

Reprint Series
9 September 1988, Volume 241, pp. 1299-1306

SCIENCE

Computational Neuroscience

TERRENCE J. SEJNOWSKI, CHRISTOF KOCH, AND PATRICIA S. CHURCHLAND

Computational Neuroscience

TERRENCE J. SEJNOWSKI, CHRISTOF KOCH, PATRICIA S. CHURCHLAND

The ultimate aim of computational neuroscience is to explain how electrical and chemical signals are used in the brain to represent and process information. This goal is not new, but much has changed in the last decade. More is known now about the brain because of advances in neuroscience, more computing power is available for performing realistic simulations of neural systems, and new insights are available from the study of simplifying models of large networks of neurons. Brain models are being used to connect the microscopic level accessible by molecular and cellular techniques with the systems level accessible by the study of behavior.

UNDERSTANDING THE BRAIN IS A CHALLENGE THAT IS attracting a growing number of scientists from many disciplines. Although there has been an explosion of discoveries over the last several decades concerning the structure of the brain at the cellular and molecular levels, we do not yet understand how the nervous system enables us to see and hear, to

learn skills and remember events, to plan actions and make choices. Simple reflex systems have served as useful preparations for studying the generation and modification of behavior at the cellular level (1). In mammals, however, the relation between perception and the activity of single neurons is more difficult to study because the sensory capacities assessed with psychophysical techniques are the result of activity in many neurons from many parts of the brain. In humans, the higher brain functions such as reasoning and language are even further removed from the properties of single neurons. Moreover, even relatively simple behaviors, such as stereotyped eye movements, involve complex interactions among large numbers of neurons distributed in many different brain areas (2-4).

Explaining higher functions is difficult, in part, because nervous systems have many levels of organization between the molecular and systems levels, each with its own important functions. Neurons are organized in local circuits, columns, laminae, and topographic maps for purposes that we are just beginning to understand (5-8).

T. Sejnowski is in the Department of Biophysics at the Johns Hopkins University, Baltimore, MD 21218. C. Koch is with the Computation and Neural Systems Program at the California Institute of Technology, Pasadena, CA 91125. P. S. Churchland is in the Department of Philosophy at the University of California at San Diego, La Jolla, CA 92093.

Properties not found in components of a lower level can emerge from the organization and interaction of these components at a higher level. For example, rhythmic pattern generation in some neural circuits is a property of the circuit, not of isolated pacemaker neurons (9, 10). Higher brain functions such as perception and attention may depend on temporally coherent functional units distributed through several different maps and nuclei (4, 8). The sources of such network properties are not accessible by the use of methods suited to investigating single neurons.

Assuming that there are emergent properties of networks, it is difficult to imagine a major advance in our understanding of brain function without a concomitant development in direct, efficient techniques for probing the mechanisms of distributed processing. New experimental techniques currently being developed include methods for simultaneously recording from multiple single units, optical recording of columnar organization in cortex with voltage- and ion-sensitive dyes, and large-scale measurements of brain structure and activity with positron emission tomography (PET), magnetoencephalogram (MEG), 2-deoxyglucose (2-DG), and magnetic resonance imaging (MRI) (11). Statistical methods are also being developed for analyzing and interpreting the information that will be collected (12, 13). Though valuable, each of these new techniques has severe limitations on its spatial or temporal resolution, and new approaches to understanding distributed processing are needed (14).

Modeling promises to be an important adjunct to these experimental techniques and is essential in addressing the conceptual issues that arise when one studies information-processing in the brain (15-17). The advantages of brain models are varied. (i) A model can make the consequences of a complex, nonlinear brain system with many interacting components more accessible. (ii) New phenomena may be discovered by comparing the predictions of simulation to experimental results and new experiments can be designed based on these predictions. (iii) Experiments that are difficult or even impossible to perform in living tissue, such as the selective lesion of particular channels, synapses, neurons, or pathways, can be simulated by the use of a model.

What kind of a computer is the brain? Mechanical and causal explanations of chemical and electrical signals in the brain are different from computational explanations (18). The chief difference is that a computational explanation refers to the information content of the physical signals and how they are used to accomplish a task. This difference is easiest to see in simpler physical systems that compute. For example, a mechanical explanation for the operation of a slide rule includes the observations that certain marks are lined up, the slider is moved, and a result is read. A computational explanation states that a slide rule computes products because the marks on the sliding wood rules correspond to logarithms, and adding two logarithms is equivalent to multiplying the corresponding pair of numbers. Thus, the physical system carries out a computation by virtue of its relation to a more abstract algorithm (19). One of the major research objectives of computational neuroscience is to discover the algorithms used in the brain.

Unlike a digital computer, which is general purpose and can be programmed to run any algorithm, the brain appears to be a collection of special purpose systems that are very efficient at performing their tasks, but are limited in their flexibility. The architecture of an efficient, dedicated system like a slide rule, or a brain system, constrains the algorithm it implements in a fashion that does not occur in digital computer (20). The clues from structure are particularly valuable because the nervous system is a product of evolution, not engineering design. The computational solutions evolved by nature may be unlike those that humans would invent, if only because evolutionary changes are always made within the context of a design and architecture that already is in place (21).

Classes of Brain Models

Realistic brain models. One modeling strategy consists of a very large scale simulation that tries to incorporate as much of the cellular detail as is available. We call these realistic brain models. While this approach to simulation can be very useful, the realism of the model is both a weakness and a strength. As the model is made increasingly realistic by adding more variables and more parameters, the danger is that the simulation ends up as poorly understood as the nervous system itself. Equally worrisome, since we do not yet know all the cellular details, is that there may be important features that are being inadvertently left out, thus invalidating the results. Finally, realistic simulations are highly computation-intensive. Present constraints limit simulations to tiny nervous systems or small components of more complex systems. Only recently has sufficient computer power been available to go beyond the simplest models.

An example of a realistic model at the level of a single neuron is the Hodgkin-Huxley model (22) of the action potential in the squid giant axon. The action potential is a stereotyped, transient, electrical event that propagates along an axon and is used for communicating information over long distances. The action potential is a result of the voltage- and time-dependent properties of several types of membrane channels. The dynamics of the membrane channels were modeled by a set of coupled, nonlinear differential equations that were solved numerically. The velocity of the action potential predicted by Hodgkin and Huxley agreed to within 10 percent of the measured value. Two important lessons can be learned from this example. First, the model was the culmination of a large number of experiments. In general, realistic models with a large number of parameters require a correspondingly large number of measurements to fit the parameters. Second, the voltage-dependent membrane channels postulated to account for the data were verified only much later with the introduction of single-channel recording techniques (23). In general, we should expect to make hypotheses that go beyond the immediate data.

An example of a realistic model at the network level is the Hartline-Ratliff model of the *Limulus* lateral eye (24). The photoreceptors in the retina have lateral inhibitory synaptic interactions with neighboring photoreceptors, and as a consequence the contrast in the firing rates at light intensity borders is enhanced. Because the interactions are approximately linear, a model of the network can be mathematically analyzed in great detail. Once all of the parameters are determined from experimental measurements, the model can accurately predict the firing rates of all the fibers in the optic nerve stimulated by an arbitrary spatiotemporal pattern of light falling on the retina. Models of networks that are nonlinear and have feedback connections are much more difficult to analyze, so that only a qualitative mathematical analysis is possible and simulations of their behavior are essential. It is worth emphasizing again the necessity of collecting a nearly complete set of experimental measurements before a realistic model can be attempted.

Simplifying brain models. Because even the most successful realistic brain models may fail to reveal the function of the tissue, computational neuroscience needs to develop simplifying models that capture important principles. Textbook examples in physics that admit exact solutions are typically unrealistic, but they are valuable because they illustrate physical principles. Minimal models that reproduce the essential properties of physical systems, such as phase transitions, are even more valuable. The study of simplifying models of the brain can provide a conceptual framework for isolating the basic computational problems and understanding the computational constraints that govern the design of the nervous system.

The class of models that is currently being investigated under the general headings of connectionist models, parallel distributed pro-

cessing models, and "neural networks" is of this second type, which we shall hereafter refer to as simplifying brain models. These models abstract from the complexity of individual neurons and the patterns of connectivity in exchange for analytical tractability (25). Independent of their use as brain models, they are being investigated as prototypes of new computer architectures (26, 27) and as models for psychological phenomena (28-30). Some of the lessons learned from these models can be applied to the brain.

One of the best studied models is the class of layered feed-forward networks. In this architecture, information is coded as a pattern of activity in an input layer of model neurons and is transformed by successive layers receiving converging synaptic inputs from preceding layers. The following three findings are of significance for the brain. (i) Even systems with only a few intermediate layers have enormous power in representing complex nonlinear functions (31, 32). (ii) The performance of a network in specific problem domains (such as visual and speech processing) depends critically on how the incoming information is represented by the neurons (such as the type of preprocessing) and the symmetries in the pattern of connections. (iii) For difficult problems, the processing units in the middle or "hidden" layers generally encode many different combinations of input variables by the use of a semidistributed type of representation (33). By combining the power of these models with further constraints from neurophysiology and neuroanatomy it may be possible to interpret some of the properties that have been observed from single-unit recordings, as we illustrate in the next section (34-36).

These simplifying brain models also make an important bridge to computer science and other disciplines that study information processing. Issues such as convergence of the network to a stable solution, the amount of time needed for the network to achieve a solution, and the capacity of networks to store information are being investigated in simplifying models in ways that are not at present feasible with realistic models (37). The scaling of these properties with the size of the network is crucially important for practical applications and for the plausibility of the model as a brain model (38-39). Many of the current models do not scale well without additional constraints on the architecture, such as restricting the connectivity to local neighborhoods.

Technology for brain modeling. Computational brain models are almost always simulated on digital computers. Computers are getting faster, but they must perform the many parallel operations of the brain one at a time and are many orders of magnitude too slow. Parallel computers with thousands of processors are being developed, but are still inadequate (40). A new approach toward simulating biological circuitry is being pioneered by Mead (20), who is constructing hardware devices that have components that directly mimic the circuits in the brain. Fast hardware can deliver the computing power necessary to evaluate the performance of a model in real time. Furthermore, the physical restrictions on the density of wires and the cost of communications imposed by the spatial layout of the electronic circuits are similar to the constraints imposed on biological circuits. This approach may lead to a "synthetic neurobiology" (20).

Mead uses analog subthreshold complementary metal oxide semiconductor VLSI (very large scale integrated) circuit technology. Several chips that implement simplifying models of visual information processing have already been produced that are highly efficient. A "retina" chip computes the spatial and temporal derivative of arbitrary images projected onto an hexagonal array of 48 by 48 phototransistors, which are approximately logarithmic over five orders of magnitude of light amplitude, coupled by means of a horizontal resistive grid and injecting current into model "amacrine" cells that compute a temporal derivative (41). Similar circuits can be designed for computing optical flow in real time (42-43).

These VLSI chips and new techniques in optical information processing may lead to a new computing technology, sometimes called artificial neural systems, or neurocomputing (44, 45). This technology for performing massively parallel computations could have a major influence on the next generation of research in computational neuroscience. For example, an analog VLSI model of a neuron that included conductance mechanisms, synaptic apparatus, and dendritic geometry could be produced in great quantities. These chips could be used as coprocessors in a conventional digital computer to greatly increase the speed of realistic simulations. If this technology is developed now, it should be possible to simulate our visual system in real time by the 21st century (40).

Specific Examples of Brain Models

We will discuss several different models that show the great variety of different levels of structure, analysis, and measurement existing in contemporary computational neuroscience. It is impossible to discuss in this article even a small fraction of the models in the literature that address a particular problem in neurobiology, so we will limit ourselves to a few examples from the invertebrate and vertebrate vision literature. This choice reflects the idiosyncrasies and research interests of the authors and in no way implies that other areas within neurobiology have not developed equally relevant models (46).

The modeling of learning and memory is an important area not covered here. Evidence for neural plasticity is accumulating at the cellular and molecular levels in a variety of systems (47-51). The conditions for plasticity can be incorporated onto large-scale network models that have properties that can be explored in simulations of realistic models (46, 52, 53) and analytically for simplifying models (54, 57). A separate review would be required to summarize all the interesting models of learning and memory in brain circuits that are now being explored (58-60).

Detecting and computing motion. Visual motion is a fundamental source of information about the world (61, 62), and motion detection is the first stage of motion processing. Reichardt's motion detection scheme was first proposed 30 years ago. Behavioral data, gained on the basis of open- and closed-loop experiments performed on beetles and flies (63, 64), indicated that a sequence of two light stimuli impinging on adjacent photoreceptors is the elementary event evoking an optomotor response. The relation between stimulus input to these two photoreceptors and the strength of the optomotor output follows the rule of algebraic sign multiplication (65, 66).

The correlation model of motion detection follows from these observations (63). The output of one photoreceptor is multiplied by a low-pass filtered signal from a neighboring receptor (Fig. 1A). The product is then integrated in time, which is equivalent to the autocorrelation of the visual input (67). Since the low-pass filter can be thought of as a delay, this direction-selective subunit (Fig. 1A) will respond with a positive signal in one direction and with a negative response in the opposite direction. This theoretical model has a number of nontrivial properties, such as phase invariance and dependence on contrast frequency (67), that have been confirmed experimentally. In humans, the psychophysical evidence favors a slight modification of the original correlation model (68, 69). Thus, the correlation model for motion detection is consistent with psychophysics in several different species. It is a realistic model at the systems level.

Motion detection has also been explored at the cellular level. Barlow and Levick (70) systematically studied the rabbit retina by recording extracellularly from the output cells of the retina. About

20 percent of all ganglion cells responded vigorously to the motion of both black and white spots in one direction but were silent when the spots moved in the opposite direction. Barlow and Levick's two principal conclusions were that inhibition is crucial for direction discrimination and that this discrimination must occur at many different sites within the receptive field. They proposed that direc-

tion selectivity is based on a scheme whereby the response in the null direction is vetoed or inhibited by appropriate neighboring inputs (Fig. 1B); directionality is achieved by a delay between the excitatory and inhibitory channels (extending from photoreceptors to ganglion cells). This proposal is a cellular level version of Reichardt's correlation detector.

Techniques are being developed to study motion detection at the biophysical level. The key experimental finding is that lateral inhibition is induced in ganglion cells for all directions of movement except for the preferred direction (71). This inhibition could be generated by mechanisms presynaptic to the ganglion cell (72). Werblin *et al.* (73) have found evidence in the retina of the tiger salamander for movement-gated lateral inhibition generated by interactions among amacrine and bipolar cells. This mechanism for motion detection has not yet been integrated into a model of directional selectivity. Another possible mechanism for motion detection is through postsynaptic interaction in the ganglion cell dendrite between an excitatory synapse and a delayed inhibitory synapse with a reversal potential close to the resting potential of the cell (shunting or silent inhibition), as shown in Fig. 1C (74, 76). The latter mechanism has been proven feasible with modeling studies, though a direct experimental test has not yet been performed in the retina.

Visual motion detection has many different uses. Reichardt and his group (77, 80) have studied the use of motion information for pattern discrimination in the house fly, *Musca domestica*. Pattern discrimination here is a special case of the more general problem of figure-ground segregation, where the figure is one pattern and the ground is the other. The fly can distinguish relative motion between two moving objects, even if both objects have a similar texture, such as random dot patterns. The cellular basis of this behavior was explored by use of a combination of exacting behavioral studies, electrophysiology, and systems analysis. The model circuitry (Fig. 1D) gives a satisfactory account of how the fly behaves under different conditions of relative motion (for example, a small figure oscillating on a background) and also explains the observed independence of the behavioral optomotor response from the spatial extent of motion—that is, the size of the moving object. Although the key neurons in the model (Fig. 1D) have been identified in the third optic ganglion of the fly (79) (Fig. 2), the proposed synaptic interactions are still under investigation.

What has been learned about visual processing from the fly? An engineer might have designed a motion system that first extracted pure velocity information. However, in the fly visual system, motion detection is accomplished by use of the same neurons that process local pattern information. The models of motion detection and motion processing have shown how populations of neurons with mixed pattern and motion signals can accomplish figure-ground segregation. We can begin to see how the fly uses time-varying visual information to control its behavior in ways that were not intuitively obvious. Models were used at the systems, network, cellular, and biophysical levels to generate hypotheses and help guide experimental work.

Orientation selectivity in visual cortex. Our second example is taken from the mammalian visual system. In an influential study, Hubel and Wiesel (81) showed that most cells in the cat striate cortex optimally respond to elongated bar stimuli, oriented in a specified direction. The receptive field of a simple cell is divided into several elongated subregions (Fig. 3A). Illumination of part or all of an excitatory region increased the cellular response, whereas such stimuli suppressed the response if projected onto inhibitory regions. The input to striate cortex from cells in the lateral geniculate nucleus also possesses such excitatory and inhibitory subdivisions but is organized in a concentric fashion (Fig. 3A). Hubel and Wiesel

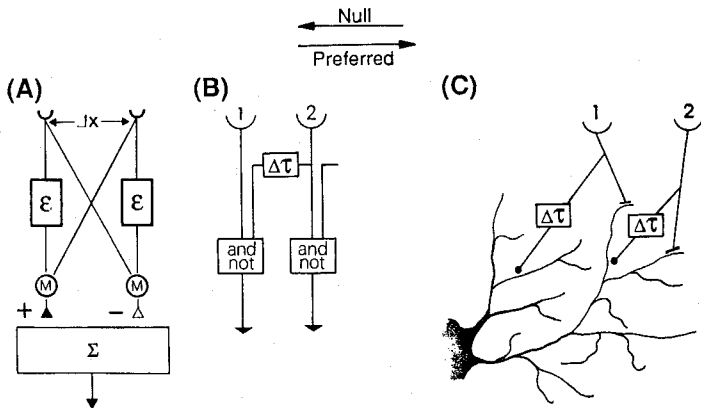


Fig. 1. Movement detection and figure-ground segregation. (A) The correlation model (63, 64). In its essential form, the two outputs from two neighboring photoreceptors are multiplied after one signal is delayed with respect to the second. If the product is averaged in time and subtracted from the time-averaged product from the other branch, the overall operation is equivalent to cross correlating the intensity on the retina. The model is symmetrical from an anatomical point of view, but is functionally antisymmetrical. The correlation model was proposed on the basis of behavioral experiments carried out in insects (65). (B) The scheme proposed by Barlow and Levick (70) to account for direction-selective ganglion cells in the rabbit retina. The output of one channel is gated in the null direction by a signal from a neighboring channel. In the preferred direction, the delayed signal from the neighboring receptor arrives too late to veto the excitatory signal. (C) A detailed cellular implementation of Barlow and Levick's scheme of direction selectivity (76), based on the proposal (75) that the nonlinear interaction between excitatory and inhibitory synaptic induced conductance changes of the silent or shunting type can approximate a multiplication. Detailed computer simulations of the cable properties of dendrites show that shunting inhibition (solid circles) can effectively veto excitation (bars) as long as the inhibitory synapses are close to the excitatory synapses or between excitation and the cell body (76). (D) The neural circuitry underlying pattern discrimination in the visual system of the house fly *Musca domestica*, as proposed by Reichardt and his group (77, 80). In each eye, two overlapping sets of retinotopic arrays of elementary movement detectors of the correlation type (63) respond selectively to either front-to-back (+) or back-to-front (-) motion. All movement detectors feed into two pool neurons (S_R and S_L). One of these cells (pr) is excited by motion in the front-to-back direction, while the second is excited by the opposite motion (► indicates excitatory synapses and ► indicates hyperpolarizing inhibitory synapses). The pool cells are also coupled with their homologues in the contralateral hemisphere. The output of both pool neurons shunts (►) the output of the movement detectors (the inhibitory synaptic reversal potential is close to the resting potential of the cell) before they either excite (►) or inhibit (►) the two output cells (X_R and X_L). The final motor output is controlled by the X cells. The key elements of this model (X and S cells) have been identified with neurons in the third optic ganglion of the fly (79, 80). Abbreviations: L, left; R, right; T, a channel that computes the running average of the output from the X cells. [Courtesy of W. Reichardt]

postulated that the orientation selectivity of simple cells was generated by a row of appropriately aligned excitatory geniculate cells converging on their cortical target neuron (Fig. 3B). The firing threshold of this cortical cell at its cell body is such that only simultaneous input from a large number of geniculate cells triggers a burst of action potentials. According to this hypothesis, the arrangement of the receptive fields of presynaptic inputs is sufficient to produce orientation selectivity (81).

Blockage of cortical inhibition by pharmacological agents leads to a partial loss of orientation selectivity in some neurons (82, 83). This observation leads to a different set of models for orientation selectivity that use intracortical inhibition (82, 84–86). In the version shown in Fig. 3C, nonoriented cortical interneurons suppress the response of the oriented simple cell to nonoptimal oriented stimuli by synaptically mediated inhibition. However, the nonoriented inhibitory interneurons postulated by this model have not been found in the cat striate cortex. Alternatively, intracortical inhibition from oriented cells could suppress firing of cells to stimuli with orthogonal orientations [cross-orientation inhibition (87)].

More recently, an “eclectic” model has been proposed that accounts for most of the experimental data by conferring orientation selectivity on all cells through massive feedback connections among cortical neurons (86, 88, 89). It assumes a Hubel-and-Wiesel-type of excitatory presynaptic arrangement superimposed on two inhibitory systems: (i) inhibition among similarly oriented cells but with

spatially nonoverlapping receptive fields and (ii) cross-orientation inhibition among cells with spatially overlapping receptive fields but differing orientation. Intracellular recordings of simple cells support the existence of at least the first two systems (90).

It is a sign of the complexity of the brain that such a seemingly simple question—what is the circuitry underlying orientation selectivity—has not been satisfactorily answered 25 years after it was first posed. Realistic models of cortical processing could be used to test the strengths and weakness of the various proposals (91). A simulation of a realistic model that includes massive inhibitory cortical interactions is presented in Fig. 4. In addition to verifying the consistency of the model with experimental measurements, the simulation also makes interesting predictions for the gain control of responses to visual stimuli of varying contrasts (91).

These models of orientation tuning for simple cells in the visual cortex provide explanations for how the response properties of neurons are generated. The question remains as to what these properties contribute to visual processing. It is believed that because simple cells respond best to bars of light, they are used to detect the

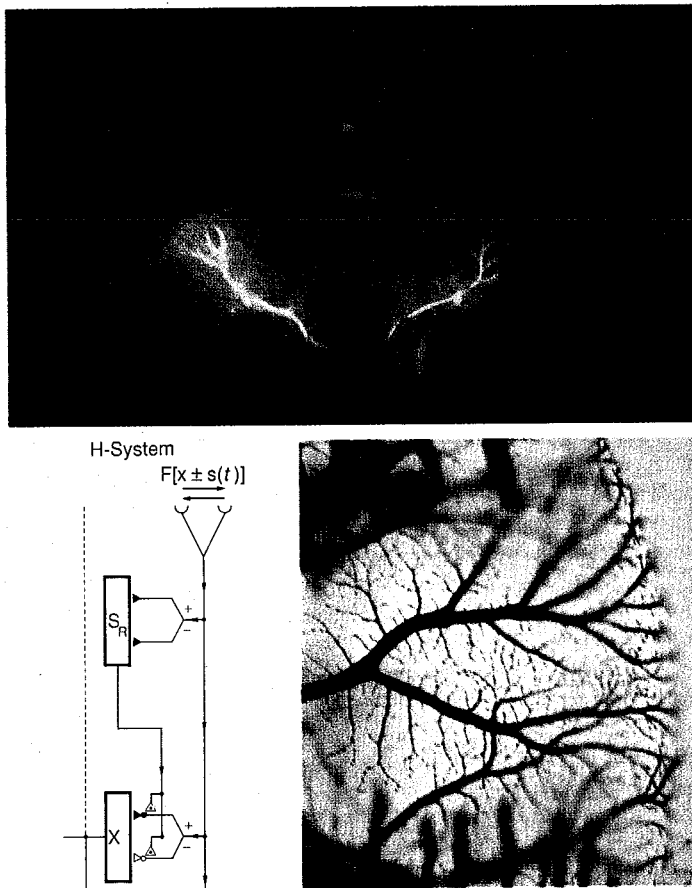


Fig. 2. Visual motion detecting system in the house fly. Two Lucifer yellow-filled horizontal cells are shown in the third optic ganglia (top) and a horizontal cell is shown at higher magnification with cobalt fill (right bottom). This cell is a possible correlate of the output x cell in the model (bottom left) that is involved in the generation of yaw torque. Additional cells controlling, for example, pitch, roll, lift, and thrust are likely to exist as well. The components of the network model are described in Fig. 1D. [Courtesy of K. Hausen and W. Reichardt].

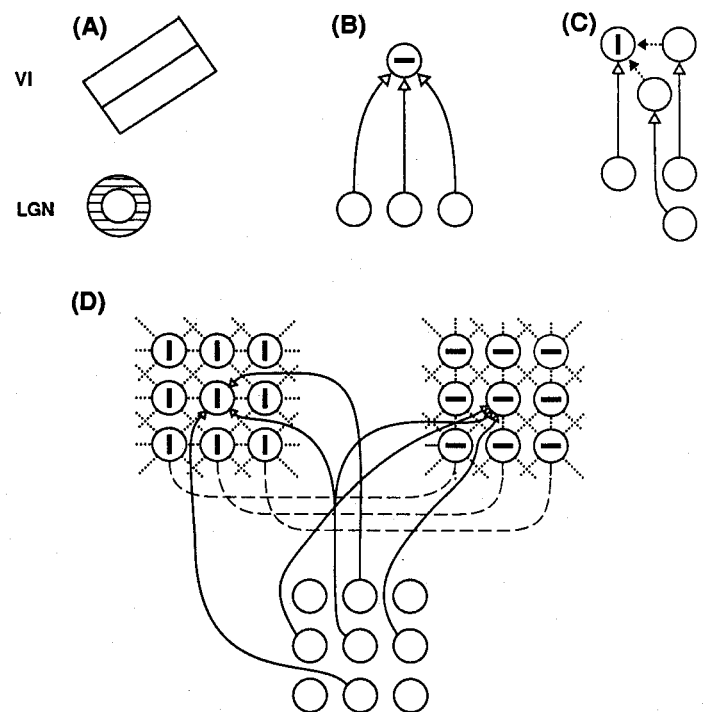


Fig. 3. Wiring underlying orientation selectivity. (A) The concentric center-surround receptive field of geniculate cells contrasts with the elongated receptive fields of their cortical target cells. (B) The excitatory-only model put forth by Hubel and Wiesel (81). A row of appropriately located geniculate cells excites the oriented cortical cell. A number of geniculate cells whose receptive fields are located along a given axis monosynaptically excite the cortical cell (open arrows). The firing threshold of the simple cell determines the degree of orientation tuning. Here and in the following, only the excitatory centers of the geniculate cells is drawn. (C) One instance of an inhibitory-only model (82, 84–86, 100). Nonoriented cortical inhibitory (filled arrows) interneurons sculpture the orientation tuning of their target cell by suppressing its response at nonoptimal orientations. For example, a horizontal bar will lead to activation of the interneuron, which will inhibit the vertically oriented cortical cell. (D) “Eclectic” model combining features of all models (88, 89). An excitatory Hubel-and-Wiesel-type of presynaptic arrangement is superimposed on two inhibitory ones: reciprocal inhibition among similarly oriented cortical cells with spatially nonoverlapping receptive fields (dotted lines) and cross-orientation inhibition among orthogonal oriented cells (dashed lines). Because of the massive feedback among the participating cortical neurons [the model can be formulated as an associative network (101)], each neuron acquires orientation selection by means of a collective computation. Abbreviations: LGN, lateral geniculate nucleus; VI, visual cortex area VI.

presence of an edge or line at a particular orientation and position in the visual field. However, the same neurons also respond to spatial frequency gratings and textured patterns, so it is not clear how to infer their actual function. The problem of inferring function from response properties applies to neurons throughout the nervous system. One way to test a functional hypothesis is to construct a network model with processing units having the same response properties as those found in the nervous system and to show that the network does perform the desired computation (92). Although a working model can help generate hypotheses, and rule some out, it cannot prove that the brain necessarily solves the problem in the same way.

Using this approach, Leaky and Sejnowski (36) have constructed a small layered feed-forward network that takes as input a small patch of shaded image, as transformed by the retina and lateral geniculate nucleus, and computes the principal curvatures and orientation of the surface from which the image was derived. The resulting network can compute surface curvatures regardless of the direction of illumination, a computational problem in which the shape of a surface is recovered from the shading in the image (93, 94). The construction of the network model was made possible by recent advances in network learning algorithms that allow complex networks with many thousands of connections to be synthesized from an input-output level description of the problem (95, 96). The network is trained by presenting images to the input layer and comparing the firing rates of the units on the output layer with the desired values. The error signals are used to correct the synaptic weights between the processing units. The learning algorithm was used solely as a technique to create the network and not as a model for learning in the brain, nor as a model for how orientation tuning actually develops. Models of development have been proposed that use unsupervised learning algorithms (53, 97).

The surprising finding was that most of the cells on the hidden layer, which received the geniculate input directly, had responses

when probed with simulated spots or bars of light that were similar to those of the simple cells in the visual cortex (Fig. 5B; see also Fig. 3A). These properties emerged through the learning procedure as the appropriate ones for extracting information about the principal curvatures, and were not put into the network directly. Curvatures were represented on the output layer by units having tuning curves for curvature magnitudes and orientations (Fig. 5A). When probed with bars of light, the output units have receptive field properties similar to complex cells in the visual cortex that are end-stopped. [The response was reduced when the end of a bar of light was extended into an inhibitory region (98).]

A finding of general significance was that the function of the model neurons in the network could not be understood solely by examination of their receptive field properties. For example, some cells with simple receptive fields were responsible not for signaling orientation to the output layer, but rather were providing information about the convexity or concavity of the surface. Examining the outputs of a simple cell, called its "projective" field by analogy with the input receptive field, was critical in uncovering its function in the network. Thus, the very same simple cells can be engaged in several functions—on one hand detecting edges, on the other hand participating in the computation of curvature from shading information—depending on the circuits to which the cell projects (36). Determining what was required for understanding the function of cells in this network model has consequences for the more difficult task of determining the function of real neurons.

This simplifying network model of shape-from-shading has generated several hypotheses about cortical processing without directly simulating the detailed connectivity of the cortex, which is not yet known. The prediction that end-stopped cells should be selectively tuned for the surface curvature can be tested by use of shaded images as stimuli along with bars of light. The model can also evolve toward a realistic model that incorporates lateral interactions between

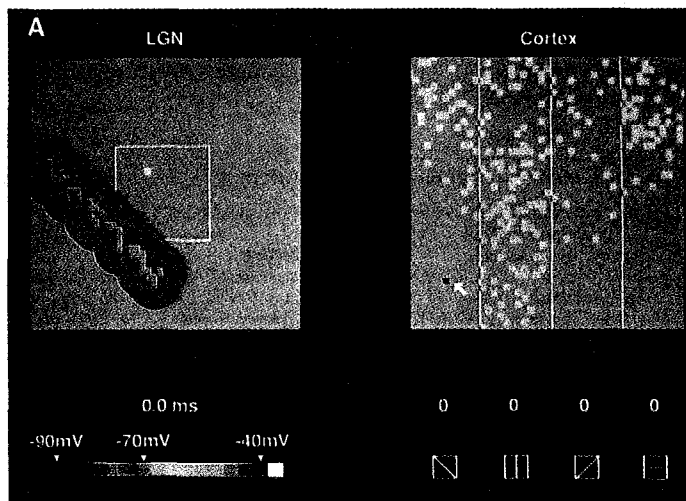
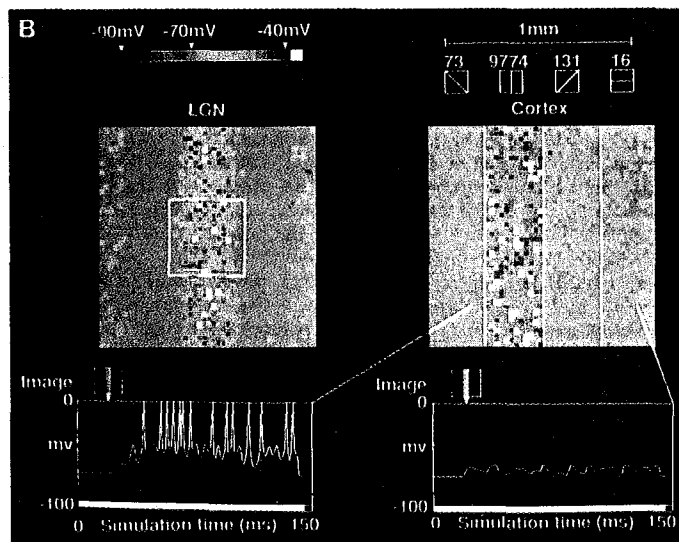


Fig. 4. Results from a realistic simulation of a $2^\circ \times 2^\circ$ patch of the early visual pathway of the adult cat. **(A)** Diagram of the divergence and convergence between the lateral geniculate nucleus (LGN) and striate cortex (cortex). A hexagonal grid of retinal X-type ganglion cells with on-center, off-surround receptive fields, modeled by differences of two Gaussian curves, project in the ratio of 1:4 to the LGN cells. A single LGN cell (small yellow square in the left central rectangle) projects to 220 cells in layer IVc of striate cortex (small yellow squares in cortex) within a disk having an area of 0.72 mm^2 , in agreement with anatomical data. A typical cortical neuron (small red square indicated by white arrow in cortex) is innervated by 40 cells in the LGN (small orange cells surrounded by circular receptive fields in blue) and has a "simple-cell"-like receptive field. The cells in cortex are arranged according to their orientation preference into vertical strips, as indicated in



the box beneath each "column". Inhibition between neighboring cortical cells in Fig. 3D is not shown. **(B)** Response of cells in the LGN and cortex to a presentation of vertical light bars to the retina, starting at 20 ms. The small squares representing neurons in the LGN and cortex are color coded by their membrane potentials at 140 ms (see calibration bar for color scheme). Simulated intracellular recordings from two cortical cells are shown in the bottom panels, one preferring a vertical orientation (left) and one preferring a horizontal orientation (right). Many cells in the vertical cortical "column" are highly excited, and these suppress the firing of cells in adjacent columns (numbers above the columns are the integrated spike counts for all of the neurons in the orientation column). Cross-orientation inhibition is of the shunting type.

circuits at neighboring patches of the visual field and interactions with circuits that use other cues for computing shape.

Conclusions

A scientific field is defined primarily by its problem space and its successful large-scale theories. Until there are such theories in computational neuroscience, the field is defined most by the problems it seeks to solve, and the general methodology and specific techniques it hopes will yield successful theories. Models of brain function driven primarily by functional considerations can provide only the most general guidance about what might be happening in the brain; conversely, models driven primarily by signal measurements and anatomy can easily miss those aspects of the signals that are relevant for information-processing. In this review we have presented several examples of models that attempt to combine insights from the functional and the implementational levels of analysis. More examples can be found in other areas of neurobiology where models are being used to explore information-processing by brain mechanisms at many different levels of structural organization (46).

Realistic and simplifying brain models have been distinguished to reveal their separate strengths and weakness. Neither type of model

should be used uncritically. Realistic models require a substantial empirical database; it is all too easy to make a complex model fit a limited subset of the data. Simplifying models are essential but are also dangerously seductive; a model can become an end in itself and lose touch with nature. Ideally these two types of models should complement each other. For example, the same mathematical tools and techniques that are developed for studying a simplifying model could well be applied to analyzing a realistic model, or even the brain itself. More accurately, the two types of models are really end points of a continuum, and any given model may have features of both. Thus, we expect future brain models to be intermediate types that combine the advantages of both realistic and simplifying models.

It may be premature to predict how computational models will develop within neuroscience over the next decade, but several general features are already emerging. First, in view of the many different structural levels of organization in the brain, and the realization that models rarely span more than two levels, we expect that many different types of models will be needed. It will be especially difficult to find a chain of models to cross the gap of at least three identified levels between the cellular and systems levels of investigation. Second, a model of an intermediate level of organization will necessarily simplify with respect to the structural properties of lower level elements, though it ought to try to incorporate as many of that level's functional properties as actually figure in the higher level's computational tasks. Thus, a model of a large network of neurons will necessarily simplify the molecular level within a single neuron.

At this stage in our understanding of the brain, it may be fruitful to concentrate on models that suggest new and promising lines of experimentation, at all levels of organization. In this spirit, a model should be considered a provisional framework for organizing possible ways of thinking about the nervous system. The model may not be able to generate a full range of predictions owing to incompleteness, some assumptions may be unrealistic simplifications, and some details may even be demonstrably wrong (99). Nevertheless, if the computational model is firmly based on the available experimental data, it can evolve along with the experimental program and help to guide future research directions.

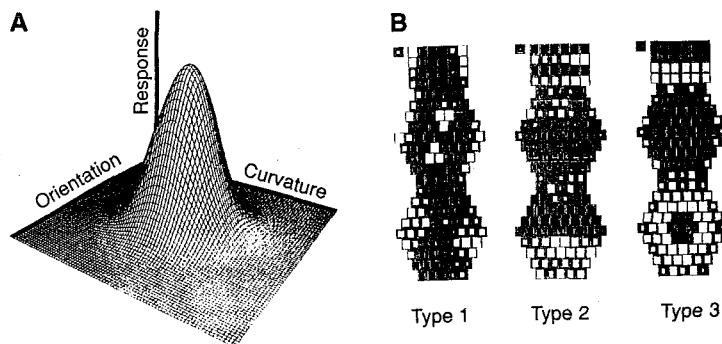


Fig. 5. Properties of model neurons in a network model that computes surface curvatures from shaded images of surfaces (36). **(A)** Representation of the tuning curve for one of the output cells in the network. Each cell is trained to respond jointly to one of the two principal curvatures and to the orientation of the long axis of the curved surface. The response of a single cell is broadly tuned and nonmonotonic. The 24 output cells have different preferred combinations of orientation and curvature with overlapping tuning curves. The pattern of responses in the population of output cells uniquely represents the curvatures and orientation. When probed with simulated bars of light these cells show some properties of end-stopped complex cells in the visual cortex. **(B)** Representative examples of three types of cells in the middle layer of a network that receive geniculate inputs and project to the output units. The hexagonal region at the bottom represents an array of inputs to the unit from on-center geniculate cells, and the hexagonal region above it are inputs from off-center geniculate cells. The white squares represent excitatory influences and the black squares represent inhibitory influences, and the area of each square is proportional to the magnitude of the synaptic strength. The patterns in the synaptic weights accurately predict the response of a cell when stimulated with simulated patterns of light and can be interpreted as the receptive field of the cell. The rectangle at the top represents the values of the weights from the cell to the 24 output cells that represent the curvature. All the cells in each column have a different preferred orientation, and each row is tuned to a different curvature, as indicated in (A). The isolated square in the upper left corner is a constant bias on the cell (negative of the threshold). Three types of cells were found in the hidden layer after training the network to produce the correct pattern on the output layer for each image. Type 1 and type 2 cells have oriented receptive fields and behave like simple cells when probed with bars of light (compare to the wiring diagram in Fig. 3A). However, the type 2 cell provides information to the output layer not about orientation, but about the convexity of the surface, as determined by pattern of projection to the cells on the output layer. The type 3 cell has a circularly symmetric receptive field and is sensitive to the ratio of the two principal curvatures.

REFERENCES AND NOTES

- See, for example, work on stretch reflexes of the spinal cord [J. C. Houk and W. Z. Rymer, in *Handbook of Physiology*, section I, *The Nervous System*, vol. 2, *Motor Control*, V. B. Brooks, Ed. (American Physiological Society, Bethesda, MD, 1981), pp. 257-323], the jamming-avoidance response in the electric fish [W. J. Heiligenberg, *J. Comp. Physiol.* **A161**, 621 (1987)], and the gill-withdrawal reflex of the sea slug [E. R. Kandel et al., in *Synaptic Function*, G. M. Edelman, W. E. Gall, W. M. Cowan, Eds. (Wiley, New York, 1987), pp. 471-518].
- S. G. Lisberger, E. J. Morris, L. Tychsen, *Annu. Rev. Neurosci.* **10**, 97 (1987).
- C. Lee, W. H. Rohrer, D. L. Sparks, *Nature* **332**, 357 (1988).
- P. S. Goldman-Rakic, *Annu. Rev. Neurosci.* **11**, 137 (1988).
- E. I. Knudsen, S. du Lac, S. D. Esterly, *ibid.* **10**, 41 (1987).
- H. B. Barlow, *Vision Res.* **26**, 81 (1986).
- E. Schwartz *ibid.* **20**, 645 (1980).
- V. B. Mountcastle, in *The Neurosciences*, G. Edelman and B. H. Smith, Eds. (MIT Press, Cambridge, MA, 1979), pp. 21-42.
- A. I. Selverston, J. P. Miller, M. Wadepuhl, *IEEE Trans. Sys. Man Cybern.* **SMC-13**, 749 (1983).
- P. A. Getting, *J. Neurophysiol.* **49**, 1017 (1983).
- Multielectrode recording: H. J. P. Reitboeck, *IEEE Trans. Sys. Man Cybern.* **SMC-13**, 676 (1983); R. R. Llinas, in *Gap Junctions*, M. E. L. Bennett and D. C. Spray, Eds. (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1985), pp. 337-353. Voltage- and ion-dependent dyes: A. Grinvald, *Annu. Rev. Neurosci.* **8**, 263 (1985); R. Y. Tsien and M. Poenic, *Trends Bio. Sci.* **11**, 450 (1986). PET: M. E. Raichle, *Trends Neurosci.* **9**, 525 (1986). MEG: S. J. Williamson, G. L. Romani, L. Kaufman, I. Modena, *Biomagnetism: An Interdisciplinary Approach* (Plenum, New York, 1983). 2-DG: L. Sokoloff, *Metabolic Probes of Central Nervous System Activity in Experimental Animals and Man* (Sinauer, Sunderland, MA, 1984). MRI: R. Bachus, E. Mueller, H. Koenig, G. Braeckle, H. Weber, in *Functional Studies Using NMR*, V. R. McCready, M. Leach, P. J. Ell, Eds. (Springer-Verlag, New York, 1987), pp. 43-60.

- New York, 1987), pp. 43-60.
12. G. L. Gerstein and A. M. H. J. Aertsen, *J. Neurophysiol.* **54**, 1513 (1985).
 13. M. Abeles, *Local Cortical Circuits* (Springer-Verlag, New York, 1982).
 14. T. J. Sejnowski and P. S. Churchland, in *Foundations of Cognitive Science*, M. I. Posner, Ed. (MIT Press, Cambridge, MA, in press).
 15. W. E. Reichardt and T. Poggio, *Theoretical Approaches in Neurobiology* (MIT Press, Cambridge, MA, 1981); F. H. C. Crick, *Sci. Am.* **241**, 219 (September 1979).
 16. M. Arbib, in *Computational Neuroscience*, E. Schwartz, Ed. (MIT Press, Cambridge, MA, in press).
 17. J. D. Cowan and D. H. Sharp, *Daedalus (Boston)* **117**, 85 (1988).
 18. Marr and Poggio [D. Marr and T. Poggio, *Science* **194**, 283 (1976); D. Marr, *Vision* (Freeman, San Francisco, CA, 1982)] have emphasized the importance of this level of analysis.
 19. P. S. Churchland, C. Koch, T. J. Sejnowski, in *Computational Neuroscience*, E. Schwartz, Ed. (MIT Press, Cambridge, MA, 1988).
 20. C. Mead, *Analog VLSI and Neural Systems* (Addison-Wesley, Reading, MA, 1988).
 21. F. Jacob, *The Possible and the Actual* (Univ. of Washington Press, Seattle, WA, 1982).
 22. A. K. Hodgkin and A. F. Huxley, *J. Physiol.* **116**, 449 (1952).
 23. A. L. Hodgkin, *ibid.* **263**, 1 (1976).
 24. F. Ratliff, *Studies on Excitation and Inhibition in the Retina* (Rockefeller Univ. Press, New York, 1974).
 25. J. J. Hopfield and D. W. Tank, *Science* **233**, 625 (1986).
 26. J. A. Feldman and D. H. Ballard, *Cognitive Sci.* **6**, 205 (1982).
 27. T. Kohonen, *Neural Networks* **1**, 3 (1988).
 28. G. E. Hinton and J. A. Anderson, *Parallel Models of Associative Memory* (Erlbaum, Hillsdale, NJ, 1981).
 29. D. E. Rumelhart and J. L. McClelland, *Parallel Distributed Processing: Explorations in the Microstructure of Cognition* (MIT Press, Cambridge, MA, 1986).
 30. S. Grossberg and G. Carpenter, *Neural Networks* **1**, 17 (1988).
 31. R. P. Lippmann, *IEEE Acous. Speech Signal Process. Mag.* **4**, 4 (1987).
 32. A. Lapedes and R. Farber, in *Neural Information Processing Systems*, D. V. Anderson, Ed. (American Institute of Physics, New York, 1988), pp. 442-456.
 33. T. J. Sejnowski and C. R. Rosenberg, *Complex Sys.* **1**, 145 (1987).
 34. D. Zipser and R. Andersen, *Nature* **331**, 679 (1988).
 35. A. C. Hurlbert and T. A. Poggio, *Science* **239**, 482 (1988).
 36. S. R. Lehky and T. J. Sejnowski, *Nature* **333**, 452 (1988).
 37. G. J. Mitchison and R. M. Durbin, *King's College Res. Center, Cambridge, U.K., Tech. Rep.* (1988).
 38. Y. S. Abu-Mostafa, *IEEE Trans. Inf. Theory* **IT-32**, 513 (1986).
 39. G. Tesauro and B. Janssens, *Complex Sys.* **2**, 39 (1988).
 40. T. J. Sejnowski, *J. Math. Psychol.* **31**, 203 (1987).
 41. M. Sivilotti, M. Mahowald, C. Mead, in *Stanford VLSI Conference*, P. Losleben, Ed. (MIT Press, Cambridge, MA, 1987), pp. 295-312.
 42. J. Tanner and C. Mead, in *VLSI Signal Processing II*, S. Y. Kung, R. Owen, G. Nash, Eds. (IEEE Press, New York, 1986), pp. 59-76.
 43. J. Hutchinson, C. Koch, J. Luo, C. Mead, *Computer*, **21**, 52 (March 1988).
 44. Y. S. Abu-Mostafa and D. Psaltis, *Sci. Am.* **256**, 88 (March 1987).
 45. Special issue on artificial neural systems, B. D. Shriver, *Computer* (March 1988).
 46. Models in motor control and vision: F. Crick, *Proc. Natl. Acad. Sci. U.S.A.* **81**, 4586 (1984); E. C. Hildreth and J. M. Holerbach, in *Handbook of Physiology*, F. Plum, Ed. (American Physiological Society, Bethesda, MD, 1987); S. Ullman, *Annu. Rev. Neurosci.* **9**, 1 (1986). Models of olfaction: W. J. Freeman, *Mass Action in the Nervous System* (Academic Press, New York, 1975); G. Lynch, in *Perspectives in Memory Research and Training*, M. S. Gazzaniga, Ed. (MIT Press, Cambridge, MA, in press); M. Wilson and J. M. Bower, in *Neural Information Processing Systems*, D. V. Anderson, Ed. (American Institute of Physics, New York, 1988), pp. 114-126. Models of sensory-motor integration: D. A. Robinson, *Annu. Rev. Neurosci.* **4**, 463 (1981); M. A. Arbib, *Behav. Brain Sci.* **10**, 407 (1987); A. Pellionisz and R. Llinas, *Neuroscience* **16**, 245 (1985); S. Grossberg and M. Kuperstein, *Neural Dynamics of Adaptive Sensory-Motor Control* (North-Holland, Amsterdam, 1986).
 47. E. R. Kandel et al., in *Synaptic Function*, G. M. Edelman, W. E. Gall, W. M. Cowan, Eds. (Wiley, New York, 1987), pp. 471-518.
 48. D. L. Alkon, *Memory Traces in the Brain* (Cambridge Univ. Press, New York, 1987).
 49. R. F. Thompson, *Science* **233**, 941 (1986).
 50. T. H. Brown, A. H. Ganong, E. W. Kariss, C. L. Keenan, S. R. Kelso, in *Neural Models of Plasticity*, J. H. Byrne and W. O. Berry, Eds. (Academic Press, New York, 1988).
 51. M. M. Merzenich, and J. H. Kaas, *Trends Neurosci.* **5**, 434 (1982).
 52. R. D. Traub, R. Miles, R. K. S. Wong, *J. Neurophysiol.* **58**, 752 (1987).
 53. L. H. Finkel and G. M. Edelman, in *Synaptic Function*, G. M. Edelman, W. E. Gall, W. M. Cowan, Eds. (Wiley, New York, 1987), pp. 711-757; J. C. Pearson, L. H. Finkel, G. M. Edelman, *J. Neurosci.* **7**, 4209 (1987).
 54. J. J. Hopfield, *Proc. Natl. Acad. Sci. U.S.A.* **79**, 2554 (1982).
 55. T. Kohonen, *Self-Organization and Associative Memory* (Springer-Verlag, New York, 1984).
 56. G. Palm, *Neural Assemblies: An Alternative Approach to Artificial Intelligence* (Springer-Verlag, New York, 1982).
 57. S. Amari, *IEEE Trans. Sys. Man Cybern.* **SMC-13**, 741 (1983).
 58. T. J. Sejnowski and G. Tesauro, in *Brain Organization and Memory*, J. L. McGaugh, N. M. Weinberger, G. Lynch, Eds. (Oxford Univ. Press, New York, in press).
 59. B. L. McNaughton and R. G. Morris, *Trends Neurosci.* **10**, 408 (1987).
 60. N. H. Donegan, M. A. Gluck, R. F. Thompson, in *Psychology of Learning and Motivation*, G. H. Bower, Ed. (Academic Press, San Diego, CA, in press).
 61. E. Hildreth and C. Koch, *Annu. Rev. Neurosci.* **10**, 477 (1987).
 62. K. Nakayama, *Vision Res.* **25**, 625 (1985).
 63. B. Hassenstein and W. Reichardt, *Z. Naturforsch.* **11b**, 513 (1956).
 64. W. Reichardt, *ibid.* **12b**, 448 (1957).
 65. D. Varju and W. Reichardt, *ibid.* **22b**, 1343 (1967).
 66. A fly eye contains about 24,000 visual receptors [N. Franceschini, *Neurosci. Res. Suppl.* **2**, S17 (1985)], and four layers of neurons between the photoreceptors at the input level and motoneurons at the output level [K. Hausen and R. Hengstenberg, *Annu. Meet. Soc. Neurosci. Abstr.* **17**, 295 (1987)].
 67. W. Reichardt, in *Sensory Communication*, W. A. Rosenbluth, Ed. (MIT Press, Cambridge, MA, 1961), pp. 303-318.
 68. J. P. H. van Santen, and G. Sperling, *J. Opt. Soc. Am.* **A1**, 451 (1984).
 69. K. Gotz, *Kybernetik* **2**, 215 (1964).
 70. H. B. Barlow and W. R. Levick, *J. Physiol.* **178**, 477 (1965).
 71. H. J. Wyatt and N. W. Daw, *Science* **191**, 204 (1976).
 72. N. M. Grzywacz and C. Koch, *Synapse* **1**, 417 (1987).
 73. F. Werblin, G. Maguire, P. Lukasiwicz, S. Eliasof, S. Wu, *Neurobio. Group Univ. CA. Berkeley, Tech. Rep.* (1988).
 74. W. Rall, in *Neuronal Theory and Modeling*, R. F. Reiss, Ed. (Stanford Univ. Press, Palo Alto, CA, 1964), pp. 73-97.
 75. V. Torre and T. Poggio, *Proc. R. Soc. London Ser. B* **202**, 409 (1978).
 76. C. Koch, T. Poggio, V. Torre, *Philos. Trans. R. Soc. London* **298**, 227 (1982).
 77. W. Reichardt and T. Poggio, *Biol. Cybern.* **35**, 81 (1979).
 78. ———, K. Hausen, *Biol. Cybern. Suppl.* **46**, 1 (1983).
 79. M. Egelhaaf, *Biol. Cybern.* **52**, 195 (1985).
 80. W. Reichardt and A. Guo, *ibid.* **53**, 285 (1986); W. Reichardt, R. W. Shloegel, M. Egelhaaf, *Naturwissenschaften*, **75**, 313 (1988); M. Egelhaaf, K. Hausen, W. Reichardt, C. Wehrhahan, *Trends Neurosci.* **11**, 351 (1988).
 81. D. H. Hubel and T. N. Wiesel, *J. Physiol. (London)* **160**, 106 (1962).
 82. A. M. Sillito, *ibid.* **250**, 305 (1975).
 83. A. M. Sillito et al., *Brain Res.* **194**, 517 (1980).
 84. L. A. Bienenstock, O. D. Creutzfeldt, H. Kuhnt, *Nature New Biol.* **238**, 124 (1972).
 85. P. Heggelund, *Exp. Brain Res.* **42**, 89 (1981).
 86. G. A. Orban, *Neuronal Operations in the Visual Cortex* (Springer-Verlag, New York, 1984).
 87. N. C. Morrone, D. C. Burr, L. Massi, *Proc. R. Soc. London Ser. B* **216**, 335 (1982); V. Braitenberg and C. Braitenberg, *Biol. Cybern.* **33**, 179 (1979).
 88. C. Koch, *Ophthalmol. Vis. Sci.* **28**, (suppl. 3), 126 (1987).
 89. D. Ferster and C. Koch, *Trends Neurosci.* **10**, 487 (1987).
 90. D. Ferster, *J. Neurosci.* **6**, 1284 (1986).
 91. U. Wehmeier, D. Dong, C. Koch, in *Methods in Neuronal Modeling*, C. Koch and I. Segev, Eds. (MIT Press, Cambridge, MA, in press).
 92. T. J. Sejnowski, in (29), vol. 2, pp. 372-389.
 93. ———, P. K. Kienker, G. E. Hinton, *Physica* **22D**, 260 (1986).
 94. D. E. Rumelhart, G. E. Hinton, R. J. Williams, *Nature* **323**, 533 (1986).
 95. K. Ikeuchi and B. K. P. Horn, *Artif. Intell.* **17**, 141 (1981).
 96. A. P. Pentland, *IEEE Trans. Pattern Anal. Mach. Intell.* **6**, 170 (1984).
 97. For models of cortical plasticity see M. F. Bear, L. N. Cooper, F. F. Ebner, *Science* **237**, 42 (1987); R. Linsker, *Proc. Natl. Acad. Sci. U.S.A.* **83**, 8779 (1986); C. von der Malsburg and D. Willshaw, *Trends Neurosci.* **4**, 80 (1981).
 98. Network models can also be constructed that extract one-dimensional curvature of lines in a plane [A. Dobbins, S. W. Zucker, M. S. Cynader, *Nature* **329**, 438 (1987); H. R. Wilson and W. A. Richards, *J. Opt. Soc. Am.*, in press].
 99. An example of an imperfect, but useful model from physics is the Bohr model of the hydrogen atom. This model served an essential role as a bridge between classical physics and quantum physics.
 100. P. O. Bishop, J. S. Coombs, G. H. Henry, *J. Physiol. (London)* **219**, 659 (1971).
 101. J. J. Hopfield, *Proc. Natl. Acad. Sci. U.S.A.* **81**, 3088 (1984).
 102. We are grateful to R. Cone, F. Crick, T. Poggio, G. Tesauro, and W. Reichardt for helpful comments on this article. T.J.S. is supported by an NSF Presidential Young Investigator Award and grants from the Sloan Foundation, Seaver Institute, the Lounsbery Foundation, the Air Force Office of Scientific Research, the Office of Naval Research, and the General Electric Corporation; C.K. is supported by an NSF Presidential Young Investigator Award and grants from the Office of Naval Research, the Advanced Engineering program at the National Foundation, the Sloan Foundation, and by the Powell Lee Foundation; P.S.C. is supported by a grant from the NSF.