Available online at www.sciencedirect.com
ScienceDirect
Neural Networks 20 (2007) 705–713
Neural Networks
www.elsevier.com/locate/neunet

Context-dependent retrieval of information by neural-network dynamics with continuous attractors

Yukihiro Tsuboshita[a,*], Hiroshi Okamoto[a,b]

[a] Corporate Research Group, Fuji Xerox Co., Ltd, 430 Sakai, Nakai-machi, Ashigarakami-gun, Kanagawa, 259-0157, Japan
[b] Laboratory for Neural Circuit Theory, RIKEN Brain Science Institute, Hirosawa 2-1, Wako, Saitama, 351-0198, Japan

Received in revised form 16 February 2007; accepted 16 February 2007

Abstract

Memory retrieval in neural networks has traditionally been described by dynamic systems with discrete attractors. However, recent neurophysiological findings of graded persistent activity suggest that memory retrieval in the brain is more likely to be described by dynamic systems with continuous attractors. To explore what sort of information processing is achieved by continuous-attractor dynamics, keyword extraction from documents by a network of bistable neurons, which gives robust continuous attractors, is examined. Given an associative network of terms, a continuous attractor led by propagation of neuronal activation in this network appears to represent keywords that express underlying meaning of a document encoded in the initial state of the network-activation pattern. A dominant hypothesis in cognitive psychology is that long-term memory is archived in the network structure, which resembles associative networks of terms. Our results suggest that keyword extraction by the neural-network dynamics with continuous attractors might symbolically represent context-dependent retrieval of short-term memory from long-term memory in the brain.

Keywords: Graded persistent activity; Continuous attractor; Bistable neuron; Information retrieval; Keyword extraction; Complex network

1. Introduction

In our daily life, ideas continually come to our mind in response to varying stimuli from external environment. Such ideas are generated in our brain where archival memory, 'long-term memory', has been largely accumulated from neonate. An idea thus generated in the brain is temporary stored until an action planned on this idea is executed. To discriminate this kind of transient memory from long-term memory, it is referred to as 'short-term memory'. Briefly, short-term memory is retrieved from long-term memory in a context-dependent manner.

It is generally considered that short-term memory of a given item is sustained by activation of an ensemble of neurons that encodes this item (Funahashi, Bruce, & Goldman-Rakic, 1989; Goldman-Rakic, 1995). Such activation can be generated as a fixed-point attractor of a bistable or multi-stable dynamic system describing a recurrent network of neurons (Amit, 1995; Amit & Brunel, 1997; Durstewitz, Seamans, & Sejnowski, 2000; Hebb, 1949; Wang, 1999, 2001). In the multi-stable neural-network system in which multiple distributed patterns are embedded, for a given pattern represented by an initial state of the network activation, one of the embedded patterns, which is the nearest to it, is retrieved (Hopfield, 1982). 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. 1(a)).

Recent neurophysiological findings of graded persistent activity, however, cast doubts on this traditional view. The firing rate of neurons recorded from the prefrontal cortex of the monkey performing a vibrotactile discrimination task varied, during the delay period between the base and comparison stimuli, as a monotonic function of the base stimulus frequency (Brody, Hernandez, Zianos, & Romo, 2003a; Romo, Brody, Hernandez, & Lemus, 1999). The firing rate of neurons in the oculomotor system of the goldfish during fixations was associated with the history of spontaneous saccadic steps (Aksay, Baker, Seung, & Tank, 2000; Aksay, Gamkrelidze, Seung, Baker, & Tank, 2001). These phenomena cannot simply

* Corresponding author. Tel.: +81 465 81 3211; fax: +81 465 81 8997.
E-mail address: Yukihiro.Tsuboshita@fujixerox.co.jp (Y. Tsuboshita).

0893-6080/$ - see front matter © 2007 Elsevier Ltd. All rights reserved.
doi:10.1016/j.neunet.2007.02.002

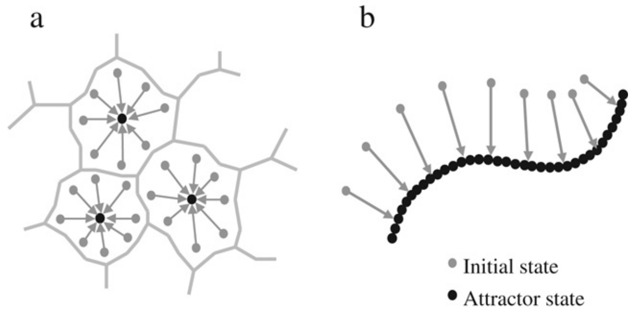

Fig. 1. (a) Discrete fixed-point attractors. (b) Continuous fixed-point attractors. Here, continuous fixed-point attractors are expressed by a line. Note that this is just a brief image to intuitively explain essential difference between continuous and discrete fixed-point attractors by a two-dimensional figure. In general, continuous fixed-point attractors in the N-dimensional phase space construct $N-1$-dimensional manifolds. Moreover, there might be more than one continuous fixed-point attractor manifold; that is, the phase space might be divided into multiple subspaces, with each having continuous fixed-point attractor manifold at its centre.

be described by multi-stable dynamic systems with discrete fixed-point attractors. They are more likely to be described by dynamic systems with continuous fixed-point attractors, in which the final state continuously depends on the initial state (Brody, Romo, & Kepecs, 2003b) (Fig. 1(b)).

The purpose of this study is to infer a novel information-processing system from supposed neural mechanisms of graded persistent activities. Several attempts have been made to build models for neural mechanisms that generate continuous fixed-point attractors (Koulakov, Raghavachari, Kepecs, & Lisman, 2002; Miller, Brody, Romo, & Wang, 2003; Seung, Lee, Reis, & Tank, 2000a, 2000b). However, their discussion has been confined to the continuity with respect to scalar quantities such as the firing rate of individual cells. We believe that rich information processing should exploit more complex properties such as the network-activation pattern represented by vector quantities. In the present study, therefore, discussion is extended to continuous fixed-point attractors with respect to the network-activation pattern. To illustrate what sort of information is retrieved in the extended system, keyword extraction from documents is examined.

2. Model architecture

2.1. Previous models of neural mechanisms for graded persistent activity

Several attempts for modelling neural mechanisms that can generate continuous fixed-point attractors have already been examined. In a recurrent network of neurons with threshold-linear input/output curves and saturating synapses, the parameter values can be tuned so that friction approximately vanishes along a certain line in the state space (Seung et al., 2000a). This line constitutes continuous fixed-point attractors (line attractors). However, continuous fixed-point attractors attained by such fine-tuning are vulnerable to slight change in parameter values. Miller et al. proposed similar mechanisms that required fine-tuning of parameters (Miller et al., 2003).

Koulakov et al. proposed mechanisms that ensured stability of continuous fixed-point attractors (Koulakov et al., 2002). They considered a recurrent network consisting of bistable neurons. Each neuron is either in the active state (‘on’ state) or in the resting state (‘off’ state). In the ‘off’ state, the firing rate is low; in the ‘on’ state, the firing rate increases with the input. In this model, the synaptic weights are essentially uniform everywhere in the network. Nevertheless, if different values of parameters defining each neuron, such as firing threshold or membrane conductance, are properly chosen, neurons 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 activation appearing in this series is stable. If the number of neurons constituting the network is sufficiently large, these stable states are approximately continuous. The firing rate of each neuron continuously depends on the number of neurons in the ‘on’ state.

In the above mechanisms, discussion is focused on the continuity of fixed-point attractors with respect to scalar quantities such as the firing rate of individual cells. This is because their main purpose is to account for present experimental observation. In usual electrophysiological recordings, the firing rate of single cells, which is represented by scalar quantities, is examined. However, rich information processing should exploit more complex properties such as those represented by vector quantities. Although not yet observed by present experimental techniques, the continuity with respect to the activation pattern of an ensemble of neurons, which will be represented by vector quantities, might underlie graded persistent activity.

2.2. Recurrent network of bistable neurons with non-uniform synaptic strengths

Consider a network of N neurons. We assume that each neuron, say neuron n for instance, has two stable states, the resting and active states, if the external input I_n is in a certain range, say $I_\theta - W/2 < I_n < I_\theta + W/2$. Let S_n be a binary variable taking 0 or 1: If neuron n is in the resting state, $S_n = 0$; if it is in the active state, $S_n = 1$. Moreover, let r_n be the firing rate of neuron n. If neuron n is in the resting state, r_n takes a low constant value r_L, and if it is in the active state, r_n is a positive linear function of I_n (Tuckwell, 1988):

$$r_n = r_L \quad \text{if } S_n = 0 \tag{1}$$

$$r_n = AI_n + B \quad \text{if } S_n = 1 \tag{2}$$

with $A > 0$ and $A(I_\theta - W/2) + B > r_L$.

Although it still remains elusive whether real neurons are bistable, it has theoretically been demonstrated that neuronal bistability is useful for stabilizing cue-position specific delay-period activity in the rotationally invariant system (Camperi & Wang, 1998) or generating psychological timing (Okamoto & Fukai, 2001). Some kind of positive feedback current, such as NMDA-receptor mediated current with sigmoidal voltage-dependent gating (Koulakov et al., 2002; Lisman, Fellous, & Wang, 1998) or spike-induced Ca^{2+}-activated after-depolarization current (Fransen, Alonso, & Hasselmo, 2002;