The Algorithm Improvement of The Neuron Model Based on Dendrites Mechanism

Zijun *Sha and Lin Hu*

Graduate School of Science and Engineering for Education, University of Toyama Gofuku 3190, Toyama-shi, Japan

Summary
In 1943, McCulloch and Pitts have created a computational model for neural networks. However, this model just only contains the spatial summation of neuron cell and threshold action. Because of this, all the neuron cells are provided with the same functions. In 2000, Tang created a neuron model based on The Koch-Poggio-Torre model whose mount of dendrites and shape are random, and synapses interact with each other. This model will eliminate the synapses which are needless, will reinforce others which are necessary, and even make synapses and dendrites` generation, so as to form a kind of dendrites holding a special function. However, this learning speed of the model is not so ideal. In this paper, the variable learning rate is introduced in order accelerate the learning speed.
Key words:
neuron model, dendrites mechanisms, learning speed, variable learning rates.

1. Introduction

Vigorous research on neural networks is being worked in order to apply the excellent information processing capabilities of the brain in engineering. Functions, such as learning and memory, are achieved by lots of interactions though neurons' synapses in the brain.

There are four parts which are a cell body, an axon, a dendrite, and a cell membrane in a nerve cell [1]. There are some differences in the branch pattern of the dendrites and the form of the cell body on nerve cells. Moreover, the functions change in keeping with differences in these structures [1, 2]. In order to understand the function of the brain as a whole and engineering application of the function it is a considerable research objective to define the functions of each cell [3]. The nerve cell model was promoted by McCulloch and Pitts in 1943, and is well known as a fundamental neuron model in neural networks [4]. And, learning is applyed by adjusting the strengths of the synapses [5]. The perceptron was devised as a learning machine which uses this model, and a learning machine was built in one layer of neurons in some pattern classification [6]. D.Marr proposed the hypothesis which was about learning as in a perceptron in the cerebellum functions in the same way. However, multilayer of the network was run because that the one-layer perceptron cannot classify a linearly unseparated pattern [9]. Methods of multilayer of the networks have been devised and some learning rules have been put forward. One of the learning rules is the back-propagation learning rule, which has many applications and was valid [10].

In recent years, there are more detailed characteristics in the detailed analysis of nerve cells. Nonlinear inhibitive interaction happens among synapses has been acknowledged [8,11]. And, the local information processing is acted in dendrites [12]. The functions of the dendrites can be acted by logical AND, OR and NOT [9]. However, rather than a single form, there are lots of forms on dendrites [13, 14]. The interaction during synapse combination is also multiple and the nature of the interaction, the existence of an interaction, and so on, cannot be decided for each synapse [15]. Then again, in the early stages, redundant synapses and dendrites are found in the nervous system, and the unnecessary ones are soon filtered out and the necessary ones are strengthened and fixed, then form the ripened neural network function [16].

In this paper, in order to accelerate the learning speed, a kind of improvement is used into a new kind neuron model. It is expected that for biologists the improvement can offer some useful advises.

2. Neural Model.

2.1. McCulloch-Pitts Model.

In 1943, a neuron model was proposed by McCulloch and Pitts and then it is recognized as the fundamental one in the neural network field. In this model, the synapses are independents of each other regardless of the local action in a dendrite, and their functions are taken as mere weight.

2.2. The Koch-Poggio-Torre model [11].

Dendrites contain receptors that take in only one particular ion, and when the ion enters the receptor, the potential

Manuscript received October 5, 2012
Manuscript revised October 20, 2012

changes. The kind of reaction of the receptor determines whether it is an excitatory synapse or an inhibitory one. It is also known that an inhibitory nonlinear interaction among synapses occurs [8, 11]. A model of this interaction is shown in Figure 1. The model was proposed by Koch, Poggio, and Torre.

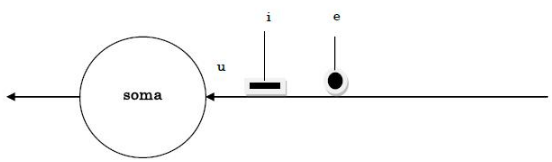


Figure 1. The interaction 1 between synapses.

2.3. Neuron model based on the dendrite mechanism.

Koch and colleagues proposed a neuron model and in this model the nonlinear interaction between the synapses on the dendrite of a retinal ganglion and the form of branch separation of a dendrite can be showed in forms of AND, OR, and NOT operations, and more advanced becomes possible. However, there is a different form in each dendrite. The interaction among the synapses is complex and there is no method to judge the existence and kind of an interaction at each synapse.And in addition, there is no middle alternation in the output of a cell,and the means is that is the output of a cell is either that the cell was excited or that it was not excited. Therefore, an output can be expressed with the value 0 or 1. Moreover, since the input will be in those two states, in which the excitation input signal enters or does not enter, it can also be expressed with value of either 0 or 1, like the output. So, in 2000, Tang et al. has made a neuron which is based on the conventional Koch model to realize the interaction among synapses and the dendrites and synapses` elimination and generation to form a kind of dendrites holding a special function[18][19]. Moreover, it also holds the movement direction selection function based on time difference [20]. The operation of a neuron is generalized as follows:

(1) The input and output are either 0 or 1.

(2) The form of an initial dendrite is arbitrary.

(3) There is an interaction among all synapses in the initial stages.

(4) The nonlinear interaction produced in a dendrite is expressed by a logical network.

(5) All matters such as the form of the dendrite and the type of interaction between synapses are decided by learning.

And, the connection states of the dendrites are shown in the Figure 2.

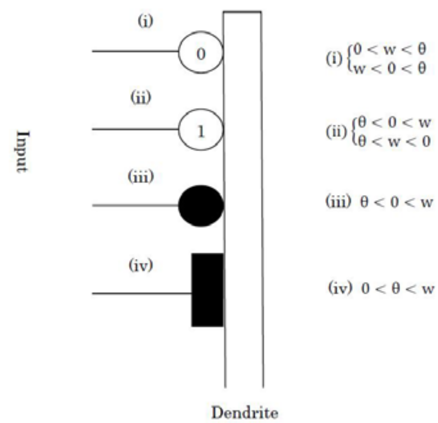


Figure 2. The kind of connection states.

3. The Improvement.

In this paper, this new model will uses back-propagation learning rule. Unlike the tradition back-propagation model, in this paper the learning rule advises back-propagates the signal from a cell body to a dendrite directly, and then decides the corrections of the connection function parameters. In recent years, it has been discovered that an impulse output also spreads retrogressively in real nerve cells, which is in accord with the presented views [17].

However, the speed of the learning is not so ideal. So, there are a lot of optimization which need to be proposed. One of the most popular optimizations is called variable learning rate. The method means a simple heuristic strategy to accelerate the convergence speed of the back-propagation algorithm with a batch update. If in previous the learning has decreased the total error function the idea behind the approach is to increase the magnitude of the learning rate. On the converse, the learning rate needs to be decreased if the error function has increased,. The algorithm can be summarized as follows [21]:

1. If the errors function over the entire training set has decreased. Increase the learning rate by multiplying it by a number$\rho > 1$. (In this paper, the number is 1.3)

If the error function has increased more than some set percentage ξ (typically a few percent, in this paper, is 1%), decreased the learning rate by multiplying it by a number $0 < \chi < 1$.(In this paper, the number is 0.7)

If the error function is increased less than the percentage ξ, the learning rate remains unchanged.

It can accelerate the convergence in the cases of smooth and slowly decreasing error functions through applying the variable learning rate to back-propagation. However, the algorithm can easily be lost in a local minimum of the error surface. In order to avoid this, the learning rate is not allowed to fall below a certain value.

4. Experimental Result.

As mentioned above, the algorithm improvement has been used into this new neural algorithm. And, an EX-OR problem with two inputs, a center symmetry problem are used to verify the improvement.

4.1. EX-OR Problem.

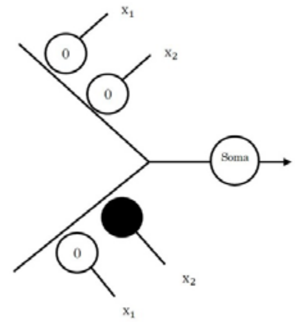
EX-OR Problem is a typical example of the liner unseperated problem.

As shown in the Figure 3, the shapes of synapses before and after learning are showed on the synaptic parameters. The shape in Figure 3(b) displays that learning was performed correctly. And the final shape will not change after the algorithm which is being improved.

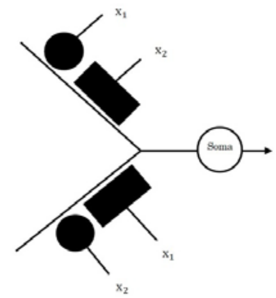
Figure 4 is the learning result. It shows that learning was performed, and that the correct answer was approached as the error decreased during learning. In Figure 4, the blue line meanted the performance of the introduced variable learning rate optimization, the red line meanted the performance of standard neural algorithm.

Although, the process of algorithm optimized by momentum and variable learning rate is not so smooth as the other, the result has gotten ideal result than the others. And the finally error is 0.022920, and less than the 0.023163 of standard one.

This shows that the result was not lost into a local minimum. It can be seen that during the simulations the likelihood of lapsing into a local minimum increased with the numbers of inputs and the final effective dendrites. The problem may be avoidable by providing a enough number of branches when the number of inputs increases. But, the effect is limited. And, although linear unseparated problems needed a multilayered neural network, the simulation result shows that this proposing model can learn the linear unseparation problem (EX-OR) with one layer of neurons, which would agree results of physiological experiments.



(a). BEFORE LEARNING



(b). AFTER LEARNING

Figure 3. The network for the ex-OR problem(input=2)before training

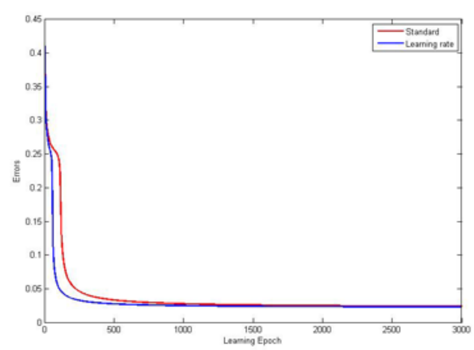


Figure 4. Sum-squared error during training.

4.2. The Center Symmetry Problem.

Table 1. shows a center symmetry detecting problem and it selected 8 AND gates for this problem. The connection shows the weights and the thresholds to the preudo-neuron were initialized randomly. Table.1 shows the set parameters. Simulation had an order from 0000 to 1111 for the 16 training sets of input data. For center symmetry

detecting problem, the learning performance is provided by Figure 5. It can be kown that the network has learned the problem correctly.

Table 1. Truth table of a center symmetry detecting problem

X_1X_2 / X_3X_4	00	01	11	10
00	1			
01				1
11			1	
10		1		

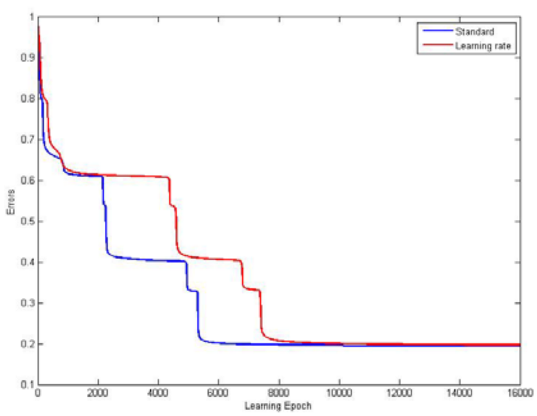
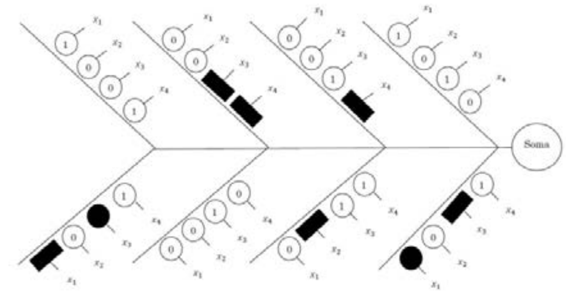


Figure 5. The learning performance for centre symmetry.

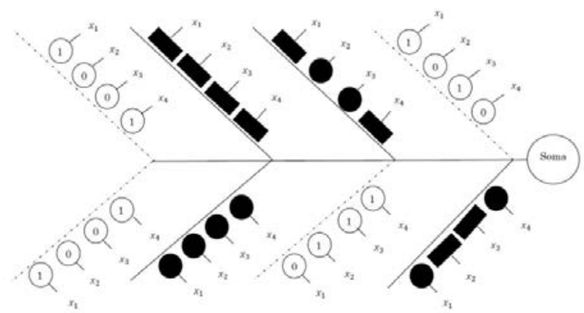
Compared with the ex-or 2-input problem,the centre symmetry problem is more complex. From FIGURE 5, after optimized the speed of the learning has been accelerated. The learning times is much more less than the standard by the variable learning rate optimization . And after optimized, the result finally decreased to 0.194280.

As can be seen from FIGURE 5, it shows that the algorithms has lapsed in a unstable point when the errors decreases to about 0.4. Owing to this situation, the learning rate will decrease when lapsing into the unstable points.

The optimized algorithm may not escape out of the unstable points zone because that the learning rate decreases continuously, So, in order to stop that the learning rate always decreased, a minimum learning rate has to be set.



(a) BEFORE LEARNING



(b)AFTER LEARNING

Figure 6. The network for the centre symmetry detecting problem before training

From Figure 6, it can be known the shapes of neuron for the centre symmetry detecting problem. Moreover, in this model, learning can provide information on what kinds of synapses exist and where they are located on a dendrite to satisfy the logical relation while a certain logical relation is supplied as a signal from outside. For physiologists to know the function of the whole dendrite, it will be very useful.

5. Conclusions.

In this paper, variable learning rate has been used into a new kind of neural algorithm with the ability of taking in function of a dendrite advanced information processing. The speed acceleration can be gotten during the learning process by the optimization. But the coefficient of the algorithm has to be decided on the situation. Moreover, during the learning, the algorithm has usually lapsed into the local minimum more easily, when the simulation turns more complex. So, the farther research is to solve the local minimum problem.

REFERENCES

[1] Fukushima K. Nerve circuits and information processing. Asakura Bookstore.

[2] Matsumoto G, Otsu N. The information processing that nerve cells perform, and its mechanism. Baifukan.
[3] Matsumoto G, Otsu N. The information processing that brain and nerve systems perform, and their models. Baifukan.
[4] McCulloch WS, Pitts W. A logical calculus of the ideas immanent in nervous activity. Bull Math Biophys 1943;4:11-133
[5] Hebb DO. The organization of behavior. Wiley; 1949.
[6] Rosenblatt F. Principles of neurodynamics. Spartan; 1961.
[7] Ishii N, Tsukada M. Neural computing: Theory and practice. Morikita Publishing.
[8] Kitajima T, Hara K. A bipolar model of the nerve cell—Dynamic characteristics of synaptic film. Trans IEICE 1987;J70-D:818-826.
[9] Minskey M, Papert S.Perception: An essay in computational geometry. MIT Press; 1943
[10] Rumelhart DE, Hinton GE, Williams RJ. Learning representations by back-propagation errors. Nature 1986;323:533-536.
[11] Koch C, Poggio T, Torre V. A functional interpretation of dendritic morphology. Philos Trans R Soc London B 298:227-264.
[12] Kagawa Y. Dictionary of bionics terminology. Japanese Standards Association.
[13] Boycott BB, Wassle H. The morphological types of ganglion cells of the retina of the domestic cat. J Physiol Landon 1974;240:397-419.
[14] Boycott BB, Peichl L, Wassle H. Morphological types of horizontal cells in the retina of the domestic cat. Proc R Soc London Ser B 1978;203:229-245.
[15] Cleland BG, Levick WR. Properties of rarely encountered types of ganglion cells in cat retina and overall classification. J Physiol London 1974;240:457-492.
[16] Komatu Y, Fujii K, Nakajima S, Umetani K, Toyama K. Electrophysiological and morphological correlates in the development of visual cortical circuitry in infant kittens. Dev Brain Res 1985;11:305-309.
[17] Japan Neural Network Society Magazine 1985;5:110-115.
[18] Zheng Tang, Hiroki Tamura, Makoto Kuratu, Okihiko Ishizuka, and Kichi Tanno, A Model of the Neruon based on Dendrite Mechanisms. Denshi Joho Tsushin Gakkai Ronbunshi, Vol. J83-A, No. 5, May 2000, p0p.486-498.
[19] Zheng Tang, Hiroki Tamura, Okihiko Ishizuka, Koichi Tanno. A Neuron Model with Interaction among Synapses. T.IEE japan, Vol. 120-C,No. 7,2000,pp.1012-1019.
[20] Hiroki TAMURA, Zheng TANG, and Masahiro ISHII. A Model of Neuron with Dendrite Mechanisms is Learning to the Movement Direction Selection Function. Denshi Joho Tsushin Gakkai Ronbunshi, A Vol.J84-A, No. 8 August 2001, pp.486-498.
[21] T. P. Vogl, J.K. Mangis, A.K. Zigler, W.T. Zink, and D.L. Alkon, “Accelerating the Convergence of the Backpropagation Method,” Bilolgical Cybernetics, vol. 59, 1988, pp. 257-63.

Zijun Sha received the B.S. degree from Hunan University of science and technology, Hunan, China in 2008 and the M.S. degree from University of Toyama, Toyama, Japan in 2012. Now, he is working toward the D.E. degree at University of Toyama, Toyama, Japan. His main research interests are multiple-valued logic, artificial neural networks and information security.

Lin Hu received the B.S. degree from Central South University of Forestry and Technology, Hunan, China in 2009 and the M.S. degree from Central South University of Forestry and Technology, Hunan, China in 2012. Now, she is working toward the D.E. degree at University of Toyama, Toyama, Japan. His main research interests are artificial neural networks, economic analysis system and eco-economic system.