MIND-1024 : A 1024 NEURONS FULLY CONNECTED REAL-TIME NEUROCOMPUTER[1]

C. GAMRAT, P. PERETTO, A. MOUGIN & O. ULRICH
CEA / CEN-Grenoble DRFMC/SP2M
BP 85X F38041 Grenoble Cedex France
e-mail gamrat@drfmc.ceng.cea.fr

Abstact : *MIND-1024 is a hardware implementation of a 1024 neurons Hopfield Network with built-in learning capabilities. The prototype of the neurocomputer uses a time-multiplexed architecture and has been built with off-the-shelf components. The architecture includes a hierarchical processor structure for learning, and a fully hardwired circuit with a noise generator for relaxing. MIND-1024 allows the use of any learning rule, synchronous or asynchronous updating schemes and real-time access to the states of the network.*

1)-Introduction

With regard to the building of real neurocomputers, a number of approaches have been put forward so far. They range from simple coprocessor boards that plug into a computer bus slot, to specialized machines using dedicated analog or digital integrated circuits. Although neural chips should be studied in order to build large neurocomputers, we feel it is not convenient to do so when we look at the poor theoretical knowledge of neural dynamics especially in learning. With this in mind, we have devised a machine that will be flexible enough to evolve according to progress in theoretical works. The project we started in 1985 is known as the MIND project. MIND is an acronym for "A Machine that Implements Neural Devices". So far, two prototypes have been built: An analog machine MIND-128 in 1987 [2] and a digital machine, MIND-1024 which we completed at the end of 1992.

MIND-1024 is a hardware implementation of a 1024 neurons fully connected Hopfield neural net. The full connectivity allows for complete flexibility over the network topologies that can be studied on the machine as long as the total number of neurons does not exceed 1024. Special hardware and a custom bus have been designed in order to get real-time access to the states and local field values of the network.

2)-Architectural options

A neurocomputer must implement two types of closely intertwined dynamics: The state dynamics also known as the recalling phase and the connections dynamics also known as the learning phase. A means of comunicating with the machine should also exist, that is the role of a host computer called the supervisor. According to this description of the functions of a neurocomputer, MIND-1024 is organized in three functional units (see figure 1). The recalling and learning units share common data by means of dual-ported memories. That is ; the weight matrix, the thresholds and the neurons states. Each of the three functional units can operate indepentendtly of the others on the same set of variables. In particular, this means that learning and recalling dynamics can occur simultaneously.

[1]This work was supported in part by a contract from the "Direction des Etudes et Recherches" (DGA/DRET).

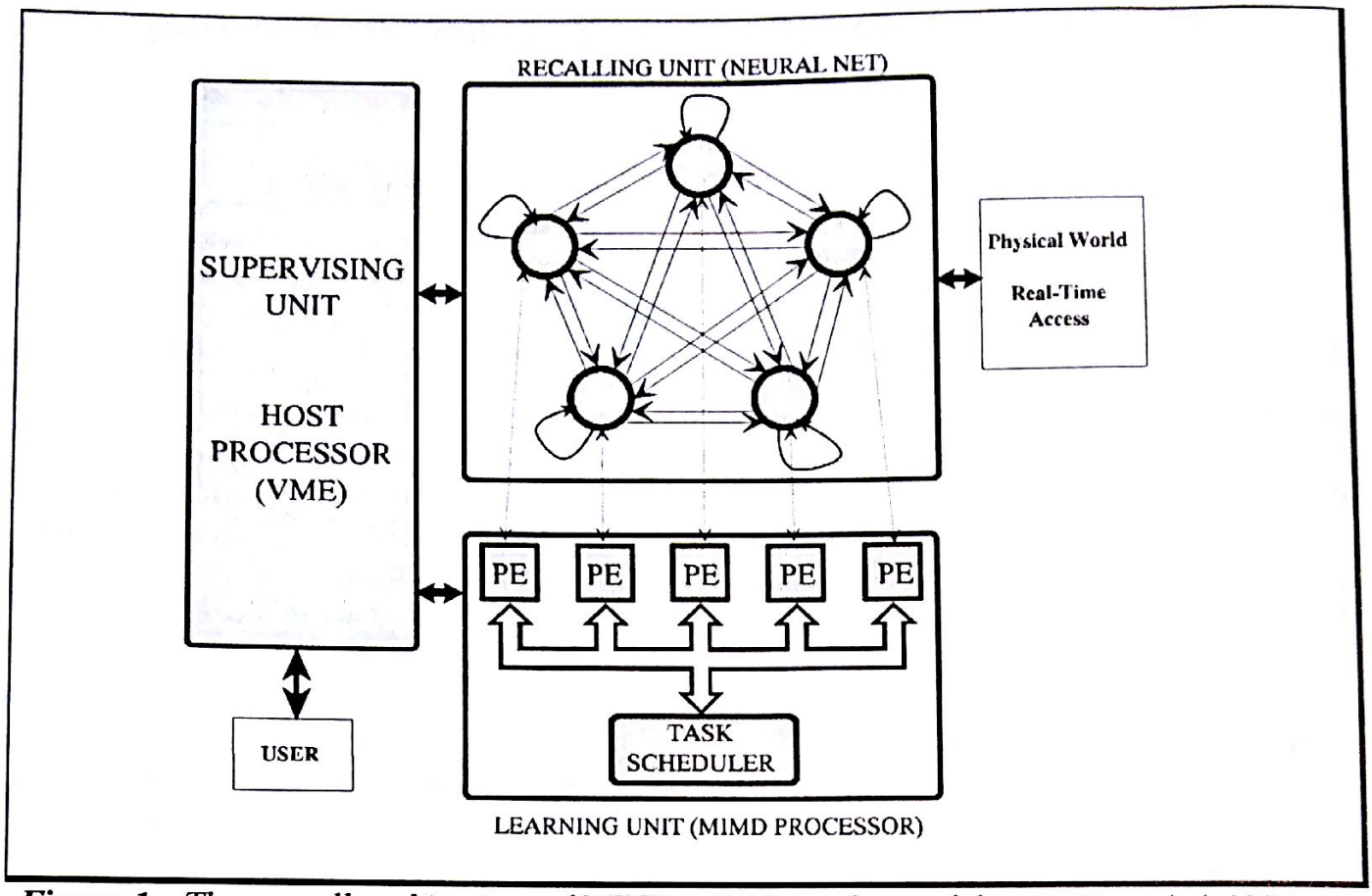

***Figure 1** - The overall architecture of MIND-1024 is made up of three units : a) A 1024 fully connected binary neuron network with real-time access. b) A 64+1 MIMD processor network for learning. c) A supervisor for the control and user interface.*

a) - Recalling unit (State dynamics).

The neurons of MIND-1024 are binary automata whose states take only two values namely Si=+/-1. Each state of the network is computed according to :

$$S_i = sign\left(f\left(\sum_{j=1}^{N} W_{ij} S_j - \theta_i + \eta_i \right) \right) \qquad (1)$$

Where the efficacy W_{ij} is the influence of neuron i on neuron j, θ_i is the threshold, η_i is a random variable (i.e. noise) whose distribution has a width T which plays the role of a temperature and f() is an arbitrary function implemented in a look-up table.

The neurons are updated in a time multiplexed fashion using a convergent semi-parallel architecture [1]. At each time step, a given neuron is fully materialized from dendrite to axone and its state is calculated using equation (1). The weighted sum is evaluated in parallel using a digital adder with 1024 inputs. The complete calculus is performed by a dedicated hardware for N=1024 in about 900 nanoseconds. This is a very convenient architecture if one wants to study, for example, a network with neurons of different kinds, since it allows for the easy implementation of neuron dependent features. Another original feature of MIND-1024, is its ability to handle "noisy neurons". The problem of noise generation, which is also fully hardwired, is given special attention. A twin digital pseudo-random generator has been designed to provide addresses to a memory in which random numbers are stored according to the desired distribution[3]. This way, the noise generator is fully programmable and can deliver more than 10^6 random numbers / second without any observable correlation.

b)- Learning unit

Naturally enough, when one thinks about performances, one will be tempted to design special hardware to perform the task as was the case for the states dynamics. However learning rules are far from being totally explained, and there exists no universal efficient learning rule. Therefore, handling of the synapses and threshold modification is deferred to an array of standard microprocessors. Every microprocessor deals with 16 neurons and their corresponding 1024 outgoing synapses. In addition, an extra processor is dedicated to the modifications of the thresholds. The processor array is basically a MIMD architecture used in a SIMD fashion since each processor runs a local copy of the same program. It is a well known problem, that inter-processor communication is the major bottle-neck when one wants to devise a parallel computer. A careful analysis of the data needed in most learning algorithms shows that apart from user defined constants, the local fields and the states of all neurons are the more frequently used data. Therefore, the states and local fields of the entire network are broadcast in real time to every learning processor by means of a dedicated bus. This way, the learning processors very seldom need communicating between each other. However, the processors are linked to a standard VME bus for non-critical communications.

c) - Supervising unit

Every functionnal part of MIND-1024 (i.e. noise generator, thresholds, synapses...) is under the control of its own microprocessor. All these processors are in turn under the control of a supervisor whose tasks are to synchronize their activities and to provide a man-machine interface. The supervisor accesses the machine ressources through a VME bus. By sending commands on the bus, the supervisor can directly control both the recalling and learning units. The supervisor is also used as the host computer and provides services for writing, debugging and executing application programs on the neurocomputer. All the operations of MIND-1024 are controlled through the use of a standard function library available for the Pascal and C programming languages.

3)- Hardware characteristics & performances

The prototype of MIND-1024 has been built using commercially available components. It is made up of 5 VME compatible crates. Four of them are filled with 16 "synapses" boards which are the basic blocks of the machine. Each of these boards houses 16 neurons with their outgoing synapses and a learning processor. The weighted sum is carried out by a digital tree of adders distibuted along the path of the signals on the backplanes of the crates. The weights are stored in dual port memories as 16 bit words. When recalling, only the 8 most significant bits are used while in learning the entire precision is considered. This reduces the hardware requirement for the summing device.

The common ressources of MIND-1024 are contained in a fifth VME crate. Apart from the supervisor which is a PC type (386 20 Mhz) VME controller, there is a board for management of the thresholds, a noise generator, a board for the transfert function, a display controller which allows for real time display of the states of the network and a sequencer which provides neuron indices and all of the timing signals. The sequencer plays a particular role since it is also used as a task scheduler for the management of the MIMD learning processor.

The two buses of MIND-1024 are the VME bus for standard dialogs and the GIBUS (General Interconnect Bus) for high speed communications. This custom bus carries all the neural network data in real-time. Therefore, it is possible to input or output neurons states or fields by simply plugging an I/O board in one of the crate's slots. Knowing that a 1024 neuron network is updated every millisecond, MIND-1024 can act as a neural controller for a physical device.

Number of binary neurons	1024
Number of synapses	1,048,576 (full connectivity)
Precision of the synapses	16 bits (learning) / 8 bits (recalling)
Precision of the thresholds	32 bits (learning) / 18 bits (recalling)
Precision of noise variable	18 bits
Type of noise distribution	Any (Gaussian, thermal and white built-in)
Cycle length (Updating time)	910 ns.
Neurons updating scheme	Asynchronous, synchronous or mixed
Number of learning processors	64 for the synapses + 1 for the thresholds
Learning processor type	Intel 80C186 at 12.5 Mhz
Learning rule	Any one programmed in C, C++ or Pascal
Recalling speed	1.1×10^9 Synapses / second
Learning speed	20×10^6 Synapses update / sec (Hebb in C)
Input/Output of neurons states	Available in real-time on a dedicated bus

***Table 1** - Hardware characteristics & performances of MIND-1024*

4)- Conclusions

It is obviously difficult to be certain that the architecture chosen in the MIND project is optimal. Nevertheless it is clear that it is very flexible and interactive. A machine like MIND-1024, which has been operationnal since november '92, offers a lot of attractive features for the neural networks theoretician.

- The full connectivity allows the implementation of any network topology. From layered structured networks to more sophisticated structures. The only limit is the number of neurons available.
- The MIMD array of learning processors plus the dedicated bus for neural data, allows the efficient programming of any learning rule.
- The updating scheme of the neurons states can be asynchronous, synchronous or mixed.
- A neural network can be composed of neurons with various noise sensitivities.
- Recalling and learning can occur simultaneously. It is a straightforward operation to implement such systems with MIND-1024, as recent work has shown their importance.[4]
- The neurocomputer is well suited for the real time command of physical devices.

Future work will put the emphasis on improving performances. It is clear that with the actual recalling cycle of 910 ns. (1.1 Mhz), we are far from any technological constraint. Changing the learning processors by state-of-the-art chips would also tremendously boost learning performances.

MIND-1024 is currently used for supporting theoretical studies on neural networks in our laboratory.

References

[1] Peretto P., Van Zurk R., Mougin A., Gamrat C. " The semi-Parallel architecture of neuro-computers", NATO ASI Series, Vol F68, Neurocomputing (1990)

[2] Gamrat C., Mougin A., Peretto P., Ulrich O. " The architecture of MIND neurocomputers", Proceedings of the 2nd intl conference on microelectronics for neural networks Munchen (1991) pp 463-469

[3] Max J. "Méthodes et techniques de traitement du signal", Vol 1, pp206-209 Masson, Paris (1972)

[4] Dong W., Hopfield J.J., "Dynamic properties of neural networks with adapting synapses" Network Vol 3 pp267-283 (1992)

[5] Peretto P. "Introduction to the modelling of neural networks" Cambridge University Press