

Spontaneous Order of Self-organizing Systems*

Mau-Hsiang Shih[†] and Feng-Sheng Tsai[‡]

As Paul Adams wrote in the *Journal of Theoretical Biology* in 1998 [1], “Two of the most influential books in the history of biology are Darwin’s *On the Origin of Species* and Hebb’s *The Organization of Behavior*,” Donald Hebb’s classic book [6] is to have a profound impact upon how populations of neurons interact collectively to perform integrative activity underlying brain process. He postulated that the changes of synaptic strengths based on coincidence detection could sustain persistent reverberatory activity in cortical circuits. The circulating neural impulses between populations of neurons would continue to circulate, forming a diffuse self-assembling structure called “cell assemblies” [6, 13, 15].

There is increasing empirical support for Hebb’s contribution to neuropsychological theory [3, 5, 11] and there also stimulates an intensive effort to promote the building of computer or network models of the brain based on Hebbian synaptic plasticity (the coincidence-detection rule) [2, 4, 7, 9, 14, 16]. But there still lacks a complete mathematical explanation, a rigorous proof, that confirms the relation between Hebbian synaptic plasticity and cell-assembly postulate.

The main problem concerning the relation between Hebbian synaptic plasticity and cell-assembly postulate can be formulated as follows:

The Consolidation Problem

Are there any organizing principles underlying the dynamical-combinatorial process of the evolutionary network that allow one to describe neural populations underlying plasticity and to probe their time- and activity-dependent interactions capturing the characteristic property of the entire ensembles of cell-assemblies?

The model of the evolutionary network we are concerned with consists of a population of distinct integrate-and-fire processing units (McCulloch-Pitts neurons or neurons) [10, 12]; each constantly integrates all incoming signals transferred from synapses on its cell body and dendrites, and fires action potentials to send signals to other neurons when

*This work was supported by grants from the National Science Council of the Republic of China and the Faculty of Science of National Taiwan Normal University.

[†]Department of Mathematics, National Taiwan Normal University, 88 Sec. 4, Ting Chou Road, Taipei 116, Taiwan (mhshih@math.ntnu.edu.tw).

[‡]Department of Mathematics, National Taiwan Normal University, 88 Sec. 4, Ting Chou Road, Taipei 116, Taiwan (fstsai@hp715.math.ntnu.edu.tw).

the combined effort reaches its threshold. Name those neurons $1, \dots, n$ and denote by the ordered pair (i, j) the evolutionary coupling linking neuron j to neuron i . To each neuron i there is associated the threshold b_i and the active state variable $x_i = 0$ or 1 , and to each evolutionary coupling (i, j) there is associated the coupling strength variable a_{ij} . The phase space of the evolutionary network of n coupled neurons is denoted by $\{0, 1\}^n$, the binary code consisting of all 01-strings $x_1 x_2 \dots x_n$ of fixed-length n . The function *hea* is the Heaviside function: $hea(u) = 1$ for $u \geq 0$, otherwise 0 . The nonlinear dynamical system of the n coupled neurons is then modeled by the following nonlinear parametric equations:

$$x(t+1) = F_{s(t),t}(x(t)), \quad t = 0, 1, \dots, \quad (1)$$

$$A(t+1) = A(t) + D_{x(t) \rightarrow x(t+1)} A, \quad t = 0, 1, \dots, \quad (2)$$

where $x(t) = (x_1(t), \dots, x_n(t))$ is the neuronal active state at time t , $A(t) = (a_{ij}(t))_{n \times n}$ is the evolutionary coupling state at time t , $s(t)$ denotes the neurons that adjust their activity at time t , and $F_{s(t),t}(x)$ is the time-and-state varying function encoding the dynamics, whose i th component is defined by

$$[F_{s(t),t}(x)]_i = x_i \quad \text{if } i \notin s(t),$$

otherwise

$$[F_{s(t),t}(x)]_i = hea\left(\sum_{j=1}^n a_{ij}(t)x_j - b_i\right).$$

The i th row and j th column $\mathcal{D}_{x(t) \rightarrow x(t+1)} a_{ij}$ of the parametric matrix $\mathcal{D}_{x(t) \rightarrow x(t+1)} A$ is called the *ultraderivative* of the evolutionary coupling state at the evolutionary coupling (i, j) , which varies with respect to the neuronal active states changing from $x(t)$ to $x(t+1)$. The ultraderivatives quantify plasticity of evolutionary couplings that allows the system as a whole to undergo spontaneous organization.

As the evolutionary network evolves, there are associated subsets $\Lambda(t), \Omega(t)$ of the space $\{1, 2, \dots, n\} \times \{1, 2, \dots, n\}$ such that

$$\Lambda(t) = \{(i, j); \mathcal{D}_{x(t) \rightarrow x(t+1)} a_{ij} > 0\},$$

$$\Omega(t) = \{(i, j); \mathcal{D}_{x(t) \rightarrow x(t+1)} a_{ij} \geq 0\}.$$

That raises the search for the algorithm that dictates the existence of $T \geq 0$ and a subset V of $\{1, \dots, n\}$ such that the condition $\mathbf{1}(x(t)) = V$ holds true for all $t \geq T$, where $\mathbf{1}(x(t))$ denotes the collection of all active neurons at time t , and that

$$\bigcap_{t \geq T} \Omega(t) \supset V \times V \supset \bigcup_{t \geq T} \Lambda(t).$$

In this case, we say that the neurons in the subset V are *synchronized* with respect to $x(t)$, accompanying a spontaneously organized distribution of the positive and negative ultraderivatives to bind themselves into a single operational unit.

The bewildering complexity with which the evolutionary network identified is inherently associated with two simple quantities to measure the driving forces of the evolutionary network's dynamics. When a given neuron j is active at time t , there are two classes emerging from the active evolutionary coupling strengths $a_{ij}(t)$ for $i = 1, \dots, n$. The first consists of those with neurons i being quiescent at time t but fired at $t + 1$, while the second consists of those with neurons i being active at time t but quiescent at $t + 1$. By classifying all the active evolutionary coupling strengths as the above fired-driven or unfired-driven character, we define the *fired-driven strength* $FS(t)$ to be the sum of all $a_{ij}(t)$'s in the first class where j is taken over all active neurons, whereas the *unfired-driven strength* $US(t)$ the sum of all $a_{ij}(t)$'s in the second class where j is taken over all active neurons. The fired-driven strength $FS(t)$ is not necessarily greater than the unfired-driven strength $US(t)$ over time, but when the discrete flow $x(t)$ behaves in the way that $x(t_*) = x(t^*) \neq x(\hat{t})$ with $t_* < \hat{t} < t^*$ (a feedback loop initiated by active neurons at time t_*), the fired-driven strength and the unfired-driven strength in the period of t_* and t^* emerge the orderliness: $FS(t_*) + FS(t_* + 1) + \dots + FS(t^* - 1) > US(t_*) + US(t_* + 1) + \dots + US(t^* - 1)$. This gives a fundamental law of pulsedynamics that combines dynamical and structural complexity of the evolutionary network.

We introduce what we call the "Hebbian evolving algorithm," a generalized learning rule analogous to a coincidence-detection rule of the Hebbian synaptic plasticity that provides a representative for the choices of the ultraderivatives in the growth dynamics of the evolutionary network. They can be expressed in the formulas: $D_{x(t) \rightarrow x(t+1)} a_{ij} \geq 0$ if neurons i and j are active simultaneously at time $t + 1$, otherwise $D_{x(t) \rightarrow x(t+1)} a_{ij} < 0$, and further we require that $|D_{x(t) \rightarrow x(t+1)} a_{ij}| \geq |D_{x(t) \rightarrow x(t+1)} a_{ji}|$ if $\delta_{ij}(t + 1) > \delta_{ji}(t + 1)$, where $\delta_{ij}(t + 1) = 0$ or 1 is the indicator that indicates whether the active of neuron j at time t has been able to change the active state of neuron i at $t + 1$.

Define $E_U(t_*, t^*)$ to be $\sum_{i \in U} \min(\{a_{ii}(t); t = t_*, \dots, t^*\})$, a quantity that measures the minimal total excitability within the group of neurons U in the period of time t_* and t^* with $t_* \leq t^*$, where the coupling strength variable a_{ii} is considered to be a measure of excitability with respect to the neuron i [8] and, according to the working of the neuron i , the increased excitability has a tendency to decrease the threshold for generating action potentials. Let us call that the discrete flow $x(t)$ generated by (1) *iterates asynchronously* if $s(t)$ is a singleton for all $t = 0, 1, \dots$.

THEOREM. Consider the evolutionary network of n coupled neurons that subjects to the dynamics (1) and obeys the Hebbian evolving algorithm. Given any initial neuronal active state $x(0)$ in the phase space $\{0, 1\}^n$, if $x(t)$ iterates asynchronously and the

minimal total excitability satisfies the assembling coordination:

$$E_U(t_*, t^*) \geq \sum_{i,j \in U} \max(\{a_{ij}(t) - a_{ji}(t) \geq 0; t = t_*, \dots, t^*\} \cup \{0\}) \quad (3)$$

for each non-empty subset U of $\{1, 2, \dots, n\}$ and for each $t_*, t^* = 0, 1, \dots$ with $t_* \leq t^*$, then there exist $T \geq 0$ and a subset V of $\{1, 2, \dots, n\}$ such that $\mathbf{1}(x(t)) = V$ for all $t \geq T$ and that

$$\bigcap_{t \geq T} \Omega(t) \supset V \times V \supset \bigcup_{t \geq T} \Lambda(t).$$

Consequently, starting with an initial neuronal active state $x(0)$ the Hebbian evolving algorithm can be used to search a group of synchronized firing neurons, and the synchronized activity leads to a spontaneously organized distribution of the positive and negative ultraderivatives that feed back to reinforce the neurons to fire in synchrony, with positive feedbacks conspiring to produce a cascade of sync-dependent neural circuits, giving rise to the consolidation of neural circuitry.

The above theorem is to formulate a consolidation problem of evolutionary networks concerning the relation between Hebbian synaptic plasticity and Hebbian cell-assembly postulate, and to demonstrate a dominant theme of connectionist unraveling a dynamical-combinatorial process in a huge, interconnected system, in which the ongoing changes of the nodal-and-coupling dynamics underlying plasticity are guaranteed to result in group synchrony and sync-dependent circuits. Our proof of the theorem shows a potential role of pulsedynamics in the stability of a nonlinear evolutionary networked system, concerning the "driving forces" derived from the evolutionary network's nodal and coupling activity, without invoking any physical "energy" to control system dynamics.

References

- [1] P. ADAMS, *Hebb and Darwin*, J. Theor. Biol., 195 (1998), pp. 419–438.
- [2] J. A. ANDERSON, *A simple neural network generating an interactive memory*, Math. Biosci., 14 (1972), pp. 197–220.
- [3] F. CRICK, *Function of the thalamic reticular complex: the searchlight hypothesis*, Proc. Natl. Acad. Sci. USA, 81 (1984), pp. 4586–4590.
- [4] S. DEHAENE, M. KERSZBERG, AND J.-P. CHANGEUX, *A neuronal model of a global workspace in effortful cognitive tasks*, Proc. Natl. Acad. Sci. USA, 95 (1998), pp. 14529–14534.

- [5] K. D. HARRIS, J. CSICSVARI, H. HIRASE, G. DRAGOI, AND G. BUZSÁKI, *Organization of cell assemblies in the hippocampus*, *Nature*, 424 (2003), pp. 552–556.
- [6] D. O. HEBB, *The Organization of Behavior*, Wiley, New York, 1949.
- [7] J. J. HOPFIELD, *Neural networks and physical systems with emergent collective computational abilities*, *Proc. Natl. Acad. Sci. USA*, 79 (1982), pp. 2554–2558.
- [8] M. KLEIN, B. HOCHNER, AND E. R. KANDEL, *Facilitatory transmitters and cAMP can modulate accommodation as well as transmitter release in Aplysia sensory neurons. Evidence for parallel processing in a single cell*, *Proc. Natl. Acad. Sci. USA*, 83 (1986), pp. 7994–7998.
- [9] J. LÜCKE, *Hierarchical self-organization of minocolumnar receptive fields*, *Neural Netw.*, 17 (2004), pp. 1377–1389.
- [10] W. S. MCCULLOCH AND W. PITTS, *A logical calculus of the ideas immanent in nervous activity*, *Bull. Math. Biophys.*, 5 (1943), pp. 115–133.
- [11] B. MILNER, L. R. SQUIRE, AND E. R. KANDEL, *Cognitive neuroscience and the study of memory*, *Neuron*, 20 (1998), pp. 445–468.
- [12] M. MINSKY, *Computation: Finite and Infinite Machines*, Prentice-Hall, New York, 1967.
- [13] M. A. L. NICOLELIS, E. E. FANSELOW, AND A. A. GHAZANFAR, *Hebb's dream: the resurgence of cell assemblies*, *Neuron*, 19 (1997), pp. 219–221.
- [14] N. ROCHESTER, J. H. HOLLAND, L. H. HAIBT, AND W. L. DUDA, *Tests on a cell assembly theory of the action of the brain, using a large digital computer*, *IRE Trans. Inf. Theory*, 2 (1956), pp. 80–93.
- [15] T. J. SEJNOWSKI, *The book of Hebb*, *Neuron*, 24 (1999), pp. 773–776.
- [16] D. J. WILLSHAW, O. P. BUNEMAN, AND H. C. LONGUET-HIGGINS, *Non-holographic associative memory*, *Nature*, 222 (1969), pp. 960–962.