN, PETER CHANCE, ADNAN CHOUDHURY, CHRISTOPHER DONHAM, RALPH ETIENNE, LISA JONES, PETER KINGET, WALTER VON KOCH, JINSOO KIM, JANE XIN.

Depts. of Electrical Engineering and Biophysics
The University Of Pennsylvania and Corticon Inc.
Philadelphia,PA 19104

Abstract

A programmable analog neural computer and selected applications are described. The machine is assembled from over 100 custom VLSI Modules containing neurons, synapses, routing switches and programmable synaptic time constants. Connection symmetry and modular construction allow expansion to arbitrary size. The network runs in real time analog mode, however connection architecture as well as neuron and synapse parameters are controlled by a digital host that monitors also the network performance through a D/A interface. Programming and monitoring software has been developed and several application examples including the dynamic decomposition of acoustical patterns are described. The machine is intended for real time, real world computations. In current configuration its maximal speed is equivalent to that of a digital machine capable of more than 10^{12} FLOPS.

We have assembled and tested a programmable Analog Neural Computer designed for real time computation of neural network problems. A photograph of the machine is shown in Fig.1.

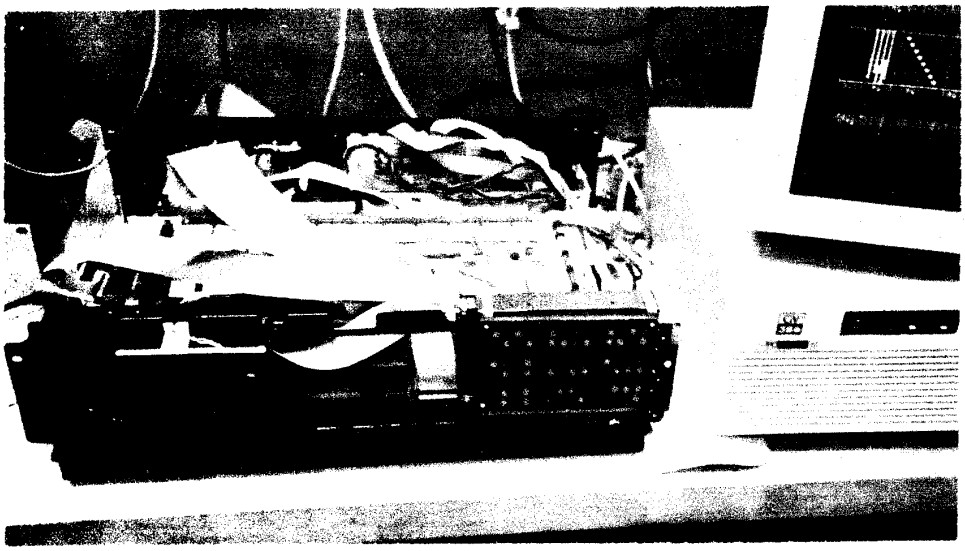

Fig.1. Photograph of the neural computer, showing the main circuit boards, I/O buffers, power supplies and LED display panel in their chassis. The cover was removed. The host computer is at right.

SYSTEM OVERVIEW

The machine contains directly interconnected arrays of **neurons, modifiable synapses, modifiable synaptic time constants and analog routing switches**. The arrays are fabricated on VLSI chips and packaged in planar chip carriers which form separate modules that are mounted on interconnected circuit boards. Neuron arrays are arranged in rows and columns and surrounded by synapse and routing switch arrays. The switches select the connections between neurons. The general architecture of the machine is shown in Fig. 2.

0-7803-0164-1/91/0000-0463$01.00©1991 IEEE

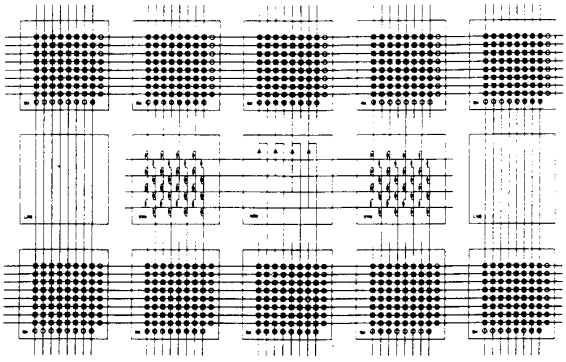

Fig.2. Schematic of the network architecture. Neuron modules receive inputs from synapse modules east and west and feed outputs into switch modules north and south. The circles on the switch modules are cut switches. Only 1/2 of the connections are drawn. The time constants which are positioned between switches and synapses are also not shown. Notice the connection symmetry.

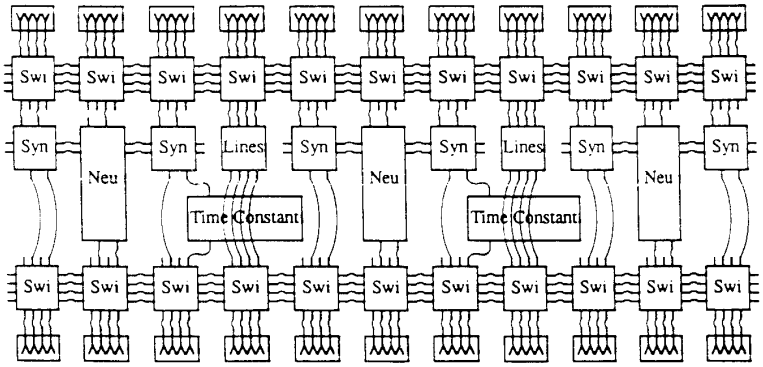

Fig.3. Block diagram of one of three processor boards that comprise the prototype. The major components and 1/4 of the analog data paths are shown. Support chips, digital data buses, and power connections have been omitted. The boards are connected with three 50 pin connectors. I/O Buffers and digital interface are on separate boards.

The computer runs entirely in analog mode. However, connection architectures, synaptic gains, time constants and neuron parameters are set by a digital host computer through an interface board either from the keyboard or from stored programs. For the implementation of learning algorithms, selected time segments of the outputs from all neurons are multiplexed on each neuron chip, digitized and stored in host memory. The multiplexing operation is independent of the analog computations.

The machine has several unique attributes. First, both the gains and the time constants of synaptic transfer are separately programmable, a feature that is indispensable for analog computation of time domain phenomena. Second, it is constructed from interchangeable modules with two-dimensional symmetrical pinout that permits easy modifications of the numerical ratios and positions of neurons, synapses and routing switches. Third, the network is expandable to arbitrary size. Finally the coupling of analog and digital hardware adds the flexibility of digital methods.

In its current configuration the machine is composed of 3 circuit boards each containing 35 directly interconnected VLSI modules (Fig.3). Each module contains a separate array of neurons, modifiable synapses, modifiable synaptic time constants or analog routing switches. The neurons have an adjustable threshold and minimum output at threshold, synapse gains are programmable from 0 to $^{+}$- 15 with logarithmic 6 bit control and the time constants from 5 ms to 2 s with 4 bit resolution. For details of the VLSI components see ref. 1.

SYSTEM OPERATION

The machine is controlled by software residing in the digital host. A schematic of the different tools is shown in Fig.4.

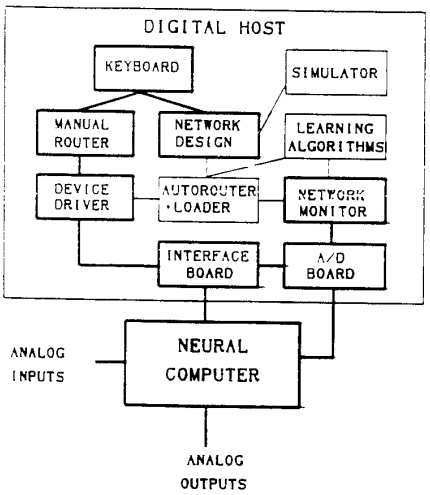

Fig. 4. Overview of the operating software for the neural computer. The portions shown in bold outline have been completed.

The connection architecture, synapse and neuron parameters are set and verified serially at 2 Mhz from a digital host through a special interface board and graphic control software. At the lowest level the network parameters are set manually through a mouse and graphic display which highlights the connections and displays the synapse parameters as seen in Fig.7. Parameters are loaded either for the entire network, a selected board or a selected module. Higher level controls, allowing the automatic transfer of conceptual networks into the machine are under development.

The network is connected directly to the outside world through parallel analog I/O buffers. In addition, each neuron chip contains an analog multiplexer that enables the digital host to monitor and store the neuron activity for graphic display and the implementation of programmed learning algorithms. The multiplexer feeds into a 250 KHz A/D board located in the host. The monitor software generates either arrays of selected time segments of each neuron output that are displayed as separate graphs as shown in Fig.6, or it generates a continuous gray scale display of activity for all neurons. Neuron activity is also displayed directly by an LED Panel.

PERFORMANCE OF SIMPLE COMPUTATIONAL TASKS

The machine has been programmed for several tasks in order to evaluate its performance. A few examples are illustrated in Figs.5-9. The figures show the conceptual networks and selected performance records. In one example the routing configuration in the neural computer is also shown. In all cases the computations are performed in real time, with response time limited only by the bandwidth of the neurons which was 30 KHz. Programs for implementation of learning algorithms are under development.

The first network (Fig.5) illustrates a fan - out , fan - in net and demonstrates that small signals generated by small synaptic weights can be transmitted and summed without distortion of the original waveform by small-signal noise.

The second example, (Fig.6), is a "Winner Take All" net which, depending on the inhibitory feedback gain, either extracts the largest output among 16 neurons or enhances the contrast between the different outputs. The network settles within the time constant of the neuron outputs (here 300 us) and showed no oscillations. Simulations of this small network on a SUN 4 are slower by a factor of 100. Since this is a fully connected net the speed ratio would scale with n^2 of the number of neurons.

Other small circuits, not illustrated here, have been programmed, including an associative network of 24 neurons, a motor control circuit that models the last stage of the control of saccadic eye movements, a neural

integrator and several circuits for the computation of time domain pattern primitives. By making use of feedback and synaptic time constants the network can generate dynamic activity patterns in the absence of external inputs. Several coupled oscillators and central pattern generators have been programmed.

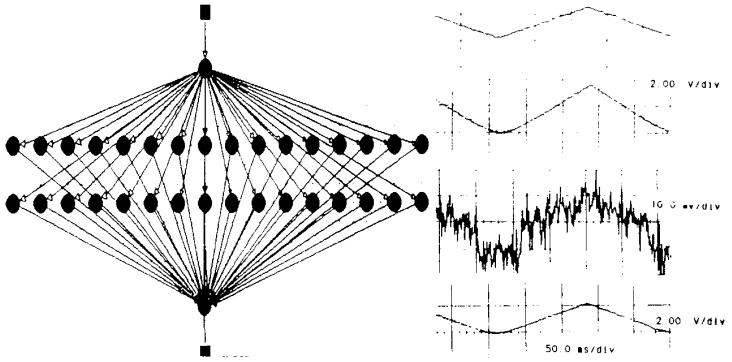

Fig. 5. A network demonstrating the fan - in / fan - out capabilities of the machine. The conceptual net is shown at left. A single external input is distributed with a synaptic gain of 0.03 from an input neuron to 32 secondary neurons and their outputs are summed at a single third stage neuron with a gain of 1. The records show from top to bottom: the input signal, the output of the first neuron, an output from a second stage neuron and the output from the final stage neuron. Notice that the low signal noise present at the second stage neurons averages out at the third stage. The noise at this stage was purposely increased.

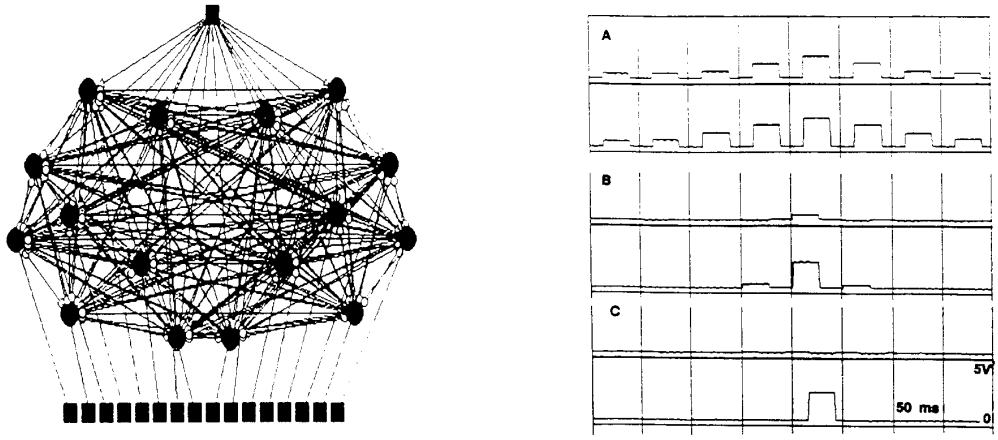

Fig.6. (left). Conceptual network and the Performance of a "Winner Take All" network. A single external input is distributed with different positive gains to 16 secondary neurons that are connected by mutual inhibitory connections. The records at right show outputs from 16 neurons in response to a square wave input of different amplitude to each neuron. In A the mutual inhibitory gains were set to zero. In B these gains were 0.5 and in C, 0.9. Notice that the 0.5 inhibition results in a contrast enhancement. The records were obtained through the analog multiplexers on the neuron modules and were displayed with the monitor software by the digital host.

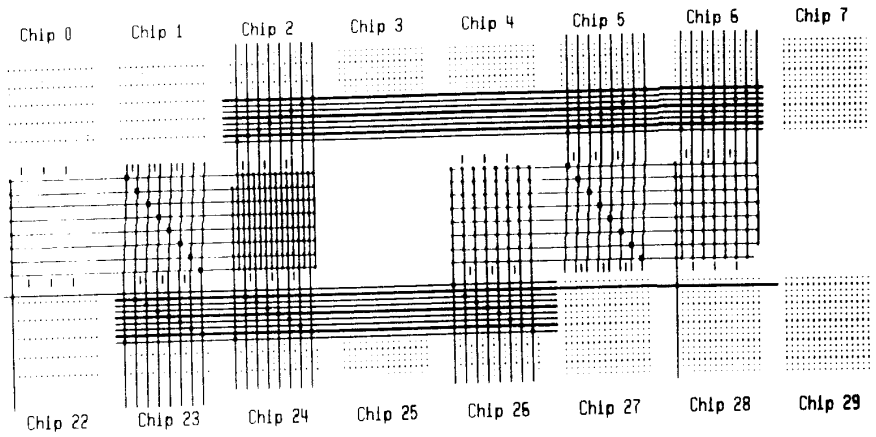

Fig.7. Routing configuration of the "Winner Take All" net. This example shows a part of the graphic displays used for routing of connections and setting of synapse parameters. Parameters are selected from graphic menus for each module.

The machine is especially suited for the real time analysis of dynamic patterns such as speech. As an example a more elaborate network was programmed for the primary decomposition of acoustical patterns. This network, involving most of the available neurons, performs an initial decomposition of acoustical patterns. It receives input from a set of 8 badness filters and extracts local maxima of amplitude vs frequency, (d^2E/dS^2), local rates of rise and fall of amplitude ($\pm dE/Dt$), and local rise and fall of frequency i.e. motion ($\pm dS/dT$) of activity along the frequency axis (see also ref. 2 and 3). The network and performance examples are illustrated in Fig.8 and 9.

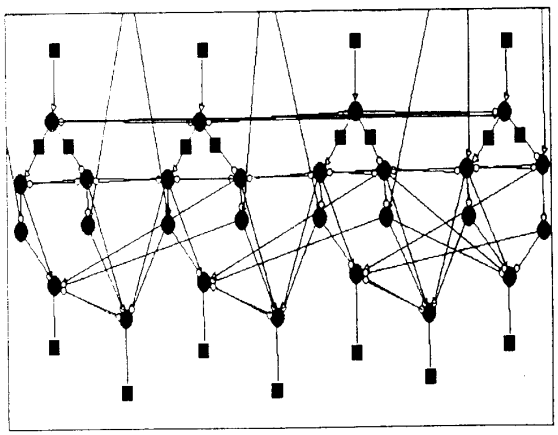

Fig. 8. Section of the conceptual design of a network for the primary decomposition of acoustical patterns. The primary neurons receive input from eight band pass filters. These neurons are connected with mutually inhibitory inputs in a center-surround scheme with spatially decaying gains. They extract the maxima of the sound amplitude. The next stage extracts separately the temporal rise and decay of the sound amplitudes. These neurons receive complementary delayed as well as undelayed excitatory and inhibitory inputs from the previous stage. The third stage neurons are normally "on" through positive bias input and compute the complement of the activity of the second stage neurons. The fourth stage units compute the changes of frequency maxima and their direction through a combination of the second and third stage neurons. In essence they are motion detectors.

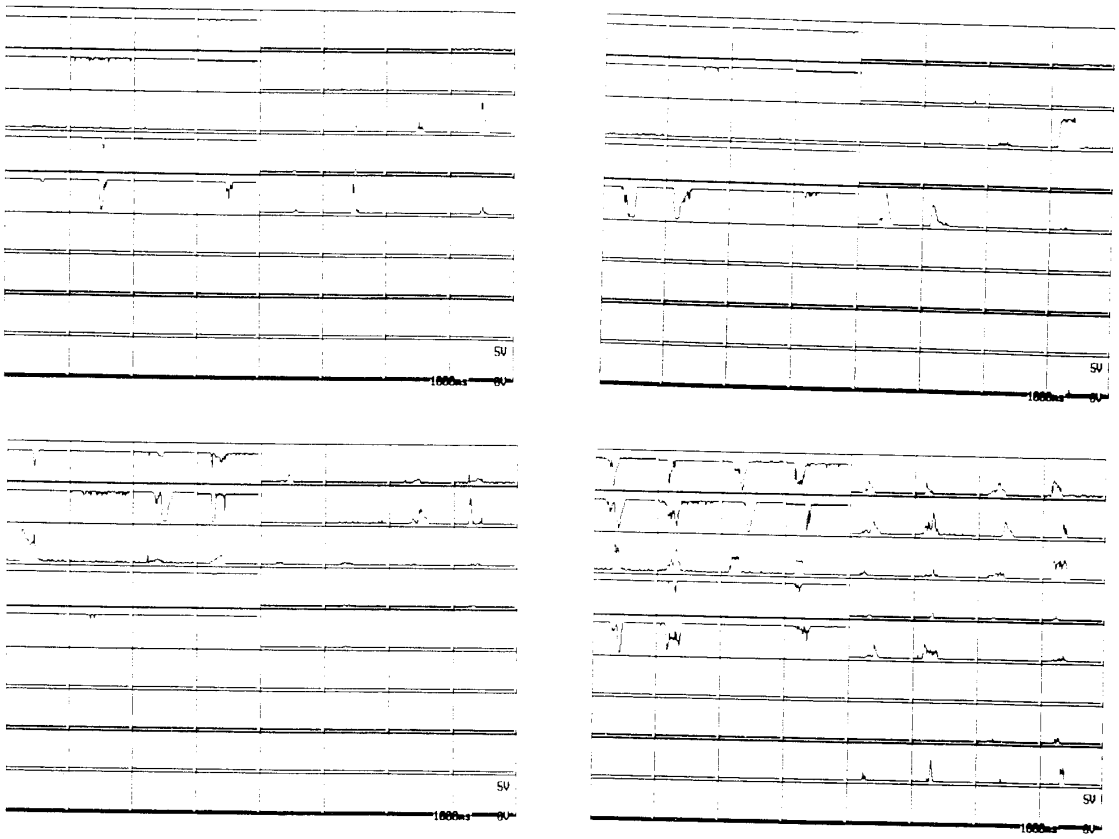

Fig.9. Neuron outputs from the network shown in Fig. 8 for four different phonemes (clockwise from top left: t,s,uh,dah), illustrating the differences of the patterns. A detailed discussion of the data is beyond the scope of this report.

Several conclusions can be drawn from these tests. First of all the machine performance has met all expectations regarding noise levels, accuracy, stability, programming flexibility and processing speed. Particularly the routing space proved entirely adequate. There is therefore no reason to project an increase in the ratio of switch modules to neuron and synapse modules for larger machines. No major design change for the neurons is needed. The synapse gain (weighting) scheme could be improved by increasing the resolution for middle range gains (between 0.1 and 1). The range and resolution of synaptic time constants seems adequate.

The major advantage of analog hardware implementations of neural computation is speed. Simulations show that the matching of real time performance of even small networks similar to that in Fig. 8 requires digital processing speeds exceeding 10^{12} FLOPS and larger machines of this type, currently on the drawing board can be expected to approach 10^{14} equivalent FLOPS.

References

1. Mueller, P., Van Der Spiegel J. et Al. Design and Fabrication of VLSI components for a General Purpose Analog Neural Computer. IEEE Workshop on VLSI and Neural Systems, Portland 1989. C. Mead Ed. Kluwer, Boston Ma.

2. Mueller, P. and Lazzaro, J. A machine for neural computation of acoustical patterns with application to real time speech recognition, AIP Conference Proceedings, 151:321-326, 1986.

3. Mueller,P., Computation of Pattern Primitives in a Neural Network for Acoustical Pattern Recognition. Proceedings of the IJCNN 1990 Wash. I 149-151 .

This work was supported by grants from the Office of Naval Research and the National Science Foundation.