

Article

Not peer-reviewed version

Emergence of Memory and Program via Functional Differentiation in Evolutionary Echo State Networks: Toward an Understanding of the Origin of Consciousness by Optimization

[Hiroshi Watanabe](#) * and [Ichiro Tsuda](#) *

Posted Date: 5 December 2025

doi: 10.20944/preprints202512.0432.v1

Keywords: self-organization of memory and program; evolutionary echo state network; Kolmogorov complexity; functional differentiation; self-organization with constraints; algorithmic consciousness; phase transitions



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a [Creative Commons CC BY 4.0 license](#), which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

Article

Emergence of Memory and Program via Functional Differentiation in Evolutionary Echo State Networks: Toward an Understanding of the Origin of Consciousness by Optimization

Hiroshi Watanabe ^{1,*} and Ichiro Tsuda ^{2,*}

¹ School of Design, Sapporo City University

² AIT Center, Sapporo City University

* Correspondence: hi.watanabe@scu.ac.jp (H.W.); i.tsuda@scu.ac.jp (I.T.); Tel.: +81-11-592-2300 (H.W.); +81-11-522-6186 (I.T.)

Abstract

In the context of “Kolmogorov consciousness”, the complexity of an object to be computed can be defined by the complexity of a computing system—typically a computer—represented as the number of bits in the minimal algorithm required to compute the object. In condensed matter systems under equilibrium and nonequilibrium conditions, macroscopic properties distinct from elementary ones can emerge from large numbers of particles. Such a large system size allows system properties to change as control parameters change, producing phase transitions. Inspired by this, it is natural to consider that a similar state transition could occur in computing systems at some critical point of Kolmogorov complexity when the system is optimized. In this paper, we propose a computational model that realizes a functional differentiation of random neural network into two subnetworks—one specialized for memory and the other for program execution—through an evolutionary change of the network. The model suggests a neural mechanism for a human-like thought process, described as inference based on extracting rules embedded in random input and storing those inputs in long- or short- term memory. Because the proposed evolutionary model with sufficiently large complexity is driven by constraints that lead to an optimized state—such as selecting minimal-complexity systems among many high-complexity alternatives—the resulting network can be regarded as an optimized “descriptor” of memory and program. The computational results indicate that the seeds of consciousness emerge when the system’s Kolmogorov complexity decreases from that of random neural networks to that of organized networks configured as descriptors of fluctuating environments. We refer to this stage as **algorithmic consciousness**.

Keywords: self-organization of memory and program; evolutionary echo state network; Kolmogorov complexity; functional differentiation; self-organization with constraints; algorithmic consciousness; phase transitions

1. Introduction

The human brain consists of multiple functional modules, each composed of various types of neurons that carry out a single or multiple functions. The hierarchical architecture of these structures is organized both biological evolution and ontogenetic development, such as those represented in Brodman’s functional map [1]. These structures reflect typical processes of **functional differentiation**. Moreover, recent studies have revealed that subareas within a functional region dynamically cooperate depending on specific tasks—a phenomenon called a **functional parcellation** [2]. Understanding the neural mechanisms underlying these differentiations is a central challenge in cognitive science and neuroscience, as it may provide insight into the neural basis of consciousness.

Nevertheless, research in this direction has not sufficiently progressed. We have previously addressed this issue within the framework of **self-organization with constraints**, which can be formulated via a variational principle [3]. Using this framework, we have proposed mathematical models that demonstrate the emergence of functional units, such as neurons and neural modules [4–6].

In this context, we here propose a new model of functional differentiation that enables the simultaneous emergence of “memory” and “program”, albeit within a specific and limited setting. The model also serves as a mathematical proposal for consciousness grounded in Kolmogorov complexity [7,8]. Kolmogorov complexity measures the minimal number of bits required to describe or compute an object. For example, a period-two periodic sequence consisting of 0 or 1, {010101...} of length $2n$ can be represented as “Repeat the sequence ‘01’ n times” with the Kolmogorov complexity on the order of $\log n$. In contrast, a random binary sequence of the same length has the complexity on the order of n , since it cannot be compressed. Accordingly, randomness corresponds to maximal complexity, including fully developed chaos. Crutchfield’s ϵ -machine theory formalizes these ideas by classifying temporal patterns in terms of computational structure [9]. We extend this viewpoint by considering observers—biological or computational—as information processing systems. A correct observation requires maximal information gain under minimal energetic cost in generating internal descriptions, where energy consumption is assumed proportional to the Kolmogorov complexity of those descriptions.

Historically, consciousness has often been associated with inference. For instance, Goerge Boole regarded consciousness as arising from inferential sequences represented by binary symbols [10]. Charles Peirce likewise characterized consciousness as an inferential process [11]. These pioneering ideas establish an early link between consciousness, logic, and Kolmogorov complexity.

We assume that memory and program are essential components of conscious activity. The model we propose generates both memory and program structures through an optimization process that naturally defines Kolmogorov complexity. For this reason, the model may illustrate fundamental mechanisms for generating conscious-like processes.

In Section 2, we introduce an adaptable neural network model that extends the echo state network (ESN) [12–14] through evolutionary modification of internal connections using a genetic algorithm, yielding what we call the **evolutionary ESN (eESN)** [4,15,16]. The system is trained on two distinct tasks—memory and program execution—, using a random input time series that represents a highly complex environment. In section 3, we present computation results that illustrate the emergence of the optimized internal structures to yield the memory and program units. Section 4 is devoted to discussion.

2. Materials and Methods

2.1. Model and Network Structure

We consider an eESN consisting of a recurrent network with N nodes. Let $\mathbf{x}(t) = (\mathbf{x}_1(t), \mathbf{x}_2(t), \dots, \mathbf{x}_N(t))^T \in \mathbb{R}^N$ denote the state vector of network at discrete time $t \in \mathbb{N}$. Here, a superscript T indicates a transpose. The network receives a scalar random input $u(t)$, drawn from a uniform distribution over $[0, 1]$, representing a maximal Kolmogorov complexity signal. The network dynamics is specified by

$$x(t+1) = (1 - \alpha)x(t) + \alpha\psi(x(t); u(t)), \quad (1)$$

where $\psi(\mathbf{x}(t); u(t))$ is a sigmoid function described by

$$\widehat{\tanh}(W\mathbf{x}(t) + W_{in}u(t)), \quad (2)$$

where $W \in \mathbb{R}^{N \times N}$ denotes a $N \times N$ matrix of internal connections in the recurrent neural network and $W_{in} \in \mathbb{R}^N$ denotes a vector whose elements represent the connections from input to internal neuronal units. The symbol $\widehat{\tanh}$ returns the respective value of \tanh for each element. Each element of the internal matrix W takes one of the three values ρ , $-\rho$, and 0 . This discrete structure allows

for faster convergence and clearer interpretation of structural changes during evolution than those of continuous one. The internal network has no self-connections, that is, the diagonal elements of \mathbf{W} are zero. We treated the case of leak rate $\alpha = 1$, since this produced an optimal behavior. The output layer consists of two units, defined by

$$y(t) = W_{out}x(t) \quad (3)$$

with $W_{out} \in \mathbb{R}^{2 \times N}$ trained by linear regression. The network architecture is shown in Figure 1.

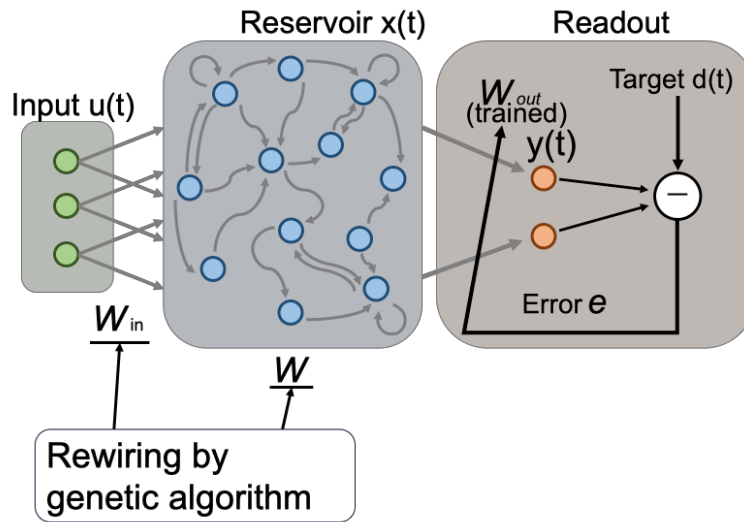


Figure 1. Network structure of eESN. The network consists of N recurrent nodes with connections from input W_{in} and internal connections W . The leak rate is fixed to $\alpha=1$ for optimal behavior.

2.2. Two Tasks Required for Consciousness

The eESN is trained on two distinct tasks: program task (logistic map prediction) and memory task (delayed logistic state recall). For the program task, the network learns to reproduce the logistic map for the input $\mathbf{u}(t)$,

$$d_1(t) \equiv au(t)(1 - u(t)), \quad (4)$$

where \mathbf{a} is a bifurcation parameter. The memory task requires reconstructing the logistic map state of the τ -steps in the past,

$$d_2(t) \equiv u(t - \tau). \quad (5)$$

Together, these tasks represent fundamental components of algorithmic inference: rule extraction (program) and information retention (memory).

2.3. Learning and Evolution

The training data consist of 1,000-step random input sequence, of which the initial 100 steps are discarded to remove transient. This length order is sufficient to guarantee a stationary random series—the case of $\mathbf{a} \sim 4.0$ yields independent-and-identical distributions. In the simulation, we used $\mathbf{a} = 3.98$. The target is described by $\mathbf{d}(t) = (\mathbf{d}_1(t), \mathbf{d}_2(t))^T$. The network is trained to achieve the least mean square error (MSE):

$$\min_{\{W_{out}\}} \sum_t \|d(t) - W_{out}x(t)\|^2. \quad (6)$$

The training data set and the target data set are provided by $\mathbf{X} = (\mathbf{x}(1), \mathbf{x}(2), \dots, \mathbf{x}(L)) \in \mathbb{R}^{N \times L}$ and $\mathbf{D} = (\mathbf{d}(1), \mathbf{d}(2), \dots, \mathbf{d}(L)) \in \mathbb{R}^{2 \times L}$, respectively. Then, the optimum solution can be obtained by $W_{out} = \mathbf{D}\mathbf{X}^T(\mathbf{X}\mathbf{X}^T)^\dagger$, where \dagger indicates Moore-Penrose inverse matrix. For the memory task, each

target series is generated for respective delay-time $\tau = 0 \sim 20$ and a total MSE was calculated for all target series.

For evolutionary optimization of W and W_{in} , we use the CHC genetic algorithm, which combines elitist selection, heterogeneous recombination, and cataclysmic mutation [5,17]. The fitness function is given by

$$-\sum_{d=0}^{20} (MSE_1(d) + MSE_2(d)) - \lambda E_W, \quad (7)$$

where $MSE_i (i = 1, 2)$ are MSEs for the two tasks. E_W is the number of nonzero internal connections, and $\lambda = 0.001$ is a sparsity penalty, whose value was empirically determined.

3. Results

Through supervised learning and evolutionary optimization, the network differentiated into two fundamental subnetworks: one responsible for memory retention and one for implementing the logistic-map rule. This differentiation is accompanied by a phase-transition-like shift from dense random connectivity to sparse and functionally organized structure (Figure 2). Internal structure looks still random, but output units differentiated into memory and program units, respectively.

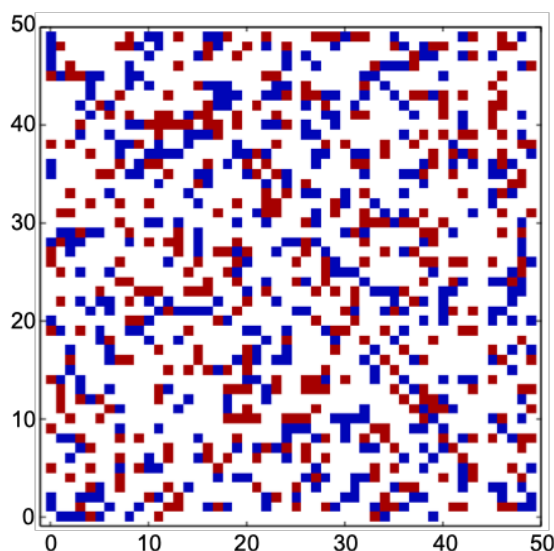


Figure 2. Functional differentiation by eESN. The internal structure evolved from a dense-random network (with two-thirds of all connections in the network) to a functionally differentiated sparse network (with around 30 % of all connections in the network). Figure shows evolved connections between nodes: red for excitatory connections and blue for inhibitory connections. Bifurcation parameter $a=3.98$, sparsity penalty $\lambda = 0.001$, and maximum memory delay $\tau = 20$. The number of nodes $N = 50$.

To highlight the advantage of eESN, we examined conventional ESNs, whose performance depends strongly on the spectral radius r . The program task achieved optimal accuracy in $r < 0.6$, whereas the memory task requires $r \sim 0.7$, where quasi-chaotic behaviors emerge. As a result, two tasks were incompatible in a single fixed reservoir (Figure 3). The eESN overcomes this limitation by evolving structural specialization.

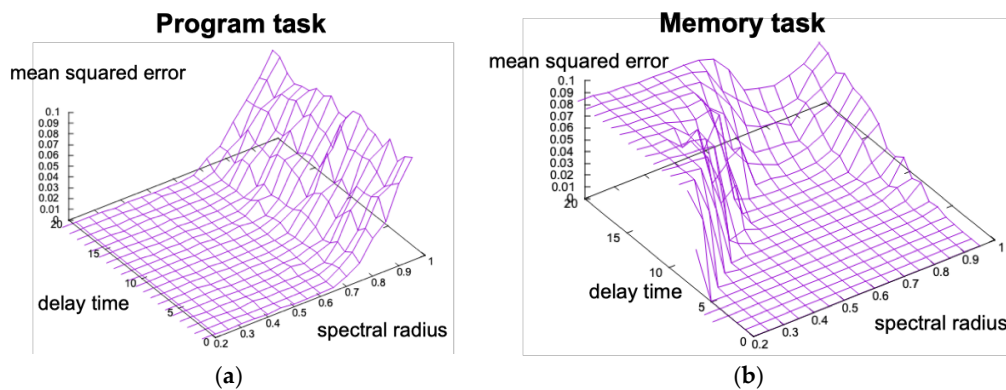


Figure 3. Optimal regions of spectral radius in the program and memory tasks are not compatible in conventional ESN. The program task achieved optimal accuracy with a spectral radius $r < 0.6$, whereas the memory task requires $r \sim 0.7$.

4. Discussion

As a first step toward a mathematical treatment of consciousness grounded in Kolmogorov complexity, we proposed a minimal model capable of generating memory and program units simultaneously via functional differentiation. Using an evolutionary ESN, we demonstrated that these two subsystems cannot be simultaneously optimized in a conventional ESN but can emerge naturally when evolutionary structural adaptation is introduced. The evolutionary process induces a phase transition. Joint minimization of output error, structural redundancy, and connection sparsity—interpreted as minimizing energy consumption—leads to an overall reduction of Kolmogorov complexity of the internal description of memory and program. This suggests that the system reaches an optimal complexity state capable of functioning as an algorithmic descriptor of an unpredictable environment. This stage of organized functionality corresponds to what we call **algorithmic consciousness**: a primitive form of consciousness characterized by inference, in line with historical accounts identifying consciousness with inferential processes [1,2]. Higher-order consciousness, including self-referential abilities, likely requires additional mechanisms such as social interaction. Such capabilities are not present in the current model. Nevertheless, the results demonstrate how key functional components of consciousness—memory and program—may emerge within a unified computational system governed by optimization and evolutionary constraints.

Acknowledgements: The authors (H.W and I.T) were partially supported by the JST Strategic Basic Research Programs (Symbiotic Interaction: Creation and Development of Core Technologies Interfacing Human and Information Environments, CREST Grant Number JPMJCR17A4). The authors were also partially supported by internal research fund of Sapporo City University.

References

1. For example, Garey LJ. *Brodman's Localisation in the Cerebral Cortex*. New York: Springer, 2006.
2. Glasser, M.F.; Coalson, T.S.; Robinson, E.C.; Hacker, C.D.; Harwell, J.; Yacoub, E.; Ugurbil, K.; Andersson, J.; Beckmann, C.F.; Jenkinson, M.; et al. A multi-modal parcellation of human cerebral cortex. *Nature* **2016**, *536*, 171–178.
3. Tsuda, I.; Yamaguti, Y.; Watanabe, H. Self-organization with constraints—A mathematical model for functional differentiation. *Entropy* **2016**, *18*, 74. <https://doi.org/10.3390/e18030074>.
4. Yamaguti, Y.; Tsuda, I. Functional differentiations in evolutionary reservoir computing networks. *Chaos* **2021**, *31*, 013137.
5. Watanabe, H.; Ito, T.; Tsuda, I. A mathematical model for neuronal differentiation in terms of an evolved dynamical system. *Neurosci. Res.* **2020**, *156*, 206–216.

6. Tomoda, Y.; Tsuda, I.; Yamaguti, Y. Emergence of functionally differentiated structures via mutual information minimization in recurrent networks. *Cogn. Neurodyn.* **2026**, *20*, 1–20. <https://doi.org/10.1007/s11571-025-10377-0>
7. Kolmogorov, A. On Tables of Random Numbers. *Theoretical Computer Science* **1998**, *207*, 387–395. [https://doi.org/10.1016/S0304-3975\(98\)00075-9](https://doi.org/10.1016/S0304-3975(98)00075-9)
8. Chaitin, G.J. (1987). *Algorithmic Information Theory*. Cambridge University Press. 1987. ISBN 9780521343060.
9. Crutchfield, J. P. The calculi of emergence: computation, dynamics and induction. *Physica D*, **1994**, *75*, 11–54.
10. Boole, G. *An Investigation into the Laws of Thought, on which are founded the Mathematical Theories of Logic and Probabilities*. UK: Walton & Maberly, 1854.
11. Peirce, C. S. On the Natural Classification of Arguments. *Proc. Amer. Acad. of Arts and Sci.* **1867**, *7*, 261–287.
12. Jaeger, H.; Haas, H. Harnessing Nonlinearity: Predicting Chaotic Systems and Saving Energy in Wireless Communication. *Science* **2004**, *304*, 78–80. <https://doi.org/10.1126/science.1091277>
13. Maass W., Natschlaeger T., and Markram H. Real-time computing without stable states: A new framework for neural computation based on perturbations. *Neural Computation* 2002, *14* (11), 2531–2560. <https://doi.org/10.1162/089976602760407955>
14. Dominey P.F. Complex sensory-motor sequence learning based on recurrent state representation and reinforcement learning. *Biol. Cybernetics* **1995**, *73*, 265–274. <https://doi.org/10.1007/BF00201428>
15. Yamaguti, Y.; Tsuda, I. Functional differentiations in evolutionary reservoir computing networks. *Chaos*, **2021**, *31*, 013137.
16. Tsuda, I.; Watanabe, H.; Tsukada, H.; Yamaguti, Y. On the nature of functional differentiation: The role of self-organization with constraints. *Entropy* **2022**, *24*, 240. <https://doi.org/10.3390/e24020240>
17. Eshelman, L.J. *The CHC Adaptive Search Algorithm: How to Have Safe Search When Engaging in Nontraditional Genetic Recombination*. In *Foundations of Genetic Algorithms*. Morgan Kaufmann Publishers: San Mateo, CA, USA, **1991**, pp. 265–283.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.