2009 International Symposium on Nonlinear Theory and its Applications
NOLTA'09, Sapporo, Japan, October 18-21, 2009

A Meso-Scopic Model of the Binding Formation in the Brain

Makito Oku† and Kazuyuki Aihara†,‡

†Department of Mathematical Informatics, Graduate School of Information Science and Technology, The University of Tokyo
7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8656, Japan
‡Institute of Industrial Science, The University of Tokyo
4-6-1 Komaba, Meguro-ku, Tokyo 153-8505, Japan
Email: oku@sat.t.u-tokyo.ac.jp

Abstract—

In a general sense, binding is dynamical coupling between distant regions in the brain. Although its detailed mechanism has not been well understood, it is suggested that binding can be perceived from the dynamical system's viewpoint. In order to examine the binding mechanism from this standpoint, here we propose a meso-scopic model that describes each region's dynamics abstractly. The model is characterized by the following two points: (1) Spatio-temporal patterns of regional activities are symbolically represented by simple limit-cycle orbits, and (2) binding is realized as a form of mutual excitation among these orbits. Preliminary numerical simulations show that various patterns of binding can be generated. Furthermore, attractor switching is also observed according to conditions. This study supports the idea that the dynamical system's point of view is effective to understand binding mechanism as well as other brain functions.

1. Introduction

In the neocortex, each region is known to be involved in different functional roles. Therefore, when distant regions need to exchange information, dynamical coupling between them have to be formed depending on the situation. This is called binding. The mechanism of binding is one of the important issues in neuroscience, which is widely known as the binding problem[1].

To understand binding, it is suggested that the dynamical system's point of view is important [2, 3](see Fig. 1). In this viewpoint, binding is regarded as a kind of attractor. Particularly, to distinguish it from attractors localized within a small region, sometimes such a cross-region attractor is also called a global attractor[2, 3]. Although this viewpoint gives us a qualitative and abstract explanation for binding, the direct correspondence to detailed biological phenomena is difficult to argue.

In this study we propose a mathematical model to explain binding from the dynamical system's viewpoint. Generally speaking, there are various levels of mathematical models of the brain. Commonly used are neuron-level models such as Hodgkin-Huxley model, leaky integrate-and-fire model, Izhikevich model, and so on. More detailed models are also used such as multi-compartment models or models those consider even the channels' distribution. On the other hand, other models put together a group of neurons and describe the abstract dynamics of the whole system[4, 5]. In distinction from former models, we call the latter meso-scopic models in this paper.

Meso-scopic models are suitable to show the relationship between binding and the concept of attractors. First, binding is essentially meso-scale phenomena, and it is explained qualitatively and abstractly by the concept of attractors. Additionally, since much about the detailed biological mechanism of binding is not known, micro-scopic models may require numbers of additional assumptions. For these reasons, we decide to use a meso-scopic model.

2. Model

Let N denotes the number of regions under consideration. In the following sentences, we call them nodes. Each

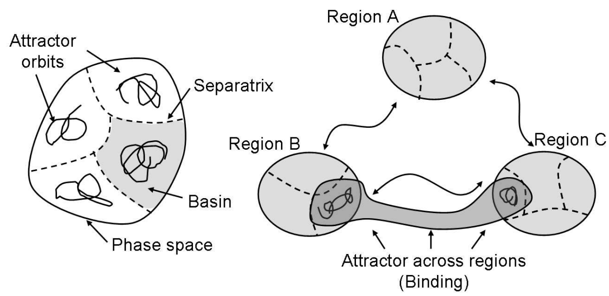

Figure 1: Meso-scale dynamics of the brain from the dynamical system's viewpoint. (left) A small region in the brain that contains thousands or millions of neurons is under consideration. If this region is seen as a dynamical system, the spatio-temporal patterns of the activities of neurons within the region, i.e., oscillation modes or sustained firing patterns, correspond to different attractor orbits. Notice that this figure is about the phase space, not the physical space. (right) It is said that binding can be interpreted as an attractor that involves multiple regions[2, 3]. Inversely, when the interaction between regions are sufficiently weak so that they behave as if they are isolated systems, the regions are unbinded.

node has a M dimensional state $\boldsymbol{x}_i \in \mathrm{R}^M$ $(i = 1, \ldots, N)$. The whole system's state is denoted by $X = \{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_N\}$, and $X_i = X \setminus \{\boldsymbol{x}_i\}$. Then, the dynamics of each node is described as follows:

$$\frac{\mathrm{d}\boldsymbol{x}_i}{\mathrm{d}t} = \boldsymbol{f}_i(\boldsymbol{x}_i; X_i) + D\boldsymbol{\eta}_i, \quad (1)$$

where $\boldsymbol{f}_i : \mathrm{R}^M \to \mathrm{R}^M$ is the flow in the phase space, which is time-varying due to the influence of other nodes X_i. Next, $\boldsymbol{\eta}_i \in \mathrm{R}^M$ is the noise term. For simplicity, uncorrelated homogeneous white noise is assumed. The parameter $D \in \mathrm{R}$ determines the noise level.

Every node has K attractor orbits $C_i^1, \ldots, C_i^K$ $(i = 1, \ldots, N)$. Assume that no two orbits are intersected each other. To extend the previous study where all attractors are equilibrium point attractors[5], in this study simple limit-cycle orbits are used.

First, let us consider the flow below. This has a stable limit-cycle attractor in the $M = 3$ dimensional phase space:

$$\boldsymbol{f}(\boldsymbol{x}) = \begin{pmatrix} -x_2 + x_1(1 - x_1^2 - x_2^2) \\ x_1 + x_2(1 - x_1^2 - x_2^2) \\ -x_3 \end{pmatrix}. \quad (2)$$

The square of the distance from a point $\boldsymbol{x} = \{x_1, x_2, x_3\}$ to the limit-cycle orbit $C = \{\{x_1, x_2, x_3\} \mid x_1^2 + x_2^2 = 1, x_3 = 0\}$ is,

$$a(\boldsymbol{x}, C) = (1 - \sqrt{x_1^2 + x_2^2})^2 + x_3^2. \quad (3)$$

Next, by randomly Affine transform the flow $\boldsymbol{f}$, we make K flows for each node $\boldsymbol{f}_i^1, \ldots, \boldsymbol{f}_i^K$ $(i = 1, \ldots, N)$ as follows:

$$\boldsymbol{f}_i^k(\boldsymbol{x}_i) = (A_i^k)^{\mathrm{T}} \boldsymbol{f}(A_i^k \boldsymbol{x}_i + \boldsymbol{b}_i^k), \quad (4)$$
$$a(\boldsymbol{x}_i, C_i^k) = a(A_i^k \boldsymbol{x}_i + \boldsymbol{b}_i^k, C), \quad (5)$$

where $\boldsymbol{f}_i^k : \mathrm{R}^3 \to \mathrm{R}^3$ is the flow that correspond to the orbit C_i^k. Next, $A_i^k \in \mathrm{R}^{3\times 3}$ and $\boldsymbol{b}_i^k \in \mathrm{R}^3$ are matrix and vector for the Affine transformation, respectively. For simplicity, we assume that A_i^k has only rotational component, and there is no scaling. Therefore, $\det A_i^k = 1$ and $(A_i^k)^{-1} = (A_i^k)^{\mathrm{T}}$.

The output of each node $\boldsymbol{y}_i = \{y_i^1, \ldots, y_i^K\}$ $(i = 1, \ldots, N)$ is described by softmax function of the values of $a(\boldsymbol{x}_i, C_i^k)$:

$$y_i^k = \frac{\exp(-\beta(a(\boldsymbol{x}_i, C_i^k)))}{\sum_{l=1}^K \exp(-\beta(a(\boldsymbol{x}_i, C_i^l)))}, \quad (6)$$

where $\beta \in \mathrm{R}$ is a parameter that determines the steepness of softmax function. Notice that the output $\boldsymbol{y}_i$ is a natural extension of that of single neuron. As if a single neuron discriminates two classes, a node in this model discriminates K classes.

The inputs from other nodes are weighted and summed up, then influence the flow of each node $\boldsymbol{f}_i$ as follows:

$$\boldsymbol{f}_i(\boldsymbol{x}_i; X_i) = \sum_{j=1}^N \sum_{k=1}^K \sum_{l=1}^K w_{ijkl} y_j^l \boldsymbol{f}_i^k(\boldsymbol{x}_i), \quad (7)$$

where weight $w_{ijkl} \in \mathrm{R}$ means a degree how $\boldsymbol{f}_i^k$ is enforced when node j's state $\boldsymbol{x}_j$ is close to the orbit C_j^l. Therefore, Eq. (1) can be regarded as a continuous relaxation of a dynamical system with switching. In other words, dominant dynamics switches continuously depending on the input from the other nodes. In fact, in the limitation of $\beta \to \infty$, the dynamics of each node switches discontinuously.

3. Simulations and Results

In the following test simulations we study the case of $N = 3$. For each node $K = 5$ limit cycles are randomly located (see Fig. 2). Then, we assign correspondence relations between orbits, which are represented by colored arrows in Fig. 2. According to the relations, we determine the value of weights w_{ijkl} as follows. If a pair of orbits C_j^l and C_i^k has a directional relation from C_j^l to C_i^k, weight w_{ijkl} is set to 1. If not, it is set to 0. Level means the number of nodes necessary to sustain the attractor. Therefore, level $L \geq 2$ attractors correspond to binding, and level 1 attractors are local regional ones. Each node has one level-1 orbit, two level-2 orbits, and two level-3 orbits.

At first, we investigate the simplest case where $D = 0$ and β is sufficiently large. A typical model's behavior in this case is shown in Fig. 3. Each node's state stays mov-

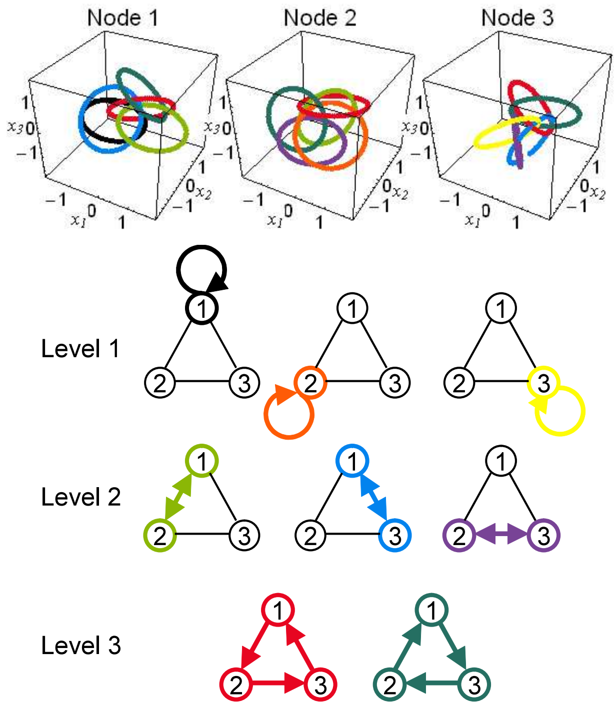

Figure 2: (top) An example of limit-cycle orbits randomly located in the $M = 3$ dimensional spaces. (bottom) Correspondence relations between orbits. The numbers indicate node indexes. Colored circles denote limit-cycle orbits, and arrows indicate the relations between orbits (see the text).

ing on a particular orbit. The value of output $\boldsymbol{y}_i$ is nearly constant, but it changes sharply where the state approaches other orbits. When the initial condition is changed, various patterns of attractors are seen. These include three level-1 attractors, combination of level-1 and level-2 ones, and a level-3 one.

Next, under different parameter settings, attractor switching is also observed as shown in Fig. 4. Each node's state moves on multiple orbits alternately. Switching is likely to occur where orbits get close. Comparing different levels of attractors, level-3 attractors appear with relatively high frequency in a wide rage of parameters, followed by the combination of level-1 and level-2 attractors. The pattern of three level-1 attractors is far less frequent.

Next, we investigate the effect of the two parameters D and β independently. If D increases, the system's behavior becomes similar to the random walk as shown in Fig. 5. On the other hand, if β decreases, the trajectory of solution gets complexly distorted. However, the maximal Lyapunov exponent in this case is not positive.

To evaluate the effect of the parameters more systematically, we consider the instantaneous entropy $H[\boldsymbol{y}_i] = -\sum_{k=1}^{K} y_i^k \ln y_i^k$ and estimate its time average $< H[\boldsymbol{y}_i] >$ and variance $< (H[\boldsymbol{y}_i] - < H[\boldsymbol{y}_i] >)^2 >$ (see Fig. 6). In all cases the analyzed statistics show single peaks. Furthermore, the

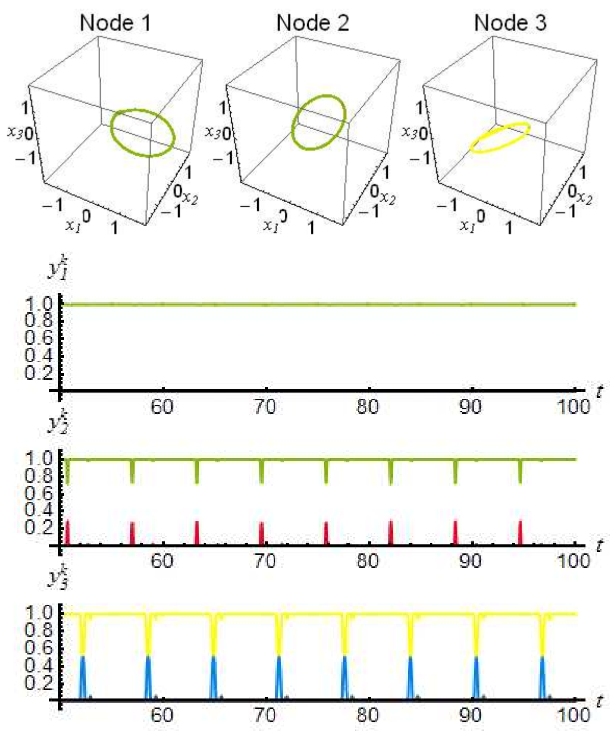

Figure 3: An example case where binding is stably maintained. Node 1 and 2 are trapped in a level-2 attractor (light green), and the remaining node 3 is in a level-1 attractor (yellow). The simulation condition is $D = 0$, $\beta = 200$.

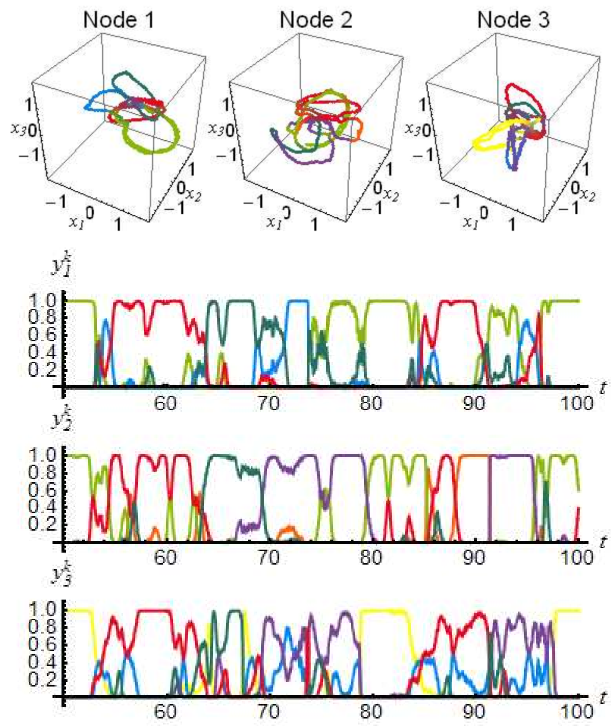

Figure 4: An example case where attractor switching occurs. Multiple attractor orbits appear alternately. For example, at $t = 50$ light green and yellow orbits are dominant. They are followed by red one at $t = 55$ and then green one at $t = 65$, and so on. The simulation condition is $D = 0.05$, $\beta = 20$.

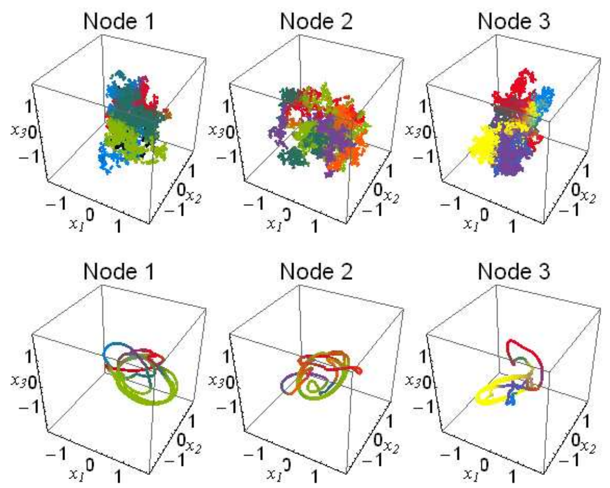

Figure 5: Effect of the parameters to the dynamics. (top) A case of large noise ($D = 0.5$, $\beta = 20$). (bottom) A case of small steepness of softmax function ($D = 0$, $\beta = 10$). Transient behavior is cut-off in both cases.

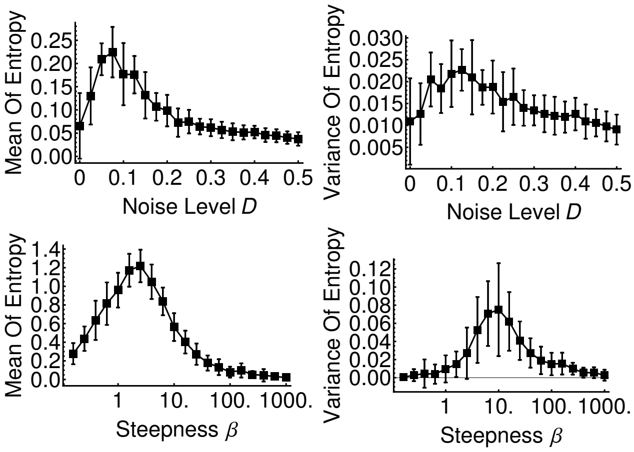

Figure 6: Plots of the statistics of instantaneous entropy against parameter change. Error bars are the standard deviations of 10 trials. In each trial orbits are randomly relocated, and the beginning 1000 steps are omitted in the analysis. In the top row β is fixed to 200, and in the bottom row D is fixed to 0.

the positions of the peaks are different between the mean and the variance.

4. Discussions

First, attractor switching as shown in Fig. 4 may correspond to the ongoing activity of the brain. In general, the brain's state changes continuously even when there is no external input[6]. To interpret this from the dynamical system's viewpoint, the spontaneous activity of the brain is itinerancy among multiple meta-stable attractors[2, 3]. Our model is consistent with this idea.

On the other hand, the situation such as Fig. 3 may correspond to that where selective attention works. In general, selective attention is said to play an important role to form a specific binding. To see this from the dynamical system's viewpoint, some sort of parameters are changed by attention, and the system's dynamics turns to an ordered phase where attractors are stably sustained[2, 3]. Our model is also consistent with this idea that associates attention with the concepts of bifurcation.

Similarly, the qualitative change of the system's dynamics in the real brain is sometimes argued in relation to phase transition and self-organized criticality (SOC)[7]. In this viewpoint, the brain's dynamics is predicted to be maximally complex in the middle between totally ordered phase and totally random phase. Our model is not fully consistent with this idea because no scale-free property, which is a relevant feature of SOC, is observed. However, the characteristic peaks shown in Fig. 6 are some sort of supportive evidences. Anyway, further discussion about this problem may require more detailed studies in the future.

5. Conclusion

In this study, we have proposed a meso-scopic model that describes regional dynamics of the brain abstractly and examined the binding mechanism from the dynamical system's viewpoint. This study shows a possibility that binding can be interpreted as a special attractor that involves multiple regions. This suggests that both meso-scopic modelling and the dynamical system's viewpoint are important to understand the function of the brain.

The next step of this research is to clarify the relationship between binding and other important issues related to it. For example, binding is suggested to have tight relations with selective attention, gamma oscillations, and neuromodulators such as acetylcholine[3]. Furthermore, some psychiatric disorders can be related with some disorder of binding mechanism. It is difficult but important to explain these issues from the mathematical viewpoint.

Acknowledgments

The authors would like to thank Prof. H. Fujii at Kyoto Sangyo University for discussions and valuable suggestions over the years. This research is partially supported by the Japan Society for the Promotion of Science, a Grant-in-Aid for JSPS Fellows (21·937).

References

[1] A. Treisman, "Solutions to the binding problem: progress through controversy and convergence.," *Neuron*, vol. 24, pp. 105–110, 111–125, 1999.

[2] H. Fujii, K. Aihara, and I. Tsuda, "Functional relevance of 'excitatory' GABA actions in cortical interneurons: A dynamical systems approach.," *J Integr Neurosci*, vol. 3, pp. 183–205, 2004.

[3] H. Fujii, K. Aihara, and I. Tsuda, "Corticopetal acetylcholine: Possible scenarios on the role for dynamic organization of quasi-attractors," *Lecture Notes In Computer Science*, vol. 4984, pp. 170–178, 2008.

[4] C. M. Bachmann, L. N. Cooper, A. Dembo, and O. Zeitouni, "A relaxation model for memory with high storage density.," *Proc Natl Acad Sci U S A*, vol. 84, pp. 7529–7531, 1987.

[5] M. Oku and K. Aihara, "A mathematical model of planning in the prefrontal cortex," *Artificial Life and Robotics*, vol. 12, no. 1, pp. 227–231, 2008.

[6] T. Kenet, D. Bibitchkov, M. Tsodyks, A. Grinvald, and A. Arieli, "Spontaneously emerging cortical representations of visual attributes.," *Nature*, vol. 425, pp. 954–956, 2003.

[7] G. Buzsáki, *Rhythms of the Brain*. Oxford University Press, USA, 2006.