IWVCC'08
@ Xian
Nov. 1-3, 2008

Feedforward Neural Networks with Fluctuation of Structure for Chaotic Time Series Learning

Yoko Uwate†, Yoshifumi Nishio‡ and Ruedi Stoop†

†Institute of Neuroinfomatics, University / ETH Zurich,
Winterthurerstrasse 190, 8057 Zurich, Switzerland,
email: {yu001, ruedi}@ini.phys.ethz.ch
‡Dept. of Electrical and Electronic Engineering, Tokushima University
2-1 Minami-Josanjima, Tokushima, Japan
Email: {uwate, nishio}@ee.tokushima-u.ac.jp

Abstract

The fluctuation observed in the real brain has attracted attentions of many researchers. We consider that the fluctuation plays an important role for the higher brain function. In this study, the neural network with fluctuation of structure is investigated. By computer simulations, we confirm that the neural network with fluctuation of structure gains better performance than the conventional neural network for generalization ability.

1. Introduction

Studies on the human brain have been carried out actively on various levels. Many modelings of the human brain with the visual or the audio sensation are reported [1]. Recently, the fluctuation observed in the real brain has attracted attentions of many researchers. The fluctuation in the brain is one of the living evidence, and the amplitude of the fluctuation decreases or disappears when human becomes sick. It can be said that the fluctuation has some relationship to the higher brain functions. We consider that it is very important to apply this fluctuation factors to artificial neural network model.

In our previous research, we have investigated the ability of the feedforward neural network when the shape of the sigmoid function of neurons is changed according to the logistic map equation [2]. We have compared its effect with the simulated annealing and have confirmed that change sometimes led the state of the network to better solutions.

We consider that the neural network with fluctuation of structure has many possibilities for the information processing. We assume that the real brain operates well when the neural circuits are shaken by some kinds of influences. In this study, the learning ability of the network with fluctuation of structure is investigated. Two methods giving fluctuation are proposed to investigate the most efficient method providing fluctuations; coherent fluctuation method and noncoherent fluctuation method. For comparison, we investigate the network shaken by additional noise signals. By computer simulations, we confirm that the proposed network with fluctuation of structure can find better solutions than the network with additional noise signals and the conventional network for generalization ability.

2. Neural Network with Fluctuation of Structure

The standard BP learning algorithm was introduced in [3]. The effectiveness of the BP learning has been confirmed in pattern recognition, system control, signal processing and so on [4]. In this study, the learning ability and generalization ability of the network with fluctuation of structure are investigated by using the batch BP learning algorithm. .

The sigmoid function has an effect on modifying connection weights and it is very important for BP learning. This function is an S shaped monotonic increasing function that has the general form as following equation:

$$f(x) = \frac{1}{1 + e^{-\varepsilon x}} \tag{1}$$

where ε is a constant that determines the steepness of the S shaped curve. Some curves of the function for different values of ε are illustrated in Fig. 1.

We apply the sequence generated by the uniform random function to ε of the sigmoid function after the following linear transform to set the average as 1.0 and control the amplitude.

$$\varepsilon(t) = A(random() - 0.5) + 1.0 \tag{2}$$

where $random()$ means the function producing uniform random value from 0.0 to 1.0 and, A corresponds to the range of ε. In this simulation, we set from $A = 0.02$ to 0.6.

Next, two fluctuation methods are proposed as follows;

2.1. Coherent Fluctuation Method

The gradient of the sigmoid function fluctuates with a coherent value for all neurons in the hidden layer at every updating.

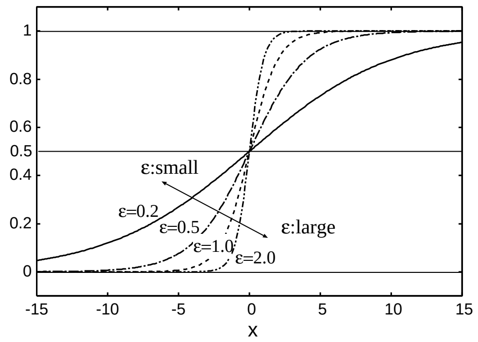
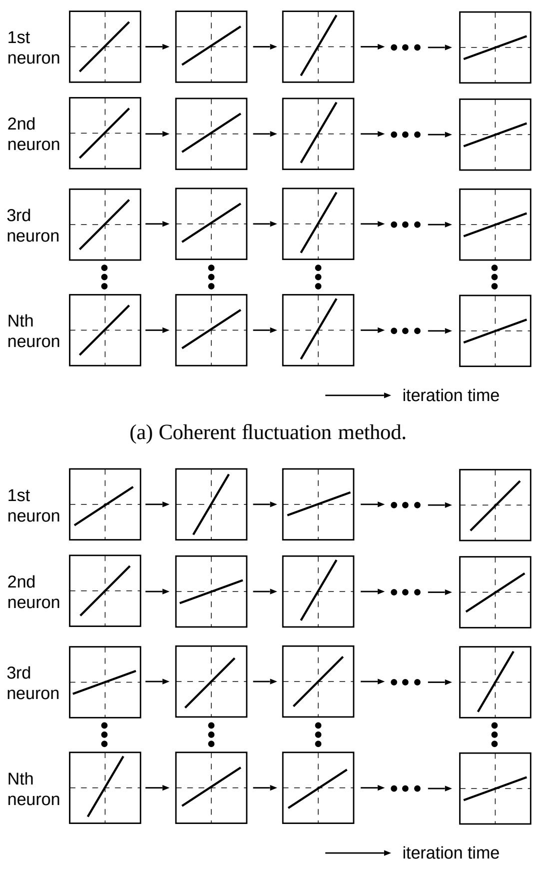


Figure 1: Sigmoid function and fluctuation ε.

2.2. Noncoherent Fluctuation Method

The gradient of the sigmoid function fluctuates with noncoherent different values for all neurons in the hidden layer at every updating by giving different initial conditions to the logistic map.

A conceptual figure of the two fluctuation methods are shown in Fig. 2.



(a) Coherent fluctuation method.

(b) Noncoherent fluctuation method.

Figure 2: A conceptual figure of two fluctuation methods.

3. Neural Network with Additional Noise Signals

For comparison, we investigate the network shaken by additional noise signals. We consider that the uniform noise (generating value form 0.0 to 1.0) is injected to neurons in the hidden layer. Figure 3 shows a conceptual neuron model for this neural network, where β limits the amplitude of the injected signals.

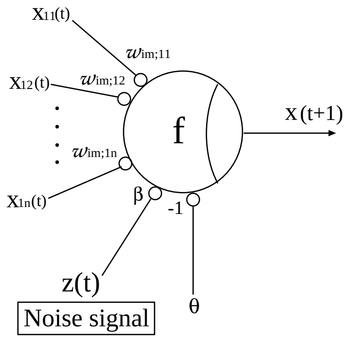


Figure 3: A neuron model with additional noise signals.

4. Simulated Results

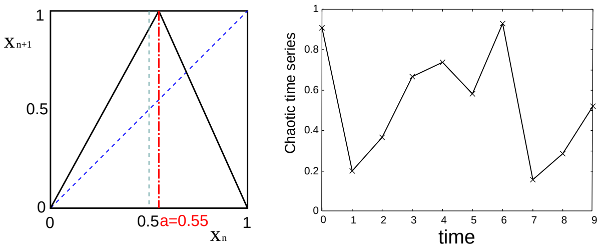
We consider that the neural network with fluctuation of structure has a generalization ability as well as a learning ability by virtue of its inefficient learning process. In this section, we investigate the learning ability and the generalization ability of our proposed networks. We consider the learning of the structure of the skew tent map by training the network to output the same time series as the input time series produced by the skew tent map.

The skew tent map and an example of time series are shown in Fig. 4. The length of chaotic time series is set to 10 and the number of learning patterns is set to 10. When the network learns 10 lengths of time series, 10 nodes are prepared in the input and the output layers. Each data is inputted to each node in the input layer. We carried out the BP learning by using the following parameters. The parameter of the inertia rate is fixed as $\eta = 0.05$ and the initial values of the weights are given between -1.0 and 1.0 at random. The learning time is set to 10000, and 12 neurons are prepared in the hidden layer. The gradient of the sigmoid function of the conventional network is fixed as $\varepsilon = 1.0$.

4.1. Learning Ability

First, we investigate the learning efficiency as the average of the total error between the output and the desired target, when the network structure of the hidden layer is changed. The “Average Error E_{ave}” for this learning example is defined by the following equation.

$$E_{ave} = \frac{1}{P}\frac{1}{N}\sum_{p=1}^{P}\left\{\frac{1}{2}\sum_{i=1}^{N}(t_{pi} - o_{pi})^2\right\}. \quad (3)$$



(a) Skew tent map. (b) Time series.

Figure 4: Skew tent map.

The simulated result of networks with fluctuation of structure is shown in Fig. 5. When the value of fluctuation A is 0.02, the methods giving fluctuation and the conventional network obtain similar performance. While the fluctuation methods becomes worse by increasing A. We compare the performance between the coherent and noncoherent fluctuation methods. When A is small value, both fluctuation methods show similar performance of E_{ave}. The noncoherent fluctuation method can keep the small error (E_{ave}) for learning data with increasing A.

Figure 6 shows the simulated result of network with noise signals. The horizontal axis is β which means the amplitude of injected noise and the vertical axis is E_{ave} for learning data. From this figure, we can confirm the noise signals method shows similar performance to the conventional network when β is set to 0.02. However, the performance of noise signals method is getting worse by increasing β value.

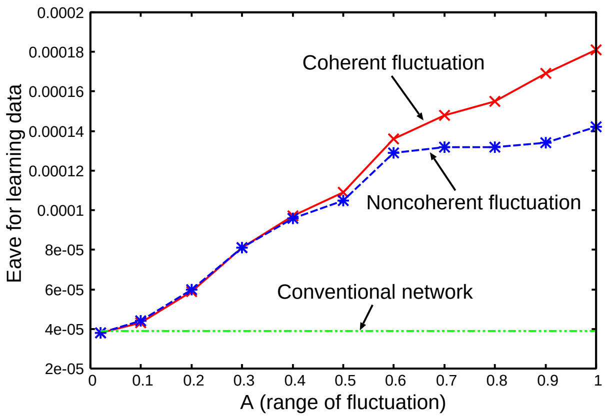


Figure 5: Network performance E_{ave} by using coherent and noncoherent fluctuation methods for learning data.

The simulated results of the fluctuation methods are summarized in Tab. 1. In this table, the average of 10 E_{ave} for different initial conditions of the weights between all layers is shown. From this table, we can confirm that E_{ave} of the proposed network with the fluctuation methods and the conventional network is similar. We consider that the performance of fluctuation methods are not so difference to the conventional network for learning ability.

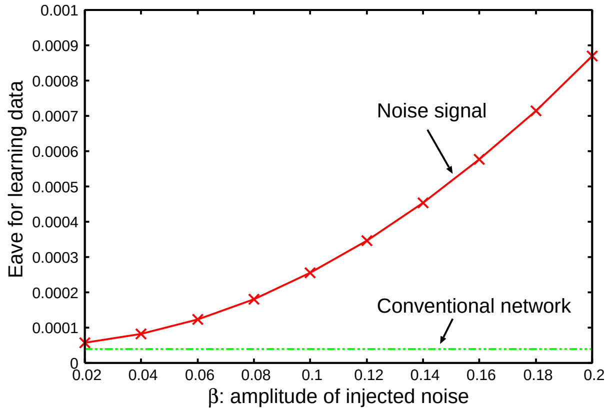


Figure 6: Network performance E_{ave} by using noise signals methods for learning data.

4.2. Generalization Ability

Next, we investigate a generalization ability of the proposed neural networks. After the above learning of 10 patterns of the time series, we input an unknown chaotic time series generated the same skew tent map as an input pattern.

The simulation result for unknown chaotic time series after learning of the network with fluctuation of structure is shown in Fig. 7. The horizontal axis is the range of fluctuation A and the vertical axis is E_{ave} for 10 different unknown input data. From this figure, the networks with fluctuation of structure gain better performance than the conventional network in every range of A. We compare the performance between the coherent and the noncoherent fluctuation methods. When A is small value, the coherent fluctuation method is better than the noncoherent fluctuation method. However, the coherent and the noncoherent fluctuation methods are getting similar performance with increasing A.

Figure 8 shows the simulated result of network with noise signals. From this figure, we can confirm the noise signals method obtains almost similar performance to the conventional network.

The simulated results of the fluctuation methods for unknown data are summarized in Tab. 2. In the table, the average of 10 E_{ave} for different initial conditions of the weights between all layers is shown. From this table, E_{ave} of the proposed network with the fluctuation of structure are small. However, the noise signal method and the conventional network do not operate well, because the average error E_{ave} of the network is large. We can see that the network with fluctuation of structure gain better performance on generalization ability than the noise signal method and the conventional network.

From these results, we consider that the fluctuation of structure are very important role to learn some characteristics or features of the given data.

Before concluding this paper, we have to say that the net-

Table 1: Learning ability for learning data.

	Coherent fluctuation	Noncoherent fluctuation	Noise signals	Conventional network
E_{ave}	0.000038	0.000038	0.000057	0.000039

Table 2: Generalization ability for unknown data.

	Coherent fluctuation	Noncoherent fluctuation	Noise signals	Conventional network
E_{ave}	0.017382	0.013418	0.025440	0.025735

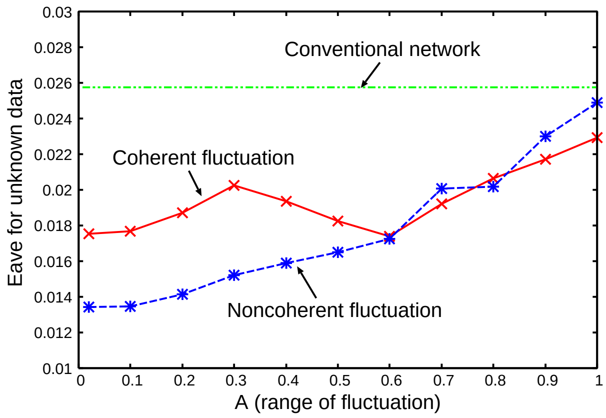


Figure 7: Network performance E_{ave} by using coherent and noncoherent fluctuation methods for unknown data.

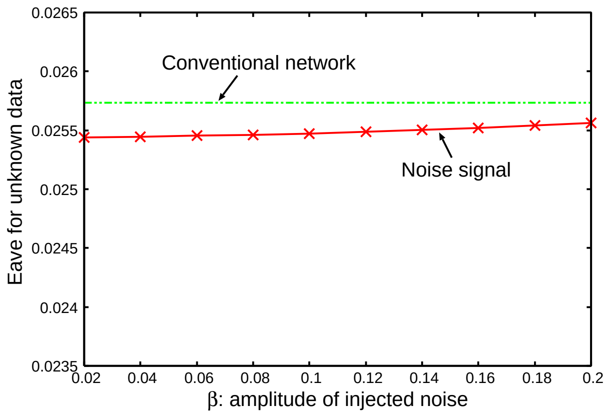


Figure 8: Network performance E_{ave} by using noise signals methods for unknown data.

work does not learn to output the same data as input data directly. In our past study, we have investigated the performance of the network when a random data is inputted as unknown data after the learning. We confirmed the network does not output the same data as input data directly [8].

5. Conclusions

In this study, we investigated learning ability and generalization ability of neural network with fluctuation of structure for back propagation learning. By computer simulations, we confirmed that the proposed network with fluctuation can produce the time series for unknown input data. We can see that the network with fluctuation of structure operates more effectively than the network with additional noise signals for generalization ability. Furthermore, the noncoherent fluctuation gains better performance than the coherent fluctuation for unknown input data. The detailed investigation of the effect of different fluctuation methods is our future work.

References

[1] R. P. Broussard, S. K. Rogers, M. E. Oxley and G. L. Tarr, "Physiologically Motivated Image Fusion for Object Detection Using a Pulse Coupled Neural Network," IEEE Transactions on Neural Networks, 10 (3), pp. 554-563, 1999.

[2] Y. Uwate and Y. Nishio, "Chaotically Oscillating Sigmoid Function in Feedforward Neural Network," *Proc. NOLTA'06*, pp. 215-218, Sep. 2006.

[3] D.E. Rumelhart, G.E. Hinton and R.J. Williams, "Learning Internal Representations by Error Propagation," *Parallel Distributed Processing*, vol. 1, MIT Press, MA, pp. 318-362, 1986.

[4] Y. LeCun, B. Boser, J.S. Denker, et al, "Handwritten Digit Recognition with a Back-Propagation Network," *Advances in Neural Information Processing Systems*, vol. 2, MIT Press, MA, 1990.

[5] T. Ueta, Y. Nishio and T. Kawabe, "Comparison between Chaotic Noise and Burst Noise on Solving Ability of Hopfield Neural Networks," *Proc. NOLTA'97*, vol. 1, pp. 409-412, Nov. 1997.

[6] Y. Uwate, Y. Nishio and A. Ushida, "Markov Chain Modeling of Intermittency Chaos and its Application to Hopfield NN," *IEICE Transactions on Fundamentals*, vol. E87-A, no. 4, pp. 774-779, Apr. 2004.

[7] Y. Uwate, Y. Nishio, T. Ueta, T. Kawabe and T. Ikeguchi, "Performance of Chaos and Burst Noise Injected to the Hopfield NN for Quadratic Assignment Problems," *IEICE Transactions on Fundamentals*, vol. E87-A, no. 4, pp. 937-943, Apr. 2004.

[8] Y. Uwate and Y. Nishio, "Performance of Affordable Neural Network for Back Propagation Learning," *IEICE Transactions on Fundamentals*, vol. E89-A, no. 9, pp. 2374-2380, Sep. 2006.