Neural Networks and Evolutionary Computation. Part II: Hybrid Approaches in the Neurosciences

Gerhard Weiß

Abstract— **This paper series focusses on the intersection of neural networks and evolutionary computation. It is addressed to researchers from artificial intelligence as well as the neurosciences.**

Part II provides an overview of hybrid work done in the neurosciences, and surveys neuroscientific theories that are bridging the gap between neural and evolutionary computation. According to these theories evolutionary mechanisms like mutation and selection act in real brains in somatic time and are fundamental to learning and developmental processes in biological neural networks.

Keywords— **Theory of evolutionary learning circuits, theory of selective stabilization of synapses, theory of selective stabilization of pre–representations, theory of neuronal group selection.**

I. Introduction

In the neurosciences biological neural networks are investigated at different organizational levels, including the molecular level, the level of individual synapses and neurons, and the level of whole groups of neurons (e.g., [11, 36, 44]). Several neuroscientific theories have been proposed which combine the fields of neural and evolutionary computation at these different levels. These are the theory of evolutionary learning circuits, the theories of selective stabilization of synapses and pre–representations, and the theory of neuronal group selection. According to these theories, neural processes of learning and development strongly base on evolutionary mechanisms like mutation and selection. In other words, according to these theories evolutionary mechanisms play in real brains and nervous systems the same role in somatic time as they do in ecosystems in phylogenetic time. (Other neuroscientific work which is closely related to these theories is described in [28, 46, 47].)

This paper overviews the hybrid work done in the neurosciences. Sections II to V survey the four evolutionary theories mentioned above. This includes a description of the major characteristics of these theories as well as a guide to relevant and related literature. Section VI concludes the paper with some general remarks on these theories and their relation to the hybrid approaches proposed in artificial intelligence.

The author is with the Institut für Informatik (H2), Technische Universität München, D-80290 München, Germany.

II. The Theory of Evolutionary Learning Circuits

According to the theory of evolutionary learning circuits (TELC for short) neural learning is viewed as the gradual modification of the information–processing capabilities of enzymatic neurons through a process of variation and selection in somatic time [12, 13]. In order to put this more precisely, first a closer look is taken at enzymatic neurons, and then the fundamental claims of the TELC are described.

The TELC starts from the point of view that the brain is organized into various types of local networks which contain enzymatic neurons, that is, neurons whose firing behavior is controlled by enzymes called excitases. (For details of this control and its underlying biochemical processes see e.g. [14, 15].) These neurons incorporate the **principle of double dynamics** [15] by operating at two levels of dynamics: at the level of readin or tactilization dynamics, the neural input patterns are transduced into chemical–concentration patterns inside the neuron; and at the level of readout dynamics, these chemical patterns are recognized by the excitases. Consequently, the enzymatic neurons themselves are endowed with powerful pattern–recognition capabilities where the excitases are the recognition primitives. Both levels of dynamics are gradually deformable as a consequence of the **structure–function gradualism** ("slight changes in the structure cause slight changes in the function") in the excitases. As Conrad pointed out, this structure–function gradualism is the key to evolution and evolutionary learning in general, and is a important condition for evolutionary adaptability in particular. (Evolutionary adaptability is defined as the extent to which mechanisms of variation and selection can be utilized in order to survive in uncertain and unknown environments [16].)

There are three fundamental claims made by the TESC: redundancy of brain tissue, specifity of neurons, and existence of brain–internal selection circuits. According to the claim for **redundany**, there are many replicas of each type of local network. This means that the brain consists of local networks which are interchangeable in the sense that they are highly similar with respect to their connectivity and the properties of their neurons. The claim for **specifity** says that the excitases are capable of recognizing specific chemical patterns and, with that, cause the enzymatic neurons to fire in response to specific input patterns. According to the third claim, the brain contains **selection**

circuits which direct the fitness–oriented, gradual modification of the local networks' excitase configurations. These selection circuits include three systems: a testing system which allows to check the consequences (e.g., pleasure or pain) of the outputs of one or several local networks for the organism; an evaluation system which assigns fitness values to the local networks on the basis of these consequences; and a growth–control system which stimulates or inhibits the production of the nucleic acids which code for the local networks' excitases on the basis of their fitness values. The nucleic acids, whose variability is ensured by random somatic recombination and mutation processes, diffuse to neighbouring networks of the same type (where they perform the same function because of the interchangeability property mentioned above). These claims imply that neural learning proceeds by means of the gradual modification of the excitase configurations in the brain's local networks through the repeated execution of the following **evolutionary learning cycle**:

1. Test and evaluation of the enzymatic neuron–based local networks. As a result, a fitness value is assigned to each network.
2. Selection of local networks. This involves the fitness–oriented regulation of the production of the excitase–coding nucleic acids, as well as their spreading to adjacent interchangeable networks.
3. Application of somatic recombination and selection to these nucleic acids. This maintains the range of the excitase configurations.

The execution stops when a local network is found which has a sufficiently high fitness. Conrad emphasized that this evolutionary learning cycle is much more efficient than natural evolution because the selection circuits enable an intensive selection even if there is hardly a difference between the fitness values of the interchangeable networks.

Finally, some references to related work. The TESC is part of extensive work focussing on the differences between the information processing capabilities of biological (molecular) systems and conventional computers; see e.g. [15, 16, 17]. A computational specification of the ESCM which concentrates on the pattern–processing capabilities of the enzymatic neurons, together with its sucessful application to a robot–control task, is contained in [29, 30, 31]. Another computational specification which concentrates on the intraneuronal dynamics of enzymatic neurons is described in [32]. A combination of these two specifications is described in [33]. Further related work being of particular interest from a computational point of view is presented in [1, 18].

III. The Theory of Selective Stabilization of Synapses

The theory of selective stabilization of synapses (TSSS for short) is presented in [7, 8]. This theory accounts for neural processes of learning and development by postulating that a somatic, evolutionary selection mechanism acts at the level of synapses and contributes to the wiring pattern in the adult brain. Subsequently the neurobiological basis and the major claims of the TSSS are depicted.

The neurobiological basis of the TSSS comprises aspects of both neurogenesis and neurogenetics. In vertebrates one can distinguish several processes of brain development. These are the cellular processes of cell division, movement, adhesion, differentiation, and death, and the synaptic processes of connection formation and elimination. (For details see e.g. [19, 20, 38].) The TSSS focusses on the "synaptic aspect" of neurogenesis; it deals with the outgrowth and the stabilization of synapses, and takes the developmental stage where **maximal synaptic wiring** exists as its initial state. The neurogenetic attidue of the TSSS constitutes a compromise between the preformist ("specified–by–genes") and the empirist ("specified–by– activity") view of brain development. It is assumed that the genes involved in brain development, the so–called genetic envelope, only specify the invariant characters of the brain. This includes, in particular, the connections between the main categories of neurons (i.e., between groups of neurons which are of the same morphological and biochemical type) and the rules of synaptic growth and stabilization. These rules allow for an activity–dependent, **epigenetic synapse formation** within the neuronal categories. (As Changeux formulated: "The genetic envelope offers a hazily outlined network, the activity defines its angles." [3, p. 193])

The TSSS makes three major claims. First, at the critical stage of maximal connectivity there is a significant but limited **redundany** within the neuronal categories as regards the specifity of the synapses. Second, at this time of so–called "structural redundany" any synapse may exist under (at least) **three states of plasticity**: labile, stable, and degenerate. Only the labile and stable synapses transmit nerve impulses, and the acceptable state transitions are those from labile to either stable or degenerate and from stable to labile. Especially, the state transition of a synapse is epigenetically regulated by all signals received by the postsynaptic soma during a given time interval. (The maximal synaptic connectivity, the mechanisms of its development, and the regulative and integrative properties of the soma are determinate expressions of the genetic envelope.) Third, the total activity of the developing network leads to the **selective stabilization** of some synapses, and to the regression of their functional equivalents. As a consequence, structural redundancy decreases and neuronal singularity (i.e., individual connectivity) increases. This provides a plausible explanation of the naturally occuring connection elimination occuring during neural development.

For further readings in the TSSS see e.g. [4, 5, 6, 10].

IV. The Theory of Selective Stabilization of Pre–Representations

The theory of selective stabilization of pre–representations (TSSP for short) can be viewed as an extension of the TSSS. This theory provides a selectionist view of neural

learning and development in the adult brain by postulating that somatic selection takes place at the level of neural networks [5, 10, 27]. Similar to the theory of neuronal group selection (section V), the TSSP may be viewed as an attempt to show how neurobiology and psychology are related to each other.

There are two major claims made by the TSSP. The first claim is that there exist **mental objects** or "neural representations" in the brain. A mental object is defined as a physical state achieved by the correlated and transitory (both electrical and chemical) activity of a cell assembly consisting of a large number of neurons having different singularities. According to the TSSP, three classes of mental objects are distinguished. First, primary percepts; these are labile mental objects whose activation depends on the direct interaction with the outside world and is caused by sensory stimulations. Second, stored representations; these are memory objects whose evocation does not demand environmental interaction and whose all–or–none activity behavior results from a stable, cooperative coupling between the neurons. And third, pre–representations; these are mental objects which are generated before and concomitant with any environmental interaction. Pre–representations are labile and of great variety and variability; they result from the spontaneous but correlated firing of neurons or groups of neurons. The second claim made by the TSSP is that learning in the adult brain corresponds to the **selective stabilization** of pre–representations, that means, the transition from selected pre–representations to stored representations. This requires, in the simplest case, the interaction with the environment, the criterion of selection is the resonance (i.e., spatial overlapping or firing in phase) between a primary percept and a pre–representation.

Further literature on the TSSP. In [9] the two selective–stabilization theories, TSSS and TSSP, are embedded in more general considerations on the neural basis of cognition. A formal model of neural learning and development on the basis of the TSSP is described in [22, 43].

V. The Theory of Neuronal Group Selection

The theory of neuronal group selection (TNGS for short) or "neural Darwinism" [23, 25] is the most rigorous and elaborate hybrid approach in the neurosciences. This theory, which has attracted much attention especially in the last few years, bridges the gap between biology and psychology by postulating that somatic selection is the key mechanism which establishes the connection between the structure and the function of the brain. As done in the preceding sections, below the major ideas of the TNGS are described.

There are three basic claims. First, during prenatal and early postnatal development, **primary repertoires** of degenerate neuronal groups were formed epigenetically by selection. According to the TNGS a neuronal group is considered as a local anatomical entity which consists of hundreds to thousands of strongly connected neurons, and degenerate neuronal groups are groups that have different structures but carry out the same function more or less well (they are nonisomorphic but isofunctional). The concept of degeneracy is fundamental to the TNGS; it implies both structural diversity and functional redundancy and, hence, ensures both a wide range of recognition and the reliability against the loss of neural tissue. Degeneracy naturally origins from the processes of brain development which are assumed to occur in an epigenetic manner and to elaborate several selective events at the cellular level. According to the regulator hypothesis, these complex developmental processes, as well as the selective events accompaning these processes, are guided by a relatively small number of cell adhesion molecules. Second, in the (postnatal) phase of behavioral experience, a **secondary repertoire** of functioning neuronal groups is formed by selection among the preexisting groups of each primary repertoire. This group selection is accomplished by epigenetic modifications of the synaptic strenghts without change of the connectivity pattern. According to the dual rules model, these modifications are realized by two synaptic rules that operate upon populations of synapses in a parallel and independent fashion: a presynaptic rule which applies to long–term changes in the whole target neuron and which affects a large number of synapses; and a postsynaptic rule which applies to short–term changes at individual synapses. The functioning groups are more likely to respond to identical or similar stimuli than the non–selected groups and, hence, contribute to the future behavior of the organism. A fundamental operation of the functional groups is to compete for neurons that belong to other groups; this competition affects the groups' functional properties and is assumed to play a central role in the formation and organization of cerebral cortical maps. Third, **reentry** – phasic signaling over re–entrant (reciprocal and cyclic) connections between different repertoires, in particular between topographic maps – allows for the spatiotemporal correlation of the responses of the repertoires at all levels in the brain. This kind of phasic signaling is viewed as an important mechanism supporting group selection and as being essential both to categorization and the development of consciousness. Reentry implies two fundamental neural structures: first, classification couples, that is, re–entrant repertoires that can perform classifications more complex than a single involved repertoire could do; and second, global mappings, that is, re-entrant repertoires that correlate sensory input and motor activity.

Some brief notes on how the TNGS accounts for psychological functions. Following Edelman's argumentation, categories do not exist apriori in the world (the world is "unlabeled"), and **categorization** is the fundamental problem facing the nervous system. This problem is solved by means of group selection and reentry. Consequently, categorization largely depends on the organism's interaction with its environment and turns out to be the central neural operation required for all other operations. Based on this view of categorization, Edelman suggests that **memory** is "the enhanced ability to categorize or generalize associa-

tively, not the storage of features or attributes of objects as a list" [25, p. 241] and that **learning**, in the minimal case, is the "categorization of complexes of adaptive value under conditions of expectancy" [25, p. 293].

There is a large body of literature on the TNGS. The most detailed depiction of the theory is provided in Edelman's book [25]. In order to be able to test the TNGS, several computer models have been constructed which embody the theory's major ideas. These models are Darwin I [24], Darwin II [26, 39, 25], and Darwin III [40, 41]. Reviews of the TNGS can be found in e.g. [21, 34, 35, 42, 37].

VI. Concluding Remarks

This paper overviewed neuroscientific theories which view real brains as evolutionary systems or "Darwin machines" [2]. This point of view is radically opposed to traditional instructive theories which postulate that brain development is directed epigenetically during an organism's interaction with its environment by rules for a more or less precise brain wiring. Nowadays most researchers agree that the instructive theories are very likely to be wrong und unrealistic, and that the evolutionary theories offer interesting and plausible alternatives. In particular, there is an increasing number of neurobiological facts and observations described in the literature which indicate that evolutionary mechanisms (and in particular the mechanism of selection) as postulated by the evolutionary theories are indeed fundamental to the neural processes in our brains.

Some final notes on the relation between the hybrid work done in the neurosciences and the hybrid work done in artificial intelligence (see part I of this paper series [45]). Whereas the neuroscientific approaches aim at a better understanding of the developmental and learning processes in real brains, the artificial intelligence approaches typically aim at the design of artificial neural networks that are appropriate for solving specific real-world tasks. Despite this fundamental difference and its implications, however, there are several aspects and questions which are elementary and significant to both the neuroscientific and the artificial intelligence approaches:

- Symbolic–subsymbolic intersection (e.g., "What are the neural foundations of high-level, cognitive abilities like concept formation?" and "How are symbolic entities encoded in the neural tissue?"),
- Brain wiring (e.g., "What are the principles of neural development?" and "How are the structure and the function of neural networks related to each other?"),
- Genetic encoding (e.g., "How and to what extend are neural networks genetically encoded?"), and
- Evolutionary modification (e.g., "At what network level and at what time scale do evolutionary mechanisms operate?" and "In how far do the evolutionary mechanisms influence the network structure?").

Because of this correspondence of interests and research topics it would be useful and stimulating for the neuroscientific and the artificial intelligence community to be aware of each others hybrid work. This requires an increased interdisciplinary transparency. To offer such a transparency is a major intention of this paper series.

References

[1] Akingbehin, K. & Conrad, M. (1989). A hybrid architecture for programmable computing and evolutionary learning. *Parallel Distributed Computing*, 6, 245–263.

[2] Calvin, W. H. (1987). The brain as a Darwin machine. *Nature*, 330, pp. 33–43.

[3] Changeux, J.–P. (1980). Genetic determinism and epigenesis of the neuronal network: Is there a biological compromise between Chomsky and Piaget? In M. Piatelli–Palmarini (Ed.), *Language and learning – The debate between Jean Piaget and Noam Chomsky* (pp. 184–202). Routledge & Kegan Paul.

[4] Changeux, J.–P. (1983). Concluding remarks: On the "singularity" of nerve cells and its ontogenesis. In J.–P. Changeux, J. Glowinski, M. Imbert, & F. E. Bloom (Eds.), *Molecular and cellular interactions underlying higher brain function* (pp. 465–478). Elsevier Science Publ.

[5] Changeux, J.–P. (1983). *L'Homme neuronal*. Fayard.

[6] Changeux, J.–P. (1985). Remarks on the complexity of the nervous system and its ontogenesis. In J. Mehler & R. Fox (Eds.), *Neonate cognition. Beyond the blooming buzzing confusion* (pp. 263–284). Lawrence Erlbaum.

[7] Changeux, J.–P., Courrege, P., & Danchin, A. (1973). A theory of the epigenesis of neuronal networks by selective stabilization of synapses. In *Proceedings of the National Academy of Sciences USA*, 70(10), 2974–2978.

[8] Changeux, J.–P., & Danchin, A. (1976). Selective stabilization of developing synapses as a mechanism for the specification of neuronal networks. *Nature*, 264, 705–712.

[9] Changeux, J.–P., & Dehaene, S. (1989). Neuronal models of cognitive functions. *Cognition*, 33, 63–109.

[10] Changeux, J.–P., Heidmann, T., & Patte, P. (1984). Learning by selection. In P. Marler & H. S. Terrace (Eds.), *The biology of learning* (pp. 115–133). Springer.

[11] Changeux, J.–P., & Konoshi, M. (Eds.). (1986). *The neural and molecular basis of learning*. Springer.

[12] Conrad, M. (1974). Evolutionary learning cicuits. *Journal of Theoretical Biology*, 46, 167–188.

[13] Conrad M. (1976). Complementary models of learning and memory. *BioSystems*, 8, 119–138.

[14] Conrad, M. (1984). Microscopic–macroscopic interface in biological information processing. *BioSystems*, 16, 345–363.

[15] Conrad, M. (1985). On design principles for a molecular computer. *Communications of the ACM*, 28(5), 464–480.

[16] Conrad, M. (1988). The price of programmability. In R. Herken (Ed.), *The universal Turing machine – A half–century survey* (pp. 285–307). Kammerer & Unverzagt.

[17] Conrad, M. (1989). The brain–machine disanalogy. *BioSystems*, 22, 197–213.

[18] Conrad, M., Kampfner, R. R., Kirby, K. G., Rizki, E. N., Schleis, G., Smalz, R., & Trenary, R. (1989). Towards an artificial brain. *BioSystems*, 23, 175–218.

[19] Cowan, W. M. (1978). Aspects of neural development. *International Reviews of Physiology*, 17, 150–91.

[20] Cowan, W. M., Fawcett, J. W., O'Leary, D. D. M., & Stanfield, B. B. (1984). Regressive events in neurogenesis. *Science*, 225, 1258–1265.

[21] Crick, F. (1989). Neural Edelmanism. *Trends in Neurosciences*, 12, 240–248. (Reply from G. N. Reeke, R. Michod and F. Crick: Trends in Neurosciences, 13, 11–14.)

[22] Dehaene, S., Changeux, J.–P., & Nadal, J.–P. (1987). Neural networks that learn temporal sequences by selection. In *Proceedings of the National Academy of Sciences USA*, 84, pp. 2727–2731.

[23] Edelman, G.M. (1978). Group selection and phasic reentrant signaling: A theory of higher brain function. In G. M. Edelman & V. B. Mountcastle (Eds.), *The mindful brain. Cortical organization and the group–selective theory of higher brain functions* (pp. 51–100). The MIT Press.

[24] Edelman, G.M. (1981). Group selection as the basis for higher brain function. In F. O. Schmitt, F. G. Worden, G. Adelman

& S. G. Dennis (Eds.), *The organization of the cerebral cortex* (pp. 535–563). The MIT Press.

[25] Edelman, G.M. (1987). *Neural Darwinism. The theory of neuronal group selection*. Basic Books.

[26] Edelman, G.M., & Reeke, G.N. (1982). Selective networks capable of representative transformations, limited generalizations, and associative memory. In *Proceedings of the National Academy of Sciences USA*, 79, 2091–2095.

[27] Heidmann, A., Heidmann, T.M., & Changeux, J.–P. (1984). Stabilisation selective de representations neuronales par resonance entre "presepresentations"spontanes du reseau cerebral et "percepts". In *C. R. Acad. Sci. Ser. III*, 299, 839–844.

[28] Jerne, N. K. (1967). Antibodies and learning: selection vs. instruction. In G. C. Quarton, T. Melnechuk & F. O. Schmitt (Eds.), *The neurosciences: a study program* (pp. 200–205). Rockefeller University Press.

[29] Kampfner, R. R. (1988). Generalization in evolutionary learning with enzymatic neuron–based systems. In M. Kochen & H. M. Hastings (Eds.), *Advances in cognitive science. Steps toward convergence* (pp. 190–209). Westview Press, Inc.

[30] Kampfner, R. R., & Conrad, M. (1983). Computational modeling of evolutionary learning processes in the brain. *Bulletin of Mathematical Biology*, 45(6), 931–968.

[31] Kampfner, R. R., & Conrad, M. (1983). Sequential behavior and stability properties of enzymatic neuron networks. *Bulletin of Mathematical Biology*, 45(6), 969–980.

[32] Kirby, K. G., & Conrad, M. (1984). The enzymatic neuron as a reaction–diffusion network of cyclic nucleotides. *Bulletin of Mathematical Biology*, 46, 765–782.

[33] Kirby, K. G., & Conrad, M. (1986). Intraneuronal dynamics as a substrate for evolutionary learning. In *Physica 22D*, 205–215.

[34] Michod, R. E. (1989). Darwinian selection in the brain. *Evolution*, (3), 694–696.

[35] Nelson, R. J. (1989). Philosophical issues in Edelman's neural Darwinism. *Journal of Experimental and Theoretical Artificial Intelligence*, 1, 195–208.

[36] Neuroscience Research (1986). Special issue 3: Synaptic plasticity, memory and learning.

[37] Patton, P., & Parisi, T. (1989). Brains, computation, and selection: An essay review of Gerald Edelman's Neural Darwinism. *Psychobiology*, 17(3), 326–333.

[38] Purves, D., & Lichtman, J. W. (1985). *Principles of neural development*. Sinauer Associates Inc.

[39] Reeke, G. N. jr., & Edelman, G. M. (1984). Selective networks and recognition automata. In *Annals of the New York Academy of Sciences*, 426 (Special issue on computer culture), 181–201.

[40] Reeke, G. N. jr., & Edelman, G. M. (1988). Real brains and artificial intelligence. *Deadalus*, 117(1), 143–173.

[41] Reeke, G. N. jr., Sporns, O., & Edelman, G. M. (1988). Synthetic neural modeling: a Darwinian approach to brain theory. In R. Pfeifer, Z. Schreter, F. Fogelman–Soulie & L. Steels (Eds.), *Connectionism in perspective*. Elsevier.

[42] Smoliar, S. (1989). Book review of [25]. *Artificial Intelligence*, 39, 121–139.

[43] Toulouse, G., Dehaene, S., & Changeux, J.–P. (1986). Spin glass models of learning by selection. In *Proceedings of the National Academy of Sciences USA*, 83, 1695–1698.

[44] Trends in Neurosciences (1988). Vol. 11(4), Special issue: Learning, memory.

[45] Weiß, G. (1993, submitted to IEEE World Congress on Computational Intelligence). Neural networks and evolutionary computation. Part I: Hybrid approaches in artificial intelligence.

[46] Young, J. Z. (1973). Memory as a selective process. In *Australian Academy of Science Report: Symposium on Biological Memory; Australian Academy of Science* (pp. 25–45).

[47] Young, J. Z. (1975). Sources of discovery in neuroscience. In F. G. Worden, J. P. Swazey & G. Edelman (Eds.), *The neurosciences: Paths of discovery* (pp. 15–46). Oxford University Press.