2010 International Symposium on Nonlinear Theory and its Applications
NOLTA2010, Krakow, Poland, September 5-8, 2010

Multi-Layer Perceptron Having Neuro-Glia Network

Chihiro Ikuta, Yoko Uwate and Yoshifumi Nishio

Dept. of Electrical and Electronic Engineering, Tokushima University
2–1 Minami–Josanjima, Tokushima, Japan
Email: {ikuta, uwate, nishio}@ee.tokushima-u.ac.jp

Abstract—Glial cells have known to exist in the brain. They watch brain's state by signal transmitting each other. Thus, many researchers have taken notice to relationships of glial cells and neurons.

In this paper, we propose a Multi-Layer Perceptron (MLP) having neuro-glia network. We give biological features of glial cells to the standard MLP. We consider that we can obtain good performance by glial cells and neurons' having the interrelation. By computer simulations for solving Two Spiral Problem (TSP), we confirm that the proposed MLP having neuro-glia network gains better performance than the conventional MLP.

1. Introduction

Back Propagation (BP) was introduced by Rumelhart in 1986 [1]. BP is used for learning algorithm of MLP and the error propagates backwards in the network. However, the solution of the network often falls into the local minimum, because BP uses the steepest descent method for the learning process. In order to avoid this problem, some methods to release the solution from the local minimum are required.

From recent research, we came to know that glial cells swap many kinds of ion with each other. And this work support a brain's systems [2][3]. Above all, it is important to change Ca^{2+} density by glial cells. Changing of Ca^{2+} density propagates in the brain like a wave. This wave influences neurons and other glial cells then their works becoming to active. This phenomenon was confirmed that blood currents of brain increased in the part of changing of Ca^{2+} density. In order to realize such biological system for the artificial neural network, we proposed a glial network connected to MLP [4]. In this network, glial cell connected to neighborhood glial cells and influence of glial cell propagates. We showed that this network gave good influence to MLP learning.

In this study, we propose MLP having neuro-glia network. Function of watching neuron's output is added to glial cell based on the glial network connected to MLP. Each glial cell makes independent random oscillation and this oscillation propagates into the glial network. Glial cells are watching neuron unit's output and glial units change amplitude of oscillation by neuron unit's output. Glial oscillations are generated in this way affects neuron unit's threshold. We apply the proposed method to the Two-Spiral Problem [5] and confirm the efficiency by computer simulations.

2. Multi-Layer Perceptron with Glial Network

In this section, we explain a concept of MLP having neuro-glia network.

2.1. Multi-Layer Perceptron

MLP is the most famous feed forward neural network. This network is used for pattern recognition, pattern classification, and other tasks. MLP has some layers, it has mainly input layer, hidden layer, and output layer. Generally, it is known that MLP can solve a more difficult task if the number of layer or neuron is increased. We consider MLP which is composed of four layers (one input layer, two hidden layers and one output layer), and MLP has the glial network in the second layer of the hidden layer. The proposed MLP having neuro-glia network structure (connected 2-20-40-1) is shown in Fig. 1.

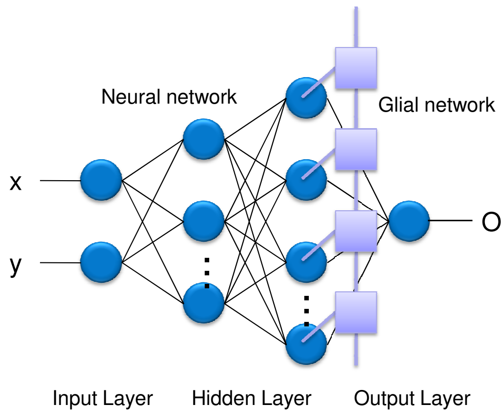

Figure 1: MLP with glial network.

2.2. Neuron Updating Rule

The updating rule of neurons in the input layer, the first hidden layer and the output layer is described by Eq. (1) which is conventional updating rule.

$$x_i(t+1) = f\left(\sum_{j=1}^{n} w_{ij}(t)x_j(t) - \theta_i(t)\right) \quad (1)$$

In the MLP having neuro-glia network, random oscillation is added to neurons in the second hidden layer. This neuron's updating rule is following as Eq. (2).

$$x_i(t+1) \quad = \quad f\left(\sum_{j=1}^{n} w_{ij}(t)x_j(t) - \theta_i(t) + \alpha\Psi_i(t)\right), \quad (2)$$

where x : input or output, w : weight parameter, θ : threshold, Ψ : oscillation of glial network, α : amplitude of noise and f : output function. And we use sigmoid for the output function as Eq. (3).

$$f(a) \quad = \quad \frac{1}{1+e^{-a}} \qquad (3)$$

2.3. Generating Oscillation by MLP with Glial Network

In the biological neural network, it is known that the glial cells affect to the neighbor neurons over a wide range by propagating in the network [6]. We considered that the glial network gives good influence to learning of neural networks. We proposed glial network connected to MLP [4] as shown Eq. 4.

$$\Psi_i(t) \quad = \quad \sum_{k=-m}^{m} \beta^{|k|}\psi_{i+k}(t-|k|), \qquad (4)$$

where ψ : uniformed random noise, β : attenuation parameter and k : the propagating range in the glial network.

2.4. Generating Oscillation by MLP Having Neuro-Glia Network

In the glial network has function of Eq. 4. Moreover, we notice glial cell's function which cool out exciting neurons [3]. In order to realize phenomena, we add random oscillation to neurons by using Eq. (5).

$$\Psi_i(t) = \sum_{k=-m}^{m} \beta^{|k|}\psi_{i+k}(t-|k|)\left\{0.5 - O_{i+k}(t-1)\right\}, \qquad (5)$$

O is the second hidden layer's neuron output. Glial cells are watching connecting neuron's output. They decide next output's amplitude by k range existing glial cells' output and before output of connecting neuron's. Random oscillation produced by the uniform random function propagates in the glial network and watching neuron's output as shown in Fig. 2.

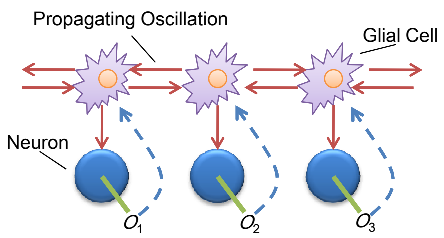

Figure 2: MLP having neuro-glia network.

3. Oscillation of Glial Network

In this section, we show that the conceptual glial network. We show three kinds of generating oscillation by glial cells.

3.1. Generating by Glial Cells

Glial cells generate uniformed random oscillation, thus all glial cells have different oscillation. We shows that neighborhood glial cells generate oscillation in the Fig. 3. We mark glial output of each time with dot and we plot moving average lines. In this figure, each moving average line has different curve. We consider that two glial cells do not have a relationship in this case.

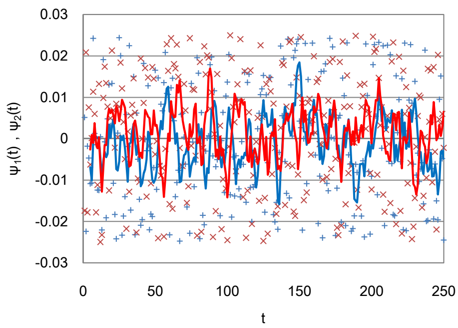

Figure 3: Uniformed random oscillation of two glial cells.

3.2. Generating by MLP with Glial Network

Figure 4 is a glial cells' output when generating oscillations pass in the glial network. The oscillations of Fig. 4 become biased oscillation by passing in the glial network. Moreover, moving average line of each changing oscillation become very similar curve. Because neighbor glial cells are influenced each other by passing in the glial network.

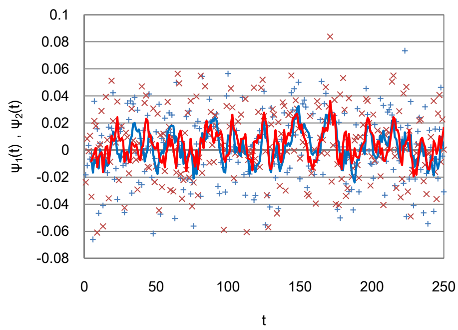

Figure 4: Propagating random oscillation in the glial network.

3.3. Generating by MLP Having Neuro-Glia Network

Figure 5 show that two kinds of oscillation using to MLP having neuro-glia network. These oscillations are similar to oscillation of Fig. 4. However, these vibrating positions are different in Fig. 5, because amplitude of oscillation is influenced by the second hidden layer neuron in the MLP having neuro-glia network. We consider that using MLP having neuro-glia network can give to each neuron unit for the adequate oscillation.

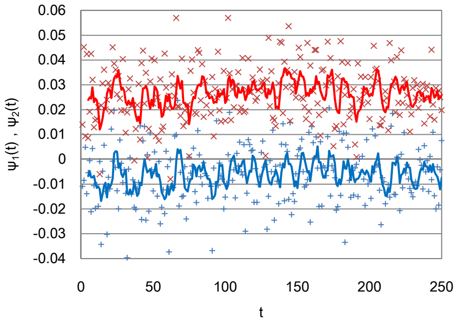

Figure 5: Propagating random oscillation in the network.

4. Simulations

In this section, the difference in the performance of our MLP; the MLP having neuro-glia network, the MLP with glial network, the MLP with random noise and the conventional MLP is compared. We confirm that the MLP having neuro-glia network is better result than others and our MLP can solve high nonlinear problem.

We apply the proposed network for solving TSP [5]. MLP learns to each point of two spirals and MLP learns by using the standard BP algorithm.

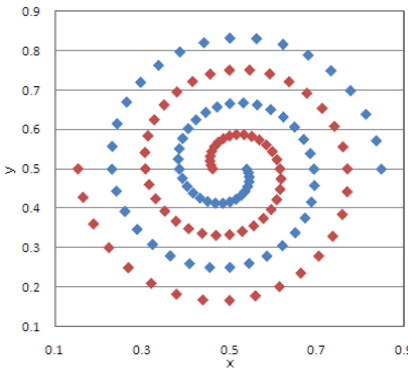

Figure 6: Two-spiral problem.

4.1. Simulation Results

We use four types of MLP structures for comparison as follows, MLP having neuro-glia network, glial network connected to MLP, MLP adding random noise and conventional MLP. In the network of MLP adding random noise, the uniform random noise is added to the threshold of the neurons in the second hidden layer. The error function is defined as Eq. (6).

$$\begin{cases} True : |t_i - O_i| <= 0.2 \\ False : |t_i - O_i| > 0.2 \end{cases} \tag{6}$$

$$E = \frac{\sum(False)}{N}, \tag{7}$$

where N denotes the number of learning point.

4.1.1. Learning 98 Points

We prepare 98 data of two spirals. The number of learning points is fixed as 500000. We use propagating range $m = 5$, attenuation parameter $\beta = 0.8$ and learning coefficient $\eta = 0.05$.

Table 1 shows simulation result of each MLP learn to 98 points. "Err = 0%" means the number of times which has memorized all the points. "Err< 10%" means the number of *False* is lower 10%. "AVG of Err" means each MLP learn the average of an error when each MLP finishes learning.

From this table, we can see that the MLP having neuro-glia network shows the highest result of all conditions. In this result, we consider that MLP MLP having neuro-glia network effective in the learning of two-spiral.

Table 1: Changing average of error as learning to 98 points by using each MLP

	Err = 0%	Err < 10%	AVG
Neuro-Glia	60%	83%	4.9%
Glial Network	35%	70%	8.4%
Random Noise	30%	75%	9.3%
Conventional	45%	80%	7.6%

Figure 7 is typical result of learning 98 points. In this result, MLP having neuro-glia network is hardly influenced by local minima. MLP having neuro-glia network and MLP with glial network are better than the conventional MLP and the MLP with random noise. Moreover, MLP having neuro-glia network is the fastest converging of learning to two-spiral.

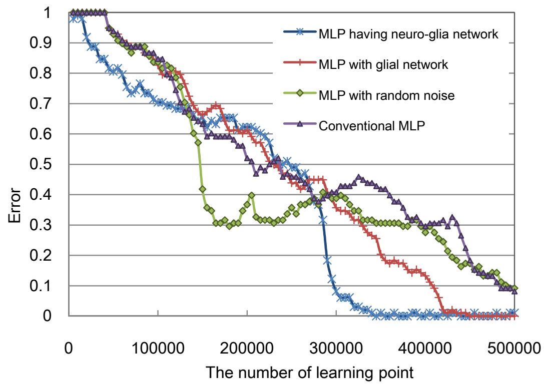

Figure 7: Error curve by four MLP networks for 98 points.

4.1.2. Learning 130 Points

We prepare 130 data of two spirals. The number of learning points is fixed as 500000. We use propagating range $m = 5$, attenuation parameter $\beta = 0.8$ and learning coefficient $\eta = 0.05$.

Table 2 shows result of each MLP learn to 130 points. MLP having neuro-glia network shows the highest percentage for all conditions. We can get a good result in the 98 points learning. However, we can see that clearly MLP having neuro-glia network is competitive solving high nonlinear problem in this result. We consider that this performance which finds nearly optimal solution is one of important characteristics for solving high nonlinear problems.

Table 2: Performance of learning to 130 points by using each MLP

	Err = 0%	Err < 10%	AVG
Neuro-Glia	27%	60%	13%
Glial Network	3.3%	30%	24%
Random Noise	0.0%	20%	26%
Conventional	6.7%	17%	29%

Figure 8 is typical result of learning 130 points. MLP with glial network, MLP adding random noise and conventional MLP are fall into the local minimum at about 250000 learning point. However MLP having neuro-glia network escape out from the local minima. In this figure, we can see learning convergence of MLP MLP having neuro-glia network is very fast.

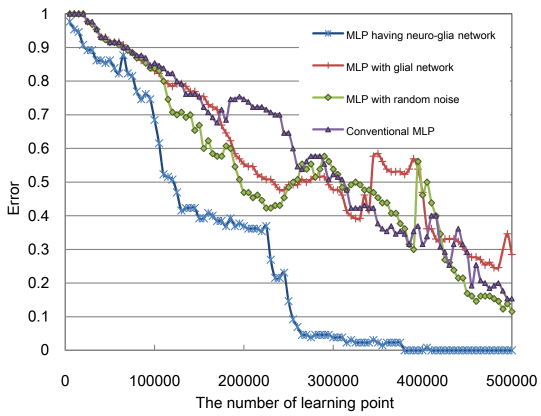

Figure 8: Error curve by four MLP networks for 130 points.

4.2. Classification of Two Spirals

In this section, MLP having neuro-glia network classifies the two spirals. The MLP having neuro-glia network learns to two spiral points. After learning, we puts points $(0.1 \leq x \leq 0.9, 0.1 \leq y \leq 0.9)$ changing little by little in MLP's input layer. Figure 9 show the simulation result. In this case, all the learning point is memorized at 500000 learning points. When MLP's output is nearer 1, this point is classified to red spiral. And when MLP's output is nearer 0, this point is classified to blue spiral. This result shows that the MLP having neuro-glia network can solve TSP.

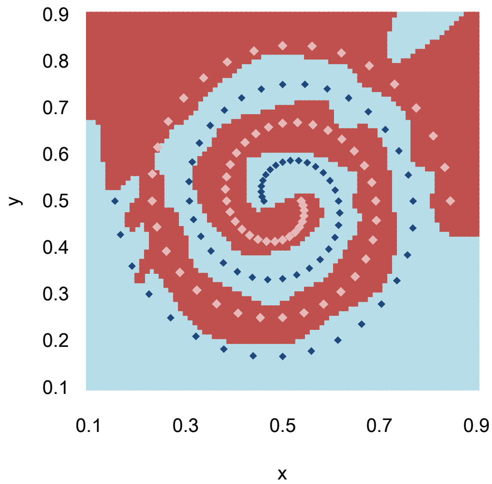

Figure 9: Classification of two spirals by MLP having neuro-glia network.

5. Conclusions

In our study, we have proposed MLP having neuro-glia network. This network gave random oscillations to the second hidden layer's neuron. The amplitude of oscillation changes by neuron's output and this random oscillation propagates to other neurons. We confirmed that MLP having neuro-glia network gains better performance than the glial network connected to MLP, MLP adding random noise and the conventional for solving TSP.

References

[1] D.E. Rumelhart, G.E. Hinton and R.J. Williams, "Learning Representations by Back-Propagating Errors," *Nature*, vol. 323-9, pp. 533-536, 1986.

[2] P.G. Haydon, "Glia: Listening and Talking to the Synapse," *Nature Reviews Neuroscience*, vol. 2, pp. 844-847, 2001.

[3] S. Ozawa, "Role of Glutamate Transporters in Excitatory Synapses in Cerebellar Purkinje Cells," *Brain and Nerve*, vol. 59, pp. 669-676, 2007.

[4] C. Ikuta, Y. Uwate and Y. Nishio, "Chaos Glial Network Connected to Multi-Layer Perceptron for Solving Two-Spiral Problem," *Proc. ISCAS'10*, May 2010. (Accepted)

[5] J.R. Alvarez-Sanchez, "Injecting knowledge into the Solution of the Two-Spiral Problem," *Neural Computing & Applications*, vol. 8, pp. 265-272, 1999.

[6] S. Koizumi, M. Tsuda, Y. Shigemoto-Nogami and K. Inoue, "Dynamic Inhibition of Excitatory Synaptic Transmission by Astrocyte-Derived ATP in Hippocampal Cultures," *Proc. National Academy of Science of the U. S. A*, vol. 100, pp. 11028-11028, 2003.