

MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

2

AD-A151 776

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER #23		2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) Recent Developments in Neural Models		5. TYPE OF REPORT & PERIOD COVERED Technical Report	
		6. PERFORMING ORG. REPORT NUMBER	
7. AUTHOR(s) Christopher L. Scofield and Leon N Cooper		8. CONTRACT OR GRANT NUMBER(s) N00014-81-K-0136	
9. PERFORMING ORGANIZATION NAME AND ADDRESS Department of Physics and Center for Neural Science Brown University Providence, Rhode Island 02912		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS NR-201-484	
11. CONTROLLING OFFICE NAME AND ADDRESS Personnel and Training Research Program Office of Naval Research, Code 442PT Arlington, Virginia, 22217		12. REPORT DATE March 8, 1985	
		13. NUMBER OF PAGES 52	
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)		15. SECURITY CLASS. (of this report)	
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE	
16. DISTRIBUTION STATEMENT (of this Report) Approved for public release; distribution unlimited. Publication in part or in whole is permitted for any purpose of the United States Government.			
17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)			
18. SUPPLEMENTARY NOTES Submitted for publication.			
19. KEY WORDS (Continue on reverse side if necessary and identify by block number) Visual Cortex; Neural Networks; Selectivity; Ocular Dominance; Distributed Memory.			
20. ABSTRACT (Continue on reverse side if necessary and identify by block number) Recent progress in a theory of the development of neural networks and their application to visual cortex is discussed. We find, that under proper assumptions, distributed memories display some of the features of animal memory. We discuss a mechanism, consistent with known neurophysiology, by which these memories develop and modify with experience. This mechanism is applied to the problem of the development of orientation selectivity and ocular dominance in visual cortex. The proposed mechanisms of synaptic evolution is based upon the competition of incoming patterns rather than			

DTIC
SELECTED
MAR 28 1985
S E D

DTIC FILE COPY

converging afferents. It is found that neurons develop maximal selectivity with respect to their signal environment, and under a variety of rearing conditions develop orientation tuning curves and ocular dominance comparable to that found in experiment. Finally, we describe a model first-order anatomy of visual cortex in which we implement this form of synaptic modification.

Accession For	
NTIS GRA&I	<input checked="" type="checkbox"/>
DTIC TAB	<input type="checkbox"/>
Unannounced	<input type="checkbox"/>
Justification	
By	
Distribution/	
Availability Codes	
Dist	Avail. and/or Special
A-1	

ORIGINAL AUTHOR-SUPPLIED KEYWORDS:

Form 1473



RECENT DEVELOPMENTS IN NEURAL MODELS

BY

Christopher L. Scofield
and
Leon N Cooper

Department of Physics
and
Center for Neural Science

Brown University
Providence, Rhode Island 02912

*Supported in Part by the U. S. Office of Naval Research,
Contract #N00014-81-K-0136

ABSTRACT

Recent progress in a theory of the development of neural networks and their application to visual cortex is discussed. We find, that under proper assumptions, distributed memories display some of the features of animal memory. We discuss a mechanism, consistent with known neurophysiology, by which these memories develop and modify with experience. This mechanism is applied to the problem of the development of orientation selectivity and ocular dominance in visual cortex. The proposed mechanism of synaptic evolution is based upon the competition of incoming patterns rather than converging afferents. It is found that neurons develop maximal selectivity with respect to their signal environment, and under a variety of rearing conditions develop orientation tuning curves and ocular dominance comparable to that found in experiment. Finally, we describe a model first-order anatomy of visual cortex in which we implement this form of synaptic modification.

INTRODUCTION

Work of the last several decades has brought a greater understanding of the properties of individual neurons in many parts of the brain. However, the manner in which large interacting networks of these nerve cells produce mental activity remains far from understood. It has become clear that it is not possible to localize any region of the nervous system in which memory is stored nor to isolate activity of single cells corresponding to higher cortical function. Study of the visual system has, however, demonstrated that it is possible to localize cellular activity corresponding to lower forms of signal processing. This work has provided an arena in which to study those mechanisms that underly the evolution of individual cell properties. All of this reinforces the belief that nervous-system properties such as storage and retrieval of memory are of unusual subtlety, involving small changes in the activities of large numbers of neurons.

A basic problem in understanding the organization of memory in a biological system is to understand how a vast quantity of information can be stored and recalled by a system composed of vulnerable and relatively unreliable elements and with no knowledge of how or where the information has been filed. The distribution of memory over a large number of cells seems, at first, very unpromising since individual memories would interfere with each other; but, as we shall see, the problem has a solution and memories so constructed can function as well as local memories.

We have been analyzing a class of neural models for the acquisition and storage of distributed memories that display, on a primitive level, features such as recognition, association, and generalization, and which suggest some of the mental behavior associated with animal memory and learning (Cooper, 1973; Anderson and Cooper, 1978). Various algorithms related to Hebb, 1949) can account for the formation of this and other classes of distributed memories (see, for example, Marr, 1969; Brindley, 1969; Anderson, 1970, 1972; Cooper, 1973; Kohonen, 1977). We suggest that it may be the same fundamental mechanism, accessible to detailed experimental investigation in primary sensory areas of the nervous system, which is also responsible for some of the higher forms of central nervous system organization.

It is now commonly believed that much of the learning and resulting organization of the central nervous system occurs due to some kind of modification of the efficacy or strength of at least some of the synaptic junctions between neurons, thus altering the relation between presynaptic and postsynaptic potentials. It is also known that small but coherent modifications of large numbers of synaptic junctions can result in distributed memories.

The neurons of the primary sensory areas, at nearly all levels, display various forms of stimulus selectivity. They may respond preferentially to a tone of a given frequency, a light spot of a given color, a light bar of a certain orientation, etc. We might, therefore, regard stimulus selectivity as a general property of sensory neurons and conjecture that the development of such selectivity obeys some general rule. Most attractive is the idea that some of the mechanisms by which selectivity develops in embryonic or early postnatal life are sufficiently general to allow a unifying theoretical treatment.

Applying our general ideas to the development of orientation selectivity and binocular interaction in area 17 of the cat visual cortex, we obtain a theory based on a single mechanism of synaptic modification that accounts for the great variety of experimental results on monocular and binocular experience in normal and various altered visual environments. In addition, we obtain some new predictions.

I. DISTRIBUTED MEMORY MAPPINGS

For a distributed memory it is the simultaneous or near-simultaneous activities of many different neurons (the result of external or internal stimuli) that are of interest. Thus a large spatially distributed pattern of neuron discharges, each of which might not be very far from spontaneous activity, could contain important, if hard to detect, information. Let us consider the behavior of an idealized neural network (that might be regarded as a model component of a nervous system) to illustrate some of the important features of distributed mappings.

Consider N neurons $1, 2, \dots, N$, each of which has some spontaneous firing rate r_{j0} . (This need not be the same for all of the neurons, nor need it be constant in time.) We can then define an N -tuple whose components are the difference between the actual firing rate r_j of the j th neuron and the spontaneous firing rate r_{j0} :

$$f_j = r_j - r_{j0} \quad (1.1)$$

By constructing two such banks of neurons connected to one another (or even by the use of a single bank which feeds signals back to itself), we arrive at a simplified model as illustrated in Figure 1.1.

Figure 1.1. An ideal distributed mapping. Each of the N incoming neurons in F is connected to each of the N outgoing neurons in G by a single ideal junction. (Only the connections to i are drawn.) (From Cooper, 1973.)

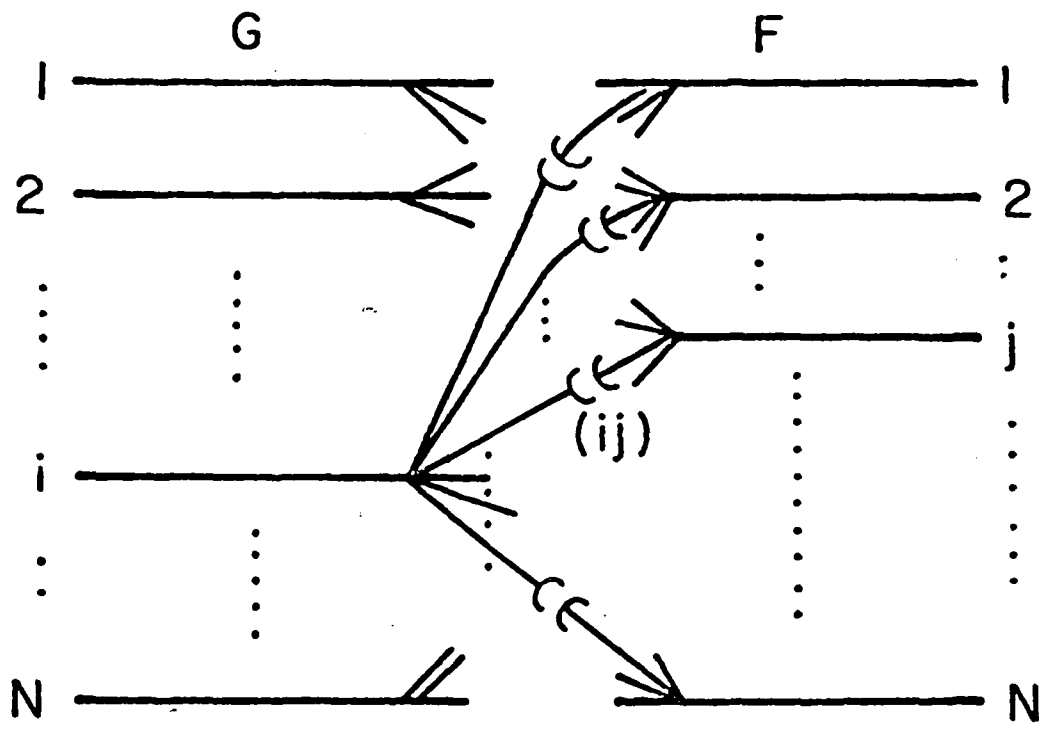


Figure 1.1

The actual synaptic connections between one neuron and another are generally complex and redundant; we have idealized the network by replacing this multiplicity of synapses between axons and dendrites by a single ideal junction which summarizes logically the effect of all of the synaptic contacts between the incoming axon branches from neuron j in the F bank and the dendrites of the outgoing neuron i in the G bank (Figure 1.2). Each of the N incoming neurons, in F, is connected to each of the N outgoing neurons, in G, by a single ideal junction.

Figure 1.2. An ideal junction (From Cooper, 1973.)

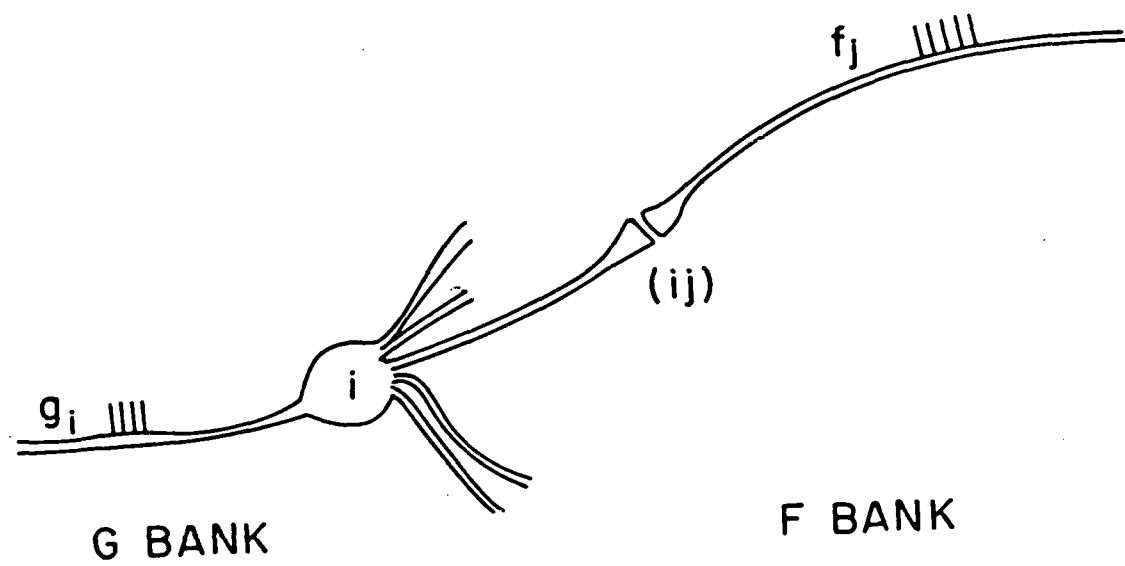


Figure 1.2

Although the firing rate of a neuron depends in a complex and non-linear fashion on the presynaptic potentials, there is usually a reasonably well-defined linear region. We focus our attention on the region above threshold and below saturation for which the firing rate g_i of neuron i in G_i is mapped from the firing rates of all of the neurons f_j in F by

$$g_i = \sum_{j=1}^N A_{ij} f_j \quad (1.2)$$

In doing this we are regarding as important average firing rates, and time averages of the instantaneous signals in a neuron (or perhaps a small population of neurons). Further, we are using the known integrative properties of neurons.

We may then regard $[A_{ij}]$ (the synaptic strengths of the N^2 ideal junctions) as a matrix or a mapping which takes us from a vector in the F space to one in the G space. This maps the neural activities $f = (f_1, f_2, \dots, f_N)$ in the F space into the neural activities $g = (g_1, \dots, g_N)$ in the G space and can be written in the compact form

$$g = Af \quad (1.3)$$

We propose that it is in modifiable mappings of the type A that the experience and memory of the system are stored. In contrast with machine memory, which is (at present) local (an event stored in a specific place) and addressable by locality (requiring some equivalent of indices and files), animal memory is likely to be distributed and addressable by content or by association. In addition for animals there need be no clear separation between memory and "logic."

It is convenient to write the mapping A in the basis of vectors the system has experienced:

$$A = \sum_{\mu\nu} c_{\mu\nu} g^\mu \times f^\nu \quad (1.4)$$

Here g^μ and f^ν are output and input patterns of neural activity while the $c_{\mu\nu}$ are coefficients reflecting the degree of connection between various inputs and outputs. The symbol \times represents the "outer" product between the input and output vectors. Although (1.4) is a well-known mathematical form, its meaning as a mapping among neurons deserves some discussion. The ij th element of A gives the strength of the ideal junction between the incoming neuron j in the F bank and the outgoing neuron i in the G bank (Figure 1.2).

Thus, if only f_j is nonzero, g_i , the firing rate of the i^{th} output neuron, is

$$g_i = A_{ij} f_j \quad (1.5)$$

Since

$$A_{ij} = \sum_{\mu\nu} c_{\mu\nu} g_{\mu}^{\nu} f_{\nu} \quad (1.6)$$

the ij th junction strength is composed of a sum of the entire experience of the system as reflected in firing rates of the neurons connected to this junction. Each experience or association ($\mu\nu$), however, is stored over the entire array of $N \times N$ junctions. This is the essential meaning of a distributed memory: each event is stored over a large portion of the system, while at any particular local point many events are superimposed.

We show below that the nonlocal mapping A can serve in a highly precise fashion as a memory that is content addressable and in which "logic" is a result of association and an outcome of the nature of the memory itself.

Recognition and Recollection

The fundamental problem posed by a distributed memory is the address and accuracy of recall of the stored patterns. Consider first the "diagonal" portion of A ,

$$(A)_{\text{diagonal}} \equiv \mathcal{R} \equiv \sum_{\nu} c_{\nu\nu} g^{\nu} \times f^{\nu} \quad (1.7)$$

If we define separated events as those which map into orthogonal vectors, then clearly a recognition matrix composed of K orthogonal vectors f^1, f^2, \dots, f^K ,

$$\mathcal{R} = \sum_{\nu=1}^K c_{\nu\nu} g^{\nu} \times f^{\nu} \quad (1.8)$$

will distinguish between those vectors contained and all vectors separated from (perpendicular to) these. Further, the response of the system to a vector previously recorded is unique and completely accurate:

$$\mathcal{R} f^{\lambda} = c_{\lambda\lambda} g^{\lambda} \quad (1.9)$$

It is important to note the change in sign of ϕ as the response c moves past a threshold θ_M , θ_M , the modification threshold, evolves as a non-linear function of the average cell response \bar{c} .

$$\theta_M = (\bar{c})^2 \quad (2.3)$$

It can be shown (BCM, 1982) that the biphasic character of ϕ is responsible for driving the cell to a state of high selectivity in a 'normal' environment. The non-linear dependence of θ_M on \bar{c} is responsible for the stability of the states of high selectivity. A detailed study of the evolution of the state of a cell, m , for a simple two dimensional signal environment, is simplified by considering the motion of m in the so-called phase-space of the cell (Figure 2.2). This space, also called the state space of the cell is simply a graphical depiction of all of the possible values that the synapses of this cell can take. Then the evolution of the cell constitutes a trajectory in the phase space. It is found that for this environment, there are only four points in phase space for which m is a constant in time. These points, called fixed points, are the only possible asymptotic states of the cell. Further, it is found that only two of the four points, m_1 and m_2 , are locally stable. This means that if the state of the cell is near a stable point it will converge to it asymptotically.

Conversely, if the state of the cell is near one of the unstable fixed points, m_1, m_2 , it will diverge from the point asymptotically, (only if m lies exactly on one of the unstable points will it stay there and be 'fixed'). Finally, it is found that only the points of maximal selectivity with respect to the signal environment are stable asymptotically. The environment of the cell, that is those signals d that are received from the LGN, is defined by the particular rearing condition one is interested in:

NORMAL REARING: The inputs d to the eye are a stochastic sequence of patterns $d_1 \dots d_K$ (these patterns represent the mapping of patterns on the retina to the response of LGN cells). It is assumed that under conditions of normal rearing, certain patterns such as edges are a repeated part of the environment. These patterns may be corrupted by noise, but they are statistically independent. Such an environment might be expressed by choosing noisy patterns from a closed one-parameter curve in the pattern space. The environment then would have the mathematical property of being circular.

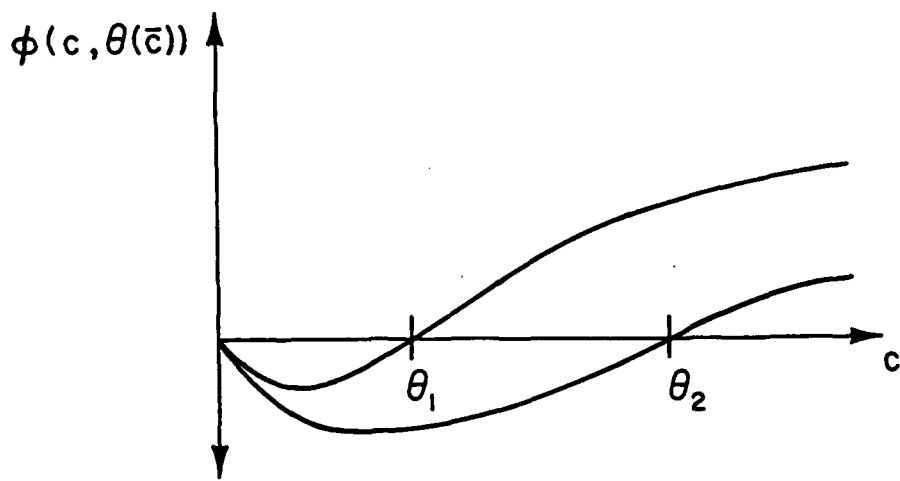


Figure 2.1. The ϕ function.

We do not claim and it is not necessary that all neurons in visual cortex be so modifiable. Nor is it necessary that modifiable neurons are especially important in producing the architecture of visual cortex. It is our hope that the general form of modifiability we require to construct distributed mappings manifests itself for at least some cells of visual cortex that are accessible to experiment. We thus make the conservative assumption that biological mechanisms, once established will manifest themselves in more or less similar forms in different regions. If this is the case, modifiable individual neurons in visual cortex can provide evidence for such modification more generally.

b) Modification of Cortical Synapses: Global and Local Variables

We now consider a single neuron that receives afferent fibers from each eye, forming synapses m_l and m_r . These fibers carry the lateral geniculate nucleus (LGN) signals d_l and d_r . Then if we assume that the cell performs spatiotemporal integration on these signals then the response of the cell is given by

$$c(t) = m_l(t) \cdot d_l(t) + m_r(t) \cdot d_r(t) \quad (2.1)$$

It is likely that the real output of the cell is a sigmoidal function of c , however, we assume that the cell will fire in the linear region in general. We propose that the modification of the left and right synaptic states of the cell, m_l and m_r , follow a generalization of the Hebbian rule:

$$\dot{m}_j = \phi(c, \bar{c}; X, Y, Z, \dots) d_j \quad (2.2)$$

Thus m_j evolves as a function of several types of variables. d_j is said to be a local variable since the signal is locally available to the j th junction m_j . The variable c is a quasi-local variable since the response of the cell is not locally available to the synapse but is physically connected to it through the cell body. For this information to be available at the j th synapse, some form of internal communication between parts of the cell is necessary. \bar{c} is a time averaged quasi-local variable. X, Y, Z are designated global variables that are intended to represent information that is present in a similar fashion at all or a large number of the cortical neurons (such as the presence or absence of a neuromodulator such as norepinephrine). The form of the function ϕ is shown below.

Orientation selectivity develops and extends to all visual cells in area 17 if the animal is reared, and behaves freely, in a normal visual environment (NR): complete "specification" and normal binocularity (about 80% of responsive cells) are reached at about 6 weeks of age (Fregnac and Imbert, 1978). However, if the animal is reared in total darkness from birth to the age of 6 weeks (DR), none or few orientation selective cells are then recorded (from 0 to 15% depending on the authors and the classification criteria); however, the distribution of ocular dominance seems unaffected (Blakemore and Mitchell, 1973; Imbert and Buisseret, 1975; Blakemore and Van Sluyters, 1975; Buisseret and Imbert, 1976; Leventhal and Hirsch, 1980; Fregnac and Imbert, 1978). In animals whose eyelids have been sutured at birth, and which are thus binocularly deprived of pattern vision (BD), a somewhat higher proportion (from 12 to 50%) of the visually excitable cells are still orientation selective at 6 weeks (and even beyond 24 months of age) and the proportion of binocular cells is less than normal (Wiesel and Hubel, 1965; Blakemore and Van Sluyters, 1975; Kratz and Spear, 1976; Leventhal and Hirsch, 1977; Watkins, et al., 1978).

Of all visual deprivation paradigms, putting one eye in a competitive advantage over the other has probably the most striking consequences. If monocular lid-suture (MD) is performed during a "critical" period (ranging from about 3 weeks to about 12 weeks), there is a rapid loss of binocularity to the profit of the open eye (Wiesel and Hubel, 1963, 1965). At this stage, opening the closed eye and closing the experienced one may result in a complete reversal of ocular dominance (Blakemore and Van Sluyters, 1974). A disruption of binocularity that does not favor one of the eyes may be obtained, for example, by provoking an artificial strabismus (Hubel and Wiesel, 1965) or by an alternating monocular occlusion, which gives both eyes an equal amount of visual stimulation (Blakemore, 1976). In what follows, we call this uncorrelated rearing (UR).

These results seem to us to provide direct evidence for the modifiability of the response of single cells in the cortex of a higher mammal according to its visual experience. Depending on whether or not patterned visual information is part of the animal's experience, the specificity of the response of cortical neurons varies widely. Specificity increases with normal patterned experience. Deprived of normal patterned information (dark-reared or lid-sutured at birth, for example) specificity decreases. Further, even a short exposure to patterned information after six weeks of dark-rearing can reverse the loss of specificity and produce an almost normal distribution of cells.

II. SINGLE CELL THEORY

a) Summary of Related Visual Cortex Experimental Data

The discussion above leads to a central issue: what is the principle of local organization that, acting in a large network, can produce the observed complex behavior of higher mental processes. There is no need to assume that such a mechanism -- believed to involve synaptic modification -- operates in exactly the same manner in all portions of the nervous system or in all animals. However, one would hope that certain fundamental similarities exist so that a detailed analysis of the properties of this mechanism in one preparation would lead to some conclusions that are generally applicable. We are interested in visual cortex because the vast amount of experimental work done in this area of the brain -- particularly area 17 of cat and monkey--strongly indicate that one is observing a process of synaptic modification dependent on the information locally and globally available to the cortical cells.

Experimental work of the last generation, beginning with the path-breaking work of Hubel and Wiesel (1959, 1962), has shown that there exist cells in visual cortex (areas 17, 18, and 19) of the adult cat that respond in a precise and highly tuned fashion to external patterns--in particular bars or edges of given orientation and moving in a given direction. Much further work (Blakemore and Cooper, 1970; Blakemore and Mitchell, 1973; Hirsch and Spinelli, 1971; Pettigrew and Freeman, 1973) has been taken to indicate that the number and response characteristics of such cortical cells can be modified. It has been observed in particular (Imbert and Buisseret, 1975; Blakemore and Van Sluyters, 1975; Buisseret and Imbert, 1976; and Fregnac and Imbert, 1977, 1978), that the relative number of cortical cells that are highly specific in their response to visual patterns varies in a very striking way with the visual experience of the animal during the critical period.

Most kittens first open their eyes at the end of the first week after birth. It is not easy to assess whether or not orientation selective cells exist at that time in striate cortex: few cells are visually responsive and the response's main characteristics are generally "sluggishness" and fatigability. However, it is quite generally agreed that as soon as cortical cells are reliably visually stimulated (e.g., at 2 weeks), some are orientation selective, whatever the previous visual experience of the animal (eg. Hubel and Wiesel, 1963; Blakemore and Van Sluyters, 1975; Buisseret and Imbert, 1976; Fregnac and Imbert, 1978).

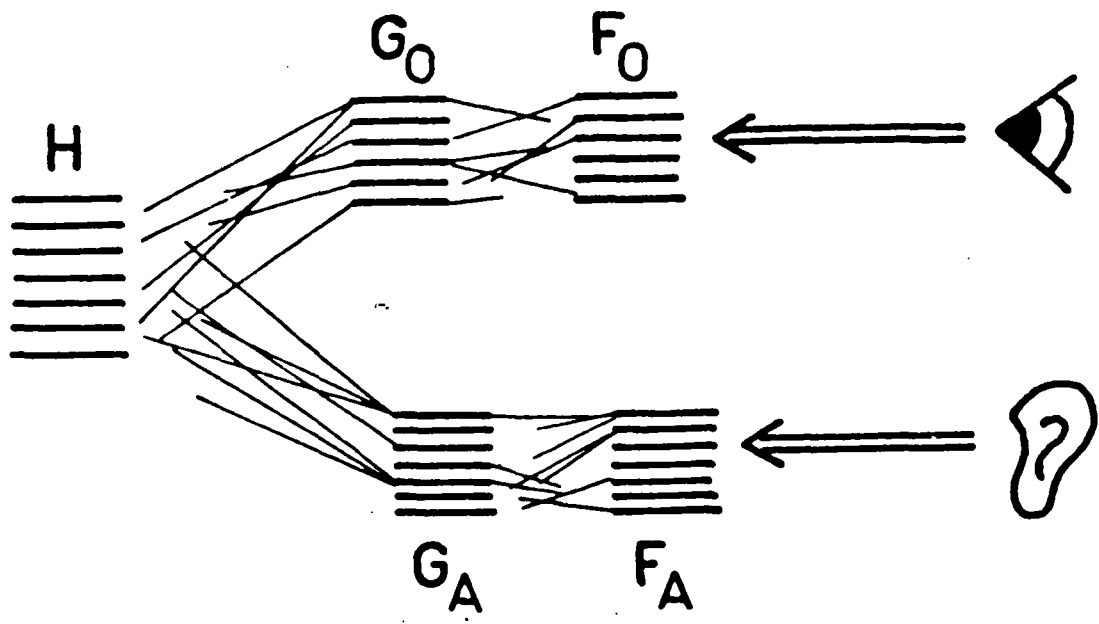


Figure 1.5. A Model Optical-Auditory System
(Cooper, 1973)

By a sequence of mappings of the form above (or by feeding the output of A back to itself) one obtains a fabric of events and connections that is rich as well as suggestive. One easily sees the possibility of a flow of electrical activity influenced both by internal distributed mappings and the external input. This flow is governed not only by direct association coefficients $c_{\mu\nu}$ (which can be explicitly learned) but also by indirect associations ^{$\mu\nu$} due to the overlapping of the mapped events. One can imagine situations arising in which direct access to an event, or a class of events, has been lost while the existence of this event or class of events in A influences the flow of electrical activity.

One problem in making the identifications suggested above is that such systems tend to form excessively large all-encompassing classes. But means have been devised to limit the extent of class formation. In fact such mappings can be made to separate classes as well as to unite them (Kohonen, 1977; Cooper, Liberman, and Oja, 1979 (CLO); Bienenstock, Cooper, and Munro, 1982 (BCM)).

Another problem is a direct consequence of the assumption of the linearity of the system. Any state is generally a superposition of various vectors. Thus one has to find a means by which events--or the entities into which they are mapped--are distinguished from one another.

There are various possibilities; neurons are so non-linear that it is not at all difficult to imagine non-linear or threshold devices that would separate one vector from another. Such separation processes compliment generalization processes in that they bring out the differences in an input environment while generalizing cells tune to the component most common to the constituent stimuli. But the occurrence of a vector in a distributed memory is a set of signals over a large number of neurons each of which is far from threshold. A basic problem, therefore, is how to associate the threshold of a single cell or a group of cells with such a distributed signal. One way this might come about has been shown by Nass and Cooper (1975). Another possibility is the stochastic process recently discussed by Hopfield (1982).

In addition to the appearance of "pontifical" cells or groups of cells, there will be a certain separation of mapped signals due to actual localization of the areas in which these signals occur. For example, optical and auditory signals are subjected to much processing before they actually meet in cortex. It is possible to imagine that identification of optical or auditory signals (as optical or auditory) occurs first from where they appear and their immediate cluster of associations. Connections between an optical and an auditory event might occur as suggested in Figure 1.5.

where n^v varies randomly; f^0 will eventually be recognized more strongly than any particular f^v actually presented. This, of course, is reminiscent of psychological properties called "generalization" or "abstraction." From such a point of view, generalization grows from the loss of detail of individual instances, a trade-off that seems characteristic of distributed systems.

We have here an explicit realization of feature abstraction. This generalizing quality might be described as the result of a built-in directive for inductive logic. The associative memory by its nature takes the step

$$f^0 + n^1, f^0 + n^2 \dots f^0 + n^k \dots f^0 \quad (1.21)$$

which one perhaps attempts to describe in language as passing from particulars: cat1, cat2, cat3 ... to the general: cat.

How fast this step is taken depends on the parameters of the system. By altering these parameters, it is possible to construct mappings which vary from those which retain all particulars to which they are exposed, to those which lose the particulars and retain only common elements--the central vector of any class.

In addition to "errors" of recognition, the associative memory also makes "errors" of association. If, for example, all (or many) of the vectors of the class $\{f^\alpha\}$, defined as a class of vectors not very separated from one another, associate some particular g^β so that the mapping contains terms of the form

$$\sum_{v=1}^K c_{\beta v} g^\beta \times f^v \quad f^v \in \{f^\alpha\} \quad (1.22)$$

with $c_{\beta v} \neq 0$ over much of $v = 1, 2, \dots, K$, then the new event e^{K+1} which maps into f^{K+1} also in the class $\{f^\alpha\}$ will not only be recognized, the inner product $(\int_B f^{K+1}, \int_B f^{K+1})$ being large, but will also associate $g^\beta, \int_B f^{K+1} = c g^\beta + \dots$ as strongly as any of the vectors $f^1 \dots f^K$ explicitly contained in (1.22).

If errors of recognition lead to the process described in language as going from particulars to the general, errors of association might be described as going from particulars to a universal: cat1 meows, cat2 meows ... all cats meow.

Whatever efficacy this inductive process has will depend on the order of the world in which the animal system finds itself. If the world is properly ordered, an animal system that "jumps to conclusions" in the sense above may be better able to adapt and react to the hazards of its environment and thus survive.

It has been shown in the above reference that with a simple form of passive modification, a system generates its own response to incoming patterns in such a way as to construct distributed mappings that can function as memories capable of recognition and association. To a limited extent these mappings can be regarded as internal representations of what has arrived from the outside world. It has further been shown (Nass and Cooper, 1975) that a form of passive modification can result in the formation of feature detectors or threshold-response units which learn to respond to repeated patterns even in the absence of any initial bias. Such units can serve to perform some nonlinear separations mentioned in the previous section. More detailed discussion of the consequences of these modification procedures and the properties of some of the mappings that result is contained in the references cited above. The application of these ideas to visual cortical cells is discussed in section II.

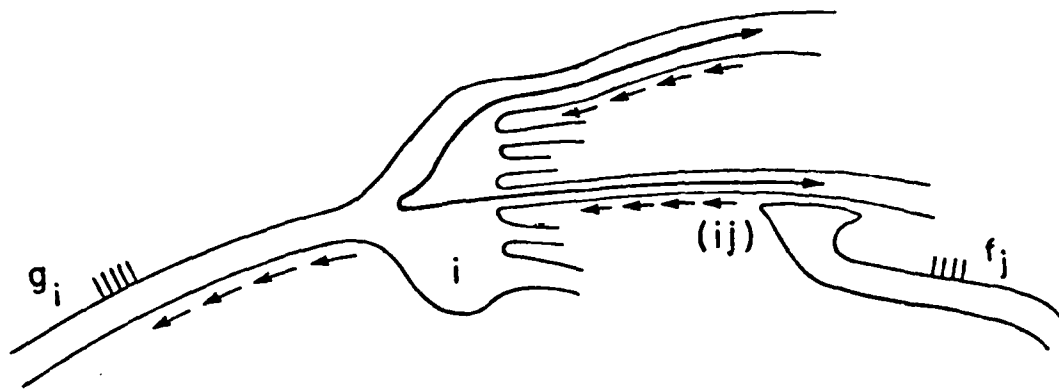
Feature Abstraction

Some networks of neurons must have the ability to extract meaningful information from a broad range of input environments. In the case of sensory input to cortex, for example, the system's range is internally constrained by the response characteristics of the sensory neurons and externally by the nature of the stimulus environment. This stimulus environment depends a great deal upon the nature of the creature's surroundings. Precise statements regarding aspects of environmental structure relevant to mathematical models are given in the next section.

Consider the recognition-association memory (1.12) described above. In actual experience, the events to which the system is exposed are not in general highly separated nor are they independent in a statistical sense. There is no reason, therefore, to expect that all vectors f^v , printed into A according to the modification rule (1.16) would be orthogonal or even very far from one another. Rather it seems likely that often large numbers of these vectors would lie close to one another. Under these circumstances, a distributed memory might be "confused" in the sense that it will respond to new events as if they were old, if the new event is close to an old one. It will "recognize" and "associate" events never, in fact, seen or associated before.

The memory will tend to categorize stimuli on the basis of the past history of the system. For example, suppose a number of vectors in the memory are of the form

$$f^v = f^0 + n^v \quad (1.20)$$



—→ INFORMATION FLOW
- - -→ SIGNAL FLOW

Figure 1.4. Two Point Modification.
(Cooper, 1973)

is a weighted sum over the j components of all mapped signals f^v and the i components of the responses g_u appropriate for recollection or association. Such a form could be obtained by additions with each input f and output g to the element A_{ij} :

$$\delta A_{ij} \sim g_i f_j \quad (1.16)$$

This δA_{ij} is proportional to the product of the differences between the actual and the spontaneous firing rates in the pre- and postsynaptic neurons i and j . [This is one realization of Hebb's form of synaptic modification (Hebb, 1949).] The addition of such changes to A for all associations $g^u \times f^v$ results finally in a mapping with the properties discussed in the previous section.

Synaptic modification dependent on n inputs alone, of the type already directly observed in Aplysia (Kandel, 1976), is sufficient to construct a simple memory-one that distinguishes what has been seen from what has not, but does not easily separate one input from another. To construct a mapping of the form above, however, requires synaptic modification dependent on information that exists at different places on the neuron membrane-what we call two-(or higher-)point modification.

In order that this take place, information must be communicated from, for example, the axon hillock to the synaptic junction to be modified. This implies the existence of a means of internal communication of information within a neuron-in the above example, in a direction opposite to the flow of electrical signals (Cooper, 1973). The junction ij , for example, must have information of the firing rate f_j (which is locally available) as well as the firing rate g_i , which is somewhat removed (Figure 1.4). One possibility could be that the integrated electrical signals from the dendrites produce a chemical or electrical response in the cell body which controls the spiking rate of the axon and at the same time communicates (by backward spiking, for example) to the dendrite ends the information of the integrated slow potential.

For such terms the presentation of event e^v will generate not only g^v (which is equivalent to the recollection of e^v) but also, and perhaps more weakly, g^μ , which should result with the presentation of e^μ .

Thus, the presentation of e^v when $c_{\mu v} \neq 0$ will also initiate this response.

We can thus divide the association matrix A into two parts:

$$A = \sum_{\mu v} c_{\mu v} g^\mu \times f^v = R + A \quad (1.12)$$

where

$$R = (A)_{\text{diagonal}} \equiv \sum_v c_{vv} g^v \times f^v \quad (1.13)$$

and

$$A = (A)_{\text{off-diagonal}} \equiv \sum_{\mu \neq v} c_{\mu v} g^\mu \times f^v \quad (1.14)$$

The $c_{\mu v}$ are then the direct recollection and association coefficients.

Network Modification, Learning

We now ask how a mapping of the type A might be put into the network. The ij th element of A ,

$$A_{ij} = \sum_{\mu v} c_{\mu v} g_{ij}^{\mu v} \quad (1.15)$$

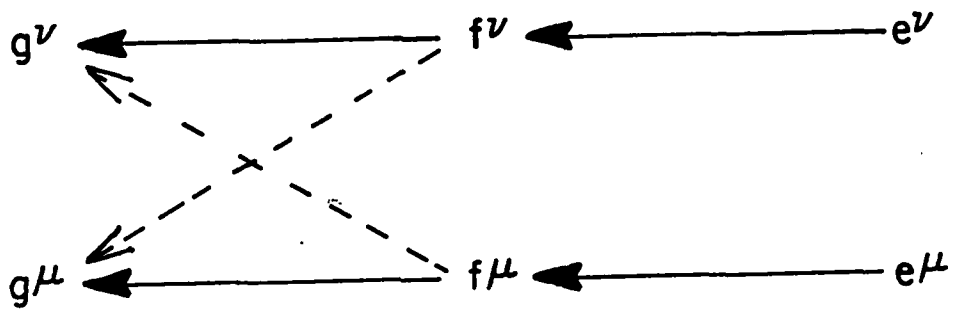


Figure 1.3

In this special situation, the distributed memory is as precise as a localized memory.

ASSOCIATION

If we now take the point of view that presentation of the event e^v which generates the vector f^v is recollected if

$$R f^v = c_{vv} g^v + \text{noise} \quad (1.10)$$

then the off-diagonal terms

$$A \equiv \sum_{\mu \neq v} c_{\mu v} g^\mu \times f^v \quad (1.11)$$

may be interpreted as containing associations between events initially separated from one another (Figure 1.3).

Figure 1.3. An Ideal Association.
(Cooper, 1973)

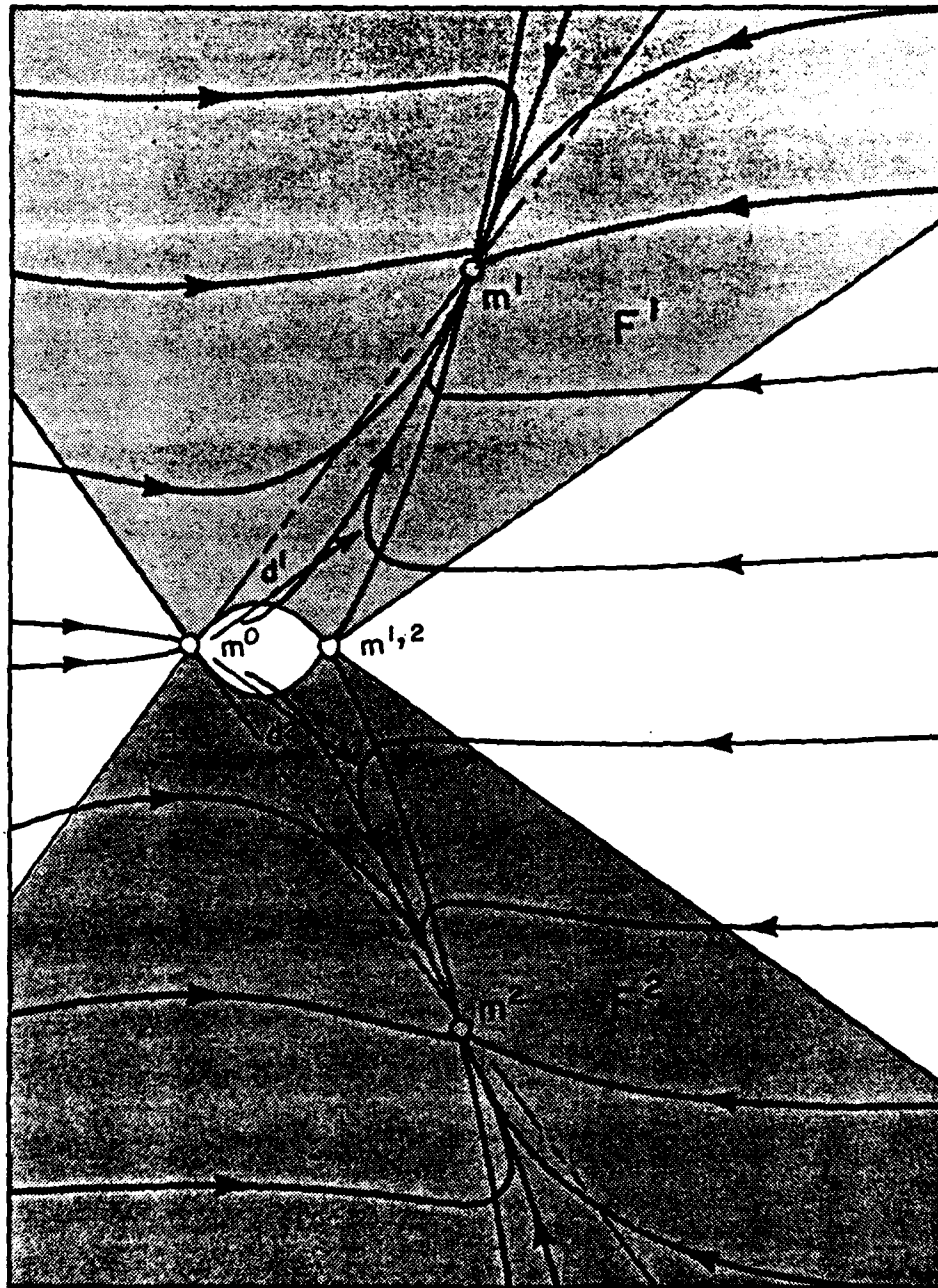


Figure 2.2 The phase plane of equation 2.2. The diagram shows the trajectories of the state of the neuron starting from different initial points. The final state of the system (m_1, m_2) is determined when the trajectory enters the corresponding "trap" (shaded) region (F^1 or F^2). (from Bienenstock, Cooper and Munro, 1982)

RESTRICTED ENVIRONMENT: The pattern space is restricted to one or a few patterns. This is intended to represent the effects of rearing in an environment consisting of perhaps only vertical or horizontal edges.

DARK REARED ENVIRONMENT: The environment consists of just random noise vectors, which represent the dark discharges of retinal cells.

Then various binocular rearing conditions can be represented as follows:

Normal Binocular Rearing (NR): $d_\ell(t), d_r(t)$ for all t , and d_r is circular.

Strabismic Rearing (UR): d_ℓ and d_r are independent identically distributed (i.i.d.): they have the same circular distribution, but no statistical relationship exists between them.

Monocular Deprivation (MD): d_ℓ is circular, d_r is a noise term, $d_r = n$.

Binocular Deprivation (BD): All components of d_ℓ and d_r are i.i.d.: d_ℓ and d_r are uncorrelated noise terms.

The evolution of the cell's response under these rearing conditions is summarized in the following figures:

It is found that simulation of the behavior of the system under these rearing conditions give the following:

NR (Fig. 2.3a) All asymptotic states are selective and binocular, with matching preferred orientations for stimulation through each eye.

BD (Fig. 2.3b) The state of the cell undergoes a random walk in phase space. The two tuning curves therefore undergo random fluctuations that are essentially determined by the second order statistics of the input d . As can be seen from the figure, these fluctuations may result sometimes in a weak orientation preference or unbalanced ocular dominance. However, the system never stays in such states very long; its average state on the long run is perfectly binocular and non-oriented. Moreover, whatever the second order statistics of d and the circular environment in which tuning curves are assessed, a regular unimodal orientation tuning curve is rarely observed, and selectivity has never exceeded 0.6. Thus, we may conclude that orientation selectivity as observed in the NR case (both experimental and theoretical) cannot be obtained from purely random synaptic weights.

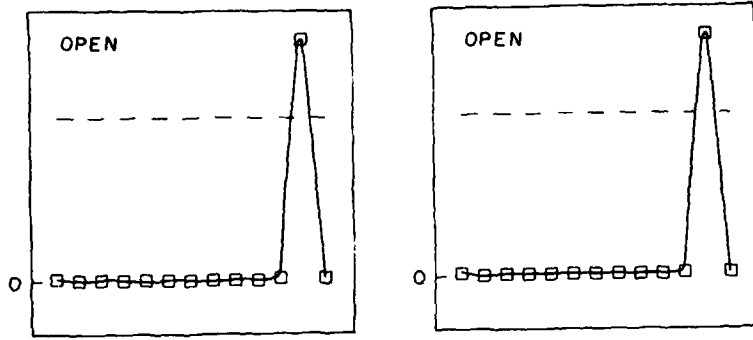
MD and RS (Fig. 2.3c and 2.3d) The only stable asymptotic states are selective and monocular. The cell will converge to one of these states whatever the initial conditions. In particular, this accounts for reverse suture experiments (Blakemore and Van Sluyters, 1974; Movshon, 1976).

UR This rearing environment causes the cell to converge to equilibrium states that are like those found in MD: the state of the cell converges to selective and monocular equilibria. However, unlike MD there are intermediate states that are selective and binocular, with mismatched orientation preferences in the two eyes.

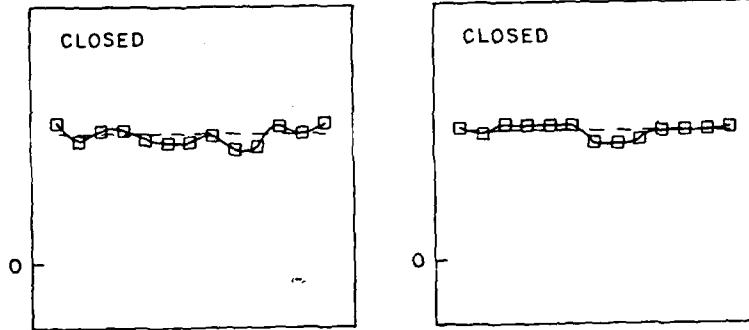
These results are in agreement with the classical experimental data in the domain of visual cortex development. Some predictions of the theory provide tests that will help to confirm some of the details of the theory. As an example we consider the correlation of orientation selectivity and the degree of monocularity of a cortical cell when reared under monocular deprivation.

Monocular Deprivation and Restricted Inputs

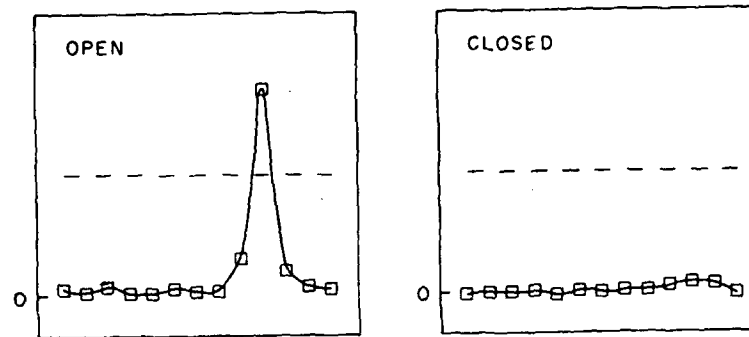
In an environment consisting of a linearly independent distribution of inputs for the open eye and a noisy environment for the closed eye the stable state is selective and monocular. This result is easiest to see if the open eye inputs are divided into three classes. The first class consists of inputs that give responses that are either far above or far below the cell modification threshold. The second and third classes consist of inputs that give a response very near the threshold and very near zero, respectively. For these inputs the value of o is near zero. The inputs to the open eye move the synaptic state of the closed eye afferents randomly for inputs in the first class, away from zero for inputs in the second class, and toward zero for inputs in the third



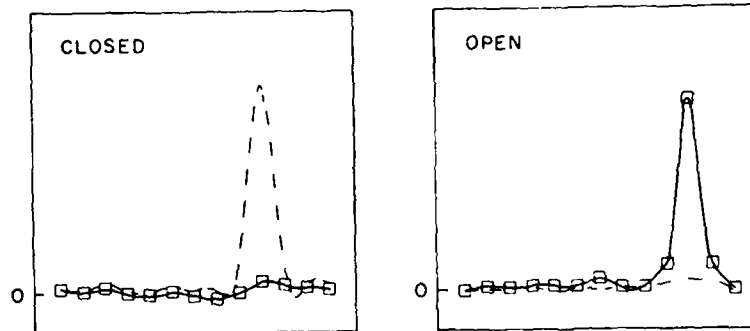
(a) NR



(b) BD



(c) MD



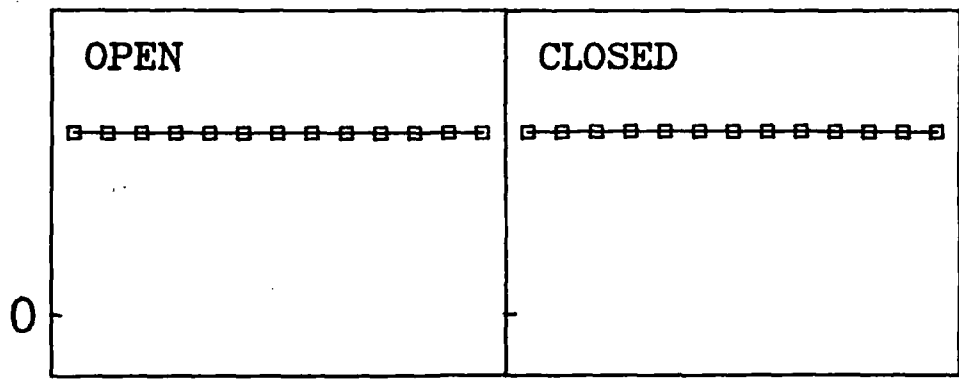
(d) RS

Figure 2.3 Computer simulations of various rearing conditions. Initial (dashed) and final (solid) responses to the two eyes are shown separately (left/right).

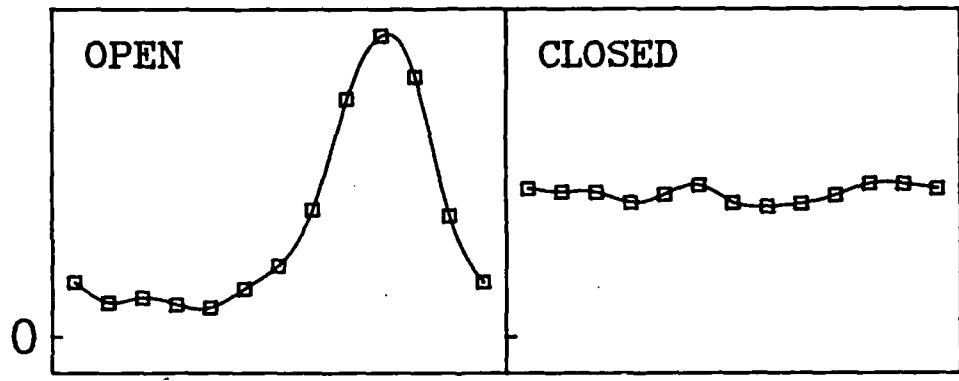
class. In the general case of a circular environment for the open eye, the more selective a cell is to inputs to the open eye, the higher will be the proportion of these inputs belonging to the third class (i.e., the class that drives the closed eye state to zero). This analysis predicts that the more selective a cell is, the more the state of the closed eye will be driven to zero, and thus the cell will be more monocular. Thus, during cell development, pattern selectivity for the open eye is achieved before the cell becomes monocular (Figure 2.4).

Final

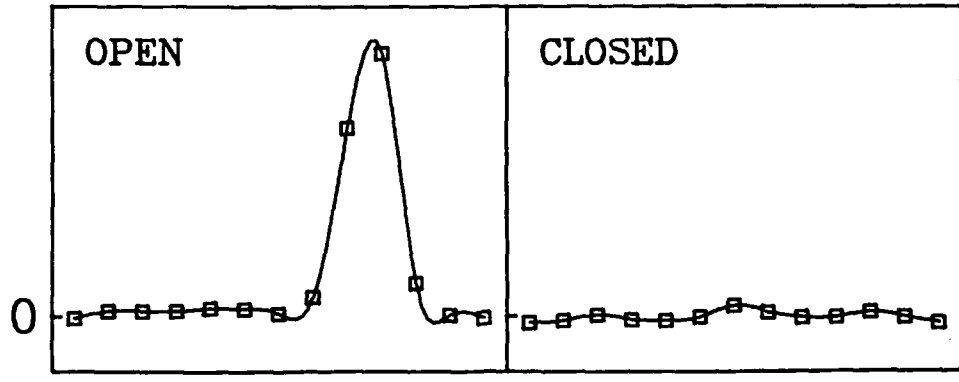
Figure 2.4 Progression of development of selectivity and ocular dominance. Note that selectivity develops for the open eye before the response to the closed eye is driven to zero.



Initial



Intermediate



Final

Figure 2.4

The non-preferred inputs presented to the open eye are a necessary part of the suppression of deprived eye responses. If inputs to the open eye are restricted to preferred patterns, selectivity may still develop but the cell will be less selective to the open eye and less monocular than it would be in a non-restricted environment. We propose to analyze in further detail the difference between normal and restricted monocularly deprived cells; this should be experimentally observable.

III. NETWORK THEORY

a) The Many-Neuron Problem

The development of selectivity in cat visual cortex has been described on the single unit level (BCM, 1982). We would like to retain these results while extending the model to include necessary anatomical detail. It is likely that the cortico-cortical inputs play an important role in the development of neuron selectivity (Creutzfeldt et al., 1974, Sillito, 1975). The development of selectivity might then be viewed as a many-neuron problem. In addition to cortical selectivity, we would like to study the development of global cortical properties such as orientation columns and the precession of preferred patterns across cortex.

The physiology described above has been observed in the simple cells of layer IV of area 17 in cat visual cortex. Although simple cells are found throughout area 17, they appear to be concentrated in layer IV (Kelly and Van Essen, 1974). Layer IV is dominated by two morphological cell types: the smooth and spiny stellate cells (Lund, 1979, Davis and Sterling, 1979). The smooth stellate cell is thought to be inhibitory and the spiny stellate is thought to be excitatory. These cells receive inputs from the LGN that have lateral spreads of up to 2 mm (Albus 1975, Ferster and Le Vay, 1978). Both cell types receive large numbers of synapses whose morphology indicates they are excitatory, and do not originate in the LGN. Finally, both cell types receive synapses whose morphology is that associated with inhibition, (Davis and Sterling, 1979). Thus in layer IV there are excitatory and inhibitory cells which receive direct input from the LGN and, very likely, synapse upon each other. Physiology indicates that these cells can be classified as simple cells.

In the following section we will describe the work conducted on what we call a first-order anatomy. First-order anatomy (FOA) is an attempt to incorporate the necessary elements of the anatomy described briefly above into an extension of the BCM model. In the FOA we clearly define the intracortical transformation of the geniculate signal. Clarification of the model anatomy for single cells allows the construction of a network of cells that can be used to study both the local and global physiology of layer IV.

b) Reduction of the Ideal Synapse

The formal neuron of BCM is separated into two cells, both of which receive the same geniculate afferents (figure 2.5). The geniculocortical synapses are strictly excitatory and thus positive. One of the cells of the circuit (S-cell) is considered to be an inhibitory interneuron and projects to the second cell (P-cell) with synaptic strength s , where $s < 0$. The second cell could be either a locally projecting cell or not, but in this case it is assumed to project out of this circuit.

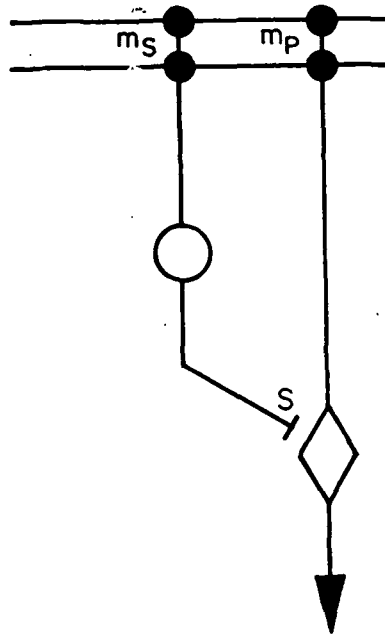


Figure 2.5. A simple circuit in which the ideal synapse is replaced with explicit intracortical inhibition. Both cells receive the same geniculate signals through the fibers entering at the left. The inhibitory cell (the S-cell represented by the circle) projects through synapse *s* to the second cell (here a P-cell represented with a diamond).

Thus each cell has a response which contains a term due to the afferent from the geniculate

$$c_s^o = m_s \cdot d \quad (2.4)$$

$$c_p^o = m_p \cdot d, \quad (2.5)$$

where m_i is a vector with positive components. In this circuit, the S-cell has no other afferents, and thus this is the total response of this cell. The second cell, here referred to as a P-cell, has a total response that is a sum of the geniculate term and the cortical term due to the inhibitory afferent from the S-cell. Then the total response of the P-cell is given by

$$c_p = c_p^o + s c_s^o \quad (2.6)$$

where since the S-cell is an inhibitory cell we have $s < 0$. This may be rewritten as

$$\begin{aligned} c_p &= m_p \cdot d + s(m_s \cdot d) \\ &= (m_p + s m_s) \cdot d. \end{aligned} \quad (2.7)$$

We now define the total state of the P-cell: $a_p = m_p + s m_s$. The state of the cell is now composed of two terms, the **excitatory geniculate synapses and the intracortical synapses that may contain both excitatory and inhibitory terms but are only inhibitory in this circuit.** Then as in BCM, we assume a linear transfer function for the cells,

$$c_s = m_s \cdot d \quad (2.8)$$

$$c_p = a_p \cdot d. \quad (2.9)$$

Under proper assumptions on the time of averaging over the environment, we may exchange summation over the cortical inputs and integration over the environment in the calculation of the average response of the P-cell, c_p . Thus

$$c_p = c_p^o + sc_s^o \quad (2.10)$$

where $c_i^o = m_i \cdot d. \quad (2.11)$

Then as expected $c_p = a_p \cdot d.$ These assumptions govern all of the response properties of this P-circuit.

The evolution of the circuit response properties is determined by the evolution equations of the synapses. Because the S-cell receives only direct LGN afferents, its asymptotic state will be determined by an equation governing only these synapses. Thus if we assume a form of m_s as in BCM or Munro (1983), this state will evolve independently of the state of the P-cell. We assume that the S-cell has reached an asymptotic state m_s^* . For simplicity we assume the intracortical synapse to have a constant value s . The geniculate afferents to the P-cell evolve as in BCM.

$$\dot{m}_p = \phi(c_p, \bar{c}_p) d \quad (2.12)$$

Then the total state of the P-cell evolves according to

$$\begin{aligned} \dot{a}_p &= \dot{m}_p + \dot{s}m_s + s\dot{m}_s & (2.13) \\ &= \dot{m}_p & \text{(asymptotically).} \end{aligned}$$

Or: $\dot{a}_p = \phi(a_p \cdot d, a_p \cdot \bar{d}) d \quad (2.14)$

We see that within a region in phase space governed by the sum of the geniculate state m_p and the intracortical terms sm_s , the total state of the P-cell evolves exactly as in BCM, (see Figure 2.6). Thus within a restricted region of phase space, it is possible to map the properties of the formal neuron of BCM into those of the P-cell. If the restricted region of phase space includes the maximal selectivity fixed points of the evolution equation for m_p then it can be concluded that the total state of the P-cell will almost surely converge to maximum selectivity. We expect that the results described in the previous section will apply to this circuit. Extensions of this circuit have been studied (Figure 2.7).

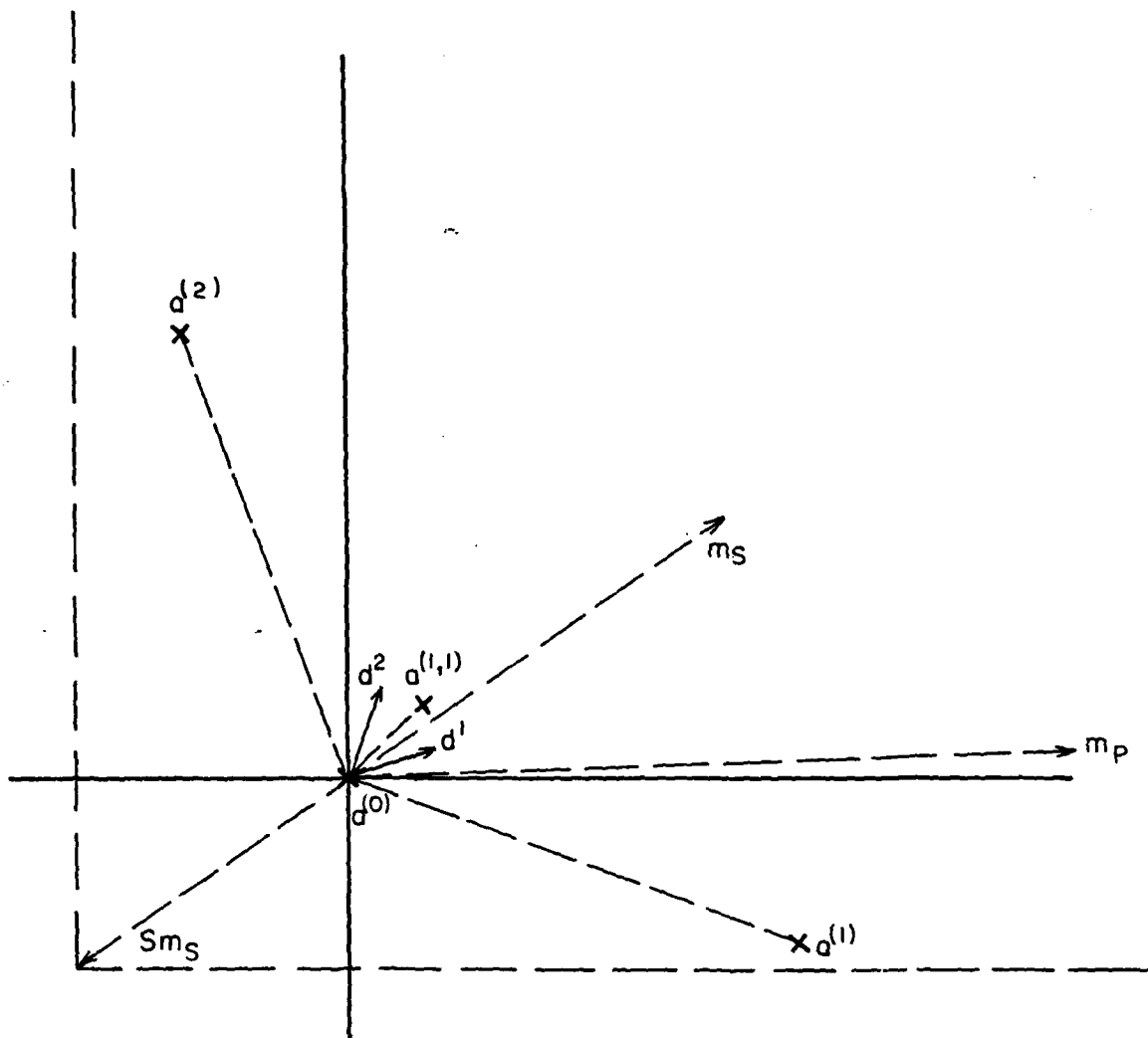


Figure 2.6 The region of phase space that is allowed to the total state of the P-cell is represented by the region to the upper right of the dotted axes. Also depicted are the geniculate states of the S-cell, m_s , m_p . The total state of the P-cell is then the sum of the vectors m_p , sm_s .

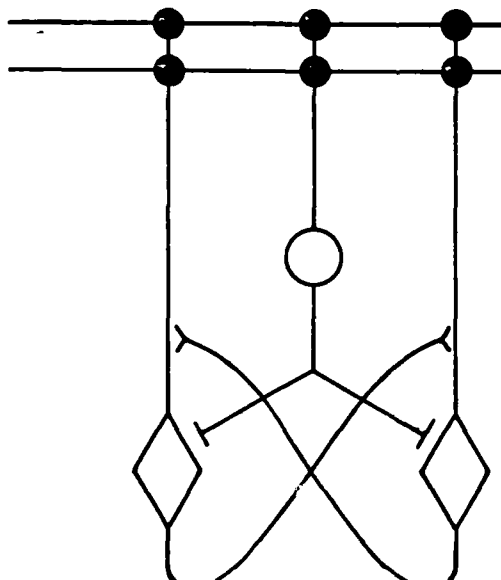
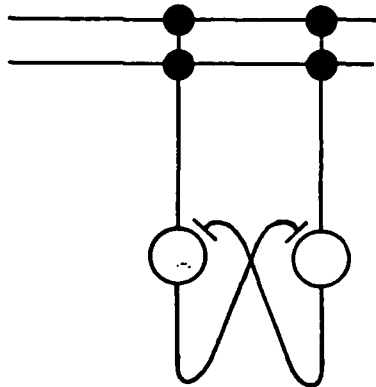


Figure 2.7. Two circuits employed in the study of the asymptotic states of the full network (fig. 2.8). a depicts two S-cells that inhibit each other through synapses s_1 and s_2 . The final states of the two cells are selective, but to different patterns. In b, mutual excitation and tonic inhibition guide the two P-cells to the same selective state.

We have seen that intracortical inhibition allows a cell to reach maximal selectivity when the geniculate state of the cell is non-selective. The degree of selectivity of the inhibition is not important for this effect. We also find (Scofield and Cooper, to be published) that mutual inhibition between selective inhibitory cells causes the cells to prefer different patterns. Intracortical excitation has the effect of causing cells to aggregate in their pattern preference. The mechanism of these results is the transformation of the geniculate phase space caused by the intracortical synapses. This transformation amounts to a vector translation, the net magnitude and direction being the vector sum of the all of the intracortical inputs to a cell.

c) Networks

The cell types described above have been incorporated into a model network that is a first-order representation of the anatomy and physiology of layer IV of cat striate cortex (Figure 2.8). This network is a generalization of the circuit shown in Figure 2.5. Analysis of the network along similar lines shows that under proper conditions on the intracortical synapses, the cells will all converge to states of maximum selectivity with respect to the environment formed by the geniculate signals. In addition, it can be shown that within a region $x \in [-R, R]$, all cells will prefer the same pattern; this defines an orientation column. Under proper conditions, the network will evolve to a state in which orientation preference is a piecewise continuous function of cortical distance.

These results for monocular environments apply to binocular environments when we make the proper extension to binocular networks. We now consider the cells of the network to be binocularly driven. Then the firing rate of the i -th cell at time t is given by

$$c_i(t) = a_i^l(t) \cdot d_l(t) + a_i^r(t) \cdot d_r(t) \quad (2.15)$$

with

$$a_i(t) = m_i(t) + \sum_j k_{ij} m_j(t) \quad (2.16)$$

where k_{ij} represents the intracortical synapse formed between a pre-synaptic cell j and the post-synaptic cell i . k_{ij} may be either positive or negative depending on whether cell j is an excitatory or an inhibitory cell, respectively.

The evolution equations that govern the development of the right and left geniculate states m_i^l, m_i^r are generalizations of that developed for the monocular cells described earlier. The environments of the left and right eyes are identical to those described for various rearing paradigms in BCM. A summary of the main results of simulations under these conditions is given below:

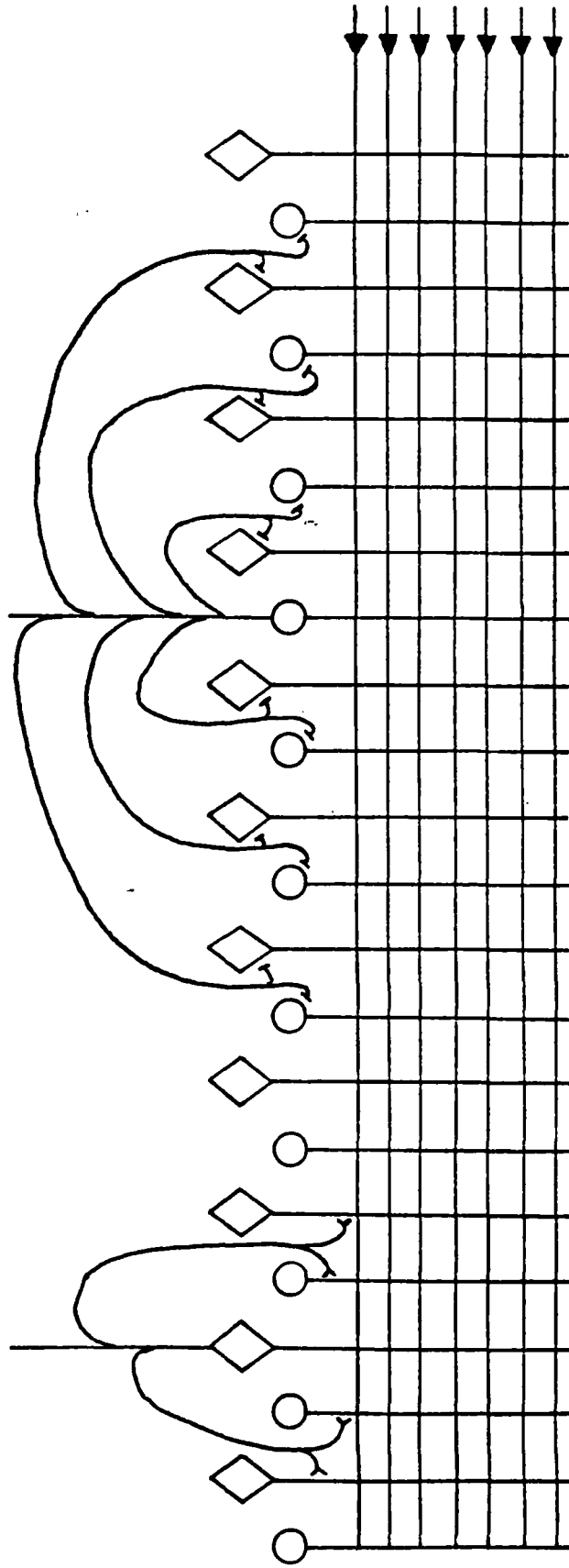


Figure 2.8. The first-order anatomy network. Shown are the two cell types (S-cells represented with circles and P-with diamonds). Geniculate afferents enter at the top of the figure and synapse with all cells in the network at the intersection of the horizontal and vertical fibers. Also shown are intracortical fibers for each cell type. The exact ratio of S-cells to P-cells is not important.

1) In an environment simulating normal binocular rearing, all of the cells in the network become orientation selective and binocular, preferring the same orientation in both eyes. Orientation preference becomes a piecewise continuous function of distance along cortex - with an index of disorder tending to zero. Column size becomes relatively constant over the pattern space.

2) In an environment simulating monocular deprivation, the cells become monocular and orientation selective, independent of the initial state. It is found that under proper conditions on the matrix of intracortical synapses, the geniculate state of the closed eye will be non-zero when the total cell states of the closed eye converge to the neighborhood of the origin. Thus it is found that suppression of the intracortical inhibition allows non-zero responses of the closed eye. This result is in correspondence with reports by Sillito (Sillito et al., 1980) and others of 'masked synapses'. The cells again form columns of orientation preference but the index of disorder is less likely to converge to zero.

3) In an environment simulating binocular deprivation (dark rearing), the cells lose orientation selectivity but remain responsive to stimulation by both eyes.

4) In an environment simulating partially correlated rearing (Paradiso, submitted for publication), the cells of the network become orientation selective and monocular. Ocular dominance columns are formed superimposed upon the orientation columns. Orientation preference becomes a piecewise continuous function of cortical distance and the disorder index tends to zero. Column width in either the orientation or ocular dominance domain is a constant function of the intracortical matrix of synapses.

CONCLUSION

One of the major goals of our program is to bring theory to a sufficiently advanced stage so that it can fruitfully interact with experiment. In this paper we have described how learning on a single cell level can lead to consequences in visual cortex where experiments can be done. We have further described how this work can be extended to networks with some of the gross anatomical features of visual cortex, a network in which inhibitory and excitatory cells receive input from LGN and from each other. The conclusions are similar to those for single cells with explicit further statements concerning the independent effects of excitatory and inhibitory neurons on selectivity and ocular dominance. For example, shutting off inhibitory cells lessens selectivity and alters ocular dominance giving 'masked synapse' effects.

Quantitative tests of progressions such as those shown in Figure 2.4 are in progress in our laboratory. We hope that such experiments can provide detailed comparisons with theory and provide us with a sensitive tool for determining synaptic modification among various classes of neurons--a possible entry to the process by which the nervous system organizes itself.

REFERENCES

- Albus, K.. (1975), *Exp. Brain Res.* 24, 159-179.
- Anderson, J. A..(1970), *Math. Biosciences* 8:137-160.
- Anderson, J. A.,(1972), *Math. Biosciences* 14:197-220.
- Anderson, J. A. and L. N Cooper,(1978), *Pluriscience, Encyclopaedia Universalis France S. A.*, pp. 168-175.
- Bienenstock, E. I., L. N Cooper, and P. W. Munro,(1982), *Jour. of Neurosci.* 2:32-48.
- Blakemore, C. and G. F. Cooper,(1970), *Nature* 228:477-478.
- Blakemore C. and D. E. Mitchell, (1973), *Nature* 241:467.
- Blakemore C. and R. C. Van Sluyters,(1974), *J. Physiol. (London)* 237:195-216.
- Blakemore, C. and R. C. Van Sluyters,(1975), *J. Physiol. (London)* 248:663-716.
- Blakemore, C.,(1976), *J. Physiol.* 261:423-444.
- Brindley, G. S.,(1969), *Proc. R. Soc. Lond. (Biol.)*, 174, 173-191.
- Buisseret, P. and M. Imbert, (1976), *J. Physiol. (London)* 255:511-525.
- Cooper, L. N.(1974), *Proc. Nobel Symposium on Collective Properties of Physical Systems*, ed. by B. Lindquist and S. Lindquist, Academic Press, New York, pp. 252-264.
- Cooper, L. N. P. Munro, and C. Scofield, (to be published).
- Creutzfeldt, O. D., U. Kuhnt, and L. A. Benevento,(1974), *Exp. Brain Res.* 21:251.
- Davis, T. I. and P. Sterling,(1979) *J. Comp. Neurol.* 188:599.
- Ferster, D. and S. LeVay.(1978), *J. Comp. neurol.* 182:923-944.
- Fregnac, Y. and M. Imbert,(1977), *J. Physiol. (Paris)* 6:Vol. 73.
- Fregnac, Y. and M. Imbert,(1978), *J. Physiol. (London)* 278:27-44.
- Hebb, D. O.,(1949), *The Organization of Behavior*. Wiley, New York, p. 62.
- Hirsch, H. V. B. and D. N. Spinelli,(1971), *Exp. Brain Res.* 12:509-527.

- Hubel, D. H. and T. N. Wiesel,(1959), J. Physiol. (London) 148:574-591.
- Hubel, D. H. and T. N. Wiesel,(1962), J. Physiol. (London) 160:106-154.
- Hubel, D. H. and T. N. Wiesel,(1963), J. Neurophysiol. 26:994-1002.
- Hubel, D. H. and T. N. Wiesel,(1965), J. Neurophysiol. 28:1041-1059.
- Imberet, M. and Y. Buisseret,(1975), Exp. Brain Res. 22:2-36.
- Kandel, E. R..(1976), Cellular Basis of Behavior: An Introduction to Behavioral Neurobiology. W. H. Freeman, San Francisco.
- Kelly, J. P.. & D. C. Van Essen,(1974), J. Physiol.. 238, 515--547.
- Kohonen, T..(1977), Associative Memory: A System Theoretic Approach. Springer-Verlag, Berlin.
- Kratz, K. E. and P. D. Spear,(1976), J. Comp. Neurol. 170:141.
- Leventhal, A. G. and H. V. B. Hirsch,(1980), J. Neurophysiol. 43:1111.
- Lund, J. S., Henry, G. H., Macqueen, C. I., & A. R. Harvey, J. Comp. Neur.. 184, 599-618.
- Movshon, J. A..(1976), J. Physiol. (London) 261:125-174.
- Munro, P. W..(1983), Doctoral Thesis, Brown University, Providence, Rhode Island.
- Nass, M. M. and L. N. Cooper,(1975), Biol. Cybernetics 19:1-18.
- Paradiso, M. A., & L. N Cooper, (submitted for publication, Journal of Neuroscience).
- Pettigrew, J. D. and R. D. Freeman,(1973), Science 182:599-601.
- Scofield, C. L., & L. N Cooper, (in preparation).
- Sillito, A. M., J..(1975), Physiol. 250: 304-330.
- Sillito, A. M., Kemp, J. A., & H. Patel. (1980), Brain Research, 41, 1-10.
- Watkins, D. W., J. R. Wilson, and S. M. Sherman. (1978). J. Neurophysiol. 41:322.
- Wiesel, T. N. and D. H. Hubel.(1963), J. Neurophysiol. 26:1003-1017.
- Wiesel, T. N. and D. H. Hubel.(1965), J. Neurophysiol. 28:1029-1040.

END

FILMED

5-85

DTIC