Systems and Computers in Japan, Vol. 32, No. 6, 2001
Translated from Denshi Joho Tsushin Gakkai Ronbunshi, Vol. J82-D-II, No. 9, September 1999, pp. 1520–1527

Effects of Neuronal Dynamics on Memory Storing in Stimulus–Response Scheme Model

Naofumi Katada and Haruhiko Nishimura

Studies of Information Science, Hyogo University of Education, Hyogo, 673-1494 Japan

SUMMARY

Previously, the authors introduced a stimulus–response scheme that supports plastic variation of synapse weights in chaotic neural networks, and analyzed how memory formation evolved under external stimulation. In this paper, the same analysis is applied to stochastic neural networks with different dynamics, and the obtained results are compared to those for chaotic networks to examine the influence exerted by dynamic parameters on the memorization process. In experiments, chaotic networks proved to be more efficient than stochastic networks. This is indicative of functionality of neural chaos in stimulus–response scheme that is close to natural learning, while further experiments are required to confirm the generality of this conclusion. © 2001 Scripta Technica, Syst Comp Jpn, 32(6): 29–35, 2001

Key words: Chaotic; stochastic; stimulus–response; neural network; learning; plasticity.

1. Introduction

In actual cerebral nervous systems, chaos is observed at several hierarchical levels, from a single neuron through active states of the entire brain, being an array of neural networks [1, 2]. Most studies of memory models based on dynamic properties of such neural chaos [3–8] have so far dealt with memory retrieval processes. Considering that memory functions are based on memorization (learning) and retrieval, there is a need for studies on functionality of neural chaos in memorization (learning) processes.

Conventional learning techniques in neural networks suggest memorization (building of weight structure) by batch processing of pattern data. On the other hand, when considering the learning process in more realistic terms (as with living organisms receiving pattern data as a sequence of external stimuli), research must be focused on dynamic properties of the neural network itself as information receiver. From such a viewpoint, the authors introduced a stimulus–response scheme with variable synapse weights into chaotic neural networks, and thoroughly analyzed how memory formation advances as a network receives external stimulating information, and how memorization is accomplished [9].

As a result, the following was confirmed for memorization in chaotic neural networks featuring hysteresis and refractoriness. First, memorizing capability grows globally with stronger stimuli, while local chaos-specific sensitivity to stimulation strength is observed as well. Second, high memorizing performance is available across a wide range of stimulation strength, as distinct from nonchaotic systems (Hopfield network point). In addition, a network's chaos characteristics observed during stimulus–response process were estimated quantitatively using (maximum) Lyapunov exponent; in so doing, fundamental properties of memorizing performance in chaotic networks were shown to be related to refractoriness and hysteresis [9].

The present study is based on the aforementioned results; chaotic neural networks are compared to stochastic networks with completely different dynamics in order to get a better understanding of the influence exerted by the network's dynamic properties on memorization. Particularly, stochastic neural networks are examined in the same experiments that were previously carried out on chaotic networks [9], and the results are used to analyze the differences between the two types of networks.

2. Stimulus-Response Scheme Model

2.1. Stimulus–response scheme

Contribution of external stimuli (denoted by S_e) is expressed through the following evolution equation for N neurons ($-1 \le X_i \le 1$, $i = 1, \ldots N$) connected by synapse weights w_{ij} (here $w_{ii} = 0$):

$$X_i(t+1) = f\big(h_i(t+1)\big) \tag{1}$$

Here

$$f(y) = tanh\left(\frac{y}{2\varepsilon}\right) \tag{2}$$

Application of a stimulus is reflected through temporary addition of external stimulus term σ_i ($i = 1, \ldots, N$) to internal state ($h_i \to h_i + \sigma_i$):

$$X_i(t+1) = f\big(h_i(t+1) + \sigma_i\big) \tag{3}$$

In other words, external information affects the neuron's firing (excitation/inhibition) through temporary variation of the threshold value (action potential).

As an external stimulus is applied, weights are not fixed but rather varied plastically:

$$w_{ij}(t+1) = w_{ij}(t) + \beta X_i(t)X_j(t) \tag{4}$$

Here β is a measure of the weight's plasticity, below referred to as plasticity parameter. To prevent infinite growth of synapse intensity, upper limit K_i is set for weight norm

$$\|w_i(t+1)\| \equiv \left(\sum_{j=1}^{N} w_{ij}^2(t+1)\right)^{1/2} \tag{5}$$

and weights are normalized as follows:

$$\tilde{w}_{ij}(t+1) = w_{ij}(t) + \beta X_i(t)X_j(t) \tag{6}$$

$$w_{ij}(t+1) = \tilde{w}_{ij}(t+1) \times \frac{K_i}{\|\tilde{w}_i(t+1)\|} \tag{7}$$

Equation (4) is a generalization of Hebb's learning rule [12] saying that when a neuron gets excited, the weight of the corresponding synapse grows simultaneously. That is, connection between input neuron i and output neuron j gets stronger through AND operation.

Consider external stimulus S_e with duration T_s applied repetitively at an interval of T_I (see Fig. 1). In so doing, only one type of stimulus can be applied during period T_s. During the nonstimulation period ($T_I - T_s$), weights are fixed ($\beta = 0$). These assumptions are used to make possible quantitative analysis of network behavior variation after each period T_s.

2.2. Neuronal dynamics

Two neuron models are considered in this study, namely, chaotic neurons and stochastic neurons. Both networks have irregular dynamical characteristics but there is a large difference in system dynamics (deterministic versus indeterministic). Both types of neurons are defined below.

2.2.1. Chaotic neuron

The chaotic neuron model [3] imitates refractoriness of actual neurons through hysteresis effect (exponential attenuation with time). The model is called chaotic because chaos is easily observed in its response characteristic. Chaotic neural network (CNN) [3, 7] is a system of interconnected chaotic neurons. In this case, the following relations correspond to Eq. (1):

$$h_i(t+1) = \eta_i(t+1) + \zeta_i(t+1) \tag{8}$$

$$\eta_i(t+1) = k_f \eta_i(t) + \sum_{j=1}^{N} w_{ij} X_j(t) \tag{9}$$

$$\zeta_i(t+1) = k_r \zeta_i(t) - \alpha X_i(t) + a \tag{10}$$

Here $a \equiv -\theta_i(1 - k_r)$, k_f (k_r) is the attenuation constant of feedback input (refractoriness), α is the refractoriness parameter, and θ_i is the threshold.

If $k_f = k_r = \alpha = 0$ in Eqs. (9) and (10), then

$$X_i(t+1) = f\left(\sum_{j=1}^{N} w_{ij} X_j(t) - \theta_i\right) \tag{11}$$

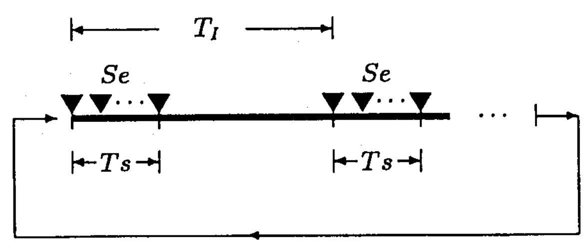


Fig. 1. Flow diagram of stimulus–response scheme.

This coincides with the neuron's evolution equation in the Hopfield neural network [10]. That is, CNN is identical to the conventional Hopfield model of static associative memory if $k_f = k_r = \alpha = 0$ [below referred to as the Hopfield network point (HNP)].

2.2.2. Stochastic neuron

Neuron states in Hopfield neural network are updated deterministically according to Eq. (11) but stochastic update is possible as follows:

$$\begin{cases} Prob\{X_i = f(h_i)\} = g_H(|h_i|) \\ Prob\{X_i = -f(h_i)\} = 1 - g_H(|h_i|) \end{cases} \tag{12}$$

In the above equation

$$g_H(h) = \frac{1}{1 + e^{-h/H}} \tag{13}$$

Here $Prob\{\}$ is the probability of the relation in brackets. The probability of neuron state inversion ($+ \Leftrightarrow -$) is determined by parameter H called temperature. Stochastic influence of this temperature is called thermal noise [11]. At the extreme of $H \to 0$, $g_H \to 1$ in Eq. (13), therefore Eq. (12) coincides with deterministic evolution equation (11) for Hopfield network.

A network composed of such stochastic neurons is called stochastic neural network (SNN).

2.3. Evaluation and measure of memorization

Denoting external stimulating data (p data points) as $\{\xi_i^\mu\} \equiv (\xi_1^\mu, \ldots, \xi_N^\mu)$, $\mu = 1, \ldots, p$; $\xi_i^\mu = 1$ or -1, the simplest relation with the external stimulus term σ_i is assumed:

$$\{\sigma_i\} = s\{\xi_i^\mu\} \tag{14}$$

Here s is an intensity parameter that determines stimulation strength.

When a network is exposed to external stimuli, weights vary plastically as in Eq. (4). Here the following measure [11] is used to evaluate the degree of memorization related to external stimuli (formation of gravisphere):

$$\Gamma_i^\mu = \xi_i^\mu \sum_{j=1}^{N} w_{ij} \xi_j^\mu \tag{15}$$

$$\gamma_i^\mu = \Gamma_i^\mu / \|w_i\| \tag{16}$$

This measure is called stability; when all N values of $\gamma_i^\mu (i = 1, \ldots, N)$ are positive for the entire network, in the case of Hopfield network [Eq. (11)], pattern $\{\xi^\mu\}$ is stable ($\{X_i(t+1)\} = \{X_i(t)\} = \{\xi_i^\mu\}$). In other words, the more neurons offer $\{\Gamma_i^\mu\}(\{\gamma_i^\mu\})$ distributed in the positive region, the more stable is memorization of the μ-th pattern.

The energy map of a network at the moment t is expressed as follows:

$$E(t) = -\frac{1}{2} \sum_{i,j}^{N} w_{ij} X_i(t) X_j(t) \tag{17}$$

This allows one to estimate depth of permeation of stimulus pattern $\{\xi_i^\mu\}$ into the network. In the general case, the existence of a function offering energy properties of a network (Lyapunov's function) is not guaranteed. Therefore, in this study, the scalar value E is referred to as the energy map.

3. Simulation

3.1. Pattern data and parameter settings

In computer-aided experiments, a network of 156 neurons was used ($N = 156$), and alphabet characters 12×13 ($= 156$) were employed as stimulus patterns $\{\xi_i^\mu\} (i = 1, \ldots, N)$. Some examples are shown in Fig. 2 (black and white dots correspond to $\xi_i^\mu = 1$ and $\xi_i^\mu = -1$, respectively).

In this study, experimental results for memorization of eight patterns (R, Z, Q, Y, X, A, T, H) are considered; overlap between these eight patterns

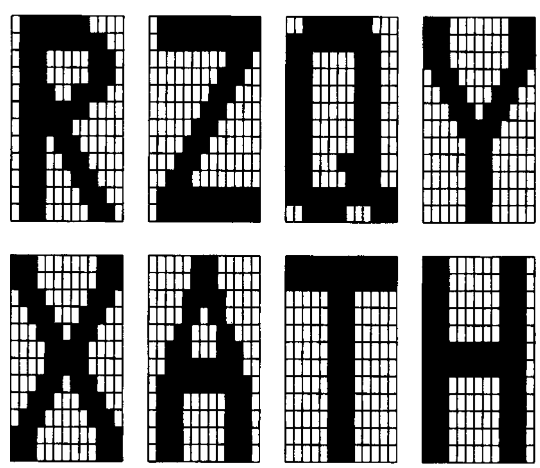


Fig. 2. Examples of stimulus patterns.

$$q_{\mu\nu} \equiv \frac{1}{N} \sum_{i=1}^{N} \xi_i^{\mu} \xi_i^{\nu} \tag{17a}$$

lies within the range of

$$-0.333(H \ vs \ T) \leqq q_{\mu\nu} \leqq 0.385(T \ vs \ Y) \tag{17b}$$

Thus, it is a correlated nonorthogonal pattern group [in the case of $q_{\mu\nu} = 0$ ($u \neq v$), the μ-th and v-th patterns are mutually orthogonal].

Now consider parameter settings. As to time setting, pattern application starts at the moment $t = t_0$ (= 100), with duration of every pattern $T_s = 10$ at the interval of $T_I = 100$. This procedure is repeated 10 times, so that the whole time is $T = t_0 + 10pT_I$, with p being the number of patterns. The initial state of the network is assumed disconnected, that is, all weights w_{ij} $(t = 0) = 0$, while the plasticity parameter in Eqs. (4) and (6) is assumed $\beta = 0.1/N$. Upper limit K_i for weight norms is obtained from Eq. (5), using w_{ij} $(t = t_0 + 10pT_I)$ after each cycle.

As regards neuron thresholds and input–output function gradient, the values $a = 0$ $(\theta_i = 0)$, $\varepsilon = 0.015$ for all neurons were determined experimentally.

3.2. Analysis of memory formation in SNN

Analysis of neuronal dynamics in chaotic networks as described in Section 2.2.1 has been performed in Ref. 9 under the same conditions as given in Section 3.1. Therefore, considered below is analysis in the case of experiments with the eight patterns fed to stochastic neural network (SNN).

Shown in Fig. 3 is the time variation of stimulus–response process for stochastic neural network with the temperature parameter set at $H = 2.0$, the overlap between network states $\{X_i\}$ and pattern I being

$$m^I(t) = \frac{1}{N} \sum_{i=1}^{N} X_i(t)\xi_i^I \tag{18}$$

Stimulation strength is set at $s = 2.20$. Figure 3(a) reflects the network's behavior under pattern application in the order of [R → Z → Q → Y → X → A → T → H], at $T_s = 10$ and $T_I = 100$, with 10 cycles of 800 each. As weights evolve due to stimulation, the network state changes dynamically from the disconnected state of $m_I = 0$, so that response develops gradually. In so doing, time settings and stimulation strength settings are the same as in similar experiments with chaotic networks [9].

In addition, time variation of the average of $1 - g_H(H_i)$ for all neurons defined as follows using $g_H(h)$ in Eq. (13)

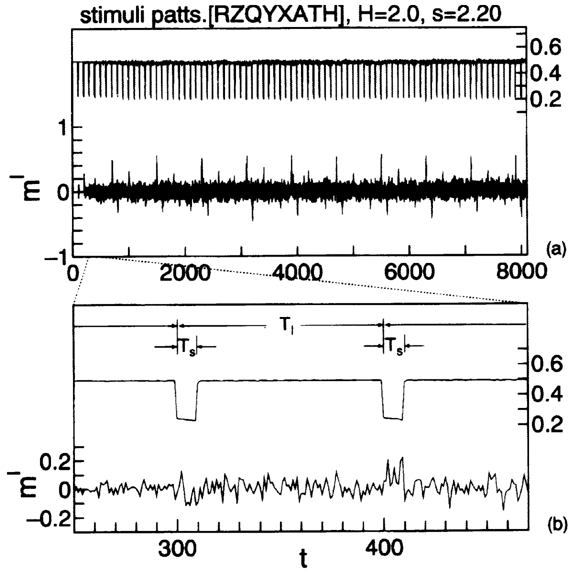


Fig. 3. Time series evolution of stochastic network with eight stimulus patterns. (a) For 10 stimulus–response cycles; (b) closeup for $t = 250 \sim 470$.

$$\rho_H(t) = \frac{1}{N} \sum_{i=1}^{N} (1 - g_H(h_i(t))) \tag{19}$$

is also shown in the diagram to estimate stochastic fluctuation in network state update. Here $h_i(t) = \Sigma_{j=1}^{N} w_{ij} X_j(t)$. This ρ_H will be referred to as stochastic activity in stochastic neural network.

Figure 3(b) gives a closeup of the region $t = 250 \sim 470$. As seen, ρ_H is close to 0.5 when no stimulus is applied while changing to about 0.2 during stimulation (T_s), which corresponds to growth of $|h_i|$ as $h_i \to h_i + \sigma_i$ takes place for all neurons. Since the upper limit K_i of weight norm $\|w_i\|$ is set for all neurons, no major changes appear in ρ_H behavior pattern, while memorization advances as variations between weight components $w_i = (w_{i1}, w_{i2}, \ldots, w_{iN})$. At nonstimulation periods with ρ_H as high as about 0.5, weights $\{w_{ij}\}$ are fixed, hence memorization progress is not hindered.

Figure 4 presents the relation between stochastic activity ρ_H and final stability (γ_i^{μ}) at $t = 8100$ obtained through repetitive experiments with the eight patterns while varying temperature parameter H from 0 through 10 at steps of 1/8. On the abscissa,

$$\langle \rho_H \rangle_s = \frac{1}{10pTs} \sum_{i=1}^{10p} \sum_{j=1}^{Ts} \rho_H(t_0 + (i-1)T_I + j) \tag{20}$$

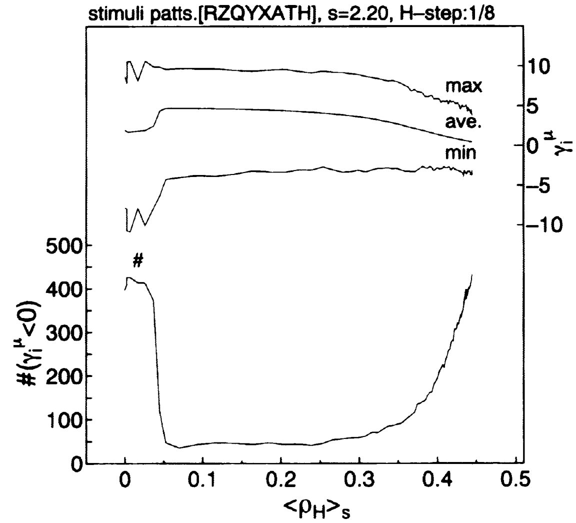


Fig. 4. Dependence of final stability coefficients on time-averaged stochastic activity.

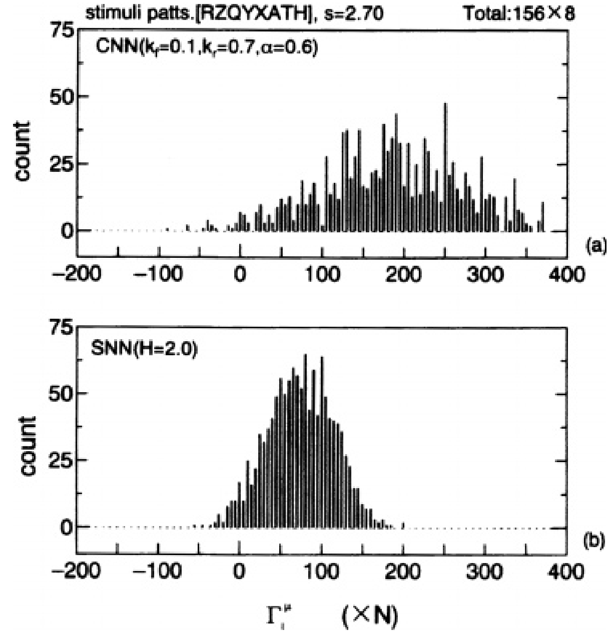


Fig. 5. Distribution of final stability coefficients for eight stimuli in case of (a) CNN and (b) SNN.

is plotted ($<\rho_H>_s = 0$ at $H = 0$) in order to estimate (time-averaged) stochastic activity of the network during stimulation period (T_s) when weights may vary plastically. Here p is the number of stimulus patterns (in our case, $p = 8$). In the area of $<\rho_H>_s = 0.05 \sim 0.25$, high memorization capability may exist (few final stability coefficients remain negative). This area corresponds to $H = 1.0 \sim 2.2$ in terms of temperature, with $H = 2.0$ as in Fig. 3 being a representative example.

4. Different Dynamics and Memorization Performance

Given below is a comparison between the results obtained here for SNN, and previous results for CNN and HNP [9].

4.1. Comparison in final stability of patterns

Stability coefficients Γ_i^μ [Eq. (15)] were calculated for all patterns using final weights w_{ij} ($t = 8100$) in experiments with stimulation strength $s = 2.70$, and the frequency distribution of the coefficients for eight patterns [total of 1248 ($= N \times p = 156 \times 8$)] was obtained as plotted in Fig. 5 (at an interval of 5). Network parameters were set so that to provide high retention (CNN: $k_f = 0.1$, $k_r = 0.7$, $\alpha = 0.6$; SNN: $H = 2.0$). The number of coefficients that remained negative $\#\,(\Gamma_i^\mu < 0)[= \#\,(\gamma_i^\mu < 0)]$ was quite different, namely, 17 for CNN, and 40 for SNN. It is particularly remarkable that the difference was more than two times for both mean value and variance (this difference, however, is not pronounced for γ_i^μ because every neuron is normalized to $\|w_i\|$).

4.2. Dependence on stimulation strength

Dependence of the distribution of final stability coefficients (γ_i^μ) on stimulation strength s in SNN was examined (data were sampled at an interval of 1/128); Fig. 6 presents

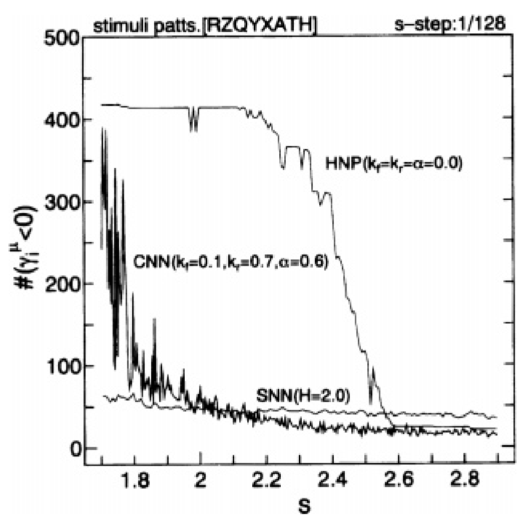


Fig. 6. Dependence of number of negative final stability coefficients on strength of stimulation.

this dependence of # ($\gamma_i^\mu < 0$) on s for SNN as well as for CNN and HNP.

In the case of SNN, # ($\gamma_i^\mu < 0$) varies globally with s, not showing such sensitivity as CNN. In the region of weak stimulation with s below 2.0, SNN outperforms CNN; the opposite is true, however, for strong stimulation. SNN offers little improvement in memorizing performance with stronger stimuli; at s = 2.6 and more, SNN offers more negative coefficients ($\gamma_i^\mu < 0$) than HNP that has no dynamic activity.

4.3. Evaluation of retention and state depth

Different results obtained for final stability distribution may also be interpreted in terms of retention in network [Eq. (11), with $\{\theta_i\} = 0$] with energy map value E for all stimulus patterns, and weights w_{ij} ($t = 8100$). Supposing the network's initial states to be stimulus patterns ($\{X_i(t=0)\} = \{\xi_i^\mu\}$), this retention is evaluated through $q^\mu = (1/N)\Sigma_{i=1}{}^N \xi_i^\mu X_i(t=50)$.

Figure 7 pertains to final weights $\{w_{ij}\}$ in the case of stimulation strength s = 2.70 that proved to offer high memorizing performance in our experiments [number of negative coefficients # ($\gamma_i^\mu < 0$) is 17 for CNN, and 40 for SNN]; neither CNN nor SNN offers complete stability ($q^\mu = 1$) for all eight patterns. However, in the case of CNN, four pattern states are nearly stable while the other pattern states are distributed at $q^\mu \geq 0.7$; on the other hand, with SNN, all pattern states but two are obviously unstable. Two types of networks also differ as to value E of energy map

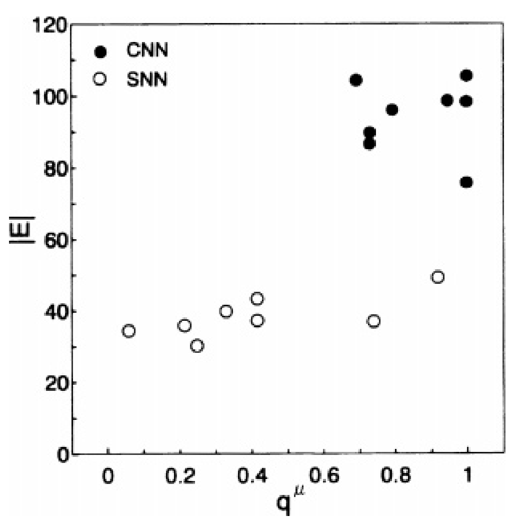


Fig. 7. Comparison of memory formation using retention (q^μ) and state depth ($|E|$).

that shows depth of the pattern's permeation; CNN offers value more than twice as high as SNN.

This difference checks well with the results of final stability analysis described in Section 4.1. The reason is that energy map $E(\mu)$ of the μ-th learning pattern and its stability Γ_i^μ are generally interrelated as follows:

$$\begin{aligned} E^{(\mu)} &= -\frac{1}{2}\sum_{i,j}^{N} w_{ij}\xi_i^\mu \xi_j^\mu \\ &= -\frac{1}{2}\sum_{i=1}^{N}\xi_i^\mu \sum_{j=1}^{N} w_{ij}\xi_j^\mu \\ &= -\frac{1}{2}\sum_{i=1}^{N}\Gamma_i^\mu \end{aligned} \tag{21}$$

In average terms for all patterns

$$(\text{average value of } E) = -\frac{1}{2}\,(\text{average value of } N\Gamma_i^\mu)$$

5. Conclusion

In this study, memory formation and fixation in stimulus–response process were analyzed for neural networks with plastic weights. In so doing, stochastic networks and chaotic networks that have very different dynamics were compared to explain previous experimental results suggesting that chaotic activity remarkably outperforms stochastic activity in stimulus–response memorization. In particular, CNN and SNN show a big difference in final number of negative coefficients # ($\gamma_i^\mu < 0$) and E value for stimulus patterns, which results directly in evaluation of retrieval performance (q^μ). This suggests the feasibility of chaotic functionality, and dynamical learning scheme using such functionality. Though the same results were obtained in experiments with varied number of patterns, further experiments seem necessary to confirm the generality of the above conclusions.

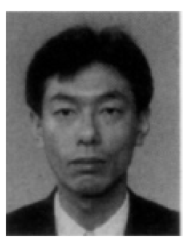
With the proposed stimulus–response scheme model describing evolution from disconnected state to synapse formation, both external stimulation strength and network fluctuations result in different maximum norm of weight $w_i = (w_{i1}, w_{i2}, \ldots, w_{iN})$ for every neuron. Except for very strong external stimulation, in the case of HNP with no source of fluctuation, uniform response for all patterns is not ensured; every neuron's norm is biased according to specific stimulus pattern. On the other hand, with a fluctuation source available, the network can respond without being engaged in the pattern's gravisphere so that specific patterns do not exert material influence on norm formation. Besides, chaotic activity with its directional sensitivity and stochastic activity with its whiteness provide networks with very different dynamics. The authors would like to continue comparative experiments with CNN and SNN to clarify mechanisms behind the differences.

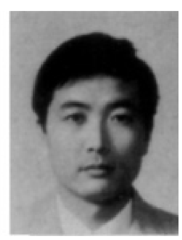
With the human brain, it is highly possible that information processing switches, as necessary, between chaotic activity and stochastic activity. In this context, it seems promising to explore hybrid chaotic–stochastic networks offering combined dynamics, in contrast to previous studies that have dealt separately with the two types of neural networks. In any case, research in new control techniques for autonomous systems, including chaotic control, is necessary to clarify relations between parameter settings and various properties of neural networks in terms of an integrated dynamical mechanism.

REFERENCES

1. Aihara K (editor). Chaos in neural systems. Tokyo Denki University Press; 1993.
2. Arbib MA (editor). The handbook of brain theory and neural networks. MIT Press; 1995.
3. Aihara K, Takabe T, Toyoda M. Chaotic neural networks. Phys Lett 1990;A144:333–340.
4. Inoue M, Nagayoshi A. A chaos neuro-computer. Phys Lett 1991;A158:373–376.
5. Tsuda I. Dynamic link of memory—Chaotic memory map in nonequilibrium neural networks. Neural Networks 1992;5:313–326.
6. Nara S, Davis P, Totsuji H. Memory search using complex dynamics in a recurrent neural network model. Neural Networks 1993;6:963–973.
7. Adachi M, Aihara K. Associative dynamics in chaotic neural network. Neural Networks 1997;10:83–98.
8. Nishimura H, Katada N, Fujita Y. Dynamic learning and retrieving scheme based on chaotic neuron model. In Nakamura R, et al., eds., Complexity and diversity. Spring-Verlag; 1997. p 64–66.
9. Nishimura H, Katada N. Dynamic learning using stimulus–response scheme in chaotic neural networks. Trans ISCIE 1997;10:518–527.
10. Hopfield J.J. Neural networks and physical systems with emergent collective computational abilities, Proc. Natl. Acad. Sci. USA,1982;79:2554–258.
11. Müller B, Reinhardt J. Neural Networks: An Introduction, Springer-Verlag, 1990.
12. Hebb D. Organization of Behaviour, Wiley, New York, 1949.

AUTHORS (from left to right)





Naofumi Katada (member) graduated from Osaka Institute of Technology (civil engineering) in 1985 and completed postgraduate studies at Hyogo University of Education in 1996. He is now at the College of Technology, Hyogo University of Education. His research interests are chaotic neural networks and hybrid systems. He is a member of the Japanese Society for Technological Education.

Haruhiko Nishimura (member) graduated from Shizuoka University (physics) in 1980 and completed his doctorate at Kobe University in 1985. He has been a professor at Hyogo University of Education since 1999. He holds a Ph.D. degree. His research interests are chaotic neural networks, cell automata, and hybrid systems. He is a member of the Japan Neuroscience Society, the Physical Society of Japan, the Information Processing Society of Japan, and the Institute of Systems, Control and Information Engineers.