P1: FHB/SPH P2: FHB/SPH QC: FHB/SPH T1: FHB
LE061-23 LE061/Kaernbach-v1.cls April 16, 2003 16:58

IV

Neural and Representational Models

P1: FHB/SPH P2: FHB/SPH QC: FHB/SPH T1: FHB
LE061-23 LE061/Kaernbach-v1.cls April 16, 2003 16:58

Guest Editorial

Ad Aertsen
Albert-Ludwigs Universität Freiburg, Germany

A model is an attempt to capture the essence of things. Hence, as a rule, a model strives to be as simple as possible, admitting complexity only as necessary. This is a very sensible rule, in fact. Approaches that disregard it tend to produce a muddle rather than a model—an insight I owe to Valentino Braitenberg. Thus, we generally use models as simplified versions of reality. They summarize our knowledge from previous experiments, allow us to make predictions to be tested in new experiments, and, above all, they enable us to make a conceptual interpretation of our results and insights.

The brain is an enormously complex system. The total number of possible states in a network comprising some 10^{11} neurons is exceedingly large. In fact, only a vanishing fraction of them will actually occur during a lifetime. This complexity has important strategic implications, both for experimental and theoretical approaches to brain function. In such a system, we cannot expect that the underlying principles will simply pop out from merely observing the neuronal activity during an experiment. A formal theory is needed to work out testable predictions regarding the functioning of the system. These predictions, in turn, lead to the design of new experiments that can critically test the theory. At the same time, a theory of such complex and only partially observable system must incorporate the relevant biological constraints. Otherwise, it runs the danger of degenerating into a sterile formal game. Thus a close interaction of experiment and theory provides the optimal research strategy to make substantial progress in brain research.

Ideas on possible principles of brain function may well be formulated on the basis of common sense. At the level of neuronal activity, however, this approach is often misleading, or even incorrect. A serious discourse on brain function could develop only several decades ago, after researchers attempted to abstractly define and actually mimic "intelligent behavior" in terms of models and simulated these on electronic computers. This started a new process of scientific inquiry in areas such as artificial intelligence, cognitive science, and computational neuroscience that, in turn, served as a rich source of inspiration for brain theoreticians. At the same time, insights and findings from brain research provided useful ideas for the understanding of complex artificial systems as studied in engineering and computer science. Thus, in our experience, the conceptual connections among modern brain science and physics, mathematics, and computer science are the most exciting and fruitful ones.

The new field of computational neuroscience presents a good case in point. A recent special issue of the journal *Nature Neuroscience* states that "Perhaps the most exciting and difficult challenge currently facing biological scientists is to understand how complex biological systems work. These systems exist at all levels of organization, from the genetic determinants of protein structure to the complex interplay of individual neurons in orchestrating behavior." In response to this challenge, computational neuroscience was developed "to provide a solid theoretical foundation and a set of technological approaches, aimed to enhance our understanding of nervous system function by providing modeling tools that describe and transcend these many different levels of organization" (National Institute of Health, 2000, p. 1161).

The issue we are dealing with here—understanding human brain function—is quite formidable, indeed. The brain enables us to actively interact with our environment. Speed, fault tolerance, adaptivity, and creativity characterize normal brain operation and ultimately guarantee that we successfully master our daily lives. The combination of these various properties is unprecedented among current technical systems. The neural networks of the brain have to integrate a steady stream of sensory inputs with previously stored experiences. Likewise, they must produce a continuous flow of output commands to control behavior that, in turn, influences sensory perception. Many of the associated brain processes run in parallel and are distributed over multiple brain areas, giving rise to fast and well-coordinated transients of neural activity. Thus, the brain represents a complex and high-dimensional dynamical system, the function of which can be fully understood only in its behavioral context.

Ever since the times of Sherrington (1941) and Hebb (1949), neurobiologists have pursued the notion that neurons do not act in isolation, but rather that they organize into assemblies for the various computational tasks involved in organizing meaningful behavior (see also James, 1890, for an early formulation of this concept). Over the years, different definitions of “neural assembly” have been proposed, each implying different functions and properties. Some of these were phrased in terms of anatomy, others in terms of shared function or shared stimulus response (see Gerstein, Bedenbaugh, & Aertsen, 1989, for a review, with many references to the original literature). One operational definition for the cell assembly has been particularly influential: near-simultaneity or some other specific timing relation in the firing of the participating neurons. As, for instance, elaborated in the concept of the “synfire chain” (Abeles, 1991), the synaptic influence of multiple neurons converging onto others in the cortical network is much stronger if they fire in (near-)coincidence (Diesmann, Gewaltig, & Aertsen, 1999). Thus, temporal coherence or synchronous firing, postulated as a mechanism for perceptual integration (Hebb, 1949), is in fact directly available to the brain as a potential neural code (Johannesma, Aertsen, van den Boogaard, Eggermont, & Epping, 1986; Perkel & Bullock, 1968).

The notion that the functional organization of the cortex is based on interactions within and among groups of cells in large neural networks is supported by the anatomical structure and, in particular, by the massive connectivity of this part of the brain (Braitenberg & Schüz, 1991). Until recently, however, few physiological data have directly addressed the cell assembly hypothesis. Neither the study of global activity in large populations of neurons, nor the recording of single-neuron activity allows for a critical test of this concept. Rather, one seeks to observe the activities of many separate neurons simultaneously, preferably in awake, behaving animals, and to analyze these multiple single-neuron activities for possible signs of (dynamic) interactions between them. Results of such analyses are then used to draw inferencs regarding the processes taking place within and between hypothetical cell assemblies. Thus, in recent years, it has become possible to study directly various phenomena associated with neuronal assemblies. The salient result of these direct assembly observations has been that the neuronal interaction structure and, hence, the membership and internal organization of the observed assemblies depend on stimulus—and behavioral context, exhibiting systematic dynamic changes on several different time scales, with time constants down to the millisecond range. (see Aertsen, Erb, & Palm, 1994, for a review and discussion on possible mechanisms involved).

These modulations of functional inter-neuron coupling form an interesting and novel feature of cortical network organization. In particular, they are the signature of an ongoing process of dynamical and activity-related “linking” and “unlinking” of neurons into modifyible, coherent groups. This process may have interesting functional implications at different levels of observation. At the single-neuron level, it may explain how even little specificity in anatomical connections could be dynamically sorted out to yield the complex functional properties that have been observed for cortical neurons. Thus, it might provide a natural mechanism for the physiologically measured context-dependence and intrinsic dynamics of receptive fields in central sensory neurons. At the multiple-neuron level, dynamic coupling may account for coherence variations in a spatially distributed neural code. Several phenomena in neocortical activity point at possible candidates for such distributed codes. One example is the observation of stimulus-specific oscillatory events in the visual cortex, with coherence properties that may extend over wide ranges of cortex (reviewed in Singer, 1999). A second is the relative exuberance of highly accurate and behavior-related spatio—temporal spike patterns in cortical activity, pointing at the presence of "synfire reverberations" (Prut et al., 1998). Finally, at the level of the organization of perception and action, modulation of functional coupling in interconnected neural networks may provide a mechanism for the selection and successive ignition of neural assemblies within and across such networks. Spatio-temporal variation of input activity, carried onto target networks by divergent-convergent projections, could effectively modulate the activity levels in these networks (Kuhn, Rotter, & Aertsen, 2003) and, hence, provide the means to select and dynamically switch from activation of one cell assembly to the next. Such "threshold control'-like (Braitenberg, 1978; Palm, 1982) mechanisms, possibly in combination with learning by means of spike-time-dependent synaptic plasticity (Bi & Poo, 2001) have, in fact, been proposed to implement the flexible generation of Hebbian "phase sequences" (Hebb, 1949) of cell assemblies and the dynamic flow of neural information associated with them (Aertsen et al., 1994; Salinas & Sejnowski, 2001).

In summary, the highly dynamic interplay of activity and connectivity in the cortical network gives rise to an ongoing process of rapid functional reorganization. Everchanging groups of neurons, each one recruited for brief periods of time, become co-activated and again de-activated, following each other in rapid succession. It is our conjecture that this dynamic reorganization provides the neural substrate to implement the computations involved in "higher brain function," including our capacity to perceive, to behave, and to learn.

The various contributions to this section on neural and representational models are firmly rooted in these recent developments in neuroscience. This holds, in particular, for Grossberg's perspective of "the complementary brain" and its relation to neural dynamics (chap. 19, this volume). Here, the author presents a readable account of his ambitious proposal for a neural-dynamics-based implementation of the classical view that the external world is paralleled by an appropriate internal representation in the brain (e.g., Craik, 1943; McCulloch, 1965). Unlike the classical view of computer-inspired collections of independent black boxes, however, his approach makes effective use of the emerging properties of a complex dynamical system of interacting feedforward and feedback processing streams. Much in the same vein, Kompass (chap. 20, this volume) addresses the temporal structure of human perception and cognition in terms of neural network dynamics. His contribution makes a case for a discrete mental time frame with an atomic unit of approximately 4.5 ms, and relates this temporal organization with the synchronization dynamics in Abelesian synfire chain networks. Leeuwenberg, finally, takes a more abstract view in his exploration of "structual information theory" as a formal vehicle to represent the perception of visual form (chap. 21, this volume). Here, the relation to a possible neuronal implementation scheme is neither sought nor is it obvious from the mathematical formalism. Thus, the intrigued reader with a background in biology is provided with the interesting challenge of working out a biologically feasible realization of this mathematical elegance.

With these contributions in a book on *psychophysics beyond sensation* we have truly come a long way from the view that "psychophysics is the scientific discipline that studies phenomena which cease to be interesting once you clearly perceive them" (a view prevalent at the time I started out as a student in neuroscience). At the same time, the chapters you are about to read underline another important development, that of "*models beyond boxology*". Biological and mathematical sophistication have matured to the extent that we can truly hope that a moderately realistic model of brain function is within reach. Whether that model (or—more likely—a collection of models) is ultimately correct is less important than that it is inspiring and capable of organizing our theoretical and experimental endeavors. Certainly, one criterion should be firmly kept in mind: The models we develop should preferably outqualify Salman Rushdie's M2C2Ds for P2C2E ("*machines too complicated to describe*" to control "*processes too complicated to explain*") (Rushdie, 1991). The proof of the present pudding is left as an exercise to the reader.

REFERENCES

Abeles, M. (1991). *Corticonics. Neural circuits in the cerebral cortex*. Cambridge, England: Cambridge University Press.

Aertsen, A., Erb, M., & Palm, G. (1994). Dynamics of functional coupling in the cerebral cortex: an attempt at a model-based interpretation. *Physica D, 75*, 103–128.

Bi, G., & Poo, M. (2001). Synaptic modification by correlated activity: Hebb's postulate revisited. *Annual Review of Neurosciences, 24*, 139–166.

Braitenberg, V. (1978). Cell assemblies in the cerebral cortex. In R. Heim & G. Palm (Eds.), *Theoretical approaches to complex systems: Vol. 21. Lecture Notes in Biomathematics* (pp. 171–188). Berlin: Springer.

Braitenberg, V., & Schüz, A. (1991). *Anatomy of the cortex. Statistics and geometry*. Berlin: Springer.

Craik, K. (1943). *The nature of explanation*. Cambridge, England: Cambridge University Press.

Diesmann, M., Gewaltig, M.-O., & Aertsen, A. (1999). Stable propagation of synchronous spiking in cortical neural networks. *Nature (London), 402*, 529–533.

Gerstein, G. L., Bedenbaugh, P., & Aertsen, A. M. H. J. (1989). Neuronal assemblies. *IEEE Transactions on Biomedical Engineering, 36*, 4–14.

Hebb, D. (1949). *The organization of behavior. A neuropsychological theory*. New York: Wiley.

James, W. (1890). *Psychology: Briefer course*. New York: Holt.

Johannesma, P., Aertsen, A., van den Boogaard, H. Eggermont, J., & Epping, W. (1986). From synchrony to harmony: Ideas on the function of neural assemblies and on the interpretation of neural synchrony. In G. Palm & A. Aertsen (Eds.), *Brain theory* (pp. 25–47). Berlin: Springer.

Kuhn, A., Rotter, S., & Aertsen, A. (2003). Higher-order statistics of input ensembles and the response of simple model neurons. *Neural Computation, 15*, 67–101.

McCulloch, W. S. (1965). *Embodiments of mind*. Cambridge, MA: MIT Press.

National Institutes of Health (2000). Sponsors' foreword: Computational neuroscience at the NIH. *Nature Neuroscience Supplements, 3*, 1161–1164.

Palm, G. (1982). *Neural assemblies. An alternative approach to artificial intelligence*. Berlin: Springer.

Perkel, D. H., & Bullock, T. H. (1968). Neural coding. *Neuroscience Research Program Bulletin, 6*, 221–344.

Prut, Y., Vaadia, E., Bergman, H., Haalman, I., Slovin, H., & Abeles, M. (1998). Spatiotemporal structure of cortical activity: properties and behavioral relevance. *Journal of Neurophysiology, 79*, 2857–2874.

Rushdie, S. (1991). *Haroun and the sea of stories*. London: Granta Books (Penguin).

Salinas, E., & Sejnowski, T. J. (2001). Correlated neuronal activity and the flow of neural information. *Nature Reviews Neuroscience, 2*, 539–550.

Sherrington, C. (1941). *Man on his nature. The Gifford lectures, Edinburgh 1937–38*. Cambridge, UK: Cambridge University Press.

Singer, W. (1999). Neuronal synchrony: A versatile code for the definition of relations? *Neuron, 24*, 49–65.