PROGRESS IN THEORETICAL BIOLOGY, VOLUME 5, 1978.
R. ROSEN AND F. SNELL (EDS.)

Communication, Memory, and Development

Stephen Grossberg*

Department of Mathematics,
Boston University,
Boston, Massachusetts

* This work was supported in part by the Advanced Research Projects Agency under ONR Contract No. N00014-76-C-0185.

ISBN 0-12-000000-0

This paper suggests that a variety of developmental mechanisms can be viewed as variations on a small number of organizational principles. These principles represent solutions to environmental problems that all living creatures must surmount. The principles are studied in several examples which all share common statistical and geometrical properties for the collective behavior of cells. The properties achieve efficient parallel processing of patterned information, which thereupon triggers a new level of system organization. This is accomplished by interactions between short-term memory (STM) and long-term memory (LTM) mechanisms, whereby cell sites are switched on and off by excitatory and inhibitory inputs, feedback signals, and couplings that obey nonlinear mass action laws. These laws unify, transform, and extend theoretical results by such authors as Turing (1952), Gustafson and Wolpert (1967), Wolpert (1969), Keller and Segal (1970), Gierer and Meinhardt (1972), Lawrence *et al.* (1972), and Meinhardt and Gierer (1974). The laws suggest that cellular systems have certain functional advantages, such as automatic gain control and its consequences (e.g., self-regulation, adaptation) that have often been omitted from previous models.

The examples considered are tuning of geniculocortical connections in the kitten, sea urchin gastrulation, slime mold aggregation and slug motion, transplantation of the cuticle in *Rhodnius*, regeneration of *Hydra's* heads, cell streaming and division, and emergence of a leadership group in a competitive intergroup interaction. Such interactive mechanisms as adaptation, filtering, contrast enhancement, tuning, nonspecific shunting, cross-correlation, and hysteresis are needed. In particular, the comparison between cortical tuning and gastrula formation suggests a self-corrective feedback mechanism whereby syncytium development can overcome a small genetic error. Analogs of STM and LTM in biochemical systems are noted in terms of antagonistic actions of the cyclic nucleotides cAMP and cGMP; hierarchies of binding strengths in the ions Na^+, K^+, Ca^{2+}, and Mg^{2+}; transmitter properties of norepinephrine, serotonin, and acetylcholine; and protein synthesis.

I. Introduction

Amid the ceaseless flux of our world are many structures that change so slowly or so regularly that we can identify them as objects. We can count ourselves and other living things among these structures. When we contemplate the development and growth of a living thing, at least two general phenomena are a source of wonder. How does one stage of development trigger the next stage? Within a given developmental stage, how does an organism's form remain so stable even as its size increases manyfold? We can also wonder whether the mechanisms for accomplishing these goals are different in different species, or between infant and adult in a given individual? Especially since the genetic code was developed and shown to be shared by all living things, it became natural to hope that developmental mechanisms in different species and in different stages of a particular individual would share many features in common. This faith makes the study of such otherwise relatively unappealing organisms as slime molds (Bonner, 1974), *Hydra* (Wilby and Webster, 1970a,b; Wolpert *et al.*, 1971), sea urchin (Gustafson and Wolpert, 1967), and *Rhodnius* (Lawrence, 1970, 1971, 1972; Lawrence *et al.*, 1972) more exciting and of general interest.

The work that will be summarized here started as a study of short-term memory (STM) in an adult. STM is the type of memory whereby a telephone number is remembered for a few seconds, but can then be forgotten forever. This seems to be a peculiar place to start a study of development. First, it discusses an adult phenomenon. Second, it involves only the nervous system. What does adult STM have to do with development?

The link is provided by the following crucial facts. STM involves the parallel processing of continuously fluctuating patterned information in the presence of noise. This is a very general problem, and it must be solved by *any* system that tries to deal with fluctuating patterns—in particular, by developing systems. In effect, the geometrical and statistical rules that must be satisfied to solve the problem are very much the same no matter how they are interpreted in special cases.

There is another reason why problems involving the nervous system can have solutions that apply to other systems. The nervous system enjoys a property of *universality*. All the data from our senses—both exteroceptive and interoceptive—are translated into a common neural language that ultimately supports a unitary personality and even an idea of God. Our brains are a kind of universal measuring device, and they are so sensitive that they can measure even a few quanta of light. Indeed, the problem of pattern processing can be restated as a problem of measurement, or of communication, between interacting cells, or states. Hence, in retrospect, it should not be too surprising that mechanisms that help a brain to process patterns in its universal language should also be relevant to the systems, both inside and outside the body, with which brains interact.

This article will review some of the main ideas that help to model pattern processing in STM. This model describes a class of networks whose cells interact by mass action laws which describe how unexcited

and excited cellular sites are switched on and off through time. The networks can adapt their total response to fluctuations in total input, suppress noise, contrast-enhance (or sharpen) a pattern of data, and store the enhanced activities indefinitely (Grossberg, 1973). Other properties of network activity include hysteresis, outward peak shifts, slow drifts, peak splits, tuning of contrast by nonspecific arousal, unstable and stable periodic waves, competition by antagonistic cells, and an adaptive resonance that can sustain activity between fields of cells whose activity patterns match, in an appropriate sense (Ellias and Grossberg, 1975; Grossberg and Levine, 1975; Levine and Grossberg, 1976; Grossberg, 1976a,b,c). In a neural context, these properties mimic various processes concerning neural coding—for example, visual illustions, such as line neutralization, tilt aftereffect, angle expansion (Levine and Grossberg, 1976), negative afterimages, spatial frequency adaptation, and the locking of binocular images (Grossberg, 1976c).

The STM models can be combined with adult long-term memory (LTM) mechanisms. LTM is the kind of memory whereby a name can be remembered for a lifetime. The result is a minimal model for the development of feature detectors in mammalian visual cortex (Grossberg, 1976a,b,c). The LTM changes describe the long-term effects of experience on a cell's ability to be activated by prescribed environmental features. This developmental model will be compared with and used to extend modeling efforts on other developing systems, such as on how *Hydra* regenerates a missing head (Gierer and Meinhardt, 1972), how slime molds aggregate in their search for food (Bonner, 1974; Meinhardt and Gierer, 1974), how the folds in the cuticle of *Rhodnius* are determined (Lawrence *et al.*, 1972), and how the sea urchin blastula becomes a gastrula (Gustafson and Wolpert, 1967). We shall hereby suggest that general properties of pattern regulation and self-organization in developing systems are captured by the neural STM and LTM model. The neural STM and LTM model brings a rigorously defined structure, derived from first principles, and with lucid mathematical properties, to the study of these systems. The neural model also captures abstract properties of more macroscopic systems, such as the emergence of leaders in idealized social systems, in which networks of communications or messages exist.

When the model is compared with data concerning the regulation of biochemical reaction rates via extracellular signals, a natural interpretation in terms of cyclic nucleotides, synergistic and antagonistic actions of ions, transmitter substances, and protein synthesis is suggested. The fact that similar organizational principles seem to operate on both microscopic and macroscopic levels illustrates, we believe, the existence

of universal developmental principles which enable the several organizational levels that have emerged via evolution to intercommunicate in a common language and thereby to stabilize each other.

II. Between the Devil and the Deep Blue Sea: Noise versus Saturation

The main problem will be stated abstractly in order to focus on an unpleasant difficulty. Suppose that n different states v_i, $i = 1, 2, \ldots, n$, are given. For definiteness, each v_i can be thought of as a cell, or as a population of cells, whichever is easier for the reader. Suppose that every v_i has a certain number of sites that can be in either an excited or an unexcited state. These sites can be thought of as small patches of cell membrane, small metabolic units, or sources of unitary messages, depending on the situation. Let B be the total number of excitable sites in each v_i, for definiteness. See Fig. 1.

Suppose that each v_i is perturbed by a continuously changing input $I_i(t)$, which will excite a certain number of v_i's sites. Think of the size of I_i at any time t as being the intensity of a coded message to v_i at that time, or in a small time interval $[t - \Delta t, t + \Delta t]$ measured from shortly before time t to shortly after time t. For example, $I_i(t)$ might be the number of unit messages received by v_i at time t, or the intensity of light received by a cell v_i in an idealized retina at time t, or the intensity of a particular feature (such as a line, color, or orientation) in a picture presented to a retina that sends signals to a cortex which contains v_i (Fig. 2).

How can such inputs $I_i(t)$ change through time? Two very different types of changes can be described in terms of the *total* input strength $I(t) = \sum_{k=1}^{n} I_k(t)$ and the *relative* input intensities $\theta_i(t) = I_i(t)/I(t)$ at each v_i. For example, let the v_i represent cells in a retina, and expose the retina to a picture drawn in shades of white, gray, and black. Then

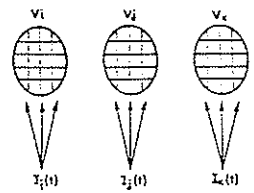


FIG. 1. Continuous inputs $I_i(t)$ perturb populations v_i each with B excitable sites.

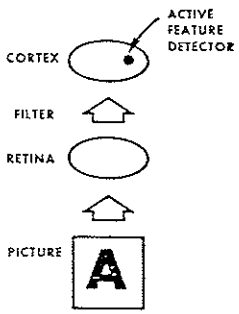


FIG. 2. Each v_i as a feature detector in a cortex perturbed by patterns at a retina.

changes in $I(t)$ describe changes in the background illumination of the picture. The picture itself is characterized by the pattern $\theta = (\theta_1, \theta_2, \ldots, \theta_n)$ of numbers, which do not change through time (Cornsweet, 1970). Thus, it is very important for a system to be able to tell what the pattern weights $\theta(t) = (\theta_1(t), \theta_2(t), \ldots, \theta_n(t))$ are whether or not the total input $I(t)$ fluctuates through time. The weights $\theta(t)$ describe the "relative figure-to-ground" of the inputs at every time t.

III. The Statistics of Switches: Gain Control and Adaptation in On-Center Off-Surround Networks

Can we design a system capable of distinguishing the pattern weights $\theta(t)$ from fluctuations in the background activity $I(t)$? If we were engineers trying to solve this problem, we would start as simply as possible. This will be our strategy. Almost immediately, however, a formidable problem will emerge. By seeing what goes wrong in our trial system, a solution to this problem will be suggested (Grossberg, 1973).

The trial system can be described in two ways. The first way describes its macroscopic properties; the second way describes its statistical interpretation. Four macroscopic properties will be imposed on the functions $x_i(t)$: (*i*) they obey a linear system; (*ii*) they are bounded above, say by B; (*iii*) they return to equilibrium, say 0, after inputs cease; (*iv*) they do not interact. Otherwise expressed, properties (*ii*) and (*iii*) say that the responses $x_i(t)$ vary within a finite dynamical range, from 0 to B, with 0 the passive equilibrium point. In terms of the time

rate of change $\dot{x}_i(t)$ of $x_i(t)$, these properties become

$$\dot{x}_i = -Ax_i + (B - x_i)I_i(t) \tag{1}$$

where $A > 0$, $0 \leq x_i(0) \leq B$, and $i = 1, 2, \ldots, n$. Term $-Ax_i$ describes the linear decay of x_i to 0. Term $(B - x_i)I_i$ is also linear as a function of x_i. It shows that, if $x_i = B$, then I_i does not influence $\dot{x}_i$. Hence $0 \leq x_i(t) \leq B$ for all $t \geq 0$.

System (1) has a natural statistical interpretation. This interpretation describes systems whose elements, or sites, can be in either of two states: for example, "excited" or "unexcited," "on" or "off," "in" or "out," "contracted" or "uncontracted." It describes statistical rules for switching a given population of sites between their two possible states. If B is the total number of excitable sites in any population v_i, then $x_i(t)$ is the number of excited sites, and $B - x_i(t)$ is the number of unexcited sites, at time t. Term $-Ax_i$ says that excited sites become unexcited at rate A. Term $(B - x_i)I_i$ says that unexcited sites become excited at a rate proportional to the input intensity. This is a mass action law. It is also sometimes called a *shunt*. In other words, system (1) describes the switching-on and passive decay of excitation by mass action.

System (1) does not suffice for the following reason. Suppose that $x_i(t)$ approaches a steady state as t increases in response to inputs I_i with fixed pattern weights θ_i and prescribed total activity I. At steady state, $\dot{x}_i = 0$, and (1) can be solved to find

$$x_i = \frac{B\theta_i I}{(A + \theta_i I)} \tag{2}$$

Now keep the θ's fixed, and vary I. In other words, study how (1) processes the *same* pattern θ, given different background activity levels. By (2), as I is increased, all x_i approach B, so that all information about θ is lost, owing to saturation. By contrast, if the system also contains noise, then as I becomes small, the weights θ are lost in the noise. The system processes θ badly both at low and at high I values. What can be done to overcome this dilemma?

We could have guessed that system (1) would fail. This is because each θ_i is defined by a mixture, or interaction, of *all* the inputs I_k, $k = 1, 2, \ldots, n$. There are, however, no interactions between inputs or populations in (1), so that the populations could not possibly compute θ. Property (*iv*) is therefore at fault. Can a system be found that preserves properties (*i*) through (*iii*) and in which interactions occur? The simplest system is readily found. Since $\theta_i = I_i(I_i + \sum_{k \neq i} I_k)^{-1}$, increasing I_i increases

θ_i, whereas increasing any I_k, $k \neq i$, decreases θ_i. In other words, I_i "excites" θ_i whereas all I_k, k $\neq$, "inhibit" θ_i. By expressing this intuition in a mass action system, we find a special case of a competitive interaction pattern that is found throughout the nervous system in some form. Let each I_i excite population v_i and inhibit all populations v_k, $k \neq i$, by mass action. The inputs then form a nonrecurrent (or feedforward) on-center off-surround interaction pattern (Fig. 3), and (1) is replaced by

$$\dot{x}_i = -Ax_i + (B - x_i)I_i - x_i \sum_{k \neq i} I_k \tag{3}$$

System (3) clearly satisfies properties (*i*) through (*iii*). The new term $-x_i \sum_{k \neq i} I_k$ says that excited sites at v_i (which number x_i) are inhibited (note the minus sign!) at a rate proportional to the total inhibitory input (which is a sum of inputs from the off-surround of v_i). Equation (3) is again a mass action law. It can also be profitably described in other ways. In neurological jargon, (3) is a passive membrane equation, with x_i the average membrane potential, and the excitatory and inhibitory inputs control membrane conductance changes that alter this potential (Hodgkin, 1964: Katz, 1966). Using engineering jargon, we can say that the off-surround automatically changes the *gain* of the system, because the inhibitory inputs multiply x_i.

How does inhibitory gain control change the system's steady state? At steady state, $\dot{x}_i = 0$, and Eq. (3) implies

$$x_i = \theta_i \frac{BI}{A + I} \tag{4}$$

In other words, no matter how large I becomes, each x_i is proportional to θ_i; there is no saturation. No matter how small B is, the effective dynamical range of the system is infinite! Furthermore, the total activity $x = \sum_{k=1}^{n} x_k$ satisfies $x = BI(A + I)^{-1} \leq B$: the maximum total activity B is independent of the number n of populations and of the total input intensity I. The off-surround hereby *normalizes*, or *adapts*, the total

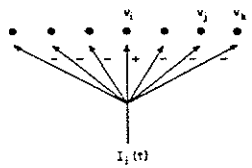


FIG. 3. Feedforward on-center off-surround input geometry.

network response to fluctuations in total input. I have elsewhere suggested that this adaptation in akin to retinal light adaptation, say as studied by Werblin in the mudpuppy retina (Grossberg, 1972d; 1977a; Werblin, 1971).

The above remarks illustrate a general conclusion that will be amply demonstrated below; namely, the statistical laws for switching sites between their excited and unexcited states in response to patterned inputs are expressed by *continuous* and *parallel* interactive mechanisms. Binary laws using serial switching rules, such as those elegantly developed by Kauffman (1971a,b), cannot capture the dynamics of the formal developmental mechanisms described below. Binary serial laws distort both the underlying statistics and the geometry of these formal developmental mechanisms. Our continuous parallel laws are capable of threshold switching behavior under suitable circumstances, as in Section IV. Serial binary laws can sometimes approximate these discrete properties of continuous parallel systems, but binary laws do not describe the mechanisms whereby, or circumstances under which, these properties will emerge. This fact is demonstrated for competitive systems in Grossberg (1978a).

IV. Contrast Enhancement and Short-Term Memory

System (3) cannot remember the pattern θ for long after the inputs are shut off, because each x_i then decays to 0. To maintain activity in the populations v_i after inputs cease, and yet be able to switch it off rapidly if a competing input pattern is delivered, recurrent (or feedback) signals among the populations v_i are needed. We shall see that, while these feedback signals are active, a pattern can reverberate in STM no matter how large the decay rate A is chosen. If an inhibitory signal breaks the reverberation, however, then the STM traces x_i can quickly decay toward equilibrium, from which they can respond to a new input pattern without bias.

How should these feedback signals be distributed? We again have to worry about saturation, so they should be distributed in an on-center off-surround anatomy, as in Fig. 4. Thus, given average activity $x_i(t)$, population v_i will generate a signal $f(x_i(t))$ to be distributed in an on-center off-surround anatomy among all the populations v_k, $k = 1, 2, \ldots, n$. Then (3) is replaced by the *nonlinear* system

$$\dot{x}_i = -Ax_i + (B - x_i)[f(x_i) + I_i] - x_i[\sum_{k \neq i} f(x_k) + J_i] \tag{5}$$

where $i = 1, 2, \ldots, n$. Term $(B - x_i)f(x_i)$ describes how a feedback

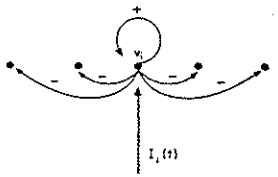


FIG. 4. Recurrent on-center off-surround signal geometry.

signal $f(x_i)$ from v_i to itself excites the unexcited sites $(B - x_i)$ by mass action. The inhibitory term $-x_i \Sigma_{k \neq i} f(x_k)$ describes the switching-off of excitation at v_i by inhibitory signals $f(x_k)$ from all v_k, $k \neq i$. Term I_i is the excitatory input, and term J_i is the inhibitory input; for example, $J_i = \Sigma_{k \neq i} I_k$. Herein, unless otherwise stated, we shall always consider the cases in which interactions occur instantaneously to demonstrate basic properties in a lucid way. Analogous properties hold when the interactions set in slowly. Of course, system oscillations may more readily be generated in the latter situations (cf. Ellias and Grossberg, 1975).

System (5) is certainly a very idealized version of a recurrent on-center off-surround network undergoing shunting interactions. Actually, much more complicated versions of (5) have also been studied, in conjunction with Ellias and Levine. But system (5), by being simple, focuses on an important problem that had to be solved before any further progress could be made. This problem is: How does system (5) know the difference between behaviorally important patterns, which should be stored in STM by the feedback signals, and behaviorally unimportant data, such as noise, which should be suppressed? In system (5), this problem becomes: How should the average signal $f(w)$ be chosen as a function of the average activity w to make the distinction between important and unimportant data? A complete answer to this question has been found (Grossberg, 1973). It is summarized in Table I in terms of the total STM trace $x = \Sigma_{k=1}^{n} x_k$ and the relative STM traces $X_i = x_i x^{-1}$. In particular, we should like to know, after a brief pattern of inputs I_i is delivered to the network, whether $x(t)$ converges to zero (no STM) or to a positive limit that is bounded above by a value that is independent of n and I (normalization)? How do the relative activities X_i change through time? Do they remember the pattern θ_i? Do they enhance certain population activities and suppress others? Do they make all population activities more similar? All of these cases can occur if $f(w)$ is suitably chosen.

TABLE I
INFLUENCE OF SIGNAL FUNCTION ON PATTERN PROCESSING

Signal function	X_i	x
1. Linear: $f(w) = Cw$	Preserves pattern	Amplifies noise
2. Slower-than-linear: $f(w) = wg(w)$, $g(w)$ decreasing	Uniform pattern	Amplifies noise
3. Faster-than-linear: $f(w) = wg(w)$, $g(w)$ increasing	Chooses population with maximal initial data	Suppresses noise; normalizes suprathreshold activity
4. Faster-than-linear-becoming-linear	Quenching threshold exists; contrast-enhances suprathreshold pattern	Suppresses noise; normalizes suprathreshold activity
5. Sigmoid	Quenching threshold exists; contrast-enhances suprathreshold pattern	Suppresses noise; normalizes suprathreshold activity

As Table I shows, if $f(w)$ is linear $[f(w) = Cw]$, then the X_i remember the pattern θ, but x amplifies noise if STM is ever possible. If $f(w)$ is slower-than-linear [for example, $f(w) = w(1 + w)^{-1}$], then all the x_i eventually become equal if STM occurs. Both of these cases are unacceptable because they amplify noise. If $f(w)$ is faster-than-linear [for example, $f(w) = w^2$], then noise is suppressed and STM is normalized. But whenever STM occurs, only the population v_i whose initial activity is maximal is stored in STM. This system therefore makes a choice (binary switch!). By attempting to suppress noise, it so vigorously contrast-enhances the pattern that it throws out the baby with the bath water!

How can we preserve the nice property of noise suppression, which implies that *some* contrast enhancement of the pattern will occur, without throwing out everything but the maximum activity? Since the three cases of linear, slower-than-linear, and faster-than-linear are exhaustive, we clearly have to cut and paste to get a hybrid signal function that has all the desirable properties and none of the bad ones. Such a hybrid signal function $f(w)$ has to be faster-than-linear at small w values, in order to suppress noise. At larger values, it will be chosen (approximately) linear $[f(w) \approx Cw, w \geq D > 0]$, as in Fig. 5a. Then a nice thing happens. A *quenching threshold* (*QT*) exists. As Figs. 5b and 5c depict, if an initial activity $x_i(0)$ is smaller than the *QT*, then the activity of v_i will be suppressed, or quenched, by the reverberation. A population's activity will be stored in STM only if its initial activity exceeds the *QT*. The pattern of suprathreshold activities is contrast-enhanced, as in Fig. 5c. The hybrid signal function in Fig. 5a is thus

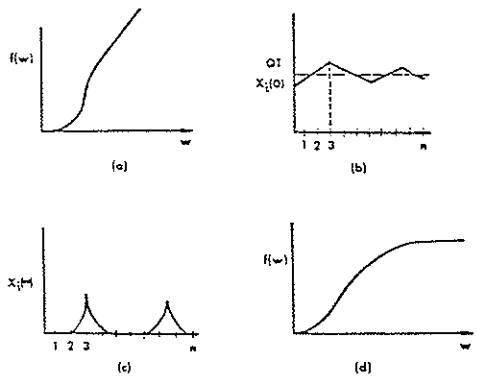


FIG. 5. Partial contrast in STM due to a sigmoid signal function.

capable of *partially* contrast-enhancing a pattern. Its QT determines the cut-off between significant and insignificant data.

Why does a QT exist? Speaking intuitively, we find that the reason is this. Suppose that a pattern of activity starts out in the faster-than-linear range of $f(w)$. The reverberation will start to contrast-enhance the pattern, and would ultimately make a choice if nothing else happened. Simultaneously, however, the total activity starts to be normalized, and drags the pattern into the linear range of $f(w)$. The linear range will preserve any pattern, including the partially contrast-enhanced pattern.

No realistic signal function can be unbounded as w increases. Hence the hybrid signal function has to level off at large activity values, and a sigmoid, or S-shaped, signal function is hereby produced, as in Fig. 5d. The width and slope of the linear range of such a signal function determines its ability to partially contrast-enhance a pattern of data. If either the QT is too small, or the slower-than-linear range is too broad, pathological effects akin to "hallucinations" or "seizures" can occur; the network can then amplify and reverberate noise or other behaviorally unimportant data.

V. The Statistics of Messages: Randomness Prevents Randomness

How can a sigmoid signal function be constructed? The slower-than-linear range is easy to get, since every realistic signal function is

bounded. If a faster-than-linear range can be guaranteed, then an intermediate linear range (perhaps very narrow) follows automatically, by continuity. Suppose, for example, that each cell in a population v_i can elicit signals, or messages, of unit size if the activity of the cell exceeds a prescribed threshold Γ. Suppose that the average threshold of the population is M, and that there are $p_M(\Gamma)$ cells in the population with threshold Γ. Then the average signal $f(w)$ produced by average activity w is a sum, or integral, over all $p_M(\Gamma)$ such that $0 \leq \Gamma < w$; namely,

$$f(w) = \int_0^w p_M(\Gamma)\, d\Gamma \tag{6}$$

If the thresholds Γ are Gaussianly distributed around the average value M, then $f(w)$ is a sigmoid function of w, as Fig. 6 depicts. Thus, a random imperfection in the choice of threshold can produce a sigmoid signal, yet the sigmoid signal acts to suppress noise (that is, randomness!) in the network. In effect, one type of randomness prevents another type of randomness because it is organized by the genetically programmed inhibitory connections of the network.

VI. Competitive Group Interactions

From the vantage point of general communication theory, we have now arrived at some interesting conclusions. Think of the v_i as sources of messages, or even as groups, or communities, that generate messages. Suppose that intragroup messages stimulate an important group activity, but that, because the groups are competing with each other, intergroup messages are designed to inhibit group activity. Let a group's ability to generate messages depend on its activity according to some random rule. Then, depending on the statistics of the rule, the group interaction can remain in balance (like a linear f), can eliminate all activities but the most intense ones (like a faster-than-linear f), can make all activities equally important (like a slower-than-linear f), or can

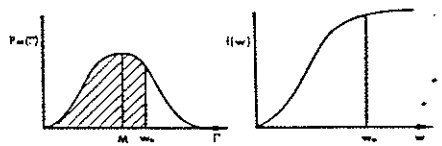


FIG. 6. A Gaussian distribution of signal thresholds produces a sigmoid signal $f(w)$.

do some combination of these. From this perspective, our systems look like a particular kind of competitive group interaction or prey–predator scheme. The abstractness of this picture, wherein particular statistical and geometrical rules determine the overall pattern of activity, is the basis of its generality. For present purposes, it does not matter whether the sources of messages are cells or people. Indeed, it has been proved that *any* competitive system induces an underlying decision scheme (Grossberg, 1978a).

The above example suggests interesting predictions about the establishment of a leader, and the realization of individual potential, in idealized competitive group interactions. Given the sigmoid signals of Fig. 6, a crucial parameter is the mean threshold M at which signals are sent. Suppose that M is large compared to the initial activities $x_i(0)$ of the respective competing groups. Speaking colorfully, we might say that group members are rather inactive and incommunicative. Because of this, the activities lie in the faster-than-linear range of $f(w)$. Recall that $f(w)$ tries to make a choice in this range, if at least one initial activity exceeds the QT. In other words, one group—the "leadership" group—will tend to get very active at the expense of all other groups. If we think of the number of excitable sites B as a kind of group "potential," this means that weak competition between groups which do not communicate easily can actualize the potential of one group at the expense of the others.

By contrast, if the mean threshold M is commensurate with the initial activities $x_i(0)$, then there is a tendency for the activities to be in $f(w)$'s linear range. Consequently, the relative activities of the groups can be approximately preserved, even as the total activity increases by normalization. In this case, competitive group interactions are mutually supportive, and all groups benefit by becoming more active.

If, however, M is small compared with the initial activities $x_i(0)$, then there is a tendency for $f(w)$ to be in its slower-than-linear range, so that all differences in group activity tend to be obliterated. Here the groups communicate too easily, considering their respective activities.

In other words, given fixed initial group activities, by varying the ease with which interpersonal communications occur, one can dramatically change the kind of intergroup activity pattern that will develop. Thus, if a nonspecific signal to all groups amplifies or suppresses the mean threshold M, then this signal can dramatically alter the eventual course of group activities by retuning the intergroup communications (Grossberg, 1973).

What happens if different groups v_i have different potentials B_i, as in

$$\dot{x}_i = -Ax_i + (B_i - x_i)f(x_i) - x_i \sum_{k \neq i} f(x_k) \tag{7}$$

This problem is formally the same as one in which each group has the same potential B, but the relative size of unit messages varies from group to group, as in

$$\dot{y}_i = -Ay_i + (B - y_i)f(C_i y_i) - x_i \sum_{k \neq i} f(C_k y_k) \tag{8}$$

Note that (8) is transformed into (7) by the change of variables $B_i = BC_i$ and $x_i = C_i y_i$.

This problem was studied with Levine in the context of how developmental and attentional biases alter STM processing (Grossberg and Levine, 1975). In effect, if a group has the largest B_i *and* the largest $x_i(0)$, then it tends to take over completely. If several groups all share the same maximal potential B_i and have larger $x_i(0)$ values than other groups, then the other groups become completely inactive—are *masked*—and the groups with maximal potential have their activities transformed as in the equal potential case. However, a group with nonmaximal B_i can suppress other groups if its initial activity is sufficiently large (Fig. 7). In more colorful terms, a tug of war exists between innate potential and initial activity such that a smaller innate potential can be compensated by a larger amount of intergroup competitive activity.

Generalizations of (7) have also been studied, and one finds interesting hysteresis effects, shifts in the populations with maximal activity, slow periodic waves of activity, unexpected emergence of new activity in groups with small initial activity, etc., that are all due to the nature of the competitive interaction. These effects have heretofore been interpreted as visual illusions and other neuropsychological phenomena (Ellias and Grossberg, 1975; Grossberg, 1976c; Levine and Grossberg, 1976), but they also constitute predictions about intergroup interactions in suitable situations. For example, we shall see how the hysteresis effect can be used to discuss a particular stage in the aggregation of a

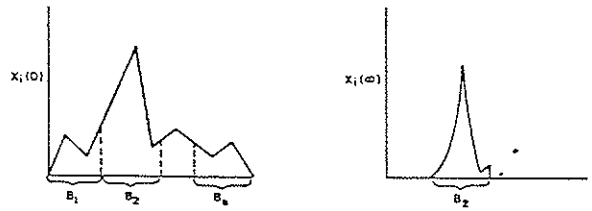


FIG. 7. Competition between number of cell sites B_i and large $x_i(0)$ can mask populations in STM.

slime mold, which is, in a clear sense, a problem about group interactions. Some of these effects are summarized below to illustrate the rich variety of properties that arise when the parameters of shunting reverberating networks are altered.

VII. Hysteresis, Peak Shifts, and Slow Drifts

A central theme unifies the discussion of these generalizations. Table I illustrates a strong tendency in the networks either to amplify noise or to make a choice. In many situations, a compromise between these two extremes is desired—namely, *partial* contrast in STM (Fig. 5c), or the possibility of simultaneously storing many activity levels in STM. For example, let a picture be the input pattern to a network retina. Let the network analyze the picture into component features, such as lines, colors, orientations. Suppose that populations in the network's cortex code particular features by becoming active when their feature is present in the picture. The amount of activity at such a population, or *feature detector*, indicates how intensely the feature is represented in the picture. To make a choice in this cortex means to ignore all but one feature even if many other features are present with almost equal intensity. Partial contrast in STM means that sufficiently weak features are ignored, but all suprathreshold features are stored after relative enhancement of the strongest features.

The special case in (5) achieves partial contrast by using the same sigmoid signal function $f(w)$ to generate both excitatory and inhibitory signals in the network. Are there other ways to achieve partial contrast in STM, say if the excitatory and inhibitory signals functions are not the same? A system in which excitatory signal strength decreases with the distance between populations more rapidly than inhibitory signal strength has this property (Ellias and Grossberg, 1975; Levine and Grossberg, 1976). Such a system describes the common neural situation where there are many feature detectors and some detectors are mutually more closely coupled than others. Such preferred couplings between populations are presumably the anatomical substrate on which innate behavioral generalization gradients are built; for example, color detectors sensitive to similar wavelengths, or line detectors sensitive to similar orientations, might excite each other more than other feature detectors.

A network of this type is described by

$$\dot{x}_i = -Ax_i + (B - x_i)\left[\sum_{k=1}^{n} f(x_k)C_{ki} + I_i\right] - x_i\left[\sum_{k=1}^{n} g(x_k)D_{ki} + J_i\right] \quad (9)$$

where $f(w)$ is the excitatory signal function, $g(w)$ is the inhibitory signal function, and C_{ki} decreases faster as a function of the distance $|k - i|$ from v_k to v_i than does D_{ki} (Fig. 8). In Fig. 8, the excitatory connection strength C_{ki} ("on-center") describes how strongly population v_k can excite v_i, whereas the inhibitory connection strength D_{ki} ("off-surround") describes how strongly population v_k can inhibit v_i. In other words, we consider networks capable of partial contrast in STM and whose feature detectors are connected by nontrivial generalization gradients.

Such a network is capable of *hysteresis*. The following experimental phenomenon, reported by Fender and Julesz (1967), illustrates this concept. Let a different vertical line be shown to each of a person's eyes. Let the two lines be positioned so that they seem to be superimposed. If the two lines are then slowly moved apart on their respective retinas, the person will still see them as one line until a critical separation is reached. Then the perceived line will seem suddenly to split into two lines that will thereafter be seen in their "real" positions. If the two lines are slowly brought back together, they will eventually be seen as one line again, but they will merge at a much smaller distance than the one at which they split apart. This is hysteresis.

A similar phenomenon occurs in (9); it is studied in Levine (1974) and Grossberg (1978b). Hysteresis has also been reported in networks studied by Wilson and Cowan (1973). If an input is presented only to one population, say v_1, and then is slowly separated into two inputs that move away from each other, the network response will have the form shown in Fig. 9a. Speaking intuitively, we find the reason for such hysteresis to be this: When the two lines are close together, the narrowly distributed on-center excitatory signals build up a central focus of excitation. This excitatory focus succeeds for a while in overcoming the effects of inhibition. When the two lines are sufficiently far apart, however, the narrow on-center cannot compete with the broad off-

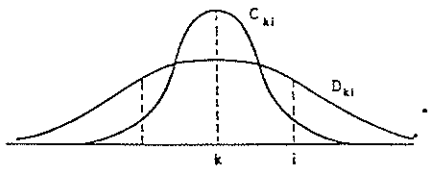


FIG. 8. Excitatory coefficients C_{ki} and inhibitory coefficients D_{ki} from v_k to v_i that depend on distance.

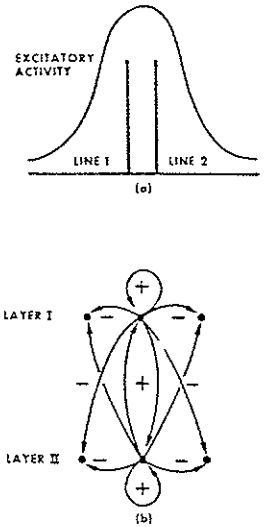


FIG. 9. Hysteresis due to competition between a narrow recurrent on-center and a broad recurrent off-surround.

surround. Inhibitory signals finally overcome the central focus of recurrent excitation, and thereby cause a split of the network's activity into two excitatory peaks. When the two lines are moved together, the inhibitory signals keep the two excitatory peaks distinguishable until the two lines are so close together that excitatory signals in the narrow on-center can overcome them. In effect, network hysteresis is due to a tug of war between a narrow on-center and a broad off-surround of recurrent interactions. It provides a mechanism for "locking together" the responses to two nearby input sources. The same phenomenon also occurs in the laminar network of Fig. 9b: each "eye" sends signals to a different layer. Consequently, one line perturbs layer I, whereas the second line perturbs layer II. The cross-excitation between layers produces a focus of sustained excitation when the two lines are nearly superimposed, just as in the network of Fig. 9a. This reverberation between layer I and layer II is an example of an *adaptive resonance* (Grossberg, 1976c).

The pattern of STM activity can also be shifted by lateral inhibition or

by an asymmetric spatial distribution of cell sites. For example, network (9) is also capable of generating an outward peak shift in the loci of maximal network responses to a pair of spatially separated inputs (Fig. 10a). This shift is due to lateral inhibition; the shift is inward in the absence of inhibition (Levine and Grossberg, 1976). If the common saturation level B in (9) is replaced by a saturation level B_i at each cell v_i, then a drift in the locus of maximal network response can be generated in response to a single fixed input (Fig. 10b). The drift moves toward the direction of increasing B_i. Its velocity becomes slower as the slope of B_i as a function of position v_i becomes flatter. This slow drift toward large B_i values is analogous to the more dramatic contrast enhancement within system (7), wherein the lateral inhibitory signals from each population v_i reach *all* populations v_k, $k \neq i$. Thus the drift is also a masking phenomenon. It helps to explain "normative" effects wherein activity drifts along a structural gradient (Grossberg, 1978b).

VIII. Ratio Processing by Antagonistic Cells

Network populations are sometimes organized into antagonistic pairs, or "dipoles," which mutually inhibit each other (Fig. 11a). A dipole is

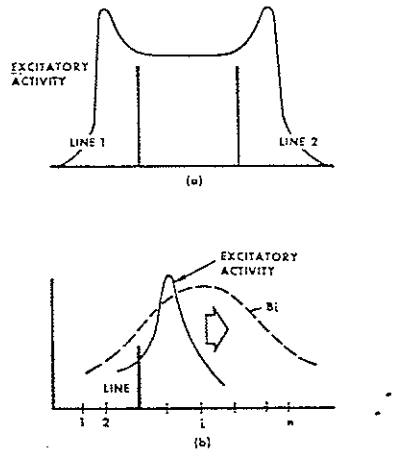
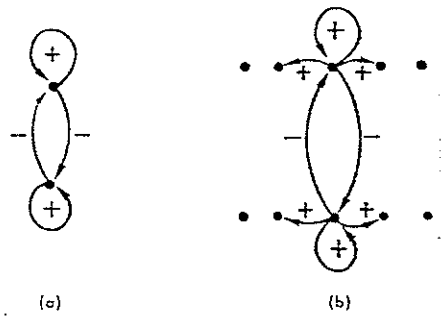


FIG. 10. (a) Outward peak shift due to lateral inhibition. (b) Drift toward larger B_i values.



(a) (b)

FIG. 11. (a) A dipole of mutually inhibitory and self-excitatory populations. (b) A field of dipoles.

simply system (5) with $n = 2$. A motivating example is a mutually antagonistic pair of motor control cells that regulate the contraction of an agonist and antagonist pair of muscles in a limb. When these dipoles are allowed to interact, two parallel fields of self-excitatory and other-inhibitory cells are found (Fig. 11b). Speaking intuitively, one might say that, if an input excites one field in an on-center off-surround configuration, then it excites the complementary field in an off-center on-surround configuration. These networks have the following important property: if a given agonist–antagonist pair is excised from the rest of the network, it can remember its pattern weight. In a motor example, this means that the pair can preserve a given posture in its agonist–antagonist muscles. This fact is illustrated by the following equations for a pair $\{v_1, v_2\}$ of antagonistic cells:

$$\dot{x}_i = -Ax_1 + (B - x_1)(I_1 + x_1) - x_1(I_2 + x_2) \tag{10}$$

and

$$\dot{x}_2 = -Ax_2 + (B - x_2)(I_2 + x_2) - x_2(I_1 + x_1) \tag{11}$$

Equation (10) says that input I_1 excites v_1, whereas I_2 inhibits v_1. Also, v_1 excites itself via the linear signal x_1, whereas v_2 inhibits v_1 via the linear signal x_2. Equation (11) describes a similar mechanism at v_2, with the roles of v_1 and v_2 reversed. The response of the relative activities $X_i = x_i(x_1 + x_2)^{-1}$ to changes in the relative inputs $\theta_i = I_i(I_1 + I_2)^{-1}$ is readily seen to obey the equations

$$\dot{X}_i = BIx^{-1}(\theta_i - X_i), \qquad i = 1, 2 \tag{12}$$

where $x = x_1 + x_2$ and $I = I_1 + I_2$. In other words, X_i approaches θ_i when the input I is turned on. When the input is off, X_i remembers θ_i, since then $\dot{X}_i = 0$. In fact, when the input is off, $x_i \equiv \theta_i(B - A)$, $i = 1, 2$. If the linear signals x_i in (11) and (12) are replaced by sigmoid signals $f(x_i)$, then contrast enhancement of the pattern of θ_i's can occur before STM storage; in particular, a quenching threshold exists (Grossberg, 1973).

IX. Slow Waves and Pacemakers

Another important generalization of (5) expresses the fact that inhibitory cells respond at a finite rate, indeed sometimes much more slowly than excitatory cells. In (5), by contrast, the assumption is tacitly made that the inhibitory cells respond so quickly that their activities can be expressed in terms of excitatory activities. More generally, consider system

$$\dot{x}_i = -Ax_i + (B - Cx_i)\left[\sum_{k=1}^{n} f(x_k)D_{ki} + I_i\right] - x_i\left[\sum_{k=1}^{n} g(y_k)E_{ki} + J_i\right] \tag{13}$$

$$\dot{y}_i = -\hat{A}y_i + (\hat{B} - \hat{C}y_i)\left[\sum_{k=1}^{n} f(x_k)\hat{D}_{ki} + \hat{I}_i\right] - y_i\left[\sum_{k=1}^{n} g(y_k)\hat{E}_{ki} + \hat{J}_i\right] \tag{14}$$

where x_i is the average activity in the ith excitatory population, and y_i is the average activity in the ith inhibitory population (Fig. 12). Analogous systems were studied with Ellias (Ellias and Grossberg, 1975). Note that excitatory cells can excite themselves and inhibitory cells (via connections D_{ki} and $\hat{D}_{ki}$), whereas inhibitory cells can inhibit themselves and excitatory cells (via connections E_{ki} and $\hat{E}_{ki}$). An important new phenomenon occurs in these systems. They are capable of generating sustained oscillations of STM activity, whose frequency can be dramatically changed as a function of system parameters. Suppose, for example, that certain populations are initially more active than other populations, and that a small uniformly distributed background input is maintained at all populations. Figures 13a and 13b illustrate two examples of this situation. In the former case, traveling waves of activity can be periodically generated in one direction (Fig. 13a). In the latter case, traveling waves of activity can be periodically generated in both directions (Fig. 13b).

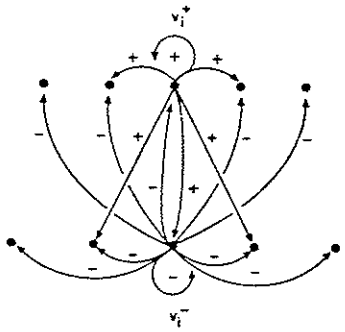


FIG. 12. Interaction of excitatory potentials x_i and inhibitory potentials y_i in a recurrent on-center off-surround geometry.

In other words, an initial imbalance in activity can trigger sustained periodic pulses of activity. Since these pulses emanate from a fixed population, this population seems to be a *pacemaker* of sorts; it seems that the pacemaker population actively and persistently produces something that the other populations do not produce, and this something maintains the traveling waves. This impression is false, however. The *interactions* among the populations produce the sustained waves no matter what brief, perhaps even random event gave the pacemaker population its initially larger activity.

These traveling waves were found while Ellias and I were searching for something else, namely sustained oscillations that could store a nonuniform spatial pattern in STM. Such an STM reverberation would have to maintain the same ordering of activities through time; for example, $x_1(t) \leq x_2(t) \leq \cdots \leq x_{n-1}(t) \leq x_n(t)$ for all $t \geq 0$. Otherwise there would be no record through time of which populations are most important in the coding by the network's feature detectors of a given external pattern of inputs. Such order-preserving STM oscillations were never found after the external input pattern was turned off. Hence it emerged that, at least in a large class of networks, sustained oscillations cannot store spatial patterns in STM unless they are supplemented by an input source. Instead, network parameters must be chosen so that a population's activity approaches a definite value as time goes on, as in Fig. 5. These "asymptotic steady states," $x_i(\infty)$, of network activity are produced by the "fast" recurrent network interactions rather than by

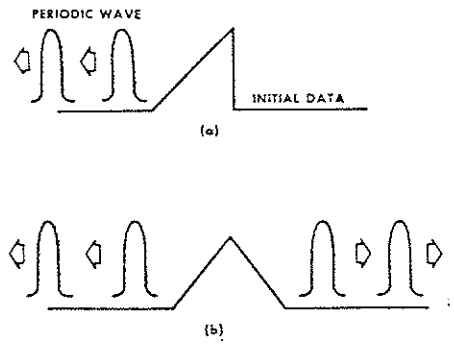


FIG. 13. Periodic traveling waves in a finite network generated by nonuniform initial data.

"slow" waves of activity. One might conjecture, however, since the asymptotic steady states are "infinitely fast" oscillations, that there exist systems in which sufficiently fast oscillations can preserve order in STM. In any case, a sustained pattern of signals from a different part of the network can serve as an "expectation" (Grossberg, 1975, 1976c). If the external input pattern matches the expected pattern, then order-preserving sustained oscillations can be generated (Ellias and Grossberg, 1975, Section 18). This idea occurs in a model of olfactory coding by the prepyriform cortex (Grossberg, 1976c).

Two kinds of variation in network parameters can change it from an asymptotic steady state to a traveling wave mode of activity: either a relative slowing down of the response rate of inhibitory potentials [small $\hat{A}$ in (14)], or a relative broadening of a population's on-center relative to its off-surround. In fact, the network's steady-state response to a prescribed input figure, such as a rectangle, can vary dramatically either as the dimensions of the figure change, such as width and height, or as the network parameters change. For example, in response to a prescribed rectangle, the network might merely contrast-enhance its boundary (Fig. 14a). As rectangle width or height increases, given fixed on-center and off-surround coefficients, extra interior peaks of excitation can be produced as a result of disinhibition at these loci (Fig. 14b). Alternatively, successively decreasing the width of the off-surround, given a fixed input rectangle, can also produce interior disinhibitory peaks, followed by the generation of periodic traveling waves.

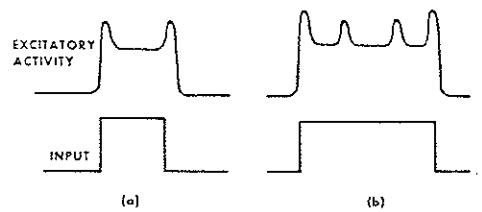


FIG. 14. Different asymptotic patterns or even traveling waves generated in response to different input patterns.

X. Long-Term Memory

Now we can begin to build a model related to development. Such experimentalists as Hubel and Wiesel (1970), Hirsch and Spinelli (1971), and Blakemore and Cooper (1970) have shown that the early visual experience of a kitten can change the responsiveness of cells in the kitten's visual cortex. Certain visual features that are not experienced lose their ability to activate cortical cells, whereas features that are experienced elicit more vigorous responses from prescribed cortical cells. These enduring changes can be modeled by a mechanism of LTM that is activated by input patterns after these patterns are transformed by systems such as (3) and (5).

This LTM mechanism has an interesting history that can be thought of as the gradual synthesis and explication of two great streams of experimental activity. After Pavlov (1927) did his remarkable studies of classical conditioning, and Ramón y Cajal (1909–1911) proved that the structural unit in the brain is the nerve cell, it became natural to expect that classical conditioning can occur in individual nerves, or at worst in populations of nerves. Classical conditioning is illustrated by the experiment in which repetitively presenting an indifferent conditioned stimulus (CS), such as a ringing bell, shortly before an unconditioned stimulus (UCS), such as food, enables the CS to elicit responses, such as salivation, which previously were under the control of the UCS but not the CS. Given that the structural unit in the brain is a nerve cell, should not repetitively presenting a CS, which excites a nerve cell v_1, shortly before a UCS, which excites a nerve cell v_2, strengthen the signals that v_1 can elicit at v_2 in response to an input of unit size? D. O. Hebb utilized this concept in his influential book of 1949, but still the idea remained purely verbal, despite the obvious need to translate it into a

precise language wherein the interactions among millions or billions of nerves could clearly be analyzed. Prior to Hebb's psychophysiological book, various other psychologists were also working out theories of how associative learning occurs, notably E. R. Guthrie (Mueller and Schoenfeld, 1954) and C. L. Hull (Koch, 1954). A generation of mathematical psychophysiological models began to emerge around 1960. Some of these models emphasize the all-or-nothing properties of nerve signals, or the presumed random connections between nerves, or the periodic behavior of nerves. The model that we need blends together deterministic and statistical aspects of nervous activity in a single individual as it occurs during the continuous flow of time.

We need the concept of a *trainable synaptic strength*, or *LTM trace*, z_{ij} in the pathway from v_i to v_j. The LTM trace z_{ij} will record how often v_i-and-then-v_j have been paired in the past, and it will influence the size of a signal received at v_j from v_i using this information. More precisely, z_{ij} computes a time average of the product of signals S_{ij} from v_i to v_j with STM traces x_j at v_j; namely, the time rate of change $\dot{z}_{ij}$ of z_{ij} satisfies an equation of the form

$$\dot{z}_{ij} = -E_{ij}z_{ij} + S_{ij}x_j \tag{15}$$

where E_{ij} is the rate of time averaging by z_{ij} of $S_{ij}x_j$. Also z_{ij} multiplicatively gates, or shunts, signals from v_i on their way to v_j; namely, the total signal to v_j from all v_i is proportional to

$$\sum_k S_{kj}z_{kj} \tag{16}$$

Using this concept, we can construct a model in which idealized cortical cells V_2 can learn to respond to prescribed, but otherwise arbitrary, pictures presented to an idealized lateral geniculate V_1 (Grossberg, 1976a,b,c). The model is summarized in its simplest form in Fig. 15. The main idea is this. When a picture is presented to the geniculate, it is characterized by the relative input intensities at all the geniculate cells. For example, if the input in the ith geniculate population v_{1i} is $I_i(t)$, then $I_i(t) = \theta_i I(t)$, where θ_i is the fixed relative intensity and $I(t)$ is the total intensity. The geniculate can normalize its inputs, as in (4). Suppose for simplicity that the normalized activity at v_{1i} is θ_i rather than $\theta_i BI(A + I)^{-1}$. By (16), the total signal from v_{1i} to the jth cortical population v_{2j} is

$$S_j = \sum_k \theta_k z_{kj} \tag{17}$$

Such an input is generated at each v_{2j}. In this fashion, the picture creates a pattern of activity across the cortex V_2. Assume that V_2 contrast-enhances its input patterns before it stores them in STM. For simplicity,

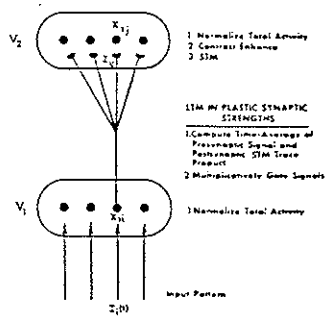


FIG. 15. The minimal feedforward geometry for retuning LTM in response to spatial pattern inputs. Feedback from STM in V_2 to LTM in V_1-to-V_2 pathways determines which LTM traces will sample normalized signals from V_1.

assume that it makes a choice. Then a population v_{2j} will become active in STM only if

$$S_j > \max\{\epsilon, S_k : k \neq j\} \tag{18}$$

where ϵ represents the quenching threshold. In particular, at most one population v_{2j} will be active in STM. Since the total activity of V_2 is normalized, the activity of the active v_{2j} will be set equal to 1.

How do these transformations of the input picture influence the LTM traces z_{ij}? We shall simplify (15) to make the main point. Replace (15) by the equation

$$\dot{z}_{ij} = (-z_{ij} + \theta_i)x_{2j} \tag{19}$$

where x_{2j} is the STM activity of v_{2j}. In other words, the picture influences z_{ij} only throgh its normalized activity, θ_i. Furthermore, *all* changes in z_{ij} come about when v_{2j} is active in STM, since otherwise $x_{2j} = 0$ and consequently $\dot{z}_{ij} = 0$.

Now consider what happens when a definite picture $\theta = (\theta_1, \theta_2, \ldots, \theta_n)$ is presented to V_1. Locate the population v_{2j} that gets the largest signal $S_j(t)$ at time $t = 0$. By (18), only v_{2j} will reverberate in STM. Consequently $x_{2j}(0) = 1$ and $x_{2k}(0) = 0$, $k \neq j$. By (19) all $\dot{z}_{ik}(0) = 0$, $k \neq j$, whereas

$$\dot{z}_{ij}(0) = -z_{ij}(0) + \theta_i \tag{20}$$

Using this information, it is easy to prove that, if the numbers $z_{ij}(0)$ are

sufficiently small, then the signal $S_j(t)$ grows through time, whereas all other signals $S_k(t)$, $k \neq j$, remain constant. In other words, pattern θ excites v_{2j} with ever-increasing vigor. In fact, learning maximizes S_j over all vectors $z^{(j)}(t) = (z_{1j}, z_{2j}, \ldots, z_{nj})$, whose length $|z^{(j)}| = (\sum_{k=1}^{n} z_{kj}^2)^{1/2}$ satisfies $|z^{(j)}| \leq |\theta|$. If we look at $\theta = (\theta_1, \theta_2, \ldots, \theta_n)$ and $z^{(j)} = (z_{1j}, z_{2j}, \ldots, z_{nj})$ as vectors in n-dimensional space, then it can be shown that presenting θ to V_1 makes $z^{(j)}(t)$ become parallel (or proportional) to θ as time goes on (Grossberg, 1976a,b). Because the choice of which vector $z^{(j)}$ will be trained in response to θ depends on *all* the LTM traces z_{ik}, as in (18), the statistical rules that generate the $z_{ik}(0)$ values before tuning begins will determine the pattern features that V_2 will try to classify. Grossberg (1976a) discusses this fact in relation to other data concerning the influence of positional gradients on the determination of interfield pathways.

How can we summarize this example in a way that will generalize to other cases? We can say that, at time $t = 0$, the pattern of signals $S_j(0)$ across V_2 might be almost uniformly distributed, as in Fig. 5b. The contrast-enhancement mechanism of V_2 converts these small differences in initial network activities into large differences in asymptotic network activities, as in Fig. 5c. If the contrast-enhanced activity gets sufficiently large to be sustained in STM, then it triggers the next development stage, such as slow changes in the LTM traces z_{ij}.

XI. Reaction-Diffusion Models in Development

Turing's (1952) classic paper on morphogenesis illustrates how a combination of chemical reactions and diffusion in a cellular system can generate and sustain spatially inhomogeneous activity patterns. Reaction-diffusion systems have become the subject of great interest because they are implicated in many experiments on developmental control processes. We shall describe a striking parallel between the dynamics of various such systems and of feedback networks undergoing mass action interactions. The reaction-diffusion models often omit important cellular properties, however, such as automatic boundedness, due to finitely many sites, and automatic gain control, due to multiplicative mass action interactions.

A class of reaction-diffusions of particular interest is found in Gierer and Meinhardt (1972) and Meinhardt and Gierer (1974). This model discusses both asymptotic steady states and traveling waves of chemical activity. A developmental example in which the former are relevant is the regeneration of *Hydra*'s heads.

XII. Regeneration of Hydra's Heads

Hydra is a small freshwater hydrozoan polyp that has a remarkable ability to regenerate amputated parts of its body. Figure 16 summarizes a series of experiments by Wilby and Webster (1970a,b) and Wolpert *et al.* (1971) that illustrate this ability. To describe these experiments, the *Hydra*'s body will be schematized as a head followed by four sections (H1234), as in Fig. 16a. If the head is cut off (Fig. 16b), another regenerates. The new head is designated by an asterisk (*) in Fig. 16b. If section 1 of a second *Hydra* is grafted onto a decapitated 1234, then one head grows (Fig. 16c). If sections 12 of a second *Hydra* are grafted onto a decapitated 1234, then two heads grow (Fig. 16d). In effect, mutual inhibition between the two 1 sections prevents growth of a second head in Fig. 16c. This inhibition cannot, however, traverse the space between the two 1 segments in Fig. 16d with sufficient strength to prevent two heads from growing. In Fig. 16e, and H12 section is grafted onto a 1234 section. No head grows at 1 in 1234. Somehow the H region in H12 can inhibit 1 in 1234 more vigorously then 1 could in 12 of Fig. 16d. Nonetheless, if H123 is grafted onto 1234, the inhibition from H is too weak to prevent head growth at 1 in 1234 (Fig. 16f). Transplantation of a

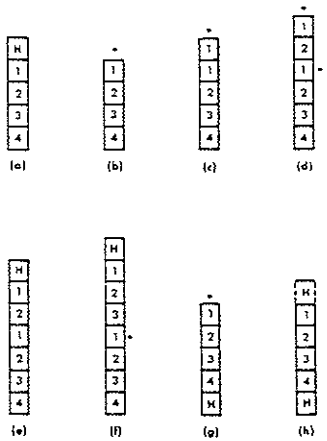


FIG. 16. Regeneration experiments on *Hydra*. The asterisk (*) describes where new heads will grow.

head from distal to proximal end (1234H) gives rise to a head at 1 (Fig. 16g), but this does not happen if a head is transplanted to the 4 end sufficiently before the original head is removed from the 1 end (Fig. 16h). Figures 16g and 16h suggest that inhibition can spread from the transplanted head to the 1 area to inhibit formation of a second head, much as the extra head in Fig. 16d can inhibit second head growth better than section 1 can in 12 of Fig. 16c. In other words, inhibition spreads over a wider region than excitation, and both excitation and inhibition can build up in prescribed regions through time.

To explain these phenomena, Gierer and Meinhardt use the computer to analyze a class of reaction-diffusion systems in which the concentrations of activators $x(w, t)$ and inhibitors $y(w, t)$ at various positions w control development through time t. One such system describes the time rates of change of x, namely $\partial x/\partial t$, and of y, namely $\partial y/\partial t$, by the equations

$$\frac{\partial x}{\partial t} = -Ax + B(w)y^{-1}f(x) + D_x \frac{\partial^2 x}{\partial w^2} + I(w) \tag{21}$$

and

$$\frac{\partial y}{\partial t} = -\hat{A}y + \hat{B}(w)g(x) + D_y \frac{\partial^2 y}{\partial w^2} \tag{22}$$

Gierer and Meinhardt are concerned with conditions under which a slight peak of initial activator concentration (for example, near section 1 in 1234) will lead to further increases of $x(w, t)$ at that position. They want these small initial concentration differences to yield large final concentration differences which are thereupon self-maintaining (Fig. 17). At a region of peak activator concentration, a new developmental stage is then triggered—for example, head formation in the *Hydra* at the 1 position in 1234.

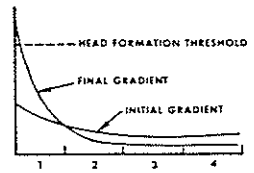


FIG. 17. Small initial spatial gradients are contrast-enhanced and thereafter sustained. Suprathreshold concentrations trigger the next developmental stage.

This proposal is strikingly similar to the network developmental mechanism described in Section X. There small differences in the pattern of geniculocortical signals are converted into large differences in excitatory cortical activity by the contrast-enhancement mechanism. These large differences are then maintained in STM, while the foci of maximal STM activity trigger the next developmental stage, which is characterized by slow changes in geniculocortical connection strengths.

A term-by-term comparison of (21) with (13) and of (22) with (14) sharpens this analogy: see Table II. Just as Gierer and Meinhardt make a distinction between the density $I(w)$ of morphogen sources and morphogen concentration [for example, $x(w, t)$], we distinguish between input intensity I_i and population activity [for example, $x_i(t)$]. They introduce activators $x(w, t)$ and inhibitors $y(w, t)$, whereas we need excitatory activities, or potentials, $x_i(t)$ and inhibitory activities $y_i(t)$. They call the mechanism whereby small differences become large differences "firing" of a gradient, whereas we call it contrast enhancement. The conditions under which firing and contrast enhancement occur are similar. For example, if their signal $f(x)$ in (21) is a power function such as $f(x) = x^r$, then firing occurs only if $r > 1$. In system (13), such a signal function induces contrast enhancement. In both kinds of system, exponential decay of concentrations or activities occur: for example, the terms $-Ax$ and $-\hat{A}y$ in (21) and (22). In both systems, activators or excitatory potentials can excite both themselves and inhibitors, whereas inhibitors can inhibit activators and possibly, but not necessarily, themselves. In both systems, mutual interactions between activators and inhibitors at different locations can occur. In (21) and (22) this occurs via the

TABLE II
COMPARISON OF REACTION-DIFFUSION AND SHUNTING NETWORK PROPERTIES

Reaction-diffusion	Shunting reverberating net
Activator $x(w, t)$	Excitatory activity $x_i(t)$
Inhibitor $y(w, t)$	Inhibitory activity $y_i(t)$
Morphogen source density	Inputs
Firing of morphogen gradient	Contrast enhancement
Maintenance of morphogen gradient	Power or sigmoid signal functions
Power or sigmoid signal functions	On-center off-surround interactions via electronic propagation of signals
Self-stabilizing distributions of morphogens if inhibitors equilibrate rapidly	Short-term memory pattern if inhibitors equilibrate rapidly
Periodic pulses if inhibitors equilibrate slowly	Periodic pulses if inhibitors equilibrate slowly
Regulation	Adaptation

diffusional terms $D_x(\partial^2 x/\partial w^2)$ and $D_y(\partial^2 y/\partial w^2)$, whereas in the networks it occurs via signals, which propagate either electronically or in nondecremental waves known as spikes. The diffusional coefficients are chosen to make inhibitor concentration spread within a wider area than activator concentration ($D_y > D_x > 0$). This constraint simulates an on-center off-surround field in the reaction-diffusion system, much as in Fig. 8 for the networks. The diffused inhibitor thereupon inhibits activators at other positions, much as the coefficients $\hat{D}_{ki}$ and E_{ki} in (13) and (14) spread inhibitory signals across an expanse of excitatory cells.

Not all choices of the numerical parameters A, $\hat{A}$, etc., in (21) and (22) yield a definite asymptotic pattern of activator concentrations as time goes on. If inhibitor concentration equilibrates rapidly, say because $\hat{A}$ is large, then an asymptotic pattern can exist. Analogously, an asymptotic pattern exists in the networks if inhibitory potentials equilibrate rapidly, thereby approximately lumping network dynamics. By contrast, if inhibitor equilibrates slowly, then traveling waves of activation can be obtained both in the reaction-diffusion scheme and in the networks. The analogy becomes even more suggestive when we realize that the partial *derivatives*, such as $D_x(\partial^2 x/\partial w^2)$, in the diffusional terms of a reaction-diffusion scheme become partial *differences*, such as

$$D_x[x(i + 1, t) - 2x(i, t) + x(i - 1, t)]$$

when the system describes interactions between contiguous cells v_i, and this difference scheme is formally a network.

XIII. The Analogy between Regulation and Adaptation

There also exists a striking analogy between the property of network adaptation and one of the most sought-after properties in reaction-diffusion systems. This is the property of *regulation*, which can be described in several ways. One way is to note that a living thing can increase its size dramatically while maintaining a remarkably constant form—for example, a growing leaf, or a teenager. A more abstract way is to ask: How can the same developmental pattern be created in two regions of different size, such as in Fig. 18. In Fig. 18 the problem for both of the cells v_1 and v_2 is to generate a red color because they both lie in the left third of their field: but how can the *cells* know this using only local information at their particular positions? This is the so-called French Flag problem of Wolpert (1969). How does each cell acquire *positional information* so that the same developmental pattern can be generated independent of the total field size? This is a *normalization*

FIG. 18. The French Flag Problem: How is the same pattern generated independent of size (Wolpert, 1969)?

property of the field that makes each cell aware of its *relative* position in the field. System (21)–(22) has approximately this property if $f(w)$ is a suitable sigmoid function of w, although the analysis of Gierer and Meinhardt does not disclose why this happens.

In shunting networks, *regulation* is replaced by *adaptation*, as illustrated by (4). Adaptation also describes a normalization property that makes the maximal total network response independent of the total number of cells of the input intensity. The adaptational mechanism preserves a record of the relative magnitude of activation at each position, just as positional information provides individual cells with indices of their relative position in a cellular array. Self-regulation in shunting networks is due to automatic gain control by intercellular signals. No such mechanism exists in (21)–(22).

Recall from Section II the observation that adaptation is needed to help a system process parallel data in the presence of noise. Developmental systems must also process such data. If the formal substrates of adaptation and regulation are to be identified, as is argued in Grossberg (1976a), then we can make a delightful statement. By solving the problems of preventing saturation and noise contamination. Nature also found a way to enable living things to grow without disturbing their form.

XIV. Blastula to Gastrula in the Sea Urchin

Gustafson and Wolpert (1967) summarize some beautiful experiments on the cellular forces that drive morphogenesis in the sea urchin embryo. Their main experimental tool was time-lapse cinematography. These experiments led them to posit the existence of a small number of mechanisms that operate at early developmental stages. They also summarize data concerning the development of other organisms in which similar mechanisms seem to exist.

We shall translate the posited mechanisms into formal terms. When

this is done, it becomes apparent that the morphogenetic mechanisms that have been suggested for sea urchin development are formally strikingly similar to the mechanisms in the model for cortical tuning in mammals. In particular, these data contain a manifest analog of the LTM traces z_{ij} in (15). This analogy suggests interesting new considerations concerning sea urchin development. For the sake of brevity, we merely summarize some highlights of one stage of development; once the formal analogies are noted, the rest of the data can be similarly analyzed.

Consider the dramatic step whereby the almost spherical blastula is deformed to produce a primitive gut in the gastrula. To discuss this transformation, Gustafson and Wolpert (1967) identify several mechanisms, which we summarize as follows: (1) cells can form pseudopods which are capable of stretching over considerable intraembryonic distances; (2) cells can be adhesive, or sticky, to other cells which contact them; (3) pseudopods can contract and thereby generate forces capable of causing cell motions. Gustafson and Wolpert describe how these properties can create a gastrula if they manifest themselves in the proper spatiotemporal pattern among blastula cells. Consider Fig. 19a. Let

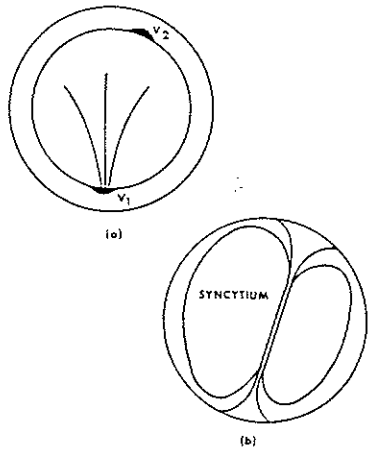


FIG. 19. (a) Pseudopods from v_1 search the cavity as cells v_2 become differentially adhesive. (b) A syncytium forms and exerts a force that draws v_1 and v_2 cells closer together.

certain mesenchymal cells v_1 start sending out pseudopods within a given time interval. These pseudopods randomly explore the blastula cavity. During an overlapping time interval, certain ectodermal cells v_2 become relatively adhesive. Consequently, with a high probability, the pseudopods from v_1 will adhere to the cells v_2. Gustafson and Wolpert argue, more generally, that pseudopods eventually aggregate in the direction of greatest adhesiveness. They show that individual pseudopods are continually making and breaking their cell contacts, but that there is a progressive fusion of the pseudopods into a syncytium. This syncytium statistically determines a strong connection between v_1 and v_2. Once this aggregate connection is established, pseudopodal contractions generate enough force to bring the cells v_1 and v_2 closer to each other and eventually form a primitive gut.

Consider the process whereby the syncytium is formed. For convenience, divide the blastula into two classes of cells, mesenchymal cells V_1 and ectodermal cells V_2, as in Fig. 19. Let V_1 contain n cells v_{1i}, $i = 1, 2, \ldots, n$, and V_2 contain m cells v_{2j}, $j = 1, 2, \ldots, m$. Suppose that v_{1i} generates $p_{1i}(t)$ pseudopods in the time interval $[t, t + \Delta t]$. Each pseudopod tends to randomly explore the blastula cavity from its vantage point. The number of pseudopods that contact v_{2j} is

$$S_{ij}(t) = p_{1i}(t)c_{ij}(t) \tag{23}$$

where $c_{ij}(t)$ is a slowly varying function that describes structural factors such as how close v_{1i} is to v_{2j}. Let $p_{2j}(t)$ be the adhesiveness of cell v_{2j} in the time interval $[t, t + \Delta t]$. If we suppose that pseudopodal connections from v_{1i} to v_{2j} are formed when a pseudopod from v_{1i} sticks to v_{2j}, then the rate of forming such pseudopods is proportional to $S_{ij}(t)p_{2j}(t)$. Choose the proportionality constant equal to one for simplicity. If we suppose that connections continue to form at this rate, then the total number $w_{ij}(t)$ of connections from v_{1i} to v_{2j} at time t satisfies

$$\dot{w}_{ij} = S_{ij}p_{2j} \tag{24}$$

since then

$$w_{ij}(t) = \int_0^t S_{ij}(v)p_{2j}(v)\,dv \tag{25}$$

if there are no connections at time $t = 0$. If, however, connections also break at a spontaneous rate, then (24) is replaced by

$$\dot{w}_{ij} = -b_{ij}w_{ij} + S_{ij}p_{2j} \tag{26}$$

where b_{ij} is the rate of breakage of pseudopods between v_{1i} and v_{2j}. Equation (26) for w_{ij} is formally identical to Eq. (15) for the LTM trace z_{ij}. In this analogy, the pseudopodal activity p_{1i} of mesenchymal cells is compared to the neural activity x_{1i} of geniculate cells. Just as x_{1i} generates neural signals at cortical cells, p_{1i} generates pseudopodal contacts at ectodermal cells. The pattern p_{2j} of adhesiveness across ectodermal cells is compared to the STM pattern x_{2j} across cortical cells. The "fast" variables p_{1i} and p_{2j} influence "slow" changes in intercellular connection strengths w_{ij}, just as the "fast" variables x_{1i} and x_{2j} influence "slow" changes in the intercellular connection strengths z_{ij}. In both cases, these slow changes embody an important step of the next developmental stage. How complete is this tantalizing analogy? Do these seemingly different systems obey the same formal laws? By comparing and contrasting them in some detail, interesting experimental and theoretical questions will emerge.

XV. Pseudopodal Signaling and Self-Corrective Feedback

Suppose for the sake of illustration that the two systems for formally identical. What does this imply? In the neural model, V_1 and V_2 are separately capable of normalizing, or adapting, their activities to produce positional information. They accomplish this by using signals within themselves that switch on and off cellular mechanisms by mass action laws. These signals are arranged in an on-center off-surround configuration to produce adaptation. Suppose that this also holds true in the sea urchin model. Then within the mesenchymal cells and the ectodermal cells there should exist excitatory and inhibitory signals in an on-center off-surround interaction pattern that switch on and off cellular mechanisms by mass action laws. Within the mesenchymal cells V_1, a large activation at v_{1i} induces intense pseudopodal activity; within the ectodermal cells V_2, a large activation at v_{2j} induces intense stickiness. Whether the signals are generated by intercellular diffusion or by another signal mechanism will generate the same qualitative results if the on-center off-surround pattern of Fig. 8 is accomplished, as was noted in Sections XII and XIII. So far, this analogy simply says that positional information selects for pseudopodal and adhesive cells. This much is easy to accept.

If the analogy is relentlessly pursued, however, then some unexpected issues emerge. For example, in the cortical model, the LTM traces z_{ij} help to determine the STM traces x_{2j} by multiplicatively gating the signal from v_{1i} to v_{2j}. The net signal from v_{1i} to v_{2j} is thus $S_{ij}z_{ij}$. At time $t = 0$, $z_{ij}(0)$ describes the a priori connection strength from v_{1i} to v_{2j} that is determined by the previous developmental stage. Arguing by analogy,

we ask whether pseudopodal contacts from v_{1i} to v_{2j} act as *signals* from v_{1i} to v_{2j} that are capable of tuning the distribution of pseudopodal contacts from V_1 to V_2? We now show that this would imply the existence of a robust self-correction procedure for growing a correctly positioned syncytium should a small genetic error occur.

To understand this mechanism, recall what happens in the geniculocortical tuning model. In that model, before tuning takes place, viable geniculocortical connections exist. A normalized pattern of activity across V_1 creates a total signal

$$S_j = \sum_{k=1}^{n} S_{kj} z_{kj}$$

at each cell v_{2j} in V_2. At time $t = 0$, the pattern of these signals across V_2 in response to a prescribed input pattern at V_1 can be almost uniform. Whenever this pattern is presented at V_1, the on-center off-surround interaction within V_2 contrast-enhances the signal pattern $\{S_1, S_2, \ldots, S_m\}$ until only certain larger signals generate sustained activity in STM. This transformation from signals at V_1 into STM at V_2 is a relatively fast process. Then the stored STM activities at V_2 cause slow changes in the LTM traces z_{ij} via (15). The slow changes in LTM traces can, in turn, shift the pattern of STM activity at V_2 until a dynamic equilibrium is established between STM at V_2 and the LTM traces from V_1 to V_2. In this way, a sustained bias in the patterns that perturb V_1 can induce a shift in the spatial distribution of STM and LTM at V_2. Such a feedback between STM and LTM is a robust and stable mechanism of developmental change.

What would the physical meaning of such a process be if it held during gastrula development? First it says that the pattern of pseudopodal contacts across V_2 can influence the relative adhesiveness of cells in V_2. In other words, pseudopodal contacts act as *signals*. The total signal

$$S_j = \sum_{k=1}^{n} S_{kj} w_{kj} \tag{27}$$

from V_1 to v_{2j} then determines how adhesive v_{2j} will start to become due to pseudopodal contacts. In other words, both z_{ij} and w_{ij} represent the *conditional expectation* that a signal will reach v_{2j} from v_{1i}. As in the cortical model, the different signals $S_j(0)$ can initially be very similar in size, because the various pseudopods are randomly exploring the blastula cavity. Suppose, however, that certain V_2 cells get an initially larger signal. Now comes a crucial point. *Why* do certain V_2 cells get an initially larger signal? They get this signal because of their position

relative to the pseudopod-generating cells in V_1. In particular, if a small genetic mistake in the position of pseudopod-generating cells in V_1 occurred, this mistake could be corrected by shifting the cells in V_2 that become sticky to the shifted pseudopods. Thus, if the signals from V_1 help to determine the pattern of adhesiveness across V_2, then the system enjoys a self-correction property.

To pursue our analogy, suppose that certain cells in V_2 initially get larger signals $S_j(0)$ from V_1. Then contrast enhancement and STM of the signal pattern will occur in V_2. That is, ectodermal cells, which receive the most pseudopods, will start inhibiting the stickiness of V_2 cells which receive fewer pseudopods via their on-center off-surround interactions. Consequently, more pseudopods will stick to these preferred cells, thereby making them even more sticky via lateral inhibition. In other words, the randomly searching pseudopods eventually form a syncytium from V_1 to V_2. This process eventually equilibrates because the total amount of adhesiveness across V_2 is normalized by its on-center off-surround interactions. In summary, the analogy with cortical tuning suggests a self-correction procedure for synchronizing mesenchymal and ectodermal interactions, and a mechanism of syncytium production due to positive feedback between pseudopodal aggregation and the adhesiveness of ectodermal cells, supplemented by adaptation, contrast enhancement, and STM of the pattern of adhesiveness.

The above mechanism can also be approached in another way. Suppose that the properties of adhesiveness and mutual cell contact are parallel properties in ectodermal cells: that is, those ectodermal cells that contact their neighbors most vigorously are also the cells that are most adhesive. Gustafson and Wolpert (1967, pp. 471–472) note compatible data: "The adhesiveness . . . is thus non-specific in the sense that it serves in the mutual contact of ectodermal cells as well as the contact between mesenchymal pseudopods and the ectoderm." The idea that pseudopods act as signals then reduces to the following statement. Stickiness can increase cell contact, *and conversely*. If more pseudopods contact particular ectodermal cells, then, other things being equal, these cells will become more sticky.

The above mechanism will operate whether or not there is an independent mechanism within V_2 for generating an adhesiveness gradient across ectodermal cells. If there is such an initial gradient across V_2, but the relative positions of pseudopodal and adhesive cells are incorrect because of some uncontrolled factor, then the tuning mechanism will tend to correct it. This is also what happens in the cortical tuning model. There, viable geniculocortical connections exist before tuning of these connections takes place. These connections are presum-

ably due to the action of positional gradients in both the geniculate and the cortex at a previous developmental stage: cf. the development of the retinotectal map in *Xenopus* (Hunt and Jacobson, 1972, 1973a,b). Nonetheless, a sustained bias of activity in the model's geniculate can induce a shift in the pattern of geniculocortical connections. If there is no mechanism within V_2 for independently generating an adhesiveness gradient—which seems unlikely, given the vital importance of gastrulation—then pseudopodal signaling will tend to produce one even if there are slight structural asymmetries that favor contacts at certain ectodermal cells above others.

XVI. Some Experimental Tests

How can these suggestions be tested? Of course, certain experimental manipulations can renormalize the field interactions, but with these problems under control, the following experiments are suggested, among others. First, if certain cells in V_2 become relatively very sticky just before pseudopods are generated, then pseudopodal signaling is at best a second-order correction procedure for a primary template of adhesiveness. Note, however, that all cells in V_2 can initially be almost equally sticky without contradicting the idea that pseudopods are primary inducers of the V_2 gradient. If pseudopods are prevented from reaching V_2, will the pattern of adhesiveness across V_2 be different than it would be if pseudopods can reach V_2; for example, does the pattern remain relatively uniform in the former case? If not, then again pseudopodal signals are not primary agents for regulating V_2's adhesiveness. If pseudopod cells are shifted relative to ectodermal cells, does the pattern of ectodermal stickiness also shift?

The above comparison of cortical tuning with gastrula formation focuses on a general property that is worthy of consideration independent of any particular model—namely, the property of self-correction, whereby a shift in pseudopodal cells can induce a corresponding shift in sticky cells. Such a property can formally be achieved without invoking pseudopodal signaling. For example, suppose that mesenchymal and ectodermal cells form part of a single normalized field of on-center off-surround signals. Let the properties of pseudopod production and adhesiveness be complementary properties in this field (dipole field!). Suppose that there is a spatial gradient of field activity from mesenchymal cells to ectodermal cells. Then a peak of pseudopodal activity at certain mesenchymal cells will correspond to a peak of adhesiveness at antipodal ectodermal cells. Consequently, a shift in the most active

mesenchymal cells will induce a shift in the locus of maximally adhesive ectodermal cells. This mechanism of self-correction operates via signals passed between contiguous cells, rather than by pseudopodal signals.

Were both mechanisms of self-correction available, there would exist a primary gradient for syncytium production stabilized by a secondary feedback mechanism.

XVII. Production versus Directed Growth

The comparison between cortical and gastrula models also suggests a formal isomorphism between certain processes of chemical production and growth. It is known that viable geniculocortical connections exist before tuning takes place (Hubel and Wiesel, 1970). In the cortical model, therefore, changes in z_{ij} need not reflect new growth or decay of neural connections, but rather merely a change in transmitter production rates and/or in the sensitivity of postsynaptic membrane sites. Of course, if more transmitter is produced, then a secondary growth of synaptic knobs can occur. In the blastula model, w_{ij} unambiguously describes the growth of connections. The fact that z_{ij} and w_{ij} obey identical laws shows that production, sensitization, and growth can all be stages of certain developmental or learning processes without changing the formal laws of these processes from stage to stage. This observation also widens the applicability of known theorems about LTM. The general theorems on LTM in networks with arbitrary anatomies and data preprocessing (Grossberg, 1971, 1972a, 1974) can be interpreted as theorems concerning the growth patterns that are generated by prescribed patterns of inducer activity. Previously, the generality of these theorems was interpreted to mean that this LTM mechanism, if invented at a prescribed time during evolution, could be utilized to learn arbitrarily complex patterns in any later evolutionary specialization.

XVIII. Biochemical Memory and the Folds of Rhodnius

Locke (1959, 1960, 1967) described a segmental gradient in the insect *Rhodnius*. This gradient controls the polarity of epidermal cells and carries positional information. Polarity is expressed by the orientation of folds in the adult epicuticle, which are parallel to the contours of the gradient. Lawrence *et al.* (1972) investigated various reaction-diffusion models to test whether one of them could match the cuticle pattern in adult insects after rotation of square pieces of epidermis in insect larvae.

They ruled out a model in which the gradient depends only on the activities of a line of source cells at one end of the segment and a line of sink cells at the other end. Such a model is compatible with Wolpert's (1969) heuristic analysis of positional information. A model in which each cell is a homeostatic unit in the gradient was found to fit the data well. In such a model, each cell participates in establishing and maintaining the gradient, as in a reverberating network. They also found that a given cell attempts to maintain its original or "set" concentration when it is transplanted to a new position in the segment. In other words, each cell "remembers" its positional information. Once moved to a new position, the cell's activity is gradually influenced by contiguous activities to produce an intermediate activity level.

The possible existence of intracellular memory of position suggests the consideration of intracellular networks of reactions. Such a memory holds in the dipole field of Fig. 11. As Eq. (12) shows, an individual dipole can store its normalized activity in STM for indefinitely long times. When such a dipole is embedded into a field of intercellular interactions, as in Fig. 11, its STM activity level will be influenced by the levels of its neighbors, as is suggested by the data of Lawrence *et al.* (1972).

The dipole model, as in Eqs. (10) and (11), requires that each cell contain two processes that excite themselves and inhibit each other, perhaps through a chain of intermediate reactions. When the inhibition reacts quickly, asymptotic steady states exist. When inhibition reacts slowly, periodic oscillations can occur, as in Section IX. These intracellular processes thereupon generate intercellular signals that influence each other in an on-center off-surround interaction pattern. A likely candidate for mutually antagonistic intracellular processes that set metabolic rates are intracellular reactions that involve cAMP and cGMP. For example, in slices of bovine sympathetic ganglion, dopamine increases cAMP but not cGMP production, whereas acetylcholine increases cGMP but not cAMP production (Kebabian *et al.*, 1975). More generally, there is a growing body of biochemical data which suggests that cAMP and cGMP are antagonists in the regulation of intracellular biochemical reactions (Libassi, 1974). In more formal terms, antagonistic pairs of recurrent processes are needed to maintain an intracellular memory of biochemical production rates, just as antagonistic pairs of recurrently interacting motor control cells can maintain a fixed limb posture. See Section XXI.

Lawrence *et al.* (1972, pp. 826–827) also find that "the homeostatic level of each cell is reset at some stage in the cell cycle to the ambient concentration at that time." In Eqs. (10) and (11) this amounts to turning on the input I during a critical time interval in the cell cycle. Such a "nonspecific shunt" or "tuning" of cell excitability also often occurs in neural networks. For example, turning on a nonspecific shunt can operate a Now Print mechanism that allows learning to occur during a critical interval (Grossberg, 1974), or trigger STM storage of a sensory pattern by overcoming the quenching threshold (Grossberg, 1973), or activate the map that causes a limb to move toward a prescribed terminal position (Grossberg, 1973, 1975, 1978b).

If a dipole does, indeed, control the "memory" that Lawrence *et al.* (1972) observe, then this memory should properly be thought of as STM, notwithstanding its longevity. STM is distinguished from LTM not by its possible duration, but by how rapidly it can be erased by competing events.

XIX. Slime Mold Aggregation and Slug Motion

Here it will be suggested that various stages of slime mold development can be discussed in terms of such interactive network properties as nonspecific shunts, contrast enhancement, and hysteresis. We shall also make some suggestions about underlying slime mold dynamics. Bonner (1967, 1969, 1974) elegantly summarizes the various stages whereby hungry amoebas aggregate into a slug, which thereupon is capable of organized movement until it forms a fruiting body. Bonner (1969) describes the experiments that implicate cAMP as the intercellularly diffusing agent that causes amoebas to stream together into a slug. In the spirit of Turing's (1952) paper, both Keller and Segel (1970) and Gierer and Meinhardt (1972) view aggregation as an instability in a reaction-diffusion system. This instability becomes the property of contrast enhancement in the analogous network. In other words, after food is removed, each amoeba begins to produce more cAMP. By chance, certain amoebas produce more cAMP initially. Contrast enhancement amplifies these small differences in initial cAMP production until they eventually become large differences. To achieve contrast enhancement, the cAMP signals $f(w)$ from individual amoebas will be assumed to be sigmoid functions of an intracellular activity level, w. For example, as an amoeba's activity level increases, more cAMP production sites can be recruited because their production thresholds are (say) Gaussianly distributed around some mean value M, as in Fig. 6. Turning on these signals at a prescribed time can be achieved by a nonspecific shunt that either amplifies activity or lowers the mean production threshold. Were such a shunt to operate in the present case, it would presumably be triggered nonspecifically across the field of cells by the absence of food.

To prevent saturation of the pattern of these intercellular signals, an antagonistic chemical signal is also needed. This chemical is called acrasinase in the classical literature (Bonner, 1969) and is known to be a phosphodiesterase (Chang, 1968). How does acrasinase work? The model suggests that, to avoid saturation, a recurrent off-surround exists. In other words, intracellular acrasinase production should be coupled to intracellular cAMP production such that cells producing the most cAMP also produce the most acrasinase, as in Fig. 8. Acrasinase would diffuse more broadly than cAMP to simulate a recurrent on-center off-surround interaction. If (say) the gain of the acrasinase production process is sufficiently small [cf. $\tilde{A}$ in Eq. (14)], then periodic pulses of cAMP can be produced, even if no amoeba acts as a pacemaker.

How does cAMP influence amoeboid motion? As in the case of sea urchin pseudopod motion, it is plausible to assume that amoebas are attracted toward the directions in which cAMP concentration is greatest (Keller and Segal, 1970). The simplest way to approximate this motion in a network is as follows. Let n amoebas $v_1, v_2, \ldots, v_n$ be distributed in a planar region at position $u_1(t), u_2(t), \ldots, u_n(t)$ at time t, where $u_i(t) = (u_{i1}(t), u_{i2}(t))$ with $u_{i1}(t)$ the abscissa and $u_{i2}(t)$ the ordinate of v_i's planar position. At time t, the total signal from all cells to v_i is

$$\sum_{k=1}^{n} f(x_k(t))C(u_k(t), u_i(t)) \tag{28}$$

where $f(w)$ is the sigmoid signal in response to activity level w, $x_k(t)$ is the activity at time t of v_k, and $C(P, Q)$ is the connection strength from position P to position Q; for example,

$$C(P, Q) = \alpha \exp[-\beta|P - Q|^2] \tag{29}$$

as in Fig. 8, with $|P - Q|$ the distance between P and Q. Expression (28) can be written more tersely as

$$\sum_{k} f_k(t)C_{ki}(t) \tag{30}$$

with $f_k(t) = f(x_k(t))$ and $C_{ki}(t) = C(u_k(t), u_i(t))$. Suppose for simplicity that the amoebas are initially close enough to each other to be able to sense each other's cAMP gradient. Then the motion of cells can be approximated by

$$\frac{\partial u_i}{\partial t} = \sum_{j=1}^{n} \sum_{k=1}^{n} f_k(C_{kj} - C_{ki}) \frac{u_j - u_i}{|u_j - u_i|} \tag{31}$$

or in component form

$$\frac{\partial u_{i1}}{\partial t} = \sum_{j=1}^{n} \sum_{k=1}^{n} f_k(C_{kj} - C_{ki}) \cos \theta_{ij}$$

and

$$\frac{\partial u_{i2}}{\partial t} = \sum_{j=1}^{n} \sum_{k=1}^{n} f_k(C_{kj} - C_{ki}) \sin \theta_{ij}$$

where θ_{ij} is as defined in Fig. 20. In other words, cell motion is induced by differences in the total signals experienced at each cell.

Several qualitative conclusions about aggregation can be read off from Eq. (31). First, suppose that a given cell v_i is much more active than other cells that surround v_i at random positions. The v_i will tend to experience an almost radially symmetric gradient, because when its large, radially symmetric signals are added to the relatively small signals produced by the other cells, the total gradient remains approximately radially symmetric. Thus, relatively active cells tend to move less rapidly than other cells, other things being equal: they act as attractors for relatively inactive cells, which tend to move most rapidly toward the highest density of signals. This much is obvious from the assumption that attractive gradients exist.

Second, an important equivalence can be observed between relative motion of two cells and the dilation or contraction of their on-center and off-surround fields (Fig. 21). Figure 21 shows two cells far apart and then closer together. The passage from Fig. 21a to Fig. 21b can be achieved either by moving the cells closer together or, equivalently, by broadening the spatial extent of their on-center and off-surround interactions. Given that Fig. 21b holds, we can now ask what happens whenever two cells get close enough together to be deeply immersed in their on-centers? The answer is: hysteresis. That is, a high central peak

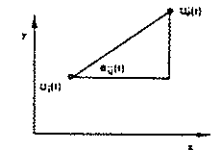


FIG. 20. Cell motions induced by spatial gradients established by the total signals at each point.

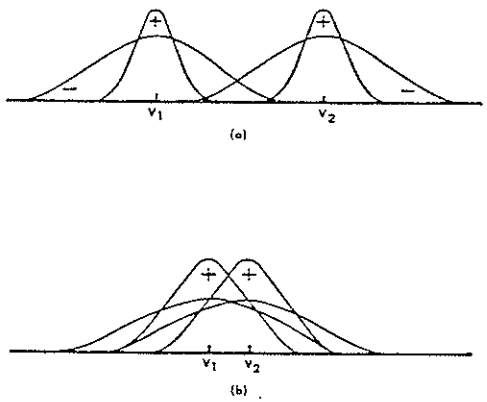


FIG. 21. Cells outside each other's on-center (a) and then inside (b), being due either to relative motion or to relative expansion of excitatory fields.

of excitation is produced by the recurrent excitation. This central peak tends to keep the cells close together. In other words, cells tend to form clumps that stick together, if only because of hysteresis. By the first remark, these clumps will tend to stream toward the locus of maximal excitation, ultimately forming a slug. This observation does not imply that mechanisms other than hysteresis, such as adhesiveness, are inoperative. Rather it suggests that field effects, such as hysteresis, tend to form a dynamic equilibrium with cellular mechanisms, such as adhesiveness.

Given that a slug is eventually formed, how does it move? To start off, suppose that the slug has a rectangular shape, in which each cell is equally active. Figure 14a shows that contrast enhancement can then create relatively large excitation at both ends of the slug. Were this the case, the amoebas would tend to move toward the two ends of the slug, thereby tending to split it into two parts. However, the amoebas are *not* equally active. The mechanism of contrast enhancement will have generated a maximum focus of activity in a small number of contiguous amoebas before the slug is totally formed. When the slug is formed, contrast enhancement will further strengthen the relative activity of these cells (Fig. 22). Consequently the amoebas will tend to move in a preferred direction, held together by hysteresis, among other factors. Indeed, sometimes the front end of the slug is raised while it moves.

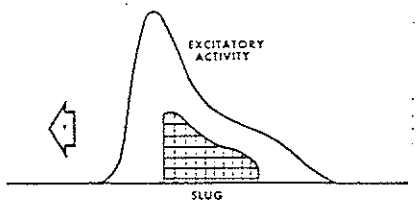


FIG. 22. Colony motion due to asymmetric excitatory gradient and hysteresis.

When this happens the slug moves faster (Bonner, 1967). Such a configuration can only increase the amount of contrast enhancement at the front end of the slug by increasing its surface area. The gradient is thereby steepened and faster movement is anticipated.

XX. Adhesiveness, Growth, Cell Streaming, and Division

This discussion of aggregation and slug motion in terms of nonspecific shunting, contrast enhancement, and hysteresis does not imply that other factors are not at work—for example, adhesiveness between amoebas. It does suggest, however, that interactive properties set the stage for, and stabilize, other mechanisms that are acting in parallel with them. Indeed, if the adhesive bonds between cells are continually being made and broken, as indicated in Section XIV, with a spatial distribution that depends on the gradient experienced by each cell, then colonies of a given cell will behave much like a liquid, as Steinberg and his colleagues have elegantly demonstrated (Johnson, 1974; Steinberg, 1970). For example, a cell near the center of a colony will tend to remain spherical because it experiences an almost spherical gradient from the cells that surround it, and will tend to flow and make intercellular contacts that match this gradient.

In a similar fashion, as a colony of cells grows, the pattern of activity across these cells due to their collective interactions can also change. This is due to a change in the spatial scale of the colony relative to the scale of its on-center off-surround interactions, much as in Fig. 14. Suppose, for example, that all cells in the colony are equally active, and that the entire colony initially fits within the breadth of the on-center of each cell. Then inhibitory interactions are relatively weak, and a unimodal net activity across populations can occur. This spatial gradient will tend to keep the cells in the colony clustered together. If, however,

the colony grows so much that it exceeds the on-center scale, then inhibitory interactions tend to contrast-enhance the activity at its boundaries, as in Fig. 14a. This will induce cell streaming toward the boundaries and a tendency to split the colony into two parts. If a similar mechanism operates in individual cells, then cell growth can help to induce mitosis by simultaneously changing the spatial gradients of membrane adhesiveness and protoplasmic streaming across the cell.

The above remarks suggest that similar formal rules hold in many developmental and adult learning situations. These common rules are generated by the common problems of systems that process patterned information which is then used to trigger a new stage of system organization. In effect, similar statistical and geometrical constraints are needed to guarantee successful processing. A casual similarity was already noticed by Gierer and Meinhardt (1972) between their reaction-diffusions and the neural model of Hartline and Ratliff (Ratliff, 1965) for lateral inhibition in the *Limulus* retina. This relationship was not pursued if only because the Hartline-Ratliff model uses additive interactions, is purely inhibitory, and is incapable of STM. In other words, both the statistics and geometry of the Hartline-Ratliff model are inappropriate for noticing the deep connections between nonlinear reaction-diffusions and nonlinear networks.

XXI. Chemical Substrates of STM and LTM: Cyclic Nucleotides, Ions, Transmitters, and Protein Synthesis

The above examples share common formal properties. Are these common formal properties realized by common cellular mechanisms—in particular, by common biochemical reactions? This section briefly suggests an affirmative answer, which will be developed more completely elsewhere. In all the examples, recurrent dipoles regulate intracellular STM. This means that pairs of antagonistic chemicals can be used to set and maintain biochemical production rates at prescribed levels until an extracellular signal quickly resets these levels. Do such ubiquitous chemical dipoles exist? The cyclic nucleotides cAMP and cGMP are implicated as such a pair by numerous experiments that depict their role as antagonistic "second messengers" capable of triggering intercellular reactions (Robison *et al.*, 1971; Schultz *et al.*, 1973a,b; Ueda *et al.*, 1973; Casnellie and Greengard, 1974; Libassi, 1974; Clement-Cormier *et al.*, 1974; Nathanson and Greengard, 1974; Garbers *et al.*, 1975a,b; Greengard, 1975; Sloboda *et al.*, 1975; Willingham and Pastan, 1975). Intercellular STM dipoles involving cAMP and cGMP can

also be inferred to exist. For example, Wise *et al.* (1973) show that reward centers employ norepinephrine and punishment centers employ serotonin as their respective transmitter substances. These authors suggest that the relative activities of these centers determine an animal's net emotional affect of any time. This data can be combined with data showing that norepinephrine (or its precursors) stimulates cAMP production, whereas serotonin stimulates cGMP production (Libassi, 1974; Kebabian *et al.*, 1975) to suggest that the relative imbalances in norepinephrine and serotonin are paralleled by relative imbalances in cAMP and cGMP in the two centers. These experiments are compatible with a neural theory of reinforcement that is derived from simple behavioral postulates (Grossberg, 1972b,c, 1975). Herein, transmitters in parallel pathways form a recurrent dipole that regulates net incentive motivation through time. The relative concentrations of the transmitters in the parallel pathways control the net emotional state any any time, as in the Wise *et al.* (1973) data. The theory derives equations for the dipole which predict how reinforcing different events will be. Given the above remarks, the equations also predict relative amounts of cAMP and cGMP in the parallel pathways.

What about LTM? Is there a common chemical mechanism for LTM in many biological systems? Is LTM also stabilized by a recurrent dipole? Recall that LTM, by contrast with STM, cannot be switched from one level to another as soon as an extracellular signal is imposed.

We suggest that the answers to all these questions is "yes." In nerve cells, we suggest that the two structural ends of the dipole are near the cell body and the synaptic knobs, and that the LTM transmitter can be acetylcholine. Such a chemical dipole was derived in Grossberg (1969) from a neural model of classical conditioning, which is reviewed in Grossberg (1974). This work analyzes the minimal intracellular mechanisms that are compatible with the LTM equations of the model. Various experimental data implicate the cholinergic synapse as the site of LTM (Deutsch, 1972), and such data are compatible with and reviewed in the reinforcement theory in Grossberg (1972c).

The minimal LTM mechanism in Grossberg (1969) was also generalized to describe a way to regulate the LTM of intracellular production rates in other situations. This minimal scheme is relevant to recent data on cyclic nucleotides, and will therefore be briefly sketched. It describes cellular sites at which hierarchies of binding strengths exist for the ions Na^+, K^+, Ca^{2+}, and Mg^{2+} acting in parallel with other (unnamed) chemicals. In this scheme, Na^+ is antagonistic to K^+, Ca^{2+} is antagonistic to Mg^{2+}, Na^+ and Ca^{2+} are synergistic cofactors, Ca^{2+} and Mg^{2+} bind much more strongly to intracellular sites than do Na^+ and K^+, and

acetylcholine causes release of unbound Ca^{2+}, which thereupon triggers events leading to LTM. Differences in the relative amounts and binding strengths of these ions at particular sites are used to explain a variety of facts pertaining to presynaptic acetylcholine synthesis, postsynaptic protein synthesis, intracellular transport, and extracellular transmitter release. The two ends of the LTM dipole coordinate the availability and spatial shifts of the ions (and parallel processes). The bound ions set reaction rates leading to LTM, and thereby constitute a type of "intermediate memory" between STM potentials and signals, and possible LTM structural changes.

Two sets of recent experiments tend to support and extend this theoretical picture. Schultz *et al.* (1973a) showed that Ca^{2+} mediates the production of cGMP in response to acetylcholine in the ductus deferens of the rat. Moreover, in mammalian organs rich in smooth muscle, cGMP initiates specific cyclic nucleotide-dependent phosphorylation of synaptic membrane proteins, which possibly alter the selective permeability properties of the membrane (Ueda *et al.*, 1973; Casnellie and Greengard, 1974; Greengard, 1975; Kebabian *et al.*, 1975). In other words, the experiments suggest how acetylcholine can initiate LTM changes, such as sensitization of the postsynaptic membrane, by causing release of unbound Ca^{2+}. Such biochemical results have recently encouraged the view that subtle shifts in the concentrations of ions can alter the rates of sustained biochemical production (Libassi, 1974). From the vantage point of the LTM theory in Grossberg (1969, 1974), the pattern of binding among the ions Na^+, K^+, Ca^{2+}, and Mg^{2+} at different cellular sites sets the rates of reaction whose end products embody LTM changes. Coordinating the availability of these ions at a variety of cellular sites requires the existence of a chemical dipole in each cell.

In summary, results on the biochemical regulation of STM and LTM in single cells tend to parallel the results on STM and LTM in networks of cells. We take these striking parallels as evidence that certain universal developmental control mechanisms are coming into view.

REFERENCES

Blakemore, C., and Cooper, G. F. (1970). *Nature (London)* 228, 477.

Bonner, J. T. (1967). "The Cellular Slime Molds," 2nd ed. Princeton Univ. Press, Princeton, New Jersey.

Bonner, J. T. (1969). *Sci. Am.* 220, 78.

Bonner, J. T. (1974). "On Development." Harvard Univ. Press, Cambridge, Massachusetts.

Casnellie, J. E., and Greengard, P. (1974). *Proc. Natl. Acad. Sci. U.S.A.* 71, 1891.

Chang, Y. Y. (1968). *Science* 160, 57.

Clement-Cormier, Y. C., Kababian, J. W., Petzold, G. L., and Greengard, P. (1974). *Proc. Natl. Acad. Sci. U.S.A.* 71, 1113.

Cornsweet, T. N. (1970). "Visual Perception." Academic Press, New York.

Deutsch, J. A. (1972). *In* "The Chemistry of Mood, Motivation, and Memory" (J. L. McGauch, ed.). Plenum, New York.

Elias, S. A., and Grossberg, S. (1975). *Biol. Cybernet.* 20, 69.

Fender, D., and Julesz, B. (1967). *J. Opt. Soc. Am.* 57, 819.

Garbers, D. L., Dyer, E. L., and Hardman, J. G. (1975a). *J. Biol. Chem.* 250, 382.

Garbers, D. L., Chrisman, T. D., Suddath, J. L., and Hardman, J. G. (1975b). *Arch. Biochem. Biophys.* 166, 135.

Gierer, A., and Meinhardt, H. (1972). *Kybernetik* 12, 30.

Greengard, P. (1975). *Adv. Cyclic Nucleotide Res.* 5, 585.

Grossberg, S. (1969). *J. Theor. Biol.* 22, 325.

Grossberg, S. (1971). *Proc. Natl. Acad. Sci. U.S.A.* 68, 828.

Grossberg, S. (1972a). *In* "Delay and Functional Differential Equations and Their Applications" (K. Schmitt, ed.), p. 121. Academic Press, New York.

Grossberg, S. (1972b). *Math. Biosci.* 15, 39.

Grossberg, S. (1972c). *Math. Biosci.* 15, 253.

Grossberg, S. (1972d). *Kybernetik* 10, 49.

Grossberg, S. (1973). *Stud. Appl. Math.* 52, 213.

Grossberg, S. (1974). *In* "Progress in Theoretical Biology" (R. Rosen and F. Snell, eds.), Vol. 3, p. 51. Academic Press, New York.

Grossberg, S. (1975). *Int. Rev. Neurobiol.* 18, 263.

Grossberg, S. (1976a). *Biol. Cybernet.* 21, 145.

Grossberg, S. (1976b). *Biol. Cybernet.* 23, 121.

Grossberg, S. (1976c). *Biol. Cybernet.* 23, 187.

Grossberg, S. (1977a). *J. Math. Biol.*, 4, 237.

Grossberg, S. (1978a). *J. Theoret. Biol.* (in press).

Grossberg, S. (1978b). This volume.

Grossberg, S., and Levine, D. S. (1975). *J. Theor. Biol.* 53, 341.

Gustafson, T., and Wolpert, L. (1967). *Biol. Rev. Cambridge Philos. Soc.* 42, 442.

Hebb, D. O. (1949). "The Organization of Behavior." Wiley, New York.

Hirsch, H. V. B., and Spinelli, D. N. (1971). *Exp. Brain Res.* 12, 509.

Hodgkin, A. L. (1964). "The Conduction of the Nervous Impulse." Thomas, Springfield, Illinois.

Hubel, D. H., and Wiesel, T. N. (1970). *J. Physiol. (London)* 206, 419.

Hunt, R. K., and Jacobson, M. (1972). *Proc. Natl. Acad. Sci. U.S.A.* 69, 780.

Hunt, R. K., and Jacobson, M. (1973a). *Proc. Natl. Acad. Sci. U.S.A.* 70, 507.

Hunt, R. K., and Jacobson, M. (1973b). *Science* 180, 509.

Johnson, K. E. (1974). *In* "Concepts of Development" (J. Lash and J. R. Whittaker, eds.). Sinauer Assoc., Stamford, Connecticut.

Katz, B. (1966). "Nerve, Muscle, and Synapse." McGraw-Hill, New York.

Kauffman, S. A. (1971a). *Curr. Top. Dev. Biol.* 6, 000.

Kauffman, S. A. (1971b). *Lect. Math. Life Sci.* 3, 000.

Kebabian, J. W., Bloom, F. E., Steiner, A. L., and Greengard, P. (1975). *Science* 190, 157.

Keller, E. F., and Segal, L. A. (1970). *J. Theor. Biol.* 26, 399.

Koch, S. (1954). *In* "Modern Learning Theory" (A. T. Poffenberger, ed.), p. 1. Appleton, New York.

Lawrence, P. A. (1970). *Adv. Insect Physiol.* 7, 197.

Lawrence, P. A. (1971). *Symp. Soc. Exp. Biol.* 25, 379.
Lawrence, P. A. (1972). *In* "Development Systems: Insects" (S. J. Counce and C. H. Waddington, eds.), Vol. 2. Academic Press, New York.
Lawrence, P. A., Crick, F. H. C., and Munro, M. (1972). *J. Cell Sci.* 11, 815.
Levine, D. S. (1974). Ph. D. thesis, Mathematics Dept., M.I.T., Cambridge, Mass.
Levine, D. S., and Grossberg, S. (1976). *J. Theor. Biol.* (in press).
Libassi, P. T. (1974). *Sciences (N.Y.)* 14, 15.
Locke, M. (1959). *J. Exp. Biol.* 36, 459.
Locke, M. (1960). *J. Exp. Biol.* 37, 398.
Locke, M. (1967). *Adv. Morphog.* 6, 33.
Meinhardt, H., and Gierer, A. (1974). *J. Cell Sci.* 15, 321.
Mueller, C. G., Jr., and Schoenfeld, W. N. (1954). *In* "Modern Learning Theory" (A. T. Poffenberger, ed.), p. 345. Appleton, New York.
Nathanson, J. A., and Greengard, P. (1974). *Proc. Natl. Acad. Sci. U.S.A.* 71, 797.
Pavlov, I. P. (1927). "Conditioned Reflexes." Oxford Univ. Press, London and New York (Reprint, Dover, New York, 1960).
Ramón y Cajal, S. (1909–1911). "Histologie du système nerveux de l'homme et des vertébrés. (transl. by L. Azoulay), 2 vols. Paris.
Ratliff, F. (1965). "Mach Bands. Quantitative Studies of Neural Networks in the Retina." Holden-Day, San Francisco, California.
Robison, G. A., Butcher, R. W., and Sutherland, E. W. (1971). "Cyclic AMP." Academic Press, New York.
Schultz, G., Hardman, J. G., Schultz, K., Davis, J. W., and Sutherland, E. W. (1973a). *Proc. Natl. Acad. Sci. U.S.A.* 70, 1721.
Schultz, G., Hardman, J. G., Schultz, K., Baird, C. E., and Sutherland, E. W. (1973b). *Proc. Natl. Acad. Sci. U.S.A.* 70, 3889.
Sloboda, R. D., Rudolph, S. A., Rosenbaum, J. L., and Greengard, P. (1975). *Proc. Natl. Acad. Sci. U.S.A.* 72, 177.
Steinberg, M. S. (1970). *J. Exp. Zool. U.S.A.* 173, 395.
Turing, A. M. (1952). *Philos. Trans. R. Soc. London, Ser. B* 237, 37.
Ueda, T., Maeno, H., and Greengard, P. (1973). *J. Biol. Chem.* 248, 8255.
Werblin, F. S. (1971). *J. Neurophysiol.* 34, 228.
Wilby, O. K., and Webster, G. (1970a). *J. Embryol. Exp. Morphol.* 24, 583.
Wilby, O. K., and Webster, G. (1970b). *J. Embryol. Exp. Morphol.* 24, 595.
Willingham, M. C., and Pastan, I. (1975). *Proc. Natl. Acad. Sci. U.S.A.* 72, 1263.
Wilson, H. R., and Cowan, J. D. (1973). *Kybernetik* 13, 55.
Wise, D. C., Berger, B. D., and Stein, L. (1973). *Biol. Psychiatry* 6, 1.
Wolpert, L. (1969). *J. Theor. Biol.* 25, 1.
Wolpert, L., Hicklin, J., and Hornbruck, A. (1971). *Symp. Soc. Exp. Biol.* 25, 391.