Systems, Man & Cybernetics
Alexandria, VA, 10/1987

A Programmable Implementation for an AI Neural Architecture*

N. El-Leithy@, R. W. Newcomb@, V. Rodellar# and P. Gómez#

@Microsystems Laboratory
Electrical Engineering Department
University of Maryland
College Park, Maryland 20742 USA
Phone: (301) 454-6869

#Facultad de Informática
Universidad Politécnica de Madrid
Ctra. de Valencia, Km. 7,00
28031 Madrid, Spain

Abstract:

This paper presents an architecture suitable for realization of electronically programmable neural-systems which can be implemented in VLSI form.

I. Introduction

Neural networks, especially the original ones of Morshita & Yajima [1] whose architecture was studied by Dimopoulos [2], have become of considerable interest to AI with their development around the formulation of Hopfield & Tank [3]. This interest results because of the ability of these neural networks to conveniently mimic some of the behavior of the brain (in the cases studied by Dimopoulos) and (in the case of Hopfield & Tank) to solve problems that appear to involve rather complicated reasoning, such as the traveling salesman problem. With the present day capabilities of VLSI it appears that such neural networks can be realized on the scale needed for making computers of adequate size to handle interesting AI problems, and attempts at such realizations have been reported [4]. However, such circuit implementations of the Hopfield architecture have led to difficulties in that a change of problem has meant a change of the system hardware. Thus, to solve a new problem a new computer must be constructed.

Although there appear to be electronic means of solving this problem using the Morshita-Dimopoulos-Hopfield type of architecture, it seems wise to look at alternate architectures, especially since a different scheme of signal processing is actually used in biological systems. Here we introduce an architecture on our neural-type circuits [5] to allow for electronic reprogamming of the neural network characteristics. This is also suitable for MOS realization of the various properties as discussed in [6].

II. The Architecture

To consider the architecture we first recall our basic building blocks, the neural-type cell, NTC, and the neural-type junction, NTJ. The former mimics the pulse handling properties of the biological cell while the latter feeds signals into the NTC, much as biological synapses do. The NTC has the ability to shape input signals into neural-type pulses as well as to pulse frequency code with the number of pulses present dependent upon the level of a constant input signal, subject to threshold and refractory period constraints. Since it appears that neural information is really contained in the frequency of neural pulses, NTCs are rather important elements for neural networks based upon the frequency of pulses present. Consequently, considerable effort has been put in the past on obtaining simple but efficient electronic NTCs suitable for integrated circuit construction. NTJs on the other hand have not received the concentration in the electronic's world even though they are equally important to the VLSI realization of neural-type systems. A NTJ also does some type of coding by combining signals to excite or inhibit a NTC. And, of most interest here, via threshold levels set to cause excitation or inhibition, NTJs can be made to adapt the junctioning of their inputs so that the NTJs can in turn adapt the responses of the NTCs in the system as desired.

To accomplish this adaptation we consider the configuration of Figure 1 where we show a NTC fed by its NTJ which is in turn fed by a set of NTCs, the latter set possibly being all the NTCs in the system

*Research supported in part by NSF Grant MIP-85-06924 and US-SPAIN Joint Committee for Scientific and Technological Cooperation Grant Nos. CCB-84002-002 and CCB-8604017.

but in most cases being a more selective set fixed by the actions desired of the right hand NTC. The adaptation is under control of adaptation circuitry, the latter being of a form to implement any of a number of adaptation rules, such as least mean square error between present results and training results or minimum energy in signal states, etc.. The results of the adaptation circuitry are fed into control circuits that adapt the NTJs. In terms of the MOS circuits that we have been developing, adaptation can occur via setting of the MOS transistor thresholds. Several means are actually available for doing this electronically, one being via the insertion of charge on a floating gate [7] and another being via a change of the substrate bias voltage [8]. In the former case the threshold can be electronically set by the insertion of charge on a floating gate and can be made to hold for very long periods. In doing this we can rely upon the proven technology of electronically programmable and electronically erasable memories. If a change is desired it can be accomplished electronically by resetting the charge, and, thus, if the neural network is desired to be programmed for another job this is easily done via programming rather than rewiring. In the case of the setting of substrate bias, this too can be easily programmed but generally would be reset to zero each time the neural system losses power. However, in the case of neural computers this may not be a disadvantage since the substrate biases could be saved as part of a program and then restored to desired values when the computer is repowered.

III. Discussion

In looking for new architectures for AI it is worth looking at nerual-type structures. The more popular structures of the day suffer from some shortcomings, a serious one being that they are not conveniently electronically reprogrammed. This can be remedied in several ways, the one which we have discussed here being to turn to an architecture most suitable to NTCs and NTJs, where these latter are taken to be the cells and junctions that have been under development for realization in MOS form. When one does this, one has available the full arsenal of electronically programmable and erasable MOS devices. Recognizing this, one is naturally led to the architecture of Figure 1 where the programming of the neural network is done via the setting of MOS device thresholds via floating gate, or similar, devices, or alternatively via the setting of substrate biases. The actual circuits to carry out the setting of the threshold voltages or substrate biases somewhat depend upon the algorithms used to determine these voltages. But for any given algorithm most likely the adaptation circuitry can be designed using MOS devices so that the full neural network including adaptation circuitry can be fabricated using VLSI techniques.

References

[1]. T. Morshita and A. Yajima, "Analysis and Stimulation of Networks of Mutually Inhibiting Neurons," Kybernetic, Vol. 11, 1979, pp. 154 - 165.

[2]. N. Dimopoulos, "Organization and Stability of a Neural Network Class and the Structure of a Multiprocessor System," Ph.D. dissertation, University of Maryland, 1980.

[3]. J. J. Hopfield and D. W. Tank, "'Neural' Computation of Decisions in Optimization Problems," Biological Cybernetics, Vol. 52, No. 3, 1985, pp. 141 - 152.

[4]. M. Sivilotti, M. Emerling, and C. Mead, "A Novel Associative Memory Implemented Using Collective Computation," Proceedings of the 1985 Chapel Hill Conference on Very Large Scale Integration, Edited by H. Fuchs, Computer Science Press, Rockville, MD, 1985, pp. 329 - 342.

[5]. N. El-Leithy and R. W. Newcomb, "Overview of Neural-Type Electronics," Proceedings of the 28th Midwest Symposium on Circuits and Systems, August 1985, Louisville, KY, pp. 199 - 202.

[6]. N. El-Leithy, R. W. Newcomb, and M. Zaghloul, "A Basic MOS Neural-Type Junction, A Perspective on Neural-Type Microsystems," Proceedings of the ICNN, to appear.

[7]. S. T. Wang, "On the I-V Characteristics of Floating-Gate MOS Transistors," IEEE Transactions on Electron Devices, Vol. ED-26, No. 9, September 1979, pp. 1292 - 1294.

[8]. Z. Czarnul, G. Kiruthi, and R. W. Newcomb, "M.O.S. Neural Pulse Modulator," Electronics Letters, Vol. 15, No. 25, December 6, 1979, pp. 823 - 824.

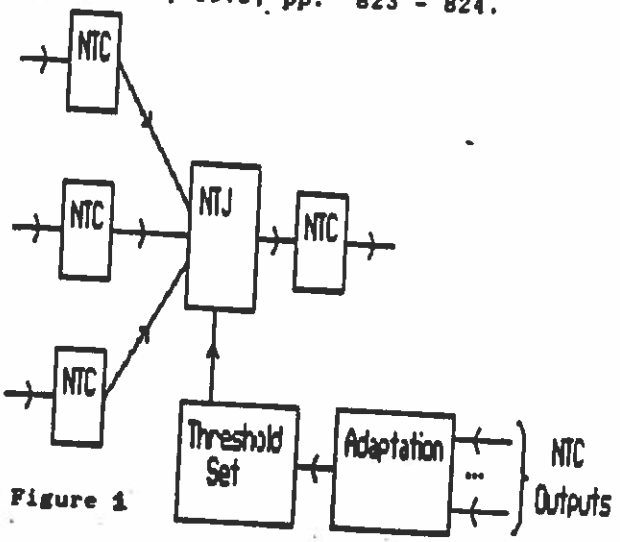

Figure 1

Architecture for Neural-Type Adaptation via MOS Threshold Settings