A study on the Convergence Condition of Chaotic Dynamic Neural Networks

Sang-Hee Kim* and Hua O. Wang**

*School of Electronic Engineering, Kumoh National Instute of Technology
Gumi Gyungbuk, Korea 730-701
** Dept. of Aerospace/Mechanical Engineering, Boston University
Boston, MA 02215, USA

Abstract

This paper analyzes on the chaos characteristics of the chaotic neural networks and presents the convergence condition. Although the transient chaos of neural network sould be beneficial to overcome the local minimum problem and speed up the learning, the permanent chaotic response gives adverse effect on optimization problems and makes neural network unstable in general. This paper investigates the dynamic characteristics of the chaotic neural networks with the chaotic dynamic neuron, and presents the convergence condition for stabilizing the chaotic neural networks.

Key Words : Chaotic dynamic neural networks. Convergence condition, Dynamic characteristics

1. Introduction

There have been many studies on biological neuron models based on biological experiment results from mammalian brains to primitive creature(squid giant axons)[1]. The major reason for the studies is to simulate biological neural network for archiving break-through on the performance of artificial neural networks. According to many extensive researches of neurology, some researchers found that biological neurons have limit cycle and chaotic activities for permanent nature or transient period[2]. Artificial neural networks were modeled from biological neuron for mimicking the efficient behavior of human neural networks.

The chaotic responses of biological neurons have been modeled quantitatively by many researchers. Chaotic neural networks consist of bunch of chaotic neurons with connection weights. The primitive model was the Hodgkin-Huxley equation. Caianiello[2,3] and Nagumo-Sato[4] modified this primitive model for making chaotic neural networks. Aihara et al. made a discrete time model with continuous output, and applied this model to make chaotic neural networks[5]. They gave some possibilities of application in solving optimization problems, which were traveling salesman problem(TSP). The effects of chaotic response have not verified yet in analytical methods. The chaotic characteristics of neuron generally give adverse effect on optimization problems, but the transient chaos of

neuron model could be beneficial to overcome the local minimum problem and speed up the learning process. Aihara proposed that the transient chaotic characteristics of neuron could be helpful for global optimization.[6,8]

Even though some modifications were made, the drawback of those previously proposed chaotic neural networks are still complicate to apply in artificial neural network, and need more dynamic characteristic in neuron itself and learning algorithm.[7] In previous paper, we presented a novel modified chaotic neural network for simplifying structure and enforcing dynamic characteristics, and applied this network to system identification and neuro control.[9]

In this paper, the traditional chaotic neuron model is studied for analyzing the dynamic characteristics. The modified chaotic dynamic neuron is presented for making the chaotic dynamic neural networks. The dynamic characteristics of proposed networks are investigated to constrain the chaotic characteristics, and the convergence conditions are presented. This convergence condition of chaotic neurons applied to chaotic neural networks for stabilizing the chaotic neural networks.

2. Chaotic Neuron Models

2.1 Traditional Chaotic Neuron Model

The traditional model of the chaotic neuron was suggested by Caianiello.[1] In the inside neuron, the past excitation inputs gives the inhibitory influence for refractory periods. This inhibitory influence of past firing decreases exponentially with

time. Under this assumption, the behavior of neuron was modeled as a nonlinear differential equation as equation (1).

$$x(k+1) = u(A(k) - \alpha \sum_{r=0}^{t} K^r x(k-r) - \theta) \qquad (1)$$

where $x(k+1)$ is the output of a neuron at discrete time $k+1$, and $x(k)$ takes either 0 or 1. $u(\cdot)$ is a unit step function, $A(k)$ is the strength of the activation input at discrete time k, K^r is the damping factor of the refractoriness having values between 0 and 1. The constant α is a positive parameter, and θ is the threshold of a chaotic neuron. If the inertial state of a neuron i at time k is assumed as follows[3],

$$y(k+1) = A(k) - \alpha \cdot \sum_{r=0}^{t} K^r x(k-r) - \theta \qquad (2)$$

then eq. (1) can be expressed as eqs. (3) and (5).

$$y(k+1) = K^r y(k) - \alpha \cdot u[y(k)] + a(k) \qquad (3)$$

where $a(t)$, a bifurcation parameter, is defined as

$$a(k) = A(k) - K^r A(k-1) - (1 - K^r) \cdot \theta \qquad (4)$$

$$x(k+1) = u[y(k+1)] . \qquad (5)$$

The conventional chaotic neuron model, suggested by Nagumo and Sato[4], has two different type of input simultaneously; input from same layer and extraneous input, and also has a refractory term, which is a self-feedback. The refractory term performs effective dynamic characteristics through repeated signal controlling as one of three terms, which affect output of the chaotic neuron. The neuron model is shown in Fig. 1.
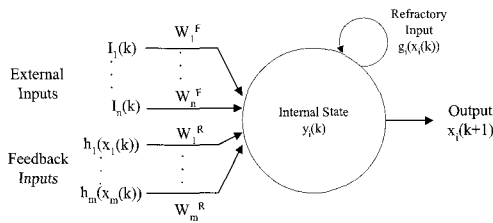

Fig. 1 Chaotic neuron unit

Generally, the dynamics of the ith chaotic neuron in networks at discrete time $k+1$ is

$$x_i(k+1) = f_N[\sum_{j=1}^{n} w_{ij}^F \sum_{r=0}^{k} K_s^r I_j(k-r)$$

$$+ \sum_{j=1}^{m} w_{ij}^R \sum_{r=0}^{k} K_m^r h_j(x_i(k-r))$$

$$- \alpha \cdot \sum_{r=0}^{k} K_r^r g_i(x_i(k-r)) - \theta_i] \qquad (6),$$

where $f_N(\cdot)$ is a sigmoid function, w_{ij}^F and w_{ij}^R are coupling coefficients(weights) from the j th external neuron and the j th feedback neuron to the i th neuron, respectively. $I_j(k-r)$ is the strength of the j th externally applied input at time $k-r$, $h_j(x_i(k-r))$ is a transfer function of the axon connected on the j th chaotic neuron, and $g_i(x_i(k-r))$ is a refractory function of the i th chaotic neuron at time $k-r$, usually an identity function. The n and m are the numbers of external and feedback inputs applied to the chaotic neuron. The decay parameters, K_s^r, K_m^r, and K_r^r are the damping factors of the external, feedback, and refractoriness, respectively. In this paper, we assumed the same values of decay parameters, K. The θ_i is the threshold of the i th chaotic neuron.

Aihara deals with the i th chaotic neuron equation in a reduced form by dividing feedback, external, and refractory terms.[5] Each term is expressed as

$$\xi_i(k+1) = K \cdot \xi_i(k) + \sum_{j=1}^{n} w_{ij}^F I_j(k) \qquad (7)$$

$$\eta_i(k+1) = K \cdot \eta_i(k) + \sum_{j=1}^{m} w_{ij}^R h_j(f(y_i(k))) \qquad (8)$$

$$\zeta_i(k+1) = K \cdot \zeta_i(k) - \alpha g_i(f_i(y_i(k))) - \theta_i(1-K) \qquad (9)$$

If the inertial state of a chaotic neuron at time $k+1$ is expressed as follows,

$$y_i(k+1) = \xi_i(k+1) + \eta_i(k+1) + \zeta_i(k+1) \qquad (10).$$

the equation (10) can be written as follows;

$$y_i(k+1) = K \cdot (\xi_i(k) + \eta_i(k) + \zeta_i(k))$$

$$+ \sum_{j=1}^{n} w_{ij}^F I_j(k) + \sum_{j=1}^{m} w_{ij}^R h_j(f_N(y_i(k))) \qquad (11).$$

$$- \alpha g_i(f_N(y_i(k))) - \theta_i(1-K)$$

Since $y_i(t)$ is defined as $y_i(t) = \xi_i(t) + \eta_i(t) + \zeta_i(t)$, the equation (11) can be expressed as

$$y_i(k+1) = K \cdot y_i(k) + \sum_{j=1}^{n} w_{ij}^F I_j(k) + \sum_{j=1}^{m} w_{ij}^R h_j(f_N(y_i(k)))$$

$$- \alpha g_i(f_N(y_i(k))) - \theta_i(1-K) \qquad (12).$$

In order to apply the continuous Hopfield neural network structure to the recurrent inputs, Aihara et al. define the symmetric structure of recurrent weights as $w_{ij}^R = w_{ji}^R$, $w_{ii}^R = 0$. This neural network uses two kinds of

learning rules in same network. Since the structure decreases the efficiency of learning and the dynamic characteristics of network, this model is not appropriate for modeling dynamic systems.

2.2 Chaotic Dynamic Neuron Model

The chaotic neuron model still has complicate dynamics for applying on neural networks. More simplification is needed in this model for reducing the computation time. This paper presents a chaotic dynamic neuron unit with same chaotic characteristics. Since the $\alpha g_i(f_N(y_i(k)))$ term in equation (12) is overlapped in the case of $i = j$ in the term of $\sum_{j=1}^{m} w_{ij}^R h_j(f_N(y_i(k)))$, the $\alpha g_i(f_N(y_i(k)))$ is abbreviated in this modified model. For more simplification, the threshold term, $\theta_i(1-K)$, is defined zero, and the nonlinear function, $h_j(\cdot)$, is defined 1. Fig. 2 shows modified chaotic neuron unit.
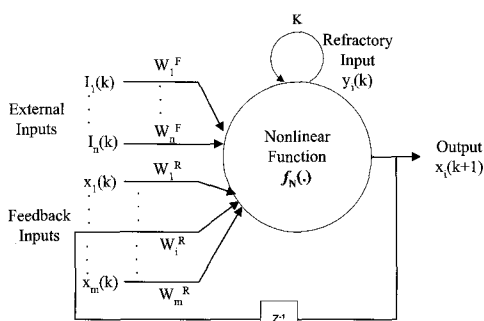

Fig. 2 Chaotic dynamic neuron unit

Then the reduced form of eq. (6) is

$$y_i(k+1) = K \cdot y_i(k) + \sum_{j=1}^{n} w_{ij}^F \cdot I_j(k) + \sum_{j=1}^{m} w_{ij}^R \cdot x_i(k) \quad (13)$$

$$x_i(k+1) = f_N[y_i(k+1)] \quad (14)$$

$$f_N[y_i(k+1)] = \frac{1}{1 + e^{-y_i(k+1)/\varepsilon}} \quad (15)$$

where ε is slope of sigmoid function.

To increase dynamic characteristics, the nonsymmetric weights are applied in recurrent inputs such as $w_{ij}^R \neq w_{ji}^R$, $w_{ii}^R \neq 0$. Each neuron sums three inputs; the refractoriness, $K \cdot y_i(k)$, the activation, $\sum_{j=1}^{n} w_{ij}^F \cdot I_j(k)$, and the recurrent input, $\sum_{j=1}^{m} w_{ij}^R \cdot x_i(k)$. The summation passes through the nonlinear sigmoid function. This model is similar to the transiently chaotic neural network(TCNN) except for the recurrent input term[6], and has similar characteristics.

2.3 Analyzing the Chaotic Characteristics of Chaotic Neuron

The single neuron model in fig. 3 is traditional chaotic neuron model suggested by Nagumo-sato and Aihara.

The simplified single chaotic neuron model is described as eq 16 and 17.

$$y(k+1) = K \cdot y(k) - x(k) + A(k) \quad (16)$$

$$x(k+1) = f_N[y_i(k+1)] = \frac{1}{1 + e^{-y_i(k+1)/\varepsilon}} \quad (17)$$

where $y(k+1)$ is internal state of chaotic neuron at discrete time $k+1$, $x(k+1)$ is output of chaotic neuron at discrete time $k+1$, K is refractory parameter, $A(k)$ is activation of chaotic neuron at discrete time k, and ε is slope of sigmoid function. As shown in eq. 16, the internal state of chaotic neuron is determined by summarizing three inputs, refractory, recurrent and activational inputs.
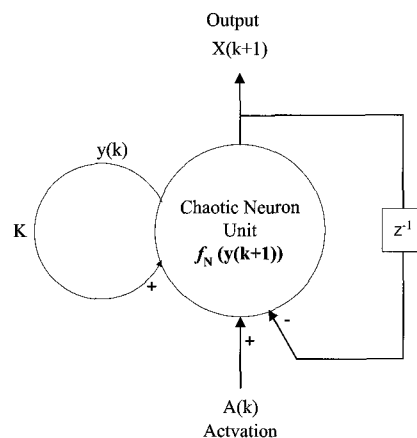

Fig. 3 Single chaotic neuron unit

The output of neuron is determined by a nonlinear sigmoid function in eq. 17. The output of neuron has different output characteristics by changing activation values as shown fig. 4. The fig. 4 is poincare map for identifying chaotic characteristics.

The slope is fixed with $\varepsilon = 0.06$, the refractory value is also fixed with $K = 0.7$, and the activation value is varied with three cases, $A = 0.2$, $A = 0.5$, $A = 0.7$, and $A = 0.9$. In cases on $A = 0.2$, and $A = 0.7$, the graphs show chaotic outputs. In $A = 0.5$, and $A = 0.9$, the outputs have fixed point. Figure 4. b shows a limit cycle, and figure 4.d shows a typical orbit spirals into the fixed point as $k \to \infty$. According to activation values, the output of chaotic neuron shows chaotic characteristics or fixed points.

(a) A=0.2   (b) A=0.5

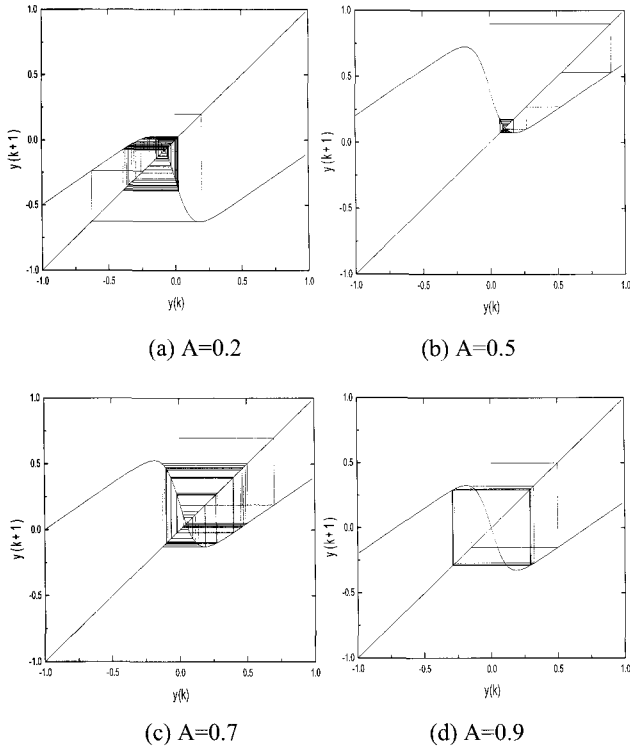(c) A=0.7   (d) A=0.9

Fig. 4 Output characteristics of CNN by activation (In case refractory K=0.7)
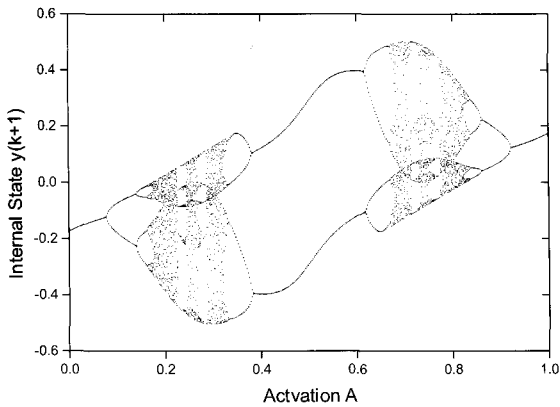

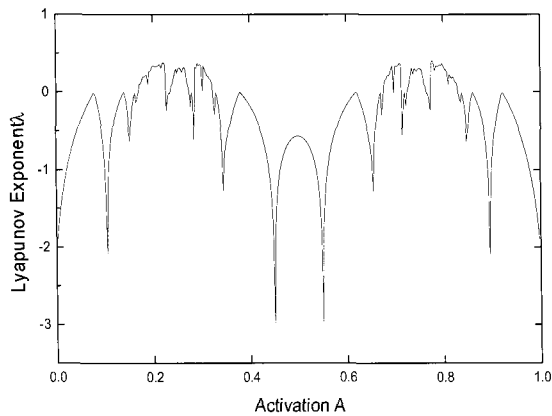Fig. 5 Bifurcation diagram of logistic map for CNN

Fig. 6 Lyapunov exponent diagram λ

Figure 5 shows bifurcation diagram for logistic map based on modified chaotic neurons. In this case, the slope is fixed with $\varepsilon = 0.06$, the refractory value is also fixed with $K = 0.7$, and the activation value is varied from 0 to 1. The output is split as period doubling bifurcation, and the map becomes chaotic intermittently.

Even though logistic map in figure 5 exhibits aperiodic orbits for certain activation values, Lyapunov exponent is more computationally useful for verifying chaotic characteristics. A positive Lyapunov exponent is a signature of chaos. Figure 6 shows Lyapunov exponent diagram. The formula is defined as,

$$\lambda = \lim_{n \to \infty} \left\{ \frac{1}{n} \sum_{i=0}^{n-1} \ln \left| f'(x_i) \right| \right\} \tag{18}$$

where λ is Lyapunov exponent, x_i is state of orbit in i th iterations, and $f(\cdot)$ is a smooth function. For limited p cycle, Lyapunov exponent can be defined as

$$\lambda \approx \frac{1}{p} \sum_{i=0}^{p-1} \ln \left| f'(x_i) \right| \tag{19}$$

In this MCNN application, Lyapunov exponent could be defined as

$$\lambda \approx \frac{1}{p} \sum_{i=0}^{p-1} \ln \left| K - \frac{x_p^2 \cdot e^{y_p / \varepsilon}}{\varepsilon} \right| \tag{20}$$

Figure 7 is three dimensional diagram of Lyapunov exponent for variable slope and activation. The results show that chaotic output is produced intermittently depending on the slope and activation values.
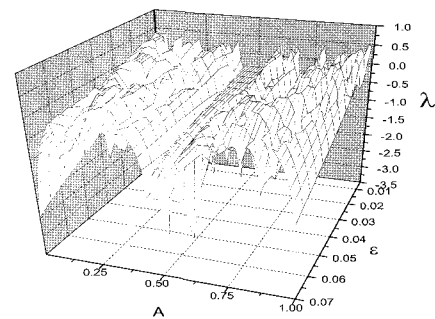

Fig. 7 Lyapunov exponents with variable slope ε

3. Convergence and Stability of Chaotic Dynamic Neural Networks

3.1 Chaotic Dynamic Neural Networks

A chaotic dynamic neural networks in figure 8 has three kind of inputs, weighted recurrent input $w^R \cdot f_N(y(k))$, sum of

weighted inputs $\sum w^F \cdot I(k)$, and refractory input $K \cdot y(k)$. The $\sum w^F \cdot I(k)$ is the weighted sum of all outside inputs from neurons including internal layer and between layers. The $f_N(\cdot)$ is a sigmoid function, $f_N[y(k+1)] = \dfrac{1}{1 + e^{-y(k+1)/\varepsilon}}$, and K, w^F and w^R are refractory variable, weights for inputs, and weight for recurrent input respectively. $I(k)$, $y(k)$, and $x(k)$ are the external inputs, the internal state, and the output of neuron respectively at time k.
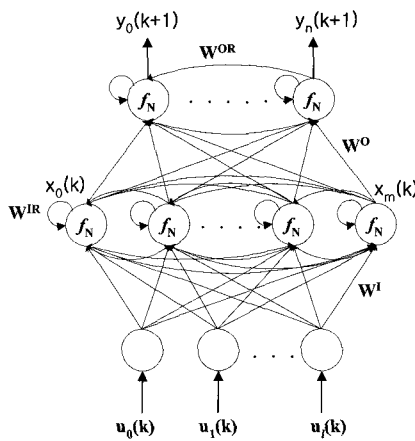

Fig. 8 Chaotic dynamic neural networks

The output is determined by a sigmoid function of weighted sum. The mathematical forms can be expressed as equation (21) and (22).

$$y(k+1) = K \cdot y(k) + \sum w^F \cdot I(k) + w^R \cdot f_N(y(k)) \quad (21)$$

$$x(k+1) = f_N[y(k+1)] \quad (22),$$

3.2 Convergence and Stable condition

Almost all neural network application such as pattern recognition, identification or prediction, and control are optimization problem. To solve optimization problems which need to almost all neural network application, the neurons in neural networks should have one stable fixed point. Chaotic characteristics or limit cycle in the output of neuron should be disappeared in permanent for solving optimization problems. So, the chaotic characteristics of chaotic neuron should be transient. Any arbitrary nonlinear continuous function(including chaotic function) with one stable fixed point shows these two convergence conditions.

1) If a fixed point $y^*(k)$ satisfies $y^*(k+1) = f(y^*(k))$ for any arbitrary nonlinear continuous function $f(\cdot)$, the $y^*(k)$ is an stable equilibrium point; for if $y(k) = y^*(k)$ then $y(k+1) = f(y(k+1))$: $f(y^*(k)) = y^*(k)$: hence the orbit remains at $y^*(k)$ for all further iterations.

2) If the derivative of any one dimensional system $y(k+1) = f(y(k))$ has a negative value at the fixed point $y^*(k)$, the fixed point is stable.

From (21) and condition 1, the internal state, $y(k)$, at fixed point can be represented as

$$y^*(k) = K \cdot y^*(k) + \sum w^F \cdot I(k) + w^R \cdot f_N(y^*(k)) \quad (23).$$

Equation 23 can be represent as

$$\frac{(1-K)}{w^R} \cdot y^*(k) - \frac{\sum w^F \cdot I(k)}{w^R} = f_N(y^*(k)) \quad (24).$$

Let $a = \dfrac{(1-K)}{w^R}$ and $b = \dfrac{\sum w^F \cdot I(k)}{w^R}$, then

$$a \cdot y^*(k) - b = f_N(y^*(k)) \quad (25).$$

Let $F(y^*(k)) = a \cdot y^*(k) - b$ and $G(y^*(k)) = f_N(y^*(k))$, then (25) can represents as

$$F(y^*(k)) = G(y^*(k)) \quad (26).$$

Theorem 1: Let w^R be the recurrent weight for the internal state of CDNN $y(k)$ and S_{f*} be defined as

$$S_{f*} = \max \left| \frac{\partial f_N(y^*(k))}{\partial y^*(k)} \right| \text{ where } f_N(y(k)) \text{ is sigmoid function.}$$

The convergence is guaranteed to only one fixed point $y^*(k)$ if w^R has the conditions as

$$-(1-K) \cdot S_{f*} < w^R \leq \frac{1-K}{S_{f*}} \quad (27).$$

Proof: If $a \geq 0$, the always stable condition of fixed point with only one fixed point can be defined using condition 2, as follows;

$$\max \left| \frac{\partial G(y^*(k))}{\partial y^*(k)} \right| \leq \frac{\partial F(y^*(k))}{\partial y^*(k)} \leq \infty \quad (28).$$

Equation (28) can be rewritten as

$$S_{f*} \leq a \leq \infty \quad (29).$$

From (29), the stability condition related to w^R can be made as

$$0 \leq w^R \leq \frac{1-K}{S_{f*}} \quad (30)$$

If $a \leq 0$, the always stable condition of fixed point with only one fixed point can be defined using condition 2, as follows;

$$-\infty \leq a < -\frac{1}{S_{f*}} \qquad (31)$$

From (31) , the stability condition related to w^R can be made as

$$-(1-K)\cdot S_{f*} < w^R \leq 0 \qquad (32)$$

From (30) and (32), the stability condition in (27) is made as follows.

$$-(1-K)\cdot S_{f*} < w^R \leq \frac{1-K}{S_{f*}} .$$

3.3 Non-convergence condition

In case $0 < a \leq S_{f*}$, the fixed points can be classified as three type: a)two fixed points case, b) three fixed points case, and c) one fixed point case.

1) Two fixed point case.

Two fixed points case can be happened when the slopes of two functions, $F(y^*(k))$, $G(y^*(k))$, have same value.

$$\frac{\partial G(y^*(k))}{\partial y^*(k)} \equiv \frac{\partial F(y^*(k))}{\partial y^*(k)} \qquad (33)$$

(23) is represents as

$$a = x^*(k)\cdot(1-x^*(k))/S_{f*} \qquad (34)$$

where $x^*(k)$ is the output of chaotic dynamic neuron at fixed point, and s is slope of sigmoid function.

From (34) $x^*(k)$ is defined as

$$x^*(k) = \frac{1 \pm \sqrt{1-4as_{f*}}}{2a} \qquad (35)$$

$$y^*(k) = -s_{f*}\cdot \ln\left[\frac{2a}{1\pm\sqrt{1-4as_{f*}}} - 1\right] \qquad (36)$$

$$b = a\cdot y^*(k) - x^*(k)$$

$$= -as_{f*}\ln\left[\frac{2a}{1\pm\sqrt{1-4as_{f*}}} - 1\right] - \frac{1\pm\sqrt{1-4as_{f*}}}{2a} \qquad (37)$$

The fixed point at (35) is half-stable, and the other fixed point is stable but it will locate in near 0 or 1.

2) Three fixed points case

$$-as_{f*}\ln\left[\frac{2a}{1-\sqrt{1-4as_{f*}}} - 1\right] - \frac{1-\sqrt{1-4as_{f*}}}{2a} < b$$

$$< -as_{f*}\ln\left[\frac{2a}{1+\sqrt{1-4as_{f*}}} - 1\right] - \frac{1+\sqrt{1-4as_{f*}}}{2a} \qquad (38)$$

Two fixed points near 0 and 1 are stable, and one fixed point between 0 and 1 is unstable.

3) One fixed point case

$$b > -as_{f*}\ln\left[\frac{2a}{1+\sqrt{1-4as_{f*}}} - 1\right] - \frac{1+\sqrt{1-4as_{f*}}}{2a} \qquad (39)$$

This case has one stable fixed point but located in near 1

$$b < -as_{f*}\ln\left[\frac{2a}{1-\sqrt{1-4as_{f*}}} - 1\right] - \frac{1-\sqrt{1-4as_{f*}}}{2a} \qquad (40)$$

This case has one stable fixed point but located in near 0

In case $-\frac{1}{S_{f*}} < a \leq 0$, the fixed points has one but unstable by condition 2.

From, the w^R can be written as

$$-\infty \leq w^R < -(1-K)\cdot S_{f*} \qquad (41)$$

3.4 Chaos and limit cycle condition

If $a = -\frac{1}{S_{f*}}$, the slopes of $F(y^*(k)) = a\cdot y^*(k) - b$ and $G(y^*(k)) = f_N(y^*(k))$ are perpendicular. The output of chaotic dynamic neuron shows stable limit cycle.

If $a \approx -\frac{1}{S_{f*}}$, the slopes of $F(y^*(k)) = a\cdot y^*(k) - b$ and $G(y^*(k)) = f_N(y^*(k))$ are almost perpendicular. The output of chaotic dynamic neuron shows chaos. The result of chaotic response shows in fig 4. a and c.

4. Conclusion

In this paper, we present the stability and convergence condition of the chaotic neural networks. Even though the chaotic response may be helpful to overcome local minimum problem, the initial chaos should be disappeared in time such as the transiently chaotic neuron. Since the chaotic neural network

has the highly dynamic characteristics and the high slope of sigmoid function, the network shows fast learning and also stability problem. To overcome this problem, the convergence and stability condition are presented in this paper.

References

[1] Sakada, C. A, Freeman, W. J, "How brain make chaos in order to make sense of the world", *Behavioral and Brain Science*, 10, 161-195, 1987.

[2] E. Caianiello, "Outline of a Theory of thought Process and Thinking Machines," *J. Theor. Biol.* 2, pp. 204-235, 1961.

[3] Caianiello, E. R., DeLuca, A., "Decision equation for binary system: Application to neuronal behavior", *Kybernetik*, 3, 33-40, 1966.

[4] J. Nagumo and S. Sato, "On a Response Characteristics of a Mathematical Neuron Model," *Kybernetik.* 10, pp. 155-164, 1972.

[5] K. Aihara, T. Takabe and M. Toyoda, "Chaotic Neural Networks," *Phys. Lett.* A144, pp. 333-340, 1990.

[6] Chen, L., Aihara, K., "Chaotic simulated annealing by a neural network model with transient chaos", *Neural Networks*, vol.8, no.6, 915-930, 1995.

[7] Ishi, S., Fukumizu, K., Watanabe, S., "A network of chaotic elements for information processing", *Neural Networks*, vol. 9, no. 1, 25-40, 1996.

[8] L. Chen and K. Aihara, "Global Search Ability of Chaotic Neural Networks", *IEEE Trans. on Circuit and System I*, Vol. 46, No. 8, pp. 974-993, 1999.

[9] S.H.Kim and W.W.Park, "An Adaptive Neuro-Controller with Modified Chaotic Neural Networks", *ICNN01*, pp. 509-514, 2001.

Sang Hee Kim

received his M.S degrees in Electrical Engineering from Yonsei University in 1985, and Ph.D degree from Texas A&M University 1992. He is professor at the school of electronic engineering at Kumoh National Institute of Technology from 1993. He performed research at M.I.T. and at Boston University as as research fellow and visiting professor in 2005-2006. His research interests are intelligent control, neural networks, nonlinear dynamics and chaos theory and Digital television.

Phone : 054) 478-7430
Fax : 054) 478-7449
E-mail : shkim@kumoh.ac.kr

Hua O. Wang

received his Ph.D. at University of Maryland. He is Associate Professor of Aerospace and Mechanical Engineering Department, College of Engineering, Intelligent Mechatronics Laboratory. He is an expert in nonlinear control and in particular in the control of systems having complex nonlinear dynamics. He is also affiliated with Boston University's Center for Information and Systems Engineering (CISE). His research interests are Control of nonlinear phenomena and bifurcations.

Phone : (617)353-8860
Fax : (617) 353-5866
Email : wangh@bu.edu