

PERCEPTUAL NEURAL ORGANIZATION: SOME APPROACHES BASED ON NETWORK MODELS AND INFORMATION THEORY

Ralph Linsker

IBM Research Division, T. J. Watson Research Center,
Yorktown Heights, New York 10598

INTRODUCTION

To understand neural processing we need to study structure and function at many levels of organization, from subcellular to systemic. We also need to understand the linkages between levels. First, what are the mechanisms at a lower level that generate structures at a higher one? Second, of all the possible structures that could be formed from the given constituents, only some are in fact generated by the lower-level mechanisms. Are the generated structures optimal, or favored over the other structures, with respect to some property? If so, we may be able to describe the lower-level mechanisms as implementing an organizing, or optimization, principle. Third, can we account for such putative organizing principles in terms of their adaptive value to the animal?

This review explores linkages between lower-level mechanisms and functional architecture in the processing of sensory information. It brings together two lines of study. The first of these is the investigation of how lower-level mechanisms can generate the types of neural structures that are found in the early processing stages of perceptual systems. This approach involves modeling the formation and modification of neuronal connections by simple rules (for example of Hebb type), expressing these rules as a mathematical procedure or algorithm, and using computer simulations or mathematical analysis to determine what structures the rules generate. The

types of structures or patterns whose formation has been studied include topographic maps, orientation-selective receptive fields, and ocular dominance and orientation columns. The second line of study consists of a set of general ideas about how data may be encoded and transformed in a perceptual system, and the informational purposes that these transformations may serve. I review both approaches, then discuss recent results that suggest how the approaches may be unified—that is, how lower-level mechanisms may be used to create perceptual processing stages that implement certain types of optimal encoding principles.

FROM SALIENT EXPERIMENTAL FEATURES TO MODELS OF SELF-ORGANIZATION

How do specific patterns of neural connectivity and functional architecture develop, how may they be plastically altered, and how do these patterns subserve perceptual functions? A great deal of experimental progress has been made concerning these issues. Theoretical work in this area has several purposes: to seek common rules and principles that may account for a range of observations, to predict new features of neural organization and cell response, and to provide a view of biological information processing that integrates several levels of organization.

Experimental evidence indeed suggests that common rules and principles may underlie important aspects of sensory processing. First, cortical regions subserving different processing functions share similar intrinsic structure (Mountcastle 1978). Second, by altering the character of the input to a sensory processing region, one can induce patterns of organization and response properties that differ from those normally found in that region, but in an apparently lawful way (e.g. Constantine-Paton & Law 1978, Kaas et al 1983, Merzenich et al 1984, Métin & Frost 1989, Rauschecker 1987, Sur et al 1988).

The biological systems of interest exhibit immense complexity. The models to be discussed are by comparison extremely simple. The purpose of this simplicity is to allow us to gain understanding of how underlying rules can generate structure and function, so that essential complexities can then be added in an insightful way.

Types of Pattern-Generating Models

When a structure or organized pattern is found in nature, various types of patterning models may account for it. Two extreme types are patterning by “explicit specification” and patterning by a process of so-called “self-organization.” In the first type of model, the pattern is directly determined or strongly influenced by a pre-existing pattern in an underlying substrate. An example of this type is Sperry’s (1943, 1963) chemoaffinity model of

topographic map formation, in which each cell has a specific chemical "address." When such a model is used, the problem of explaining the original pattern's emergence is replaced by the problem of explaining how the underlying pattern of "addresses" itself arises from lower-level rules. In some models, an explicit specification may determine the pattern in detail; in others, it may determine only certain features of the pattern, such as its overall orientation or its coarse-grained arrangement.

In a model of self-organization, on the other hand, the pattern develops from an initially homogeneous structure as a result of processes that incrementally change each element of the system according to a relatively simple set of rules (Turing 1952). Typically such rules are local; that is, each incremental change in one element depends only upon the state of a few other elements. In the generation of certain properties such as topographic maps, both self-organization and a partial form of explicit specification appear to play a role (Udin & Fawcett 1988). This review focuses mainly on the self-organizing aspects of patterning models.

Basic Structure of the Models

The models discussed here can all be understood with reference to a common basic structure and set of patterning rules, although the details vary and not all of the components are present in each model. The structure consists of a "source" and a "target" layer of cells, with feedforward connections from source to target and lateral connections between target cells. Each connection is characterized by a number called its "strength." The patterning process consists of repeatedly modifying the connection strengths (and in some cases creating or destroying connections) according to the rules until a final configuration develops.

Each connection strength is modified in a way that depends upon the "state" of the connected cells. In "marker-based" models the state of a cell is defined as the amount of a marker substance of some type (there may be more than one type of marker). In "activity-dependent" models the state is defined as a measure of neuronal signaling activity such as a firing rate.

The patterning rules consist of three parts: (a) a "transmission" rule that determines how each target cell's state depends upon the states of the source cells connected to it (the target cell's state is typically related to an average of the source cells' states weighted by their connection strengths); (b) a "lateral interaction" rule that describes how the states of nearby target cells are modified by interactions within the target layer; (c) an "update" rule that modifies each feedforward (and in some cases lateral) connection according to the degree to which the states of the connected cells are similar.

In activity-dependent models, the update rule typically changes each connection strength in a way that depends upon the presynaptic signaling activity and the postsynaptic firing rate or depolarization potential. In this case, the change usually depends upon the degree of correlation between the pre- and postsynaptic quantities, with stronger correlation causing an increase in strength. I refer to this type of neural activity-dependent rule as “Hebb-like” (see Brown et al 1990 for review), while recognizing that Hebb’s original proposal (Hebb 1949) referred to cell firing, not depolarization, and that it provided no statement of the conditions under which strengths could decrease (cf. Stent 1973).

Note that both Hebb-like rules and marker-based rules play similar roles in the patterning process, although the mechanisms to which they refer are very different. In each case a connection is strengthened when the pre- and postsynaptic cells are correlated with each other—either in their signaling activity or in the possession of similar amounts of a marker.

It is striking that many salient features of perceptual neural organization emerge in models containing the basic elements described above: a *positive feedback* process (whether marker-based or Hebb-like) in which large connection strength causes the states of the connected source and target cells to be more similar, and greater similarity tends further to increase the connection strength; and a *lateral interaction* process that causes the states of nearby target cells to tend to become either more or less similar to each other, depending upon the interaction.

Models of neural self-organization developed since the early 1970s have emphasized different aspects of these rules and invoked different constraints or assumptions concerning connectivity and input and output signaling activity. They have been directed toward various goals such as elucidating basic principles of pattern formation, showing how specific features emerge, or modeling specific underlying mechanisms in greater detail. The remainder of this section reviews some key ways in which these models have increased our understanding of how simple rules lead to complex structures that resemble those found biologically.

Self-Organizing Models of Topographic Map Formation

In a topographic map, source and target cells are connected in such a way that positional ordering is preserved. Experimental results on topographic map formation have been recently reviewed (Udin & Fawcett 1988). How are the main features of topographic maps generated in self-organizing models?

NEIGHBOR-PRESERVING MAPS For marker-based models, the essential idea

is described by the so-called "tea trade analogy" (von der Malsburg & Willshaw 1977, Willshaw & von der Malsburg 1979). In this model, one assumes that the states of nearby source cells are similar. Each target cell acquires a state that is the average of the source cells' states weighted by their connection strengths. The lateral interaction then "blends" the states of nearby target cells so that these states become more similar to each other. The update rule strengthens connections between source and target cells having similar states. The result of this process is that nearby source cells map to nearby target cells.

How can a topographic map arise in an activity-dependent model (Willshaw & von der Malsburg 1976)? Start with initially random non-negative feedforward connection strengths. (For simplicity, all pairs of source and target cells can be assumed to be connected. A strength that declines to and remains at zero corresponds to an absent connection in a more realistic model.) Many patterns of signaling activity in the source layer are presented to the network, one at a time. Each input presentation consists simply of a localized region, or "spot," of activity against a quiet background. If certain constraints (specific to the model) are met, each such presentation will cause a localized group of target cells to be active. (In particular, the lateral interactions are set up so that the active target cells tend to form a localized region, rather than being dispersed.) The connections between pairs of active source and target cells are then strengthened by a small amount (this is the Hebb-like rule), and compensating reductions are made in the strengths of other connections. After many such spot patterns are presented, the resulting pattern of connection strengths will be topographic; that is, nearby source cells will make their strongest connections to nearby target cells.

Kohonen (1982a,b) has described a related algorithm in which a simple geometric computation is substituted for the more detailed properties of the Hebb-like rule and lateral interactions, and a topographic map emerges. For each presentation of an input activity "spot," this algorithm finds the particular target cell that fires most strongly, and then changes the target cells' response properties so that the maximally responding target cell and its neighbors will respond more strongly to a spot at the same input location in the future. The algorithm also applies more generally to cases in which the input patterns are ordered according to properties other than spatial position.

In both marker-based and activity-dependent models, the degree of selectivity of a target cell—that is, the range of different input patterns to which the cell responds strongly—can depend upon (a) the update rule (e.g. Bienenstock et al 1982); (b) constraints on total connection strength, or on some other function of the strengths, for each source or target cell

(Willshaw & von der Malsburg 1976); and (c) the particular form of lateral interaction (Grossberg 1976).

Let us consider several important experimental and theoretical questions concerning topographic maps that go beyond the basic issue, of mapping neighbors to neighbors, discussed above.

OVERALL MAP ORIENTATION The self-organizing model rules do not by themselves favor one map orientation over another. In biological systems, temporal or chemical mechanisms that favor certain regions to be connected first, or preferentially, may induce an orientation bias. This roughly corresponds to weakly biasing either the initial connection strengths (Willshaw & von der Malsburg 1976, 1979) or the patterning rules (White-law & Cowan 1981, Fraser 1985). Also, the shapes of the source and target layers may favor a particular orientation, owing to boundary effects. More generally, source and target layer boundary conditions can influence map formation in significant ways that are not limited to overall orientation (Schwartz 1977). Models having such biases can be thought of as hybrids of self-organizing and explicit-specification models. Experimental manipulations (reviewed by Udin & Fawcett 1988) are important for determining the extent to which the biological mechanisms fit either type of model.

LOCALLY OPTIMAL MAPS THAT ARE NOT GLOBALLY OPTIMAL If separate regions of the network independently become organized, the resulting maps may be out of register along their boundaries or may have conflicting orientations. To avoid this, one can start with a coarse topographic map that is subsequently refined by a local process. As an example, a two-step process in which the refinement stage is activity-dependent is found during map regeneration in goldfish (Schmidt 1985). Alternatively, the lateral interaction distance in the target layer can be made initially large, then decreased during development (Kohonen 1982a,b). In a biological system, growth of the target layer relative to the lateral interaction distance could achieve a similar result. Yet another approach is to limit early map formation to a single region of nucleation (Willshaw & von der Malsburg 1976).

CONTINUOUS VS. DISCRETE MAPS Some self-organizing models generate topographic maps in which the receptive fields of a sequence of target cells either (a) shift in a continuous and overlapping manner, as described above, or (b) form clusters with discontinuous jumps as one crosses from one column-like target region to another (Takeuchi & Amari 1979). Discontinuous mappings also can emerge in models with more complex dynamics (Pearson et al 1987).

MAP MAGNIFICATION FACTORS In certain models (Amari 1980, Kohonen

1982a,b, Pearson et al 1987, Linsker 1988b), if one region of the source layer is stimulated by an activity "spot" more often than another, that source region is mapped onto a larger region of the target layer. From one theoretical standpoint (see A PRINCIPLE OF MAXIMUM PRESERVATION below) it can be desirable for the magnification factor of each source region to be proportional to the frequency of "spot" stimulation, so that each target region of given size is activated an equal fraction of the time. Although the Kohonen algorithm described above was originally thought to have this property (Kohonen 1982a,b), it has since been shown not to (Ritter & Schulten 1986, Kohonen 1988). A model that does generate maps having this proportionality property is described by Linsker (1989b), and has been generalized to cases in which the activity patterns are more complex and the magnification factor does not depend simply upon frequency of stimulation.

Ocular Dominance

Banded regions of ocular dominance are found in cat and monkey striate cortex (LeVay et al 1975, Hubel & Wiesel 1977, Shatz & Stryker 1978) and in frog tectum following implantation of a third eye (Constantine-Paton & Law 1978). Their formation depends upon correlated electrical activity (Stryker & Harris 1986, Reiter & Stryker 1988).

Two questions addressed by self-organizing models are:

1. How does the process of segregation into ocularity domains interact with the process of topographic map formation?
2. Why do patterns ranging from regular stripes to less regular bands or patches form, and what determines their characteristic dimensions?

The interaction between topographic map and ocularity domain formation is studied in a marker-based model (von der Malsburg 1979) having three classes of markers: one for ocularity and one each for horizontal and vertical position. The multiple-marker model is an extension of topographic map models to a case in which similarities in properties other than position affect the mapping. The overall "similarity" between the marker content of two cells is defined as a combined measure of the similarities of the markers in each class. The overall effect is to map source to target in such a way that the similarity between nearby target cells is maximized on average. This type of mapping criterion is either explicitly introduced or emergent in other self-organizing models, both marker-based and activity-dependent, as well (see for example Kohonen 1982a,b, 1988, Linsker 1988b, Durbin & Mitchison 1988).

Models that incorporate a more detailed interaction between topog-

raphy and ocularity in map formation include an activity-dependent model (von der Malsburg & Willshaw 1976) in which same-eye input activities are locally correlated and opposite-eye activities are anti-correlated, and a model (Fraser 1985) involving markers that mediate an adhesive interaction, activity-dependent modification, and global positional biasing effects.

Why do ocular dominance stripes and related patterns form? Suppose we ignore visual field position, and simply distinguish between inputs from the two eyes. Each target cell receives connections of initially random strength from each eye. Now let the target cells interact, and their connection strengths change incrementally, according to the following assumed rule (Swindale 1980): Each cell's ocularity preference changes to be more nearly like the average preference of its near neighbors, and to be more nearly opposite the average preference of its midrange neighbors. The result is that regions of each ocularity preference form and segregate into (a) locally parallel stripes (with forks and bends) if the lateral interaction rules are unbiased between the two eyes, or (b) islands of one ocularity in a "sea" of the other if the rules favor the latter ocularity to a sufficient degree. Stripes tend to run into the layer boundary at a right angle, consistent with the observed perpendicularity of stripes at the area 17/18 border in, for example, macaque (Hubel & Wiesel 1977). In addition to these simulation results, mathematical analysis of the onset of segregation shows that as stripes start to form, their width has a particular preferred value that depends upon the scale of the lateral interaction (Swindale 1980). Stripe formation has also been analyzed and simulated in a more detailed activity-dependent model in which each target cell receives inputs from source cells lying in a fixed topographic arrangement (Miller et al 1986, 1989).

Orientation-Selective Cells and Columns

Cells selectively responsive to edges or bars having a particular orientation, and arranged in columns containing cells of similar orientation preference, were discovered by Hubel & Wiesel in cat (Hubel & Wiesel 1962, 1963, reviewed by Frégnac & Imbert 1984) and macaque (Hubel & Wiesel 1968, 1977). In macaque, they are present at birth (Wiesel & Hubel 1974) prior to structured visual experience. Recent advances in optical imaging of electrical activity (Blasdel & Salama 1986) have provided a more detailed picture of the columnar structure in macaque. The detailed mechanisms that mediate orientation-selective response are still unclear (see review by Ferster & Koch 1987).

The early experimental findings motivated attempts to explain how orientation selectivity could be generated by activity-dependent self-organ-

izing models. To see what has been learned from these models, it is useful to think of the problem in two parts.

1. When the input patterns presented to the network (during the connection modification process) differ from one another with respect to only one property, how does each target cell learn to respond selectively to patterns having a particular value (or range of values) of that property? That is, how does selectivity per se emerge when there is only one pattern selection criterion?
2. When the input patterns differ from one another in many respects, what determines the particular property, or combination of properties, to which each target cell will learn to respond selectively? In particular, how can orientation selectivity emerge in self-organizing models either in the absence of structured visual experience, corresponding to the prenatal development of well-formed selectivity in macaque, or under the influence of realistic postnatal experience?

HOW DOES SELECTIVITY EMERGE? In the activity-dependent models of topographic map formation discussed above, topographic ordering arises when each input activity pattern consists of a localized spot. Two input activity patterns are similar to each other if their spots are at similar positions. The update rule, and in some models the lateral interactions, cause each target cell to become selective for a subset of similar patterns, and the lateral interactions cause nearby target cells to have similar response properties.

In the first self-organizing models to address orientation selectivity (von der Malsburg 1973, Bienenstock et al 1982), each input pattern consists of an arbitrarily oriented bar of activity against a quiet background. (There can be added random noise that does not affect the results.) The bar's shape and center position are the same for all the input patterns. The only measure of pattern similarity is thus bar orientation similarity.

Von der Malsburg (1973) considers a layer of target cells with fixed center-surround lateral connections. Each source cell is connected to each target cell (topographic mapping is neither present initially nor does it emerge). The result is that most of the target cells become selectively responsive to bars lying at or near some orientation, and nearby target cells tend to become selective for bars of similar orientation.

Bienenstock et al (1982) consider a single target cell whose synaptic inputs can be thought of as forming a circular ring whose center is the same as that of each bar pattern. A version of a Hebb-like rule is used in which the change in strength depends upon the difference in value between the postsynaptic activity and an adjustable threshold. This threshold is

important in their model for ensuring that the target cell becomes selective for one bar pattern or a subset of similar patterns.

Because of the simple centered-bar form chosen for the input patterns, the emergence of orientation selectivity in both these models closely parallels the emergence of positional selectivity in the formation of topographic maps. In fact, the activity pattern on a semicircular portion of the ring of synapses in the latter model consists of an interval (or one-dimensional "spot") of activity whose position is given by the bar orientation. The development of orientation selectivity in this model thus corresponds exactly to the development of positional selectivity in a network having a one-dimensional source "layer" (that has been "wrapped around" so that its endpoints meet).

COLUMN FORMATION We have noted that von der Malsburg's (1973) model generates target cells whose orientation preferences are similar within a region of the target layer; that is, the cells form orientation "columns." In the model studied by Swindale (1982), each target cell is labeled by a vector whose magnitude and direction are modified by lateral interactions with the labels at other sites. The approach is the same as that used for ocular dominance stripe formation (Swindale 1980), in which each cell was labeled by its degree of left-eye and right-eye preference. The same type of center-surround lateral interaction rule is used, so that each label incrementally changes to become more nearly like its near neighbors and unlike its midrange neighbors. The result is that irregularly banded regions of similarly labeled cells emerge. In this model, the questions of why orientation-selective receptive fields arise, and how the particular lateral interaction between labels may arise, are not addressed.

By invoking a similar center-surround type of lateral interaction, von der Malsburg & Cowan (1982) show that if one postulates that different groups of target cells have, for example, horizontal and vertical orientation preference, then the lateral interactions induce the formation of a sequence of cells having the intermediate orientation preferences.

WHY ORIENTATION SELECTIVITY? Linsker (1986a-c) has analyzed the development of a network in which random uncorrelated signaling activity is preprocessed by a layer of cells having topographically arranged center-surround receptive fields. The output from this "source" layer of cells is in turn provided as input to a target layer via Hebb-modifiable connections. Even in the absence of lateral connections, a layer of orientation-selective target cells emerges under certain conditions. Since the architecture and development rules possess no orientational bias, the emergence of orientation selectivity is an example of a "symmetry-breaking" process (Linsker 1986b, 1988c). Under different conditions a layer of center-surround cells

emerges. The two parameters that determine which cell type develops are the radius of the source region that provides input to each target cell, and the degree of correlation needed to cause strength increase via the Hebb-like rule.

If the parameter values lie in a range that leads to orientation-selective cell formation, then adding weak lateral connections of fixed strength to the model causes orientation columns to form. In the model, feedforward connections suffice to generate orientation-selective cells. The simulations discussed do not address the question of the extent to which lateral interactions can influence the formation of orientation-selective cells (as opposed to their columnar organization).

The emergence of orientation-selective cells and columns in this model is explained by analyzing how the Hebb-like rule creates geometric patterns of connection strengths when the pairwise signal-activity correlations in the source layer depend upon the relative positions of the source cells (Linsker 1986b, 1987). These activity correlations arise because of the center-surround preprocessing. Even if one uses a Hebb-like rule different from the particular one used in this model, orientation-selective cells can emerge (see for example Kammen & Yuille 1988). The mathematical reason is that the orientation-selective solutions are formed from combinations of the first few eigenfunctions (those having the largest eigenvalues) of the covariance matrix of activity in the source layer (Linsker 1987).

Experimentally, one finds a series of several layers of cells having substantially center-surround receptive field properties in retina and LGN of cat and macaque monkey, and layer IV of macaque striate cortex, followed by the onset of well-tuned orientation selectivity. Although the feedforward connections in the model can in general be of both excitatory and inhibitory types, orientation selectivity emerges even when the connections from the center-surround layer are constrained to be excitatory (Linsker 1987, and unpublished results). This is of interest since the geniculocortical connections which immediately precede the first well-tuned orientation-selective stage in cat are thought to be exclusively excitatory.

In Linsker's model, the center-surround cells responsible for the preprocessing themselves emerge by a self-organizing process, given only random input and topographically arranged connections to each cell from a neighborhood of cells of the previous layer (Linsker 1986a). The means by which the center-surround layer is generated—whether by an activity-dependent self-organizing process as in the model, or by retinal interactions mediated by anatomically complex connections—does not affect the emergence of orientation selectivity in the later processing stage of the model.

The model as a whole shows that a sequence of feature-analyzing cell types of progressive complexity can emerge via a Hebb-like rule in a multistage feedforward network. Structured rather than random uncorrelated input can also be provided to the first layer, and the developmental results studied (Linsker 1988a).

Models relating to orientation selectivity that are not self-organizing models of cell formation are outside the scope of this review; two examples are (a) a nonadaptive model of cell formation (Braitenberg 1985) and (b) a demonstration, using a “learning by error correction” algorithm, that orientation-selective elements perform useful intermediate processing functions in a network computation of curvature in a scene (Lehky & Sejnowski 1988).

NETWORK MODELS, OPTIMIZATION PRINCIPLES, AND INFORMATION THEORY

The results of the preceding section provide a stimulus for investigating why a Hebb-like rule, with or without lateral connections, generates a sequence of feature-analyzing cell types in a layered network. This section explores several senses in which the types of patterning rules studied can create structures that are optimal from an information processing standpoint.

What Does a Hebb-like Rule Optimize?

First consider a single cell with feedforward but no lateral connections. To understand intuitively the effect of a Hebb-like rule, an analogy is useful: Imagine a committee whose recommendation (“output activity”) on any issue is the weighted average of its members’ opinions (“input activities”), each opinion being weighted according to the member’s voting strength. A member who consistently agrees (disagrees) with the committee’s opinion receives an increase (decrease) in voting strength. The operation of this Hebb-like rule transforms a committee of members having random voting strengths, whose averaged output is rarely a strong recommendation (either positive or negative), into a committee whose members of high voting strength tend to agree more often, and whose recommendations tend to be strong (in either direction) more of the time.

What does this result mean for a model cell (Linsker 1988a)? Consider a cell whose output activity is a linear function of its input activities. A histogram of the output values over a range of input presentations will have a spread, or variance, that depends upon the connection strengths. If the Hebb-like rule has a term that changes each strength by an amount proportional to the product of (a) the input activity at that connection

(minus its mean value) and (b) the cell's output activity (minus its mean value), this term will tend to increase the cell's output variance. Depending upon the other terms in the update rule, the net result can be to maximize the variance subject to some constraint, for example on the sum of the strengths (Linsker 1986b, 1988a,c) or the sum of the squares of the strengths (Amari 1977, Oja 1982). A cell with low output variance tends to be nonselective—its responses to different input presentations are similar. High variance tends to correspond to greater selectivity. Even when the cell's output variance is itself constrained (e.g. by a maximum allowed firing rate), or when the cell's response is nonlinear, a suitable Hebb-like rule can act to maximize the cell's responsiveness to statistical structure (or "features") in the input presentations, and to minimize its responsiveness to uncorrelated inputs or random processing noise (Linsker 1988a).

Under certain conditions, a Hebb-like rule generates cells that are optimal in several ways. As we have seen, such a rule produces cells that respond selectively to statistically significant properties of the input presentations. In particular, a suitably constructed rule (Oja 1982) produces a cell that performs a standard computation in statistical feature extraction called "principal component analysis." [Watanabe (1985) discusses in detail the application of this and related statistical methods to pattern recognition.] A cell produced by such a Hebb-like rule also has an "optimal inference" property: The average error incurred when using the cell's output value to estimate its input values is less for such a cell than for any other linear response cell. Finally, under certain conditions, the output from such a cell conveys maximum information about its input activity values, compared with cells having arbitrary connection strengths (Linsker 1988a).

The Role of Lateral Interactions

If a Hebb-like rule can cause individual target cells to develop so as to optimize certain properties, can the addition of lateral interactions allow an assembly, or an entire layer, of target cells to develop in an optimal or near-optimal way with respect to some property that is important for information processing?

When both feedforward and lateral connections are present, target cells not only can develop feature analyzing properties by means of a Hebb-like rule (as shown above), but can "cooperate" or "compete" with each other in the formation of these properties (see, for example, von der Malsburg 1973, Grossberg 1976, Kohonen 1982a,b, Rumelhart & Zipser 1985, Linsker 1986c, 1988a,b, 1989b, Pearson et al 1987, Hinton 1989). An optimization method (Durbin & Willshaw 1987) based on the "tea trade" model (von der Malsburg & Willshaw 1977) has been developed.

In some of these models the process of modifying the connection strengths corresponds closely to a statistical method for classifying or detecting regularities in data sets (see Lippmann 1987 for review). For example, a network algorithm for forming clusters of related input presentations, developed in connection with the "adaptive resonance theory" (Carpenter & Grossberg 1988), is related to the "sequential leader clustering" algorithm in statistics. Kohonen's (1982a,b) algorithm for "feature map" formation, discussed above, is related to a version of the "k-means" algorithm (MacQueen 1967) for partitioning a set of data points into a number of groups such that the members of each group are similar to each other according to a particular measure. This type of partitioning is useful for extracting common features from raw data, and also for the data compression problem of communicating an item of data, by stating which group it belongs to, in such a way that the original datum can be inferred with minimum error. [Cf. the "optimal inference" result of the previous section, the literature on "vector quantization" (reviewed by Gray 1984), and the work of Watanabe (1969, 1985) on feature extraction.] I discuss the choice of "similarity" measure further in the next section.

Might some type of statistical algorithm be playing an important role in the biological processing of sensory data? If so, what property if any is being optimized, how can the optimization process be realized by a biological network, and of what value is the process to the animal? Although these questions are necessarily speculative, I discuss here a concrete approach to them, in which some basic ideas from information theory play an important role (Linsker 1988a). To place recent connections between information theory and perceptual network algorithms in perspective, it is important to review the history, since the 1950s, of the idea that information theory may relate to the organization of sensory and perceptual processing. The relevant information-theoretic ideas are discussed here in an informal way. For the classic development of information theory see Shannon (1949); a tutorial treatment of some of the ideas as they relate to sensory processing is given in Linsker (1988a).

Information Theory and Perception

REDUNDANCY REDUCTION Sensory input, such as a time sequence of visual scenes, is not random and uncorrelated, but contains structure and regularities. Knowing certain aspects of the input allows us to infer other aspects; the input is partially redundant. Attneave (1954, p. 189) proposed that

A major function of the perceptual machinery is to strip away some of the redundancy of stimulation, to describe or encode incoming information in a form more economical than that in which it impinges on the receptors.

The suggested approach is “equivalent to that of a communications engineer” who exploits regularities in an input scene, in order to encode the input signals so that they may be sent along an essentially noise-free channel of limited capacity and reconstructed by the recipient with high fidelity (low distortion). [For examples of how the computations performed by cells during early visual processing can be analyzed from the standpoints of data compression and reconstruction, see Daugman (1988) and Oğuztöreli & Caelli (1986).]

Attneave’s recognition that encoding for redundancy reduction is related to the identification of specific features such as edges or corners in a scene, and that a redundancy-reducing strategy might be important for perceptual processing, is an important insight. The idea is related to Craik’s (1943) view of the brain as building a model of the external world that incorporates the world’s lawful regularities and constraints. Attneave’s idea of sensory processing as data compression, however, leaves some crucial questions unanswered: What role if any is played by the “recipient” in the communications analogy? If the original scene is never “reconstructed” by the brain (and why should it be?), what is the meaning of the fidelity criterion? That is, if not all input information can be preserved, what criterion determines what to discard? Finally, how do biological constraints and costs influence what types of encoding are “economical,” and how does a biological system carry out the encoding?

In a series of papers, Barlow has studied several of the physiological correlates of redundancy reduction as a perceptual processing strategy:

1. In a model having discrete (e.g., binary) inputs, with no processing noise to induce coding or transmission errors, one can encode the input “message” by removing correlations (a type of redundancy) and thus produce an output that is more compact than the input, yet is able to regenerate the input in a completely reversible way (Barlow 1959, 1961). The compaction process can be driven by a constraint limiting the average firing rate to a low value in later processing stages (Barlow’s principle of “economy of impulses”). Note that as long as the code is fully reversible, the question of what fidelity measure to use does not arise.

2. When the input signals are continuous-valued (or can take on a very large number of values), it is important to treat processing elements having limited resolution (Barlow 1969). This, however, leaves open the question of when two stimuli should be considered “similar” enough not to be resolved by the network, and of how “similarity” with respect to different properties should be compared. The issue is related to the choice of appropriate fidelity measure (above), and recurs in connection with Marr’s work (below).

3. Building on Hebb’s (1949) discussion of the importance of identifying

correlated inputs, one can regard a Hebb-like rule as forming cells that detect "suspicious coincidences" (Phillips et al 1984, Barlow 1985). These "coincidences" are sets of events that signal the existence of structure (e.g. coherent motion of an object), since they would not jointly occur in the absence of such structure. Legény (1970, 1975) discusses a similar idea, that of forming cells that respond selectively to "surprising" combinations of stimuli, in the context of a suggestion that individual cells may possess complex memory and processing capabilities to carry out the requisite computations.

4. The existence of multiple cortical areas with different mappings allows different types of regularities to be detected, if one assumes that a constraint on connection length requires cells to be relatively close together for correlations in their activity to be detected, and consequently for redundancy to be reduced (Barlow 1986).

THE INFORMON MODEL An information-theoretic idea is explicitly used as the motivation for a synaptic learning rule in the "informon" model of Uttley (1970, 1979). The "mutual information" (Shannon 1949) between two messages is, informally, a measure of the information that either message "conveys about" the other. For example, if two signals tend to co-occur, their mutual information is high. If knowing one signal has no effect on one's expectation of what the other signal is, their mutual information is zero. In Uttley's model both variable and fixed strength synapses provide input to a cell. Each variable strength is incrementally adjusted so that the cell's output becomes as nearly decorrelated with the synaptic input as possible. The update rule is motivated by the idea of minimizing the mutual information between input and output at each synapse, although in practice a function different from the mutual information is used. This learning rule is very similar to the Rescorla-Wagner conditioning rule and to the Widrow-Hoff "LMS" algorithm for supervised learning (Widrow & Stearns 1985). The rule adjusts the strengths so that the net effect of the inputs at the variable-strength synapses is to cancel, as nearly as possible, the net reinforcement or "teacher" signal at the fixed-strength synapses. The output of the "informon" cell is then equal to the residual difference between the two sets of signals. This rule differs from a Hebb-like rule, in which correlations between input and output cause synaptic strength to increase. Nonetheless, the rule causes the cell to detect structure among the inputs to the variable-strength synapses, and to use such structure to match and cancel the "teacher" signal most effectively. I return to the use of a mutual information criterion below (Linsker 1988a) in a different context.

INFORMATION THEORY AND BIOLOGICAL UTILITY Marr's (1970) discussion

of redundancy reduction in sensory processing addresses an issue (introduced above) that is crucial for any connection between information-theoretic or statistical classification schemes and perception. To form generalizations, a system must be able to decide when two events (input presentations) are "similar"—that is, belong in the same class. This "lumping process," however, discards information about the events. By what criterion should this, rather than other, information be discarded? Marr notes that system reliability offers one such criterion: If two events correspond to sufficiently similar sets of signal activities, the events could be confused by the network, and therefore should not be classified differently. He does not, however, consider this a fundamental criterion.

Marr seeks to integrate information-theoretic ideas with biological utility in the following way. His criterion for discarding information is that the loss of that information should not impair the system's ability to diagnose whether the event possesses features of a certain type. The type of feature that is favored is a feature whose presence or absence (in an event) is not sensitive to small changes in the set of other, already-diagnosed features that are present (in the same event). Thus the presence of a feature (or type of redundancy) as such is not sufficient ground for classification; the feature must tend to co-occur with a sufficiently large class of other features (see the "Fundamental Hypothesis," Marr 1970, p. 182).

It is intuitively plausible that the detection of a feature is likely to have greater biological utility if that feature tends to be associated with others. Marr's (1970) theory is developed substantially beyond this point to make neurophysiologic predictions concerning the roles of various cell types. Important questions that are left open include: Can the theory be used to make specific predictions of feature-analyzing properties? If not, would the incorporation of additional biological constraints of some type suffice to allow one to make such predictions? Is the proposed feature classification criterion definite enough that it can be used, even in principle, to construct a perceptual system (or the early stages of one) whose functioning might be compared with what is biologically observed?

A PRINCIPLE OF MAXIMUM INFORMATION PRESERVATION Motivated by the finding (discussed above) that a Hebb-like rule under certain conditions generates a cell whose output conveys maximum information about its input activities, Linsker (1988a) has proposed an information-theoretic principle for the organization of a biological sensory processing stage having source and target cell layers. This principle of "maximum information preservation" states that the transformation of sensory signals from source to target layer should be chosen such that the target cell activities jointly convey maximum information about the source cell activi-

ties. The quantity to be maximized is the average (over time) of the mutual information between the two sets of activities.

The choice of transformation is subject to biological “hardware” constraints. These may include constraints on (a) the type of function each cell can compute (e.g. a linear combination of inputs, or a sigmoid function of the linear combination, or some more complicated function), (b) the spatial extent of lateral connections, and (c) the reliability of signal transmission and processing.

The concept of maximizing average mutual information between input and output can be made more concrete by an analogy to a “guessing game” (cf. Shannon 1949) such as “Twenty Questions.” In that game, the questioner’s proper strategy is to ask, at each stage, a question whose answer will convey, on average, the maximum possible amount of information (in that case, one bit of information) about the unknown object. In this analogy, the “questioner” is the cell, the “question” is the computation being performed by the cell upon its inputs (e.g. a center-surround cell “asks” the “question”: “What degree of contrast is present between the central and peripheral regions of my receptive field?”), and the “answer” is the output activity of the cell. An unreliable “answer” corresponds to the effect of noise associated with the processing stage.

We want to know what set of processing functions, or computations, emerges for an entire layer of cells, according to the principle of maximum information preservation. This is analogous to asking: How should a set of questioners, asking questions *in parallel* about the same object (the input presentation), choose their questions optimally?

If the activity values could take on only discrete values, and if noise could be ignored, the proposed principle would lead to data compression with redundancy reduction along the lines discussed by Attneave and Barlow (above). Processing noise is unavoidable, however, and it has important consequences. It helps to determine which types of information about the input pattern are preserved or discarded, and how redundancy is to be introduced—as well as removed—during processing. The latter point is important if one is to account for the similar response properties of nearby cells in the context of an “optimal encoding” principle.

The optimal choice of cell response properties, according to the principle of maximum information preservation, depends upon the types of statistical regularities present in the set of input presentations, and on the processing constraints. The principle has been applied to simple model networks under various conditions. It generates features that are qualitatively similar to those found in biological systems. These features include topographic maps (Linsker 1988b, 1989b; also see MAP MAGNIFICATION FACTORS above), center-surround cells and cells sensitive to temporal variations in

input activity (Linsker 1989a), orientation-selective cells, and column-like assemblies of cells having similar response properties (Linsker 1988a,b, 1989b). In each case the generic quantity being optimized—the average mutual information between source and target layer—is the same.

Note that the quantity being maximized is a function of the processing being carried out by many target cells. This global optimization criterion stands in contrast to other proposed criteria that refer only to an information-theoretic quantity at a single connection (Uttley 1970, 1979) or cell (Pearlmutter & Hinton 1986). A different global information-theoretic quantity is optimized in the “Boltzmann machine” network discussed by Hinton & Sejnowski (1983).

The idea of choosing system parameters so as to maximize an appropriately defined mutual information has been used in several nonbiological contexts, including the choice or placement of sensors in physical systems (e.g. Phua & Dillon 1977 and references therein, Luttrell 1985, Fraser & Swinney 1986) and the choice of acoustic-processing models in speech recognition (Bahl et al 1987). Much remains to be learned about the patterns of information processing that can emerge from this type of information-theoretic principle, particularly in nonlinear systems having multiple processing stages. The potential value of cross-fertilization between neuroscience and nonbiological fields, in advancing our understanding of these issues, is great.

Although the average mutual information between two layers of cells is a global property having a complex mathematical form, a local algorithm that implements the proposed principle has recently been developed for certain types of model networks having feedforward and lateral connections (Linsker 1989b). By a local algorithm we mean one that modifies connection strengths in a way that depends only upon signals available at that connection or cell.

This algorithm, or learning rule, was derived by asking: How should the response properties of each cell be incrementally adjusted, so as to increase the average mutual information between the source and target layer activities by the greatest amount? (It is an example of a “gradient ascent” learning rule.) It is striking that the resulting rule (Linsker 1989b) exhibits properties of Hebb-like modification and cooperative and competitive learning, combined in a particular way, even though no assumptions were made concerning the form of the learning rule or its component properties.

The principle of maximum information preservation has some theoretically attractive features (Linsker 1988a, 1989b), and the existence of local algorithms increases our confidence that biological networks may be capable of implementing it. It does not follow from this that nature in fact makes use of the principle. It will be necessary to see what detailed pre-

dictions the principle generates in biologically realistic situations, and to subject these to experimental test.

The proposed principle may be usefully extended in various ways. For example, the principle as stated uses processing noise as the criterion for classifying two patterns as "similar." (That is, "similar" patterns are those which cannot be reliably discriminated from each other.) It can be biologically useful to introduce a different criterion of pattern "similarity." A suitable choice of criterion can facilitate the learning of generalizations or allow resolution to be varied as the focus of attention is shifted. The type of local algorithm mentioned above can still be used to generate near-optimal processing stages provided the "similarity" criterion satisfies certain conditions (Linsker 1989b). [For examples of how the choice of resolution or "similarity" measure is handled in other approaches to learning and perception, see Carpenter & Grossberg's (1988) use of an adjustable "vigilance" parameter, and the "regularization" method for reconstructing aspects of a scene from its image (Poggio & Koch 1985).]

SUMMARY AND FUTURE DIRECTIONS

This review has focussed on a set of ways in which experimental work, modeling, and theory have interacted in advancing our understanding of the organization of sensory processing systems. These interactions are not of a rigid "bottom-up" or "top-down" type. They involve interplay among—rather than isolated study of—the "computational task," algorithmic, and hardware levels of description (Marr 1982).

The path traced here comprises the following steps:

1. The finding of specific salient experimental features motivates the search for models, especially self-organizing models, that can generate these features.
2. Common elements of architecture and patterning rules are found to underlie a variety of these models.
3. These elements are found, in certain cases, to be associated with optimization principles that appear significant from an information-processing standpoint. In essence, the models are treated as objects of study in their own right, and are found to have properties that might have been difficult or impossible to infer directly from observation of the much more complex biological system. This study leads to links with principles and methods in statistical analysis and information theory.
4. To be able to test whether these putative principles are biologically relevant, it is necessary to develop biologically plausible algorithms that can implement them, and to generate explicit predictions of cell

properties and neural maps that can be compared with experiment. This step has been taken for simple model cases. More work is required to extend the principle and algorithms to networks and sets of input presentations that are of biologically realistic complexity.

If some of the theoretical ideas discussed withstand future experimental tests, it will become important to extend the work in two directions. First, new experimental findings concerning neuronal dynamics (e.g. Gray & Singer 1989) could allow "optimal encoding" strategies to be implemented in new ways. Second, even if the principles discussed are found to describe some of the main effects of feedforward and lateral interactions within a processing stage of a sensory system, it will be important to understand how these principles relate to higher levels of neural organization. The study of attentional mechanisms, other effects of feedback from later processing states, integration across sensory modalities, and sensorimotor integration are examples of such higher-level organizational issues to which the type of approach explored here might be fruitfully extended.

More sophisticated experimental techniques (e.g. for recording and analyzing complex spatial and temporal patterns of activity), increased computational power, and refinement of theoretical ideas are likely to lead to closer and more fruitful contact between experimental and theoretical work. For the theorist, it is important to use experimental findings to (a) constrain the essential features of biological models, (b) inspire and inform the choice of key questions, and (c) enable comparisons with predicted structure and behavior. For the experimentalist, model predictions may suggest appropriate candidate stimuli for studying receptive field properties and neural maps, and may help to guide the analysis and interpretation of complex activity data. Such guidance becomes valuable as increasingly complex sets of response properties are found, and as the characterization of a cell as a detector or analyzer of a well-defined "feature" may become more difficult or untenable.

More generally, if we are to gain an understanding of how and why neural systems are organized in particular ways, it is important to supplement descriptive principles with a developing theoretical framework that can be used to generate testable predictions and to relate structure and function at a variety of organizational levels. Multifaceted interactions among experiment, modeling, and theory will play an important role in achieving these goals.

ACKNOWLEDGMENTS

I thank Drs. Eric Schwartz, Roger Traub, and Christoph von der Malsburg for their comments on the manuscript.

Literature Cited

- Amari, S.-I. 1977. Neural theory of association and concept-formation. *Biol. Cybern.* 26: 175-85
- Amari, S.-I. 1980. Topographic organization of nerve fields. *Bull. Math. Biol.* 42: 339-64
- Attneave, F. 1954. Some informational aspects of visual perception. *Psychol. Rev.* 61: 183-93
- Bahl, L. R., Brown, P. F., de Souza, P. V., Mercer, R. L. 1987. Speech recognition with continuous-parameter hidden Markov models. *Comput. Speech Lang.* 2: 219-34
- Barlow, H. B. 1959. Sensory mechanisms, the reduction of redundancy, and intelligence. In *Mechanisation of Thought Processes, Natl. Phys. Lab. Symp.* 10: Paper 4-1. London: HMSO
- Barlow, H. B. 1961. Possible principles underlying the transformations of sensory messages. In *Sensory Communication*, ed. W. A. Rosenblith, pp. 217-34. Cambridge: MIT Press and New York/London: Wiley
- Barlow, H. B. 1969. Trigger features, adaptation and economy of impulses. In *Information Processing in the Nervous System*, ed. K. N. Leibovic, pp. 209-26. New York: Springer-Verlag
- Barlow, H. B. 1985. Cerebral cortex as model builder. In *Models of the Visual Cortex*, ed. D. Rose, V. G. Dobson, pp. 37-46. New York/London: Wiley
- Barlow, H. B. 1986. Why have multiple cortical areas? *Vision Res.* 26: 81-90
- Bienenstock, E. L., Cooper, L. N., Munro, P. W. 1982. Theory for the development of neuron selectivity: Orientation specificity and binocular interaction in visual cortex. *J. Neurosci.* 2: 32-48
- Blasdel, G. G., Salama, G. 1986. Voltage-sensitive dyes reveal a modular organization in monkey striate cortex. *Nature* 321: 579-85
- Braitenberg, V. 1985. Charting the visual cortex. In *Cerebral Cortex*, ed. A. Peters, E. G. Jones, 3: 379-414. New York/London: Plenum
- Brown, T. H., Kairiss, E. W., Keenan, C. L. 1990. Hebbian synapses: Biophysical mechanisms and algorithms. *Annu. Rev. Neurosci.* 13: 475-511
- Carpenter, G. A., Grossberg, S. 1988. The ART of adaptive pattern recognition by a self-organizing neural network. *Computer* 21(3): 77-88
- Constantine-Paton, M., Law, M. I. 1978. Eye-specific termination bands in tectum of three-eyed frogs. *Science* 202: 639-41
- Craik, K. J. W. 1943. *The Nature of Explanation*. Cambridge: Cambridge Univ. Press
- Daugman, J. G. 1988. Complete discrete 2-d Gabor transforms by neural networks for image analysis and compression. *IEEE Trans. Acoust. Speech Signal Process.* 36: 1169-79
- Durbin, R., Mitchison, G. 1988. Dimension reduction models for topographical maps in the cortex. *Neural Networks* 1(Suppl. 1): 491 (Abstr.)
- Durbin, R., Willshaw, D. J. 1987. An analogue approach to the travelling salesman problem using an elastic net method. *Nature* 326: 689-91
- Ferster, D., Koch, C. 1987. Neuronal connections underlying orientation selectivity in cat visual cortex. *Trends Neurosci.* 10: 487-92
- Fraser, A. M., Swinney, H. L. 1986. Independent coordinates for strange attractors from mutual information. *Phys. Rev. A* 33: 1134-40
- Fraser, S. E. 1985. Cell interactions involved in neuronal patterning: An experimental and theoretical approach. In *Molecular Bases of Neural Development*, ed. G. M. Edelman, W. E. Gall, W. M. Cowan, pp. 481-507. New York: Wiley
- Frégnac, Y., Imbert, M. 1984. Development of neuronal selectivity in primary visual cortex of cat. *Physiol. Rev.* 64: 325-434
- Gray, C. M., Singer, W. 1989. Stimulus-specific neuronal oscillations in orientation columns of cat visual cortex. *Proc. Natl. Acad. Sci. USA* 86: 1698-1702
- Gray, R. M. 1984. Vector quantization. *IEEE ASSP Mag.* 1(2): 4-29
- Grossberg, S. 1976. On the development of feature detectors in the visual cortex with applications to learning and reaction-diffusion systems. *Biol. Cybern.* 21: 145-59
- Hebb, D. O. 1949. *The Organization of Behavior*. New York: Wiley
- Hinton, G. E. 1989. Connectionist learning procedures. *Artif. Intell.* 40: 185-234
- Hinton, G. E., Sejnowski, T. J. 1983. Optimal perceptual inference. *Proc. IEEE Conf. Comput. Vision*, pp. 448-53
- Hubel, D. H., Wiesel, T. N. 1962. Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *J. Physiol.* 160: 106-54
- Hubel, D. H., Wiesel, T. N. 1963. Shape and arrangement of columns in cat's striate cortex. *J. Physiol.* 165: 559-68
- Hubel, D. H., Wiesel, T. N. 1968. Receptive fields and functional architecture of monkey striate cortex. *J. Physiol.* 195: 215-43
- Hubel, D. H., Wiesel, T. N. 1977. Functional

- architecture of macaque monkey visual cortex (Ferrier lecture). *Proc. R. Soc. London Ser. B* 198: 1-59
- Kaas, J. H., Merzenich, M. M., Killackey, H. P. 1983. The reorganization of somatosensory cortex following peripheral-nerve damage in adult and developing mammals. *Annu. Rev. Neurosci.* 6: 325-56
- Kammen, D. M., Yuille, A. L. 1988. Spontaneous symmetry-breaking energy functions and the emergence of orientation selective cortical cells. *Biol. Cybern.* 59: 23-31; (erratum) 59: 430
- Kohonen, T. 1982a. Self-organized formation of topologically correct feature maps. *Biol. Cybern.* 43: 59-69
- Kohonen, T. 1982b. Analysis of a simple self-organizing process. *Biol. Cybern.* 44: 135-40
- Kohonen, T. 1988. *Self-Organization and Associative Memory*. Berlin/Heidelberg/New York: Springer-Verlag. 2nd ed.
- Legéndy, C. R. 1970. The brain and its information trapping device. In *Progress of Cybernetics*, ed. J. Rose, 1: 309-38. London/New York/Paris: Gordon & Breach
- Legéndy, C. R. 1975. Three principles of brain function and structure. *Int. J. Neurosci.* 6: 237-54
- Lehky, S. R., Sejnowski, T. J. 1988. Network model of shape-from-shading: neural function arises from both receptive and projective fields. *Nature* 333: 452-54
- LeVay, S., Hubel, D. H., Wiesel, T. N. 1975. The pattern of ocular dominance columns in macaque visual cortex revealed by a reduced silver stain. *J. Comp. Neurol.* 159: 559-76
- Linsker, R. 1986a. From basic network principles to neural architecture: Emergence of spatial-opponent cells. *Proc. Natl. Acad. Sci. USA* 83: 7508-12
- Linsker, R. 1986b. From basic network principles to neural architecture: Emergence of orientation-selective cells. *Proc. Natl. Acad. Sci. USA* 83: 8390-94
- Linsker, R. 1986c. From basic network principles to neural architecture: Emergence of orientation columns. *Proc. Natl. Acad. Sci. USA* 83: 8779-83
- Linsker, R. 1987. Towards an organizing principle for perception: Hebbian synapses and the principle of optimal neural encoding. *IBM Res. Rep. RC12830*. IBM Res., Yorktown Heights, NY
- Linsker, R. 1988a. Self-organization in a perceptual network. *Computer* 21(3): 105-17
- Linsker, R. 1988b. Towards an organizing principle for a layered perceptual network. In *Neural Information Processing Systems*, Denver, CO, 1987, ed. D. Z. Anderson, pp. 485-94. New York: Am. Inst. Phys.
- Linsker, R. 1988c. Development of feature-analyzing cells and their columnar organization in a layered self-adaptive network. In *Computer Simulation in Brain Science, Copenhagen, 1986*, ed. R. M. J. Cotterill, pp. 416-31. Cambridge/New York: Cambridge Univ. Press
- Linsker, R. 1989a. An application of the principle of maximum information preservation to linear systems. In *Advances in Neural Information Processing Systems, Denver, CO, 1988*, ed. D. S. Touretzky, 1: 186-94. San Mateo, Calif: Morgan Kaufmann
- Linsker, R. 1989b. How to generate ordered maps by maximizing the mutual information between input and output signals. *Neural Comput.* 1: 402-11
- Lippmann, R. P. 1987. An introduction to computing with neural nets. *IEEE ASSP Mag.* 4(2): 4-22
- Luttrell, S. P. 1985. The use of transinformation in the design of data sampling schemes for inverse problems. *Inverse Problems* 1: 199-218
- MacQueen, J. 1967. Some methods for classification and analysis of multivariate observations. In *Proc. 5th Berkeley Symp. Math. Stat. Prob.*, ed. L. M. LeCam, J. Neyman, 1: 281-97. Berkeley: Univ. Calif. Press
- Marr, D. 1970. A theory for cerebral neocortex. *Proc. R. Soc. London Ser. B* 176: 161-234
- Marr, D. 1982. *Vision*. San Francisco: Freeman
- Merzenich, M. M., Nelson, R. J., Stryker, M. P., Cynader, M. S., Schoppmann, A., Zook, J. M. 1984. Somatosensory cortical map changes following digit amputation in adult monkeys. *J. Comp. Neurol.* 224: 591-605
- Métin, C., Frost, D. O. 1989. Visual responses of neurons in somatosensory cortex of hamsters with experimentally induced retinal projections to somatosensory thalamus. *Proc. Natl. Acad. Sci. USA* 86: 357-61
- Miller, K. D., Keller, J. B., Stryker, M. P. 1986. Models for the formation of ocular dominance columns solved by linear stability analysis. *Soc. Neurosci. Abstr.* 12: 1373
- Miller, K. D., Keller, J. B., Stryker, M. P. 1989. Ocular dominance column development: Analysis and simulation. *Science* 245: 605-15
- Mountcastle, V. B. 1978. An organizing principle for cerebral function: The unit module and the distributed system. In *The Mindful Brain*, ed. G. M. Edelman, V. B. Mountcastle. Cambridge: MIT Press
- Oğuztöreli, M. N., Caelli, T. M. 1986. An inverse problem in neural processing. *Biol. Cybern.* 53: 239-45

- Oja, E. 1982. A simplified neuron model as a principal component analyzer. *J. Math. Biol.* 15: 267-73
- Pearlmutter, B. A., Hinton, G. E. 1986. G-maximization: An unsupervised learning procedure for discovering regularities. In *Neural Networks for Computing*, ed. J. S. Denker, AIP Conf. Proc. 151. New York: Am. Inst. Phys.
- Pearson, J. C., Finkel, L. H., Edelman, G. M. 1987. Plasticity in the organization of adult cerebral cortical maps: A computer simulation based on neuronal group selection. *J. Neurosci.* 7: 4209-23
- Phillips, C. G., Zeki, S., Barlow, H. B. 1984. Localization of function in the cerebral cortex. *Brain* 107: 327-61
- Phua, K., Dillon, T. S. 1977. Optimal choice of measurements for state estimation. In *1977 Power Industry Computer Applic. Conf., Toronto*, pp. 431-41. New York: IEEE
- Poggio, T., Koch, C. 1985. Ill-posed problems in early vision: From computational theory to analogue networks. *Proc. R. Soc. London Ser. B* 226: 303-23
- Rauschecker, J. P. 1987. What signals are responsible for synaptic changes in visual cortical plasticity? In *Imprinting and Cortical Plasticity*, ed. J. P. Rauschecker, P. Marler, pp. 193-219. New York: Wiley
- Reiter, H. O., Stryker, M. P. 1988. Neural plasticity without postsynaptic action potentials: Less-active inputs become dominant when kitten visual cortical cells are pharmacologically inhibited. *Proc. Natl. Acad. Sci. USA* 85: 3623-27
- Ritter, H., Schulten, K. 1986. On the stationary state of Kohonen's self-organizing sensory mapping. *Biol. Cybern.* 54: 99-106
- Rumelhart, D. E., Zipser, D. 1985. Feature discovery by competitive learning. *Cogn. Sci.* 9: 75-112
- Schmidt, J. T. 1985. Formation of retinotopic connections: Selective stabilization by an activity-dependent mechanism. *Cell. Mol. Neurobiol.* 5: 65-84
- Schwartz, E. L. 1977. The development of specific visual connections in the monkey and the goldfish: Outline of a geometric theory of receptotopic structure. *J. Theor. Biol.* 69: 655-83
- Shannon, C. E. 1949. In *The Mathematical Theory of Communication*, ed. C. E. Shannon, W. Weaver. Urbana: Univ. Illinois Press
- Shatz, C. J., Stryker, M. P. 1978. Ocular dominance in layer IV of the cat's visual cortex and the effects of monocular deprivation. *J. Physiol.* 281: 267-83
- Sperry, R. W. 1943. Visuomotor coordination in the newt (*Triturus viridescens*) after regeneration of the optic nerve. *J. Comp. Neurol.* 79: 33-55
- Sperry, R. W. 1963. Chemoaffinity in the orderly growth of nerve fiber patterns and connections. *Proc. Natl. Acad. Sci. USA* 50: 703-10
- Stent, G. S. 1973. A physiological mechanism for Hebb's postulate of learning. *Proc. Natl. Acad. Sci. USA* 70: 997-1001
- Stryker, M. P., Harris, W. A. 1986. Binocular impulse blockade prevents the formation of ocular dominance columns in cat visual cortex. *J. Neurosci.* 6: 2117-33
- Sur, M., Garraghty, P. E., Roe, A. W. 1988. Experimentally induced visual projections into auditory thalamus and cortex. *Science* 242: 1437-41
- Swindale, N. V. 1980. A model for the formation of ocular dominance stripes. *Proc. R. Soc. London Ser. B* 208: 243-64
- Swindale, N. V. 1982. A model for the formation of orientation columns. *Proc. R. Soc. London Ser. B* 215: 211-30
- Takeuchi, A., Amari, S.-I. 1979. Formation of topographic maps and columnar microstructures in nerve fields. *Biol. Cybern.* 35: 63-72
- Turing, A. M. 1952. The chemical basis of morphogenesis. *Philos. Trans. R. Soc. London Ser. B* 237: 37-72
- Udin, S. B., Fawcett, J. W. 1988. Formation of topographic maps. *Annu. Rev. Neurosci.* 11: 289-327
- Uttley, A. M. 1970. The informon: A network for adaptive pattern recognition. *J. Theor. Biol.* 27: 31-67
- Uttley, A. M. 1979. *Information Transmission in the Nervous System*. London/New York/San Francisco: Academic
- von der Malsburg, C. 1973. Self-organization of orientation sensitive cells in the striate cortex. *Kybernetik* 14: 85-100
- von der Malsburg, C. 1979. Development of ocularity domains and growth behaviour of axon terminals. *Biol. Cybern.* 32: 49-62
- von der Malsburg, C., Cowan, J. D. 1982. Outline of a theory for the ontogenesis of iso-orientation domains in visual cortex. *Biol. Cybern.* 45: 49-56
- von der Malsburg, C., Willshaw, D. J. 1976. A mechanism for producing continuous neural mappings: Ocularity dominance stripes and ordered retino-tectal projections. *Exp. Brain Res.* 1: 463-69 (Suppl.)
- von der Malsburg, C., Willshaw, D. J. 1977. How to label nerve cells so that they can interconnect in an ordered fashion. *Proc. Natl. Acad. Sci. USA* 74: 5176-78
- Watanabe, S. 1969. *Knowing and Guessing*. New York: Wiley
- Watanabe, S. 1985. *Pattern Recognition: Human and Mechanical*. New York: Wiley
- Whitelaw, V. A., Cowan, J. D. 1981. Specificity and plasticity of retinotectal con-

- nections: A computational model. *J. Neurosci.* 1: 1369-87
- Widrow, B., Stearns, S. D. 1985. *Adaptive Signal Processing*. Englewood Cliffs, NJ: Prentice-Hall
- Wiesel, T. N., Hubel, D. H. 1974. Ordered arrangement of orientation columns in monkeys lacking visual experience. *J. Comp. Neurol.* 158: 307-18
- Willshaw, D. J., von der Malsburg, C. 1976. How patterned neural connections can be set up by self-organization. *Proc. R. Soc. London Ser. B* 194: 431-45
- Willshaw, D. J., von der Malsburg, C. 1979. A marker induction mechanism for the establishment of ordered neural mappings: Its application to the retinotectal problem. *Proc. R. Soc. London Ser. B* 287: 203-43