Neural Networks 62 (2015) 3–10

Contents lists available at ScienceDirect

Neural Networks

journal homepage: www.elsevier.com/locate/neunet

2015 Special Issue

Mathematical modeling for evolution of heterogeneous modules in the brain

Yutaka Yamaguti *, Ichiro Tsuda

Research Institute for Electronic Science, Hokkaido University, Kita-12, Nishi-7, Kita-ku, Sapporo, Hokkaido 060-0812, Japan
Research Center for Integrative Mathematics, Hokkaido University, Kita-12, Nishi-7, Kita-ku, Sapporo, Hokkaido 060-0812, Japan

ARTICLE INFO

Article history:
Available online 1 August 2014

Keywords:
Modularity
Coupled-oscillator
Information flow
Transfer entropy
Evolution
Symmetry breaking

ABSTRACT

Modular architecture has been found in most cortical areas of mammalian brains, but little is known about its evolutionary origin. It has been proposed by several researchers that maximizing information transmission among subsystems can be used as a principle for understanding the development of complex brain networks. In this paper, we study how heterogeneous modules develop in coupled-map networks via a genetic algorithm, where selection is based on maximizing bidirectional information transmission. Two functionally differentiated modules evolved from two homogeneous systems with random couplings, which are associated with symmetry breaking of intrasystem and intersystem couplings. By exploring the parameter space of the network around the optimal parameter values, it was found that the optimum network exists near transition points, at which the incoherent state loses its stability and an extremely slow oscillatory motion emerges.

1. Introduction

Modular architecture is one of the most frequently referenced concepts of neural organization (Felleman & Van Essen, 1991; Mountcastle, 1997; Szentágothai, 1983). In biological systems, such modular architecture has been self-organized in ontogeny, influenced by phylogeny (Wagner, Pavlicev, & Cheverud, 2007). In contrast, the modules and their interactions of artificial machines are designed by humans. Understanding the mechanism giving rise to functionally differentiated modules in biological systems, such as cortical module architectures, is of great interest.

In theoretical studies, it has been proposed that maximizing information transmission between subsystems can be used as a guiding principle for understanding the development and evolution of complex brain networks. Linsker (1988, 1989, 1997) shows that information transmission between successive layers of feed-forward networks is a viable principle for designing functional neural networks. Recently, Tanaka, Kaneko, and Aoyagi (2009) showed that the learning algorithm that maximizes information retention in recurrent networks also gives rise to the appearance of biological structures, such as cell assemblies, and even to dynamics, such as spontaneous activity of synfire chain and critical neuronal avalanches. One of the present authors has proposed the idea of a new self-organization principle based on a variational principle (Kaneko & Tsuda, 2001; Tsuda, 1984, 2001): components (or elements) are self-organized according to constraints that act on the whole system, whereas according to the usual self-organization principles (ex. Haken, 1980; Nicolis & Prigogine, 1977), macroscopic order is self-organized via interactions among microscopic elements. Actually, neuron-like elements as components of a system have been obtained under the condition of maximum transmission of information across a whole system (Ito & Tsuda, 2007; Watanabe, Ito, & Tsuda, in preparation). In the brain, bidirectional information transmission among modules is crucial for information processing. In fact, most connections between cortical modules are known to be bidirectional (Felleman & Van Essen, 1991). Yamaguti, Tsuda, and Takahashi (2014) have found intermittent switching of direction of information flow in two heterogeneously coupled chaotic systems. Such a mechanism may be applicable to bidirectional information transmission in neural systems.

Motivated by these studies, here we propose a mathematical model for functional differentiation induced by selection based on maximizing bidirectional information transmission. We try to extract the essence of evolutionary dynamics by investigating a coupled-oscillator network model. We construct randomly coupled-oscillator networks, each consisting of two sub-networks, by using a discretized version of the Kuramoto phase oscillator

---

[*] Corresponding author at: Research Institute for Electronic Science, Hokkaido University, Kita-12, Nishi-7, Kita-ku, Sapporo, Hokkaido 060-0812, Japan. Tel.: +81 11 706 2593; fax: +81 706 2413.
E-mail address: yyama@es.hokudai.ac.jp (Y. Yamaguti).

http://dx.doi.org/10.1016/j.neunet.2014.07.013
0893-6080/© 2014 Elsevier Ltd. All rights reserved.

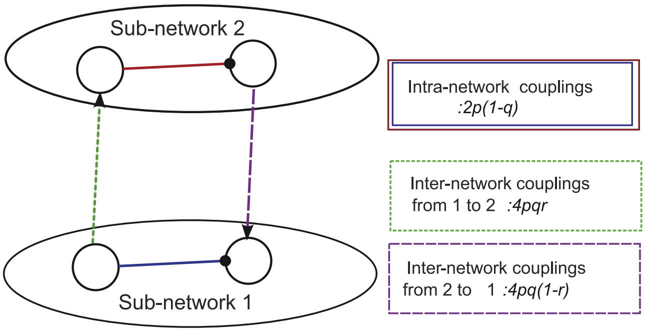

Fig. 1. Schematic representation of the model network and the parameters that regulate coupling probabilities.

model (Kuramoto, 1984). Each sub-network consists of N oscillators, whose states are represented by phase variables. A genetic algorithm is used to modify the parameters to the direction of better fitness. Transfer entropy (Schreiber, 2000), which measures directed information transfer, is estimated to quantify information flow between the two sub-networks in both directions and their product is regarded as fitness in the evolutionary process.

In Section 2, the proposed network model, analytic method, and procedure for evolution by the genetic algorithm are described. In Section 3, we give numerical results of this evolutionary process. In Section 4, we study the fitness landscape around the optimum network, exploring the parameter space as a means of characterizing the evolutionary process from the viewpoint of dynamical systems. In Section 5, dependences of the coupling density on heterogeneity are investigated. Section 6 is devoted to summary and discussions.

2. Models and methods

2.1. Network model

We consider a network of oscillators, which consists of two sub-network systems. Each sub-network consists of $N = 200$ phase oscillators, whose dynamics is described by a discrete-time version of the Kuramoto model (Barlev, Girvan, & Ott, 2010; Daido, 1986; Kuramoto, 1984). It is known that the dynamics of weakly coupled oscillators that have stable limit-cycles can be reduced to a coupled equation of phase oscillators by using a reduction technique (Kuramoto, 1984). The Kuramoto model corresponds to the simplest case in which the coupling term can be represented by the lowest order term of its Fourier series. Therefore, it is reasonable to expect that the results presented here will be robust, even for other coupled-oscillator systems.

The numbers of couplings between oscillators in the networks are provided by three parameters, p, q, and r. Roughly, p is the mean probability for a coupling between two oscillators in a whole network, q is the proportion of internetwork couplings to whole couplings, and r is the proportion of internetwork couplings from system 1 to 2 to whole internetwork couplings (see Fig. 1). The coupling probability from an oscillator to another oscillator in the same sub-network (intranetwork coupling) is given by $2p(1-q)$, and the coupling probability from an oscillator in sub-network 1 (resp. 2) to an oscillator in sub-network 2 (resp. 1) (internetwork coupling) is given by $4pqr$ ($4pq(1-r)$). The existence of coupling between any two oscillators is independent of other couplings.

The dynamics of the k-th oscillator in sub-network i ($i = 1, 2$, $k = 1, \ldots, N$) is described by

$$\theta_{t+1}^{(i,k)} = \theta_t^{(i,k)} + \omega + \frac{\alpha}{2Np} \sum_{(j,l)\in G^{(i,k)}} \sin\left(\theta_t^{(j,l)} - \theta_t^{(i,k)} - \psi_{kl}^{ij}\right) + \beta_t^{(i,k)}, \tag{1}$$

where ω is a natural frequency, α is a coupling strength, $\beta_t^{(i,k)}$ denotes the Gaussian noise term independently imposed on each oscillator, and $G^{(i,k)}$ denotes a set of labels (j, l) for oscillators connecting to oscillator (i, k). In numerical simulations, ω and α are fixed to 1 and 0.1, respectively, and the standard deviation of $\beta_t^{(i,k)}$ is set to 0.05.

For each ψ_{kl}^{ij}, one of two possible values, $(m-1)\pi$ with $m = 1$ or 2, is randomly assigned according to given probabilities $p_m^{(ij)}$, which satisfy the probability conditions $p_m^{(ij)} \geq 0$ and $p_1^{(ij)} + p_2^{(ij)} = 1$. In the case of a two-oscillator system, these correspond to phase locking, with 0 as "in-phase" and π as "anti-phase".

To characterize the macroscopic states of the sub-networks, we define a complex order parameter (Kuramoto, 1984; Landau & Ginzburg, 1950) $R_t^{(i)}$ for each sub-network as follows:

$$R_t^{(i)} = \frac{1}{N} \sum_{k=1}^{N} \exp(\sqrt{-1}\theta_t^{(i,k)}).$$

Its magnitude $|R_t^{(i)}|$ can be used as a measure of phase coherence within the sub-network, and its angle $\Theta_t^{(i)} = \arg R_t^{(i)}$ represents a mean phase for each sub-network. The phase differences given by $\Phi_t = \Theta_t^{(2)} - \Theta_t^{(1)}$ may be considered to be another order parameter which conveys information on phase modulations.

2.2. Information flow as an index of optimization

We used transfer entropy (TE), an information-theoretic measure, to quantify directed information transfer from one sub-network to another (Kaiser & Schreiber, 2002; Paluš & Vejmelka, 2007; Schreiber, 2000; Vicente, Wibral, Lindner, & Pipa, 2011; Yamaguti et al., 2014). Let $X = \{X_t\}$ and $Y = \{Y_t\}$, $(t = 0, 1, \ldots)$ be two discrete random processes with states x_t and y_t selected from a countable alphabet A. Let $x_t^{(m)} = (x_t, x_{t-\tau}, \ldots, x_{t-(m-1)\tau})$ and $y_t^{(l)} = (y_t, y_{t-\tau}, \ldots, y_{t-(l-1)\tau})$ be the delay-coordinate vectors with time-step τ. We write probabilities and conditional probabilities as $p(\cdot)$ and $p(\cdot|\cdot)$, respectively. If the future state of Y, denoted by $y_{t+\tau}$, depends on $y_t^{(l)}$ but not on $x_t^{(m)}$, then the generalized Markov property

$$p(y_{t+\tau}|y_t^{(l)}, x_t^{(m)}) = p(y_{t+\tau}|y_t^{(l)}), \tag{2}$$

holds. If there is any dependence on $x_t^{(m)}$, it can be quantified by the transfer entropy (TE), which is defined as the Kullback–Leibler divergence between the two probability distributions on each side of Eq. (2):

$$T_{X\to Y}(\tau) \equiv \sum_{y_{t+\tau}, y_t^{(l)}, x_t^{(m)}} p(y_{t+\tau}, y_t^{(l)}, x_t^{(m)}) \times \log\left(\frac{p(y_{t+\tau}|y_t^{(l)}, x_t^{(m)})}{p(y_{t+\tau}|y_t^{(l)})}\right), \tag{3}$$

where the sum runs over the all possible orderings of alphabets. In this way, TE quantifies a directional dependence between two time series. TE can also be represented by using conditional mutual information (Matsumoto & Tsuda, 1988; Paluš & Vejmelka, 2007):

$$\begin{aligned} T_{X\to Y}(\tau) &= I\left(y_{t+\tau}; x_t^{(m)}|y_t^{(l)}\right) \\ &\equiv \sum_{y_t^{(l)}} p(y_t^{(l)}) \sum_{y_{t+\tau}, x_t^{(m)}} p(y_{t+\tau}, x_t^{(m)}|y_t^{(l)}) \\ &\quad \times \log \frac{p(y_{t+\tau}, x_t^{(m)}|y_t^{(l)})}{p(y_{t+\tau}|y_t^{(l)})p(x_t^{(m)}|y_t^{(l)})}. \end{aligned}$$

When the two processes cannot be assumed to be Markov processes, TE is regarded as just an approximate measure of information transfer. Hereafter we treat the case $l = m = 1$.