2004 International Symposium on Nonlinear
Theory and its Applications (NOLTA2004)
Fukuoka, Japan, Nov. 29 - Dec. 3, 2004

Information retrieval based on a neural-network system with continuous attractors

Yukihiro Tsuboshita and Hiroshi Okamoto

Corporate Research Laboratory, Fuji Xerox Co., Ltd.
430 Sakai, Nakai-machi, Ashigarakami-gun, Kanagawa 259-0157, Japan
Email: Yukihiro.Tsuboshita@fujixerox.co.jp, hiroshi.okamoto@fujixerox.co.jp

Abstract– Memory retrieval in neural networks has traditionally been described by dynamical systems with discrete attractors. Evidence raised by recent neurophysiological findings of graded persistent activity, however, suggests that information retrieval in the brain is more likely to be described by dynamical systems with continuous attractors. We propose a neural-network system that has continuous attractors with respect to the network-activation pattern. In this system, the attractor pattern continuously depends upon the initial pattern; it also reflects learned patterns. The usefulness of information encoded by the attractor pattern is demonstrated by applying our system to key extraction from a document. Thus, our model presents a novel information-retrieval design inspired by neuroscience.

1. Introduction

It is traditionally considered that storage of short-term memory in the brain is formed by sustained reverberatory activation of an ensemble of neurons. Such activation can emerge as an attractor of a multi-stable dynamical system describing a neural network in which multiple distributed patterns are embedded [1, 2].

In the multi-stable dynamical system, for a given pattern represented by an initial state, one of the embedded patterns, which is the nearest to it, is retrieved. The state space is divided into multiple attractor basins; the attractor on which the state point settles depends upon the basin to which it initially belongs (Fig. 1a).

However, recent neurophysiological findings of graded persistent activity challenge to this traditional paradigm. The firing rate of neurons recorded from the prefrontal cortex of the monkey performing vibrotactile discrimination task varied, during the delay period between the base and comparison stimuli, as a monotonic function of the base stimulus frequency [3]. The firing rate of neurons in the oculomotor system of the goldfish during fixation was associated with the history of spontaneous saccadic steps [4]. These phenomena cannot be described simply by multi-stable dynamical systems with discrete attractors. They are more likely to be described by dynamical systems with attractors that continuously depend upon the initial state (Fig. 1b). (For overall survey of graded persistent activity, see [5]).

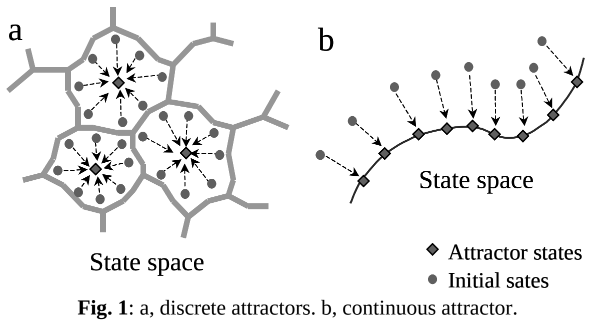

Fig. 1: a, discrete attractors. b, continuous attractor.

The purpose of this study is to infer new information-processing design from supposed neural mechanisms of graded persistent activities. Several attempts have already been made to build models for neural mechanisms that generate continuous attractors [6, 7]. However, their discussion is confined to the continuity with respect to scalar quantities such as the firing rate of individual cells. Nevertheless, rich information processing should exploit more complex quantities such as the network-activation pattern represented by a vector. Here we propose a model for neural mechanisms that generate continuous attractors with respect to the network-activation pattern. To demonstrate the functional significance of this process and its technological implication, our system is applied to a typical document-processing task, keyword extraction from a document.

2. Model

2.1. Previous models

Several models for neural mechanisms that can generate continuous attractors have already been examined in previous works. In the mechanisms proposed by Seung et al. [6], the parameter values are finely tuned so that friction vanishes along a certain line in the state space. This line constitutes continuous attractors (line attractors). However, continuous attractors attained by their mechanisms are marginally stable and vulnerable to slight change in parameter values.

Koulakov et al. [7] proposed mechanisms that ensured stability of continuous attractors. They considered a recurrent network consisting of bistable neurons. Each neuron is either in the 'off' state or in the 'on' state. In the ‘off’ state, the firing rate is low; in the ‘on’ state, the firing rate increases with the input from presynaptic neurons. In this model, the synaptic weights are uniform everywhere in the network. Nevertheless, if different parameter values are properly associated with each neuron, they can move to the ‘on’ state not simultaneously but one by one in a fixed order. Owing to the neuronal bistability, each state of the network appearing in this series is stable. If the number of neurons constituting the network is sufficiently large, continuous attractors can approximate this series of stable states. Therefore, the firing rate of each neuron or the network activity defined by the average firing rate across neurons is continuously associated with the number of neurons in the ‘on’ state.

In the above mechanisms, discussion is focussed on continuous attractors with respect to scalar quantities such as the firing rate of individual cells. This is because their main purpose is to account for experimental observation of graded persistent activity. In usual neurophysiological experiment, the firing rate of a single cell, which is represented by a scalar quantity, is examined.

However, rich information processing should exploit more complex quantities such as vector quantities. It is highly probable that, although not observed by present experimental techniques, the continuity with respect to the activation pattern of an ensemble of neurons, which can be represented by a vector, underlies observed graded persistent activity.

2.2. Our model

In this short article, we provide an outline description of our model. A detailed description of the model is given in [8]. Following Koulakov et al. [7], we consider a network of N bistable neurons. For minimal modelling of the neuronal bistability, each neuron is described by a two-spin Ising system [9]. The strength of connection between neuron i and neuron j is defined by the covariance-learning rule [10]:

$$T_{ij} = \frac{1}{P}\sum_{p=1}^{P}\left(\xi_i^{(p)} - <\xi_i>\right)\left(\xi_j^{(p)} - <\xi_j>\right) \quad (1)$$

where $\xi_i^{(p)}$ is the i-th component of the p-th pattern and $<\xi_i> = \sum_{p=1}^{P}\xi_i^{(p)} \Big/ P$.

The network-activation pattern is represented by a vector $\vec{S}$. The i-th component S_i represents the state of neuron i: If the neuron is in the active state and transmitting signals to postsynaptic neurons, $S_i = 1$; otherwise, $S_i = 0$. The activity of neuron i itself is defined by the input from presynaptic neurons [7]:

$$I_i = \sum_{i=1}^{N} T_{ij} S_j \quad (2)$$

To examine the continuity with respect to the network-activation pattern, we calculate the correlation between the current pattern $\vec{S}$ and an initial pattern $\vec{S}^{(ini)}$:

$$C\left(\vec{S}, \vec{S}^{(ini)}\right) = \frac{\sum_{i=1}^{N}\left(S_i - <S>\right)\left(S_i^{(ini)} - <S^{(ini)}>\right)}{\sqrt{\sum_{i=1}^{N}\left(S_i - <S>\right)^2}\sqrt{\sum_{i=1}^{N}\left(S_i^{(ini)} - <S^{(ini)}>\right)^2}} \quad (3)$$

where $<S> = \sum_{i=1}^{N} S_i \Big/ N$.

3. Results

3.1. Dynamical properties of our model

First, we confirm that our model can produce the continuity of attractors with respect to the vector properties. For simplicity, we consider a simple network in which only a single pattern $\vec{\xi}$ is embedded:

$$T_{ij} = \left(\xi_i - <\xi>\right)\left(\xi_j - <\xi>\right) \quad (4)$$

where $<\xi> = \sum_{i=1}^{N}\xi_i \Big/ N$.

For different patterns of $\vec{S}^{(ini)}$, the time course of $C\left(\vec{S}, \vec{S}^{(ini)}\right)$ was calculated by computer simulation. The results obtained show that $C\left(\vec{S}, \vec{S}^{(ini)}\right)$ continuously depends on $\vec{S}^{(ini)}$ in the attractor state (Fig. 2a). This means that the attractor pattern reflects information of the initial pattern in a graded manner. Thus, our model does generate continuous attractors.

We also examined the time course of the correlation between $\vec{S}$ and $\vec{\xi}$:

$$C\left(\vec{S}, \vec{\xi}\right) = \frac{\sum_{i=1}^{N}\left(S_i - <S>\right)\left(\xi_i - <\xi>\right)}{\sqrt{\sum_{i=1}^{N}\left(S_i - <S>\right)^2}\sqrt{\sum_{i=1}^{N}\left(\xi_i - <\xi>\right)^2}} \quad (5)$$

The $C\left(\vec{S}, \vec{\xi}\right)$ increases with time and then saturates (Fig. 2b). This means that attractor states reflect information of the embedded pattern. However, unlike in the case of the original Hopfield model [1], the attractor pattern is not the embedded pattern itself. Thus, our model presents a novel design to retrieve information in a graded manner from a neural network system.

3.2. Application to a keyword-extraction task

To demonstrate the usefulness of information encoded by the attractor state thus retrieved, we next apply our

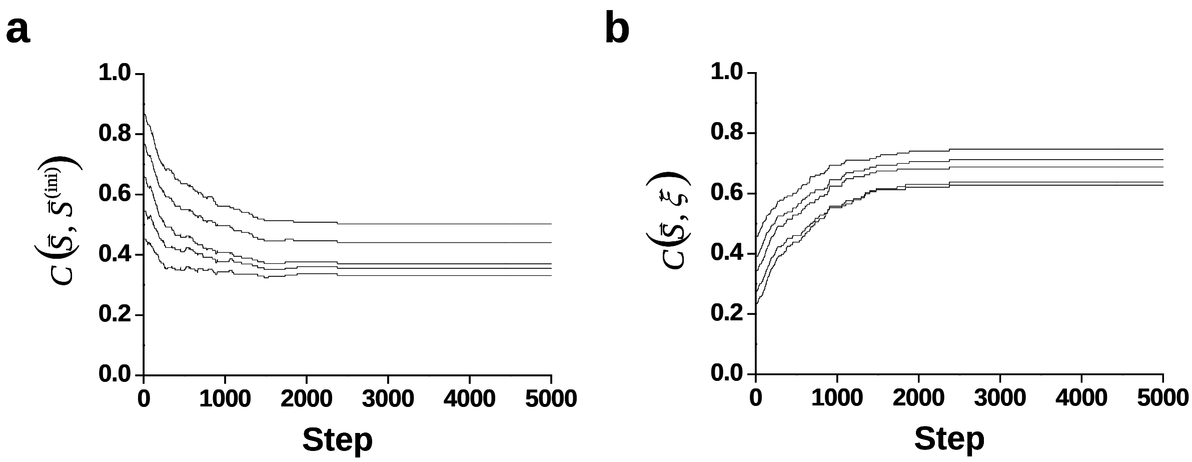

Fig. 2: a, time courses of the correlation between the network state $\vec{S}$ and the initial pattern $\vec{S}^{(\mathrm{ini})}$. b, time course of the correlation between the network state and the embedded pattern $\vec{\xi}$.

system to keyword extraction from a document. As an instance, Medline 1033, a data set consisting of 1033 medical paper abstracts (available at [11]), was used as a document corpus.

Each document in this corpus is represented by a vector in the space spanned by N terms; i.e., for the p -th document, $\vec{\xi}^{(p)} = \left(\xi_1^{(p)}, \cdots, \xi_N^{(p)}\right)$ with $\xi_i^{(p)}$ being the relative importance of term i in the p -th document.

General relation between terms was acquired by the covariance-learning rule through all the documents in the corpus; i.e., the relation between term i and term j is given by T_{ij} defined by the equation (1).

For a given document, set its vector representation as an initial state of the network-activation pattern. Then, update the network-activation pattern according to the dynamics of our model. The obtained attractor pattern will be associated with specific features of the document because it continuously depends on the initial pattern. Furthermore, it will also reflect general relation between terms stored in the network. Therefore, the attractor pattern is considered to represent appropriate keywords extracted from the document. The results of keyword extraction by our system were compared with those by TFIDF (term frequency inverse document frequency), an ordinary keyword extraction method [12].

Typical results obtained for No. 808 document (Box 1) are shown in Table 1 and Table 2, where top ten high-scored terms extracted by TFIDF (Table 1) and by our system (Table 2) are listed. Comparison between them indicates that our model can extract keywords that appropriately represent the underlying meaning of the document. The terms that are closely relevant to the meaning of document No.808 but do not appear in this document or are lower rated by TFIDF, for example, “emotion”, “ment (mental)” or “behavior”, are extracted or higher rated by our system (Table 2). The terns that are extracted by TFIDF but have little relevance to the meaning of the document, such as “history”, “today” or “term”, are removed.

It might be interesting to visualize the activity pattern ($\vec{I} = \left(I_1, \cdots, I_N\right)$, see the equation (2)) of the attractor state (Fig. 3) [13]. Terms that are reciprocally activating each other tend to exhibit higher activity.

Autistic reactions in early childhood: Differential diagnostic considerations
The term "autism" is frequently used today in the differential diagnosis of the severe emotional disturbances of early childhood. However, to label a child as "autistic" presents some formidable problems with regard to definition of the term, the specific etiological-diagnostic implications, and treatment considerations for any given child so designated. The purpose of this paper is to briefly review some of the historical psychiatric background of the term "autism", its more recent ramifications, and our clinical experiences in this field.

Box 1: Full text of No 808 in Medline1033

Term	TFIDF
autism	8.91
histor	8.43
autist	8.25
today	8.01
term	7.89
diagnost	7.68
background	7.01
child	6.74
implic	6.21
brief	5.93

TABLE 1: HIGH SCORED TERMS BY TFIDF

Term	Activity
emotion	26.87
child	25.93
autist	22.27
*ment	21.56
*behavior	21.50
*disturb	20.67
autism	20.31
*social	19.22
psych	18.40
*mutual	17.23

TABLE 2: HIGH ACTIVATED TERMS BY OUR SYSTEM. TERMS WITH ASTERISKS (*) ARE ABSENT IN THE DOCUMENT NO. 808.

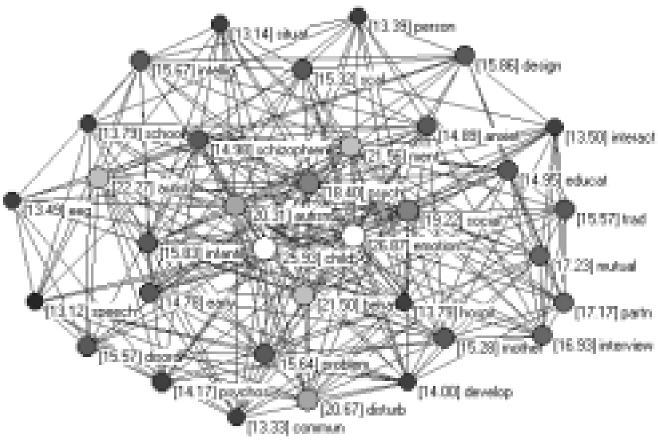

Fig. 3: The relation of terms in the attractor state: Each value in brackets indicates the activity (calculated by the equation (2)). This figure was created with Pajek, which is free software for analyzing and visualizing large networks (available at [13]).

4. Discussion

We have inferred a novel information-retrieval design based on a neural-network system with continuous attractors from neurphysiological findings of graded persistent activity. For each query encoded by the initial state of the network-activation pattern, our system can retrieve information encoded by the attractor state in a graded manner. The process might be useful for keyword extraction from a document and other document processing tasks.

Nevertheless, there is still room for improvement. In the present system, the output from the neuron i is binary ($S_i = 1$ or 0, see the equation (2)). To make it possible to deal with more detailed information, the system should be modified so that the output is represented by graded values. Our preliminary study suggests that this can be achieved by the use of supposed cellular mechanisms to account for recent experimental evidence indicating that graded persistent activity is formed at a single-cell level [14, 15].

References

[1] Hopfield, J. J., “Neural networks and physical systems with emergent collective computational abilities”, Proc. Natl. Acad. Sci. USA 79, 2554-2558 (1982).

[2] Durstewitz, D., Seamans, J. K. & Sejnowski, T. J., “Neurocomputational model of working memory”, Nature neurosci. 3 supplement, 1184-1191 (2000).

[3] Romo, R., Brody, C. D., Hernandez, A. & Lemus, L., “Neuronal correlates of parametric working memory in the prefrontal cortex”, Nature 399, 470-473 (1999).

[4] Aksay, E., Gamkrelidze, G., Seung, H. S., Baker, R. & Tank, D. W., “In vivo intracellular recording and perturbation of persistent activity in a neural integrator”, Nature neurosci. 4, 184-193 (2001).

[5] Brody, C. D., Romo, R. & Adam Kepecs, “Basic mechanisms for graded persistent activity: discrete attractors, continuous attractors, and dynamic representations”, Curr. Opin. Neurobiol. 13, 204-211 (2003).

[6] Seung, H. S., Lee, D. D., Reis, B. Y. & Tank, D. W., “Stability of the memory of eye position in a recurrent network of conductance-based model neurons”, Neuron 26, 259-271 (2000).

[7] Koulakov, A. A., Raghavachari, S., Kepecs, A. & Lisman, J. E., “Model for a robust neural integrator”, Nature neurosci. 5, 775-710 (2002).

[8] Tsuboshita, Y., Okamoto, H., “Extracting information in a graded manner from a neural-network system with continuous attractors”: in the proceedings of IJCNN2004 (2004).

[9] Okamoto, H. & Fukai, T., “Neural mechanism for a cognitive timer”, Phys. Rev. Lett. 86, 3919-3922 (2001).

[10] Amari, S., “Neural Theory of Association and Concept-Formation”, Biol, Cybernetics 26, 175-185 (1977).

[11] ftp://ftp.cs.cornell.edu/pub/smart/med/

[12] Porter, M. F., “An algorithm for suffix stripping”, Program 14, No.3, 130-137, (1980).

[13] http://vlado.fmf.uni-lj.si/pub/networks/pajek/

[14] Egorov, A. V., Hamam, B. N., Fransen, E., Hasslmo, M. E. & Alonso, A. A., “Graded persistent activity in entorhinal cortex neurons”, Nature 420, 173-178 (2002).

[15] Golman, M. S., Levine, J. H., Major, G., Tank, D. W. & Seung, H. S., “Robust persistent neural activity in a model integrator with multiple hysteretic dendrites per neuron”, Cereb. Cortex 13, 1185-1195 (2003).