Chapter 1

Reverse Engineering for Biologically-Inspired Cognitive Architectures: A Critical Analysis

Andreas Schierwagen

Abstract Research initiatives on both sides of the Atlantic try to utilize the operational principles of organisms and brains to develop biologically inspired, artificial cognitive systems. This paper describes the standard way bio-inspiration is gained, i.e. decompositional analysis or reverse engineering. The indisputable complexity of brain and mind raise the issue of whether they can be understood by applying the standard method. Using Robert Rosen's modeling relation, the scientific analysis method itself is made a subject of discussion. It is concluded that the fundamental assumption of cognitive science, i.e. complex cognitive systems are decomposable, must be abandoned. Implications for investigations of organisms and behavior as well as for engineering artificial cognitive systems are discussed.

1.1 Introduction

> *Wer will was Lebendig's erkennen und beschreiben,*
> *Sucht erst den Geist heraus zu treiben,*
> *Dann hat er die Teile in seiner Hand,*
> *Fehlt, leider! nur das geistige Band.*
>
> J.W. GOETHE, Faust, Erster Teil[a]

For some time past, computer science and engineering devote close attention to the functioning of the brain. It has been argued that recent advances in cognitive science and neuroscience have enabled a rich scientific understanding of how cognition works in the human brain. Thus, research programs have been initiated by leading research organizations on both sides of the Atlantic to develop new cogni-

---

Institute for Computer Science, Intelligent Systems Department
University of Leipzig
Leipzig, Germany
e-mail: schierwa@informatik.uni-leipzig.de
`http://www.informatik.uni-leipzig.de/~schierwa`

tive architectures and computational models of human cognition [1, 2, 3, 4, 5, 6] (see also [7], and references therein).

Two points are emphasized in the research programs: First, there is impressing abundance of available experimental brain data, and second, we have the computing power to meet the enormous requirements to simulate a complex system like the brain. Given the improved scientific understanding of the operational principles of the brain as a complexly organized system, it should then be possible to build an operational, quantitative model of the brain. Tuning the model could be achieved using the deluge of empirical data.

The main method used in empirical research to integrate the data derived from the different levels of the brain organization is *reverse engineering*. Originally a concept in engineering and computer science, reverse engineering involves as first step the process of detailed examination of a functional system and its dissecting at the physical level into component parts, i.e. *decompositional analysis*. In a second step, the (*re*–)construction of the system is attempted, see below. This principle is usually not much discussed with respect to its assumptions, conditions and range[1] but see [10, 11, 12].

Together, according to the prevailing judgement there is nothing *in principle* that we do not understand about brain organization. All the knowledge about its ’building blocks’ and connectivity is present (or can be derived), and needs only to be put into the model. This view is widely taken; it represents the belief in the power of the reverse engineering method. As I am going to show in this paper, there is, however, substantial evidence to question this belief. It turns out that this method in fact ignores something fundamental, namely that biological and engineered systems are basically different in nature.

The paper is organized as follows. Section 1.2 presents the fundamental assumption employed in the cognitive and brain sciences, i.e. the assumption that both brain and mind are decomposable. In Section 1.3, the concepts of decompositional analysis, reverse engineering and localization are reviewed. The following Section 1.4 is devoted to modularization and its relation to the superposition principle of system theory. Then, Section 1.5 shortly touches on Blue Brain and SyNAPSE, two leading reverse-engineering projects. Both projects are based on the hypothesis of the columnar organization of the cortex. The rationale underlying reverse engineering in neurocomputing or computational neuroscience is outlined. New findings are mentioned indicating that the concept of the basic uniformity of the cortex is untenable. Section 1.6 ponders about the claim that non-decomposability is not an intrinsic property of complex systems but is only in our eyes, due to insufficient mathematical techniques. For this, Rosen’s modeling relation is explained which enables us to make the scientific analysis method itself a subject of discussion. It is concluded that the fundamental assumption of cognitive science must be abandoned. We end

---

[1] Only recently, differences between proponents of reverse engineering on how it is appropriately to be accomplished became public. The prominent heads of two reverse engineering projects, Markram [2] and Modha [8], disputed publicly as to what granularity of the modeling is needed to reach a valid simulation of the brain. Markram questioned the authenticity of Modha’s respective claims [9].

the paper by some conclusions about the relevance of Rosen's [13, 14, 15] work for the study of organisms and behavior as well as for engineering artificial cognitive systems.

1.2 Conceptual Foundations of Cognitive and Brain Sciences

Brains, even those of simple animals, are enormously complex structures, and it is a very ambitious goal to cope with this complexity. The scientific disciplines involved in cognition and brain research (Fig. 1.1) are committed to a common method to explain the properties and capacities of complex systems. This method is decompositional analysis, i.e. analysis of the system in terms of its components or subsystems.

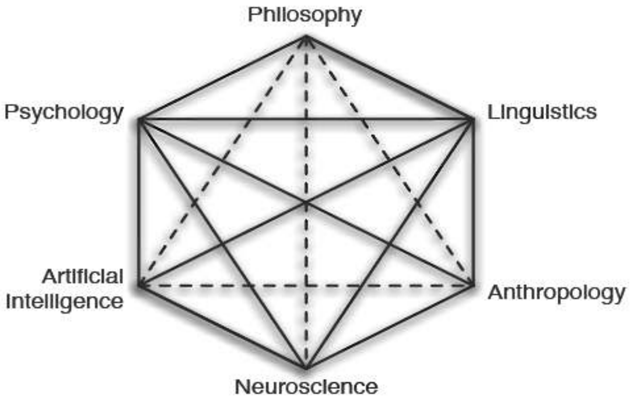

Fig. 1.1 The Cognitive Hexagon (as of 1978 [16]). Cognitive science comprised six disciplines, all committed to decompositional analysis as the basic research method.

Since Simon's influential book "The Sciences of the Artificial" [17], (near-) decomposability of complex systems has been accepted as fundamental for the cognitive and brain sciences (CBS). We call this the *fundamental assumption* for the cognitive and brain sciences. Simon [17], Wimsatt [18] and Bechtel and Richardson [19], among others, have further elaborated this concept. They consider decomposability a continously varying system property, and state, roughly, that systems fall on a continuum from aggregate (full decomposable) to integrated (non-decomposable). The *fundamental assumption* states that cognitive and brain systems are non-ideal aggregate systems; the capacities of the components are internally realized (strong intra-component interactions), and interactions between components do not appreciably contribute to the capacities, they are much weaker than the intra-component interactions. Hence, the description of the complex system as a set of weakly interacting components seems to be a good approximation. This property of complex systems, which should have evolved through natural selection, was called near-decomposability and characterized as follows [17]:

Near-decomposability.

1. In a nearly decomposable system, the short-run behavior of each of the component subsystems is approximately independent of the short-run behavior of the other components;
2. in the long run the behavior of any one of the components depends in only an aggregate way on the behavior of the other components [17, p.100].

Thus, if the capacities of a near-decomposable system are to be explained, to some approximation its components can be studied in isolation, and based on their known interactions, their capacities eventually combined to generate the systems behavior. In other words, the aforementioned *fundamental assumption* represents the conceptual basis for reverse engineering the brain and mind.

Let us summarize this assumption because it is of central importance in the following:

Fundamental assumption for cognitive and brain sciences.

> Cognitive and brain systems are non-ideal aggregate systems. The capacities of the components are internally realized (strong intra-component interactions) while interactions between components are negligible with respect to capacities. Any capacity of the whole system then results from superposition of the capacities of its subsystems. This property of cognitive and brain systems should have evolved through natural selection and is called near-decomposability.

1.3 Decompositional Analysis, Localization and Reverse Engineering

The primary goal of cognitive science and its subdisciplines is to understand cognitive capacities like vision, language, memory, planning etc. Capacities are considered as dispositional properties which can be explained via decompositional analysis, see Fig. 1.2. In CBS, two types of decompositional analysis are differentiated, i.e. functional analysis and structural analysis [20, 21, 22]. *Functional analysis* is the type of decompositional analysis that proceeds without reference to the material composition of the system. It is concerned with the sub-functions of some hypothesized components of the whole system which enable this whole system to have certain capacities and properties and to realize corresponding functions.

Structural analysis involves to attempt to identify the structural, material components of the system. Thus, the material system S can be decomposed into context-independent components S_j, i.e. their individual properties are independent of the decomposition process itself and of S's environment.

Functional analysis and structural analysis must be clearly differentiated, although in practice, there is a close interplay between them (as indicated by the double arrow in Fig. 1.2). This is obvious in the *localization approach* which combines both analysis types, i.e. a specific component function is linked with a specific structural component. Functional analysis should also be differentiated from capacity analysis. The former is concerned with the functions performed by components of the whole system which enable this whole system to have certain capacities and properties. The latter is concerned with the dispositions or abilities of the whole system, whereas functional and structural analysis is concerned with the functional and structural bases of those dispositions or abilities.

Understating the case, the localization approach is sometimes described as a hypothetical identification which is to serve as research heuristics [19]. In fact, however, the majority of cognitive scientists considers it as fundamental and indispensable (e.g. [28]).

Obviously, decompositional analysis and reverse engineering are closely related. Reverse engineering is a two-step method: It has the decompositional analysis of the original system as the first, basic step, while the second step consists in creating duplicates of the original system, including computer models.

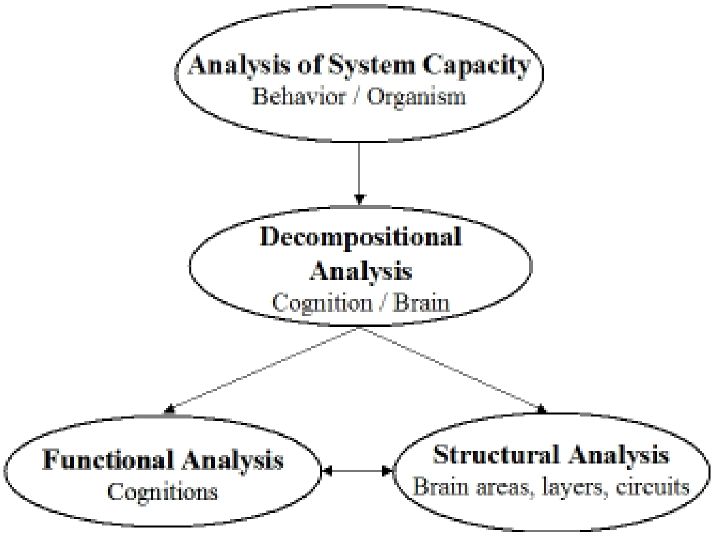

Fig. 1.2 View on decompositional analysis of brain and cognition. See text for details.

It should be noticed that there is no reason to assume that functional and structural components match up one-to-one! Of course, it might be the case that some functional components map properly onto individual structural components - the dream of any cognitive scientist working as 'reverse engineer'. It is rather probable, however, for a certain functional component to be implemented by non-localized, spatially distributed material components. Conversely, a given structural component may implement more than one distinct function. According to Dennett [24, p. 273]: "In a system as complex as the brain, there is likely to be much 'multiple, superimposed functionality'." With other words, we cannot expect specific functions to be mapped to structurally bounded neuronal structures, and vice versa. It is now well known that Dennett's caveat has been proved as justified (e.g. [25]). Thus the value of the localization approach as 'research heuristics' seems rather dubious [26, 27].

1.4 Complex Systems and Modularization

In CBS and in other fields of science, the components of near-decomposable systems are called modules. This term originates from engineering; it denotes the process of decomposing a product into building blocks - modules - with specified interfaces, driven by the designer's interests and intended functions of the product. Modularized systems are linear in the sense that they obey an analog of the superposition principle of linear system theory in engineering [29]. The behavior of a decomposable system results from the linear combination of the behavior of the system modules. In some respects, this principle represents a formal underpinning of the constructive step in reverse engineering[2] (see Sections 1.1, 1.3, 1.5). The terms 'linear' and 'nonlinear' are often used in this way: 'Linear' systems are decomposable into independent modules with linear, proportional interactions while 'nonlinear' systems are not[3] [29, 30].

Applying this concept to the systems at the other end of the complexity scale, the integrated systems are basically not decomposable, due to the strong, nonlinear interactions involved. Thus, past or present states or actions of any or most subsystems always affect the state or action of any or most other subsystems. In practice, analyses of integrated systems nevertheless try to apply the methodology for decomposable systems, in particular if there is some hope that the interactions can be linearized. Such linearizable systems have been above denoted as nearly decomposable. However, in the case of strong nonlinear interactions, we must accept that decompositional analysis is not applicable.

Already decades ago this insight was stressed. For example, Levins [31, p.76 ff.] proposed a classification of systems into aggregate, composed and evolved systems. While the aggregate and the composed would not cause serious problems for scientific analyses, Levins emphasized the special character of evolved systems:

---

[2] A corresponding class of models in mathematics is characterized by a theorem stating that for homogeneous linear differential equations, the sum of any two solutions is itself a solution.

[3] We must differentiate between the natural, complex system and its description using modeling techniques from linear system theory or nonlinear mathematics.

> A third kind of system no longer permits this kind of analysis. This is a system in which the component subsystems have evolved together, and are not even obviously separable; in which it may be conceptually difficult to decide what are the really relevant component subsystems.... The decomposition of a complex system into subsystems can be done in many ways... it is no longer obvious what the proper subsystems are, but these may be processes, or physical subsets, or entities of a different kind.

The question then arises: Should we care about integrated systems, given the *fundamental assumption* that all relevant systems are nearly decomposable? Non–decomposability then would be only in our eyes, and not an intrinsic property of strongly nonlinear systems, and – as many cognitive and computer scientists believe – scientific progress will provide us with the new mathematical techniques required to deal with nonlinear systems. We will return to this problem in Section 1.6.

1.5 Reverse Engineering the Brain and Neurocomputing

1.5.1 The Column Concept

A guiding idea about the composition of the brain is the hypothesis of the columnar organization of the cerebral cortex. It was developed mainly by Mountcastle, Hubel and Wiesel, and Szenthágothai (e.g. [32, 33, 34]), in the spirit of the highly influential paper "The basic uniformity in structure of the neocortex" published in 1980 [37]. According to this hypothesis (which has been taken more or less as fact by many experimental as well as theoretical neuroscientists), the neocortex is composed of 'building blocks' of repetitive structures, the 'columns' or neocortical microcircuits, and it is characterized by a basic canonical pattern of connectivity. In this scheme all areas of neocortex would perform identical or similar computational operations with their inputs.

Referring to and based on these works, several projects started recently, among them the *Blue Brain Project* [2] and the *SyNAPSE Project* [5]. They are considered to be "attempts to reverse-engineer the mammalian brain, in order to understand brain function and dysfunction through detailed simulations"[2] or, more pompous, "to engineer the mind"[5]. The central role in these projects play cortical microcircuits or columns. As Maas and Markram [35] formulate, it is a "tempting hypothesis regarding the computational role of cortical microcircuits ... that there exist genetically programmed stereotypical microcircuits that compute certain basis function." Their paper well illustrates the modular approach fostered, e.g. by [36, 38, 39, 12]. Invoking the localization concept, the tenet is that there exist fundamental correspondences among the anatomical structure of neuronal networks, their functions, and the dynamic patterning of their active states. Starting point is the 'uniform cortex' with the cortical microcircuit or column as the structural component. The question for the functional component is answered by assuming that there is a one-to-one relationship between the structural and the functional component (see Section 1.3). Together, the modularity hypothesis of the brain is considered to be both structurally and functionally well justified. As quoted above, the goal is to examine the hypothesis that there exist genetically programmed stereotypical microcircuits that compute certain basis function.

1.5.2 Neocortical Microcircuits and Basis Functions

The general approach to cognitive capacities takes for granted that "cognition is computation", i.e. the brain produces the cognitive capacities by computing functions[4]. According to the scheme formulated in Section 1.3, reverse engineering of the cortex (or some subsystem) as based on the column concept and performed from a neurocomputational perspective then proceeds as follows.

Reverse engineering the cortex.

1. Capacity analysis: A specific cognitive capacity is identified which is assumed to be produced through the brain by computing a specific function.
2. Decompositional analysis:
   a. Functional (computational) analysis: From mathematical analysis and approximation theory it is well-known that a broad class of practically relevant functions f can be approximated by composition or superposition of some basis function. If we assume that some basis functions can be identified, they provided the components of a hypothetical functional decomposition.
   b. Structural analysis: Provide evidence that cortical microcircuits are the anatomical components of the cortex.
3. Localization: The next step consisted in linking the component functions with the component parts by suggesting that the basis function are computed by the structural components (columns or cortical microcircuits).
4. Synthesis/Superposition: The specific cognitive capacity or function under study now can be explained by combining the basis functions determined in step 2.a. The composition rules were implicitly contained in the interconnection pattern of the circuits, thus enabling the brain system under study to generate the specific cognitive capacity.

The question now is, however - Are the assumptions and hypotheses made appropriate, or must they considered as too unrealistic? In fact, most of the underlying hypotheses have been challenged only recently. To start with the assumptions about the structural and functional composition of the cortex, the notion of a basic uniformity in the cortex with respect to the density and types of neurons per column

---

[4] See [7] for discussion of the computational approaches (including the neurocomputational one) to brain function, and their shortcomings.

for all species turned out to be untenable (e.g. [40, 41, 42]). It has been impossible to find the cortical microcircuit that computes specific basis function [43]. No genetic mechanism has been deciphered that designates how to construct a column. It seems that the column structures encountered in many species (but not in all) represent spandrels (structures that arise non-adaptively, i.e. as an epiphenomenon) in various stages of evolution [44].

If we evaluate the column concept of the cortex employed in theories of brain organization, it is obvious that – employing the localization concept mentioned in Section 1.3 – hypothesized structural components (cortical columns) have been identified with alike hypothetical functional components (basis function).

There is evidence, however, for a certain functional component to be implemented by spatially distributed networks and, vice versa, for a given structural component to implement more than one distinct function. With other words, it is not feasible for specific functions to be mapped to structurally bounded neuronal structures [25, 40, 41, 42].

This means, although the column concept is an attractive idea both from neurobiological and computational point of view, it cannot be used as an unifying principle for understanding cortical function. Thus, it has been concluded that the concept of the cortex as a 'large network of identical units' should be replaced with the idea that the cortex consists of 'large networks of diverse elements' whose cellular and synaptic diversity is important for computation [45, 46, 47].

It is worth to notice that the reported claims for changes of the research concept completely remain within the framework of reverse engineering. A more fundamental point of criticism concerns the methods of decompositional analysis and reverse engineering themselves and will be discussed in the next section.

1.6 Complex Systems and Rosen's Modeling Relation

In Section 1.4, we concluded that integrated systems are basically non-decomposable, thus resisting the standard analysis method. We raised the question: Should we at all care about integrated systems, given the *fundamental assumption* that all relevant systems are nearly decomposable?

According to the prevalent viewpoint in CCN, non-decomposability is not an intrinsic property of complex, integrated systems but is only in our eyes, due to insufficient mathematical techniques (e.g. [48, 49, 50, 51]). Bechtel and Richardson, instead, warn that the assumption according to which nature is decomposable and hierarchical might be false [19, p. 27]: "There are clearly risks in assuming complex natural systems are hierarchical and decomposable."

Rosen [14, 15] has argued that understanding complex, integrated systems requires making the scientific analysis method itself a subject of discussion. A powerful method of understanding and exploring the nature of the scientific method, and in particular, reverse engineering, provides his modeling relation. It is this relation by which scientists bring "entailment structures into congruence" [14, p. 152]. This can be explained as follows.

The modeling relation is the set of mappings shown in Figure 1.3 [13, 52]. It relates two systems, a natural system N and a formal system F, by a set of arrows depicting processes or mappings. The assumption is that this diagram represents the various processes which we are carrying out when we perceive the world. N is a part of the physical world that we wish to understand (in our case: organism, brain), in which things happen according to rules of causality (arrow 1). On the right, F represents symbolically the parts of the natural system (observables) which we are interested in, along with formal rules of inference (arrow 3) that essentially constitute our working hypotheses about the way things work in N, i.e. the way in which we manipulate the formal system to try to mimic causal events observed or hypothesized in the natural system on the left. Arrow 2 represents the encoding of the parts of N under study into the formal system F, i.e. a mapping that establishes the correspondence between observables of N and symbols defined in F. Predictions about the behavior in F, according to F's rules of inference, are compared to observables in N through a decoding represented by arrow 4. When the predictions match the observations on N, we say that F is a successful model for N.

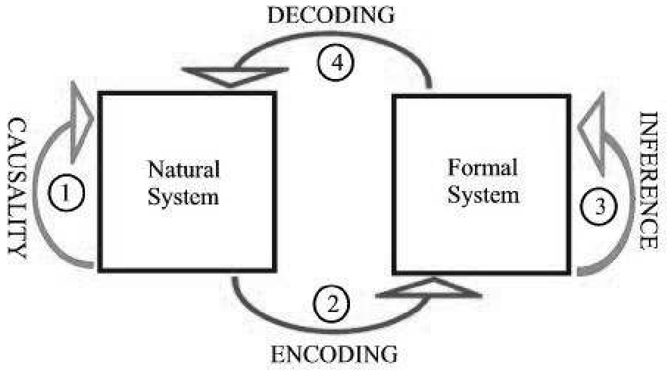

Fig. 1.3 Rosen's Modeling Relation. A natural system N is modeled by a formal system F. Each system has its own internal entailment structures (arrows 1 and 3), and the two systems are connected by the encoding and decoding processes (arrows 2 and 4). From http://www.panmere.com.

It is important to note that the encoding and decoding mappings are independent of the formal and natural systems, respectively. In other words, there is no way to arrive at them from within the formal system or natural system. That is, the act of modeling is really the act of relating two systems in a subjective way. That relation is at the level of observables; specifically, observables which are selected by the modeler as worthy of study or interest.

Given the modeling relation and the detailed structural correspondence between our percepts and the formal systems into which we encode them, it is possible to make a dichotomous classification of systems into those that are *simple* or *predicative* and those that are *complex* or *impredicative*. This classification can refer to formal inferential systems such as mathematics or logic, as well as to physical systems. As Rosen showed [13], a simple system is one that is definable completely by algorithmic method: All the models of such a system are Turing-computable or simulable. When a single dynamical description is capable of successfully modeling a system, then the behaviors of that system will, by definition, always be correctly predicted. Hence, such a system will be *predicative* in the sense that there will exist no unexpected or unanticipated behavior.

A complex system is by exclusion not a member of the syntactic, algorithmic class of systems. Its main characteristics are as follows. A complex system possesses non-computable models; it has inherent impredicative loops in it. This means, it requires multiple partial dynamical descriptions - no one of which, or combination of which, suffices to successfully describe the system.

It is not a purely syntactic system, it necessarily includes semantic elements, and is not formalizable. Complex systems also differ from simple ones in that complex systems are not simply summations of parts - they are non-decomposable. This means, when a complex system is decomposed, its essential nature is broken by breaking its impredicative loops.

This has important effects. Decompositional analysis is inherently destructive to what makes the system complex - such a system is not decomposable without losing the essential nature of the complexity of the original system! In addition, by being not decomposable, complex systems no longer have analysis and synthesis as simple inverses of each other. Building a complex system is therefore not simply the inverse of any analytic process of decomposition into parts. In other words, reverse engineering the brain – a complex, integrated and thus non-decomposable system – must necessarily fail and will not provide the envisaged understanding!

It should be stressed that simple and complex systems after Rosen's definition cannot be directly related to those sensu Simon (Sections 1.2, 1.4). While Rosen's approach yields a *descriptive* definition of complexity, Simon's is *interactional*, see [53]. It seems clear, however, that Rosen's 'simple systems' comprise Simon's full- and near-decomposable systems, and Rosen's 'complex systems' correspond to Simon's non-decomposable, integrated systems. No matter which definition is applied, the conclusion about the brain's non–decomposability remains valid.

1.7 Conclusions

If one attempts to understand a complex system like the brain it is of crucial importance if general operation principles can be formulated. Traditionally, approaches to reveal such principles follow the line of decompositional analysis as expressed in the *fundamental assumption* of cognitive and computational neuroscience, i.e. cognitive systems like other, truly complex systems are decomposable. Correspondingly, reverse engineering has been considered the appropriate methodology to understand the brain and to engineer artificial cognitive systems.

The claim was discussed that non-decomposability is not an intrinsic property of complex, integrated systems but is only in our eyes, due to insufficient mathematical techniques. For this, the scientific analysis method itself was considered. Referring to results from mathematics and system theory, I have presented arguments for the position that the dominant complexity concept of cognitive and computational neuroscience underlying reverse engineering needs revision. The updated, revised concept must comprise results from the nonlinear science of complexity, and insights expressed, e.g., in Rosen's work on life and cognition. It was concluded that the decomposability assumption of cognitive science must be abandoned.

Organisms and brains are complex, integrated systems which are non–decomposable. This insight implies that there is no 'natural' way to decompose the brain, neither structurally nor functionally. We must face the uncomfortable insight that in cognitive science and neuroscience we have conceptually, theoretically, and empirically to deal with complex, integrated systems which is much more difficult than with simple, decomposable systems of quasi–independent modules! Thus, we cannot avoid (at least in the long run) subjecting research goals such as the creation of 'brain-like intelligence' and the like to analyses which apprehend the very nature of natural complex systems.

---

[a] **Translation of Goethe's Verse by George Madison Priest**

Who'll know aught living and describe it well,
Seeks first the spirit to expel.
He then has the component parts in hand
But lacks, alas! the spirit's band.

J.W. GOETHE, Faust, First Part

References

1. Biologically-Inspired Cognitive Architectures, Proposer Information Pamphlet (PIP) for Broad Agency Announcement 05-18. DARPA Information Processing Technology Office, Arlington, VA (2005)
2. Markram, H.: The Blue Brain Project. Nature Rev. Neurosci. **7**, 153–160 (2006)
3. Albus, J.S., Bekey, G.A., Holland, J.H., Kanwisher, N.G., Krichmar, J.L., Mishkin, M., Modha, D.S., Raichle, M.E., Shepherd, G.M., Tononi, G.: A proposal for a Decade of the Mind initiative. Science. **317**, 1321 (2007)
4. Perry, W., Broers, A., El-Baz, F., Harris, W., Healy, B., Hillis, W.D., et al.: Grand challenges for engineering. National Academy of Engineering, Washington, DC (2008)
5. Systems of Neuromorphic Adaptive Plastic Scalable Electronics (SyNAPSE). DARPA / IBM (2008)
6. European Commission, ICT Call 6 of the 7th Framework Programme, Objective 2.1: Cognitive Systems and Robotics (2009)

7. Schierwagen, A.: Brain Organization and Computation. In: J. Mira and J.R. Alvarez (eds.) IWINAC 2007, Part I: Bio-inspired Modeling of Cognitive Tasks, LNCS 4527, pp. 31-40 (2007)
8. Ananthanarayanan, R., Esser, S.K., Simon, H.D., Modha, D.S.: The Cat is Out of The Bag: Cortical Simulations with 10^9 neurons and 10^{13} synapses. Supercomputing 09, Proc. ACM/IEEE SC2009 Conference on High Performance Networking and Computing, Nov 14–20, 2009, Portland, OR (2009)
9. Brodkin, J.: IBM cat brain simulation dismissed as ‘hoax’ by rival scientist. Network World, November 24 (2009)
10. Dennett, D.C.: Cognitive Science as Reverse Engineering: Several Meanings of ’Top Down’ and ’Bottom Up,’. In D. Prawitz, B. Skyrms, D. Westersthl (eds.) Logic, Methodology and Philosophy of Science IX, pp. 679–689. Elsevier Science, Amsterdam (1994)
11. Marom, S., Meir, R., Braun, E., Gal, A., Kermany, E., Eytan, D.: On the precarious path of reverse neuro-engineering. Front. Comput. Neurosci. 3 (2009) doi:10.3389/neuro.10.005.
12. Gurney, K.: Reverse engineering the vertebrate brain: methodological principles for a biologically grounded programme of cognitive modelling. Cognit. Computat. **1**, 29–41 (2009)
13. Rosen,R. : Anticipatory Systems: Philosophical, Mathematical and Methodological Foundations. Pergamon Press, Oxford (1985)
14. Rosen, R.: Life Itself: A Comprehensive Inquiry into the Nature, Origin, and Fabrication of Life. Columbia University Press, New York (1991)
15. Rosen, R.: Essays on Life Itself. Columbia University Press, New York (2000)
16. Keyser, S.J., Miller, G.A., and Walker, E.: Cognitive Science in 1978. An unpublished report submitted to the Alfred P. Sloan Foundation, New York (1978)
17. Simon. H.: The Sciences of the Artificial. MIT Press, Cambridge, MA (1969)
18. Wimsatt, W.: Forms of aggregativity. In: A. Donagan, A.N. Perovich, M.V. Wedin (eds.) Human Nature and Natural Knowledge, pp. 259-291. D. Reidel, Dordrecht (1986)
19. Bechtel, W., Richardson, R. C.: Discovering complexity: Decomposition and localization as strategies in scientific research. Princeton University Press, Princeton, NJ (1993)
20. Cummins, R.: The Nature of Psychological Explanation. MIT Press, Cambridge, MA (1983)
21. Cummins, R.: ”How Does It Work” Versus ”What Are the Laws?”: Two Conceptions of Psychological Explanation. In: F. Keil, R. A. Wilson (eds.) Explanation and Cognition, pp. 117–145. MIT Press, Cambridge, MA (2000)
22. Atkinson, A.P.: Persons, systems and subsystems: The explanatory scope of cognitive psychology. Acta Analytica **20**, 43–60 (1998)
23. Rosen, R.: The mind-brain problem and the physics of reductionism. In L. Rocha (ed.) Self-Reference in Cognitive and Biological Systems, CC-AI 12 (1-2) pp. 2944 (1995)
24. Dennett, D. C.: Consciousness explained. Little, Brown and Co, Boston (1991)
25. Price, C.J., Friston, K.J.: Functional ontologies for cognition: the systematic definition of structure and function. Cogn. Neuropsychol **22**, 262-275 (2005)
26. Uttal, W. R.: The New Phrenology. The Limits of Localizing Cognitive Processes in the Brain. Cambridge, Mass., MIT Press (2001)
27. Henson, R.: What can functional neuroimaging tell the experimental psychologist? Quart. J. Exper. Psychol. **58A**, 193–233 (2005)
28. Ross, E.D.: Cerebral Localization of Functions and the Neurology of Language: Fact versus Fiction or Is It Something Else? Neuroscientist **16**, 222–243 (2010)
29. Schierwagen, A.: Real neurons and their circuitry: Implications for brain theory. iir–reporte, pp. 17–20. Akademie der Wissenschaften der DDR, Institut f’ur Informatik und Rechentechnik), Eberswalde (1989)
30. Forrest, S.: Emergent computation : self-organizing, collective, and cooperative phenomena in natural and artificial computing networks. Physica D **42**, 1–11 (1990)
31. Levins, R.: Complex Systems. In: C. H. Waddington (ed.) Towards a Theoretical Biology, Vol. 3, pp. 73–88. University of Edinburgh Press, Edinburgh (1970)
32. Hubel, D. H., Wiesel, T. N.: Shape and arrangement of columns in cats striate cortex. J. Physiol. **165**, 559-568 (1963)

33. Mountcastle, V. B.: The columnar organization of the neocortex. Brain **120**, 701-722 (1997)
34. Szenthágothai J.: The modular architectonic principle of neural centers. Rev. Physiol. Bioche. Pharmacol. **98**, 11–61 (1983)
35. Maass, W., Markram, H.: Theory of the computational function of microcircuit dynamics. In: S. Grillner and A. M. Graybiel (eds.) The Interface between Neurons and Global Brain Function, Dahlem Workshop Report 93, pp. 371–390. MIT Press, Cambridge, MA (2006)
36. Arbib, M., Érdi, P., Szenthágothai, J.: Neural Organization: Structure, Function and Dynamics. MIT Press, Cambridge, MA (1997)
37. Rockel A.J., Hiorns R.W., Powell T.P.S.: The basic uniformity in structure of the neocortex. Brain **103**, 221-244 (1980)
38. Bressler,S.L., Tognoli,E.: Operational principles of neurocognitive networks, Intern. J. Psychophysiol. **60**, 139-148 (2006)
39. Grillner, S., Markram,H., De Schutter,E., Silberberg, G., LeBeau,F.E.N.: Microcircuits in action from CPGs to neocortex. Trends Neurosci. **28**, 525–533 (2005)
40. Horton, J. C., Adams, D. L.: The cortical column: a structure without a function. Phil. Trans. R. Soc. B **360**, 386–62 (2005)
41. Rakic, P.: Confusing cortical columns. Proc. Natl. Acad. Sci. USA **105**, 12099–12100 (2008)
42. Herculano-Housel, S., Collins, C.E., Wang, P., Kaas, J.: The basic nonuniformity of the cerebral cortex. Proc. Natl. Acad. Sci. USA **105**, 12593-12598 (2008)
43. de Garis, H., Shuo, C., Goertzel, B. and Ruiting, L.: A world survey of artificial brain projects PartI: Large-scale brain simulations, Neurocomputing, doi:10.1016/j.neucom.2010.08.004
44. Gould, S.J., Lewontin, R.C.: The spandrels of San Marco and the Panglossian paradigm: a critique of the adaptationist programme. Proc. Roy. Soc. London B **205**, 581–598 (1979)
45. Destexhe, A., Marder, E.: Plasticity in single neuron and circuit computations. Nature **431**, 789–795 (2004)
46. Frégnac, Y. et al.: Ups and downs in the genesis of cortical computation. In: S. Grillner and A.M. Graybiel (eds.) Microcircuits: The Interface between Neurons and Global Brain Function, Dahlem Workshop Report 93, pp. 397–437. MIT Press, Cambridge, MA (2006)
47. Bullmore E., Sporns O.: Complex brain networks: graph theoretical analysis of structural and functional systems, Nature Rev. Neurosci. **10**, 186–198 (2009)
48. Bechtel, W.: Dynamics and decomposition: Are they compatible? Proceedings of the Australasian Cognitive Science Society (1997)
49. Bechtel, W.: Decomposing the brain: A long term pursuit. Brain and Mind **3**, 229–242 (2002)
50. Poirier, P.: Be There, or Be Square! On the importance of being there. Semiotica **130**, 151–176 (2000)
51. van Vreeswijk, C.: What is the neural code? In: 23 Problems in System neuroscience, pp. 143–159. van Hemmen, J.L. and Sejnowski, T.Jr. (eds.) Oxford University Press, Oxford (2006)
52. Mikulecky,D.C.: Robert Rosen: the well posed question and its answer – why are organisms different from machines? Syst. Res. **17**, 419-432 (2000)
53. Wimsatt, W.C.: Complexity and Organization. Proc. Biennial Meeting Philos. Sci. Ass. **1972** 67–86 (1972)

Index