

**Can stochastic renewal of maps
be a model for cerebral cortex?**

Ichiro Tsuda

Series #224. December 1993

HOKKAIDO UNIVERSITY
PREPRINT SERIES IN MATHEMATICS

- # 194: F. Hiroshima, Scaling limit of a model of quantum electrodynamics, 52 pages. 1993.
- # 195: T. Ozawa, Y. Tsutsumi, Global existence and asymptotic behavior of solutions for the Zakharov equations in three space dimensions, 34 pages. 1993.
- # 196: H. Kubo, Asymptotic behaviors of solutions to semilinear wave equations with initial data of slow decay, 25 pages. 1993.
- # 197: Y. Giga, Motion of a graph by convexified energy, 32 pages. 1993.
- # 198: T. Ozawa, Local decay estimates for Schrödinger operators with long range potentials, 17 pages. 1993.
- # 199: A. Arai, N. Tominaga, Quantization of angle-variables, 31 pages. 1993.
- # 200: S. Izumiya, Y. Kurokawa, Holonomic systems of Clairaut type, 17 pages. 1993.
- # 201: K.-S. Saito, Y. Watatani, Subdiagonal algebras for subfactors, 7 pages. 1993.
- # 202: K. Iwata, On Markov properties of Gaussian generalized random fields, 7 pages. 1993.
- # 203: A. Arai, Characterization of anticommutativity of self-adjoint operators in connection with Clifford algebra and applications, 13 pages. 1993.
- # 204: J. Wierzbicki, An estimation of the depth from an intermediate subfactor, 7 pages. 1993.
- # 205: N. Honda, Vanishing theorem for the tempered distributions, 11 pages. 1993.
- # 206: T. Hibi, Betti number sequences of simplicial complexes, Cohen-Macaulay types and Möbius functions of partially ordered sets, and related topics, 25 pages. 1993.
- # 207: A. Inoue, Regularly varying correlations, 23 pages. 1993.
- # 208: S. Izumiya, B. Li, Overdetermined systems of first order partial differential equations with singular solution, 9 pages. 1993.
- # 209: T. Hibi, Hochster's formula on Betti numbers and Buchsbaum complexes, 7 pages. 1993.
- # 210: T. Hibi, Star-shaped complexes and Ehrhart polynomials, 5 pages. 1993.
- # 211: S. Izumiya, G. T. Kossioris, Geometric singularities for solutions of single conservation laws, 28 pages. 1993.
- # 212: A. Arai, On self-adjointness of Dirac operators in Boson-Fermion Fock spaces, 43 pages. 1993.
- # 213: K. Sugano, Note on non-commutative local field, 3 pages. 1993.
- # 214: A. Hoshiga, Blow-up of the radial solutions to the equations of vibrating membrane, 28 pages. 1993.
- # 215: A. Arai, Scaling limit of anticommuting self-adjoint operators and nonrelativistic limit of Dirac operators, 35 pages. 1993.
- # 216: Y. Giga, N. Mizoguchi, Existence of periodic solutions for equations of evolving curves, 45 pages. 1993.
- # 217: T. Suwa, Indices holomorphic vector fields relative to invariant curves, 10 pages. 1993.
- # 218: S. Izumiya, G. T. Kossioris, Realization theorems of geometric singularities for Hamilton-Jacobi equations, 14 pages. 1993.
- # 219: Y. Giga, K. Yama-uchi, On instability of evolving hypersurfaces, 14 pages. 1993.
- # 220: W. Bruns, T. Hibi, Cohen-Macaulay partially ordered sets with pure resolutions, 11 pages. 1993.
- # 221: S. Jimbo, Y. Morita, Ginzburg Landau equation and stable solutions in a rotational domain, 32 pages. 1993.
- # 222: T. Miyake, Y. Maeda, On a property of Fourier coefficients of cusp forms of half-integral weight, 12 pages. 1993.
- # 223: I. Nakai, Notes on versal deformation of first order PDE and web structure, 34 pages. 1993.

Can stochastic renewal of maps be a model for cerebral cortex?

Ichiro Tsuda*

Department of Mathematics,
Faculty of Science,
Hokkaido University,
Sapporo, 060, Japan

Abstract

We introduce a new type of stochastic dynamics as stochastic renewal of maps, relating to the neurodynamics of cortical memory process. This stochastic dynamics can be reformulated by a skew product transformation of two kinds of variables, one of which describes an underlying dynamical system and the other describes chaotic dynamics, say, Bernoulli shift. The feature of orbits in phase space is investigated in the particular case of neurodynamics model for cortical chaotic memories. A new computational result on the functional role of cortical chaos is obtained. We also present a neurobiological interpretation of psychological perception and memories by means of the notion of chaotic itinerancy.

Keyword: stochastic renewal of dynamics, skew product transformation, cortical chaotic memory, dynamic learning, chaotic itinerancy.

* Leaving Kyushu Institute of Technology(Iizuka).

1. Introduction

Since Freeman's finding of chaos in cortical neuron assemblies[1],[2], the brain research from dynamical system's viewpoints has been highlighted. The present paper focuses on the dynamic features of semi-biological neural network model which was made to investigate the cortical memory dynamics, as well as its relation with synaptic learning.

We deal with a stochastic model in the sense of stochastic renewal of neurodynamics, but this stochasticity can be replaced with chaotic systems by means of skew product transformation. Thus, we reach a stage to analyze the whole network in terms of dynamical systems.

We, furthermore, investigate the change of basin structure in phase space by additional learning. This finding leads us to a neurobiological interpretation of dynamics of perception.

2. Cortical chaotic memory(CCM)

In general, voluntary movements trigger a dynamics of memories, since the motor command must utilize a varying sensory information and, thus, it depends on the past integrated events, *i.e.*, history. A single association of memory with an instant sensory input is not sufficient for voluntary movements. During the movements, visual information changes continually which is again fed back to motor-control systems. Then, the activated memory system cannot be fixed in one definite state, but be driven as changing its states. Some causal rule must be created in this variability of accessible states, according to motor command. Thus, it will not be a random change, and it must not also be fixed to stereotypic motions, for the inputs will be changed unexpectedly. Hence, it will be reasonable that the causal rule itself obeys deterministic chaos, thereby the varieties of highly ordered behaviors are assured. Even if a synaptic learning could take place in short time, this type of memory dynamics would appear as a basis of reorganization of memory space, too.

The notion of input-driven change of memory states has also been highlighted in the problem of motivated recognition of odor, which has been studied most extensively in

Freeman's laboratory [1], [2], [3], [4]. We briefly review the main issues of the olfactory information processing, based on the Freeman's findings. Odor information, which is still vague in the pattern separability at the receptor level, activates neuron assemblies in the olfactory bulb, and its activities are sent to the olfactory cortex and also to other parts of cortex via afferent pathways. A neuron, mainly an inhibitory neuron, in the bulb receives feedback signals from olfactory cortex, whose feedback signals contain the information of motivation, whereby the input-output relation at the assemblies' activity level in the bulb changes. Inputs from receptor cells bring about a spatial coherency of patterns representing the dendritic potential of neuron assemblies in the bulb. However, such a coherent pattern is not a representation of odorant, but rather it is context-dependent in the sense of the dependence on a series of odorant.

Furthermore, Freeman established that the pattern is chaotic, namely odor memories are mediated by weakly chaotic activities of neuron assemblies. He called this activity state a "wing". The following findings are essential. A new odor input triggers chaotic activities which links wings, where if it is compiled, it forms another wing. Since a spatial pattern itself is not a representation of odor memories, no form of wing is any representation of memories. A process of creating wings seems to correspond to the memory representation. If it is true, memories are created at the bulb's level. Freeman has argued the dynamic recursive process of motor command, reafference and perception, where chaotic activities were placed as a dynamic machinery to serve readiness for voluntary movement and self-controlled perception [5].

The linking process of memories in Freeman's experiment can be interpreted in terms of chaotic itinerancy. Several examples about chaotic itinerancy have been proposed in neural network model [6], [7], [8], coupled map lattices [9], and ring resonator optical systems [10], [11]. This is the behavior that dynamical orbits exhibit a chaotic transition among the ruin with past history. The orbits are attracted, at the first stage, in every eigendirection, but destabilized, at the second stage, in some subspace. Since the system is dynamical, such a state-dependent destabilization must occur in subspace, not in whole space. Thus, the system can be described by a few dominant variables when the orbits are attracted to some manifold, but during the transition process, the

system needs a large dimension.

In theoretical framework, chaotic itinerancy can be compared with a random transition triggered by external noise. Actually, an itinerant motion can be observed also in the system driven by external noise. In general, however, a stochastic process is represented by Langevin equation, where the dynamical part of the motion and the heat bath are well separated, whereas in chaotic itinerancy, both components seem to be related, namely the interference between the dynamical and the noise terms is inevitable. The dynamical orbit will be driven by noise, noise will be reflected by dynamical motion, and consequently, a self-consistent interplay between dynamical motion and noise would take place[12]. Phenomenologically, chaotic itinerancy can be characterized by a composite of low dimensional attractors and low dimensional unstable manifold distributed in high dimensional phase space.

The focus here is on the possibility that chaotic itinerancy serves for the efficient works in biological neural networks. Actually, Freeman has interpreted in the olfactory system that the itinerant motion is directed by inputs biased from sensory receptors and from other parts of cortex to the bulb [4], [13]. In order to memorize inputs, the effective search in phase space for not only looking for the input-related memory but also forming a new attractor at an adequate position in phase space can be performed by chaotic itinerancy.

3. A model for CCM

We proposed a skeleton neural network for CCM[8]. Here, a state of the network is represented by a vector whose component indicating the activities of each neuron takes a continuous value on $[-1,1]$. The input pattern is supposed to be transformed to a specific state of the network, i.e. a vector composed of $+1$, or -1 . Thus, if the network which are responsible for inputs and/or outputs consist of n neurons, patterns memorized or to be memorized are represented by vertices of n -cube, and then the phase space is the internal including boundary of n -cube. A model is made, based on neuro-anatomical[14] and neuro-physiological data[15]. In making a model for CCM at the level of neural networks, we took the following principle. The common feature of the

cortical dynamics emerges, based on the common structure of the networks to all areas of the cortex, whereas a specific feature emerges, based on a specific structure in each specific area such as visual cortex, auditory cortex, and so on. Since neural events associated with the dynamics of memory would be common to every kind of sensory perception, what we should first do is to find the common structure in the neocortex. Scientific endeavor of Szentagothai over recent thirty years leads us to find such a structure. Figure 1 is taken from his recent work[16].

— Fig.1 —

The main structure of the network is a recurrent net composed of the pyramidal cells, whose recurrence is formed by the axon collateral of the pyramidal cells. By $S_i(t)$ we denote the activity of the i -th neuron monitored at discrete time with fixed time interval of the order of 10 milliseconds. Synaptic connections among these neurons are denoted by connection matrix $(C_{ij}(t))$. Several other types of neurons are seen in Fig.1. Among others, we took explicitly in the model two types of neurons — spiny stellate and axonal tuft cells, which are denoted by $R_i(t)$ and $\phi(t)$, respectively. The connections from the spiny stellate to the pyramidal cells, d_i and those from the pyramidal to the stellate cells, e_i are fixed randomly, distributed from some negative to positive values. This choice of values stems from the effect of the random coupling via basket cells, whose coupling with the above two kinds of neuron is monosynaptic or multisynaptic. The axonal tuft cells are assumed to work specifically in the sense that it detects the stationary state of the dendritic potential of pyramidal cells.

The neural dynamics we took is as follows (eqs. (1)—(6)).

$$S_i(t+1) = f^P \left(\sum_{j=1}^{2m} C_{ij}(t) S_j(t) + d_i R_i(t) - \delta_i \phi(t) \right), \quad (1)$$

where $\phi(t) = x(t_1)$, $t_1 = \max_{t > s} \{s \mid x(s) = x(s-1)\}$, $\delta_i = 0$ if $i = 1 \sim m$ and 1 if $i = m+1 \sim 2m$,

and $x(t)$ is given by $x(t) = \sum_{j=m+1}^{2m} C_{ij}(t) S_j(t)$.

$$R_{i+l}(t+1) = f^P \left(\sum_{j=1+l}^{m+l} e_j S_j(t) \right), \text{ here } l=0 \text{ or } m. \quad (2)$$

The symbols S_i and R_i denote stochastic variables determined by the probability law,

eq.(6).

The learning algorithm is as follows.

$$C_{ij}(t+1) = C_{ij}(t) + \epsilon S_i(t)S_j(t) \text{ for } 1 \leq i, j \leq m \text{ or } m+1 \leq i, j \leq 2m, \quad (3)$$

and

$$C_{ij}(t+1) = C_{ij}(t) + \epsilon S_i(t)S_j(t) \text{ for } 1 \leq i \leq m \text{ and } m+1 \leq j \leq 2m, \\ \text{or for } m+1 \leq i \leq 2m \text{ and } 1 \leq j \leq m. \quad (4)$$

When the learning is completed at the initial stage, we take

$$C_{ij}(0) = \sum \mu S_i^{(\mu)} S_j^{(\mu)}, \quad (5)$$

where the vector $S^{(\mu)}$ is the μ -th stored pattern.

When the learning is performed from the initial stage, we take

$$C_{ij}(0) = \text{random numbers over some positive to negative values.} \quad (5')$$

In either case, we reinforce the net by n -vector patterns as the same patterns are stored in two nets, $i=1 \sim m$ and $i=m+1 \sim 2m$. We also assume $n = m$ in all simulations.

The probability law adopted in S_i and R_i is given by

$$y(t+1) = \begin{cases} f^p(x(t)) = \tanh(\gamma x(t)) & \text{with probability } p, \\ y(t) & \text{with probability } 1-p, \end{cases} \quad (6)$$

where y denotes S or R , x a dendritic potential, and γ a steepness of the sigmoid function. We call this type of dynamics a stochastic renewal of maps. Two maps are contained in eq.(6): nonlinear contraction map and linear (identity) map.

The learning is performed by a Hebbian type (eqs.(3)&(4)), thereby a memory is represented by a fixed point. Furthermore, the same scale between the dynamics of neurons and that of synapses is assumed. This is equivalent to assuming a fast synapse [17] during perception, and even learning.

4. Stochastic renewal of neurodynamics

The neurodynamics described by eq.(6) stems from a stimulus-induced stochastic

release of synaptic vesicles. In general, there are two types of stochasticity for a neuron placed in networks, except for an intrinsic stochasticity like chaos. One is stochasticity due to a randomly leaked current of adjacent neurons. This case is described by the equation,

$$y(t+1) = f(x(t) + \sigma), \quad (7)$$

where σ is a random noise. If σ obeys a Gaussian distribution, the probability of the neuron being active ($x(t) + \sigma > 0$) is given by an error function. This type of stochastic dynamics has been well studied and so it is not necessary to repeat here.

Another type is due to a *synaptic noise*. A spontaneous quantal emission of neurotransmitter was widely investigated by Eccles and his colleagues [18], where quantal emission obeys a stochastic process. This spontaneous release of vesicles, however, can only influence the post-synaptic membrane by the order of μV , which is not sufficient for activation of the membrane. On the other hand, it has been observed that the release of synaptic vesicles is still stochastic even when a large stimulation is given the pre-synaptic membrane. This stimulus-induced stochastic release of synaptic vesicles affects the post-synaptic membrane by the order of mV , which can sufficiently participate in the membrane activation. This has been observed in the goldfish Marthner cells [19], in motor neurons of mammals [20], and in mammalian hippocampal pyramidal cells [21]. Since it is not difficult to conceive the similarity in fundamental properties of neurons in neocortex, archaecortex, and subcortex, we assume this type of synapse in the present model, though no such experimental evidence has been presented so far in cerebral neocortex.

Now, let us estimate the value of probability in the recurrent nets. We discuss the simplest case. For the axon(s) of the i -th neuron, we assign the probability p_{ij} to j -th synaptic connection, keeping the normalization condition, $\sum_{j=1}^n p_{ij} = 1$. The first assumption is that the release of one synapse is sufficient for the initiation of dynamics of connected neuron. Then, the probability of the i -th neuron executing dynamics is given by

$$P_i(n) = 1 - \prod_{k=1}^n (1 - p_{ki}), \quad (8)$$

where n denotes a number of neuron. The second assumption is the uniform assignment of probability, $p_{ij} = p_i = 1/n$. Then, the equation

$$P_i(n) = 1 - (1 - 1/n)^n \quad \text{for all } i \quad (9)$$

is obtained for finite system size. If we take an infinite system size, we obtain

$$P_T = \lim_{n \rightarrow \infty} P_i(n) = 1 - e^{-1} = 0.63\cdots. \quad (10)$$

The following inequality is obvious.

$$P_T \leq P_i(n), \quad (11)$$

where the equality holds only in the case of an infinite system size.

Comparatively, in the Hopfield dynamics for a single autoassociative memory, a single spin flip in each time step must be assumed in order that steepest descent by dynamics on the potential surface makes sense. This means that one neuron executes dynamics at each time step and all the other neurons keep themselves in the same state as one step before. In such a case, $P_i(n) = 1/n$ which we call the Hopfield limit in finite system size, hence $P_H = \lim_{n \rightarrow \infty} P_i(n) = 0$.

—Fig.2(a), (b)—

Figure 2(a) shows a phase diagram of the present model, eq.(1)—(6), parametrized by the assigned probability. The presence of a critical value of probability denoted by P_c was found in numerical simulation. Below P_c , a multi-basin of fixed points appears, thereby a single autoassociation for a single stationary input is achieved, which indicates a formation of a classical autoassociative memory. Above P_c , a chaotic itinerancy takes place. The chaotic itinerancy here means a chaotic transition among fixed points with a past history. The chaotic transition in n -dimensional phase space is schematically drawn in Fig.2(b). The state of chaotic itinerancy appears in the finite interval of the probability value, $[P_c, 1)$. At $P = 1$, the chaotic itinerancy disappears and coexistence of fixed points and limit cycles emerges. The other critical value P_T given in eq.(10) was empirically larger than P_c . Thus, all the allowed values of probability lead chaotic itinerancy, which

implies successive retrieval of memories by means of chaos in the cortex.

5. Information theory for CCM

In this section, we introduce information theory for CCM. The information theory in chaotic systems were established [22]~[26], where the notion of information flow was defined and its relation to the Lyapunov exponent and K-S entropy was clarified. Especially, Matsumoto and Tsuda argued the fluctuation of information flow and showed the information-theoretical significance of mutual information from the viewpoint of fluctuation of information flow [24]. Let us, for the first time, review the essence of the theory, and then argue its application to the present neural network model.

For simplicity, we deal with one-dimensional chaotic maps, but the theory is applicable to higher dimensional maps and flows if an invariant probability distribution exists. The basic information is Kullback information,

$$I_k(p) = \sum_i p_i \ln p_i / q_i \quad (12)$$

for two probability distribution p and q . A summation is taken with respect to the subintervals of an appropriate partition. In the continuum limit, the summation in eq.(12) is replaced by integral. This gives the information content measured relative to some standard distribution q . These two distributions may be created by a single chaotic map f . In dynamical systems, an orbit is forwarded by f , while probability distribution is forwarded by Frobenius-Perron operator F . This operator is defined by

$$Fp(x) = \sum_{f^{-1}(x)=y} p(y) / |df(y)/dy|. \quad (13)$$

The information flow is defined by $I_k(p) - I_k(Fp)$. In the numerical simulation, one can replace this quantity by the summation with respect to subintervals I_i , namely $\sum_i I_k(p_i) - I_k(Fp_i)$, where p_i is the part of invariant distribution $p_0(x)$ of f on the interval I_i . One can prove the following relation.

$$I_k(p) - I_k(Fp) = \int p_0(x) \sum_i p_i(x) \ln p_i(x) dx, \quad (14)$$

where $p_i(x) = Fp_i(x) / Fp_0(x)$ that means the information gain by inverse mapping. Moreover, one can obtain the following equation if $p_0(x)$ is an absolutely continuous

invariant measure with respect to the Lebeague measure dx .

$$I_k(p) - I_k(Fp) = \int p_0(x) \ln |df/dx| dx = \text{Lyapunov exponent.} \quad (15)$$

This is just the information loss rate per one iteration of map. If $p_0(x)$ is an absolutely continuous invariant measure and f is ergodic, then this information loss is exactly the same as K-S entropy.

If one expands the state variable x into a binary digit, the information flow in chaotic systems provides an average flow of information from microscopic to macroscopic digits. Since we would like to know the detailed structure of information, we introduce another significant notion, that is, "window" given by a partition. It is natural to assume that one observes the motion of the variable only within the window. Mutual information is only one quantity by which one calculates all the information escaping from the window. Mutual information between i and j states is defined as follows.

$$I(i; j) = \sum_j p(j) \ln p(j)^{-1} - \sum_i \sum_j p(i)p(j|i) \ln p(j|i)^{-1}, \quad (16)$$

where $p(j|i)$ is a conditional probability of the orbit x_n entering in the interval j , provided the orbit having entered in the interval i one step before.

One can argue whether the fluctuation of information flow is large or small in terms of characteristic function of distribution $\chi_i(x)$, which equals $p_i(x)$ in the interval i but zero in other subintervals. If $F(\chi_i(x)) = p_0(j|i) \chi_j(x)$ holds for any x , the fluctuation is exactly zero. Hence, if this equality approximately holds, one can conclude the small fluctuation of information flow. In this condition, one can prove the following.

$$I_k(p) - I_k(Fp) \sim \sum_i \sum_j p(i)p(j|i) \ln p(j|i)^{-1}. \quad (17)$$

Therefore, in case of small fluctuation of information flow it is sufficient to know the Lyapunov exponent in order to clarify the system's information structure. Nothing new happens in such a case. However, if the fluctuation is large, namely if the relation (17) does not hold, the Lyapunov exponent, thus the K-S entropy too does no longer represent the whole structure of information, and then mutual information shows much more rich structure of information.

One can extend a mutual information $I(i; j)$ to the time-dependent one, which we

call the time-dependent mutual information.

$$I^{(n)}(j; j) = \sum_j p(j) \ln p(j)^{-1} - \sum_i \sum_j p(i) p^{(n)}(j/i) \ln p^{(n)}(j/i)^{-1}, \quad (18)$$

where $p^{(n)}(j/i)$ is a conditional probability of the orbit x_n entering in the j -th interval, provided the orbit having entered in the i -th interval n steps before. The exactly zero fluctuation means the flow of information such as the first-in-first-out. Then, the same information content is lost per each iteration, hence a linear decay of mutual information in time is expected. On the other hand, the large fluctuation of information flow allows an information mixing in, say, binary space, which gives rise to the same ratio of mutual information in time will be expected. We, indeed, showed these characteristics in logistic map and B-Z (Belousov-Zhabotinskii reaction) map [24], [25].

We calculated the time-dependent mutual information eq.(18) also in our present model. Since we obtained one-dimensional representation of chaotic itinerancy in the neural network model for cortical chaotic memory, we calculated the right hand side of eq.(18) in that one-dimensional map [8]. The overall behavior of this was an exponential decay in time, irrespective of a kind and a number of stored patterns. Thus, the information mixing is apparently seen also in the neural networks.

An effective transmission of information was assured in the case of linear chain of the chaotic maps that possess the mixing property of information [25]. This will be crucial for the arguments concerning encoding, decoding, and storage of any information which is fed from outside of the underlying systems in the brain. As is shown in Fig.3, the mechanism that assures effective transmission of information becomes clearer in unidirectionally coupled maps. Any information fed from the outside is distributed in every binary digit of the variable within the window. Due to such a mixing property, a complete information is contained in each digit. If the complete information is transmitted to the neighboring chaos by the coupling before a decay of information, then the neighboring chaos is able to reconstruct the whole information. Namely, since the survival of one digit is enough for keeping the whole information, in principle, because of the mixing property of information, the complete information is correctly transmitted to neighboring chaos and then reconstructed completely if the decay of information is slower than the

propagation speed of information. This condition holds easily in the case of the exponential or the power decay of mutual information in time. Apparently, the linear decay of mutual information violates the condition. Thus, one can conclude that any information can be stored in the dynamic behavior of coupled chaotic systems if elementary chaos has the mixing property of information. It may be useful to note that *virtual* networks of information in binary space emerge, even though *real* coupling is linear.

— Fig.3 —

Our network model is made so as to implement one column in cerebral neocortex. In the neocortex, both forward and backward connections via cortico-cortical fibers exist, but these are not symmetric. Though a columnar organization in neocortex is well-known, it is still questionable whether or not a column is a functional unit. However, if one recognizes a well-defined dynamic behavior as a dynamical unit, that is, as a dynamical order parameter, then it would be plausible to think that the chaotic behaviors such as those obtained above can be coupled each other asymmetrically in the neocortex and the composed system can work for the dynamic storage of any information, in spite of chaos. A more important point is that such a chaotic network can work for reorganization of memory space, too. This will be discussed in section 8.

6. Skew product transformation

A stochastic renewal of maps can generally be reformulated by a skew product transformation. An idea lies in a redefinition of the probability in terms of a dynamical system. We are a follower of Boltzmann, in the above sense.

Let us define neural networks as n -dimensional dynamical system on $X = [-1, 1]^n$.

$$U_i(t+1) = h_i(U_1(t), U_2(t), \dots, U_n(t)), \quad (19a)$$

or

$$U_i(t+1) = g_i(U_1(t), U_2(t), \dots, U_n(t)) \quad (19b)$$

for $i = 1, 2, \dots, n$. The probability p for the stochastic renewal process is created by chaotic systems, say, Bernoulli shift which is defined on $Y = [0, 1)$. The skew product

transformation is defined on $T = X \times Y$. Since properties of the probability have not been described in the original model, one can have various kinds of models of stochastic process. Here, we will argue two simple models: Bernoulli and Markov models.

The Bernoulli model is as follows.

$$T(U(t), y(t)) = \begin{cases} (h(U(t)), y/p) & \text{if } 0 \leq y < p \\ (g(U(t)), (y-p)/(1-p)) & \text{if } p \leq y < 1. \end{cases} \quad (20)$$

Here, neural dynamics h or g are coupled with Bernoulli shift $B(p, 1-p)$. The symbolic dynamics in Bernoulli shift which is defined by the choice of decision point, $y = p$, creates an independent symbol-sequence of any size.

The second is the Markov model.

$$T(U(t), y(t)) = \begin{cases} (h(U(t)), 2y \pmod{1}) & \text{if } 0 \leq y < p \\ (g(U(t)), 2y \pmod{1}) & \text{if } p \leq y < 1. \end{cases} \quad (21)$$

Here, neural dynamics h or g are coupled with Bernoulli shift $B(1/2, 1/2)$, but the symbolic dynamics defined by the decision point $y = p$ does no longer create an independent symbol-sequence unless $p = 1/2$ but create a Markov sequence of n -Markov, generally.

A special case of a skew product transformation has been argued by Tanaka and Ito [27]. They obtained a condition of chaos in the following transformation.

$$T_{a,b}(x, y) = \begin{cases} (f_a(x), 2y) & \text{if } 0 \leq y < 1/2 \\ (f_b(x), 2y-1) & \text{if } 1/2 \leq y < 1. \end{cases} \quad (22)$$

Two tent maps of each having different values of parameter α , $f_\alpha = \alpha x$ ($0 \leq x < 1/2$), and $-\alpha x + \alpha$ ($1/2 \leq x < 1$) are coupled with Bernoulli shift $B(1/2, 1/2)$. They proved a theorem that T is ergodic and has an absolutely continuous invariant measure if and only if $ab > 1$. This condition is obvious, because the condition apparently means the positiveness of an averaged logarithm of absolute values of derivatives of maps, that is, $1/2 \log a + 1/2 \log b > 0$, hence a positive Lyapunov exponent.

By this argument, it is also obvious that the skew product between neurodynamics and Bernoulli shift cannot assure a positive Lyapunov exponent if both h and g are a contraction map. In our model, at least one of h and g is a complicated function, thereby positive Lyapunov exponents can appear, as will be shown below. Even though both h

and g are a contraction map, however, a chaos-driven contraction map creates a curious fractal - like structure on an attractor [28]. Having this new topic in complex chaotic systems into mind, one may expect new mathematical theorems also in the condition of $ab < 1$. Moreover, one may obtain positive Lyapunov exponents even in the case of both functions being a contraction map, if one takes a stationary set composed of pieces of transient motions. This case was actually found in a periodic or a random switch of bifurcation parameters of logistic map [29].

If the noise is added to the term of dendritic potential of each neuron, the overall dynamics is not given by skewproduct transformation, but by direct product transformation. Here, a direct product transformation is defined by $h = g$ for all $U(t)$ in eqs.(20) and (21). This means, therefore, that each one of the transformations is independent of the other transformation's variables. In the present context, the skew product transformation describes the stochastic behaviors created by neurodynamics dependent on the noisy variable which may stem from noisy synapse, while the direct product transformation describes the stochastic behaviors created by noisy dynamics composed of neurodynamics and dendritic noise, both of which are independent. Thus, the two types of stochastic dynamics can be formulated by two distinct transformations, thereby the dynamic behavior itself will be distinguished.

7. Phase space structure

It would be valuable to see a phase space structure of a single neuron's dynamics, though the overall behavior is created by the interactions of such a dynamics. In order to see the fundamental vector field, we rewrite the dynamics concerning the deterministic parts only, eqs. (1) - (5), viewed a dendritic potential of neuron as a new variable. The R - unit plays a role of a self-consistent field for the S - unit if the coupling constants $\{d_i\}$ and $\{e_{ij}\}$ are weak, relative to the other couplings $\{C_{ij}\}$. Now suppose that this effect is expressed as simply a time-dependent external field and denote it by $\sigma(t)$. Replacing $\sum_i C_{ki} S_i$ by \bar{u}_k , we obtain the following map.

$$u_k(t+1) = \sum_j C_{ki} f(u_j(t)) - 1/N_{t,t'} \sum_{\tau=t'}^t \theta(du_k(\tau)/d\tau) u_k(\tau) + \sigma_k(t), \quad (23)$$

where $\theta(du_k(t)/dt) = 1$ if $du_k(t)/dt = 0$, and otherwise zero, N denotes a number of $\theta = 1$ during the time interval $[t', t]$, and t' is the nearest time of t among t'' such that both $\theta(du_k(t)/dt |_{t=t'}) = 1$ and $\theta(du_k(t)/dt |_{t=t'-1}) = 0$ are satisfied. Furthermore, replacing the argument of f by $C_j V_j(t)$, where $C_i \equiv C_{ij}$, we obtain the following dynamics.

$$V_k(t+1) = W_k(t)$$

$$W_k(t+1) = \alpha(t, t') W_k(t) + \alpha(t, t')/C_i \sum_j C_{ij} [f(C_j W_j(t)) - f(C_j V_j(t))] + \varepsilon \sigma_k(t), \quad (24)$$

where $\alpha(t, t') = N_{t,t'} / N_{t,t'+1}$ when $\theta(du_k(\tau)/d\tau |_{\tau=t+1}) = 1$, and $\alpha(t, t') = 1$ when $\theta(du_k(\tau)/d\tau |_{\tau=t+1}) = 0$, and ε is a renormalization factor of the fields.

Neglecting the field term and taking into account only a single neuron, one can investigate a phase space structure. A single neuron dynamics is as follows under the above assumption.

$$V(t+1) = W(t)$$

$$W(t+1) = \alpha(t, t') W(t) + \alpha(t, t') [f(C W(t)) - f(C V(t))], \quad (25)$$

where $1/2 \leq \alpha(t, t') \leq 1$.

In the case of $\alpha(t, t') = 1$, which we call a chaotic regime, there is a fixed line $V = W$, on which eigen values are 1 and $C\gamma / \cosh^2(C\gamma V)$ if we take $f(W) = \tanh(\gamma W)$. Jacobian J is given by $J = C\gamma / \cosh^2(C\gamma V)$, hence the volume expansions in the region of $J > 1$ occurs in the finite region centered at the origin, and the volume contraction in the region $J < 1$ occurs in the outside region of $J > 1$. Accordingly, one direction is unstable in the region $J > 1$ and stable in $J < 1$. The direction of the oblique line $V = W$ is always marginally stable. Two curves where eigen values are degenerate appear, and on which curves $J = 1$ holds, too. These curves are here called a "degenerate curve". These features are shown in Fig. 4.

- Fig.4 -

In the case of $\alpha(t, t') = N_{t,t'} / N_{t,t'+1}$, which we call a pulsatile regime, there are neither fixed points nor fixed lines. Jacobian is given by $J = \alpha(t, t') C\gamma / \cosh^2(C\gamma V)$. The degenerate curves change their positions in a pulsatile fashion, according to the development of the dynamics.

These two regimes are exchanged, depending on the history of the orbits. Thus, the overall deterministic dynamics allows the reinjection of orbits into the unstable region by the pulsatile movements of the degenerate curves, which may give rise to the positive Lyapunov exponent. The possible orbits projected in two-dimensional space, $V - W$, are shown in Fig. 5. These motions will be coupled in the case of n - coupled neurons.

- Fig.5 -

Furthermore, these complex dynamics are contaminated by the stochastic renewal of dynamics. We will show that the skew product transformation can give much more efficient performance for learning than the direct product transformation, if the above structure is constructed in phase space. Such a structure may assure an "intelligent" noise that enhances the efficiency of performance, by rectifying the fluctuations.

8. Biological significance of CCM

8.1 Chaotic novelty filter

We showed in the previous paper [8] that the chaotic itinerancy in our model can work as a novelty filter. In conventional model for associative memory, incomplete information which is put in the network is completed; the complete memory state is the output of the network. However, the network state is often pinned to metastable states which has nothing to do with any memory states, where a metastable state is called a parasitic state. Hence, one cannot have any indications of whether the input pattern is close to either memory state or not, unless one has obtained, in advance, a complete information of memories. Moreover, in such a static network, it is not easy to recover the correct output, once the network state is pinned to parasitic state.

The dynamics in our model shows a different performance. Suppose one pattern is fed into the network when it shows chaotic itinerancy. If such a pattern is within a basin

of attraction of some memory state, the network output is just such a memory state. While, if the input pattern is outside any basin of memory state, the output is again a chaotic itinerancy. Thus, the state of chaotic itinerancy is a "novelty filter" which can judge whether or not the input should be viewed as one of the stored patterns. Moreover, no states pinned to parasitic states appear, since the system's behavior is always dynamic. In this dynamic process, the closeness of patterns is automatically determined, hence the closeness is not practically controllable. This might become a weak point of our model for practical use, but could provide a plausible interpretation for mental judgement.

8.2 Enhancement of learning ability

In the previous paper [8], we showed that the network can learn even in the chaotic itinerancy via Hebbian learning algorithm. This implies that chaos does not hamper the learning. Here, we report a more drastic result of computation.

We performed a comparative study in numerical experiments concerning the possibility of additional Hebbian learning. For the first time, several patterns are learned up to the critical number of storage that depends on the system size, i.e., the number of neurons participated in the synaptic plasticity. Therefore, in this condition, an additional learning is impossible in the sense that either of the followings holds: a new pattern is learned but such a learning destroys some or all stored patterns, or any new patterns cannot be learned, preserving all stored patterns. This preliminary learning was done only in the deterministic terms of the model. Then, we performed two types of experiments.

The first type of experiments is associated with the direct product transformation, namely we add a noise term to the term of the dendritic potential in place of applying the stochastic renewal of dynamics. It is not easy to tune the noise amplitude for keeping a continual transition among memory states. If the noise amplitude is small, it will take an extremely long time for the transition, thus the noise amplitude must be raised for a rapid transition, but almost the cases break keeping the memory states. We have to choose, therefore, an intermediate size of noise. After a choice, we have the network perform the

additional learning. The result is shown in Fig. 6(a). As is expected, additional learning was impossible in the above sense. In the figure, the phase III denotes that no new patterns are learned, but all stored patterns are preserved, and the phase I denotes that a new pattern is learned, but such a learning destroys all stored patterns. The region between these two phases is divided into subregions showing the hierarchy such that the number of destroyed patterns accompanied with a success of new learning increases toward the phase I.

- Fig. 6(a), (b) -

The second type of experiment is associated with the skew product transformation, namely the stochastic renewal of dynamics, eq.(6) is used. All the other conditions are the same as in the first type of experiments. As is shown in Fig. 6(b), the new phase II appears, where new patterns are correctly learned and all stored patterns are still preserved. In the experiments with 32 effective plastic neurons, three more patterns were additionally learned for six patterns, here the number six is a practical memory capacity in both conventional static neural nets and the noisy nets of the first type of experiment. Other preliminary experiments (still in small system) showed the same tendency. Thus, the memory capacity, or the learning capability of the nets can be enhanced by an emergence of chaotic itinerancy. However, we have not yet investigated how much percentage of memory capacity increases in case of large size of the networks. Furthermore, no mathematical theory which explains this feature is known.

The following is a rough sketch of the inherent dynamics. We used a Hebbian learning algorithm, which is local in phase space. By this locality, the stability of fixed points representing memories tend to be strengthened in conventional static neural nets.

If the noise is added in the additive form (like the first type of experiment), the phase space place representing a current network state can also be a candidate for its stability being strengthened since transition probabilities to any place in phase space are almost equal. On the other hand, the learning during the chaotic itinerancy must be different. The transition is directional, namely the transition probabilities are no longer equal. Thus, the notion of orbits is partially recovered, in spite of noisy dynamics. Consequently, the *orbits* linking memory states can also be stabilized, in spite of the

local learning rule. Due to this reinforcement of the orbits, the memory states remain stable even if an additional pattern is learned. The overall dynamics behaves as if appropriate place for new memories is searched efficiently by chaotic itinerancy.

The possibility of this additional learning leads necessarily to the *reorganization of basin boundaries*, which brings about the reorganization of memory space. Namely, a new itinerant motion after additional learning takes place among all stored patterns, including newly learned patterns. However, no clear change was seen in one-dimensional chaotic map during additional learning. Thus, one-dimensional representation (see [8]) is too coarse-grained one in order to see the change of basin boundaries. An elaborated analysis is necessary for this.

9. Concluding remarks

We finally argue one problem of perception developed in cognitive psychology, in relation with chaotic itinerancy found in the neural network model which was made for explaining CCM. The schema theory of Bartlett [30] includes not only static viewpoints but also dynamic ones. On one hand, schema is defined as an organized set of past experiences and responses. Based on this definition, memories are simply a storeroom of traces. On the other hand, however, schema can develop continually, accompanied with a dynamic change of the above organization. This is caused by performance, for instance. In this dynamic process, the recall of memories, or remembering is also perception [31], for the recall is one of events evoked in accordance with the modalities of subject's conscious mind. This leads to the notion of perceptual drift [30] accompanied with memory dynamics.

In the present neural network model, the overall activities of the network as a response to some stimulus obeys a self-organized dynamic rule. From megal- (super coarse-grained) viewpoint, this rule can be expressed by chaotic map [8]. Furthermore, the network exhibits chaotic itinerancy as the network macroscopic state in the process of recall. The basin structure in phase space changes if the learning is set up, thereby the chain of retrieved memories also changes. This assures a variability of brain activity patterns discussed widely by Freeman [5].

By the equivalence of perception and recall of memories, the *dynamical orbits* of chaotic itinerancy can be a representation of perception, whereas the fixed points in phase space represent elementary memories. Namely, elementary memories constitute a static part of schema and the linking process of memories organizes its dynamic part. Thus, the notion of chaotic itinerancy provides a neurobiological basis of dynamic perception.

Acknowledgements

I would like to express my special thanks to Walter Freeman and Kunihiro Kaneko for continual encouragement and valuable discussions. To Janos Szentagothai I also express my special thanks for giving a freedom of use of his artistic work. I also thank Kazuhiko Shite for double checking of the numerical experiments discussed in the section 8. I thank Kazuhiro Matsuo for hospitality and stimulation.

References

- [1] W. J. Freeman, *Biol. Cybernetics* **56** (1987) 139.
- [2] C. A. Skarda and W. J. Freeman, *Behav. and Brain Sci.* **10** (1987) 161.
- [3] Y. Yao and W. J. Freeman, *Neural Networks* **3** (1990) 153.
- [4] W. J. Freeman, in this volume.
- [5] W. J. Freeman, *Scientific American* **264** (1991) 78.
- [6] I. Tsuda, *World Futures* **32** (1991) 167.
- [7] I. Tsuda, *Neurocomputers and Attention I*, eds. A. V. Holden and V. I. Kryukov (Manchester University Press 1991) 405.
- [8] I. Tsuda, *Neural Networks* **5** (1992) 313.
- [9] K. Kaneko, *Physica D* **41** (1990) 137.
- [10] K. Ikeda, K. Otsuka, and K. Matsumoto, *Prog. Theor. Phys. Suppl.* **99** (1989) 295.
- [11] P. Davis, *Jpn J. Appl. Phys.* **29** (1990) L1238.
- [12] K. Ikeda, in *The Proc. of Intl. Symp. on Information Physics*, eds. I. Tsuda and K. Takahashi (Kyushu Institute of Technology 1992) 101.
- [13] W. J. Freeman, in *The Proc. of the 2nd Intl Conf. on Fuzzy Logic & Neural Networks* (Fuzzy Logic Systems Institute, Iizuka 1992) 933.
- [14] J. Szentagothai, *Rev. of Physiol., Biochem., and Pharmacol.* **98** (1983) 11.
- [15] See, for example, J. C. Eccles, *The Understanding of the Brain* (McGraw-Hill, New York 1977).
- [16] J. Szentagothai, *J. Theor. Medicine* **14** (1993) 101.
- [17] C. von der Malsburg, in *Brain Theory*, eds. G. Palm and A. Aertsen (Springer-Verlag, Berlin Heidelberg) 161.
- [18] See, for example, P. Fatt and B. Katz, *J. Physiol.* **117** (1952) 109.
- [19] H. Korn, A. Mallet, and D. S. Faber, *J. Neurophysiol.* **48** (1982) 679.
- [20] J. C. Eccles, in *Complex Systems — Operational Approaches*, ed. H. Haken (Springer-Verlag, Berlin Heidelberg 1985) 81.
- [21] C. Yamamoto (unpublished).
- [22] R. Shaw, *Z. Naturf.* **36a** (1981) 80.
- [23] J. P. Crutchfield and N. H. Packard, *Intl. J. of Theor. Phys.* **21** (1982) 433.

- [24] K. Matsumoto and I. Tsuda, *J. Phys. A: Math. Gen.* **18** (1985) 3561.
- [25] K. Matsumoto and I. Tsuda, *Physica D* **26** (1987) 347.
- [26] K. Matsumoto and I. Tsuda, *J. Phys. A: Math. Gen.* **21** (1988) 1405.
- [27] S. Tanaka and S. Ito, *Tokyo J. Math.* **58** (1982) 463.
- [28] O. E. Röessler, C. Knudsen, J. L. Hudson and I. Tsuda, *Intl. J. of Intell. Systems* (in press 1993).
- [29] J. Röessler, M. Kiwi, B. Hess and M. Markus, *Phys. Rev. A* **39** (1989) 5954.
- [30] F. C. Bartlett, *Remembering : A Study in Experimental and Social Psychology* (Cambridge Univ. Press, Cambridge 1932).
- [31] J. J. Gibson, *The Ecological Approach to Visual Perception* (Houghton Mifflin, Boston, MA 1979).

Figure Captions

Fig. 1

Typical cells and networks in mammalian neocortex, drawn by Janos Szentagothai (handed over in June 30th, 1992). Reproduction through his courtesy. Original publication with black-and-white in *J. Theor. Medicine* 14 (1993) 101-116: Kluwer Academic Publishers Group. Reprinted by permission of Kluwer Academic Publishers. Color coded by Janos Szentagothai.

Fig. 2

(a) Phase diagram. Abscissa denotes assigned probability value and ordinate a set of orbits. Details in the text.

(b) Schematic drawing of itinerant motion among fixed points. Unstable manifold is formed temporally in subspace of n -dimensional phase space.

Fig. 3

Schematic drawing of information structure in linear chain of the chaotic map that possesses mixing property of information.

Fig. 4

The feature of vector field in two-dimension of the single deterministic neuron. Arrows indicate stability. The dotted curve denotes "degenerate curve" explained in the text.

Fig. 5

Allowed orbits in the single neuron dynamics affected by the stochastic renewal.

Fig. 6

The success of additional learning with chaotic behavior. Abscissa indicates learning time and ordinate learning strength, i.e., the increment of synaptic learning.

(a) the case of additive noise.

(b) the case of stochastic renewal of dynamics.

WIDE RANGE
> 1mm

MEDIUM RANGE
< 0.5mm

NARROW COLUMNS
~ 50 μ m

SPOTS (GABA)
~ 200 μ m DIAM.

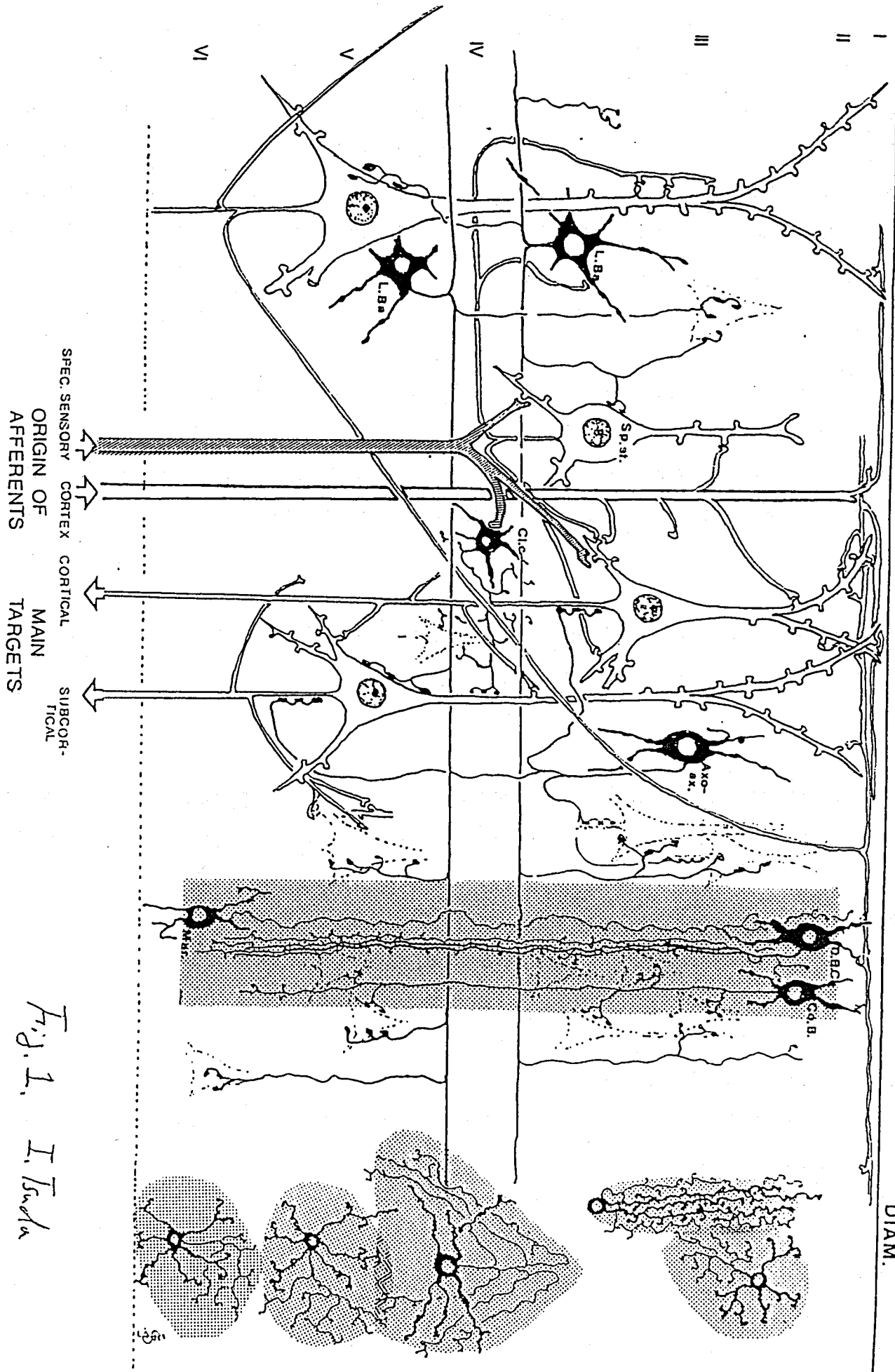


Fig. 1. I. Isoda

(a)

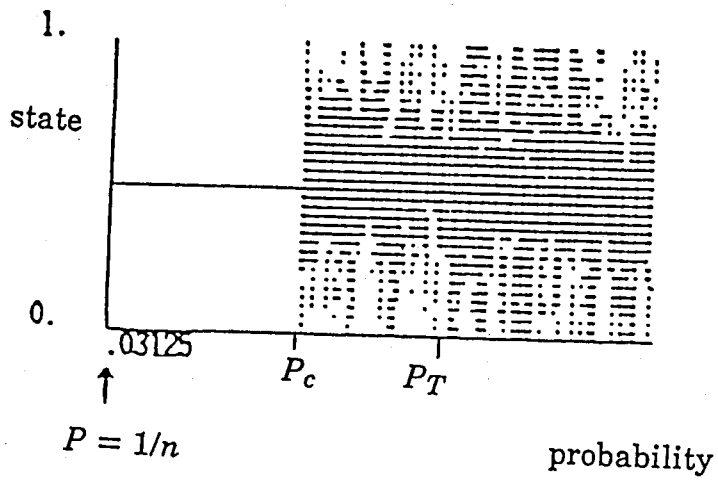


Fig 2 (a)

I. Tsuda

(b)

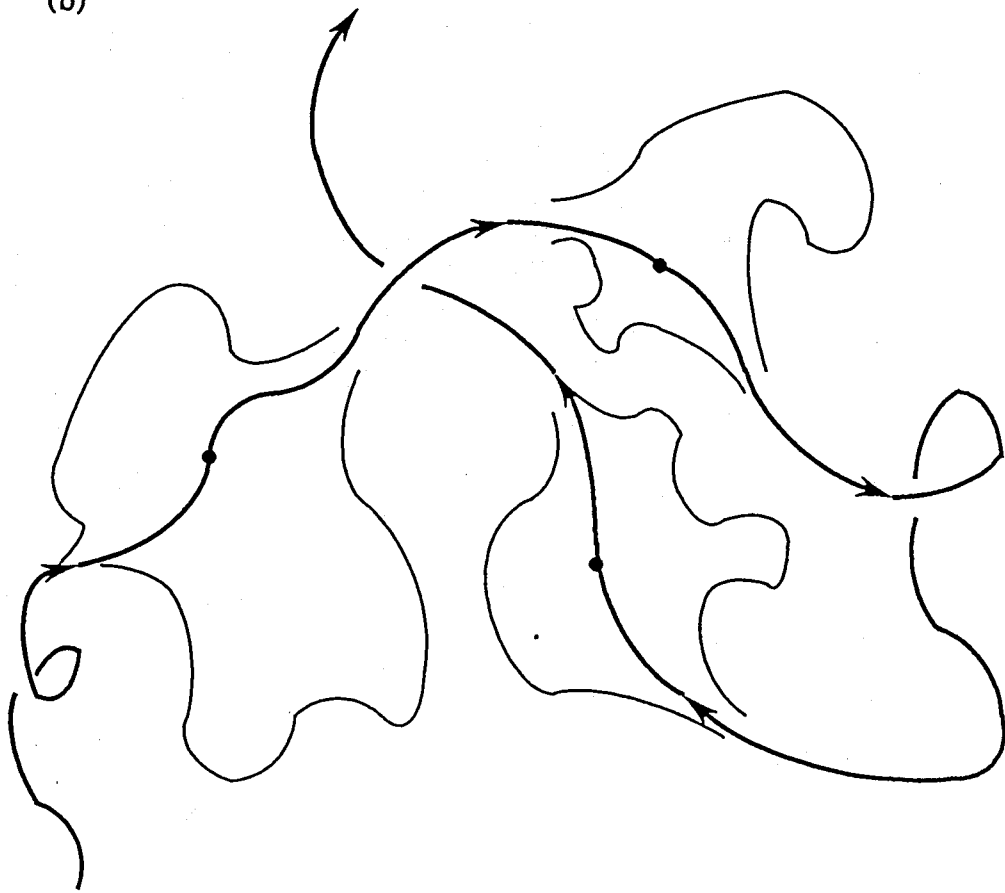


Fig. 2 (b) I. Tsuda

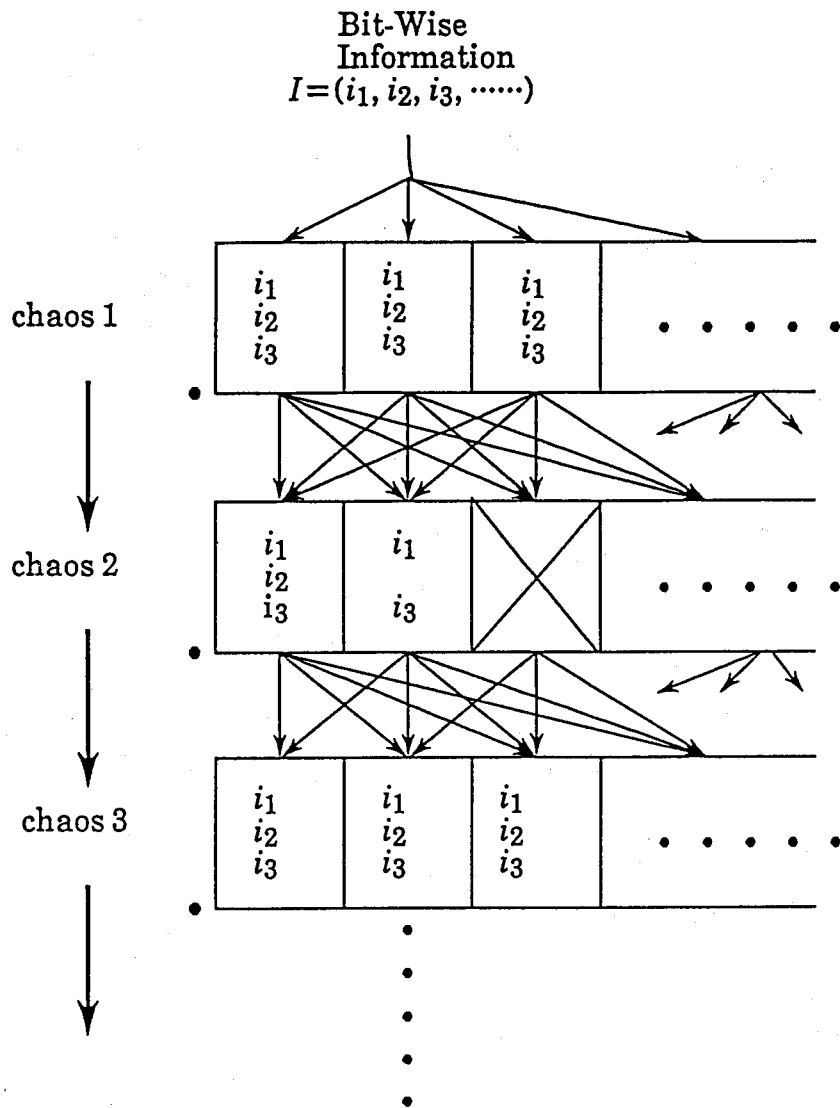


Fig. 3. I. Tsuda

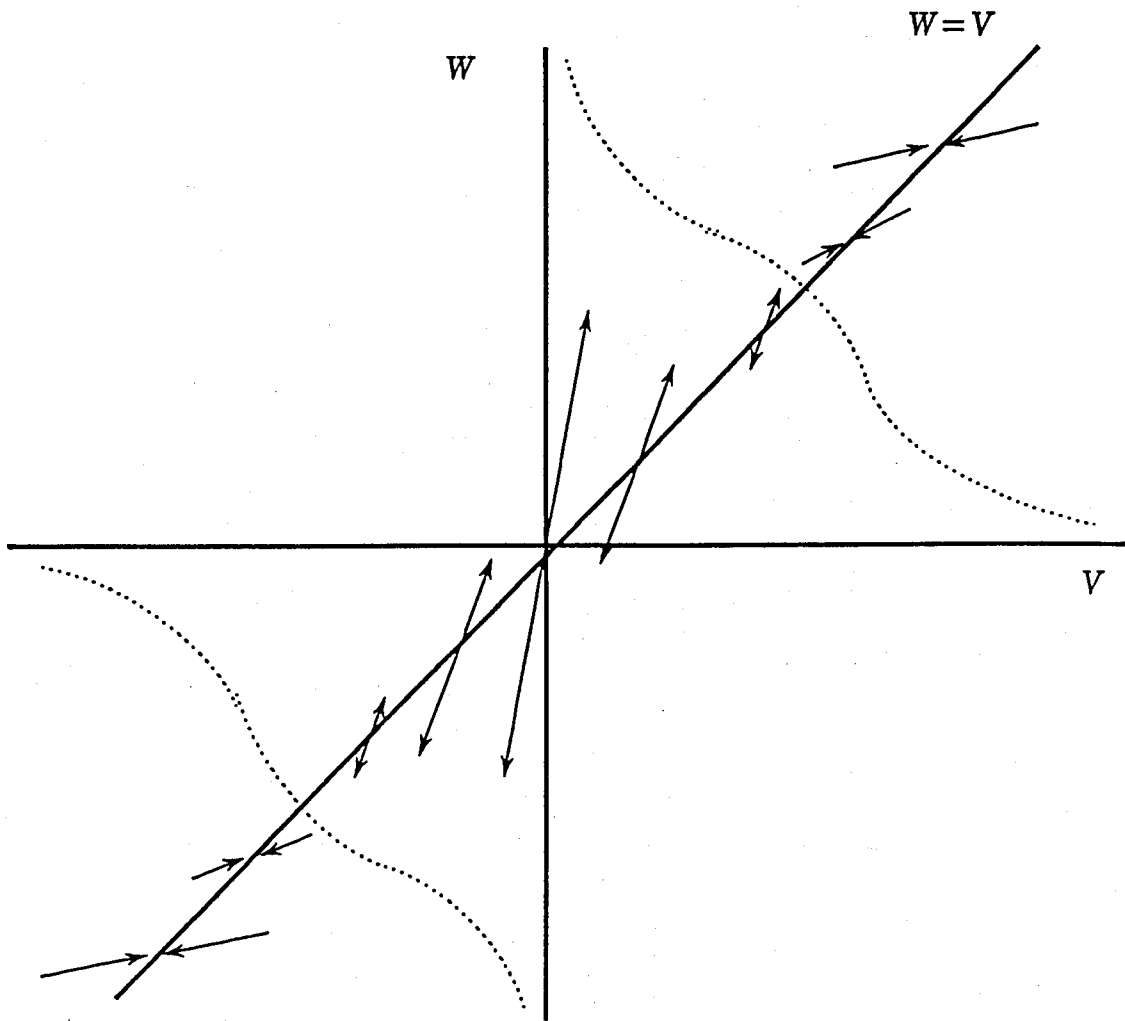


Fig 4. I, Tsuda

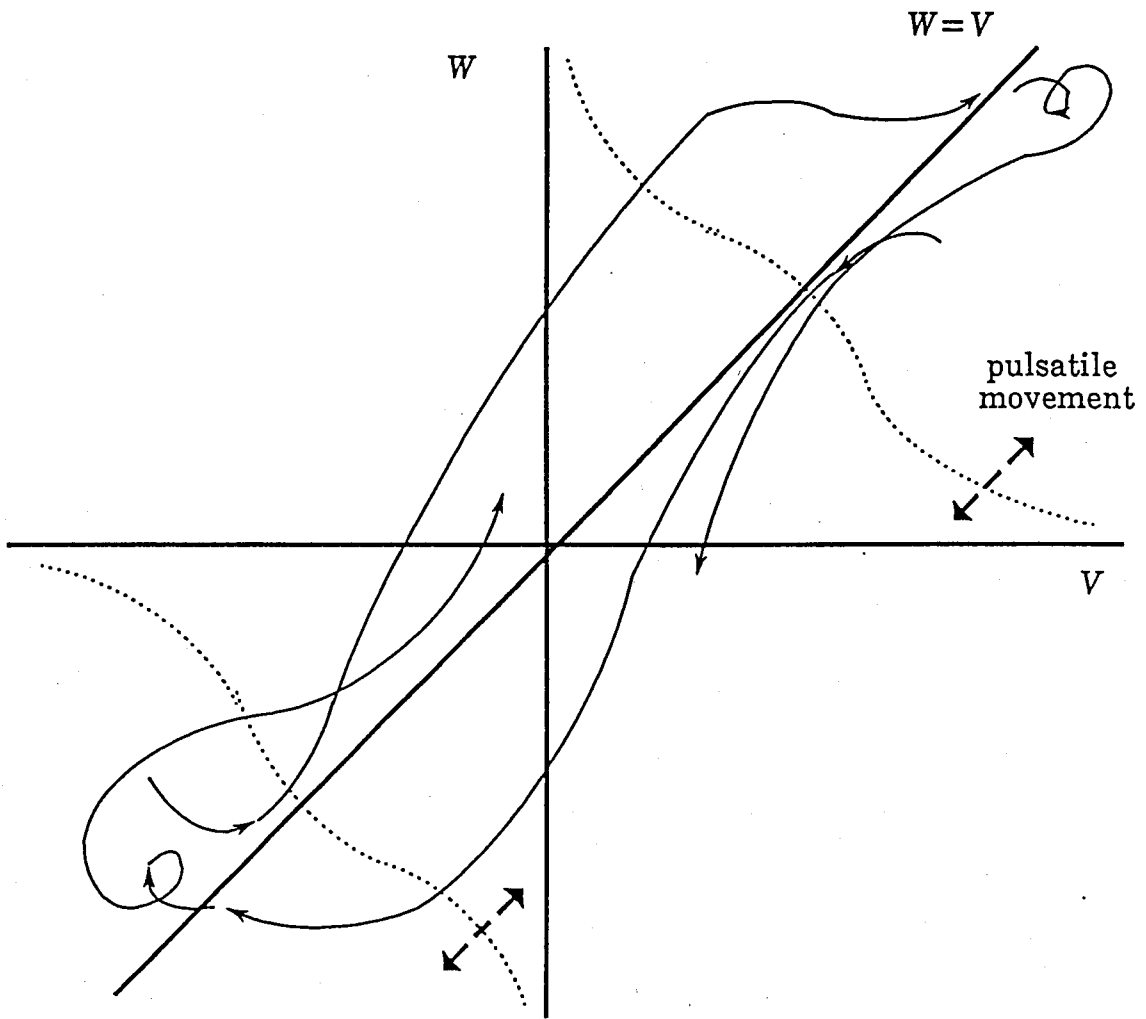


Fig. 5 I. Tsuda

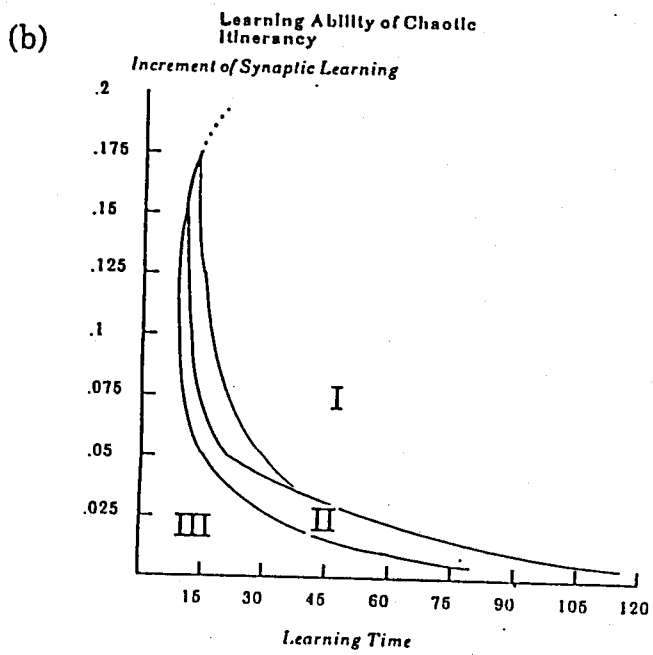
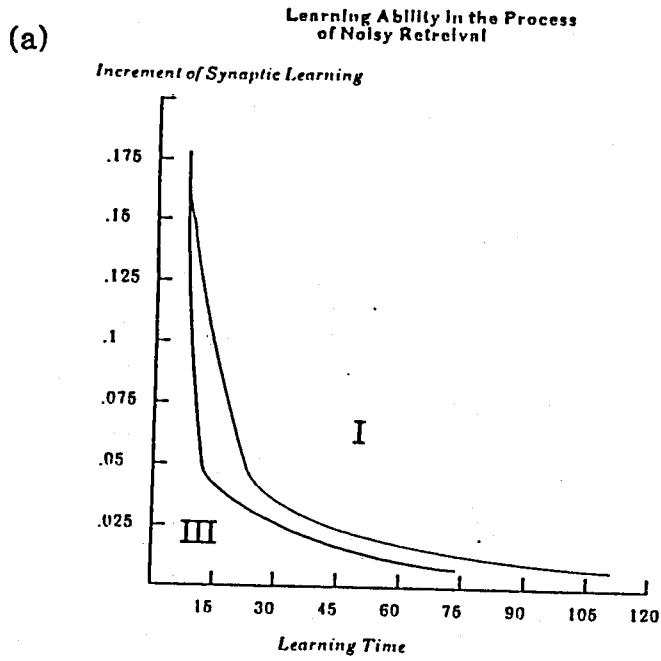


Fig 6. (a) & (b)

I, Tsuda