Extended Abstracts of the 1994 International Conference on Solid State Devices and Materials, Yokohama, 1994, pp. 385-387

Superior Generalization Capabilities of Neuron-MOS Neural Networks in Mirror-Symmetry Problem Learning

Shuhei KONDO, Tadashi SHIBATA, and Tadahiro OHMI

Department of Electronic Engineering, Tohoku University
Aza-Aoba, Aramaki, Aobaku, Sendai 980 JAPAN
PHONE. 022-224-2649/FAX. 022-224-2549

We have studied the self-learning performance of Neuron-MOS (υMOS) neural networks in solving mirror symmetry problems using computer simulation. Despite the inherent restrictions imposed on Hardware-Backpropagation (HBP) learning algorithm directly implemented on υMOS neural networks, a superior generalization capability (the ability to solve new problems not shown during the learning phase) has been demonstrated for HBP by optimizing the circuit parameters.

I INTRODUCTION

Hardware implementation of self-learning neural networks on silicon is now very actively pursued[1]. However, learning algorithms directly implemented on integrated circuit hardware have a lot of handicaps as compared to software learning algorithms running on computers. An example is the range of synaptic weight adjustment, which is $\pm\infty$ in software but is certainly limited in hardware learning. In Hardware-Backpropagation (HBP) algorithm[2], a hardware-oriented learning algorithm we have developed for Neuron MOS (υMOS) neural networks[3] by simplifying the original backpropagation (BP)[4], the weight value is limited only in the range of 0 to V_{DD}. Such a restriction to the original BP severely degrades its learning performance. Furthermore, the sigmoid function and its derivative are approximated by a step function and a window function in HBP, respectively, in order to facilitate their circuit implementation using υ MOS[5]. The purpose of this paper is to investigate the hardware-learning performance of υMOS neural networks that suffer from such restrictions using computer simulation. For mirror symmetry problem solving, We will demonstrate a superior generalization characteristics of HBP to that of original BP.

LEARNING SIMULATION

The two-dimensional mirror symmetry problem solving was employed for learning simulation. The mirror symmetry problem requires detecting which one of three possible axes of symmetry is present in 4×4 pixel (binary) input(see Fig 1). The network architecture used for simulation is depicted in Fig. 2, where each output neuron was trained to get fired corresponding to respective axis. If more than two symmetry axis were

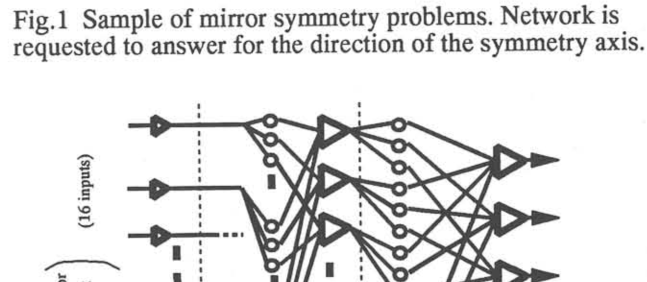

Fig.1 Sample of mirror symmetry problems. Network is requested to answer for the direction of the symmetry axis.

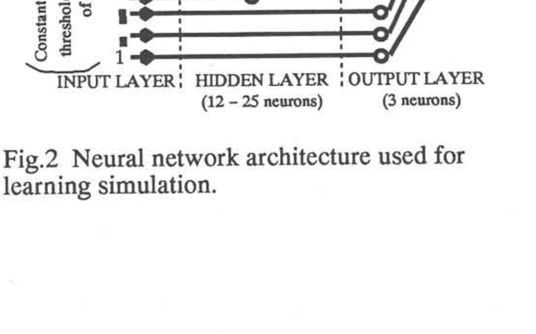

Fig.2 Neural network architecture used for learning simulation.

present in a pattern, only one axis was toughed to the network. Typically 100~200 problems were generated for each learning session. After learning iteration of 10^4/sample, the percentage of correct answers for learned problems was evaluated. The weight value of a synapse was set in the range of ±1, which corresponds to the voltage range of 0 to V_{DD} on the hardware. The initial weights were set at random values in the range of ±0.25. In this simulation, the number of hidden-layer neurons were 12 for the data in Figs. 4 and 5, and 15 in Fig. 7.

In Fig. 3, the HBP algorithm is schematically compared to BP, where NET = −1, 0, and 1 on the abscissa correspond to the dendrite potentials of 0, $V_{DD}/2$, and V_{DD} on the hardware, respectively. In HBP, weight modification is done only when the NET (weighted sum of inputs for a neuron) falls into the range of ±ε. The learning region(ε) and the learning rate(η) are the two major parameters determining the learning performance of HBP. Optimization of pertinent parameters were carried out by computer simulation and the results are presented in the following.

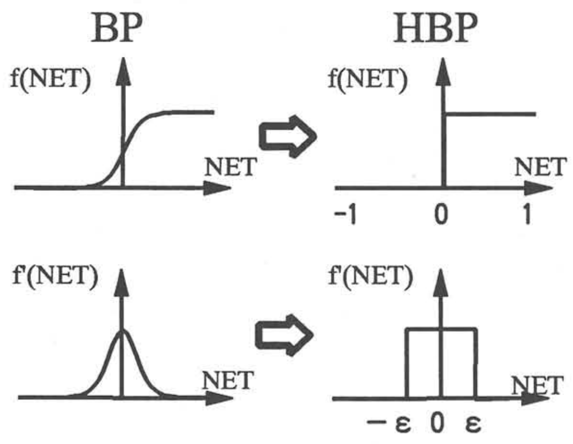

Fig.3 Hardware Backpropagate (HBP) algorithm in comparison with original BP.

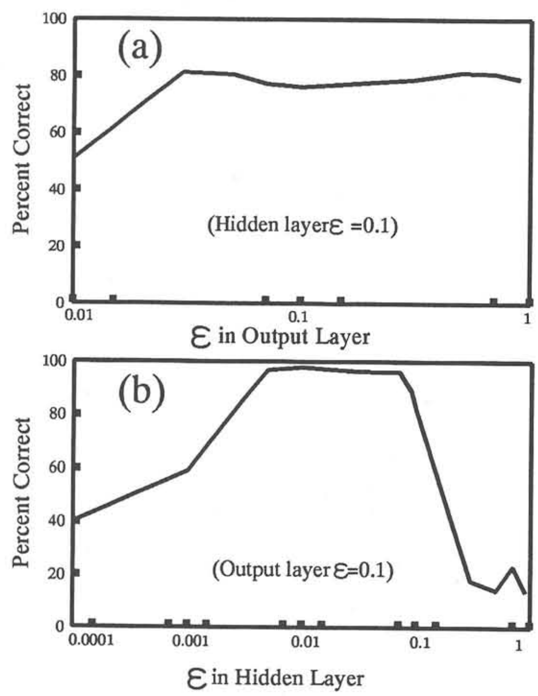

Fig.4 Learning performance of a network as a function of ε in the Output layer(a) and ε in the Hidden layer(b)

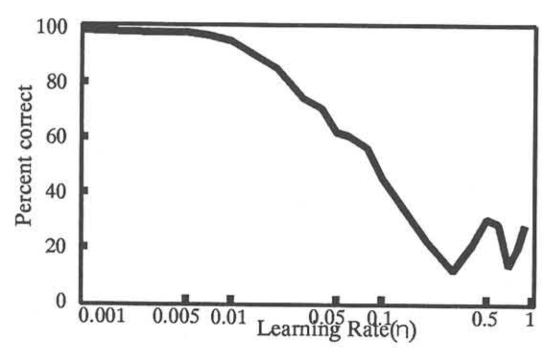

Fig.5 Percent correct for 200 learned problems as a function of learning rate.

II RESULTS AND DISCUSSION

Figure 4 demonstrates the optimization procedure of ε for output-layer neurons(a) and for hidden-layer neurons(b). For output-layer neurons, the learning performance is largely the same when ε is larger 0.03. On the other hand, ε in the hidden layer certainly has an optimum range of values between ε=0.01 and 0.1, where almost 100% learning performance is achieved. It is believed that the role of hidden layer neurons is to classify input patterns according to their characteristic features. When the learning region width(±ε) is too large (ε>0.1), very frequent updating of connection weights occurs in each hidden-layer neuron, which is not desirable for a hidden-layer neuron to establish its resonant response to specific kind of patterns. If the width is made too small (ε<0.01), the neurons can hardly learn. This is why the ε in the hidden-layer has an optimum range. For ε in the output-layer neurons, however, frequent weight updating is rather favorable because the neurons must classify a number of excitation patterns in the hidden-layer neurons into only three categories. From these results, the values of ε were determined as 0.1 in the output layer and 0.01 in the hidden layer.

The optimization of the learning rate(η) is shown in Fig. 5, indicating the degraded learning for larger learning rate. This is related to the restriction of the weight range of ±1, resulting in a small number of available weight values when η becomes large. For this reason a small value of η = 0.01 was adopted.

Figure. 6 shows the percent correct for 200 learned

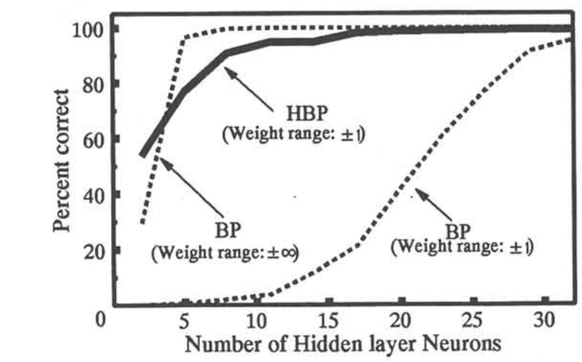

Fig.6 Percent correct for 200 learned problems as a function of the number of Hidden Layer neurons.

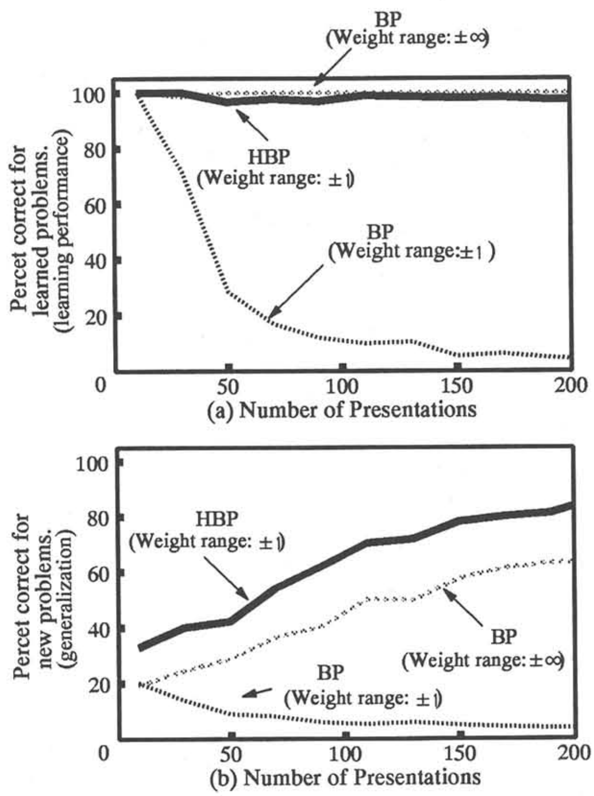

Fig.7 Comparision of learning performance(a) and genaralization(b) for HBP and BP.

problems as a function of the number of hidden-layer neurons. From this result, we employed 15 neurons in the hidden layer for HBP. The results for original BP are also shown in the figure for comparison. Original BP in which weight can take values in the range of ±∞, less number of neurons are sufficient. However, if the weight range is restricted to within ±1 in BP, the learning performance severely degrades. More than 30 hidden layer neurons are required for BP to learn 200 problems. BP algorithm loses its powerful learning ability by just limiting its range of weight adjustment. HBP has a very powerful learning ability in spite of the weight range restriction as well as of its hard-limiting squashing function.

Under the condition that all pertinent parameters are optimized and the noise margins (marginal region)[3] are also taken into account, the learning performances of HBP and original BP are compared in Fig. 7. The percent correct for problems that were used in learning sessions is almost the same for HBP and BP as shown in Fig. 7(a). However, if the weight range is restricted to ± 1 in BP as in the case of HBP, learning performance is severely degraded. Most interesting is the result shown in Fig. 7(b), where the generalization capabilities of the networks are compared. The generalization is the ability to solve new problems not shown during the learning phase. The percent correct for 800 unlearned problems are largest for HBP. HBP is equivalent to original BP in the learning ability, but is superior to BP in the generalization capability.

Ⅲ CONCLUSIONS

The learning performance of the hardware-oriented learning algorithm(HBP) developed for υMOS neural networks, has been evaluated by computer simulation. When the pertinent parameters(ε,η etc.) were optimized, HBP shows the same learning performance as that of original BP despite its inherent restrictions. In particular, HBP is superior to original BP in the generalization capability. It was also found that both learning and generalization performances are severely degraded in BP when the range of weight adjustment is limited. We can conclude that HBP is a very powerful learning algorithm which can be directly implemented on integrated circuits hardware.

Acknowledgment

This work was partially supported by the Ministry of Education, Science and Culture of Japan under Grand-in-Aid for Developmental Scientific Research 05505003.

Reference

[1]IEEE Trans on Neural Networks Vol.4, No.3, Special Issue on Neural Network Hardware (1993).

[2] H. Ishii, T. Shibata, H. Kosaka, and T. Ohmi, 1992 IEDM Tech. Dig., p.435.

[3] T. Shibata, H. Ishii, H. Kosaka, and T. Ohmi, this, conference

[4]D.E.Rumelhart et al. in *Explorations in the Microstructures of Cognition, Vol. I* ,318 Ed. by D.E.Rumelhart and J.L.MacCelland (Cambridge: MITpress).

[5] H. Ishii, T. Shibata, and T. Ohmi, this conference