

A Theory of Human Memory: Self-Organization

and Performance of Sensory-Motor Codes,

Maps, and Plans*

1. Introduction

A psychophysiological theory of the self-organization and performance of sensory-motor codes, maps, and plans is derived herein. This general topic includes a variety of phenomena in many species, ranging from the imprinting of sensory-motor coordinates in an infant to complex goal-oriented serial behavior in an adult. The theory approaches the problem of biological diversity by seeking organizational principles that have evolved in response to environmental pressures to which a surviving species must adapt. These organizational principles are translated into explicit neural networks that realize the principles in a minimal way. Once the properties of the minimal mechanisms are understood, variations of them can be more readily recognized and analyzed in species-specific examples.

The result is a small number of robust principles and mechanisms that form a common substrate for coping with many tasks. These mechanisms are useful because their collective, or interactive, behavior causes most of their interesting properties. In effect, when simple principles are joined interactively, they can generate enormously complicated properties. Such collective properties often appear in psychophysiological data, and often generate erroneous interpretations because they are not obvious consequences of the dynamics of single cells. An analysis of single cells alone does not reveal which single-cell properties generate important collective properties, or whether prescribed collective properties are insensitive to wide variations in certain single-cell properties. The types of parallel, nonlinear, self-organizing, hierarchical, and feedback interactions that are commonplace in psychophysiological data must be explicitly modeled in order to derive accurate conclusions about them.

The present theory suggests a way of thinking no less than a series of

* This work was supported in part by the Advanced Research Projects Agency of the Office of Naval Research (N00014-70-C-0350).

mechanisms. In particular, it suggests how the probabilistic and computer models that have been used, in somewhat complementary fashion, to analyze memory data can be modified, unified, and strengthened. Psychophysiological models provide a natural framework for this synthesis because they must routinely deal with the evolution of patterned activities within hierarchically organized networks. Such models also synthesize serial and parallel processing properties into a unified framework, and weave together phenomena about development, perception, learning, and cognition into an interactive portrait. If nothing else, the method of deriving complex phenomena and predictions from simple environmental pressures confronts us with unexpected and nontrivial consequences of our present beliefs, and provides a rigorous and transparent conceptual superstructure with whose aid new concepts can be more effectively fashioned.

Another basic property of much psychophysiological data is their evolutionary character, whether due to the development of species, the development of individuals, or individual learning. The theory tries to respect the wisdom of evolution by imitating it. At each successive stage of theory construction, prescribed environmental pressures determine a definite class of network principles and mechanisms, and mathematical analysis shows what these mechanisms can and cannot do. As more sophisticated pressures are considered, the earlier principles and mechanisms provide a substrate on which newer principles and mechanisms are superimposed. Similarly, by imposing ever-more-demanding variations on the same problem, we find a sequence of related networks capable of ever-higher levels of behavioral sophistication. Such a sequence illustrates the evolution of a network principle in response to an environmental pressure. Of particular interest in the present work is the evolution of serial order in behavior.

The paper's structure imitates this evolutionary method, subject to space limitations. It is self-contained and written for an audience of nonspecialists. The remainder of this section motivates some central themes of the paper in a heuristic fashion.

A. DOES MEMORY PRESERVE ORDER?

We shall consider a maze as illustrative of the many situations in which there exists a succession of choice points leading to a goal, such as in walking from one room of a house to another room at the other end of the house, or from home to school. (see Fig. 1). Suppose that one leaves the filled-in start box and is rewarded with food in the vertically hatched goal box. Every successful transit from the start box to the goal

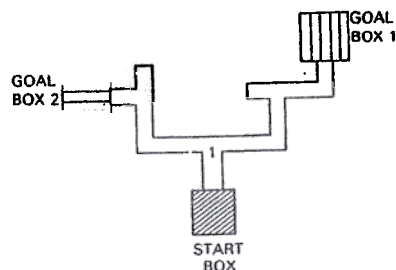


FIG. 1. Correct performance from the start box to a goal box is always order-preserving, with the goal box occurring last.

box requires the same sequence of turns at choice points in the maze. Correct performance is therefore *order-preserving*, and the goal box *always* occurs *last*. In some sense, therefore, correct performance requires that our memory traces remember the order in which events occur. The most naive possibility is that choice points are somehow organized in a chain, as in Fig. 2. Such an encoding is clearly insufficient, however, if the sequence of choices is triggered within the start box by the desire to attain the goal. For example, if I am sitting in my office and decide to go to the cafeteria for lunch, I can then elicit a characteristic series of sensory-motor coordinations that end by eating lunch. This could never happen using the mechanism of Fig. 2. In such a world, if a friend stopped me while I was walking down the hall and asked where I was going, I could only say, "I don't know. I'll tell you when I get there," because the goal in Fig. 2 always occurs last and is inaccessible to me until I reach the last link in the chain. In Figure 2, there is no behavioral *plan*. In goal-directed behavior, by contrast, an internal representation of the goal occurs *first*, and this representation somehow triggers the behavior that can lead to goal attainment.

This state of affairs can be rephrased as the *Goal Paradox*: How can the goal representation occur both last and first? More precisely, in *all* of our experiences with the goal, it is the last event to occur. This makes it plausible that our memory-traces order our choices so that behavior appropriate to the goal occurs last. However, if these memory traces are

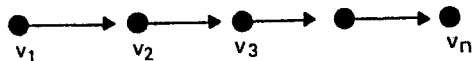


FIG. 2. A chain of associations can accurately code order, but it is insufficient to achieve goal-oriented serial behavior.

order-preserving, as is necessary to actually reach the goal, and if the goal always occurs last, how can an internal goal representation also be activated first, as is necessary for this representation of the goal to trigger a compatible behavioral plan?

We want an internal representation of the goal to trigger a plan that controls a sequence of acts leading to the goal. What we are demanding is schematically drawn in Fig. 3a, where we indicate by points and arrows, respectively, the minimal dimensions of the problem and directed influences between these dimensions. The events (for example, choices in a maze) have internal representations that are designated by states $v_1, v_2, v_3, \dots, v_n$, where n is the index of the goal. The plan is a state that somehow organizes the order in which the events will occur—hence the arrows from plan to events. The state corresponding to the plan must be determined by the events themselves, since during a correct sequence of choices on a learning trial, only these events occur. This dependence is schematized by the upward-directed arrow in Fig. 3b. Thus the events determine the state that will represent the plan, and this state thereupon gains control of the event-representations themselves. Simultaneously, an internal trace of the goal gains control of the plan. Given such a picture, albeit vague at this stage, several definite design problems emerge:

(a) What mechanism maintains the activity of the plan throughout the

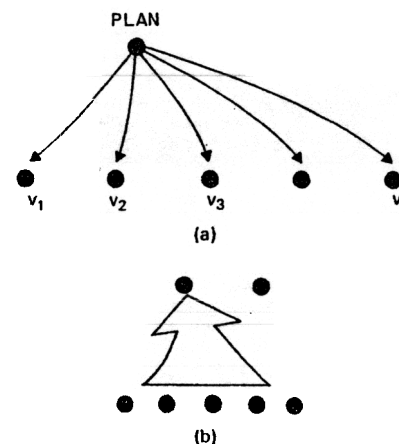


FIG. 3. (a) The goal representation organizes the individual commands $i = 1, 2, \dots, n$; (b) the individual commands help to choose the plan.

presentation of all the events so that the plan knows which events to control?

(b) What mechanism tells the plan which of the events came first, so that it will be able to perform them in the proper order?

Such questions apply to a host of related situations. To illustrate the breadth of the problem, we consider a similar problem in language learning and performance. Suppose that I wish to say a long word, such as MAGNETOHYDRODYNAMIC. There is a clear intuitive sense in which I am ready to say the whole word at a given instant of time; that is, the "idea" or "plan" of saying the word is active in my mind at that instant, yet the actual elicitation of the plan occurs serially, one behavioral unit after another. A similar problem arises when we say the name of a familiar person or object that is visually perceived. How can I establish a command that can "see" the whole word at a given time, yet also organizes the serial performance of its parts? Clearly, a picture such as Fig. 3 is again called for, and the same formal problems must be solved in synthesizing a mechanism that justifies the picture.

In the hall-walking problem, we are considering how to control muscles in our arms, legs, eyes, etc., using visual and proprioceptive feedback, etc. In the word-elicitation problem, we are considering how to control muscles in our mouth, throat, larynx, diaphragm, etc., using auditory and proprioceptive feedback, etc. Both situations address common problems about how sensory-motor loops between particular modalities are serially organized by command structures or plans. A related problem is playing a piece on the piano, in which the sensory-motor loops that develop clearly depend on both visual and auditory modalities during various phases of reading and performing the piece.

B. FREE VERSUS FORCED PARAMETERS IN A PLAN

Suppose that I am navigating a maze (for example, a hallway) on my way to lunch. On the way, I can stop for a drink of water or to chat with a friend. Or, more simply still, I can walk quickly to the cafeteria, or I can stroll leisurely, pausing along the way to rest. These events are not preprogrammed by my plan to get lunch, but they are compatible with the plan. By contrast, I must not make the wrong turn at a choice point, or I will never reach the goal. Thus, some of my behavior is under tight control, such as how to respond to prescribed choice-point cues, but the rest of my behavior is quite undetermined by the plan, in particular the velocity with which the plan is executed.

A similar temporal freedom occurs when I say a word or play a piece

on the piano. Within rather broad limits, I can say certain syllables more slowly than others, or play certain passages more slowly than others. The *order* information in the plan is not tightly coupled to the *velocity* of performance.

Uncontrolled parameters in a plan are also of a more subtle type. For example, a common phrase in language might have a rigidly controlled order, yet the plan might allow a wide choice of nouns or verbs to be fit into the phrase between the rigidly controlled items—comparable to the rigid control of choice points in a hallway versus the freedom to do other things between the choice points. Or a certain number of equivalent phrases can be chosen to express an idea, just as a variety of techniques for executing the correct turns in the hallway will all lead to the goal. Similar remarks can be made about playing a piece on the piano: Certain passages can be played as tightly coupled units, so much so that it is difficult to start playing them in the middle.

C. CIRCULAR REACTIONS

To execute a sequence of sensory-motor coordinations, one must first be able to execute one member of the sequence. Even at the level of individual sensory-motor acts, there is a decoupling of order information (or positional information) from velocity information. For example, I can plan to move my hand to a fixed terminal position, and can move it there at a wide range of velocities. "Knowing" where I want to move my hand and "willing" it to move are not the same operation. Similarly, being "ready" to turn right in the hallway when I see a certain cue does not determine how fast I will turn right.

What is the cue that tells my hand where to move? Suppose that my eyes are focused on a certain object. The tilts of my neck, head, and eyes, along with the vergence of my eyes, etc., establish proprioceptive coordinates that determine the relative position of the object from my body. Somehow these coordinates must get translated into commands to the muscles in my hand and arm that correspond to the correct terminal coordinates of the hand on the object. In other words, the *proprioceptive map* of the head, neck, eyes, etc., excites a *terminal motor map* of the hand. "Willing" the hand to move releases the information in the motor map, and makes the hand move.

How does the transformation between maps get established? Because there exists so many individual differences in body parameters between individuals, it seems clear that much of the transformation must be learned. Piaget (1963) has carefully observed the development of the ability in young children. He notes that at first an infant's hand makes a

series of unconditional motions, which the infant's eyes unconditionally follow. As the hand occupies a variety of positions that the eye fixates upon, a map is learned from the proprioceptive coordinates of the hand-arm system to the motor coordinates of the eye-head-neck system, and *conversely*, from the proprioceptive coordinates of the eye-head-neck system to the motor coordinates of the hand-arm system. Using the map from eye-head-neck proprioception to hand-arm motor coordinates, we can move our hands to a fixed object. During the learning trials, the eyes try to continuously follow the motions of the hand, or if they fall behind, they must try to catch up by leaping to the correct position via saccades. Since the eye always tries to fixate the present position of the hand, the two transformations between proprioceptive maps and motor maps code only the (approximately) present motor positions. During performance trials, the transformation from eye-head-neck proprioception (or where we are now looking) to hand-arm motor coordinates therefore determines only the *terminal* position of the hand (or where we *want* the hand to go). If the initial position of the hand is very different from the desired terminal position, then the directed motion of the hand can be viewed as a saccade of the hand. To say that positional and velocity information are decoupled translates into the statement that the saccadic velocity is not preprogrammed in this system.

The above observations can be reformulated to emphasize an important point. Since the terminal motor map suffices to guide the hand throughout its trajectory from initial to terminal position, all intermediate positions of the arm-hand system must be derivable from this information. What auxiliary feedback mechanisms within the hand-arm system translate the terminal motor map into a physically realizable trajectory of this system? More precisely, the transformation between maps codes only where the hand is destined to go, but not how it gets there. It ignores the properties of the arm-hand system as a mechanical system, and codes only the plan. In particular, on each performance trial aimed at extending the hand to a fixed position, any of the free parameters—such as hand velocity—can be chosen differently, and can thereby alter the forces on the system, even though the plan remains the same. Somehow these varying mechanical properties must be controlled by auxiliary mechanisms, which average them away, so that the invariant plan can be realized.

D. THE INTERNAL STRUCTURE OF MAPS AND THEIR TRANSFORMATIONS

If directed reaching for an object is controlled by a transformation between maps, then a tremendous reduction in the amount of informa-

tion that must be stored has been achieved. Indeed, suppose that the transformation has been learned up to a given level of accuracy on a finite number of learning trials. Then, without *any* further learning, any of the infinitely many reachable positions of an object can be approached by the hand, under eye-head-neck guidance, up to this level of accuracy.

This assertion tacitly assumes that, if a set of proprioceptive coordinates \mathcal{P}_1 has been associated with a terminal motor map \mathcal{M}_1 , and a different set of proprioceptive coordinates \mathcal{P}_2 has been associated with a terminal motor map \mathcal{M}_2 , then a new set of proprioceptive coordinates \mathcal{P}_3 will have the following effects:

- It partly excites \mathcal{P}_1 and \mathcal{P}_2 both, with an intensity that depends on how similar \mathcal{P}_3 is to \mathcal{P}_1 and \mathcal{P}_2 .
- \mathcal{P}_1 and \mathcal{P}_2 will excite \mathcal{M}_1 and \mathcal{M}_2 , respectively, with an intensity that depends on how excited they are by \mathcal{P}_3 .
- The mixture of motor excitation will form a hybrid terminal motor map \mathcal{M}_3 that is between \mathcal{M}_1 and \mathcal{M}_2 , which moves the arm closer to the position that excited \mathcal{P}_3 than either \mathcal{M}_1 or \mathcal{M}_2 could have separately (see Fig. 4). In other words, each proprioceptive representation has a *generalization gradient*. Representations excite each other with an intensity that depends on how close they lie with respect to each other on their gradients. Each motor map also has a generalization gradient. The fact that \mathcal{M}_3 can be synthesized from \mathcal{M}_1 and \mathcal{M}_2 to yield a position close to the one determined by \mathcal{P}_3 means that the transformation from

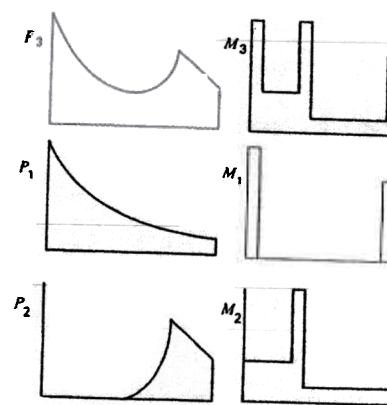


FIG. 4. Filtering of \mathcal{P}_3 as a weighted average of \mathcal{P}_1 and \mathcal{P}_2 followed by synthesis of \mathcal{M}_3 as a weighted average of \mathcal{M}_1 and \mathcal{M}_2 .

proprioceptive maps to terminal motor maps preserves (at least approximately) the distances between representations that are defined by the generalization gradients.

An instructive example occurs in language learning. Again there exists a period during which unconditional behavior is emitted (Fry, 1966). Instead of unconditional motions of the eye-hand system, there exists a period when the infant babbles various simple sounds. Instead of the eye following the hand so that a transformation from proprioceptive feedback to terminal motor position can be learned, the internal trace of the auditory sensory feedback elicited by a babbled sound is conditioned to the motor coordinates that produced the sound. The babbling phase cannot go on forever; if it did, the unconditional urge to babble would forever interfere with the desire to say something interesting. When babbling stops, a certain number of connections exist between prescribed auditory representations and the motor controls that produced them. Exciting one of these representations can elicit the appropriate sound, so that simple imitation begins to be possible. How does an infant learn more complex sounds than the ones that occurred during babbling? One way is to suppose that more complex sounds are decomposed into weighted combinations of the simpler sounds that already are capable of eliciting speech sounds. If these simpler sensory representations map into a motor speech space that preserves their mutual distances, then the speech sound that is synthesized in this way will be closer to the heard sound than any of the simpler sounds. The system can hereby try to imitate more complex sounds than are originally in its repertoire, and to build internal representations of these.

E. CONTEXT-DEPENDENT CHOICES

In the maze of Fig. 1, let there be more than one goal box. For example, let a food reward be in the vertically hatched box and a sexual reward be in the horizontally hatched box. At choice point number 1, a right turn leads to food and a left turn leads to sex, in response to the same external sensory cues. The choices are controlled by different plans, which create different contexts in which to respond to the cues. Figure 5 schematizes this situation. In Fig. 5, plan 1 excites a given pathway in response to the i th cue, whereas plan 2 excites a different pathway in response to the i th cue. How does the convergence between a given event and different plans excite different responses as a result of learning trials?

This problem also occurs in many situations. For example, how can

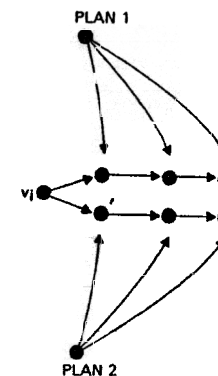


FIG. 5. Plans 1 and 2 excite different pathways in response to the i th cue representation v_i .

we serially learn both lists ABCDEFGH and ABAFALAK? In the two lists, the responses to B are different even though both lists have the same beginning AB, and in the second list the letter A is a precursor of four different letters. Clearly the stimulus for saying the next letter is not merely the previous letter, or even the previous two letters. Longer subsequences of letters somehow distinguish the two lists in our memory. Again Fig. 5 is called to mind.

F. CHUNKING AND FEEDBACK

The above remarks can be rephrased by saying that the units that control behavior are not necessarily representations of individual stimuli or responses, but can be built up from lists or other aggregates of these units. Such composites, or chunks (Miller, 1956), make possible a recoding of memory that enables ever-more-complex commands to form. For example, if an adult had to pay attention while walking to every step taken as a complex juxtaposition of motor events in different joints, guided by sensory feedback, then it would be difficult to pay attention to anything else while walking. Once these events are organized by higher-order commands or chunks, by contrast, the details of walking become simple, and attention can be devoted to other tasks. Similarly, once a long word is organized by a higher-order command, it can be treated as a single behavioral unit rather than as a complicated series of mouth, tongue, larynx, diaphragm, and related motor activities, with attendant sensory feedback, whose conscious control would interfere with thinking about other things.

Lashley (1951) noted that a pianist can play successive notes so fast

that feedback from the last note played could not possibly influence how the next note would be played. Clearly, however, making a mistake can influence our playing of later passages. Also there are many data on the importance of auditory feedback on speech production (Lenneberg, 1967), even though it is clear that the commands controlling language are not individual letters.

How can feedback be unimportant in some cases and important in others? There is no paradox if we say that feedback is important when it reorganizes the structure of the plan. A single command can guide the performance of an entire word, or of a sequence of notes on the piano, without feedback. But feedback can reorganize which commands will be active after this sequence is played, or even before it is played in its entirety.

G. PERFORMING THE FIRST OR THE LAST ITEM

A picture such as Fig. 3 does not tell us how events are performed in a given order. Daily experience suggests some important constraints on this mechanism. Suppose that the letters ABARFD are said to me one at a time. I can be told to do any of at least four things:

- (a) Repeat each letter aloud as soon as I hear it.
- (b) Listen passively to the list and, after hearing it, repeat the list items in their proper order.
- (c) Repeat each letter aloud as soon as I hear it, and repeat the list items in their proper order after the whole list is presented.
- (d) After completing task (a), I can be asked to repeat the whole list in its proper order.

In task (a), I am being asked to repeat the *last* thing that I have heard. In some sense, the last item must have the greatest "weight," so that I can choose it above all others. By iterating this requirement, each item has greater weight than the preceding items. In task (b), I am being asked to repeat the *first* thing that I have heard. In some sense, the first item must also have the greatest "weight," so that I can choose it above all others (see Fig. 6a). Furthermore, after saying the first item, its "weight" must be decreased, so that I can say the second item, whose weight is then greater than all other weights. After saying the second item, its "weight" must decrease, etc. But this means that the last item has the *least* weight in task (b). How can it also have the *greatest* weight in task (a)? Moreover, tasks (c) and (d) mix tasks (a) and (b), so that the

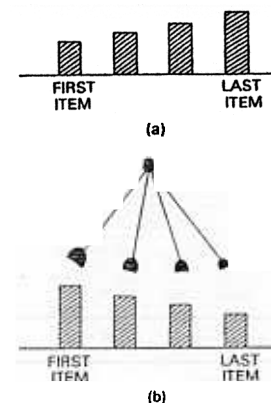


FIG. 6. (a) The most recent item has the greatest weight in task (a); (b) the least recent item has the greatest weight in task (b).

last item must be able to have the least and the greatest weight in the same situation! How can this be?

It is intuitively clear that the most recent item "should" have the greatest weight in the sense of task (a). Surely the most recent events are more "salient" than earlier events. It is also clear that something like a command, or plan, develops to perform tasks (b), (c), or (d), since the letter A leads to B or R, depending on the context. The dilemma of performing tasks (a) and (b) therefore can be rephrased as follows: How does activation of the plan *reverse* the weights of its individual events? (See Fig. 6b.)

In the above example, I could also be told to do the following:

- (e) Repeat every pair of successive letters when it occurs. Then repeat the whole list in its proper order.

To repeat a given pair of items, the first item in the pair must have the greater weight. This is true for *every* pair, so that we can no longer talk about weights that increase monotonically or decrease monotonically with list position as in Fig. 6. It is tempting to instead draw a picture such as that shown in Fig. 7. Task (a) differs from task (e) only in their rehearsal strategies. How does rehearsal reorganize plans as in Figs. 6 and 7? Actually, familiar letter sounds, such as the sound of A, can be composed of more than one sound component, as slow pronunciation clearly indicates. Prior experience has organized these components into a single letter via a suitable command. In this sense, Fig. 7 is a

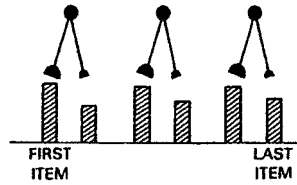


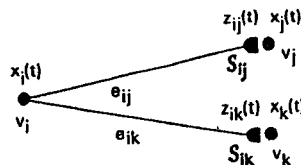
FIG. 7. Rehearsal organizes plans and their order information

refinement of Fig. 6 even in task (a); in this situation, the paired arrows are the commands for eliciting individual sound components of the same letter. Thus the general problem is: How does rehearsal group individual motor acts into a hierarchy of commands?

2. Stimulus Sampling of Spatial Patterns

The neural networks of this paper are built out of network components that have been derived elsewhere from psychological postulates and mathematically analyzed (see, for example, Grossberg 1967, 1969a,b, 1970a, 1971a, 1972a). They are sketched herein for completeness. The material in Sections 2 through 5 is more completely reviewed in Grossberg (1974).

The first stage of the theory analyzes the simplest concepts of classical conditioning: How does pairing of a conditioned stimulus (CS) with an unconditioned stimulus (UCS) on learning trials enable the CS to elicit a conditioned response (CR), or UCS-like event, on performance trials? This analysis yields a psychophysiological theory operating in real time. Psychological inputs, or stimuli, representing particular experiments perturb a neural network that elicits definite outputs or responses. The network dynamics are described by interactions between the *short-term memory* (STM) traces $x_i(t)$ of cell body populations v_i , and the *long-term memory* (LTM) traces $z_{jk}(t)$ of the axonal pathways e_{jk} from v_j to v_k , as in Fig. 8. The simplest realization of these interactions among n

FIG. 8. STM traces $x_i(t)$ and LTM traces $z_{jk}(t)$ in cell body populations and axonal pathways, respectively.

populations v_1, v_2, \dots, v_n is given by the system

$$\dot{x}_i = -A_i x_i + \sum_{k=1}^n B_{ki} z_{ki} - \sum_{k=1}^n C_{ki} + I_i(t) \quad (1)$$

and

$$\dot{z}_{jk} = -D_{jk} z_{jk} + E_{jk} x_k \quad (2)$$

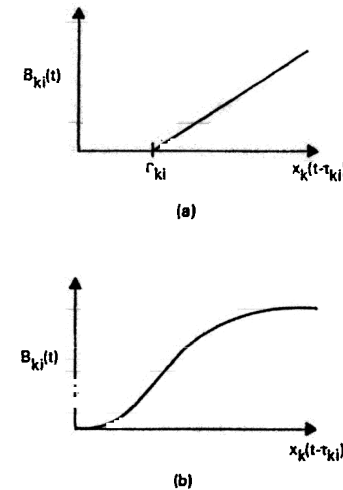
where $i, j, k = 1, 2, \dots, n$. The terms in (1) and (2) have the following interpretation. Function A_i in (1) is the decay rate of the STM trace x_i . Function B_{ki} in (1) is a performance signal from v_k to the synaptic knobs S_{ki} of e_{ki} . Two typical choices of B_{ki} are

$$B_{ki}(t) = b_{ki} [x_k(t - \tau_{ki}) - \Gamma_{ki}]^+ \quad (3)$$

where $[\xi]^+ = \max(\xi, 0)$ for any number ξ , or

$$B_{ki}(t) = f(x_k(t - \tau_{ki})) b_{ki} \quad (4)$$

where $f(\xi)$ is a sigmoid (S-shaped) function of ξ with $f(0) = 0$. In (3), signals leave v_k only if x_k exceeds the signal threshold Γ_{ki} (Fig. 9a) and reach S_{ki} after τ_{ki} time units; in (4), the signal threshold Γ_{ki} is replaced by attenuation of the signal at small x_k values (Fig. 9b). Such a population signal from v_k is generated, for example, if the signal thresholds of cells in v_k are Gaussianly distributed around a mean

FIG. 9. (a) Signal $B_{ki}(t)$ is positive only if $x_k(t - \tau_{ki})$ exceeds the threshold Γ_{ki} ; (b) signal $B_{ki}(t)$ is attenuated at small $x_k(t - \tau_{ki})$ values.

threshold value. Term $B_{ki}z_{ki}$ in (1) says that the signal B_{ki} from v_k to S_{ki} interacts with the LTM trace z_{ki} at S_{ki} . In particular, z_{ki} *gates* signal B_{ki} , so that the signal strength that perturbs x_i at v_i is $B_{ki}z_{ki}$ rather than B_{ki} . Thus, even if an input to v_k excited equal signals B_{ki} in all the pathways e_{ki} , only those v_i such that z_{ki} is large will be appreciably excited by v_k . All such gated signals from populations v_k combine additively at v_i in term $\sum_{k=1}^n B_{ki}z_{ki}$. The term $\sum_{k=1}^n C_{ki}$ in (1) describes the total effect of inhibition from cells v_k upon v_i . The choice

$$C_{ki}(t) = g(x_k(t - \sigma_{ki}))c_{ki}$$

with $g(\xi)$ a sigmoid signal function, is illustrative. The function $I_i(t)$ in (1) is the input corresponding to presentations of the i th event through time; $I_i(t)$ is large when the stimulus is presented, and otherwise equals zero. In all, the STM trace x_i can spontaneously decay, be excited by external stimuli, and interact with other populations via sums of gated excitatory signals and inhibitory signals. The net size of x_i after all these processes operate determines whether v_i will generate an output.

Function D_{jk} in (2) is the decay rate of the LTM trace z_{jk} . Function E_{jk} in (2) describes a learning signal from v_j to S_{jk} which drives the LTM changes in z_{jk} at S_{jk} . In other words, v_j *samples* v_k by turning on E_{jk} . In the simplest cases, E_{jk} is proportional to B_{jk} , but this is not necessary. It is only necessary to prevent B_{jk} from being large over a sustained time interval if E_{jk} is small over that interval (Grossberg, 1969b, 1971a, 1972a). This occurs automatically if the LTM trace is computed at S_{jk} ; since the signal from v_j passes through S_{jk} on its way to v_k , the thresholds for B_{jk} are then no smaller than the thresholds for E_{jk} . The term $E_{jk}x_k$ in (2) shows how a pairing of the j th and k th events influences the growth of the LTM trace z_{jk} . By joining together terms $-D_{jk}z_{jk}$ and $E_{jk}x_k$, we conclude from (2) that the LTM trace is a time average of the product of learning signals from v_j to S_{jk} with STM traces at v_k . When z_{jk} changes in size, it thereupon alters the gating of signals from v_j to v_k via term $B_{jk}z_{jk}$ in (1).

Two facts are of crucial importance in these systems:

- A. The unit of LTM is a *spatial pattern*.
- B. There exists a *stimulus sampling* operation.

By (A) we mean the following: Consider the network in Fig. 10a. It has the minimal anatomy capable of learning by classical conditioning. The network represents a population v_0 that receives a CS-activated input. Population v_0 can send signals to its axon collaterals, which abut on the UCS-activated populations v_1, v_2, \dots, v_n . The LTM traces z_{0i}

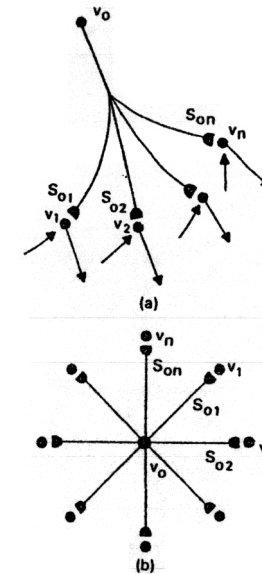


FIG. 10. (a) The CS-activated population v_0 samples populations v_1, v_2, \dots, v_n ; (b) the *outstar* is the minimal network capable of classical conditioning.

are computed at the synaptic knob terminals S_{0i} of the v_0 axon collaterals e_{0i} . Such a network is called an *outstar* because it can be redrawn as in Fig. 10b.

An outstar can learn an arbitrary spatial pattern. A *spatial pattern* is a UCS to the cells v_1, v_2, \dots, v_n whose intensities have a fixed relative size through time; that is, $I_i(t) = \theta_i I(t)$, for some $\theta_i \geq 0$ such that $\sum_{k=1}^n \theta_k = 1$. For example, suppose that the UCS is a picture playing across the "retina" of cells v_1, v_2, \dots, v_n . The total intensity of white light that illuminates the picture can be varied through time without changing the picture itself. The *relative* intensities of light (or *reflectances*) reflected from various points in the picture characterize it, and these remain constant through time (Cornsweet, 1970). The function $I(t)$ is the total UCS input intensity, which can fluctuate wildly through time. The constant relative intensities $\theta = (\theta_1, \theta_2, \dots, \theta_n)$ characterize the spatial pattern. In short, an outstar can learn an arbitrary spatial pattern of relative figure-to-ground. Thus, the unit of LTM cannot be determined by measurements from just one population v_i ; *parallel* measurements are needed to test whether the *relative* intensities are changing through time.

The *stimulus sampling probabilities* of an outstar are the relative LTM traces

$$Z_{0i} = z_{0i} \left(\sum_{k=1}^n z_{0k} \right)^{-1} \quad (5)$$

As CS–UCS pairing takes place, the functions Z_{0i} approach θ_i . During later performance trials, a CS input to v_0 creates signals in the e_{0i} axons. These signals are gated by the LTM traces z_{0i} . Since the z_{0i} are proportional to θ_i , the gated signal to v_i is proportional to θ_i . The CS hereby elicits responses in the STM traces x_i that are proportional to θ_i . In short, after CS–UCS pairing, the CS can reproduce the pattern θ .

Stimulus sampling means that the functions Z_{0i} can change *only* when signals from v_0 reach the synaptic knobs S_{0i} . Unless the CS perturbs these knobs, their LTM traces cannot “see” what UCS patterns are received at the cells v_1, v_2, \dots, v_n . This is because the learning signals E_{0i} in (2) vanish unless a CS perturbs v_0 . This interpretation of stimulus sampling in an outstar can be extended to a more general neural interpretation of stimulus sampling that modifies and generalizes Estes’ theory of amplifier elements (Grossberg, 1972b,c).

3. Sensory Codes and Motor Synergies

The outstar is a general-purpose device. It can learn a spatial pattern of activity playing across whatever cells its knobs S_{0i} sample. For example, suppose that the cells v_1, v_2, \dots, v_n are feature detectors in a sensory cortex of a network. By this we mean the following. When a picture is presented to the network’s retina, the picture is analyzed in such a way that each v_i responds most vigorously to particular features (for example, color, orientation, disparity) in a prescribed retinal region. Each picture hereby generates a spatial pattern of activity across the feature detectors of v_1, v_2, \dots, v_n . This pattern is a coded internal representation of the picture. An outstar can learn and reproduce *any* such representation with complete fidelity.

Alternatively, suppose that the cells v_1, v_2, \dots, v_n are motor control cells. In this case, each v_i can excite a particular group of muscles, and a larger signal from v_i causes a faster contraction of its target muscle group. A spatial pattern across v_1, v_2, \dots, v_n then codes fixed relative contraction rates across many muscle groups: for example, playing a chord on the piano with prescribed fingers; or withdrawing a hand with fixed relative speeds of wrist, elbow, and shoulder motion; or forming a particular configuration of lips and tongue when uttering a sound. In

other words, the outstar can learn any motor synergy in which prescribed relative rates exist across a family of muscle groups. An increase of the CS input speeds up all the muscle contractions at their fixed relative rates; that is, the CS can perform the synergy at an arbitrary rate.

In summary, a single outstar can coordinate, through its parallel pathways e_{0i} , the learning and reproduction of a distributed sensory code or a synergistic pattern of motor commands. Not all sensory and motor acts have fixed relative figure-to-ground, but using the results on outstars shows how to approach the coding and performance of arbitrary sequences of events.

4. Ritualistic Learning of Arbitrary Acts

The properties of stimulus sampling and of encoding in spatial pattern units show how to learn an arbitrary act, such as a piano recital, a dance, or a sequence of sensory images, in a minimal way (Grossberg, 1969c, 1970a, 1974). The simplest example describes a ritualistic encoding, wherein performance is insensitive to environmental feedback. In this case, only one cell is needed to encode the memory of an arbitrary act. This fact shows that the encoding of complexity per se is relatively easy. In fact, nervous systems with few cells can activate complicated behaviors, as is well known in invertebrates (Dethier, 1968; Kennedy, 1968; Willows, 1968). The ritualistic construction is also universal; such a cell can encode *any* act. The genetic code for such a cell need not concern itself with which act will eventually be encoded. The ritualistic construction focuses our attention on deeper questions concerning the global organization of memory when environmental feedback is operative, and suggests mechanisms of encoding that are sensitive to environmental feedback.

Suppose that the act to be learned is controlled by the cells $v_1^{(1)}, v_2^{(1)}, \dots, v_n^{(1)}$ in a field of cells $\mathcal{F}^{(1)}$. Each $v_i^{(1)}$ might be a feature detector, a motor control cell, a hormonal source, an interneuron—anything you like. The number n of cells being controlled can be chosen arbitrarily large. Let each cell $v_i^{(1)}$ receive a nonnegative and continuous input $I_i(t)$, $t \geq 0$, $i = 1, 2, \dots, n$. Any such input is covered by our analysis. A particular choice

$$J(t) = (I_1(t), I_2(t), \dots, I_n(t)), \quad t \geq 0$$

of inputs controls a given act. In intuitive terms, $J(t)$ describes a moving picture playing on the cells $v_1^{(1)}, v_2^{(1)}, \dots, v_n^{(1)}$ through time. The

movie shall be learned and performed as a sequence of still pictures that are smoothly interpolated in time.

Because each $I_i(t)$ is continuous, the functions

$$\theta_i(t) = I_i(t) \left[\sum_{k=1}^n I_k(t) \right]^{-1} \quad (6)$$

are also continuous. As in the case of moving pictures, any continuous function $\theta_i(t)$ can be arbitrarily well approximated by a sequence of its values

$$\theta_i(\zeta), \theta_i(2\zeta), \theta_i(3\zeta), \theta_i(4\zeta), \dots$$

sampled every ζ time units, if ζ is chosen so small that the functions $\theta_i(t)$ do not change too much in a time interval of length ζ . For every fixed k , the numbers

$$\theta^{(k)} = (\theta_i(k\zeta), \quad i = 1, 2, \dots, n)$$

sampled across all the cells $v_1^{(1)}, v_2^{(1)}, \dots, v_n^{(1)}$ at time $t = k\zeta$ form a spatial pattern. To learn and perform the movie $J(t)$, $t \geq 0$, it therefore suffices to learn and perform the spatial patterns $\theta^{(1)}, \theta^{(2)}, \theta^{(3)}, \dots$ in the correct order. This can be done if a sequence of outstars O_1, O_2, O_3, \dots is arranged so that O_k samples just spatial pattern $\theta^{(k)}$ on successive learning trials, and is then briefly activated in the order O_1, O_2, O_3, \dots on performance trials. An *avalanche*-type anatomy, such as that in Fig. 11, accomplishes this by using the minimum number of spatial dimensions. In Fig. 11, a brief CS-activated sampling pulse travels along the long axon of cell $v_1^{(2)}$ from left to right, and down its serially arranged

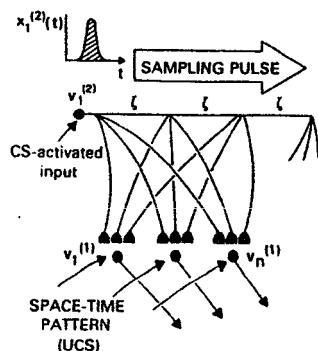


FIG. 11. An avalanche is the minimal network that can ritualistically learn any space-time pattern.

bouquets of axon collaterals. Each bouquet (really an outstar) can learn a spatial pattern, and successive bouquets are activated every ζ time units. No bouquet can see more than one pattern $\theta^{(k)}$, because of the property of stimulus sampling. On performance trials, the CS-activated pulse runs along the axon, serially exciting the bouquets and their learned spatial patterns; the space-time pattern is hereby synthesized from its ordered spatial components $\theta^{(k)}$. Thus one cell, albeit in possession of many axon collaterals, can code the memory of an arbitrary act.

5. Nonspecific Arousal as a Command

Once a pulse is emitted by $v_1^{(2)}$, there is no way to stop it in Fig. 11. If, for example, the avalanche controlled the performance of a long dance, and the stage on which the dance was being performed began to burn, there would be no way to stop the dance in mid-course to escape the flames. Sensitivity to environmental feedback is possible only if the pulse can be abruptly terminated as it travels along the avalanche axon (Grossberg, 1969c, 1970a, 1971b, 1974). By considering the minimal way to do this, we find an anatomy that is isomorphic with that discovered for command neurons in various invertebrates—for example, in the control of the rhythmic beating of crayfish swimmerets (Stein, 1971).

The avalanche must be modified so that performance can be terminated at loci all along the axon of the CS-activated cells $v_1^{(2)}$. Consider Fig. 12a. In Fig. 12a, cell bodies $v_1^{(2)}, v_2^{(2)}, v_3^{(2)}, \dots$, forming a field of $\mathcal{F}^{(2)}$ of cells, are interpolated at every outstar source. Performance still cannot be terminated if a signal from $v_1^{(2)}$ suffices to fire a signal at $v_{i+1}^{(2)}$. Figure 12b remedies this situation in a minimal way. The new population $v_1^{(3)}$ can supply a signal that reaches all the populations $v_1^{(2)}, v_2^{(2)}, \dots$ (approximately) simultaneously. Require that $v_{i+1}^{(2)}$ can fire a signal *only* if it receives an input from $v_i^{(2)}$ and $v_1^{(3)}$ simultaneously. Withdrawal of the signal from $v_1^{(3)}$ can therefore abruptly terminate output from the avalanche, since $v_{i+1}^{(2)}$ cannot fire even if it receives input from $v_i^{(2)}$. In this sense, $v_1^{(3)}$ supplies an *arousal* input to the avalanche. Because this input is delivered to all populations $v_1^{(2)}, v_2^{(2)}, \dots$, it is a *nonspecific* arousal input. Population $v_1^{(3)}$ supplies a *command* signal that prepares the avalanche to fire in response to the CS input to $v_1^{(2)}$.

Such command neurons are familiar in the control of behavioral acts by invertebrates (Dethier, 1968; Kennedy, 1968; Willows, 1968). If changes in the LTM traces of the avalanche in Fig. 12 are prevented [set $D_{jk} = E_{jk} = 0$ in (2)], then this network is capable of performing

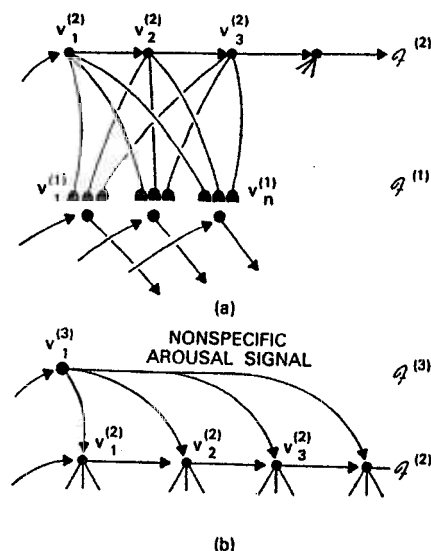


FIG. 12. (a) Interpolating cells $v_i^{(2)}$ along the avalanche cannot terminate performance unless (b) a nonspecific signal is also needed to fire the next cell.

arbitrary acts whose pattern weights are permanently encoded in its synaptic knobs.

Note that varying the size of the arousal signal through time can continuously modulate the speed of performance. Suppose that a given cell $v_i^{(2)}$ starts to receive a signal from $v_{i-1}^{(2)}$ when arousal from $v_i^{(3)}$ is large. Then the total input to $v_i^{(2)}$ is large, so its STM trace grows rapidly, and quickly exceeds its signal threshold. By contrast, if $v_i^{(2)}$ receives a signal from $v_{i-1}^{(2)}$ when arousal from $v_i^{(3)}$ is small, then the total input to $v_i^{(2)}$ is small, so its STM trace grows slowly, and takes longer to exceed its signal threshold. If the arousal signal is too small, the signal threshold is never reached, and the reaction time of $v_i^{(2)}$ is infinite.

6. Self-Organization of Codes and Order Information

Several important themes are made evident by the avalanche example. They are introduced below to motivate our later network constructions.

A. INSTRUMENTAL CONDITIONING

Not every competing event should be able to switch off nonspecific arousal. To cite a colorful example: It is one thing to stop dancing at

your debut if the stage is consumed by flames, and quite another to risk your career because a mosquito is hovering above. Only more important events should be able to shut off the arousal that supports a given act. Knowing what is important to an organism requires, in particular, that we know what events are rewarding or punishing to the organism. What is the relationship between reinforcement and a cue's ability to trigger nonspecific arousal? Grossberg (1971b, 1972a,b, 1975) develops a reinforcement theory that suggests an answer to this question.

From the discussion of arousal as a command, we expect a process akin to that depicted in Fig. 13. In Fig. 13, cue CS_1 excites arousal source A_1 plus its avalanche $v_1^{(1,2)}, v_2^{(1,2)}, v_3^{(1,2)}, \dots$ to elicit sequential performance of its encoded act. When CS_2 occurs, it excites arousal source A_2 plus its avalanche $v_1^{(2,2)}, v_2^{(2,2)}, v_3^{(2,2)}, \dots$. Arousal sources A_1 and A_2 mutually inhibit each other. If A_2 is excited more than A_1 , performance of CS_1 's act abruptly terminates and performance of CS_2 's act commences. How do the cues CS_1 and CS_2 gain control over their arousal sources A_1 and A_2 in cases where such control is not genetically preprogrammed?

B. SERIAL LEARNING OF ORDER INFORMATION

In the avalanche of Fig. 12, a chain of connections from $v_1^{(2)}$ to $v_2^{(2)}$ to $v_3^{(2)}$, and so on, exists in the network at all times. This chain determines the order in which spatial patterns will be performed. Such preprogrammed chains of cells do not generally exist before we learn a sequence of successive acts, such as a piano sonata. For example,

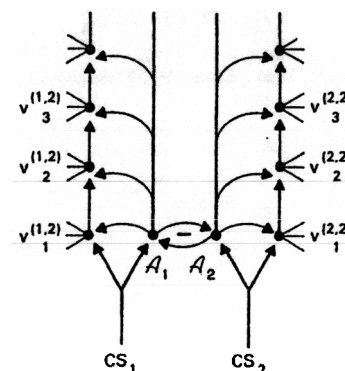


FIG. 13. Competition between arousal sources decides which avalanche will be triggered.

suppose that each $v_k^{(2)}$ controls a different chord in the sonata. Surely there does not exist in every mind a chain corresponding to every possible sequence of chords! This order information must be learned. There exist two conceptually distinct ways in which this can happen. In Fig. 14a, each $v_i^{(2)}$ is initially connected to all cells $v_k^{(2)}$, and eventually becomes differentially connected to $v_{i+1}^{(2)}$ as a result of practicing the sequence $v_1^{(2)}, v_2^{(2)}, v_3^{(2)}, \dots$. This is a problem in serial learning. In Fig. 14b, somehow a higher-order cell population looks at the sequence $v_1^{(2)}, v_2^{(2)}, v_3^{(2)}, \dots$ as it is practiced, and learns to reproduce a spatial pattern of activity across these cells such that the earliest cells have the largest activity. When these differential activity levels are translated into speed of performance, $v_1^{(2)}$ is performed before $v_2^{(2)}$, $v_2^{(2)}$ before $v_3^{(2)}$, and so on. This is again a problem of serial learning. Which strategy of serial encoding is used? The theory of serial learning in Grossberg (1969d) and Grossberg and Pepe (1971) provides a foundation for answering this question.

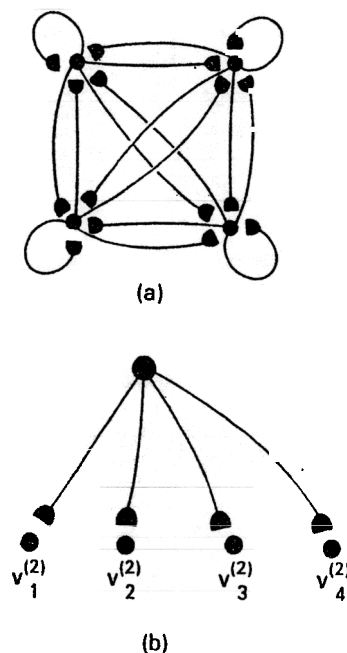


FIG. 14. (a) All populations initially interact equally, but serial learning differentially weights a chain of LTM traces; (b) a command population elicits a graded pattern of activity across a population field, which translates the activities into order of performance.

C. SELF-ORGANIZATION OF CONTEXTUAL COMMANDS OR CHUNKS

If the order information from $v_1^{(2)}$ to $v_2^{(2)}$ to $v_3^{(2)}$, and so on, is not genetically preprogrammed, then the command $v_1^{(3)}$ in Fig. 12 that nonspecifically arouses these cells is also not genetically preprogrammed. In other words, given that the particular sequence $v_1^{(2)}, v_2^{(2)}, v_3^{(2)}, \dots$ depends on the act that it controls, the command $v_1^{(3)}$ that arouses this sequence must be chosen by the sequence itself. Otherwise, there would exist innately prewired commands for every possible sequence, which is absurd. How does a particular sequence code a particular command? How does this coding occur so that different orderings of the same set of cells $v_i^{(2)}$ do not all code the same command?

Given that the command $v_1^{(3)}$ is adaptively coded by its controlled sequence $v_1^{(2)}, v_2^{(2)}, v_3^{(2)}, \dots$, how does $v_1^{(3)}$ learn what sequence it codes? This question immediately calls Fig. 14b to mind. Is the higher-order cell in Fig. 14b a command population $v_1^{(3)}$? Does the learning process that teaches $v_1^{(3)}$ which sequence it controls *automatically* encode the order information needed to perform this sequence correctly? In more abstract terms, are adaptive coding (or chunking) and the learning of order information *dual* processes in a feedback system?

Given that sequence $v_1^{(2)}, v_2^{(2)}, v_3^{(2)}, \dots$ adaptively codes its command, we must also realize that there is nothing special about this particular sequence. Every subsequence of $v_1^{(2)}, v_2^{(2)}, v_3^{(2)}, \dots$ is also a sequence, and must be able to code a command. Not every subsequence will be able to code a command with equal ease. Nonetheless, at every time, there will exist a field $\mathcal{F}^{(3)}$ of command cells, each excited to a different degree by its generating sequence. Each command cell in $\mathcal{F}^{(3)}$ samples the activity patterns that prevail across $\mathcal{F}^{(2)}$ while it is active. Simultaneously, the pattern of activity across $\mathcal{F}^{(3)}$ continually shifts through time as new events are rehearsed, and thereby generates new subsequences to be coded by $\mathcal{F}^{(3)}$. At every time, the total signal from the command field $\mathcal{F}^{(3)}$ to $\mathcal{F}^{(2)}$ helps to determine the order information among the cells in $\mathcal{F}^{(2)}$. Grossberg (1976a,b,c) provides a conceptual foundation for synthesizing the adaptive feedback relationships between the generating sequences and the commands that organize them.

Speaking intuitively, the command cells provide the *context* in which a particular pattern is performed. For example, after playing the first few bars of a piano sonata, a pianist is ready to play the next several bars. While the second movement of the sonata is being played, the first few bars no longer control performance; more recently played notes provide the command context for determining the next notes to be played.

Similarly, if I emerge from my bedroom door, I can decide to turn left to brush my teeth, or to turn right to go directly to breakfast. The difference is decided not by the sensory cues that confront me whenever I am at the door, but by the context, command, or plan that modulates these cues.

The above remarks clarify an important distinction. An adaptively coded command in $\mathcal{F}^{(3)}$ provides a type of *cognitive* context, or arousal, for the populations in its sequence in $\mathcal{F}^{(2)}$. This type of arousal is different from the *motivational* arousal that determines whether performance of a given act continues to be in the best interests of the network. With this distinction, the elicitation of performance by an avalanche would employ an anatomy such as that in Fig. 15. In Fig. 15, the command cell is activated by a particular cue plus motivational arousal. Once activated, the command cell determines a cognitive context that arouses certain populations in preparation for their firing in the correct order.

D. ADAPTIVE CODING OF SENSORY-MOTOR RELATIONSHIPS

Similar coding problems occur on a more microscopic level. In particular, how is it determined which $v_k^{(2)}$ will be chosen to learn a particular chord? For example, while a piano piece is being learned, a $v_k^{(2)}$ is presumably excited by some combination of visual cues, from reading the piano music and seeing the keyboard, and auditory feedback cues, from having played the chord. At this point, characterizing the exact cue combination is unimportant. What is important is that $v_k^{(2)}$ is adaptively coded by its cues in a manner that is strikingly similar to the adaptive coding of a command by a sequence of $v_k^{(2)}$'s. Thus, we are dealing with a problem concerning the *hierarchical* organization of adaptive codes, or chunks, and the feedback signals that order these codes, among emergent fields $\mathcal{F}^{(1)}$, $\mathcal{F}^{(2)}$, $\mathcal{F}^{(3)}$,

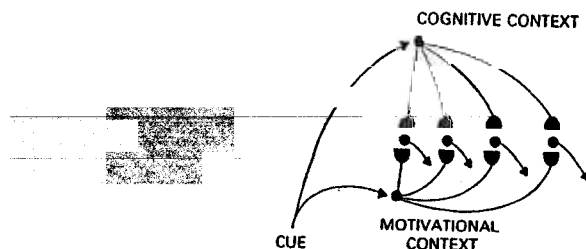


FIG. 15. A cue can excite a cognitive command as well as a nonspecific source of incentive-motivation.

E. TEMPORAL DISCRIMINATION AND FEEDBACK INHIBITION

The above remarks set the stage for analyzing how representations are turned on at appropriate times. How are they turned off?

For example, suppose that the CS that activates $v_1^{(2)}$ in Fig. 12 has a very long duration. If, consequently, each $v_1^{(2)}$ fires for a long time, then each synaptic knob will sample many spatial patterns, and will learn a weighted average of all the patterns (that is, "noise") rather than any particular pattern in the act. How is a prolonged input translated into a brief sampling signal? Grossberg (1970b) proves that a feedforward inhibitory interneuron can create a brief signal that is turned on by rapid changes in input level (Fig. 16b). A feedback inhibitory interneuron can turn off the signal only temporarily at best (Fig. 16c) and can allow a steady leakage of signal if inhibition is too weak (Fig. 16d). None of these mechanisms prevent a second input pulse from reactivating the avalanche while it is performing a later stage of the act. Then cells such as $v_1^{(2)}$ can sample and learn spatial patterns very much out of their correct order. Clearly a feedback inhibition mechanism is needed which prevents premature reactivation, or other perturbations of avalanche performance, unless more urgent environmental demands occur. In cases where the order information is not genetically preprogrammed,

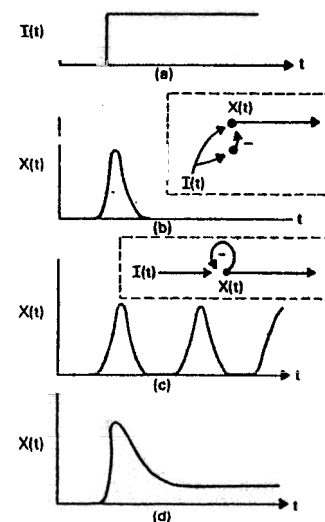


FIG. 16. The prolonged input $I(t)$ in (a) can be shut off by a feedforward inhibitory interneuron (b), but not by a feedback inhibitory interneuron, (c) or (d).

these inhibitory mechanisms must have an anatomy that is independent of any particular ordering that might be learned (cf. Grossberg, 1969c).

With the above heuristic remarks as motivation, we shall now construct a class of networks capable of adaptively synthesizing codes, maps, and plans.

7. Instrumental Conditioning

The mechanisms of drive, reward, and motivation that are needed herein can be derived from postulates about classical conditioning (Grossberg, 1971a, 1972a,b, 1975). This procedure demonstrates that classical and instrumental conditioning share certain mechanisms in common. In effect, these mechanisms embed, or buffer, the cells capable of learning in a network that prevents sampling except under appropriate circumstances. It is remarkable that explicit mechanisms for such nontrivial phenomena as self-stimulation, partial reinforcement acquisition effect, peak shift and behavioral contrast, and novelty as a reinforcer can be derived from such seemingly innocuous postulates as:

- A. The time interval between CS and UCS presentation on successive learning trials can differ; and
- B. The CS alone can elicit a CR on performance trials.

Postulate (A) describes the obvious fact that successive stimulus presentations under natural conditions in real time are not always perfectly synchronized; postulate (B) simply describes the outcome of classical conditioning. To cope with these postulates in a world wherein events continually buffet our senses, and wherein our long-term memories are spatially, albeit nonlocally, coded requires additional network structure.

To see this, suppose that an outstar \mathcal{O}_1 attempts to learn a prescribed spatial pattern $\theta^{(1)}$ in a sequence $\theta^{(1)}, \theta^{(2)}, \theta^{(3)}, \dots$ of spatial patterns by practicing the sequence on successive learning trials. If postulate (A) holds, then the time lag between the CS that excites \mathcal{O}_1 's sampling population $v_1^{(2)}$ and the UCS sequence $\theta^{(1)}, \theta^{(2)}, \theta^{(3)}, \dots$ can be different on successive learning trials. If $v_1^{(2)}$ fires whenever the CS occurs, then \mathcal{O}_1 can sample a different pattern $\theta^{(k)}$ on every learning trial. \mathcal{O}_1 will consequently learn an average pattern that is derived from all the sampled patterns—that is, “noise.” How does \mathcal{O}_1 know when to sample the “important” pattern $\theta^{(1)}$? Somehow, the onset of sampling by $v_1^{(2)}$ and the arrival of the UCS at the field $\mathcal{F}^{(1)} = \{v_1^{(1)}, v_2^{(1)}, v_3^{(1)}, \dots\}$ of

sampled cells must be synchronized. This can happen only if the UCS lets $v_1^{(2)}$ know when it will arrive at $\mathcal{F}^{(1)}$ by sending a signal to $v_1^{(2)}$. Also $v_1^{(2)}$ must be prevented from eliciting a sampling signal unless a large CS and UCS signal converge at $v_1^{(2)}$. This UCS signal must arrive at $v_1^{(2)}$ before the UCS pattern activates $\mathcal{F}^{(1)}$, since $v_1^{(2)}$ must be able to send a signal to $\mathcal{F}^{(1)}$ in time to sample $\theta^{(1)}$. In other words, the UCS activates a bifurcating pathway; one branch *arouses* $v_1^{(2)}$, and the other branch delivers the UCS pattern a little while later. The same argument holds for every cell $v_i^{(2)}$ that is capable of being activated by a CS, since it is not known a priori which CS-UCS combination will be learned. Thus the UCS *nonspecifically* arouses the field $\mathcal{F}^{(2)} = \{v_1^{(2)}, v_2^{(2)}, v_3^{(2)}, \dots\}$ of sampling cells. In summary, simultaneous convergence of the CS input and the UCS nonspecific arousal at a sampling cell are needed to fire this cell. This mechanism synchronizes the onset of CS-activated sampling from $\mathcal{F}^{(2)}$ and the arrival of UCS patterns at $\mathcal{F}^{(1)}$ on successive learning trials. Convergence of a specific input and a nonspecific input is also needed to fire sampling cells $v_i^{(2)}$ in the avalanche of Fig. 12b. It is the same mechanism derived from different considerations.

Postulate (B) shows that conditioning of the CS to the UCS arousal pathway occurs during learning trials. This is the basis for the emergence of “conditioned reinforcers” or “secondary reinforcement” in the networks. Conditioned arousal is necessary, since otherwise the CS alone could not elicit a CR on performance trials. This is because sampling cells can be fired only by the convergence of a CS input and an arousal input. Since the UCS is not present on performance trials to fire the arousal pathway, the CS must gain control over the arousal pathway by being paired with the UCS. An analysis of the minimal mechanism capable of conditioned arousal is shown in Fig. 17, wherein each

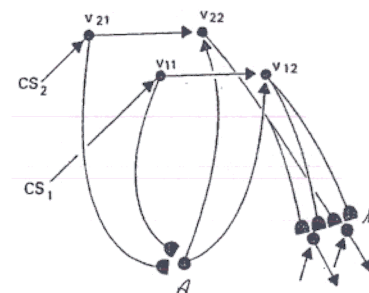


FIG. 17. Conditioning occurs in parallel at the arousal source \mathcal{A} and at the motor command cells \mathcal{M} .

discriminated cue CS_i has a two-stage network representation $\{v_{i1}, v_{i2}\}$, $i = 1, 2, \dots$. Consider the network response to a prescribed CS. Let the CS fire v_{i1} . Then v_{i1} sends signals to v_{i2} and to the arousal population \mathcal{A} . Nothing else happens until the UCS arrives at v_{i2} . This is because v_{i2} can fire only if it receives an input from v_{i1} and from \mathcal{A} , but the signal from v_{i1} to \mathcal{A} is initially too small to fire \mathcal{A} . When the UCS perturbs v_{i2} , v_{i2} sends a signal to v_{i1} and to \mathcal{A} . The $v_{i2} \rightarrow \mathcal{A}$ signals are large enough to fire \mathcal{A} , because the cue firing v_{i2} is a UCS. When \mathcal{A} fires, it releases nonspecific signals to all cells $v_{i2}, v_{i2}, v_{i2}, \dots$. Now three things happen. First, since v_{i1} and \mathcal{A} are both active, the LTM traces in the synaptic knobs of $v_{i1} \rightarrow \mathcal{A}$ axons get stronger. When these traces get strong enough, the CS alone will be able to fire v_{i2} . Second, the arousal signal from \mathcal{A} combines with the UCS-derived signal from v_{i2} at v_{i2} , thereby firing signals from v_{i2} to \mathcal{M} . These signals elicit the UCS pattern in the populations of \mathcal{M} . Third, because the arousal signal from \mathcal{A} is nonspecific, it also combines with the CS-derived signal from v_{i1} at v_{i2} , thereby firing signals from v_{i2} to \mathcal{M} . These signals sample the UCS-elicited pattern at \mathcal{M} . Consequently, the CS begins to acquire UCS properties, both by learning to control the arousal pathway \mathcal{A} , and by learning to elicit the UCS-induced pattern at \mathcal{M} .

The following psychological terms can be used to interpret the above interactions. The arousal pathway supplies "incentive motivation" to the cells v_{i2} . As a population v_{i1} gains control over the arousal pathway, it becomes a "conditioned reinforcer." As v_{i2} samples a pattern at \mathcal{M} , it is said to learn a "habit." Thus, a cue that excites v_{i1} can learn to control incentive motivation via the pathway $v_{i1} \rightarrow \mathcal{A} \rightarrow v_{i2}$ as it simultaneously learns to control a habit via $v_{i1} \rightarrow v_{i2} \rightarrow \mathcal{M}$.

Grossberg (1971b, 1972a,b, 1975) continues this derivation by imposing other simple psychological postulates that act as environmental pressures on an evolving network. These postulates lead to networks in which functional analogs of familiar neural regions appear, such as hippocampus, reticular formation, hypothalamus, septum, and cerebral cortex. Each of these regions emerges because the network tries to deal with environmental feedback in an ever-more-sophisticated way. The present paper cannot review these developments for lack of space.

8. STM Reverberation until Reward Influences LTM

For present purposes, our discussion of instrumental conditioning makes two essential points: (1) A cue can generate an STM response at certain cells v_{i1} without firing the cells v_{i2} ; (2) the cells v_{i2} can be fired

only if the specific signals from v_{i1} are supplemented by nonspecific arousal.

In many situations, an STM trace of a previous event must be kept active after the event itself terminates. For example, in an instrumental conditioning paradigm, the STM traces of previous events must be kept active long enough for later rewards to influence their storage in LTM. Yet these STM traces must also be capable of rapid decay if competing events occur. The two properties of sustained STM activity and rapid induced decay cannot both be achieved by a slow passive decay of STM (Grossberg, 1971b). An active reverberation from v_{i1} to excitatory interneurons v_{i3} and then back to v_{i1} has these two properties (Fig. 18). Excitation in the $v_{i1} \leftrightarrow v_{i3}$ loop can sustain itself indefinitely, even if the passive decay rates of the v_{i1} and v_{i3} populations separately are fast. If one of the links v_{i1} or v_{i3} in the loop is inhibited, then the potentials in v_{i1} and v_{i3} can rapidly decay. The reverberation in $v_{i1} \leftrightarrow v_{i3}$ can go on indefinitely without influencing any LTM changes. Only when arousal reaches v_{i2} can v_{i2} fire and induce sampling by its LTM traces of patterns at \mathcal{M} .

The very virtue of this mechanism introduces a difficulty. Unless the reverberation is inhibited, the loop will continue to reverberate even after arousal allows v_{i2} to fire. In short, reverberation can keep STM traces on, but cannot turn them off. When such a mechanism is used in a network trying to learn order information, chaos results unless there exist sources of inhibition to shut off the STM reverberations at appropriate times. To see how to do this, we embed the functional units of Fig. 18 into the simplest anatomy capable of learning order—namely, the avalanche.

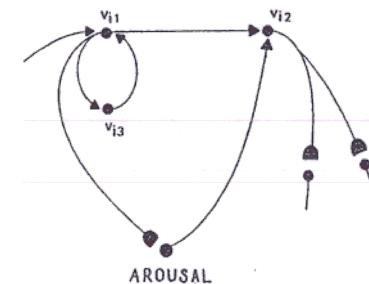


FIG. 18. Reverberation in the $v_{i1} \leftrightarrow v_{i3}$ loop keeps the STM trace active and permits it to be rapidly shut off by inhibition.

9. Rehearsal Resets STM Order Information Using Feedback Inhibition and Decouples Order and Velocity Information

The cells $v_i^{(2)}$ in Fig. 12 are analogous to the cells v_{i2} in Fig. 18. In effect, Sections 7 and 8 refine our understanding of how the cells $v_i^{(2)}$ are influenced by arousal. To construct the simplest example of how to combine the properties of sequential performance, STM reverberation, and modulation by arousal, we replace the populations $v_i^{(2)}$ of Fig. 12 by the units $\{v_{i1}, v_{i2}\}$, along with their interaction pathways, as in Fig. 19. In this figure, the following pathology occurs. When a signal from $v_{i-1,2}$ activates v_{i1} , the STM reverberation $v_{i1} \leftrightarrow v_{i3}$ is switched on, and is never shut off. The signal $v_{i1} \leftrightarrow v_{i3}$ plus arousal keeps the population $v_{i2} \leftrightarrow v_{i3}$ firing at all future times. Consequently, v_{i2} will sample *every* pattern that reaches \mathcal{M} after v_{i1} is switched on, thereby learning nothing. This is intolerable. The $v_{i1} \leftrightarrow v_{i3}$ loop can be allowed to reverberate until v_{i2} performs its LTM pattern, but the reverberation must then be inhibited, or else v_{i2} 's pattern will be washed away by the tide of future events. Thus, when v_{i2} fires, it not only sends excitatory signals to \mathcal{M} . It must also send feedback inhibitory signals to either v_{i1} or v_{i3} that terminate their STM reverberation (Fig. 20). As v_{i2} fires, it also excites $v_{i+1,1}$, which reverberates with $v_{i+1,3}$ and sends signals to $v_{i+1,2}$. Population $v_{i+1,2}$, in turn, fires when it is aroused. This system thus provides a simple example of how order and velocity information are decoupled;

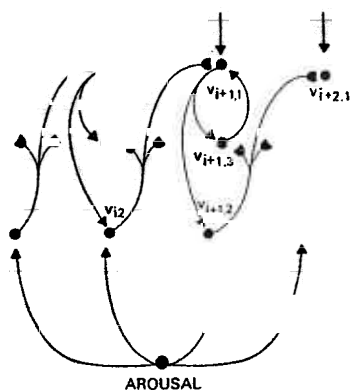


FIG. 19. Minimal synthesis of learned sequential performance, STM reverberation, and modulation by arousal.

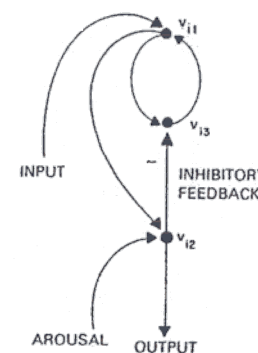


FIG. 20. Feedback inhibition elicited by performance signal prevents permanent sampling.

that is, the same sequence of acts can be performed with different rhythms by varying the size of the rehearsal wave on successive trials.

10. An Emergent Neocortical Analog

The units $\{v_{i1}, v_{i2}, v_{i3}\}$, $i = 1, 2, \dots$, along with their auxiliary interneurons and pathways, form network "modules" or "chips" that carry out important "data processing" tasks. When many of these units are placed alongside each other, a natural laminar structure emerges, with the same type of processing going on in all the cells of a given layer. A column of layers is a functional unit in this laminar field.

There are many reasons to interpret this structure as an emergent neocortical analog. The ubiquitous laminar organization of neocortical tissue (Crosby *et al.*, 1962) and its organization into columnar functional units (Hubel and Wiesel, 1962, 1963) are two of the most casual similarities. Furthermore, the cells v_{i2} are often polyvalent cells, or cells that respond to more than one modality, such as an auditory CS (tone) and a visual UCS (visual presentation of food). These cells fire only in response to the sum of CS plus UCS inputs, and are importantly implicated in plastic network changes. John (1966, 1967) reports the existence of analogous polyvalent cells in neocortical tissue. Grossberg (1971b) summarizes related data.

In vivo, a slab of neocortex that has been isolated by cutting through the underlying white matter can maintain sustained reverberating activity (Burns, 1958). A similar persistence of reverberation occurs in the $v_{i1} \leftrightarrow v_{i3}$ excitatory loop if inhibitory $v_{i2} \rightarrow v_{i3}$ feedback is prohibited by

cutting off arousal inputs to v_{12} . This fact clarifies an interesting paradox. Arousal in Fig. 20 has two functions. First, it excites the cells v_{12} , which thereupon rehearse their patterns. Second, it indirectly inhibits the loop $v_{11} \leftrightarrow v_{13}$, and thereupon resets STM.

Other functional properties have been cited for identifying this structure as a neocortical analog. In the theory of attention in Grossberg (1975), a conditionable excitatory feedback pathway modulates excitability in this structure. This feedback pathway is analogous to the source of *contingent negative variation* (CNV), a slow cortical potential shift that has been associated with an animal's expectancy, decision (Walter, 1964), motivation (Irwin *et al.*, 1966; Cant and Bickford, 1967), volition (McAdam *et al.*, 1966), preparatory set (Low *et al.*, 1966), and arousal (McAdam, 1969). Walter (1964) hypothesized that the CNV shifts the average base line of the cortex by depolarizing the apical dendritic potentials of its pyramidal cells and thereby priming the cortex for action. The arousal pathway of Fig. 19 is this CNV analog, although we have not, for lack of space, reviewed why this pathway should be conditionable. In brief, this conditionable pathway establishes a "psychological set" without which inappropriate acts can be elicited by any motivational source (Grossberg, 1969c, 1971b, 1975). Grossberg (1975) also notes that more than one type of arousal exists; for example, incentive-motivational arousal and the arousal triggered by novel events are conceptually and anatomically different mechanisms that, in fact, often compete with each other. The former system focuses attention on cues that have in the past yielded expected consequences; it blocks, or overshadows, irrelevant cues. The latter system frees irrelevant cues from overshadowing when unexpected consequences occur.

The constructions in later sections will further develop this analog. For the present, we draw Fig. 21 in a way that emphasizes known neocortical structure.

11. Control of Performance Duration by STM and Arousal

In Fig. 20, once a state v_{12} receives a signal from v_{11} plus arousal, it fires and thereupon accomplishes three things:

1. It inhibits $v_{11} \leftrightarrow v_{13}$ reverberation, and thus it fires only for a brief time.
2. It excites $v_{i+1,1}$, which can reverberate with $v_{i+1,3}$ for an indefinite interval of time, until arousal combines with the $v_{i+1,1} \rightarrow v_{i+1,2}$ signal, and thereupon inhibits the reverberation via the $v_{i+1,2} \rightarrow v_{i+1,3}$ inhibitory pathway.

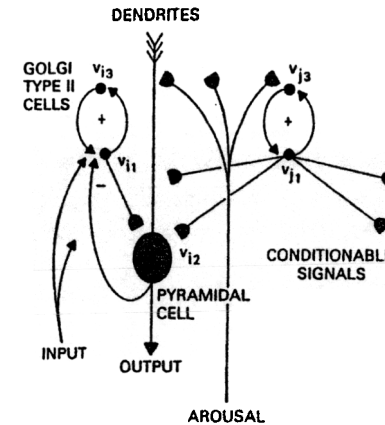


FIG. 21. A cortical analog of the minimal network module suggests interactions between phasic inputs, nonspecific arousal, pyramidal cells, and cortical interneurons—for example, Golgi type II cells.

3. It briefly excites an output pathway that leads, perhaps multisynaptically, to certain muscles.

Suppose that, after v_{12} fires, arousal is terminated. Then v_{12} 's brief signal moves down the output pathway, and the $v_{i+1,1} \leftrightarrow v_{i+1,3}$ reverberation labels the next pattern to be elicited without releasing this pattern. How does the brief v_{12} output signal generate a new output configuration that can be maintained until arousal releases the next output pattern? For example, how can one *hold* a tone, or a note on a keyboard, or a phrase, until the next pattern is released? Somehow the brief output signal imposes a new pattern on the motor controls, and this new pattern also reverberates in STM until it is supplanted, or reset, by the next pattern. In other words, the motor controls maintain a *posture* until a new command changes the posture or terminates the reverberation. Figure 22 schematizes this relationship in terms of descending control by v_{12} of pairs of agonist-antagonist muscle groups. The descending signal inhibits the excitatory reverberation due to previous patterns, and imposes a new pattern on all the pairs. This new pattern thereupon reverberates until a new disturbance occurs, such as a competing command or the removal of arousal.

The above discussions make plain the need to study two kinds of processes in greater detail. First, how is order information embedded in LTM when a sequence of events is presented to a network? Second,

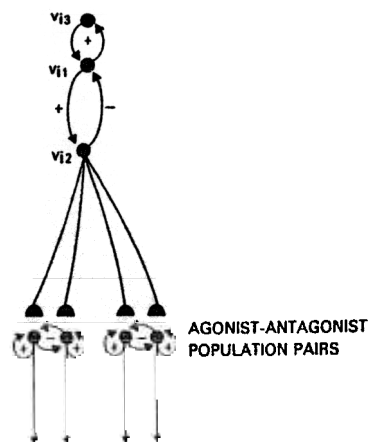


FIG. 22. The brief command from v_{12} is stored in STM until it is reset by a new disturbance.

how is a graded pattern of STM activity maintained across many populations? We shall begin by summarizing some facts about the first problem, and then the second, before combining these facts to reconsider the first. The two problems are intertwined, and so therefore is our discussion of them.

12. Serial Learning and STM \rightarrow LTM Order Reversal

This section summarizes the simplest results on how order information is transferred from STM to LTM. The basic references are Grossberg (1969d) and Grossberg and Pepe (1971). Grossberg (1974, Section VII) reviews some of these results. They show how temporal sequences of events are coded as spatial patterns of LTM activity.

Two types of serial learning are included in this analysis. In both types, order information is not innate. In type I (Fig. 14a) many $v_{12} \rightarrow v_{11}$ pathways of comparable strength exist before learning occurs. Somehow a sufficient amount of serial learning embeds a directed chain-like structure $v_{12} \rightarrow v_{21}$, $v_{22} \rightarrow v_{31}$, $v_{32} \rightarrow v_{41}$, \dots into this anatomy. Actually, we shall see that spatially distributed LTM patterns exist even in this case. Type II (Fig. 14b) exhibits another anatomical substrate of serial learning. Here a command state learns order information by sampling populations as they are sequentially activated. This latter anatomy has the advantage that serial order in $\mathcal{F}^{(2)}$ can be reorganized by

changing which commands are active at any time. The chains in Fig. 14a, by contrast, rigidly constrain the possible performance order once they are entered (cf. Lenneberg, 1967, Chapter 3). Many of the same LTM patterns are learned in both types of anatomy. This fact is important, because, when both types of phenomena are operative, they yield self-consistent order information. The two cases are schematized more completely in Fig. 23. Figures 23a and 23b depict variants of the type of serial learning in Fig. 14a. In Fig. 23a, every state v_i is connected to all *other* states v_j by conditionable pathways. In Fig. 23b, every state v_i is connected to all states v_j by conditionable pathways. In both cases, we let the states v_1, v_2, \dots, v_L be sequentially excited by a list of

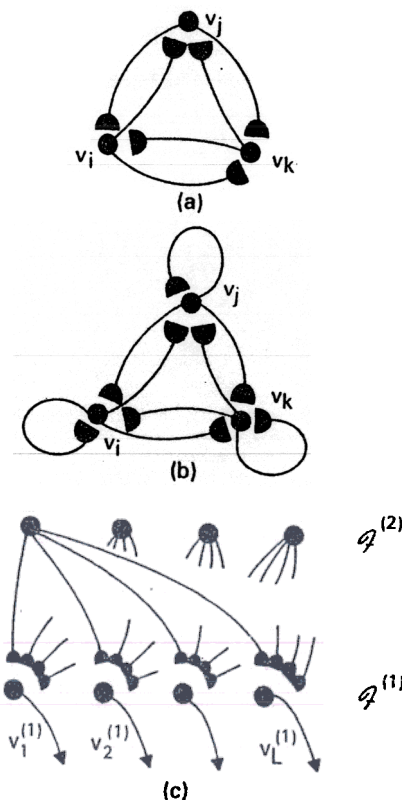


FIG. 23. Two anatomies in which serial learning builds up chains of LTM associations, (a) and (b). In (c), command states $v_j^{(2)}$ learn serial order of inputs presented to populations $v_k^{(1)}$.

inputs, with a time lag of w (intratrial interval) between successive list items. See Grossberg (1969d, 1974) and Grossberg and Pepe (1971) for a discussion of how w , L , and the intertrial interval between list presentations influence LTM. In Fig. 23c, the states $v_1^{(1)}, v_2^{(1)}, \dots, v_L^{(1)}$ in $\mathcal{F}^{(1)}$ are sequentially excited by a list of inputs. The states in $\mathcal{F}^{(2)}$ are command states, or chunks, that are either directly excited by the list inputs, or indirectly excited by sets of active states in $\mathcal{F}^{(1)}$. For the remarks below to hold, it only matters that disjoint subsequences in $\mathcal{F}^{(2)}$ are sequentially excited at a uniform rate and with a uniform intensity through time. Generalizations of this situation will be clear once the basic mechanism is understood.

For definiteness, we shall consider Fig. 23c when an input to state $v_i^{(1)}$ also excites $v_i^{(2)}$. We shall discuss LTM under the following assumptions:

1. The system starts out at equilibrium.
2. The states are serially excited every w time units.
3. The STM traces decay at an exponential rate after they are excited.
4. The LTM traces z_{jk} add up the products of signals B_{jk} from $v_j^{(2)}$ to $v_k^{(1)}$ and STM traces $x_k^{(1)}$. For simplicity, we let $B_{jk} = [x_j^{(2)} - \Gamma]^+$, but the results also hold if B_{jk} is a sigmoid function of $x_j^{(2)}$. We ignore the rate of LTM decay, which is assumed to be slow on each learning trial.
5. The stimulus sampling probabilities

$$Z_{jk} = z_{jk} \left(\sum_m z_{jm} \right)^{-1}$$

measure how strong the (j, k) th association is relative to competing (j, m) th associations, $m \neq j$.

In all,

$$\dot{x}_i^{(1)} = -A_1 x_i^{(1)} + I_i(t), \quad x_i^{(1)}(0) = 0 \quad (7)$$

$$\dot{x}_i^{(2)} = -A_2 x_i^{(2)} + I_i(t), \quad x_i^{(2)}(0) = 0 \quad (8)$$

$$\dot{z}_{jk} = B[x_j^{(2)} - \Gamma]^+ x_k^{(1)}, \quad z_{jk}(0) = \alpha > 0 \quad (9)$$

and

$$Z_{jk} = z_{jk} \left(\sum_m z_{jm} \right)^{-1} \quad (10)$$

Figure 24a shows how STM activity is distributed across the field $\mathcal{F}^{(1)}$ through time. The last item to have received an input is always most intensely active, and successively earlier items have progressively weaker STM traces (STM recency effect). A similar distribution of activity holds for the STM traces of the command states in $\mathcal{F}^{(2)}$; the last

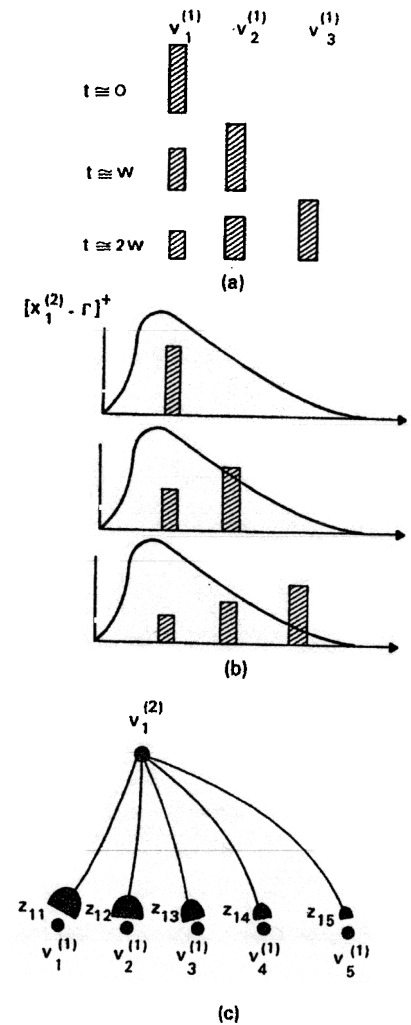


FIG. 24. Whereas STM shows a recency effect, LTM (at least as sampled by $v_1^{(2)}$) shows a primacy effect.

item to have received an input is always the most active, etc. What about the sizes of the LTM traces z_{jk} ? By (9), z_{jk} grows faster when either $x_j^{(2)}$ or $x_k^{(1)}$ is increased in size, other things being equal, but does not grow if either $x_j^{(2)} \leq \Gamma$ or $x_k^{(1)} \approx 0$. Note, however, that the stimulus sampling probability Z_{jk} in (10) can decrease even when z_{jk} grows, if the competing terms $\sum_{m \neq j} z_{jm}$ grow more quickly.

Consider the distribution of the probabilities Z_{it} emanating from $v_i^{(2)}$ after a single list presentation. Three facts control this distribution:

1. Each STM trace $x_i^{(1)}$ is the same as the previously excited trace $x_{i-1}^{(1)}$, except that its growth and decay are delayed in time by w because of (7).
2. Trace $x_i^{(2)}$ starts to decay after the input to $v_i^{(2)}$ terminates, by (8).
3. Trace z_{it} adds up the products $[x_i^{(2)} - \Gamma]^+ x_i^{(1)}$ through time, by (9).

In Fig. 24b, the STM traces $x_1^{(1)}, x_2^{(1)}, x_3^{(1)}, \dots$ are superimposed on the signal $[x_i^{(2)} - \Gamma]^+$. The products $[x_i^{(2)} - \Gamma]^+ x_i^{(1)}$ are clearly made smaller, other being things equal, as i increases. Hence, after the learning trial is over $z_{11} > z_{12} > z_{13} > \dots$. This fact is illustrated in Fig. 24c by drawing largest the synaptic knobs of the largest LTM traces. In particular, if a rehearsal input perturbs $v_i^{(2)}$ after the STM traces have decayed, then the STM traces $x_i^{(1)}$ in $\mathcal{F}^{(1)}$ that are activated by the LTM traces z_{it} have the distribution shown in Fig. 24c. Whereas the largest STM traces in Fig. 24a correspond to the *most* recently inputted states (recency effect), the largest traces in Fig. 24c correspond to the *least* recently inputted states (primacy effect). The LTM traces of the command state $v_i^{(2)}$ have *reversed* the order of STM trace strength, as was discussed in Section 1, G! Compare the data on STM recency and LTM primacy as reviewed in Atkinson and Shiffrin (1971).

Recent studies of serial learning emphasize the interplay of "structural" versus "transient" factors in memory (Bjork, 1975; Craik and Jacoby, 1975; Estes, 1972; Hogan, 1975; Hogan and Hogan, 1975; Shiffrin, 1975). In the present context, structural factors are those that use the LTM traces to generate an STM distribution on which performance is based. Transient factors are those STM properties that are *directly* induced by the experimental manipulations, without intervention of feedback signals that are gated by LTM traces. The opposite tendencies in LTM and STM orders shown above will be used below to explain various bowing phenomena in performance.

Before doing this, two related kinds of phenomena will be summarized: first, the relative LTM learning rates at different list positions; and second, the shape of the generalization gradients, or STM spatial patterns, controlled by the LTM traces at different list positions.

To discuss the learning rates at different list positions, we define the function

$$G(i, \Gamma, t) = Z_{it}(tw) \quad (11)$$

and consider $G(i, \Gamma, i)$. This latter function measures the "correct

association" from $v_i^{(2)}$ to $v_i^{(1)}$ one time unit after sampling signals from $v_i^{(2)}$ reach $\mathcal{F}^{(1)}$, and thus before "incorrect future associations" such as $z_{i,i+k}$, $k \geq 1$, can develop. Figure 24a shows that, as i increases, there are more active "past" STM traces $x_{i-1}^{(1)}, x_{i-2}^{(1)}, \dots$ to compete with the growth of Z_{it} after the input to $v_i^{(2)}$ occurs. The function $G(i, \Gamma, i)$ correspondingly has the graph in Fig. 25a, at any threshold value Γ for which some learning occurs. By contrast, let t in $G(i, \Gamma, t)$ be allowed to increase to values that correspond to times long after the last list item is presented. Then $G(i, \Gamma, t)$ is no longer monotone decreasing. The nonoccurrence of any more list items after the L th item is presented facilitates LTM growth at positions near the end of the list. If $\Gamma = 0$ and t is allowed to become arbitrarily large, this facilitation propagates backward through list items until the middle of the list is reached. That is, the minimum of $G(i, 0, \infty)$ occurs at $i = L/2$ or $(L-1)/2$, whichever is an integer (Fig. 25b). The middle of the list is consequently hardest to learn in this case (bowing); the proof is in Grossberg (1969d). If t increases by a finite amount beyond $(L-1)w$, then a curve between Fig. 25a and Fig. 25b is obtained by continuous interpolation (Fig. 25c). That is, *skewing* of the bowed curve occurs, if only because the intertrial interval is of finite duration. Skewing can also be caused by choosing Γ

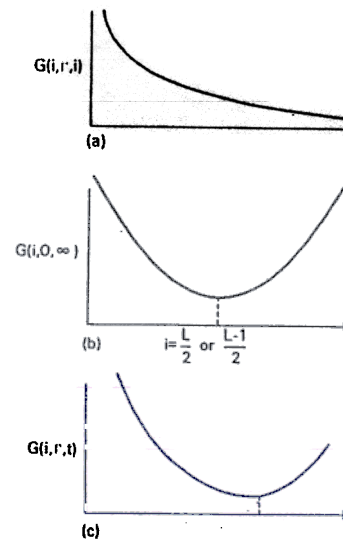


FIG. 25. (a) If no future field interference occurs, then the end of the list is hardest to learn; (b) if maximal future field interference occurs, then the middle of the list is hardest to learn (bowing); (c) if the signal threshold is finite, then skewing occurs.

> 0 . This happens because, as Γ increases, or equivalently, nonspecific arousal decreases, population $v_i^{(2)}$ in $\mathcal{F}^{(2)}$ can sample fewer "future" populations $v_{i+2}^{(1)}, v_{i+3}^{(1)}, \dots$ in $\mathcal{F}^{(1)}$, but can still sample all active populations $v_1^{(1)}, v_2^{(1)}, \dots, v_{i-1}^{(1)}$ in the "past field" of $v_i^{(2)}$. Grossberg and Pepe (1971) prove, in addition, that, whereas the beginning of the list is easier to learn, given "normal" arousal or threshold levels (primacy dominates recency), eventually the end of the list is easier to learn at overaroused levels (recency dominates primacy). They compare these results to attentional problems that learning subjects have when they are overaroused. The result that the relative LTM trace strengths at the end versus the beginning of the list reverse as arousal level increases does not include many important STM and LTM interactions. It nonetheless emphasizes the importance of performance variables on what is encoded, and suggests a convenient learning measure by which the degree of overarousal might be quantified (cf. Section 27).

The LTM generalization gradients at various list positions have the familiar position-dependent form shown in Fig. 26. There is a forward distribution of associations $Z_{11}, Z_{12}, \dots, Z_{1L}$, a backward distribution of associations $Z_{L1}, Z_{L2}, \dots, Z_{LL}$, and a two-sided distribution of associations $Z_{L/2,1}, Z_{L/2,2}, \dots, Z_{L/2,L}$ (L even).

These results illustrate some important facts that will reappear in a suitable form when more complex structures are considered. First, no matter what anatomy in Fig. 23 is perturbed by serial inputs, similar distributions of STM and LTM patterns evolve through time, and suggest mechanisms of bowing, skewing, primacy/recency balance, response generalization, and STM \rightarrow LTM order reversal. Second, there is a delicate interplay between STM and LTM factors at $\mathcal{F}^{(1)}$. For example, let recall begin immediately after the list is presented. Then the STM pattern that is produced at $\mathcal{F}^{(1)}$ by serial inputs exhibits a recency effect (Fig. 24a). By contrast, LTM produces a strong primacy effect at $\mathcal{F}^{(1)}$. This is due to several factors. At times t right after the list is presented, $G(i, \Gamma, t)$ is monotone decreasing as a function of i . In other words, the LTM traces $Z_{11}, Z_{12}, \dots, Z_{1L}$ with $i \equiv 1$ code a more differentiated pattern than do the LTM traces with $i \equiv L$. Thus, a signal to $\mathcal{F}^{(1)}$ from a $v_i^{(2)}$ with $i \equiv 1$ will have a more pronounced effect on the relative sizes of the STM traces $x_1^{(1)}, x_2^{(1)}, \dots, x_L^{(1)}$ than will a signal to $\mathcal{F}^{(1)}$ from a $v_i^{(2)}$ with $i \equiv L$. But the LTM traces $Z_{11}, Z_{12}, \dots, Z_{1L}$ with $i \equiv 1$ exhibit a strong primacy effect, by Fig. 26. Thus, signals from $\mathcal{F}^{(2)}$ to $\mathcal{F}^{(1)}$, gated by their LTM traces, will produce a strong primacy effect in the STM pattern at $\mathcal{F}^{(1)}$.

On the other hand, if recall begins after the STM pattern at $\mathcal{F}^{(1)}$ has decayed, then a recency effect can still be obtained, but it is due to LTM

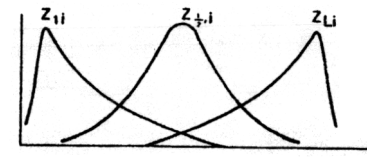


FIG. 26. Generalization gradients as a function of list position.

rather than STM. More precisely, if the STM traces at $\mathcal{F}^{(1)}$ are allowed to decay without experimental interference, then the LTM traces $Z_{ii}(t)$ will be bowed as a function of i . In other words, LTM traces near the end of the list can be large. Moreover, the distributions $Z_{11}, Z_{12}, \dots, Z_{1L}$ with $i \equiv L$ exhibit a recency effect by Fig. 26. Signals to $\mathcal{F}^{(1)}$ from states $v_i^{(2)}$, $i \equiv L$, will be gated by these LTM traces and will, therefore, produce a recency effect at $\mathcal{F}^{(1)}$. This recency effect will be balanced by the primacy effect due to signals to $\mathcal{F}^{(1)}$ from states $v_i^{(2)}$, $i \equiv 1$, whose LTM traces $Z_{11}, Z_{12}, \dots, Z_{1L}$ code a primacy effect. Different experimental manipulations can differentially activate the states $v_i^{(2)}$ during recall trials. Such manipulations act as *probe* stimuli, in the sense that they determine which LTM patterns in the $\mathcal{F}^{(2)} \rightarrow \mathcal{F}^{(1)}$ pathways will be activated and thereby influence STM at $\mathcal{F}^{(1)}$. The net STM pattern elicited across $\mathcal{F}^{(1)}$ in a given time frame by all stimuli will determine the performance controlled by $\mathcal{F}^{(1)}$.

These remarks address some aspects of the complex interplay between activity at $\mathcal{F}^{(1)}$ that is due to serial inputs, and activity at $\mathcal{F}^{(1)}$ that is due to feedback signals from $\mathcal{F}^{(2)}$ that are gated by LTM traces. The remarks also raise some questions. Do the STM traces spontaneously decay? If not, what factors cause their ultimate termination? How are STM traces of command states in $\mathcal{F}^{(2)}$ activated in the first place? We now turn to the fundamental question of how STM activities are regulated.

13. Storing Spatial Patterns in STM

Several papers have appeared that analyze how spatial patterns of activity are stored in STM (Grossberg, 1973; Ellias and Grossberg, 1975; Grossberg and Levine, 1975; Levine and Grossberg, 1976). The general conclusion is that, in a field of populations capable of this task, each population excites nearby populations (recurrent on-center) and inhibits a broad expanse of populations (recurrent off-surround) by mechanisms that obey mass action laws. All these operations describe *parallel*

computations that go on in a distributed fashion across the field at any time. Such mechanisms can prevent the stored pattern from being distorted by noise or saturation.

To see how these networks arise, suppose that a field \mathcal{F} of populations v_i , $i = 1, 2, \dots, n$, is given. Let each v_i have a certain number B of sites that can be in either an excited or an unexcited state. Let the STM activity $x_i(t)$ be the number of excited sites in v_i at time t . 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. How can such inputs $I_i(t)$ change through time? As Section 2 notes, 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 in shades of white, gray, and black. Then $I(t)$ describes the intensity of background illumination of the picture. This intensity can vary wildly through time. The picture itself is characterized by the spatial pattern $\theta = (\theta_1, \theta_2, \dots, \theta_n)$ of numbers (the *reflectances*), 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 $\theta(t) = (\theta_1(t), \theta_2(t), \dots, \theta_n(t))$ 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 .

Another reason for distinguishing $\theta(t)$ from $I(t)$ is that the unit of LTM is a spatial pattern. When an outstar performs a spatial pattern θ on a field \mathcal{F} , it can do so in response to a CS input that fluctuates wildly through time. Somehow certain network cells must be able to "read" θ independent of the fluctuations in $I(t)$, or else the pattern could never be decoded. Such considerations originally motivated the construction of network filters that can discriminate relative figure-to-ground (Grossberg, 1970b, 1972d, 1976b). Some of the minimal filters have anatomies that are strikingly "retinal," and they are capable of formal analogs of such perceptual constancies as hue, brightness, and lightness constancy (Grossberg, 1972d).

14. Gain Control and Adaptation in On-Center Off-Surround Networks

How can we design a system capable of distinguishing the pattern weights $\theta(t)$ from fluctuations in the background activity $I(t)$? First we consider a trial system to show what the difficulties are. If B is the total number of excitable sites in any population v_i , and $x_i(t)$ is the number of excited sites, then $[B - x_i(t)]$ is the number of unexcited sites at time t .

Suppose that there is an equilibrium point (say 0), such that, as excited sites spontaneously become unexcited at rate A , $x_i(t)$ approaches 0. Also suppose that unexcited sites $[B - x_i(t)]$ are excited at a rate jointly proportional to their number and the input intensity $I_i(t)$. This is a mass action law. Then

$$\dot{x}_i = -Ax_i + (B - x_i)I_i(t) \quad (12)$$

with $0 \leq x_i \leq B$. In other words, system (12) 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 with fixed pattern weights θ_i and total activity I . A steady state, $\dot{x}_i = 0$, and (12) implies that

$$x_i = \frac{B\theta_i I}{A + \theta_i I} \quad (13)$$

Now keep the θ 's fixed and vary I . In other words, study how (12) processes the *same* pattern θ given different background activity levels. By (13), as I increases, all x_i approach B , and all information about θ is lost because of 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 at both low and high I values. How can this be corrected?

System (12) fails because there are no interactions among the v_i . Each θ_i is defined by an interaction of *all* the inputs I_k , $k = 1, 2, \dots, n$. Since neither the populations nor their inputs interact, they cannot possibly compute θ . What interactions are needed? Writing $\theta_i = I_i(I_i + \sum_{k \neq i} I_k)$, it is clear that increasing I_i increases θ_i and that increasing any I_k , $k \neq i$, decreases θ_i . In other words, I_i "excites" θ_i , whereas all I_k , $k \neq i$, "inhibit" θ_i ; the inputs compete in order to prevent saturation and thereby to compute relative figure-to-ground. When this intuition is translated into mass action dynamics, we find the simplest example of a type of system that occurs throughout the nervous system—namely, a feedforward on-center off-surround network undergoing shunting, or passive membrane dynamics. Thus let each I_i excite population v_i and inhibit all populations v_k , $k \neq i$. The inputs then form a nonrecurrent (or feedforward) on-center off-surround interaction pattern (Fig. 27), and (12) is replaced by

$$\dot{x}_i = -Ax_i + (B - x_i)I_i - x_i \sum_{k \neq i} I_k \quad (14)$$

The new term $-x_i \sum_{k \neq i} I_k$ says that excited sites at v_i (which number x_i)

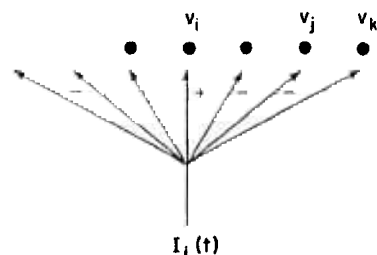


FIG. 27. Feedforward (nonrecurrent) on-center off-surround anatomy

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 (14) is again a mass action law, or passive membrane equation, in which 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. (14) implies

$$x_i = \theta_i \frac{BI}{A + I} \quad (15)$$

In other words, no matter how large I becomes, each x_i is proportional to θ_i ; there is no saturation. The system has an infinite dynamical range no matter how small B is! Furthermore, the total activity $x = \sum_{k=1}^n x_k$ satisfies $x = BI(A + I)^{-1} \leq B$; the maximal 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 network response to fluctuations in total input. I have elsewhere suggested that this adaptation is akin to retinal light adaptation, say as studied by Werblin (1971) in the mudpuppy retina (Grossberg, 1972d), with the responses in (14) analogous to potential changes in retinal bipolar cells. Indeed, bipolar cell potential [cf. x_i in (15)] is sensitive to the ratio of on-center to off-surround excitation [cf. θ_i in (15)] and obeys a type of Weber-Fechner psychophysical law [cf. $BI(A + I)^{-1}$ in (15)]. Furthermore, bipolar potential is shifted to the right as a function of the logarithm of on-center input intensity when the off-surround input is parametrically increased. Similarly, rewrite (15) as $x_i(K, L) = Be^K(A + e^K + L)^{-1}$, where $K = \log I_i$ and $L = \sum_{k \neq i} I_k$. Check that, if L is changed from L_1 to L_2 , then $x_i(K + S, L_1) = x_i(K, L_2)$ for all $K \geq 0$, where the shift $S = \ln[(A + L_1)(A + L_2)^{-1}]$.

The above example illustrates a point that repeatedly reappears in this

paper. The teleology, or principles, of design that lead to network mechanisms are all based on a world in which continuous fluctuations exist. Binary codes are singular limits of the continuous models, and lose most of the properties, and the rationale, of the continuous models.

15. Contrast Enhancement in STM

System (14) cannot remember the pattern θ for long after the inputs are shut off, because each x_i then decays to 0. As was suggested in Section 8, recurrent (or feedback) signals among the populations v_i are needed to ensure STM. How should these feedback signals be distributed? In Section 8, we worried only about how to keep STM active until it is shut off. Now we have to worry also about what spatial pattern across \mathcal{F} will be stored in STM. In particular, we have to prevent saturation of this pattern, so the signals should be distributed in an on-center off-surround anatomy, as in Fig. 28. The excitatory on-center signals $v_{i1} \leftrightarrow v_{i3}$ in Figs. 18 through 21 must be supplemented by inhibitory off-surround signals to prevent pattern saturation. 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 , $k = 1, 2, \dots, n$. Then (14) is replaced by the nonlinear system

$$\dot{x}_i = -Ax_i + (B - x_i)[f(x_i) + I_i] - x_i \left[\sum_{k \neq i} f(x_k) + J_i \right] \quad (16)$$

$i = 1, 2, \dots, n$. Term $(B - x_i)f(x_i)$ describes how a feedback signal $f(x_i)$ from v_i to itself excites the unexcited sites $B - x_i$ by mass action. The inhibitory term $-x_i \sum_{k \neq i} f(x_k)$ 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 total excitatory input and term J_i is the total inhibitory input at v_i . Often J_i is an off-surround input such as $J_i = \sum_{k \neq i} I_k$.

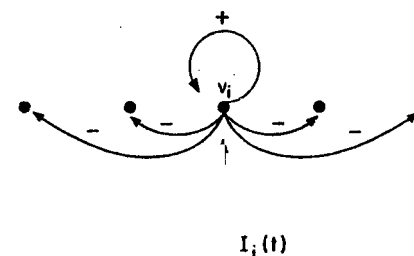


FIG. 28. Feedback (recurrent) on-center off-surround anatomy

Much more complex versions than (16) of recurrent on-center off-surround networks have been studied (cf. Ellias and Grossberg, 1975; Levine and Grossberg, 1976; Grossberg, 1977a,b,c). But system (16), by being simple, focuses on a very important problem whose solution was a prerequisite to further progress. How does system (16) 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? How should the average signal $f(w)$ be chosen as a function of the average activity w to distinguish between important and unimportant data? Grossberg (1973) solves this problem. The solution is summarized in Table I in terms of the total STM trace $x = \sum_{k=1}^n x_k$ and the relative STM traces $X_i = x_i x^{-1}$. After a brief pattern of inputs (I_i, J_i) is delivered to the network, does $x(t)$ converge to 0 (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 ? Do they enhance certain population activities and suppress others? Do they become more similar through time? All these cases can occur if $f(w)$ is suitably chosen.

A sigmoid, or S-shaped, $f(w)$ has all the desirable physical properties that we need. It is proved that $f(w)$ must grow faster-than-linearly at

TABLE I
INFLUENCE OF SIGNAL FUNCTION $f(w)$ ON PATTERN TRANSFORMATION AND STM STORAGE

$f(w)$	$g(w)$	$X_i(\infty) = \frac{X_i(t \rightarrow \infty)}{\sum_k X_k(\infty)}$	$X(\infty) = \sum_k X_k(\infty)$
			AMPLIFIES NOISE
			AMPLIFIES NOISE
			QUENCHES NOISE
			QUENCHES NOISE

small w values in order not to amplify noise; that is, the function $g(w) = w^{-1}f(w)$ must be strictly increasing at small w values, as when $f(w) = w^2$. To prevent the system from choosing the population with maximal initial activity for STM storage and suppressing all other populations, $f(w)$ must be approximately linear—that is, $f(w) \approx Cw$ —at intermediate w values. At large w values, all signal functions are necessarily bounded. Piecing the three regions together yields a sigmoid $f(w)$. Grossberg and Levine (1975) review in detail the reasons for using a sigmoid signal function. Other signal functions, such as faster-than-linear signal functions, can produce a network that behaves like a finite-state machine.

Given a sigmoid signal function, a *quenching threshold* (QT) exists. As Figs. 29a and 29b depict, if an initial activity $x_i(0)$ is smaller than the QT , then the activity of v_i will be quenched, or masked, by the STM 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. 29b. Thus, a sigmoid $f(w)$ is capable of *partially* contrast-enhancing a pattern. Its QT determines the cutoff between significant and insignificant data. If the QT is pathologically small, then the network can bootstrap into STM disturbances that do not represent behaviorally meaningful inputs; that is, a "seizure" occurs. Grossberg (1973) computes the QT in a special case and shows what parameters can lower it.

16. Tuning of STM and Releasing Subliminal Maps by Arousal

Given the existence of a QT , varying the arousal level can dramatically change what will be stored in STM, if arousal modulates the

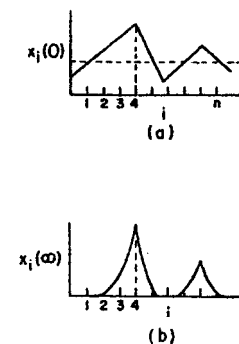


FIG. 29. Influence of the quenching threshold (QT) on STM storage of patterns.

excitability of cell responses to inputs. Arousal either can directly amplify the response of excitatory cells to inputs, or can shunt the excitability of recurrent off-surround interneurons to yield an equivalent effect. In the former case, arousal is excitatory; in the later case, it is inhibitory, and turning it on disinhibits excitatory cells.

One important example shows how a choice can be made in one parallel processing step among complex data. For example, suppose that inputs from several sensory sources stream into the network and excite a complicated pattern of activity across billions of populations, where each population can be thought of as a feature detector to fix ideas. How can the dominant feature be chosen in one step of parallel processing? If the arousal level is lowered until only one population's activity exceeds the QT , then the network chooses that population for STM storage and suppresses all other activities.

By contrast, if arousal suddenly increases in response to an unexpected event, then all recently presented cues can be amplified and stored in STM until the network can use all the available data to cope with the unexpected event. Grossberg (1975) applies this property to the analysis of various attentional processes, such as overshadowing. This analysis suggests that two feedback systems continually retune STM during behavior, as noted in Section 10. One system—an incentive motivational and CNV system—focuses attention on cues that are expected to generate prescribed consequences of behavior. This system can overshadow irrelevant cues. The competing system is triggered by unexpected events (novelty) and allows the network to redefine the set of relevant cues to avoid unexpected consequences. Overarousal of either system can yield attentional deficits, but the exact nature of the deficit and its proper treatment depends on the particular system that is overaroused. For example, a schizophrenic-like syndrome of punning, fuzzy response categories (Section 12), and blocking can be elicited by overarousal of the incentive-motivational system, but would not necessarily be cured by a depressant that acted primarily on the novelty (reticular formation) system.

The above remarks indicate that varying the arousal level can determine which sensory chunks will reverberate in STM and thereupon influence behavior. Similar properties hold in the motor system. For example, let an active population send a sustained pattern of signals to a field of motor control cells. These signals represent a subliminal motor map. If arousal nonspecifically arouses the field, then the pattern is bootstrapped into STM and elicits signals that determine motor output. Section 48 will discuss how such motor maps are learned. Here we merely note the following example to fix ideas. If I look at a given object

that is within my reach, I can decide to touch it with either hand, either foot, my nose, etc. Somehow the proprioceptive coordinates of my eyes, head, and neck can be mapped into a terminal position of either hand, either foot, my nose, etc. "Willing" to move my right hand arouses my right hand-arm system, which thereupon moves to its terminal position.

A more sophisticated version of the tuning process is depicted in Fig. 30, which shows that there can be several (in fact, any number) of equilibrium points of total STM activity $x(t)$. As Fig. 30 shows, every other solution E_2, E_4, \dots of the equation $g(w) = A(B - w)^{-1}$ represents a stable equilibrium point of total STM activity. As the total initial input size increases, so can the asymptotic activity $x(\infty)$.

In each population v_i , successive pairs of roots of $g(x_i) = A(B - x_i)^{-1}$ are produced by bumps in the graph of the signal function $f(x_i)$. Each bump corresponds to a subpopulation of v_i ; for example, the signal thresholds of cells in a population can be distributed around several preferred mean values. Then the population can fire preferentially not only to particular features, but also to particular energy levels to which these features are excited. Grossberg (1977b,c) proves that a definite STM pattern is achieved no matter how many random factors exist within each population, no matter how many populations interact, and no matter how the average level of interpopulation competition is chosen. However, the decision, or enhancement, steps whereby the asymptotic STM pattern is reached can be incredibly complicated, and in general cannot be computed. This lack of computability does not

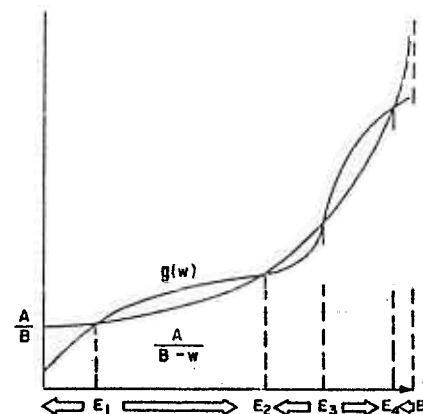


FIG. 30. The even solutions E_2, E_4, \dots of $f(w) = Aw(B - w)^{-1}$ are stable equilibrium points of $x(\infty)$. If $x(0) < E_1$, then $x(\infty) = 0$; thus E_1 defines the level below which $x(t)$ is treated as "noise."

prevent the asymptotic STM pattern from reliably being read into LTM. Here is thus an example wherein individual differences can yield subtle behavioral changes, yet the individual differences are expressions of a robust and stable principle of behavior.

The network (16) is a simple version of more complex networks that capture further properties of interacting neurons (cf. Ellias and Grossberg, 1975; Levine and Grossberg, 1976). In these networks, the properties of contrast enhancement and normalization go through in a modified form, say because the strength of recurrent excitatory and inhibitory interactions both decrease as a function of distance. When this is true, feature detectors are enhanced or suppressed only by prescribed subsets of other feature detectors. For present purposes, we note that the STM traces in the field $\mathcal{F}^{(1)}$ of Fig. 24a are not normalized. We must check how normalization of total STM activity, and tuning of it by shifting the arousal level, change our discussion of serial learning.

17. Adaptive Coding and the Emergence of Command Chunks

Before reanalyzing how normalization and tuning influence serial LTM, we shall discuss a mechanism whereby chunks are adaptively synthesized, or coded, by their defining sequences. This must be done first, because we can then ask what kinds of sequential operations on STM leave the code invariant. In particular, we can discuss whether, and how, normalization and tuning leave the code invariant.

The adaptive coding model is an outgrowth of a model that describes how experience can retune cells in the kitten visual cortex (Grossberg, 1976a,b,c). This model uses *adult* STM and LTM mechanisms in a minimal way. Formally isomorphic mechanisms can also be used to discuss the development of nonneural structures in many species (Grossberg, 1978a): That similar formal mechanisms seem to arise in many species and stages of individual development is ascribed to the fact that all living creatures must solve certain environmental problems in order to survive. These commonly shared problems suggest universal developmental principles which impose statistical, geometrical, and dynamical constraints on all systems capable of solving the problems. From this perspective, the self-organization of commands is a special case of the emergence through evolutionary processes of ever-more-complex, or "higher," degrees of order.

The developmental model shows how any number of arbitrarily complex spatial patterns can be adaptively coded by emergent command populations, and indeed recorded into any other spatial patterns using

the command populations as sampling sources. In particular, the network does not have to preprogram *all* the codes that it will ever need; they can be adaptively synthesized by the environmental demands that confront each individual. Each command population responds to a prescribed convex set of spatial patterns. This convex set is determined, in part, by the affinities of all the other command populations in a field of populations; the populations in a field compete to determine which patterns will succeed in exciting them. The convex set to which a given command population can respond defines the *features* associated with the population. The command population is thus a *feature detector* in a well-defined sense. Since the population responds to a convex set of features, it automatically responds to *average* features chosen from the set, even if the average features have never been experienced. Experience can retune a command population to respond to different features. This is accomplished by changing the competitive balance of STM and LTM activity across the field of populations in response to input patterns. In effect, the model illustrates how experience can generate and globally organize a field of command states, or chunks.

The feedforward version of the model describes the interaction via trainable synaptic pathways (LTM traces) from a field of cell populations $\mathcal{F}^{(1)}$ (for example, lateral geniculate nucleus) to a field of cell populations $\mathcal{F}^{(2)}$ (for example, visual cortex). Fields $\mathcal{F}^{(1)}$ and $\mathcal{F}^{(2)}$ are separately capable of normalizing and contrast enhancing their activity, but $\mathcal{F}^{(2)}$ can also store the contrast enhanced pattern in STM (Fig. 31). In the simplest case, $\mathcal{F}^{(1)}$ consists of a nonrecurrent on-center off-surround anatomy undergoing mass action interactions, as in (14). Denote the STM trace of $v_i^{(1)}$ by $x_i^{(1)}$. Then

$$\dot{x}_i^{(1)} = -Ax_i^{(1)} + [B - x_i^{(1)}]I_i - x_i^{(1)} \sum_{k \neq i} I_k \quad (17)$$

with $0 \leq x_i^{(1)} \leq B$. As in (15), at equilibrium

$$x_i^{(1)} = \theta_i \frac{BI}{A + I}$$

Normalization in $\mathcal{F}^{(1)}$ by (17) occurs gradually in time, as each $x_i^{(1)}$ adjusts to its new equilibrium value, but it will be assumed below to occur instantaneously with $x_i^{(1)}$ approaching θ_i rather than $\theta_i BI(A + I)^{-1}$ to avoid unimportant details. This assumption of instantaneous normalization is tenable because the normalized pattern at $\mathcal{F}^{(1)}$ drives slow changes in the LTM traces that gate the signals from $\mathcal{F}^{(1)}$ to $\mathcal{F}^{(2)}$. Instantaneous normalization means that the pattern at $\mathcal{F}^{(1)}$ normalizes

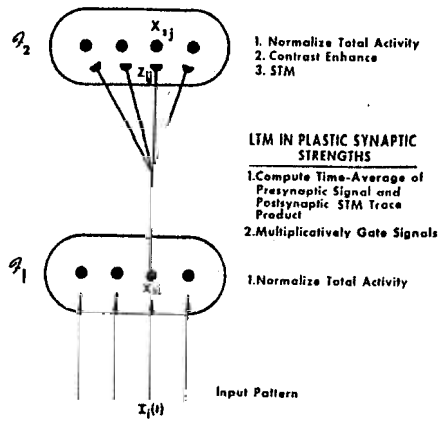


FIG. 31. The input pattern is normalized at \mathcal{F}_1 . The normalized signals are gated and added on their way to each population in \mathcal{F}_2 . The \mathcal{F}_2 inputs compete (adaptation and contrast enhancement) for storage in STM. The STM activities drive LTM traces via feedback, so that there appears to be a competition for synaptic sites.

itself before the LTM traces have a chance to change substantially; that is, it is "fast" relative to the "slow" LTM time scale.

Let the synaptic strength of the pathway from $v_i^{(1)}$ to $v_j^{(2)}$ be denoted by $z_{ij}(t)$. By (1), the total signal to $v_j^{(2)}$ due to the normalized pattern $\theta = (\theta_1, \theta_2, \dots, \theta_n)$ at $\mathcal{F}^{(1)}$ and the vector $z_j(t) = (z_{1j}(t), z_{2j}(t), \dots, z_{nj}(t))$ of synaptic strengths is [in the simplest case of $B_{ij}(w) = w$]

$$S_j(t) = \theta \cdot z_j(t) = \sum_{k=1}^n \theta_k z_{kj}(t) \quad (18)$$

Since $z_j(t)$ determines the size of the input to $v_j^{(2)}$, given any pattern θ , it is called the *classifying vector* of $v_j^{(2)}$ at time t . Every $v_j^{(2)}$, $j = 1, 2, \dots, N$, in $\mathcal{F}^{(2)}$ receives such a signal when θ is active in $\mathcal{F}^{(1)}$. In this way, θ creates a pattern of activity across $\mathcal{F}^{(2)}$.

Suppose that $\mathcal{F}^{(2)}$ is endowed with a recurrent on-center off-surround network undergoing mass action interactions. In particular, let $\mathcal{F}^{(2)}$ normalize and contrast-enhance its signals before they are stored in STM. In effect, the populations in $\mathcal{F}^{(2)}$ compete for STM activity. These STM operations also occur much faster than the slow changes in connection strengths z_{ij} ; hence, it is assumed below that these operations occur instantaneously in order to focus on the slow changes in z_{ij} . These slow changes in the z_{ij} 's will pick out populations in $\mathcal{F}^{(2)}$ —the feature detectors—to code the spatial patterns perturbing $\mathcal{F}^{(1)}$.

We now summarize the coding process. Suppose for simplicity that all changes in z_{ij} are driven by the *feedback* within the excitatory recurrent loops in $\mathcal{F}^{(2)}$ that establish STM storage. In other words, the "fast" competition for STM activity feeds back as a "slow" competition for LTM activity, or for synaptic site efficacy. Then (2) becomes

$$\dot{z}_{ij} = (-z_{ij} + \theta_i) x_j^{(2)} \quad (19)$$

where $\sum_{k=1}^N x_k^{(2)}(t) = 1$ if STM in $\mathcal{F}^{(2)}$ is active at time t , whereas $\sum_{k=1}^N x_k^{(2)}(t) = 0$ if STM in $\mathcal{F}^{(2)}$ is inactive at time t . Two important cases can be distinguished.

If tuning is adjusted so that $\mathcal{F}^{(2)}$ chooses a population for storage in STM, then

$$x_j^{(2)} = \begin{cases} 1 & \text{if } S_j > \max\{\epsilon, S_k : k \neq j\} \\ 0 & \text{if } S_j < \max\{\epsilon, S_k : k \neq j\} \end{cases} \quad (20)$$

except if two or more populations have a maximum signal; in the latter case, share the total STM activity ($= 1$) among the maximally excited populations. The parameter ϵ in (20) is the *QT*; if no signal exceeds the *QT*, no STM storage occurs.

If *partial contrast* in STM holds at $\mathcal{F}^{(2)}$, then the dynamics of the network can be approximated by a rule of the form

$$x_j^{(2)} = \begin{cases} f(S_j) [\sum_{S_k > \epsilon} f(S_k)]^{-1} & \text{if } S_j > \epsilon \\ 0 & \text{if } S_j < \epsilon \end{cases} \quad (21)$$

In (21), ϵ represent the *QT*; $f(w)$ controls how suprathreshold signals will be contrast-enhanced [for example, $f(w) = w^2$]; and the ratio of $f(S_j)$ to $\sum\{f(S_k) : S_k > \epsilon\}$ expresses the normalization of STM.

18. Feature Detectors

To see how these rules classify patterns, first hold all z_{ij} constant. Then Eqs. (18) and (20) reduce to the statement that population $v_j^{(2)}$ is stored in STM if

$$S_j > \max\{\epsilon, S_k : k \neq j\} \quad (22)$$

where $S_j = \theta \cdot z_j$ and $\theta_i = I_i (\sum_k I_k)^{-1}$. In other words, $v_j^{(2)}$ codes all patterns θ such that (22) holds. Stated alternatively, $v_j^{(2)}$ is a *feature detector* in the sense that all patterns

$$P_j = \{\theta : \theta \cdot z_j > \max\{\epsilon, \theta \cdot z_k : k \neq j\}\} \quad (23)$$

are classified by $v_j^{(2)}$. The set P_j defines a *convex cone* C_j in the space of

nonnegative input vectors $J = (I_1, I_2, \dots, I_n)$, since if two such vectors J_1 and J_2 are in C_j , then so are all the vectors αJ_1 , βJ_2 , and $\gamma J_1 + (1 - \gamma)J_2$, where $\alpha > 0$, $\beta > 0$, and $0 < \gamma < 1$. The convex cone C_j defines the *features* coded by $v_j^{(2)}$.

The classification rule in (22) has an informative geometrical interpretation in n -dimensional Euclidean space. The signal $S_j = \theta \cdot z_j$ is the inner, or dot, product of θ and z_j (Thomas, 1968). Letting $\|\xi\| = (\sum_{k=1}^n \xi_k^2)^{1/2}$ denote the Euclidean length of any real vector $\xi = (\xi_1, \xi_2, \dots, \xi_n)$, and $\cos(\eta, \omega)$ denote the cosine of the angle between two vectors η and ω , it is elementary that

$$S_j = \|\theta\| \|z_j\| \cos(\theta, z_j)$$

The signal S_j is thus the length of the projection of the normalized pattern θ on the classifying vector z_j times the length of z_j . In effect, θ is projected on all the cells v_{z_j} by their classifying vectors z_j . Thus if all z_j , $j = 1, 2, \dots, N$, have equal length, then (22) classifies all patterns θ in P_j whose angle with z_j is smaller than the angles between θ and any z_k , $k \neq j$, and is small enough to satisfy the ϵ -condition. In particular, patterns θ that are *parallel*, or proportional to z_j , are classified in P_j . The choice of classifying vectors z_j hereby determines how the patterns θ will be divided up. Section 19 describes how the adaptive coding mechanism (18) through (20) makes the z_j vectors more parallel to prescribed patterns θ , and thereupon changes the classifying sets P_j . In summary, given (20):

(i) The number of populations in $\mathcal{F}^{(2)}$ determines the maximum number N of pattern classes P_j .

(ii) The choice of classifying vectors z_j determines how different the sets P_j can be; for example, choosing all vectors z_j equal will generate one class that is redundantly represented by all $v_j^{(2)}$.

(iii) The size of ϵ determines how similar patterns must be to be classified by the same $v_j^{(2)}$ if the z_k 's are not the same.

If the choice rule (20) is replaced by the partial contrast rule (21), then an important new possibility occurs, which can be described by studying STM responses either to all θ at fixed $v_j^{(2)}$, or to a fixed θ at all $v_j^{(2)}$. In the former case, each $v_j^{(2)}$ has a *tuning curve*; namely, a maximal response to certain patterns, and submaximal responses to other patterns. In the latter case, each pattern θ is *filtered* by $\mathcal{F}^{(2)}$, or generates a *generalization gradient*, in a way that shows how close θ lies to *each* of the classifying vectors z_j . The pattern will be classified by $v_j^{(2)}$ —that is, stored in STM—only if it lies sufficiently close to z_j for its signal to exceed the quenching threshold of $\mathcal{F}^{(2)}$. It can be shown that the

existence of tuning curves in a given cortical field $\mathcal{F}^{(i)}$ increases the discriminative capabilities of the next cortex $\mathcal{F}^{(i+1)}$ in a hierarchy (cf. Grossberg, 1976a).

19. Development of an STM Code

System (18) through (20) will be discussed below because it illustrates important properties of the coding process in a lucid way. First, consider how the system responds to a single pattern θ that is iteratively presented through time. Suppose that there exists a unique j in $\mathcal{F}^{(2)}$ such that

$$S_j(0) > \max\{\epsilon, S_k(0): k \neq j\}$$

That is, $v_j^{(2)}$ is the population that initially receives the largest signal from $\mathcal{F}^{(2)}$. In particular, the signal $S_j(0)$ might be only slightly larger than the other signals $S_k(0)$, $k \neq j$. Then, as θ is presented to $\mathcal{F}^{(1)}$, the angle between $z_j(t)$ and θ , as n -vectors, monotonically decreases to zero. In other words, coding makes z_j parallel to θ . Simultaneously, the signal $S_j(t)$ monotonically approaches $\|\theta\|^2$; in other words, the coding process maximizes the inner product signal $S_j(t) = \theta \cdot z_j(t)$ over all possible choices of z_j such that $\|z_j\| \leq \|\theta\|$. Thus, whereas the initial signal $S_j(0)$ to $v_j^{(2)}$ might be only slightly larger than the signals $S_k(0)$, $k \neq j$, after coding takes place, θ generates a maximal signal to $v_j^{(2)}$. The other signals $S_k(t)$, $k \neq j$, remain constant.

By changing $z_j(t)$, the coding process changes the class of patterns θ that will be coded by $v_j^{(2)}$. For example, patterns θ that were originally more parallel to $z_j(0)$ than any $z_k(0)$, $k \neq j$, and which are therefore coded by $v_j^{(2)}$ at $t = 0$, can become more parallel to some $z_k(0)$ than $z_j(T)$, and are therefore coded by $v_k^{(2)}$ at $t = T$. Conversely, patterns that were originally more parallel to some $z_k(0)$ than $z_j(0)$, and which are coded by $v_k^{(2)}$ at $t = 0$, can become more parallel to $z_j(T)$ than $z_k(0)$, and are therefore coded by $v_j^{(2)}$ at $t = T$. In effect, presenting pattern θ at $\mathcal{F}^{(1)}$ has shifted the convex set of patterns—the features—that will be coded by $v_j^{(2)}$.

In the terminology of development biology, the z_j define *positional gradients* between $\mathcal{F}^{(1)}$ and $\mathcal{F}^{(2)}$. The initial positional gradients determine what features $\mathcal{F}^{(2)}$ will try to classify. Thus, if $\mathcal{F}^{(1)}$ feeds signals in parallel to several adaptive fields, each with different positional gradients, then each field will try to classify different features of the data base at $\mathcal{F}^{(1)}$, even though all fields use the same computational machinery. A serial hierarchy $\mathcal{F}^{(1)}$, $\mathcal{F}^{(2)}$, $\mathcal{F}^{(3)}$, \dots of adaptive fields can also be

constructed, since each field $\mathcal{F}^{(i)}$ possesses the requisite condition of normalization that is used to prepare data for field $\mathcal{F}^{(i+1)}$. As i increases, the data that $\mathcal{F}^{(i)}$ feeds to $\mathcal{F}^{(i+1)}$ include increasingly abstract convex sets of features (cf. Rocha-Miranda *et al.*, 1975). For example, what will a field $\mathcal{F}^{(3)}$ code if it receives patterned signals from $\mathcal{F}^{(2)}$? A population in $\mathcal{F}^{(3)}$ will code convex sets of spatial patterns across $\mathcal{F}^{(2)}$. Each population in $\mathcal{F}^{(2)}$ codes a set of features, so a spatial pattern across $\mathcal{F}^{(2)}$ is a global construct that describes how much of each feature is in the input pattern to $\mathcal{F}^{(1)}$. A convex set of spatial patterns across $\mathcal{F}^{(2)}$ describes the tolerated changes, or "fuzziness," in each feature's activity that are compatible with unchanged coding at $\mathcal{F}^{(3)}$. Cells in $\mathcal{F}^{(3)}$ can hereby generate stable responses as the input patterns to $\mathcal{F}^{(1)}$ undergo significant global transformations. If the positional gradients from $\mathcal{F}^{(i)}$ to $\mathcal{F}^{(i+1)}$ only excite a localized cluster of populations, then the populations in $\mathcal{F}^{(i+1)}$ will respond to features only in prescribed regions of $\mathcal{F}^{(i)}$ —for example, in the simplest case, prescribed retinal regions. If also the on-center and off-surround interactions within $\mathcal{F}^{(i+1)}$ decrease with distance, so that normalization only holds approximately and among prescribed subsets, or channels, of populations, then features in prescribed regions will be enhanced and/or suppressed only by particular classes of related features in nearby regions. Changing the distribution of positional gradients and on-center off-surround interactions can thus dramatically influence which features are computed and which features mutually influence one another.

What happens if several different spatial patterns $\theta^{(k)} = (\theta_1^{(k)}, \theta_2^{(k)}, \dots, \theta_n^{(k)})$, $k = 1, 2, \dots, M$, all perturb $\mathcal{F}^{(1)}$ at different times? How are changes in the z_j 's due to one pattern prevented from contradicting changes in the z_j 's due to a different pattern? The choice-making property of $\mathcal{F}^{(1)}$ helps to do this; it acts as a sampling device that often prevents contradictions from occurring. The following argument suggests how sampling works. This argument is, however, not entirely correct. For definiteness, suppose that M spatial patterns $\theta^{(k)}$ are chosen, $M \leq N$, such that their signals at time $t = 0$ satisfy

$$\theta^{(k)} \cdot z_k(0) > \max\{\epsilon, \theta^{(k)} \cdot z_j(0); j \neq k\} \quad (24)$$

for all $k = 1, 2, \dots, M$. In other words, at time $t = 0$, $\theta^{(k)}$ is coded by $v_k^{(2)}$. Let $\theta^{(1)}$ be the first pattern to perturb $\mathcal{F}^{(1)}$. By (24), population $v_1^{(2)}$ receives the largest signal from $\mathcal{F}^{(1)}$. All other populations $v_j^{(2)}$, $j \neq 1$, are thereupon inhibited by the off-surround of $v_1^{(2)}$, whereas $v_1^{(2)}$ reverberates in STM. By (19), none of the synaptic strengths $z_j(t)$, $j \neq 1$, can learn while $\theta^{(1)}$ is presented. Presenting $\theta^{(1)}$ makes $z_1(t)$ more parallel to $\theta^{(1)}$ as t increases. Consequently, if a different pattern, say $\theta^{(2)}$, perturbs

$\mathcal{F}^{(1)}$ on the next learning trial, then it will excite $v_2^{(2)}$ more than any other $v_j^{(2)}$, $j \neq 2$: it cannot excite $v_1^{(2)}$ because the coefficients $z_1(t)$ are more parallel to $\theta^{(1)}$ than before; and it cannot excite any $v_j^{(2)}$, $j \neq 1, 2$, because the $v_j^{(2)}$ coefficients $z_j(t)$ still equal $z_j(0)$. In response to $\theta^{(2)}$, $v_2^{(2)}$ inhibits all other $v_j^{(2)}$, $j \neq 2$. Consequently none of the $v_j^{(2)}$ coefficients $z_j(t)$ can learn, $j \neq 2$; learning makes the coefficients $z_2(t)$ become more parallel to $\theta^{(2)}$ as t increases. The same occurs on all learning trials. By inhibiting the postsynaptic part of the learning mechanism in all but the chosen $\mathcal{F}^{(2)}$ population, the on-center off-surround network in $\mathcal{F}^{(2)}$ samples one vector $z_j(t)$ of LTM traces at any time. In this way $\mathcal{F}^{(2)}$ can learn to classify as many as N patterns if it contains N populations. (When tuning curves or resonant feedback exist, each population can share in the coding of several patterns.)

This argument is almost correct. It fails, in general, because by making (say) $z_1(t)$ more parallel to $\theta^{(1)}$, it is also possible to make $z_1(t)$ more parallel to $\theta^{(2)}$ than $z_2(0)$ is. Thus when $\theta^{(2)}$ is presented, it will be coded by $v_1^{(2)}$ rather than $v_2^{(2)}$. In other words, practicing one pattern can recode other patterns. This property can be iterated to show how systematic trends in the sequence of practiced patterns can produce systematic drifts in recoding (Grossberg, 1976b). Moreover, if the statistical structure of the practice sequences continually changes, then there need not exist a stable coding rule in $\mathcal{F}^{(2)}$. This is quite unsatisfactory.

By contrast, if there are few, or sparse, patterns relative to the number of populations in $\mathcal{F}^{(2)}$, then a stable coding rule does exist, and the STM choice rule in $\mathcal{F}^{(2)}$ does provide an effective sampling technique. In effect, given any fixed class of patterns at $\mathcal{F}^{(1)}$ and sufficiently many populations in $\mathcal{F}^{(2)}$, the $\mathcal{F}^{(1)}$ patterns can induce a stable STM code in $\mathcal{F}^{(2)}$. By contrast, the problem of stabilizing the STM code given a fixed number of cells in $\mathcal{F}^{(2)}$ and arbitrarily many patterns θ at $\mathcal{F}^{(1)}$ requires additional network mechanisms. This problem is studied in Grossberg (1976b,c). In this case the cells in $\mathcal{F}^{(2)}$ can continually be recoded by patterns at $\mathcal{F}^{(1)}$. No stable hierarchy of codes could develop using only this mechanism, since the coded meaning of the signals from one level to the next would be continually changing. Below are reviewed relevant aspects of how a developing code can be stabilized in an arbitrary environment. In passing, we note that rules such as (20) and (22) define discriminant functions of a type that is familiar in pattern classification studies, and is related to Bayesian decision rules that make choices to minimize risk (Duda and Hart, 1973, Chapter 2). If terms such as $\theta \cdot z_j$ are generalized to $f(\theta) \cdot z_j$, where $f(\theta) = (f_1(\theta_1), f_2(\theta_2), \dots, f_n(\theta_n))$, then the decision boundaries of the discriminant

functions are not necessarily convex. Our method departs from the classical development in several ways. One way is described by the rules whereby the $z_j(t)$ vectors shift, owing to learning. A more fundamental way is described in the next section, which shows how local discriminants, or features, are synthesized into a global code by adaptive resonance between two fields of cells. The local properties of the field, by themselves, neither define feature detectors nor when adaptation of feature detectors will occur. The functional unit of coding is a global feedback module that I call an adaptive resonance.

20. Stabilizing the STM Code: Expectation, Resonance, Rebound, and Search

To stabilize the code, it suffices to use attentional mechanisms. These mechanisms were introduced in Grossberg (1975). The reverse statement is also true: the minimal mechanisms for stabilizing the STM code can also generate various attentional phenomena. Why is there a relationship between code stability and attention? The next example motivates this relationship.

Suppose that a population $v_j^{(2)}$ in $\mathcal{F}^{(2)}$ already codes a given class of patterns \mathcal{P}_j at $\mathcal{F}^{(1)}$, and that a pattern θ not in this class succeeds in activating $v_j^{(2)}$. If this activation is not rapidly terminated, then recoding of the LTM traces will occur, since by (19) z_j can learn θ while $v_j^{(2)}$ is active. Somehow, sustained activation of $v_j^{(2)}$ by an erroneous pattern θ must be prevented; activity in $v_j^{(2)}$ is somehow inhibited. This can happen only if the network can determine that $v_j^{(2)}$ codes a pattern class that is incompatible with θ . Furthermore, the operation that inhibits $v_j^{(2)}$ cannot inhibit all populations in $\mathcal{F}^{(2)}$; otherwise θ could not find any population in $\mathcal{F}^{(2)}$ to code it. Somehow the network selectively inhibits the erroneously activated populations $v_j^{(2)}$ before it searches for an uncommitted population with which to code θ . Given these remarks, it is not surprising that STM code stability is related to attention; stability requires the network to selectively activate populations whose codes are compatible with the sensory data of the moment.

The mechanisms in Grossberg (1976c) describe how a test pattern θ at $\mathcal{F}^{(1)}$ can tentatively activate feature detectors in $\mathcal{F}^{(2)}$, which thereupon generate feedback signals either to $\mathcal{F}^{(1)}$ or to a field $\mathcal{E}^{(1)}$ that acts in parallel with $\mathcal{F}^{(1)}$. These feedback signals represent an expectation, or template, with which the afferent test pattern at $\mathcal{F}^{(1)}$, or its parallel representation at $\mathcal{E}^{(1)}$, is compared (cf. Section 44). This expectation is a spatial pattern that can be learned by the $\mathcal{F}^{(2)} \rightarrow \mathcal{F}^{(1)}$ (or $\mathcal{F}^{(2)} \rightarrow \mathcal{E}^{(1)}$) LTM traces at the same time that chunks are being coded by the $\mathcal{F}^{(1)} \rightarrow \mathcal{F}^{(2)}$

LTM traces. If the test pattern matches the expectation, then the patterned STM activity in $\mathcal{F}^{(1)}$ and $\mathcal{F}^{(2)}$ is amplified and can resonate between $\mathcal{F}^{(1)}$ and $\mathcal{F}^{(2)}$. This resonant activity can activate the STM of other fields, say of higher-order feature detectors or of motor commands; it can also drive slow LTM changes in synapses that sample $\mathcal{F}^{(1)}$ or $\mathcal{F}^{(2)}$.

Suppose, however, that the expected and test patterns are very different. This mismatch means that an erroneous classification has occurred at $\mathcal{F}^{(2)}$. An alarm system is thereupon triggered that generates a nonspecific wave of input activity across $\mathcal{F}^{(2)}$. The alarm system acts nonspecifically because, at the place where the mismatch is computed, no data is available concerning which populations in $\mathcal{F}^{(2)}$ have erroneously been activated. The nonspecific signal must somehow selectively inhibit, or reset, the active populations of $\mathcal{F}^{(2)}$ without preventing inactive populations from being tentatively activated during the next time interval. In effect, the populations whose activity set off the alarm must have been erroneously classified, so they should be selectively suppressed. This idea realizes a kind of probabilistic logic operating in real time, with activity level replacing truth value.

The inhibition of active populations must be enduring as well as selective. Otherwise, the inhibited populations could immediately be reactivated by the pattern θ . Given such an inhibitory mechanism, the network automatically searches for a population which is not compatibly classified. When one is found, say $v_i^{(2)}$, STM at $v_i^{(2)}$ can stay on long enough to drive the "slow" coding process in the z_i LTM traces.

How is a selective and enduring inhibition at $\mathcal{F}^{(2)}$ effected? I suggest that it is due to the organization of $\mathcal{F}^{(2)}$ into antagonistic pairs, or dipoles, of populations. In effect, the antagonistic population to an active population $v_i^{(2)}$ is turned on when the nonspecific alarm goes off, and thereupon selectively inhibits $v_i^{(2)}$; that is, if "yes" at $v_i^{(2)}$ is wrong, turn on "no" at $v_i^{(2)}$, but in a graded fashion.

The idea of population dipoles was not originally introduced to stabilize the STM code, although this is a fundamental reason for their existence. Originally the idea arose in a neural theory of reinforcement, wherein cell dipoles regulate net incentive motivation through time (Grossberg, 1972b,c, 1975); also see Wise *et al.* (1973) for compatible data. Properties of the rebound from positive (negative) incentive to negative (positive) incentive through time are analogous to many paradoxical phenomena about reinforcement—for example, how an amphetamine can calm an agitated syndrome that is really a form of underaroused emotional depression, whereas overaroused depression can yield indifference to the emotional meaning of cues.

More generally, suppose that the on-cell of a dipole is activated persistently by the presence of its external cue, whereas the off-cell is activated transiently by the offset of the cue. Otherwise expressed, offset of the cue elicits a transient antagonistic rebound. This transient activity can be used to sample STM patterns at the synaptic knobs of the off-cell and encode these patterns in LTM. Hereby the offset of a cue can elicit learned behavior.

When the antagonistic rebound is explicitly modeled, one is led to postulate the existence of slowly varying transmitter substances that multiplicatively gate all signals before they can reach the on-cells and off-cells. Among these signals are a tonically active nonspecific arousal that is distributed uniformly across all the cells. The arousal signal regulates the size of the off-cell rebound when the cue to the on-cell terminates. This happens as follows. When the cue is on, the total signal in the on-cell channel exceeds the total signal in the off-cell channel. Both signals are gated by transmitter before they reach their targets. Because the on-cell signal is larger, transmitter is depleted more in the on-cell channel than in off-cell channel. The on-cell nonetheless receives the larger input because of the multiplicative effect of signal and transmitter on the cells: the equilibrium transmitter level has the form $A[B + CS]^{-1}$ in response to a steady signal S , and therefore decreases as S increases; but the equilibrium input has the form $ADS[B + CS]^{-1}$, which increases as S increases. When the cue is removed, equal arousal signals remain in both channels. Since transmitter level changes slowly, there is more transmitter in the off-cell channel. The multiplicative coupling of arousal signal to transmitter now gives the off-cell a larger input, thereby causing the rebound. Gradually, in response to the equal arousal signal in both channels, the transmitter levels also equalize, and both channels receive equal inputs, so that the rebound eventually terminates.

A property of this system is that a rapid increment in nonspecific arousal can, by itself, reverse or rebound the relative activities in a dipole. Thus, if an on-cell is active when arousal increases, then it can be inhibited by its off-cell, whereas if neither cell is active when arousal increases, then neither on-cell nor off-cell receives any relative advantage. The rebound therefore selectively inhibits active populations. If the on-cells are now hooked into a recurrent network capable of STM, and the off-cells are similarly organized, then it follows that a transient arousal increment can selectively, and in a graded fashion, inhibit active populations by shifting the STM pattern across both fields. When this mechanism acts on various fields of formal feature detectors, phenomena analogous to negative afterimages and spatial frequency adaptation

are found (Grossberg, 1976c). The properties of antagonistic rebound in a dipole of populations are useful for understanding many psychophysiological processes in which rapid shifts of specific cues and/or nonspecific arousal signal important events.

21. Pattern Completion, Hysteresis, and Gestalt Switching

The concept of two fields $\mathcal{F}^{(1)}$ and $\mathcal{F}^{(2)}$ joined together by reciprocal trainable signal pathways is also relevant to many psychophysiological processes, if only because it describes the minimal network module that can stabilize its STM code in a rich input environment. The two LTM processes—code-learning in $\mathcal{F}^{(1)} \rightarrow \mathcal{F}^{(2)}$ LTM traces and template-learning in $\mathcal{F}^{(2)} \rightarrow \mathcal{F}^{(1)}$ LTM traces—are partners in establishing a stable state of resonant STM activity in $\mathcal{F}^{(1)}$ and $\mathcal{F}^{(2)}$ when the active LTM channels are compatible. I call this module an *adaptive resonance*. Grossberg (1976c) summarizes some examples of this concept in olfactory coding, in the regulation of attention by the matching of presently available cues (conditioned reinforcers) with feedback from compatible drive sources (expressed through the contingent negative variation), and in a search and lock mechanism for stabilizing eye position. The next two sections describe several other important examples of this concept. These examples illustrate how closely related “perceptual” and “cognitive” properties can be.

First, feedback from $\mathcal{F}^{(2)}$ to $\mathcal{F}^{(1)}$ can deform what “is” perceived into what “is expected to be” perceived. Otherwise expressed, the feedback is a prototype, or higher-order Gestalt, that can deform, and even complete, activity patterns across lower-order feature detectors. For example, suppose that a sensory event is coded by an activity pattern across the feature detectors of a field $\mathcal{F}^{(1)}$. The $\mathcal{F}^{(1)}$ pattern is then coded by certain populations in $\mathcal{F}^{(2)}$. If the sensory event has never before been experienced, then the $\mathcal{F}^{(2)}$ populations that are chosen are those whose codes most nearly match the sensory event, because the pattern at $\mathcal{F}^{(1)}$ is projected onto $\mathcal{F}^{(2)}$ by the positional gradients in the $\mathcal{F}^{(1)} \rightarrow \mathcal{F}^{(2)}$ pathways. If no approximate match is possible, then these $\mathcal{F}^{(2)}$ populations will be inhibited and a search procedure will be elicited. If an approximate match is possible, however, then feedback signals from $\mathcal{F}^{(2)}$ to $\mathcal{F}^{(1)}$ will elicit the template of the sensory events that are *optimally* coded by the $\mathcal{F}^{(2)}$ pattern. These feedback signals gradually deform the $\mathcal{F}^{(1)}$ pattern until this pattern is a mixture of feedforward codes and feedback templates. Otherwise expressed, $\mathcal{F}^{(2)}$ tries to *complete* the $\mathcal{F}^{(1)}$ pattern using the prototype, or template, that its active populations

release. I suggest that many Gestalt-like pattern completion effects are special cases of this feedback mechanism, and therefore consequences of STM stability.

Two important manifestations of the completion property are hysteresis and Gestalt switching. Once an STM resonance is established, it resists changing its codes in response to small changes in the sensory event; this is hysteresis. Hysteresis occurs because the active $\mathcal{F}^{(2)} \rightarrow \mathcal{F}^{(1)}$ template keeps trying to deform the shifting $\mathcal{F}^{(1)}$ pattern back to one that will continue to code the $\mathcal{F}^{(2)}$ populations that elicit this template.

If, however, the sensory event changes so much that the mismatch of test and template patterns becomes too great, then the alarm-and-reset mechanism is triggered and a new code in $\mathcal{F}^{(2)}$ will be activated. This new code feeds its template back to $\mathcal{F}^{(1)}$ and deforms the $\mathcal{F}^{(1)}$ pattern toward its optimal pattern. A dramatic switch between global percepts can hereby be effected. The global nature of the switch is due to the fact that $\mathcal{F}^{(2)}$ contains codes that synthesize data from many feature detectors in $\mathcal{F}^{(1)}$, and the templates of these codes can reorganize large segments of the $\mathcal{F}^{(1)}$ field. I suggest that an analogous mixture of hysteresis and switching is operative in various visual illusions, such as Necker's cube (Graham, 1966). In fact, one can think of such illusions in the following light. Often when one slowly scans a visual scene, small shifts in perspective imply small shifts in higher-order codes. These illusions are *designed* so that small shifts in perspective imply large shifts in higher-order codes after the hysteresis range is exceeded and the reset mechanism is triggered. Also of interest are the "spontaneous" switches that can occur when ambiguous figures are persistently examined. By Section 20, a prolonged STM reverberation at a given population can partially deplete its transmitter, and thereby shift the relative balance of on-cell to off-cell excitation across the field in favor of other populations. If this shift is sufficiently large, it can induce a cyclic drift in activity across the field as populations cyclically deplete and accumulate their transmitter stores (cf. Section 40).

Hysteresis can also occur between two reciprocally connected fields that are not hierarchically organized. In particular, suppose that each eye activates a field of monocularly coded feature detectors. Suppose that each monocular field is endowed with a recurrent on-center off-surround anatomy. Also let corresponding detectors be capable of sending each other signals. It does not matter *what* features are coded by these detectors to draw the following conclusion. Once a resonance is established between the two fields, hysteresis will prevent small changes in input pattern from changing the coded activity. Julesz (1971) introduced a field of physical dipoles to model binocular hysteresis. Reso-

nance between two recurrent on-center off-surround networks undergoing mass action dynamics provides a neural model of the phenomenon.

In passing, we note that "reverberation" and "resonance" are not interchangeable concepts. Reverberation between $\mathcal{F}^{(1)}$ and $\mathcal{F}^{(2)}$ occurs even as $\mathcal{F}^{(2)}$ is continually reset in search of an admissible population. Resonance occurs only when the test pattern at $\mathcal{F}^{(1)}$ and the active codes at $\mathcal{F}^{(2)}$ are compatible.

22. Context-Dependent Coding and Restricted Conditions for Recoding

Can a field of feature detectors uniquely classify a sensory event even if no single detector is uniquely activated by that event? The answer is "yes" if feedback exists within the system. In particular, a resonant state of activity need not be established between $\mathcal{F}^{(1)}$ and $\mathcal{F}^{(2)}$ unless the pattern of activity across feature detectors in $\mathcal{F}^{(1)}$ correctly classifies the sensory event. Otherwise expressed, the code is context-dependent; each $\mathcal{F}^{(1)}$ population shares in "multiple meanings" by participating in the coding of many events (Chung *et al.*, 1970).

Not all sensory patterns need recode the network's feature detectors. Only patterns that generate a resonant state of activity can generate such a change. The resonant state explicates the intuitive idea that the network is attending to the pattern and has stored it in STM, whereupon it can induce recoding via LTM changes. In particular, mere passive presentation of patterns need not recode any feature detectors. Adaptational differences between passive and active responses to cues have been experimentally described (Held and Hein, 1963) and might be one factor that explains why certain experiments seem to reliably recode feature detectors, whereas others do not (Stryker and Sherk, 1975). From this perspective, terminating the critical period for developmental plasticity seems to depend on the switching on of attentional factors in addition to the possible switching off of a chemical agent. The paper by Grossberg (1977a) discusses this possibility for the visual system. Some themes are differential amplification or attenuation of lateral geniculate nucleus (LGN) activity depending on whether afferent sensory data match or mismatch cortical feedback to LGN; lateral inhibition in LGN as a matching mechanism, and thus the growth of LGN inhibitory pathways as a precursor of critical period termination; temporal modulation of catecholaminergic cortical arousal as a rebound trigger mechanism; and cortical dipole organization, possibly realized by feature detector pairs that code complementary features, and embedded in

recurrent on-center off-surround networks, as a mechanism to maintain the rebounded STM pattern.

23. Reset, Reaction Time, and P300

When a mismatch between test pattern and expected pattern occurs, STM is reset by a nonspecific mechanism that inhibits active populations until an admissible population is found. Several properties of the reset mechanism are of interest. First, STM does not spontaneously decay at a rapid rate, but can endure until it is actively reset by a variety of mechanisms: competing inputs, change of expectations, shift in arousal level, etc. The durability of STM in a recurrent network is in marked contrast to the rapid passive decay of activity in nonrecurrent networks. *In vivo*, one expects to find recurrent networks more centrally (for example, in neocortex) and nonrecurrent networks more peripherally (for example, in retina).

Second, unexpected events trigger STM reset by a nonspecific mechanism that inhibits ongoing activity. In average evoked potential experiments, one often finds an inhibitory wave, the P300, that accompanies unexpected events (Rohrbaugh *et al.*, 1974; Squires *et al.*, 1976). To discuss the model's relationship to P300, we consider an idealized example in Fig. 32. The active population $v_j^{(2)}$ in Fig. 32 generates a subliminal pattern at $\mathcal{F}^{(1)}$ that acts as a sensory expectation. Suppose that $v_j^{(2)}$ also generates a subliminal pattern at the motor control cells \mathcal{M} . This latter pattern acts as a motor expectation, or subliminal motor map.

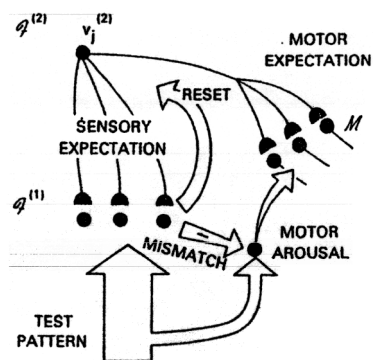


FIG. 32. Sensory and motor expectations are reset or released by mismatch or match, respectively, of sensory expectation with sensory test pattern.

Let a test pattern perturb $\mathcal{F}^{(1)}$. If the test pattern does not match the sensory expectation, then nonspecific arousal is triggered. How does this happen? We assume (Grossberg, 1975) that *every* test pattern can activate this arousal source. If the test pattern matches the sensory expectation, then a signal is released by $\mathcal{F}^{(1)}$ which inhibits the arousal on its way to $\mathcal{F}^{(2)}$. We also suppose that there is a complementary arousal system in Fig. 32. If the test pattern matches the sensory expectation, then the motor map in \mathcal{M} is aroused, and therefore sends signals to motor effectors. If the test pattern does not match the sensory expectation, then the motor arousal is inhibited. In summary, a mismatch can excite the reset arousal system as it inhibits the goal-oriented motor arousal system (it says "no"), whereas a match can inhibit the reset arousal system as it excites the goal-oriented motor arousal system (it says "yes"). There are many variations on this theme; for example, the sensory expectation might merely predispose $\mathcal{F}^{(1)}$ to excite a particular $\mathcal{F}^{(2)}$ population which thereupon excites \mathcal{M} ; or a match between patterns might activate a signal that inhibits a nonspecific inhibition, which in turn disinhibits pacemakers in $\mathcal{F}^{(2)}$ that drive dipole rebounds. The main points are independent of such details.

If the reset mechanism is indeed related to P300, then P300 is conceptually different from the model's CNV, which subserves a conditionable "psychological set." Donchin *et al.* (1975) report compatible data. Furthermore, if P300 resets STM, it can do so without eliciting motor activity. See Donchin *et al.* (1972) for compatible data.

If the test pattern matches the sensory expectation, then the system's reaction time is less than if the test pattern does not match the sensory expectation. This is due to two effects working together: In the former case, the motor pattern can be released without a prior reset of STM, and the subliminal expectation reduces the reaction time for suprathreshold signals to be emitted from $\mathcal{F}^{(1)}$. Moreover, the reaction time will be a monotone increasing function of the inhibitory rebound size, since the latter index measures how extensively STM must be reset before an appropriate response to the test pattern can be elicited. Analogously, reaction time is an increasing function of P300 size (Squires *et al.*, 1976). Finally, Section 36 shows that a given population can code a sequence of events, that longer sequences can be coded by more cellular sites, at least up to some maximal length, and that these long-sequence codes are harder to inhibit than short-sequence codes. This leads to the prediction that resetting the population of a long sequence can take longer than resetting the population of a short sequence, and that P300 should again be a monotone increasing function of reaction time. Analogous data have been reported by Remington (1969) and Squires *et al.* (1976).

Various other properties of adaptive resonances can be cited, but go beyond the scope of this paper. One of the deepest and least understood concerns the periodic oscillations of STM activity that can arise in an adaptively resonating network. These oscillations are due to response lags of the inhibitory cells that generate off-surround signals relative to the excitatory cells that generate on-center signals. These periodic oscillations are formal analogs of brain rhythms in the model. They are discussed in Grossberg (1977a). Section 35 discussed a possible effect of these slow waves on interference due to delayed auditory feedback.

24. Hierarchical Critical Periods and Retrograde Amnesia

Several fundamental points should be emphasized about the role of feedback in stabilizing the STM code. The coded data are not locally defined; they are expressed by patterns across many cells. Hence no local signal is sufficient to determine *what* is being coded, or whether the code is a useful one behaviorally. Nonetheless, Grossberg (1974, Section VII) shows that each network cell that is capable of LTM is a chemical dipole. This dipole can, in principle, determine when LTM-like activity is turned on, and can therefore turn it off a fixed time later using an internal cellular clock. Once this happens, no reset of LTM can occur. Switching off LTM changes would end the cell's critical period of developmental plasticity.

Such a clock mechanism would seem to be satisfactory only if the coded data do not depend on individual experience. Within fields $\mathcal{F}^{(i)}$ that are at least partly determined by experience, I prefer the view that LTM stabilization by a clock is supplemented, or even supplanted, by signals from a higher-order field $\mathcal{F}^{(i+1)}$ which has learned templates of feedback signals. The present paper argues that this feedback mechanism is useful in generating a self-consistent code whether or not a chemical clock exists. Moreover, it explains several facts parsimoniously: why it takes a while for LTM in $\mathcal{F}^{(i)} \rightarrow \mathcal{F}^{(i+1)}$ synapses to become stabilized (critical period); why the highest-order code remains plastic; why lower-order codes can be reset by sufficient disconfirmation despite their stability; why the oldest memories are the most stable, since they are the best buffered by feedback; and thus (one reason!) why newer memories are often easier to erase than older ones (retrograde amnesia).

Given the above results on adaptive coding, the question of how sequences of events can be coded by prescribed populations can be broached.

25. Invariance of the Past Code under Future Sequential Inputs

We now make a series of crucial observations. Suppose as in Fig. 24 that a certain number of states $v_1^{(1)}, v_2^{(1)}, \dots, v_i^{(1)}$ have already received serial inputs, and therefore have positive STM traces. At a given time, these traces define a spatial pattern. In other words, the temporal order of inputs is coded spatially. Imagine that $\mathcal{F}^{(1)}$ sends trainable signals to $\mathcal{F}^{(2)}$, as in the coding model of Sections 17 through 19. Then a certain population in $\mathcal{F}^{(2)}$ will be excited by this particular spatial pattern, and will reverberate in STM. Suppose that the same states had been perturbed by a different serial ordering. Then the activities of the populations $v_1^{(1)}, v_2^{(1)}, \dots, v_i^{(1)}$ would be permuted, and a very different population in $\mathcal{F}^{(2)}$ could be excited by the new spatial pattern. In other words, the spatial pattern of activity across these cells carries temporal, or order, information, whereas the adaptive coding process whereby each $v_i^{(1)}$ is excited carries item information. The fact that different spatial patterns on the same set of cells in $\mathcal{F}^{(1)}$ can excite different cells in $\mathcal{F}^{(2)}$ is a crucial one in our analysis. Knowing only *that* the cells $v_1^{(1)}, v_2^{(1)}, \dots, v_i^{(1)}$ are excited does *not* determine the order in which they were excited, nor which cells in $\mathcal{F}^{(2)}$ will fire. The enormous loss of information when a graded spatial pattern is replaced by a binary on-off code weakens algebraic theories of sequential coding by preventing activity per se from generating order information. Our parallel code for order information must then be replaced by a serial code that depends on an item's position in a serially organized buffer (cf. Atkinson and Shiffrin, 1968). As soon as item codes must move through a buffer, one is beleaguered by problems concerning how the code shifts along the buffer, how the code can elicit the same motor output at different buffer positions, and how a given buffer position can elicit the motor output of every item. None of these difficulties arises in the present theory. More generally, the entire design of the system must be changed when its code is changed.

Once a sequence $v_1^{(1)}, v_2^{(1)}, \dots, v_i^{(1)}$ is already presented, its spatial pattern represents "past" order information. Presenting a new input to $v_{i+1}^{(1)}$ can reorganize the *total* pattern of coded STM activity at $\mathcal{F}^{(2)}$, but we shall assume that it does not recode that part of this coded activity which involves only past order information. In other words, new inputs can weaken the strength of past codes, but do not deny the fact that the past events did occur. If this were not the case, the code in $\mathcal{F}^{(2)}$ would become very unstable through time, since every little perturbation at $\mathcal{F}^{(1)}$ could destroy the entire past record of events. We therefore impose the following basic constraint on STM coding.

Invariance Principle. The spatial patterns at $\mathcal{F}^{(1)}$ are generated by sequential inputs in such a way as to leave the codes in $\mathcal{F}^{(2)}$ of past sequences invariant.

This principle implies rules for generating spatial patterns across $\mathcal{F}^{(1)}$ in response to input sequences. Suppose that at time t , the STM activities at $v_1^{(1)}, v_2^{(1)}, \dots, v_i^{(1)}$ form a pattern $P_i(t) = (x_1^{(1)}(t), x_2^{(1)}(t), \dots, x_i^{(1)}(t))$. By (18), the signal to any cell $v_j^{(2)}$ in $\mathcal{F}^{(2)}$ is $S_j(t) = P_i(t) \cdot z^{(2)}(t)$. In order not to recode any of these populations in response to this sequence, $P_i(t)$ must not change its direction, as a Euclidean i -vector, at later times t . In other words,

$$P_i(t) = P_i(T)f(t, T) \quad (25)$$

where $f(t, T)$ is a scalar function of $t \geq T$. The invariance principle thus implies that, after $x_1^{(1)}, x_2^{(1)}, \dots, x_i^{(1)}$ are excited by sequential inputs, they thereafter undergo proportional changes through time. Table II describes rules for generating such changes. It displays STM activities at discrete times which are synchronized with the input presentation rate. In Table II, μ_i is the activity of the last item to be presented. Thus, when $v_1^{(1)}$ is first excited, it reaches activity μ_1 . Then $v_2^{(1)}$ is excited with activity μ_2 , and $v_1^{(1)}$'s activity is changed from μ_1 to $\mu_1\omega_2$ to satisfy the Invariance Principle. Then $v_3^{(1)}$ is excited with activity μ_3 . The "past" information is scaled down by ω_3 to yield activity $\omega_3\omega_2\mu_1$ at $v_1^{(1)}$ and activity $\omega_3\mu_2$ at $v_2^{(1)}$. And so on, until an item's STM trace falls below the QT , whence it is quenched. Physical intuition suggests, at least in cases where all the list items are equivalent in all important respects, that μ_k and ω_k are functions of the total activity in the $(k-1)$ st time frame, and of a parameter ν that represents input strength; that is,

$$\mu_k = U(\nu, S_{k-1}) \quad (26)$$

and

$$\omega_k = V(\nu, S_{k-1}) \quad (27)$$

TABLE II
INVARIANCE PRINCIPLE CONSTRAINS POSSIBLE
SEQUENTIAL STM VALUES

$t \equiv 1$	μ_1	0	0	0
$t \equiv 2$	$\mu_1\omega_2$	μ_2	0	0
$t \equiv 3$	$\mu_1\omega_2\omega_3$	$\mu_2\omega_3$	μ_3	0

where the total activity of the STM field in the j th time frame is

$$S_j = \sum_{m=1}^j \mu_m \prod_{r=m+1}^j \omega_r \quad (28)$$

Since these relations are merely approximate guides to construct the networks that actually perform the computations, we content ourselves with discussing some special cases that provide important physical insights.

First we note what the rules do *not* say. They do not say whether or not STM activity at $\mathcal{F}^{(1)}$ spontaneously decays. Spontaneous decay is permitted, but not required. For example, if the $\mathcal{F}^{(1)}$ activities exponentially decay, then (25) is satisfied, with $f(t, T) = \exp[-A(t - T)]$. If no decay occurs, then all changes in $\mathcal{F}^{(1)}$ are induced by the presentation of new items. Individual cells in $\mathcal{F}^{(2)}$ then code order information, but not the time spacing between items. If spontaneous decay does occur, then different time spacings of the same items in the same order can generate different spatial patterns across $\mathcal{F}^{(1)}$. When this occurs, individual cells in $\mathcal{F}^{(2)}$ can code a mixture of order and spacing information.

If two different input sequences at $\mathcal{F}^{(1)}$ create similar (distinct) STM patterns at $\mathcal{F}^{(2)}$, then the "distance" between these sequences is small (large). The above remarks illustrate how different distance functions on the space of input sequences at $\mathcal{F}^{(2)}$ can be generated by different decay rules for $\mathcal{F}^{(1)}$.

26. Bowing of the STM Pattern

Two extreme cases of the sequence-generating rules will now be considered. More realistic cases will often be mixtures of the two extremes, as Section 27 illustrates.

Case I. Let items be presented at a fixed rate such that the most recent item has unit strength, and each item's strength decays at a fixed rate in time. Then all $\mu_i = \mu$ and $\omega_i = \omega$. Set $\mu = 1$ for convenience. The distribution in Table IIIA is found. This distribution describes the situation in Fig. 24a, in which new items can excite their states with a strength that is independent of the amount of prior STM activity in the field.

A very different result occurs if total field activity is normalized, say due to a recurrent on-center off-surround network. Let all $\omega_i = \omega$, and choose the total activity equal to 1. Then Table IIIB is generated. Parts A and B of Table III share certain similarities. In both, the last item in any list of length at least two has the same STM activity no matter how

TABLE III
(A) PASSIVE DECAY OF STM TRACES

$t \equiv 1$	1	0	0	0
$t \equiv 2$	ω	1	0	0
$t \equiv 3$	ω^2	ω	1	0

(B) NORMALIZATION OF AN STM PATTERN THAT OBEYS THE INVARIANCE PRINCIPLE

	$x_2^{(1)}$				
$t \equiv 1$	1	0	0	0	0
$t \equiv 2$	ω	$1 - \omega$	0	0	0
$t \equiv 3$	ω^2	$(1 - \omega)\omega$	$1 - \omega$	0	0
$t \equiv 4$	ω^3	$(1 - \omega)\omega^2$	$(1 - \omega)\omega$	$1 - \omega$	0

many prior items are in the list, which makes good intuitive sense. Also, if $0 < \omega < \frac{1}{2}$, the earliest items have the weakest STM strengths.

An important new phenomenon occurs if $\frac{1}{2} < \omega < 1$. Then the STM pattern that is produced by serial inputs can develop a bow without any intervention of LTM. To see this, consider time step $t = 2$. Since $\omega > 1 - \omega$, the *earlier* item has a larger STM trace. If $t = 3$, then $\omega^2 > (1 - \omega)\omega < 1 - \omega$, so that there is a bow in the STM pattern at the second list item. An important parameter is the maximal k such that $\omega^k > 1 - \omega$. Denote it by K . The longest list length for which the first item has a larger STM activity than any other item is then $K + 1$. Every list length $k \leq K + 1$ is said to exhibit a *primacy* effect. Lists of length $k > K + 1$ exhibit a *recency* effect (Fig. 33).

27. Regulation of STM Primacy, Recency, and Bowing by Lateral Inhibition

The size of parameter ω in Table IIIB measures the relative balance between STM maintenance by recurrent intrafield interactions and STM

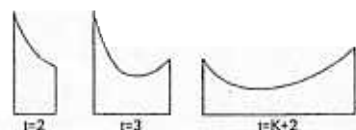


FIG. 33. Bowing of the STM pattern at list length 3, and recency effect at list length $K + 2$.

reset due to inhibition triggered by new inputs. A small value of ω represents powerful reset inhibition. Strengthening reset inhibition decreases ω , which decreases K , which enhances recency by making it easier for the STM traces of recent items to become stronger than the STM traces of early items. In the absence of STM reverberation, only a recency effect is possible, since the earliest items will always have decayed in strength more than recent items. Because the Invariance Principle constrains inhibition to change the past field activities by a multiplicative factor, it argues for *shunting* (mass action) rather than *subtractive* inhibition.

In general, an STM pattern that mixes properties of Tables IIIA and IIIB can be expected. This can happen, for example, if normalization holds only partially. In Eq. (16), normalization holds because each $v_i^{(1)}$ in $\mathcal{F}^{(1)}$ excites only itself and inhibits *all* other $v_k^{(1)}$, $k \neq i$, with equal strength. Often the strength of recurrent excitation and inhibition from $v_i^{(1)}$ to other populations $v_k^{(1)}$ decreases with the interpopulation distance (cf. Elias and Grossberg, 1975; Levine and Grossberg, 1976). These distance-dependent connections form the anatomical substrate of the generalization gradients that join feature detectors together. If a given $v_i^{(1)}$ does not inhibit *all* other $v_k^{(1)}$, then certain $v_k^{(1)}$ can build excitation without interference from $v_i^{(1)}$, so that the normalization property is weakened. If partial normalization holds, then a bow in the STM pattern can coexist with a gradual increase in total STM activity as more items are presented. (Also recall the multiple STM equilibrium points in Fig. 30.) Furthermore, the bow does not, in general, have to occur at the second list item. By comparison, the statistical model of Hogan and Hogan (1975) for structural and transient components of memory describes data in which the transient memory distribution is bowed at list positions other than the second (see their Fig. 3).

We now solve a coding process that marries the Invariance Principle with the Partial Normalization Property to show that the bow can occur at any list position if system parameters are properly chosen. More precisely, consider codes for which:

1. The Invariance Principle holds.
2. The last item to be presented has an STM strength that is independent of list length.
3. The total STM activity grows in a negatively accelerated fashion from an initial value μ to a finite asymptotic value M that is independent of list length, but not necessarily of the stimulus materials from which a given class of lists is constructed.

Property 2 is a natural approximation because, no matter how many past items have occurred, it should still be possible to attend to the most recent item under normal stimulus conditions. Postulate 3 merely interpolates the simplest continuous curve between the initial and asymptotic values of total STM strength. Independent evidence that total STM strength has such a qualitative curve is found in studies of pupillary dilation in short-term memory tasks (Kahneman and Beatty, 1966). The properties of this code are described by the following theorem.

Theorem 1 (STM Buffer): Let all $\mu_i = \mu$. Let the total STM strength in a list of length i satisfy

Invariance Principle:

$$S_i = \mu \sum_{m=1}^i \prod_{r=m+1}^i \omega_r \quad (29)$$

as in Table II, as well as

Partial Normalization:

$$S_i = \lambda^{i-1}\mu + M(1 - \lambda^{i-1}) \quad (30)$$

for some λ such that $0 < \lambda < 1$, and $\mu < M$. Then letting $R = \mu^{-1}M$, for every $i > 1$,

$$\omega_i = \frac{\lambda^{i-1} + R(1 - \lambda^{i-1}) - 1}{\lambda^{i-2} + R(1 - \lambda^{i-2})} \quad (31)$$

and the STM strength in a list of length j of the k th item, $k < j$, is

$$x_{kj} = \mu \prod_{m=k+1}^j \omega_m \quad (32)$$

Every STM pattern $(x_{1j}, x_{2j}, x_{3j}, \dots, x_{jj})$ is either monotone decreasing, monotone increasing, or bowed. The longest list length J for which the STM pattern is monotone decreasing is given by the maximal j such that

$$(R - 1)(1 - \lambda)\lambda^{j-2} > 1 \quad (33)$$

In every list of length greater than J , the bow occurs at list position J .

The proof is found by equating (29) and (30) and solving for $\omega_2, \omega_3, \dots$, etc., by iteration. The theorem shows that a bow can occur at any list position J if R and λ are properly chosen. Also, given this code and prescribed stimulus materials that are homogeneous with respect to one another, the bow always occurs at the same list position, independ-

ent of list length. Moreover, since the last item always has strength μ , and the total STM strength never exceeds M , a strong recency effect develops as the list length becomes large. This result is generalized in Grossberg (1978b).

Do experimental operations that change the relative strengths of arousal and inhibition determine whether and where an STM bow occurs *in vivo*? Grossberg and Pepe (1971) show that variations in arousal level can change the list position where the bow occurs in network LTM traces (cf. Section 12). Theorem 1 suggests that arousal can change the list position at which the network's STM pattern bows by shunting inhibitory interaction strengths. Indeed, if some of the weights ω_i exceed 1, then STM at a given population can *increase* through time. This is a form of behavioral contrast due to lateral inhibition; cf. Grossberg (1975, Section 12), where peak shift and behavioral contrast of a generalization gradient also are explained by shunting inhibition. These network results suggest that performance variables, such as motivational or attentional state, can influence information processing constraints, such as primacy and recency.

28. Feedback Inhibition by Rehearsal in an Opaque STM Field

We now broach the question of how order information is read out of an STM pattern. To motivate the discussion, consider the task of repeating a telephone number that you have just heard. At no time are all the digits simultaneously rehearsed, and there can exist times during which no digit is consciously in mind. Moreover, the telephone number can be rehearsed at various rates, which can be controlled at will, within limits. At times when no digit is consciously available, the sequence of digits is *opaque* to the individual (Estes, 1972). Somehow the STM buffer organizes order information so that, when a rehearsal act perturbs the buffer, the correct item is elicited. Because the STM code is opaque, rehearsal *nonspecifically* activates all the possible item representations. The internal organization of activity patterns in the buffer codes the order information, and the nonspecific activation translates this activity into output signals. This use of arousal is analogous to its use in the avalanche of Fig. 12. There also arousal controls the readout of ordered signals.

Once a given item is rehearsed, the buffer must reset its activity so that the next item can be rehearsed; otherwise, the nonspecific rehearsal wave would cause an endless repetition of the first item. In some way, rehearsal of an item deletes its STM trace from the buffer (cf. Section 9).

The new STM activity pattern is then stored (or decays, etc.) until the next rehearsal wave perturbs it.

Consider the case in which order information is stored by a spatial pattern. Let each population v_i in a field \mathcal{F} code a certain command. The relative sizes of the STM activities x_i then determine in what order these commands will be elicited. Suppose that $x_1 > x_2 > x_3 > \dots > x_n$. Whenever a nonspecific rehearsal wave perturbs \mathcal{F} , all the populations begin to emit signals to the next processing stage. The most active population v_1 reaches its firing threshold soonest, so that it begins to fire earliest. It is likely that emergent signals can inhibit each other via feedforward on-center off-surround interactions. This would prevent leakage of signals corresponding to later list items. Using this mechanism, v_1 wins the lateral inhibitory competition, since its STM trace x_1 is largest. Thus v_1 fires its command signal first. As v_1 fires, it also activates a feedback inhibitory signal to its STM source. This feedback inhibition continues to act until it self-destructs by quenching suprathreshold STM activity at v_1 , but not necessarily subthreshold activity (Fig. 34). After v_1 is deactivated, the population v_2 has the largest STM activity. It can therefore fire signals through the feedforward on-center off-surround network, and it continues to do so until it self-destructs via feedback inhibition. The process continues until either all items are rehearsed, or arousal is terminated.

29. Transient Memory Span and Free Recall

Given the above rehearsal mechanisms, items that are presented earlier must have larger STM activities in order to be rehearsable in their

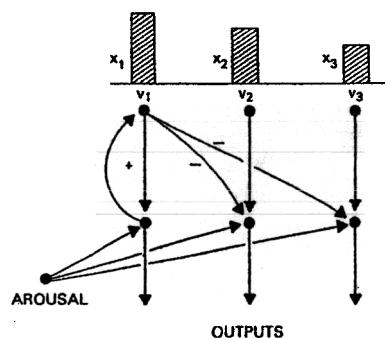


FIG. 34. STM traces compete before the maximal trace elicits a signal and self-inhibits via feedback inhibition. The process then repeats itself.

correct order. Below we first consider how this can happen if no LTM effects occur. That is, we consider only the primary effect of serial inputs on the STM pattern. By Sections 26 and 27, an earlier item can then have a larger STM activity only if inputs to the field are influenced by lateral inhibitory signals from previously stored items; without at least partial normalization, earlier items will always have smaller STM traces. In order to be reliably rehearsed in its correct order, after just one presentation, using only transient STM data, a list must be so short that it has not yet bowed. We define the *transient memory span* (TMS) to be the length of the longest list with a monotone decreasing transient STM pattern. This length is J in Theorem 1. The TMS can depend on the materials from which the list is constructed, since these will determine the strength of inputs and the distribution of lateral inhibition among items, as well as performance variables, such as the magnitude of shunting arousal.

If a list is so long that STM bowing takes place, then there will be a tendency to guess items near the beginning and the end of the list before guessing items in the middle, as is often observed in free recall experiments (Hogan and Hogan, 1975, Fig. 3). As ever longer lists are chosen, a strong recency effect develops, as Theorem 1 illustrates.

30. Parallel versus Serial Search in STM

The activation of order information using a nonspecific rehearsal wave is a parallel operation. In a normalized STM field, however, each item in the field—except perhaps the last—has a smaller STM trace if a longer list perturbs the field. Since an item's reaction time depends on its STM activity, as in the avalanche of Section 5, reaction time can vary with list length. This fact has often been used to support the idea that recognition memory is realized by a serial scanning process that exhaustively searches all stored items (Sternberg, 1966). Townsend (1974) has pointed out that the longer reaction times do not unambiguously implicate a serial process, by describing statistical parallel processing completion times that are indistinguishable from their serial processing counterparts. Whereas items are read out of the field in a prescribed order, the field operations that accomplish this are parallel rather than serial operations. Sections 60 and 61 discuss the questions of reaction time and (apparently) serial processing in greater detail.

31. The Influence of Rehearsal on Chunking

Given that rehearsal deletes an item from STM, it follows that rehearsal organizes which combinations of events will generate codes, or chunks. Naive experience also suggests that rehearsal should have such an effect, since rehearsing a particular combination of events defines that combination as a behaviorally meaningful unit.

Imagine that a series of events sequentially activates populations $v_{i_1}^{(1)}, v_{i_2}^{(1)}, \dots, v_{i_k}^{(1)}$ before these items are rehearsed. These events establish a spatial pattern of STM activity in $\mathcal{F}^{(1)}$. This spatial pattern generates $\mathcal{F}^{(1)} \rightarrow \mathcal{F}^{(2)}$ signals which begin to code a population $v_{i_1 i_2 \dots i_k}^{(2)}$ in $\mathcal{F}^{(2)}$. Simultaneously, $v_{i_1 i_2 \dots i_k}^{(2)}$ can send signals to $\mathcal{F}^{(1)}$ which learn the spatial pattern of activity across $v_{i_1}^{(1)}, v_{i_2}^{(1)}, \dots, v_{i_k}^{(1)}$. This pattern represents a sensory expectation, as shown in Section 20. What happens when a rehearsal wave perturbs $\mathcal{F}^{(1)}$? If $k < \text{TMS}$, then $x_{i_1} > x_{i_2} > \dots > x_{i_k}$, so that the individual items can be rehearsed in their proper order. Two remarks are pertinent here. First, by grouping items into sublists of length less than the TMS, it is possible to rehearse them in their correct order. Second, we must ask how chunk $v_{i_1 i_2 \dots i_k}^{(2)}$ learns to reproduce this order. Does the chunk activate order information only via its learned signals to $\mathcal{F}^{(1)}$ (sensory expectation), or can the chunk directly sample motor representations? The next sections will show that the latter alternative must hold. After the chunk $v_{i_1 i_2 \dots i_k}^{(2)}$ is adaptively coded and learns to reproduce the order of its coded sequence, it acts as a new functional unit of the network.

After all the items $v_{i_1}^{(1)}, v_{i_2}^{(1)}, \dots, v_{i_k}^{(1)}$ are rehearsed, suppose that another series of events occurs which serially activates the populations $v_{i_{k+1}}^{(1)}, v_{i_{k+2}}^{(1)}, \dots, v_{i_{k+m}}^{(1)}$. Since the first k items no longer reverberate in STM, signals from $\mathcal{F}^{(1)}$ to $\mathcal{F}^{(2)}$ will begin to adaptively code a chunk $v_{i_{k+1} i_{k+2} \dots i_{k+m}}^{(2)}$ that depends only on the second series of events. Simultaneously $v_{i_{k+1} i_{k+2} \dots i_{k+m}}^{(2)}$ begins to learn the sensory expectation that characterizes the second series of events. Suppose that $m < \text{TMS}$, so that the second series of events can be rehearsed in its correct order. As this happens, each item's STM representation is deleted and the chunk learns order information by sampling the items' motor representations.

Two important phenomena occur together in this scheme. Grouping items into lists of length less than the TMS allows them to be rehearsed in their correct order, and simultaneously defines a code for the list that is capable of learning to perform the list in its correct order. Time enters this mechanism in a subtle fashion. Rehearsal occurs *after* the time interval in which $v_{i_1 i_2 \dots i_k}^{(2)}$ is coded by its defining sequence and

prevents *future* events from being chunked with this sequence. However, the rehearsal act presents $v_{i_1 i_2 \dots i_k}^{(2)}$ with the data for sampling that enables it to control overt behavior on future trials.

32. Immediate Memory Span, and Readout of LTM Order Information by Feedback Signals from Commands

Section 27 shows that a bowed STM pattern can be elicited by serial inputs across a field $\mathcal{F}^{(1)}$ whose populations are joined by a recurrent on-center off-surround network. Section 12 shows that an STM bow can also be elicited by a combination of serial inputs and feedback signals. These feedback signals are generated by a field $\mathcal{F}^{(2)}$ and are gated on their way to $\mathcal{F}^{(1)}$ by LTM traces. Section 17 shows that the $\mathcal{F}^{(2)}$ populations that generate the feedback signals to $\mathcal{F}^{(1)}$ can be activated by STM patterns at $\mathcal{F}^{(1)}$. The adaptive coding process that accomplishes this also uses LTM traces. The process whereby event sequences at $\mathcal{F}^{(1)}$ are coded at $\mathcal{F}^{(2)}$, and codes at $\mathcal{F}^{(2)}$ learn order information at $\mathcal{F}^{(1)}$, is clearly a special type of adaptive resonance. The order information is the expectation, or template, of the resonance.

Before further analyzing this adaptive coding process, we can generalize the serial learning model of Section 12 by using the Invariance Principle. We want to see what kinds of feedback patterns from $\mathcal{F}^{(2)}$ to $\mathcal{F}^{(1)}$ can arise. In particular, under what circumstances does a population $v_i^{(2)}$ correctly code the order with which populations in $\mathcal{F}^{(1)}$ were excited? This code is carried by the LTM pattern $(Z_{i1}, Z_{i2}, \dots, Z_{in})$ of stimulus sampling probabilities Z_{ij} from $v_i^{(2)}$ to $v_j^{(1)}$.

To discuss this problem, we introduce some convenient nomenclature. Consider the LTM pattern $(Z_{i1}(t), Z_{i2}(t), \dots, Z_{in}(t))$ at any time t . Let $j_i(t)$ be the smallest j such that $Z_{ij}(t) > 0$, and let $J_i(t)$ be the largest j such that $Z_{ij}(t) > 0$. The integers $j_i(t)$ and $J_i(t)$ define the range of positive $\mathcal{F}^{(1)}$ activities that $v_i^{(2)}$ has sampled by time t . Restrict attention to the LTM pattern formed by $(Z_{ij}(t): j_i(t) \leq j \leq J_i(t))$. If the function $M(j; i, t) \equiv Z_{ij}(t)$, for fixed i and t , is monotone decreasing in j , for $j_i(t) \leq j \leq J_i(t)$, we say that $v_i^{(2)}$'s LTM pattern is *monotone decreasing* (Fig. 35a). If $M(j; i, t)$ has a single maximum, neither at $j_i(t)$ or at $J_i(t)$, we say that $v_i^{(2)}$'s LTM pattern is *unimodal* (Fig. 35b); if $M(j; i, t)$ has two local maxima, we say that $v_i^{(2)}$'s LTM pattern is *bimodal* (Fig. 35c). Population $v_i^{(2)}$ codes the order in which a list, or sublist, perturbs $\mathcal{F}^{(1)}$ only if its LTM pattern is monotone decreasing at all times after the list has been presented.

At any given time, these LTM patterns influence $\mathcal{F}^{(1)}$ by gating signals

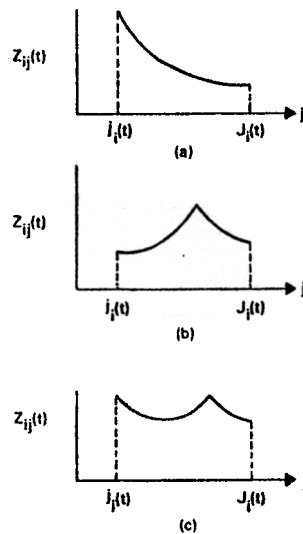


FIG. 35. (a) Monotone decreasing LTM pattern; (b) unimodal LTM pattern; (c) bimodal LTM pattern.

from $\mathcal{F}^{(2)}$. Let $F_j(t)$ be the total feedback signal from $\mathcal{F}^{(2)}$ to $v_j^{(1)}$ at time t . Signal $F_j(t)$ perturbs $v_j^{(1)}$ and thereupon influences the STM trace $x_j^{(1)}$ of $v_j^{(1)}$. It is possible for the pattern $(F_1(t), F_2(t), \dots, F_n(t))$ to code the order in which the list was presented, even though the LTM patterns of individual $v_i^{(2)}$ do not. That is, the total feedback pattern can be monotone decreasing, with $F_1(t) > F_2(t) > \dots > F_n(t)$, even though certain LTM patterns $(Z_{i1}, Z_{i2}, \dots, Z_{in})$ are not monotone decreasing. Each $v_i^{(2)}$ has a local view of the serial experiment at $\mathcal{F}^{(1)}$. The global synthesis of all these views is expressed by the signal pattern $(F_1(t), F_2(t), \dots, F_n(t))$, which controls behavior due to $\mathcal{F}^{(2)}$ at $\mathcal{F}^{(1)}$. Thus we shall ask what types of adaptive resonances code correct order information.

In Section 29, the concept of transient memory span (TMS) was introduced to discuss the effect of serial inputs on the STM pattern at $\mathcal{F}^{(1)}$. When feedback signals from $\mathcal{F}^{(2)}$ also perturb $\mathcal{F}^{(1)}$, they can change the list length at which the total STM pattern of $\mathcal{F}^{(1)}$ becomes bowed. We define the immediate memory span (IMS) to be the maximal list length at which the STM pattern at $\mathcal{F}^{(1)}$ is monotone decreasing when feedback signals from $\mathcal{F}^{(2)}$ are operative.

Our first results apply to the case wherein a list is presented once to the network, which thereupon tries to repeat it in its correct order. We

study the LTM patterns that develop in $\mathcal{F}^{(2)} \rightarrow \mathcal{F}^{(1)}$ synapses, but do not yet let these patterns influence STM at $\mathcal{F}^{(1)}$. The first result shows that the IMS can only be longer than the TMS, under weak conditions. In effect, feedback signals can only make it easier for lists to be repeated in their correct order.

Theorem 2 (Primacy). Let $\mathcal{F}^{(1)}$ obey the Invariance Principle. Also let the STM activities of populations in $\mathcal{F}^{(2)}$ decrease when new inputs perturb $\mathcal{F}^{(1)}$. Then the IMS is at least as long as the TMS, because every active $v_i^{(2)}$ codes a monotone decreasing LTM pattern after a list no longer than the TMS is presented to $\mathcal{F}^{(1)}$. In fact, this conclusion holds if the Invariance Principle is replaced by the weaker conditions

$$x_i^{(1)}[(i+m)\tau] \geq x_{i+k}^{(1)}[(i+k+m)\tau], \quad k > 0, \quad m \geq 0 \quad (34)$$

where τ is the intratrial interval.

If the list is longer than the TMS, then not all LTM patterns can be monotone decreasing. For example, in Fig. 35b a population $v_i^{(2)}$ that is excited after many list items have been presented will code more recent items more strongly than early items. The next result generalizes the statements of Section 12 about STM \rightarrow LTM order-reversal to the case where $\mathcal{F}^{(1)}$ obeys the Invariance Principle. After that, we shall suppose that $\mathcal{F}^{(2)}$ obeys the Invariance Principle to show how feedback signals can lengthen the IMS. Theorem 3 is conveniently stated in terms of the following definition.

Definition. The past (future) field of $v_i^{(2)}$ is the set of populations $v_j^{(1)}$ that are excited by inputs before (after) $v_i^{(2)}$ is excited.

Theorem 3 (Generalization Gradients). Let $\mathcal{F}^{(1)}$ obey the Invariance Principle. Let its invariant parameters $\mu_j^{(1)}$ and $\omega_k^{(1)}$ satisfy

$$\mu_1^{(1)} \geq \mu_2^{(1)} \geq \mu_3^{(1)} \geq \dots \quad (35)$$

and

$$\omega_2^{(1)} \geq \omega_3^{(1)} \geq \dots \quad (36)$$

Also let the STM activities of populations in $\mathcal{F}^{(2)}$ decrease when new inputs perturb $\mathcal{F}^{(1)}$. Let a list of length L serially perturb $\mathcal{F}^{(1)}$. If $L \leq \text{TMS}$, then every active $v_i^{(2)}$ population has a monotone decreasing LTM pattern. If $L > \text{TMS}$, then the LTM pattern of each $v_i^{(2)}$ is either monotone decreasing, unimodal, or bimodal. In all cases, each $v_i^{(2)}$ codes in LTM the ordering in its past field, and a monotone decreasing LTM pattern of its future field. In particular, the LTM pattern of $v_i^{(2)}$ can be monotone decreasing no matter how large L is, or how small the TMS is, if $v_i^{(2)}$ is excited before $v_2^{(1)}$ is excited (primacy). A unimodal pattern can be generated only if the past field of $v_i^{(2)}$ is monotone

increasing, either because there is no lateral inhibition within $\mathcal{F}^{(1)}$, or because the STM traces of the TMS items have become subthreshold. The maximum Z_{ij} then occurs at the $v_j^{(1)}$ that is excited most simultaneously with $v_i^{(2)}$. A bimodal pattern occurs if the past field of $v_i^{(2)}$ has a bowed STM pattern. Then one local maximum occurs at Z_{i1} , and the other occurs as in the unimodal case.

The above theorem makes no assumption about $\mathcal{F}^{(2)}$, except that its STM activities decrease when new inputs perturb $\mathcal{F}^{(1)}$, owing either to competitive inhibition by newly activated $\mathcal{F}^{(2)}$ populations, or to passive decay. It is natural to assume, moreover, that $\mathcal{F}^{(2)}$ obeys the invariance principle, if only because $\mathcal{F}^{(2)}$ can be a source of adaptively coded inputs to another field in a hierarchy of codes, and a stable coding of the past field in each level of the hierarchy must be guaranteed. Section 41 will show, however, that $\mathcal{F}^{(2)}$ cannot in general be constructed from a single, homogeneous pool of populations; several distinct population types can exist in a single field. In the simple case that only one population type exists, $\mathcal{F}^{(2)}$ possesses a single set of invariant parameters $\mu_j^{(2)}$, $j = 1, 2, \dots$, and $\omega_k^{(2)}$, $k = 2, 3, \dots$. In the next, more general case, each population type, or subfield, of $\mathcal{F}^{(2)}$ possesses its own set of invariant parameters. By defining invariant parameters for $\mathcal{F}^{(2)}$, we temporarily sidestep the adaptive coding problem. Instead of studying how $\mathcal{F}^{(2)}$ populations are activated by the i th serial input to $\mathcal{F}^{(1)}$, we demand that the i th population $v_i^{(2)}$ in $\mathcal{F}^{(2)}$ is excited according to the rule of Table II. Given this rule, we can study the total feedback pattern $(F_1(t), F_2(t), \dots, F_n(t))$ as it evolves through time. In particular, we can study how the feedback pattern tends to produce a primacy effect that balances the recency effect that is produced directly at $\mathcal{F}^{(1)}$ by a long list.

Whenever $\mathcal{F}^{(2)}$ has one population type that obeys the Invariance Principle, it also possesses a transient memory span (TMS_2). Denote the transient memory span of $\mathcal{F}^{(1)}$ by TMS_1 to avoid confusion. To understand how $\mathcal{F}^{(2)}$ can lengthen the IMS of $\mathcal{F}^{(1)}$, suppose that $\text{TMS}_2 \geq \text{TMS}_1$. This makes sense intuitively, because often the populations in $\mathcal{F}^{(2)}$ will represent commands that stay active for a long time in order to sample long sequences of events. Consequently the STM pattern in $\mathcal{F}^{(2)}$ can be monotone decreasing long after the STM pattern in $\mathcal{F}^{(1)}$ has bowed. In particular, the populations that are excited earlier in $\mathcal{F}^{(2)}$ will have greater STM activity than those excited later on. These early populations also tend to code a monotone decreasing LTM pattern, by Theorems 2 and 3. Thus, they sample a monotone decreasing STM pattern at $\mathcal{F}^{(1)}$ until the list exceeds TMS_1 . After that time, populations such as $v_1^{(2)}$ continue to sample a monotone decreasing LTM pattern. The later $\mathcal{F}^{(2)}$ populations have smaller STM activities, so they sample

the bowed STM activities at $\mathcal{F}^{(1)}$ less vigorously. Since $\mathcal{F}^{(2)}$ has a long TMS_2 , the STM signals from the early populations in $\mathcal{F}^{(2)}$ will be stronger than those from $\mathcal{F}^{(2)}$ populations that are excited later on. Hence the total feedback signal from $\mathcal{F}^{(2)}$ will more heavily weight the LTM patterns coded by the early $\mathcal{F}^{(2)}$ populations. This tends to make the $(F_1(t), F_2(t), \dots, F_n(t))$ pattern monotone decreasing (primacy effect). When these feedback signals act at $\mathcal{F}^{(1)}$, they tend to make the IMS longer than the TMS.

Another point of interest can be made here before it is developed in Section 61. The STM buffer at $\mathcal{F}^{(1)}$ can, in principle, store a list much longer than the TMS, albeit with STM activities that code incorrect order information. Given any search task that must be performed under time pressure before STM is reset, many of these items can be masked by items with larger STM activities.

33. A Minimal Model of Structural versus Transient Components of Memory

The above analysis suggests how feedback signals from $\mathcal{F}^{(2)}$ can induce a primacy effect at $\mathcal{F}^{(1)}$ even if presenting a long list to $\mathcal{F}^{(1)}$ tends to produce a recency effect. The total feedback pattern $(F_1(t), F_2(t), \dots, F_n(t))$ induces an STM response at $\mathcal{F}^{(1)}$ that can be called the *structural* component of memory. It is that part of the total input to $\mathcal{F}^{(1)}$ that is controlled by LTM, whether via $\mathcal{F}^{(1)} \rightarrow \mathcal{F}^{(2)}$ adaptive coding, or via $\mathcal{F}^{(2)} \rightarrow \mathcal{F}^{(1)}$ readout of order information. By contrast, the *transient* component of memory is the STM response at $\mathcal{F}^{(1)}$ due to serial inputs. Many papers have tried to understand the interplay of structural memory with transient memory (for example, Atkinson and Shiffrin, 1968; Estes, 1972; Hogan and Hogan, 1975). All these theories are weak in at least one respect. None of them gives an explicit description of how STM and LTM patterns are generated, coded, and mutually transform one another in real time. The present theory suggests a class of minimal models that is capable of approaching this task.

The simplest model discretizes and generalizes Eqs. (7) and (8) of Section 12. This generalization assumes that both $\mathcal{F}^{(1)}$ and $\mathcal{F}^{(2)}$ obey the Invariance Principle. By Theorem 1, the STM activities $x_j^{(1)}(k)$ and $x_i^{(2)}(k)$ of $v_j^{(1)}$ and $v_i^{(2)}$ at time $t = k$ satisfy

$$x_j^{(1)}(k) = \mu_j^{(1)} \prod_{m=j+1}^k \omega_m^{(1)} \quad (37)$$

and

$$x_i^{(2)}(k) = \mu_i^{(2)} \prod_{m=i+1}^k \omega_m^{(2)} \quad (38)$$

where $\prod_{m=i}^k \omega_m^{(p)} = 0$ if $i > k$, $p = 1, 2, \dots$. Consequently the LTM trace $z_{ij}(k)$ from $v_i^{(2)}$ to $v_j^{(1)}$ at time $t = k$ satisfies

$$z_{ij}(k) = z_0 + \sum_{m=1}^k f(x_i^{(2)}(m))x_j^{(1)}(m)D(k-m) \quad (39)$$

where the signal function $f(w)$ either is a sigmoid function of w , or describes a threshold cutoff at Γ ; $D(k-m)$ describes any LTM decay that might occur between time $t = m$ and $t = k$ (cf. Grossberg, 1974, Section IV); and $z_{ij}(0) = z_0$. The total feedback signal from $\mathcal{F}^{(2)}$ to $\mathcal{F}^{(1)}$ at time $t = k$ is

$$F_j(k) = \sum_{i=1}^N f(x_i^{(2)}(k))z_{ij}(k) \quad (40)$$

The pattern $(F_1(k), F_2(k), \dots, F_N(k))$, $k = 1, 2, \dots$, describes the effect of structural memory on $\mathcal{F}^{(1)}$ at any time $t = k$. The patterns generated at $\mathcal{F}^{(1)}$ according to Table II describe the transient memory through time. A weighted average of the two patterns describes the total STM pattern at $\mathcal{F}^{(1)}$. This latter pattern determines the order in which items will be rehearsed from $\mathcal{F}^{(1)}$ in response to a nonspecific rehearsal wave.

This total STM pattern replaces the probabilities of performance that are used in statistical learning models. In Section 25, we noted that, in the Atkinson-Shiffrin model, a computer analogy suggested a binary code for an item to be either in the STM buffer or not. Then item representations had to move through the buffer to remember order information. The binary code did not, however, meet STM order data. Somehow recency and primacy effects had to be generated. Recency gradients were generated by supposing that there exists a probability for the item to fall out of the buffer. Then the probability that an item is in the buffer decreases as a function of how long ago it entered the buffer (recency). Thus a hybrid mixture of binary and probabilistic concepts was wed together to achieve order information and a recency gradient. However, in each individual this hybrid scheme predicts that an item is either in the buffer with unit strength, or not in the buffer. The recency gradient is an intersubject construct. In the present theory, each item can be in the buffer of a single individual with variable STM activity,

and item motion is unnecessary to code order information. These two conceptions can be differentiated by experiments that test whether an item's ability to influence probe stimuli depends on its position in the buffer.

When a nonspecific rehearsal wave perturbs an STM buffer, the buffer's order information manifests itself. This order information is the net effect of all STM and LTM interactions among buffer item representations and higher-order commands. We now turn to the problem of globally synthesizing these interactions to achieve the order needed to perform prescribed tasks. Two general classes of task impose different requirements on the buffer. The task of prediction looks into the future. The task of imitation looks into the past. The task of naming stands somewhere between. By endeavoring to harmonize these demands, a deeper insight into the global structure of the field of command populations is achieved. The prediction task is particularly useful as a probe of this structure.

34. Prediction

Suppose that a pianist has learned to play a long series of chords. After having played several of the chords, how does the pianist know what chords come next? How does playing the previous chords generate commands capable of eliciting the future chords in their proper order? Consider Fig. 36. Suppose that the pianist learns the piece by playing one chord at a time. For the moment, let motor commands for playing each chord already be coded at $\mathcal{F}^{(4)}$. As each chord is played, it generates a sensory feedback pattern at the sensory field $\mathcal{F}^{(1)}$. This pattern is then adaptively coded and stored in STM by a population in $\mathcal{F}^{(2)}$, as in Sections 17 through 19. Intuitively, $\mathcal{F}^{(2)}$ codes spatial patterns of sensory data, or item information.

As a long sequence of chords is elicited by the motor commands $v_1^{(4)}, v_2^{(4)}, \dots, v_i^{(4)}$, their sensory feedback creates a spatial pattern across $\mathcal{F}^{(2)}$. By the Invariance Principle, once this spatial pattern is established, it does not subsequently change. It therefore activates a well-defined command population $v_{12\dots i}^{(3)}$ in $\mathcal{F}^{(3)}$ by adaptive coding. Intuitively, $\mathcal{F}^{(3)}$ codes temporal sequences of sensory data, or order information. Population $v_{12\dots i}^{(3)}$ thereupon reverberates in STM and begins to emit sampling signals via trainable pathways. The sampling signals to $\mathcal{F}^{(2)}$ learn a sensory expectation that codes order information. How is motor order information learned? Suppose for definiteness that $\mathcal{F}^{(3)}$ can sample $\mathcal{F}^{(4)}$. This will not ultimately be tenable, and the following argument shows

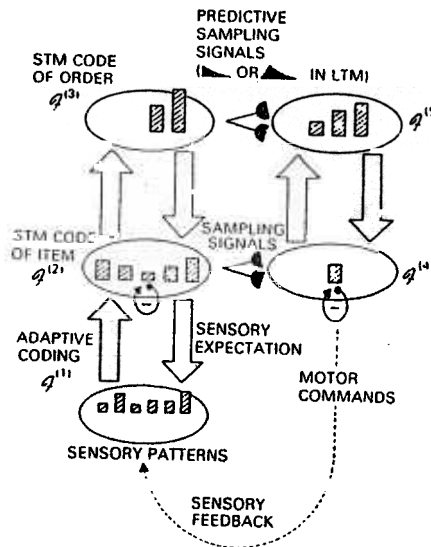


FIG. 36. Auditory feedback to $\mathcal{F}^{(1)}$ is coded at $\mathcal{F}^{(2)}$. Order information at $\mathcal{F}^{(2)}$ is coded at $\mathcal{F}^{(3)}$. Active populations in $\mathcal{F}^{(3)}$ can sample their future field of motor commands at $\mathcal{F}^{(4)}$.

why. Since rehearsal deletes motor representations after they are performed, population $v_{12...i}^{(3)}$ can sample all the future motor commands $v_{i+1}^{(4)}, v_{i+2}^{(4)}, \dots$ that are activated while it reverberates in STM. By contrast, $\mathcal{F}^{(3)}$ can also sample STM from past events in $\mathcal{F}^{(2)}$, since $\mathcal{F}^{(3)}$ is excited by sensory feedback owing to rehearsal at $\mathcal{F}^{(4)}$ (cf. Section 31). How does $v_{12...i}^{(3)}$ encode the order information of the motor representations $v_{i+1}^{(4)}, v_{i+2}^{(4)}, \dots$ in its LTM traces?

As in (2), the LTM trace from $v_{12...i}^{(3)}$ to $v_j^{(4)}$ time-averages the product of the signal from $v_{12...i}^{(3)}$ to $v_j^{(4)}$ with the STM trace of $v_j^{(4)}$. This signal is a monotone increasing function $f(w)$ of STM activity $w = x_{12...i}^{(3)}$ at $v_{12...i}^{(3)}$. As in Table II, the STM activity $x_{12...i}^{(3)}$ decreases monotonically through time after it is excited. This is due, for example, to lateral inhibition from other chunks in $\mathcal{F}^{(3)}$ that are activated as new chords are played. Hence the signal $f(x_{12...i}^{(3)})$ also decreases through time. Suppose, moreover, that $\mathcal{F}^{(4)}$ is normalized. Then each motor command at $\mathcal{F}^{(4)}$ has unit STM activity during its brief activation interval. The activation interval is brief because each motor command self-destructs its STM activity via feedback inhibition, as in Section 9. Consequently, the LTM trace from $v_{12...i}^{(3)}$ to $v_{i+1}^{(4)}$ is larger than the trace to $v_{i+2}^{(4)}$, the LTM trace to $v_{i+2}^{(4)}$ is larger than the trace to $v_{i+3}^{(4)}$, and so on. The LTM pattern from $v_{12...i}^{(3)}$ to the set $\{v_{i+1}^{(4)}, v_{i+2}^{(4)}, \dots\}$ is thus monotone decreasing. On a later per-

formance trial, after the chords $v_1^{(4)}, v_2^{(4)}, \dots, v_i^{(4)}$ are played, $v_{12...i}^{(3)}$ is activated by sensory feedback and thereupon elicits across $\mathcal{F}^{(4)}$ a monotone decreasing STM pattern that codes the correct order of future chord performance. The network is then ready to play a sequence of future chords. Turning on the motor rehearsal wave at $\mathcal{F}^{(4)}$ releases the chords in their correct order.

35. Sensory Feedback and Interference by Its Delay

The above mechanism works if sensory feedback due to motor performance does not excite other $\mathcal{F}^{(3)}$ populations and thereby change the total pattern of $\mathcal{F}^{(3)} \rightarrow \mathcal{F}^{(4)}$ signals. How is order information organized when sensory feedback continually excites new $\mathcal{F}^{(3)}$ populations? This question will lead to the conclusion that $\mathcal{F}^{(3)}$ samples $\mathcal{F}^{(5)}$ rather than $\mathcal{F}^{(4)}$, where $\mathcal{F}^{(5)}$ codes sequences of motor items just as $\mathcal{F}^{(3)}$ codes sequences of sensory items. To see what goes wrong if $\mathcal{F}^{(3)}$ samples $\mathcal{F}^{(4)}$, consider Fig. 37. Suppose that a given $\mathcal{F}^{(3)}$ population excites a monotone decreasing STM pattern at $\mathcal{F}^{(4)}$. Let motor arousal at $\mathcal{F}^{(4)}$ initiate performance of these items. The sequential motor perform-

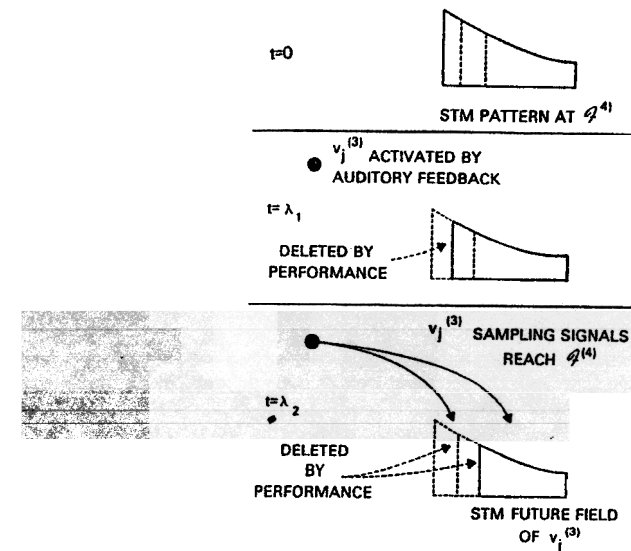


FIG. 37. $\mathcal{F}^{(3)}$ populations sample a monotone decreasing future field irrespective of how long it takes auditory feedback to activate them.

ance generates sensory feedback at $\mathcal{F}^{(2)}$, which is gradually coded by populations in $\mathcal{F}^{(3)}$. These $\mathcal{F}^{(3)}$ populations thereupon send sampling signals to $\mathcal{F}^{(4)}$. What patterns do these sampling signals see? Since motor performance at $\mathcal{F}^{(4)}$ deletes its motor command, they see only the activities of commands that have not yet been performed. The $\mathcal{F}^{(3)}$ populations therefore start to learn the monotone decreasing pattern that codes the correct order of the future items. If this trend continued through time, all would be well. It does not continue because, as each future item is rehearsed, its representation at $\mathcal{F}^{(4)}$ is deleted. As the $\mathcal{F}^{(3)}$ population continues to sample, it therefore tends to learn a monotone increasing pattern, rather than a monotone decreasing pattern, across its future field.

This difficulty arises only because the $\mathcal{F}^{(4)}$ representations are deleted when they are rehearsed. Somehow $\mathcal{F}^{(3)}$ must be able to sample representations that are linked to the motor commands in $\mathcal{F}^{(4)}$, but are not deleted when rehearsal occurs.

Let $\mathcal{F}^{(5)}$ code order information from $\mathcal{F}^{(4)}$. In particular, when an individual command $v_i^{(4)}$ is active at $\mathcal{F}^{(4)}$, it generates a code $v_i^{(5)}$ at $\mathcal{F}^{(5)}$, which in turn learns to activate $v_i^{(4)}$ via its feedback template (cf. Section 47). As a sequence of commands $v_1^{(4)}, v_2^{(4)}, \dots, v_i^{(4)}$ is rehearsed at $\mathcal{F}^{(4)}$, it elicits a monotone decreasing STM pattern across $v_1^{(5)}, v_2^{(5)}, \dots, v_i^{(5)}$. Then $v_{i+1}^{(3)}$ is turned on, and samples this STM pattern. Population $v_{i+1}^{(3)}$ also samples the monotone decreasing pattern that is elicited across its future field as the commands $v_{i+1}^{(5)}, v_{i+2}^{(5)}, \dots$ are sequentially rehearsed. By Theorem 3, the LTM traces from $v_{i+1}^{(3)}$ to $\mathcal{F}^{(5)}$ are unimodally distributed, and the maximal LTM trace abuts $v_{i+1}^{(5)}$.

Given this background, now let a sequence $v_1^{(3)}, v_{i+1}^{(3)}, \dots, v_{i+1}^{(3)}$ of $\mathcal{F}^{(3)}$ populations be activated by sensory feedback. The LTM pattern across each population's future field is monotone decreasing. In particular, all the LTM patterns that abut the populations $v_{i+1}^{(5)}, v_{i+2}^{(5)}, \dots$ are monotone decreasing. Consequently the total STM pattern due to sensory feedback is monotone decreasing across the unrehearsed $\mathcal{F}^{(5)}$ populations. This STM pattern elicits at $\mathcal{F}^{(4)}$ a monotone decreasing STM pattern across the unrehearsed populations $v_{i+1}^{(4)}, v_{i+2}^{(4)}, \dots$. The signals from $\mathcal{F}^{(5)}$ to the rehearsed populations $v_1^{(4)}, v_2^{(4)}, \dots, v_i^{(4)}$ do not generate large STM activities, because motor arousal at $\mathcal{F}^{(5)}$ bootstraps subliminal signals into STM by contrast-enhancing activities that are already large—namely, the unrehearsed activities—and quenching the small activities. The total STM pattern at $\mathcal{F}^{(4)}$ is therefore monotone decreasing across the unrehearsed items and negligible across the rehearsed items. The correct order of performance is hereby predicted, no matter how many $\mathcal{F}^{(3)}$ populations are activated by sensory feedback.

The above mechanism is robust because it does not depend on any particular rate of generating sensory feedback in response to motor performance. It requires only that the feedback delay on performance trials be the same feedback delay that is experienced during learning trials. If feedback is artificially delayed by electronic means, then the motor command elicited by the sensory feedback can interfere with proper motor performance.

The amount of delay will influence how severe the interference is. Suppose, for example, that when the feedback is not artificially delayed, it excites $v_i^{(4)}$ as the next command to be performed. Then the feedback enhances performance of the $v_i^{(4)}$ command. If feedback is slightly delayed, then it tends to prolong performance of $v_i^{(4)}$. Figure 38 shows how this happens. Before $v_i^{(4)}$ is performed, the STM pattern in $\mathcal{F}^{(4)}$ is given by Fig. 38a. Suppose that performance of $v_i^{(4)}$ begins, and with it feedback inhibition of activity at $v_i^{(4)}$ (Fig. 38b). Once the activity at $v_i^{(4)}$ is less than that at $v_{i+1}^{(4)}$, performance at $v_{i+1}^{(4)}$ can begin. Suppose that the delay in sensory feedback to $v_i^{(4)}$ is so brief that feedback re-excites $v_i^{(4)}$ just before this can happen (Fig. 38c). Then $v_i^{(4)}$'s activity stays maximal for a longer time than usual and prolongs performance of the $v_i^{(4)}$ chunk.

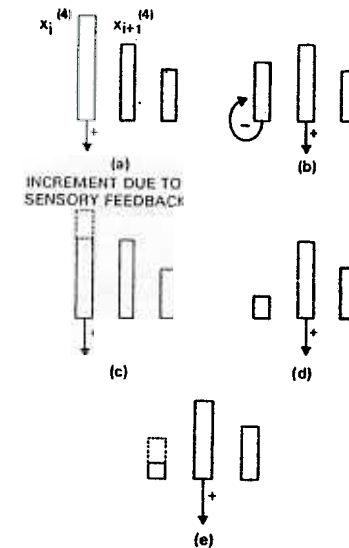


FIG. 38. Sensory feedback in (c) suffices to prolong performance at $v_i^{(4)}$, but in (e) $v_{i+1}^{(4)}$ continues to control performance.

If, however, self-inhibition of $v_i^{(4)}$ is almost complete when sensory feedback to $v_i^{(4)}$ arrives (Fig. 38d), then $v_{i+1}^{(4)}$ still has the maximal activity after the feedback signals take effect (Fig. 38e), so that the interference due to a longer delay is smaller. This phenomenon is analogous to the Lee effect, wherein auditory feedback causes maximal interference with speech production—primarily a drawing out of vowels—if it is delayed by about 180 msec, but less interference at smaller or later delays (Lenneberg, 1967, p. 109).

Other possible effects of delayed feedback go beyond the scope of the present discussion, because they all involve rhythmic activity in the network. As Section 23 notes, periodic oscillations in network activity can occur in recurrent networks due to the lagging behind of inhibitory activities relative to their excitatory counterparts (Ellias and Grossberg, 1975; Grossberg, 1977a). Whenever this happens, different effects can be elicited by a delay of sensory feedback that coincides with the waxing rather than the waning of STM activity. If we interpret the rhythmic STM activity as an analog of brain waves, then different effects can occur when feedback is in or out of phase with brain activity. The following effect is of this type. Let a sensory feedback cue reach $\mathcal{F}^{(1)}$ when the feedback expectation from $\mathcal{F}^{(2)}$ to $\mathcal{F}^{(1)}$ is waxing. Then the nonspecific alarm is not set off if the feedback is compatible with the expectation. By contrast, if the sensory feedback cue reaches $\mathcal{F}^{(1)}$ when the expectation from $\mathcal{F}^{(2)}$ is waning, then a tendency to reset STM at $\mathcal{F}^{(2)}$ is at least momentarily elicited, even if the feedback is compatible with the expectation. Thus a mismatch of activation *phases* can cause interference in sensory coding whether or not the expectations and the feedback cues are compatible. A similar remark can be made about the reset of motor coding when proprioceptive feedback is delayed, or about the rate with which terminal motor maps are reset by new motor commands. This fact is compatible with data concerning the rhythmicity of language production (Lenneberg, 1967, Chapter V), in particular with the idea that speech is broken up into “syllables” or “breath pulses” that have a rhythmicity of 6 or 7 cycles per second. A proper understanding of these phase versus informational phenomena requires a study of how feedback cues can reset the endogenous STM rhythm. Even without such a study, the above remarks show the importance of distinguishing whether interference is due primarily to phase or to informational mismatch in particular cases. For example, in Fig. 38e, a long feedback delay prevents $v_i^{(4)}$ from receiving sufficient feedback to dominate $v_{i+1}^{(4)}$. Can this be prevented by appropriate amplification of the feedback signal, or is the reduction in interference due primarily to a

matching of phases? Testing this alternative unambiguously will be difficult, if only because amplification of the feedback signal might also reset the phase of the STM rhythm.

The above mechanisms work well as far as they go. They clarify some points, but also raise new questions. On the clarification side, they begin to show why sensory feedback is not needed to play the *next* chord, but is nonetheless important. After a command such as $v_{12\dots i}^{(3)}$ is activated, it can predict the correct performance order of a long sequence of chords $v_{i+1}^{(4)}, v_{i+2}^{(4)}, \dots$. Thus, sensory feedback from $v_{i+1}^{(4)}$ is not needed to perform $v_{i+2}^{(4)}$. Nonetheless, sensory feedback does determine which commands, such as $v_{12\dots i}^{(3)}$, will be activated. A sequence from the past hereby determines a sequence in the future.

On the question side, we note the following:

1. Population $v_{12\dots i}^{(3)}$ is not the only active chunk in $\mathcal{F}^{(3)}$. Every active subsequence across $\mathcal{F}^{(2)}$ can, in principle, be coded at $\mathcal{F}^{(3)}$. A spatial pattern of activity therefore exists across $\mathcal{F}^{(3)}$ at every time. How should the relative activities of these chunks be determined so that the *total* signal from $\mathcal{F}^{(3)}$ to $\mathcal{F}^{(4)}$ correctly codes the order information of future chords?
2. What advantages are gained by using higher-order chunks like $v_{12\dots i}^{(3)}$ rather than $v_i^{(3)}$? The LTM traces of chunk $v_i^{(3)}$ can also encode the future commands $v_{i+1}^{(4)}, v_{i+2}^{(4)}, \dots$ in their proper order. Why bother using higher-order commands at all?

36. Greater Weight and Longer Duration of Higher-Order Chunks

These questions suggest a general principle with far-reaching consequences. Consider question (2) to start off. Chunk $v_i^{(3)}$ can be activated by *every* chord sequence in which chord $v_i^{(4)}$ is played. Chunk $v_i^{(3)}$ might therefore already code a different chord sequence at $\mathcal{F}^{(4)}$ before the new piece is learned. By contrast, chunk $v_{12\dots i}^{(3)}$ is better defined by the piece that the pianist is playing, and can therefore more reliably predict the correct sequence of future chords. For example, if I hear the word “C” (= “see”), I can reply “See what?”; but if I hear “ABC,” it is much more likely that I will reply “D.” Higher-order chunks should therefore have greater STM activity than lower-order chunks in $\mathcal{F}^{(3)}$.

There is a related reason for giving higher-order chunks greater weight. It is desirable to keep the chunks that code longer sequences active for a longer time, so that they can predict far into the future. Such chunks can dramatically compress the encoding of data by each predict-

ing a long sequence of chords. In particular, once a chunk in $\mathcal{F}^{(3)}$ enters STM, it can be actively quenched only by lateral inhibition due to later activation of other chunks in $\mathcal{F}^{(3)}$, assuming that arousal level and other performance variables stay fixed. Shutting off the $\mathcal{F}^{(2)}$ populations that originally activated an $\mathcal{F}^{(3)}$ chunk does not inhibit this chunk. This property is similar to the maintenance of STM activity by populations in $\mathcal{F}^{(2)}$ after their defining input patterns at $\mathcal{F}^{(1)}$ are deleted by rehearsal and replaced by new input patterns. These constraints are summarized by the following rule.

Self-Similar Coding Rule: Other things being equal, higher-order chunks have greater STM activity and longer duration than lower-order chunks.

37. Spatio-temporal Self-Similarity and the Resolution of Uncertainty

The above idea is a special case of the principle of *spatio-temporal self-similarity*, or STSS (Grossberg, 1969e). The network takes a risk by allowing any chunk to remain active for a long time. What if the chunk codes erroneous information? It can then cause errors until its activity can finally be quenched. The risk is minimized by letting the highest-order chunks remain active the longest. Since these chunks are better characterized by the temporal sequence in which they are embedded, it is more likely that this output will be the correct one. Thus, in the present situation, STSS means that chunks that are coded by long sequences can remain active for a long time to sample and control performance of long sequences.

How is STSS physically realized? The STSS concept constrains the *global* rules for building a field; what are the *local* rules whereby its individual populations are constructed? A mechanism is suggested by considering the concept of an *STSS cell type*. Suppose, statistically speaking, that a given cell type has a characteristic shape, including dendritic tree, cell body, and axons. If the cell type is STSS, then a small cell of this type is transformed into a large cell by blowing up all its spatial dimensions. What effects does this scale change have on a cell's functional capabilities? We now discuss certain neocortical cells using the teleology of STSS. Consider a cortical Betz cell that has a large dendritic tree (Crosby *et al.*, 1962). Such a dendritic tree gets a good view of input data in the vicinity of the cell. The firing of such a cell can thus be precisely controlled by prescribed patterns of sensory inputs. Because the cell fires only at (essentially) the correct times, it can be permitted to send a long axon all the way down to spinal motor centers

(pyramidal tract), where it directly activates motor output. By contrast, if a cortical pyramidal cell has a small dendritic tree, then it can be fired by a much wider variety of input data. Its axon cannot be allowed to be a final command pathway to motor centers. By giving such cells a shorter axon, their effects on motor outputs will be manifested only over multisynaptic pathways (extrapyramidal tract), wherein further computations can be undertaken. In a similar fashion, if longer axons are also wider, then they can carry their signals more quickly, and thereby achieve the same transmission time lag as shorter and thinner axons. STSS is thus a structural constraint on individual cells that helps the aggregate network dynamics to resolve uncertain input data without taking untoward risks.

38. Order-Preservation in the Future Field of Motor Commands

The principle of SSTS suggests that higher-order chunks should be coded either by more cell sites—either more or bigger cells in a population—and/or by stronger signal pathways than lower-order chunks. This design will lead to fields \mathcal{F} whose anatomy is in a dynamic equilibrium with the average spatial distribution of inputs that perturb \mathcal{F} through time.

Before making this construction, we note that predicting the order of future motor commands could be achieved under weak constraints on $\mathcal{F}^{(4)}$, were it not for the fact that low-order chunks can sample many different chord sequences. This was demonstrated in Section 34, wherein it was proved that every chunk in $\mathcal{F}^{(3)}$ codes a monotone decreasing pattern across its future field in $\mathcal{F}^{(4)}$, independent of what items or sequences of items are coded by the chunk. Thus $\mathcal{F}^{(3)}$ can elicit the correct performance order at $\mathcal{F}^{(4)}$, no matter how many classes of chunks sample $\mathcal{F}^{(4)}$. We wish to suppress lower-order chunks only because they can sample and perform too many incompatible sequences.

39. Masking of STM by More Cell Sites or Amplified Signals

What STM patterns can exist in networks wherein some populations have more cell sites, stronger signals, or broader tuning curves than other populations? It is shown below that such networks have exactly the properties needed to satisfy STSS. Examples of such networks have been studied by Grossberg and Levine (1975) and Levine and Grossberg (1976). The former paper considers recurrent on-center off-surround

networks of the form

$$\dot{x}_i = -Ax_i + (B_i - x_i)[f(x_i) + I_i] - x_i \left[\sum_{k \neq i} f(x_k) + J_i \right] \quad (41)$$

where the populations v_i can have different total numbers B_i of cell sites. System (41) is formally equivalent to

$$\dot{u}_i = -Au_i + (C - u_i)[f(D_i u_i) + I_i] - u_i \left[\sum_{k \neq i} f(D_k u_k) + J_i \right] \quad (42)$$

where both the excitatory and inhibitory signals $f(D_i u_i)$ from each population v_i are amplified by a scaling factor D_i . System (41) is transformed into (42) by the substitutions $x_i = D_i u_i$ and $B_i = C D_i$. A system

$$\dot{w}_i = -Aw_i + (C_i - w_i)[f(D_i w_i) + I_i] - w_i \left[\sum_{k \neq i} f(D_k w_k) + J_i \right] \quad (43)$$

in which both types of asymmetry exist is thus equivalent to system (41) with $x_i = D_i w_i$ and $B_i = C_i D_i$. In particular, if $C_1 \leq C_2 \leq \dots \leq C_n$ and $D_1 \leq D_2 \leq \dots \leq D_n$, then $B_1 \leq B_2 \leq \dots \leq B_n$.

Such asymmetries introduce a new type of contrast enhancement into the system. For example, in (41), populations v_i with the largest B_i values tend to quench, or mask, the STM activity in populations v_j with smaller B_j values. If higher-order chunks are given larger B_i values in (41), then they will mask lower-order chunks, as we desire by STSS. This is the main idea.

More precisely, suppose that $B_1 \leq B_2 \leq \dots \leq B_n$ in (41). If $f(w) = Ew$, then $x_i(\infty) = 0$ if $B_i < B_n$, and $x_i(\infty) = Kx_i(0)$ if $B_i = B_n$. In other words, given a linear signal function, all populations with nonmaximal B_i are masked, and the STM pattern of all populations with maximal B_i (that is, $B_i = B_n$) is stored faithfully in STM. No states are masked if all $B_i = B_n$. If $f(w)$ is, more realistically, chosen to be a sigmoid signal function, then an interesting phenomenon occurs. Once again there is the tendency for populations v_i with maximal B_i to mask other populations. In particular, if some populations v_i with $B_i = B_n$ get relatively large inputs, then all states v_j with $B_j < B_n$ will be masked. In general, however, there is a competition between the relative sizes of the B_i 's and the relative sizes of the initial activities $x_i(0)$, the latter in turn being determined by the relative sizes of inputs to v_i before time $t = 0$. In all cases, only the STM traces corresponding to one B_i can be stored in STM. If certain v_i with $B_i < B_n$ have sufficiently large $x_i(0)$ compared with the $x_j(0)$ values of all v_j with $B_j = B_n$, then the subfield of populations with the nonmaximal weight B_i can mask all other populations. The STM pattern of this

subfield is simultaneously contrast-enhanced and stored in STM. Grossberg and Levine (1975) interpret the competition between B_i and $x_i(0)$ in terms of developmental and attentional biases in the field. A developmental bias can, for example, give certain feature detectors larger B_i values than others. An attentional shunt can amplify the signals of one subfield more than others via larger D_i values. Either operation biases the field in favor of some subfield. In (43), the developmental biases C_i and the attentional biases D_i can create a complicated tug of war that favors the particular subfield having maximal $B_i = C_i D_i$ for STM storage. Nonetheless, a population with nonmaximal B_i can be stored if its features are present in the input display with relatively large saliency, or are coded by relatively strong pathways that amplify its inputs.

40. STM Drift toward a Norm: Primary Gradient Induces Secondary Gradient

The tendency of populations with maximal B_i to *totally* mask all other populations is due to the fact that each v_i can inhibit *all* v_k , $k \neq i$, with equal strength in (41). When the strength of recurrent excitatory and inhibitory signals decreases as a function of interpopulation distance, as in Section 27, then the masking effect can be partial, and can generate a slow drift by the spatial locus of maximal STM activity toward the populations having the largest B_i values. In the case that each v_i codes particular features, then the falloff with distance of recurrent signals defines generalization gradients between the feature detectors, and the detectors with the largest B_i act as "norms" toward which activity drifts across these generalization gradients.

For example, Levine and Grossberg (1976) study networks of the form

$$\dot{x}_i = -Ax_i + (B_i - x_i) \left[\sum_{k=1}^n f(x_k) C_{ki} + I_i \right] - (x_i + D) \left[\sum_{k=1}^n f(x_k) E_{ki} + J_i \right] \quad (44)$$

where the excitatory coefficients C_{ki} and inhibitory coefficients E_{ki} both decrease as a function of interpopulation distance $|i - k|$, with excitation ("on-center") decreasing faster than inhibition ("off-surround"). Suppose in addition that the B_i 's are normally distributed around a given population v_i ; that is, $B_i = B e^{-\lambda|i-I|^2}$. Then if an input perturbs a population v_i , $i \neq I$, the locus of maximal STM activity drifts toward v_I . The drift rate depends on how steep the slope of the function B_k is as a function of v_k for k values between i and I . If the slope is small, the drift

rate is slow; if the slope is large, the drift rate is fast. Levine and Grossberg (1976) suggest that such a drift is responsible for the line neutralization effect that is perceived when a nearly vertical or horizontal line is inspected for a sufficiently long time (Gibson, 1933). In summary, if there exist generalization gradients among feature detectors, and if certain detectors are coded by more sites or broader tuning curves than other detectors, then STM activity drifts toward the nearest populations having the most sites or the most highly amplified signals.

The above mechanism can be described in terminology from developmental biology. The B_i 's define a *primary* gradient; the input is an inducing stimulus; and the STM drift is a *secondary* gradient that is generated by field interactions in response to the inducing stimulus.

41. Masking of Lower-Order Codes

Which chunks in $\mathcal{F}^{(3)}$ will be masked? The answer depends on at least three factors. First, it depends on the spatial distribution of LTM traces across the $\mathcal{F}^{(2)} \rightarrow \mathcal{F}^{(3)}$ pathways. The LTM vectors z_i in these pathways define positional gradients that determine how close together in $\mathcal{F}^{(3)}$ are the populations that code two different sequences of events; for example, how close is the $\mathcal{F}^{(3)}$ code for sequence ABC to the code for ABCD? Second, it depends on how many sites code each population. If ABCD is coded by more sites than ABC, then ABCD's code will tend to mask ABC's code. Third, it depends on the breadth of recurrent excitatory and inhibitory signals within $\mathcal{F}^{(3)}$. Even if the codes for ABC and ABCD lie next to each other in $\mathcal{F}^{(3)}$, they are far away from each other, functionally speaking, if they do not fall within each other's generalization gradients. Section 39 shows that more sites and broader tuning curves have the same effect on STM masking. Hence we expect the chunks that have the most sites to have the broadest generalization gradients. This is immediately guaranteed, given STSS, if the largest cells are in the populations having the most sites.

It remains to determine how many sites will be given to $\mathcal{F}^{(3)}$ populations that code sequences of prescribed length. We seek a law that can plausibly be realized by simple rules of neuronal growth before the stage of adaptive coding takes place. The qualitative features of this law are already apparent. For example, given sequences A, AB, ABC, ABCD, . . . , it follows by STSS that the number of sites should increase monotonically with list length until a maximal length is reached. Thereafter, the number of sites should decrease with list length to prevent infinitely long (and infinitely unlikely) sequences from being

coded (see Fig. 39). The simplest rule of this type is the Poisson distribution. Suppose that a population in $\mathcal{F}^{(2)}$ contacts a certain population in $\mathcal{F}^{(3)}$ with a prescribed small probability p . Let λ be the mean number of such contacts on all the cells of $\mathcal{F}^{(3)}$. Then the probability that exactly k contacts perturb a given population is

$$P_k = (\lambda^k / k!) e^{-\lambda} \quad (45)$$

(Parzen, 1960). If K is chosen so that $K < \lambda < K + 1$, then P_k increases for $1 \leq k \leq K$ and decreases for $k > K$. More sophisticated but related distributions, such as the hypergeometric distribution

$$P_{mk} = \binom{m}{k} \binom{N-m}{n-k} / \binom{N}{n} \quad (46)$$

where $\binom{N}{n}$ is the binomial coefficient, are also discussed in Parzen (1960).

Consider the Poisson distribution in (45) for definiteness. Given this rule, sequences of length K will generate maximal STM activity at $\mathcal{F}^{(3)}$, other things being equal. For example, suppose that $K = 4$, and consider network response to the sequence of events ABCD As each item is processed, it excites a code (A); then codes (B), (AB), and (BA); then codes (C), (ABC), (BAC), (BC), . . . ; and so on. How close are the codes (A), (B), and (AB) to each other in $\mathcal{F}^{(3)}$? The code (AB) differs from the codes (A) and (B) separately only by one item. Since the items A and B are coded by being *projected* onto the LTM traces z_i in $\mathcal{F}^{(2)} \rightarrow \mathcal{F}^{(3)}$ pathways, (A) and (B) are closer to (AB) than, say, (F) is. This argument can be refined by taking into account the phonetic similarities that cause items to be similarly coded at the $\mathcal{F}^{(1)} \rightarrow \mathcal{F}^{(2)}$ stage of filtering; such extensions can be supplied once the main idea is clear. In summary, (A) and (B) lie close to (AB) in $\mathcal{F}^{(3)}$. Since (AB) has greater weight, it tends to mask (A) and (B) by lateral inhibition. Similarly, (ABC) masks (AB), and (ABCD) masks (ABC). Since $K = 4$, (ABCD) also masks (ABCDE), (ABCDEF) and so on. Thus, chunks of length four tend to dominate the STM activity in $\mathcal{F}^{(3)}$.

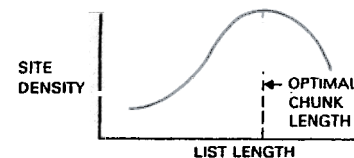


FIG. 39. STSS suggests that longer sequences are coded by more sites, up to some maximal sequence length.

42. Clustering, Compression, Spacing, and Completion

The above masking process has many interesting implications. Most obviously, it dramatically reduces, or compresses, the number of active chunks that are needed to control motor activity by suppressing the populations that code nondominant subsequences. It is important to recall here that the codes for A, AB, ABC, etc., are not wired into the network. Thus if a population for ABCD masks many of the populations that *would be* codes for BCD, CDB, CD, etc., before these codes can be learned, then the masked populations remain uncommitted and can be coded by other events; in particular, by events in which the items A, B, and C occur separately or in other dominant subsequences.

A second interesting consequence is a spacing effect in the chunks that get stored in $\mathcal{F}^{(3)}$. After the events A, B, C, and D are presented, (ABCD) actively masks the populations (BC), (CD), (BCD), etc. When event E occurs, there is a tendency for (BCDE) to be dominant, but this tendency is offset by two factors: First, all the populations related to (BCDE), such as (CD), which could supply it with recurrent excitation, have been suppressed; second, (ABCD) is sufficiently close to (BCDE) to suppress it, at least partially, by recurrent inhibition. This argument must be made with care, because it depends on the detailed choice of network parameters. It is, for example, possible for (BCDE) to be in the on-center of (ABCD), whence it is enhanced rather than suppressed, whereas (CDEF) is in the off-surround of (ABCD). The main robust point is clear, however. There is a tendency for the codes of dominant subsequences to be *spaced*, owing to mutual recurrent inhibition by their off-surrounds. For example, only the codes (ABCD), (DEFG), (GHIJ), . . . might have significant STM activity after recurrent inhibition acts. This spacing effect accomplishes a further compression of data encoding by $\mathcal{F}^{(3)}$.

The spacing effect generates a tendency to cluster responses into subsequences in order to maximize learning rate. As a sequence of events is presented to the network, it tends to generate spaced chunks of a dominant length. When these chunks control performance, they tend to group responses into the functional units coded by these chunks. Suppose, for example, that $K = 4$. Then if four consecutive items are listened to before rehearsing them, these items will be coded by a chunk of maximal weight. A longer sequence will tend to be broken into two or more subsequences because of the spacing between dominant chunks within $\mathcal{F}^{(3)}$. Feedback from dominant chunks in $\mathcal{F}^{(3)}$ to $\mathcal{F}^{(2)}$ and $\mathcal{F}^{(4)}$ thereupon tends to group items into rehearsal units of length four.

Does the length of the TMS at $\mathcal{F}^{(2)}$ influence the number K at $\mathcal{F}^{(3)}$ by

determining the maximal length of sequences that can be rehearsed from the STM buffer with no LTM feedback? As Section 32 notes, the TMS is shorter than the IMS. Is the TMS $\cong 4$ and the IMS $\cong 7$ in most individuals?

In a similar fashion, suppose that free responding is controlled by dominant chunks in $\mathcal{F}^{(4)}$. Suppose that item A has just been elicited by chunk (ABCD). Since chunk (ABCD) is soon suppressed by nearby chunks via recurrent inhibition, there is a tendency not to say A again until after (ABCD) is released from inhibition. This creates an apparent refractoriness for emitting the same item again while its controlling chunk is suppressed by chunks that are activated by sensory feedback owing to performance of later items.

Estes (1972) suggests a coding theory in which inhibition is used to control the clustering effect. His theory does not, however, study the dynamics of coding, or the real-time mechanisms whereby the inhibition organizes itself across emergent codes. The above remarks also use inhibition, but provide a different and more complete theory of how inhibition works. Estes' paper also reviews various data that are related to the above mechanisms.

Finally, the masking mechanism provides a deeper insight into pattern completion. For example, let a spatial pattern at $\mathcal{F}^{(2)}$ be coded by a nondominant population in $\mathcal{F}^{(3)}$. The STM activity in $\mathcal{F}^{(3)}$ can then drift toward the nearest dominant population. The drift "completes" the pattern, or in the present case, the sequence, by activating a higher-order code. Then the dominant population can send its template of feedback signals back to $\mathcal{F}^{(2)}$ where they reorganize STM at $\mathcal{F}^{(2)}$ to code the "completed" pattern.

43. The Magic Number Seven and Self-Similar Coding

The coding of sequences by patterned activity across spaced dominant chunks accomplishes several tasks at once. One is to suppress predictive sampling by lower-order chunks that are excited by sensory feedback during motor performance. The discussion in Sections 36 through 41 shows that this mechanism is a consequence of STSS.

STSS also implies that the duration of higher-order chunks exceeds that of lower-order chunks. After recurrent inhibition causes spacing within the field of dominant chunks, it creates a TMS among the spaced dominant chunks. If a self-similar scaling of recurrent interaction strengths holds within every subfield of chunks, then the TMS of the spaced dominant chunks will be commensurate with that of the chunks

(A), (B), (C), Then the chunk for (ABCD) can remain active much longer than the chunk for (D), even if the recurrent interactions within each subfield of spaced chunks have the same effect on STM activity when a population in their subfield is excited. This is true because, as the length of the subsequences that are coded by a subfield increases, so does the time interval between successive activations of its spaced chunks.

By STSS, each subfield of chunks in a prescribed sensory field has the same TMS, *other things being equal*. Otherwise expressed, if sensory data are recoded by a different subfield, then they have the same TMS in the new code as they had in the old code. The existence of a commonly shared "magic number seven, plus or minus two," for the immediate memory span of various codes (Miller, 1956) thus supplies indirect support for STSS as a principle of code synthesis.

44. Suppression of Uniform Patterns and Edge Detection

The principle of STSS provides one mechanism whereby erroneous signals from lower-order chunks can be eliminated: Lower-order chunks, and their signals, are inhibited by higher-order chunks. Section 36 suggested one reason for doing this: If a given event occurs in different contexts, it should be able to elicit different responses. In particular, if a lower-order chunk, such as $v_i^{(3)}$, were allowed to vigorously sample every chord $v_j^{(5)}$ that succeeded it, then eventually $v_i^{(3)}$ could encode a nearly uniform pattern of activity across its LTM traces. Signals from $v_i^{(3)}$ to $\mathcal{F}^{(5)}$ would then be uninformative, or irrelevant, since they would not discriminate any population in $\mathcal{F}^{(5)}$ from any other. STSS helps to prevent this by restricting the circumstances under which sampling can occur.

Often cues should be allowed to sample even if their signals are uninformative. For example, before a chunk learns a pattern, its signals to $\mathcal{F}^{(5)}$ are uniformly distributed. If the chunk is not allowed to sample $\mathcal{F}^{(5)}$, then it can never learn a pattern. The problem is to allow uninformative cues to sample $\mathcal{F}^{(5)}$ under appropriate circumstances, and yet to prevent their uniform, or "noisy," signals from destroying the patterns that are driving $\mathcal{F}^{(5)}$. Whether a given chunk controls a uniform pattern or not in its LTM traces can be decided only after these traces elicit signals at $\mathcal{F}^{(5)}$. Moreover, even if each active chunk codes a nonuniform pattern, the total input pattern to $\mathcal{F}^{(5)}$ can be uniform, and therefore uninformative. Hence some mechanism within $\mathcal{F}^{(5)}$ must exist to deal with the noise in its total input pattern.

We now show how recurrent mass action networks suppress the "uniform part," or noise, in their total input patterns, and generate suprathreshold responses only to spatial differences, or "discriminations," in these patterns. This property implies that LTM sampling occurs among chunks when the sampling paths carry useful information, but not otherwise.

The simplest version of the uniform quenching property occurs in system

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

with $-C \leq x_i \leq B$. If $C > 0$, by contrast with (14), the equilibrium response ($\dot{x}_i = 0$) to a pattern $I_i = \theta_i I$ is

$$x_i = \frac{(B + C)I}{A + I} \left(\theta_i - \frac{C}{B + C} \right) \quad (47)$$

If, for example, $B = (n - 1)C$, then $C/(B + C) = 1/n$. Now let the input pattern be uniform. Then all $\theta_i = 1/n$, so that no matter how intense I is, all $x_i = 0$. If not all $\theta_i = 1/n$, then the network quenches the "uniform part" of the pattern. More generally, whenever $B \leq (n - 1)C$, the x_i 's are suppressed even more vigorously by inhibition than when $B = (n - 1)C$. Consequently only values of $\theta_i > C/(B + C) > 1/n$ can generate a supraequilibrium response. Increasing C hereby contrast-enhances the network's response to input patterns. It has been suggested that this contrast-enhancement property can influence the size of certain visual illusions, such as tilt aftereffect and angle expansion (Levine and Grossberg, 1976).

The quenching of uniform patterns is due to a competitive balance between a narrow on-center I_i that interacts with a relatively large excitatory saturation point B , and a broad off-surround that interacts with a relatively small inhibitory saturation point $-C$; such a relative size scaling between B and C often occurs in passive membranes (Hodgkin, 1964). This conclusion generalizes to systems

$$\dot{x}_i = -Ax_i + (B - x_i) \sum_{k=1}^n I_k C_{ki} - (x_i + D) \sum_{k=1}^n E_{ki} \quad (48)$$

wherein inputs I_i can excite populations v_j near to v_i via the coefficients C_{ij} ("on-center") and can inhibit populations v_j over a broad expanse of cells via the coefficients E_{ij} ("off-surround"). Since the equilibrium

point of (48) is

$$x_i = \frac{I \sum_{k=1}^n \theta_k (BC_{ki} - DE_{ki})}{A + I \sum_{k=1}^n \theta_k (C_{ki} + E_{ki})} \quad (49)$$

a uniform pattern (all $\theta_i = 1/n$) is quenched (all $x_i \leq 0$) for any $I \geq 0$ whenever

$$B \sum_{k=1}^n C_{ki} \leq D \sum_{k=1}^n E_{ki}, \quad i = 1, 2, \dots, n \quad (50)$$

The breadth of excitatory and inhibitory interactions across the network determines the input patterns to which a population will respond. For example, let a vertical bar of light perturb the network. Suppose that the breadth of on-center and off-surround interactions is less than that of the bar. Then cells near the center of the bar will perceive a uniform field. Also, cells far away from the bar will perceive a uniform field. Both types of cells will be incapable of generating suprathreshold responses. Only cells near the transition regions of light and dark will respond. Such a network detects the edges of the bar.

The above mechanism can also be used as a matching mechanism, as in Section 20. To see this, consider the following question: Given a spatial pattern θ , how can a maximally mismatched pattern $\bar{\theta}$ be generated? Clearly $\bar{\theta}$ should be large where θ is small, and conversely. If both θ and $\bar{\theta}$ are input patterns to the network, then their mismatched peaks and troughs will add to create an almost uniform net pattern. Network activity is consequently suppressed. By contrast, if $\bar{\theta}$ is proportional (parallel) to θ , then the patterns add to amplify network activity.

45. The Growth of On-Center Off-Surround Connections

How can a balance between the ratio of excitatory and inhibitory saturation points, and between the distribution of on-center and off-surround coefficients, be effected? If it is not, and (say) the off-surround is too strong, then by (49) essentially all patterns will be suppressed. For example, suppose that lateral inhibition in the LGN not only contrast-enhances afferent sensory patterns, but also differentially amplifies LGN activity depending on how well the sensory patterns match cortical

feedback. How is the excitatory-inhibitory balance that is needed for matching generated? Before considering structural substrates of this mechanism, we note that differential shunting of on-center or off-surround interactions can retune the network by shifting its criterion of how uniform a pattern must be to be suppressed (cf. Elias and Grossberg, 1975, Section 3).

Simple growth rules are sufficient to formally explain some qualitative features of this balance. These are stated to illustrate how local data at each cell, such as the ratio BD^{-1} , can determine intercellular connections. Two related properties are desired: (1) the quenching of uniform patterns; (2) a narrow on-center and broad off-surround. By (50), the quenching of uniform patterns occurs if the ratio BD^{-1} equals all the ratios $E_i C_i^{-1}$, where C_i (E_i) is the total strength of excitatory (inhibitory) connections to v_i , $i=1, 2, \dots, n$. This can be achieved if "opposites attract." That is, suppose excitatory sites at v_i (whose number is proportional to B) support a process which attracts growing inhibitory connections, and inhibitory sites at v_i (whose number is proportional to D) support a process which attracts growing excitatory connections. Otherwise expressed, let the excitatory sites and inhibitory sites support processes which generate attractive *gradients* for inhibitory and excitatory connections, respectively. Then all the ratios $E_i C_i^{-1}$ will approach BD^{-1} if there exist enough intercellular connections to match the attracting sources. The hypothesis that cell growth follows some type of spatial gradient is a familiar one in developmental biology (cf. Gustafson and Wolpert, 1967; Grossberg, 1978a). Moreover, because $B > D$, the gradient attracting inhibitory connections will, other things being equal, be more uniform across space than the gradient attracting excitatory connections. A narrow on-center and broad off-surround will hereby tend to be produced. If this mechanism exists, then a change in BD^{-1} during the growth period, whether natural or experimentally controlled, should alter the relative spread of excitatory and inhibitory connections.

46. Goal Gradient and Plans

Using STSS and the quenching of uniform patterns, we can now study how associations develop among the chunks in a hierarchically coded field of populations. The need for such associations is clear from a variety of examples. In fact, we have already been using a simple version of this mechanism. In Section 21, the reciprocal trainable pathways between two fields $\mathcal{F}^{(1)}$ and $\mathcal{F}^{(2)}$ can be reinterpreted as a special case of reciprocal trainable pathways between two subfields of

chunks in a recurrent network. The code-learning in $\mathcal{F}^{(1)} \rightarrow \mathcal{F}^{(2)}$ LTM traces is distinguished from template-learning in $\mathcal{F}^{(2)} \rightarrow \mathcal{F}^{(1)}$ LTM traces only by the fact that tuning in $\mathcal{F}^{(2)}$ might contrast-enhance its patterns more than tuning in $\mathcal{F}^{(1)}$ does, but even this distinction need not hold in general.

Sections 1,A and 6,C provide other examples of this concept. For example, how does the internal representation of a goal object, such as an apple, trigger a plan to get an apple from the refrigerator? Somehow the code for the apple, which is not as "abstract" as the code for the plan, can activate the code for the plan based on prior experience. This occurs, we assume, because the apple representation can sample the representation of the plan during learning, and can activate this representation during performance.

Several remarks are important here. First, conditioning from representations of a goal to a plan should be possible given essentially any goal and any realizable plan. Thus, the possible sampling pathways should be distributed broadly throughout the network. This would create a devastating noise problem were it not for STSS and the quenching of uniform patterns. The quenching of uniform patterns eliminates all effects of sampling across irrelevant cues, and STSS amplifies the codes that are most informative in a given context of events. Second, if the goal representation is a lower-order chunk, its signals among other lower-order chunks can be suppressed by the quenching of uniform patterns, but its sampling of higher-order representations, such as plans, will not be quenched because of the built-in distinctiveness, or informativeness, of the activity patterns across higher-order chunks. Third, even if the plan is coded by a spatial pattern across higher-order populations, rather than by undifferentiated activity within a population, it can be accurately sampled by the goal representation. Fourth, when the goal occurs, it generates incentive-motivation that can amplify sampling by its representation of the plan (cf. Grossberg, 1975). At the time this occurs, the goal is the last event to have occurred. By Section 27, even if the goal is preceded by many prior events, it still has enough activity to elicit sampling when this activity is supplemented by incentive motivation. Fifth, because the incentive motivation is nonspecific, *all* active lower-order chunks can sample the plan, but with a strength that depends on their STM activities. For example, suppose that the STM pattern is monotone increasing across the most recent event representations, owing to the occurrence of many prior events. Then a classical goal gradient will be learned such that the most recent items will elicit the plan with the greatest efficacy. Sixth, this mechanism shows how *partial avalanche* structures can be embedded in the network, as in Fig. 40. The

partial avalanches blend together, in a self-consistent fashion, descending order information from commands with ritualistic sequence information due to reliable, and therefore unquenched, occurrence of the same sequence through time. These partial avalanches are a type of primitive syntactic structure in the network. Figure 40 illustrates the importance of adaptation, or competition, of STM activity throughout the hierarchy of chunks. This competition prevents any population from receiving too large an input from any one source, and thereby preserves decision rules for signal generation that require convergence of specific and nonspecific inputs (Section 7). Finally, the hierarchy tunes itself. For example, suppose that chunks (A), (B), (C), . . . are activated during a particular stage of learning. Then plans and partial avalanches can be gradually built on these chunks during this learning stage. If, however, chunks (ABC), (DEF), . . . are activated during a later learning stage, then the earlier plans and partial avalanches will be masked by STSS. This observation can be expanded to explain various properties of transfer from serial learning to paired-associate learning, and vice versa. In effect, the rehearsal strategy helps to choose the base code on which higher codes and feedback relations will develop.

The above properties all depend on a rapid normalization, or adaptation, of STM activity throughout the hierarchy of chunks, followed by LTM sampling of informative STM patterns throughout the hierarchy. Competition within the STM hierarchy has been used to explain differences between simultaneous versus successive contrast, respectively, in the visual illusions of angle expansion and tilt aftereffect (Levine and Grossberg, 1976). These qualitative properties must eventually be supplemented by a mathematical classification of how parametric differences in the intrafield and interfield interactions of particular classes of feature detectors generate different STM and LTM patterns. This analysis must include systems of the form

$$\begin{aligned} \dot{x}_i = & -A_i x_i + (B_i - C_i x_i) \left[\sum_{k=1}^n f_k(x_k) D_{ki} z_{ki} + I_i \right] \\ & - (E_i + F_i x_i) \left[\sum_{k=1}^n g_k(x_k) G_{ki} + J_i \right] \end{aligned} \quad (51)$$

and (2), which generalize (1) and (2).

47. STM Order Reversal: Item Learning versus Order Learning

Having discussed aspects of the self-organization of codes, we must now focus more closely on what is coded. This problem can be

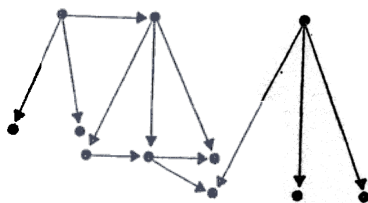


FIG. 40. Partial avalanches combine self-consistent order information from descending (contextual) commands with associational chains.

motivated by using the discussion in Section 31. Suppose that items are rehearsed one at a time as they occur. Let the $\mathcal{F}^{(2)}$ populations $v_{i_1}^{(2)}, v_{i_2}^{(2)}, \dots, v_{i_k}^{(2)}, \dots$ be sequentially activated by these items, and let rehearsal delete suprathreshold STM activity at $\mathcal{F}^{(2)}$ before the next item is presented. As this occurs, the populations $v_{i_1}^{(3)}, v_{i_2}^{(3)}, \dots, v_{i_k}^{(3)}, \dots$ in $\mathcal{F}^{(3)}$ are also sequentially activated (cf. Section 35). Although rehearsal deletes STM at $\mathcal{F}^{(2)}$, it does not delete STM at $\mathcal{F}^{(3)}$. If $k < \text{TMS}$, then order information is accurately coded at $\mathcal{F}^{(3)}$ after rehearsal terminates. A nonspecific rehearsal wave to $\mathcal{F}^{(3)}$ can then trigger a readout from $\mathcal{F}^{(3)}$ of the items with their proper relative activities. This example illustrates how each item can be rehearsed during list presentation (repeat the last item), followed by a repetition of the whole list in its correct order (repeat the first item), even if no population codes order information in its LTM traces. This capability is important. Otherwise a telephone number could not be repeated unless it had already been encoded in LTM. This example again suggests the usefulness of studying how performance variables alter the arousal level and thus the form (decreasing, increasing, or unimodal) of the transient STM pattern.

Contrast the above experiment with one in which the items are presented in the same order, but they are rehearsed two at a time. Then a population $v_{i_1 i_2}^{(3)}$ in $\mathcal{F}^{(3)}$ will gradually code the first two items, and will learn order information as these items are rehearsed. As $v_{i_1 i_2}^{(3)}$ is coded, it gradually masks $v_{i_1}^{(3)}$ and $v_{i_2}^{(3)}$ if STSS holds. Eventually the chunks $v_{i_1 i_2}^{(3)}, v_{i_2 i_3}^{(3)}, \dots$ will control readout from $\mathcal{F}^{(3)}$ of order information after rehearsal takes place. This argument shows how a combination of prior coding and the present rehearsal strategy determine which chunks will be active at $\mathcal{F}^{(3)}$; for example, how the network automatically groups items in a given rehearsal unit into familiar chunks. Thus, if the familiar letters Q and L are rehearsed together, then each letter can excite a previously coded $\mathcal{F}^{(3)}$ chunk even if the present rehearsal strategy starts to code a higher-order chunk. The dominant codes at a given time learn

order information, but what codes are dominant can change as item learning continues.

The following sections consider the spatial analog of this temporal chunking process—namely, the self-organization of maps.

48. Circular Reaction and Map Formation

This section explicates the heuristic themes in Sections 1,C and 1,D. To fix ideas, imagine that an infant's hand-arm system is endogenously active and that its eye-head system tends to follow the motions of its hands. How does this unconditional process generate learned maps that can guide the hand-arm system to a terminal position, never before experienced, that is focused by the infant's eye-head system? Similarly, after an infant unconditionally babbles simple sounds, how does it imitate sounds that are more complicated than those it babbled? A complete answer to these questions would require a thorough analysis of the neural controls of motor behavior. Herein we note a minimal synthesis of resonance, sampling, and nonspecific arousal mechanisms acting on proprioceptive and terminal motor data that suffices to learn and perform sensory-motor maps and to maintain descending postural commands. Then we note that this model is really a variation of the adaptive coding model.

The following construction holds independently of what data are coded by a particular motor map. It focuses on the minimal operations that are needed to learn maps effectively. For example, we shall ignore the fact that different combinations of eye and head position can focus the eyes on the same physical position relative to the body. The construction will, however, be motivated by a familiar example: pursuit of the endogenously active hand by unconditional eye motions.

Let the image of the hand move across the retina. Suppose that its position on the retina (after compensatory computations of head position) determines a terminal eye position that will move the eye until the hand is viewed by the fovea. Clearly, no learned correlation between eye and hand position should be initiated until the eye foveally fixates the hand. Otherwise, an arbitrary eye position could continuously be correlated with all possible hand positions. How does the eye-head system know that it is fixating something? This happens when its terminal motor coordinates match its proprioceptive motor coordinates. The terminal coordinates code where the system *wants* to go; the proprioceptive coordinates code where the system *is*. When the eye

actively fixates an object, these two sets of coordinates code the same position.

The following mechanisms explicate the idea that the eye is fixating something. First, the cells that code terminal coordinates send signals to the eye muscles that hold the eyes in position. As in Section 16, we assume that specific inputs to these cells determine their relative activities, and that a nonspecific arousal source shunts these activities into the suprathreshold range. In particular, let these cells be joined by on-center off-surround interactions so that the shunt can work. Second, the system must somehow know that the proprioceptive and terminal coordinates agree. Somehow data from the two coordinate sets must be brought together, and a characteristic dynamic state must be generated only when the two sets match. In the present theory, such a match triggers a resonant state. In all, sampling signals are emitted only if a resonant state exists between proprioceptive and terminal coordinates, and this resonant state is maintained as long as the nonspecific shunt is kept on.

The same conditions exist when a previously learned map is performed. Suppose that an eye fixates on an object that is to be touched. The eye-hand terminal coordinates then match their proprioceptive coordinates. If arousal is turned on, then the eye-hand system can send learned signals to the hand-arm system. These signals code the terminal hand position that was correlated with the eye position during learning trials.

Another constraint is needed before the minimal mechanism can be described. Often a motor system is directed to fixate on a different position from its present one. Then its terminal coordinates (where it wants to go) do not match its proprioceptive coordinates (where it is). Indeed, this is the typical situation when the hand is directed to touch an object. Obviously the motor system can do this, and it does it without interference from its present position. Thus, during performance trials, the terminal coordinates can release motor signals even if they do not match the proprioceptive coordinates. Arousal suffices to release signals from the terminal coordinates in this case.

In summary, sampling signals can be released only when terminal and proprioceptive coordinates match and are sustained by arousal; yet if terminal coordinates are activated by signals from another system, they can generate performance signals when they are activated by arousal, even if they do not match their proprioceptive coordinates. These constraints are summarized in Fig. 41. Figure 41 depicts two systems, (I) and (II), of proprioceptive and terminal motor coordinates; for example,

A THEORY OF HUMAN MEMORY

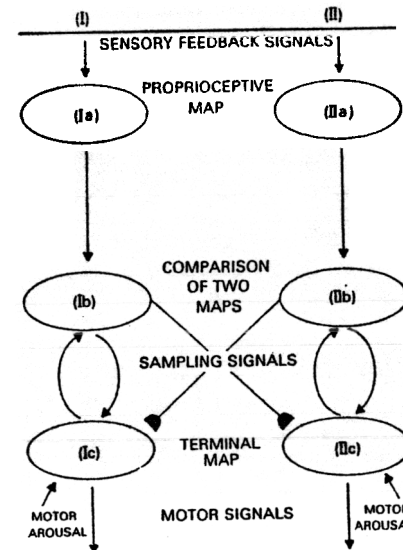


FIG. 41. Sampling signals are elicited only if proprioceptive and terminal maps agree. Motor signals are elicited if a subliminal terminal map is aroused.

let (I) be an idealization of the eye-head system, and let (II) idealize the hand-arm system. Consider (I) for definiteness. Proprioceptive coordinates are coded by the populations in field (Ia), and thereupon send signals to field (Ib). Terminal coordinates are coded in field (Ic), and can send signals in two directions if they are aroused. Signals can descend as motor commands to appropriate muscles. Signals can also go to (Ib). At (Ib), the proprioceptive and terminal maps are compared, as in Section 44, by an on-center off-surround network. If the maps match, then (Ib) amplifies their commonly shared pattern. If the maps do not match, then activity in (Ib) is suppressed. When the maps match, two things happen. Feedback signals go to (Ic), whereupon a resonance is established between (Ib) and (Ic). This resonance sustains sampling signals from (Ib) to the terminal coordinates at (IIC). In all, when proprioceptive coordinates at (Ia) match suprathreshold terminal coordinates at (Ic), then sampling signals are emitted from (Ib).

What happens to these sampling signals? Suppose that (IIC) receives signals from (Ib), which are thereupon amplified by arousal. Then (IIC) can emit motor commands to its muscles. Field (IIC) also sends signals to (IIb), but if the hand is not yet at the desired terminal position, the mismatch between proprioceptive and terminal signals at (IIb) quenches

any possible feedback signals from (IIb) to (IIc). Thus the terminal map at (IIc) can emit motor commands to its muscles without interference due to proprioceptive feedback.

49. Quenching of Irrelevant Cues

An important property is now evident. Suppose that irrelevant signals reach (IIc); that is, suppose that several populations in (Ib) send uniform signal patterns to (IIc). The on-center off-surround network at (IIc) will quench these patterns, as in Section 44. Thus, any number of irrelevant cues can sample (IIc) and learn a map imposed by a relevant cue without distorting the performance of this map.

The property of irrelevant cue quenching is crucial wherever many cues are capable of sampling learned data. Without it, the omnipresent existence of such cues would rapidly erode the LTM pattern. This property also shows how relevant cues can be included in a plan, whereas irrelevant cues, such as slowing down, speeding up, taking a drink, etc., need not control behavior if they vary across trials.

50. Feedforward Reset of Sequential Terminal Maps

What happens if a sequence of motor commands iteratively perturbs (IIc)? How is the terminal map at (IIc) reset by the next command? Moreover, if no new commands occur, how can the terminal map at (IIc) be stored while its command is being executed? Sections 9 and 29 pointed out that a command is deleted after it is released to prevent continual iteration of the same command. Somehow the command data, which are now explicated as a terminal map, are then stored while performance proceeds.

Both the reset and STM storage properties follow from the uniform quenching property if the network at (IIc) is made recurrent (see Fig. 42). Suppose that no signals are being emitted by (IIb), because the terminal map at (IIc) has not yet been executed. Let a given terminal map be received at (IIci) via external signals. This map is then reproduced at (IIcii) via the (IIci) \rightarrow (IIcii) pathways. If the external signals are shut off, then the terminal map can resonate at (IIc) as long as shunting arousal is maintained. Suppose, however, that a new terminal map is imposed at (IIci) by external signals. Feedback signals from (IIcii) to (IIci) still carry the old terminal map. If the two maps do not match, then activity across (IIc) will be momentarily quenched. All traces of the old terminal map signals are hereby eliminated, and the

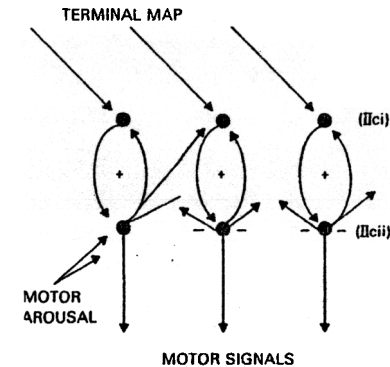


FIG. 42. Delayed feedback from (IIcii) to (IIci) briefly inhibits STM at (IIci) if the new terminal map does not match the old terminal map. The new terminal map then resonates in STM, and reset is completed.

new terminal map is instated. In short, delayed interneuronal feedback within a network that possesses the uniform quenching property implies the existence of an STM reset mechanism. Several variations on the theme of Fig. 42 exist. For example, *in vivo* do inhibitory interneurons also descend from (IIci) to (IIcii)? All these variations work better if the time needed to self-inhibit a descending command exceeds the time needed to reset a terminal map.

51. Posture, Isometrics, Saccades, and Feedforward Motor Control

The above construction suggests many insights and questions about motor control.

It suggests a mechanism of posture. There exist two main types of systems that must be distinguished by experiment in particular cases. The first type was described in Section 50. Stage (IIc) can maintain a terminal map using recurrent signals even if the descending motor command is silent. In this system, posture is the terminal position of a persistently aroused terminal map.

The second type cannot store a terminal map at stage (IIc). It requires a persistent source of terminal map signals during map performance. Is eye-head to hand-arm control of this type? For example, let a limb be controlled by system (II). Suppose that the limb moves to a desired position under the direction of sustained eye fixation at that position. How can this position be maintained after the source of its terminal

signals is shut off? When the limb reaches the desired position, a resonance is established between (IIb) and (IIc). After resonance sets in, shut off the source of terminal signals to (IIc)—that is, the motor command. Field (IIb) continues to supply the correct terminal coordinates to (IIc), and arousal at (IIc) keeps the map suprathreshold. In other words, proprioceptive feedback plus arousal can sustain the resonance, and thus the signals to muscles that maintain the posture. When arousal is shut off, the resonating loop is broken, and the postural position is terminated.

The minimal mechanism also has isometric properties when a limb is passively dragged into a sequence of terminal positions. If arousal is off, the proprioceptive data flows from (IIa) to (IIb) to (IIc). There is no feedback to (IIb). When arousal is turned on, however, then the resonance between (IIb) and (IIc) resists further change due to hysteresis. Simultaneously, signals from (IIc) to the muscles tend to hold the present terminal position. Thus the size of the arousal signal determines the amount of isometric tension in this system.

Some light is also shed on the controversy about whether the motor commands of terminal position are feedforward, or are determined by proprioceptive feedback (Bizzi *et al.*, 1975). In fact, the answer even in the minimal network is subtle. After a signal is received at (IIc) from (Ib), then the motor command leaves (IIc) for the muscles. Proprioceptive feedback at (IIa) is blocked at (IIb) to prevent it from distorting the terminal map at (IIc). This system is therefore capable of reaching its terminal position by using feedforward control. However, as the limb approaches the desired terminal position, resonance is established between proprioceptive feedback and the terminal map. Then proprioceptive cues support the terminal map even though they are not needed to reach the terminal position. The subtlety arises because the existence of proprioceptive input at the terminal map cells does not imply the necessity of this input for computing terminal position. Rather it sets the stage for map learning and postural maintenance.

The systems also suggest interesting questions about saccadic versus continuous map control. While a map is being *learned*, the arm and hand move together or are at rest. Only in this way can resonances be established between proprioceptive and terminal coordinates. Yet, while a map is being *performed*, its proprioceptive map almost never agrees with its terminal map. Indeed, when you fixate an object that your hand is already on, there is no need to turn on a map to move your hand onto the object. The situation during learning can be described as a *continuous* motion wherein proprioceptive and terminal coordinates are always very close. The situation during performance can be described as a

saccadic motion wherein proprioceptive coordinates do not influence the motor act. This distinction is often made in the literature on motor control (Robinson, 1964; Yarbus, 1967). In the minimal network, the continuous and saccadic systems share some cells in common—for example, the (Ic) cells that code terminal maps. The present construction therefore suggests some interesting questions about the interpretation of oculomotor data.

The cells in (Ib) will fire only when the system moves in its continuous mode—that is, when proprioceptive and terminal coordinates are close to one another. Are these among the cells that are usually included in the continuous system? If they are, then cells such as those in (Ic) should exist that are active in both the continuous and the saccadic modes. A likely place to search for such cells in appropriate species is the superior colliculus (Goldberg and Wurtz, 1972a,b; Stryker and Schiller, 1975; Wurtz and Goldberg, 1972a,b), where maps of visual sensory data into motor eye-movement data have been experimentally described.

In the system as it stands, it is possible to learn spurious correlations between eye and hand positions. In principle, the eye can fixate one point and the hand can be held at another point while map learning takes place. However, since an infant's eye is drawn strongly toward moving objects (Piaget, 1963), it is very unlikely that this will ever happen under normal developmental conditions. Also, the tendency is not eliminated by building up maps using cells that are activated by dynamic limb motions rather than by spatial patterns. For example, the eye is often fixated on an object that it wants to touch; the sampling cells thus cannot be activated only by eye motion. Also the arm is often called upon to move to a prescribed terminal point from an arbitrary initial point; thus, the terminal map cannot be determined by directionally sensitive cells. Although the above system can in principle be fooled, it can also delete correlations that lead to erroneous, and therefore unexpected, consequences using attentional mechanisms, as in Section 20 and in Grossberg (1975).

The system can also be fooled, as it stands, because (Ib) can sample a suprathreshold pattern at (IIc) even if system (II) has not yet reached its terminal position. This property creates no difficulty if sampling usually occurs in the continuous mode.

The above shortcomings suggest possible limitations on the accuracy of map learning in infants, and emphasize the importance of an infant's sensitivity to moving objects. The last shortcoming can easily be formally overcome. Insert an interneuron in the path from (IIb) to (IIc), and let (Ib) sample the interneuron. Since the interneuron is inactive

unless resonance holds between (IIb) and (IIc), map learning between (I) and (II) can then occur only if both systems are resonating. Such variations are not very useful in the absence of data that can discriminate between them.

There is a variation that is useful to consider, however. In principle, the proprioceptive and terminal coordinates of a given position need not generate the same pattern of neural activity. Where this occurs, stage (Ib) can be used to learn the match between proprioceptive and terminal maps that code the same position. In fact, stages (Ia), (Ib), and (Ic) then emerge as an adaptive coding scheme for bridging the gap between the proprioceptive and terminal coordinates within a given motor system. By way of illustration, suppose that the (Ia) \rightarrow (Ib) pathways contain LTM traces as in Section 17. Then (Ib) can learn a code for the proprioceptive coordinates at (Ia). Feedback pathways from (Ib) to (Ia) would then exist to stabilize this code using learned templates. The coded proprioceptive map at (Ib) could then learn the corresponding terminal map at (Ic) using LTM traces in the (Ib) \rightarrow (Ic) pathways, while the terminal map at (Ic) can learn the proprioceptive code at (Ib) using LTM traces in the (Ic) \rightarrow (Ib) pathways.

Given any such adaptive rules within the systems (I) and (II), map learning *between* these systems then becomes a continuation of map learning *within* each system separately. The between-system stage cannot begin until resonances can be generated within each system. As in Section 24, each stage in this formal developmental sequence obeys similar principles, even though different stages code different behaviors.

Finally, we note that, if the networks learn only terminal positions, then auxiliary systems must exist that enable limbs to reach these positions. The α - γ system is a classical example of such an auxiliary system (Granit, 1966; Grillner, 1969; Thompson, 1967, Chapter 12) but it goes beyond the scope of this paper.

52. Feedforward versus Feedback Control of Sequential Map Performance

What is remembered when we learn a series of motor acts, such as a dance? Below we suggest that serially ordered terminal maps generate and are sampled by adaptive sensory-motor codes (cf. Sections 11, 50, and 51) under the guidance of internal and external feedback. A model of this process will now be described using mechanisms that have already been introduced.

First we suppose that a serial STM buffer for motor activity exists. In Fig. 36, fields $\mathcal{F}^{(1)}$, $\mathcal{F}^{(2)}$, and $\mathcal{F}^{(3)}$ model some stages of a sensory, in

particular an auditory, STM buffer. Henceforth these fields are denoted by $\mathcal{F}_S^{(1)}$, $\mathcal{F}_S^{(2)}$, and $\mathcal{F}_S^{(3)}$. Now expand Fig. 36 as follows. Denote $\mathcal{F}^{(4)}$ in Fig. 36 by $\mathcal{F}_M^{(2)}$, since it is functionally analogous to $\mathcal{F}_S^{(2)}$, which also codes data from a given time frame. Let $\mathcal{F}_M^{(1)}$ denote the field of terminal motor maps, and denote $\mathcal{F}^{(5)}$ by $\mathcal{F}_M^{(3)}$, since it codes order information among these maps. Trainable pathways from $\mathcal{F}_S^{(3)}$ to $\mathcal{F}_M^{(3)}$ are introduced, as in Section 35, to enable sensory order information to activate corresponding motor order information. These sensory-motor signals will automatically be tuned at $\mathcal{F}_M^{(3)}$ by the properties of self-similarity and uniform quenching within the motor STM buffer (see Fig. 43). Indeed, all coding operations in the sensory buffer are assumed to have analogs in the motor buffer.

How are learned data read out of the motor buffer? To see this, suppose that a monotone decreasing pattern exists across the populations $v_1^{(2)}$, $v_2^{(2)}$, $v_3^{(2)}$, . . . of $\mathcal{F}_M^{(2)}$. When $\mathcal{F}_M^{(2)}$ is aroused by a nonspecific rehearsal wave, $v_1^{(2)}$ fires until its activity self-destructs via inhibitory feedback. The spatial pattern of $v_1^{(2)}$'s signals generates a terminal map at $\mathcal{F}_M^{(1)}$. As in Section 50, this pattern becomes suprathreshold and is stored in STM if $\mathcal{F}_M^{(1)}$ receives motor arousal. The terminal map thereupon elicits descending signals to the muscles that will execute it.

The above mechanism highlights a difficulty of traditional probabilistic models of behavior. How would a probability theorist interpret the control of motor commands? Suppose that STM activity at $v_1^{(2)}$ is interpreted as the probability of executing a terminal map. After $v_1^{(2)}$

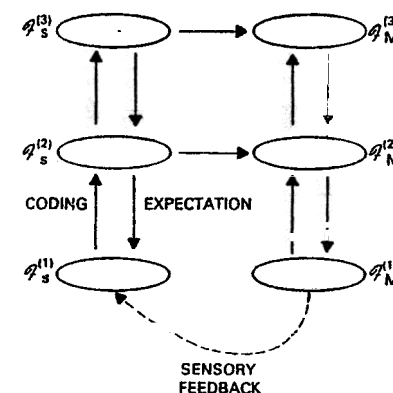


FIG. 43. Code learning and template learning in and between sensory and motor modalities.

self-inhibits, where does the probability go? The terminal map is still active, so the probabilistic modeler might say that $v_1^{(2)}$ stays active. This is, however, false and yields incorrect intuitions about the process by obscuring the distinction between coding of order and coding of terminal position.

How does field $\mathcal{F}_M^{(2)}$ know when the $v_1^{(2)}$ act is nearing completion? Otherwise expressed, how does $\mathcal{F}_M^{(2)}$ know when to turn on the next motor map? There are two answers in the model.

The first answer describes a feedforward control that is analogous to rapid arpeggio playing on a piano (cf. Lashley, 1951). If rehearsal arousal is kept on at $\mathcal{F}_M^{(2)}$, then each population $v_1^{(2)}$, $v_2^{(2)}$, $v_3^{(2)}$, . . . is excited in its proper order as the previous population self-inhibits. Each successive spatial pattern of signals descends to $\mathcal{F}_M^{(1)}$, where it controls motor action until it is reset by the next pattern. Feedback influences the system only indirectly via sensory feedback that is chunked before generating signals to $\mathcal{F}_M^{(2)}$.

The second answer describes a feedback control that is analogous to executing slow dance motions. Let $\mathcal{F}_M^{(2)}$ be briefly aroused. Then $v_1^{(2)}$ fires and self-inhibits. Motor arousal is maintained at $\mathcal{F}_M^{(1)}$, which stores the first terminal map and activates its muscles accordingly. As the act is executed, the proprioceptive coordinates approach the terminal coordinates. As the match between proprioceptive and terminal coordinates improves, a nonspecific rehearsal wave perturbs $\mathcal{F}_M^{(2)}$. This rehearsal wave was mentioned in Section 23, where the existence of a competition between two antagonistic motor arousal systems was postulated. The first system is allowed to fire when a mismatch of test pattern and expected pattern occurs. The second system is triggered by output from the match between test pattern and expected pattern. The first arousal system energizes orientation toward an unexpected event. The second arousal system energizes goal-directed consummation of an expected event. In the present motor example, the expected pattern is the terminal map, and the test pattern is the proprioceptive map.

Thus, as the proprioceptive map approaches the terminal map, a nonspecific rehearsal wave is released and reaches $\mathcal{F}_M^{(2)}$. Population $v_2^{(2)}$ thereupon fires and sends its spatial pattern of terminal signals to $\mathcal{F}_M^{(1)}$ as it self-inhibits. When this terminal map reaches $\mathcal{F}_M^{(2)}$, it resets $\mathcal{F}_M^{(2)}$ as in Section 50. The proprioceptive map of the first command does not match the terminal map of the second command. Hence the rehearsal wave terminates so that no further output emerges from $\mathcal{F}_M^{(2)}$. The second terminal map can therefore control performance until its goal is reached. Then the cycle of matching; resonance and rehearsal wave;

sampling and self-inhibition; and reset and mismatch begins again. After there are no further STM data in $\mathcal{F}_M^{(2)}$, the posture of the last gesture is held until motor arousal is withdrawn from $\mathcal{F}_M^{(1)}$.

The above mechanism can now be joined with previous results (for example, Section 35) concerning the reorganization of the future field due to sensory feedback. All the results go through because they hold for arbitrary spatial patterns, irrespective of what these patterns code.

53. Sequential Switching between Sensory and Motor Maps

The same general mechanisms seem to hold in many examples of goal-oriented behavior. Another example is briefly sketched below to illustrate how switching between sensory and motor map systems can occur when a plan is executed.

Suppose that the command "Touch the yellow ball" is given. I suggest that the verbal command is encoded and does (at least) two things. The verbally decoded message elicits signals that are gated by LTM traces. The gated signals subliminally activate sensory feature detectors (for example, in $\mathcal{F}_S^{(2)}$) which code a yellow, ball-like object because of prior learning. Since there is no match with this field, the message can also trigger a series of orienting reactions by activating the eye-head motor buffer. This series is perhaps subject to the spacing effect of Section 42, which tends to inhibit perseverative search in the same position. The search continues until the yellow ball is seen. The ball is then visually decoded and (approximately) matches the subliminal sensory pattern. The match induces a resonance that inhibits orienting motor arousal and, in a complementary fashion, triggers goal-oriented motor arousal.

What motor system is hereby aroused? When orienting arousal is inhibited, the eye-head system is fixated on the yellow ball. It can thereby generate a terminal map in the hand-arm systems. One of these is activated by motor arousal. The chosen hand then moves toward the ball until it touches it. In all, the plan has been executed by a sensory match, a motor match, and a switchover from orienting motor arousal to goal-oriented motor arousal.

54. Map Reversal by Antagonistic Rebound

Many details have been overlooked in the above description in order to emphasize its essential simplicity. Also, many new questions are raised. For example, after the hand surrounds the ball, how does it

retrieve the ball? There exists an elementary answer that is probably most true in infants, but it also highlights how different adult control can be. After the hand surrounds the ball, suppose that an antagonistic rebound is triggered throughout the arm-hand system. For example, if touching the ball elicits nonspecific arousal to the arm-hand terminal map, then the agonist-antagonistic cells throughout the map will transform flexion into extension, and conversely, in a graded fashion across the affected muscle groups. In particular, the hand will close and the arm will be retrieved.

The main point is that retrieval need not, in principle, recode the act cell by cell, which would require a high-dimensional control. A one-dimensional increment in nonspecific arousal can be used instead.

55. Imprinting, Imitation, and Sensory-Motor Algebra

Consider the babbling behavior of an infant (Fry, 1966) as motivation for the following construction. Suppose that terminal maps at $\mathcal{F}_M^{(1)}$ are endogenously activated, analogous to the babbling of sounds. The execution of these maps elicits sensory feedback, such as sounds, which feed back to $\mathcal{F}_S^{(1)}$ and are coded at $\mathcal{F}_S^{(2)}$ on successive trials. The terminal motor map is adaptively coded at $\mathcal{F}_M^{(2)}$, and the motor code at $\mathcal{F}_M^{(2)}$ simultaneously learns the terminal motor map at $\mathcal{F}_M^{(1)}$ as a feedback template. As this happens, $\mathcal{F}_M^{(2)}$ is sampled by its sensory feedback code at $\mathcal{F}_S^{(2)}$. If more than one motor code is active at $\mathcal{F}_M^{(2)}$ —say during predictive performance—then $\mathcal{F}_S^{(2)}$ samples its motor command in a context of temporally contiguous motor commands. Performance of the same command can thus be different in different contexts; compare the performance of a given speech sound in different words. As codes within $\mathcal{F}_S^{(2)}$ and $\mathcal{F}_M^{(2)}$ develop, repeated sequences of motor commands can be adaptively coded as sequential motor chunks at $\mathcal{F}_M^{(3)}$ and as sequential sensory feedback chunks at $\mathcal{F}_S^{(3)}$. Thereupon LTM traces from $\mathcal{F}_S^{(3)}$ to $\mathcal{F}_M^{(3)}$ can encode associations from the sensory to motor sequential chunks.

Endogenously active arousal sources supply the motivational support that drives these LTM processes. The combined effect of all these LTM processes is to *complete*, or *close*, the sensory-motor feedback loops that are endogenously activated during the babbling phase. Hereby the network can imitate a sound that is supplied by an external source, if that sound lies on the generalization gradient of some sound that was endogenously babbled. It does this by activating the sequence of fields

$$\mathcal{F}_S^{(1)} \rightarrow \mathcal{F}_S^{(2)} \rightarrow \mathcal{F}_M^{(2)} \rightarrow \mathcal{F}_M^{(1)}$$

if the sound is coded by a spatial pattern, or the sequence of fields

$$\mathcal{F}_S^{(1)} \rightarrow \mathcal{F}_S^{(2)} \rightarrow \mathcal{F}_S^{(3)} \rightarrow \mathcal{F}_M^{(3)} \rightarrow \mathcal{F}_M^{(2)} \rightarrow \mathcal{F}_M^{(1)}$$

if the sound is coded by a sequence of spatial patterns.

The endogenously active arousal sources *imprint* developing maps by using the same mechanisms that are driven by reactively activated motivational sources in the "adult" network. Imprinting in the young network and learning in the adult network are similar processes using different motivational triggers.

How does the network learn to imitate more complex sounds than those that are endogenously produced? This process follows automatically from properties of the adaptive coding model. A new sensory pattern at $\mathcal{F}_S^{(1)}$ is filtered, or decomposed, by $\mathcal{F}_S^{(1)} \rightarrow \mathcal{F}_S^{(2)}$ signals into a set of simpler patterns whose sensory-motor loops have been endogenously closed. Each simpler sensory pattern activates its corresponding motor pattern at $\mathcal{F}_M^{(1)}$. The total pattern that is hereby synthesized at $\mathcal{F}_M^{(1)}$ can produce a sound that is close to the new sensory pattern.

More precisely, suppose that m motor patterns O_i have been endogenously activated at $\mathcal{F}_M^{(1)}$ and thereby elicit the sensory feedback patterns I_i at $\mathcal{F}_S^{(1)}$, $i = 1, 2, \dots, m$. Let I_i be adaptively coded by population $v_{S_i}^{(2)}$ in $\mathcal{F}_S^{(2)}$, and let O_i be adaptively coded by population $v_{M_i}^{(2)}$ in $\mathcal{F}_M^{(2)}$. Also let each population learn its template of feedback signals, and its $\mathcal{F}_S^{(2)} \rightarrow \mathcal{F}_M^{(2)}$ associations. Given this learned substrate, how does the network learn to imitate a sensory pattern I which has never been endogenously elicited at $\mathcal{F}_M^{(1)}$? It is shown below how input I at $\mathcal{F}_S^{(1)}$ is filtered by the $v_{S_i}^{(2)}$ and then resynthesized by the $v_{M_i}^{(2)}$ to produce at $\mathcal{F}_M^{(1)}$ a motor pattern O which elicits the sound I . In summary, during imprinting, endogenous motor commands at $\mathcal{F}_M^{(1)}$ elicit their sensory commands at $\mathcal{F}_S^{(1)}$ as external feedback; during imitation, sensory commands at $\mathcal{F}_S^{(1)}$ elicit their motor commands at $\mathcal{F}_M^{(1)}$ via network filtering.

The filtering mechanism uses elementary vector space properties (Thomas, 1968). As adaptive coding proceeds in response to a sensory pattern I , the classifying vector z_i becomes proportional to I_i , and the signal from $\mathcal{F}_S^{(1)}$ to $v_{S_i}^{(2)}$, namely $S_i = I_i \cdot z_i$, becomes proportional to $|I_i|^2$. Suppose that the set of I_i , $i = 1, 2, \dots, n$, *spans* the vector space of input patterns I at $\mathcal{F}_S^{(1)}$; then any I can be written a linear combination $I = \sum_{k=1}^n \alpha_k I_k$ of the I_i 's given suitable coefficients α_k . If, moreover, the I_i are mutually orthogonal (that is, $I_i \cdot I_j = 0$, $i \neq j$), then the signal to $v_{S_i}^{(2)}$ in response to I , namely $S_i = I \cdot z_i$, is proportional to α_i . Thus each $v_{S_i}^{(2)}$ is excited by a signal that is proportional to how close I_i is to I . If the signals from $\mathcal{F}_S^{(2)}$ to $\mathcal{F}_M^{(2)}$ are linear, then the signal to $v_{M_i}^{(2)}$ is also

proportional to α_i . Since $v_{M_i}^{(2)}$ encodes pattern O_i in its LTM traces, $v_{M_i}^{(2)}$ generates a pattern at $\mathcal{F}_M^{(1)}$ that is proportional to $\alpha_i O_i$. The total pattern at $\mathcal{F}_M^{(1)}$ that is generated by $\mathcal{F}_M^{(2)}$ is therefore close to $O = \sum_{k=1}^n \alpha_k O_k$.

Does O elicit a sound close to I ? This is true if the mapping from motor patterns at $\mathcal{F}_M^{(1)}$ to sensory feedback patterns (approximately) preserves the weights α_i . This important property should be tested experimentally.

Several aspects of this mechanism deserve comment. Most remarkably, the signal law $\sum_{k=1}^n B_{ki} z_{ki}$ in (1), which was originally derived from simple classical conditioning postulates, also implies a map formation property.

The crux of the argument is that, no matter what motor pattern is endogenously active at $\mathcal{F}_M^{(1)}$, and no matter what its sensory feedback pattern is at $\mathcal{F}_S^{(1)}$, the sensory feedback pattern gets associated via $\mathcal{F}_S^{(2)}$ and $\mathcal{F}_M^{(2)}$ with the motor pattern. In vector space terms, an arbitrary set of vectors in the vector space at $\mathcal{F}_M^{(1)}$ can be associated with an arbitrary set of vectors in the vector space at $\mathcal{F}_S^{(1)}$. This property lets each network adapt to individual differences in the structure of its sensory and motor modalities. Given these associations, imitation is achieved if the network can map the weights α_i as signal sizes from $\mathcal{F}_S^{(2)}$ to $\mathcal{F}_M^{(2)}$. This is relatively easy to do, since it reduces a global mapping problem between two multidimensional vector spaces to a local rule for signal transmission.

What kinds of coding difficulties can occur? First, the patterns I_i need not span the space of input patterns. Inputs I , which cannot be represented as a linear combination $\sum_{k=1}^n \alpha_k I_k$, will then be filtered, or projected, by $\mathcal{F}_S^{(1)} \rightarrow \mathcal{F}_S^{(2)}$ signals as the closest pattern I^* that can be represented in this way. Second, if the patterns I_i are not mutually orthogonal, then each signal $S_i = I \cdot z_i$ will include interference terms of the form $I_j \cdot I_i$. These terms will distort the relative activities of the $\mathcal{F}_S^{(2)}$ populations. How can these and other distortions be corrected? The next section suggests an answer.

56. Self-Tuning and Multidimensional Inference in a Parallel Processor

The size of the QT in $\mathcal{F}_S^{(2)}$ and $\mathcal{F}_M^{(2)}$ will determine which of the populations $v_{S_i}^{(2)}$ and $v_{M_i}^{(2)}$ will be supraliminally excited. In order to map the weights α_i accurately, the QT must be small. This can occur in two main ways: structurally or dynamically. Either lateral inhibition is weak, owing to a sparsity of inhibitory interneurons during the filtering stage, or shunting arousal is large. The latter mechanism is a special case of the

self-tuning process in Section 20. It embodies a search procedure, or attentional mechanism, whereby the correct level of map tuning is stabilized. To illustrate the main idea, let input $I = \sum_{k=1}^n \alpha_k I_k$ be presented to $\mathcal{F}_S^{(1)}$, and suppose that $\alpha_1 > \alpha_2 > \dots > \alpha_m$. If arousal starts out low, then only $v_{S_1}^{(2)}$ might initially be stored in STM. Consequently $v_{S_1}^{(2)}$ releases a template to $\mathcal{F}_S^{(1)}$ that is proportional to I_1 . The mismatch between I_1 and I increases arousal at $\mathcal{F}_S^{(2)}$. The activity of $v_{S_1}^{(2)}$ starts to decrease owing to antagonistic rebound, and $v_{S_2}^{(2)}$ starts to become active. Then $v_{S_1}^{(2)}$ and $v_{S_2}^{(2)}$ both release templates to $\mathcal{F}_S^{(1)}$, and this hybrid template is approximately proportional to $\alpha_1 I_1 + \alpha_2 I_2$, which matches I better. Thus, if arousal increases again, the increment is smaller, and a hybrid template close to $\alpha_1 I_1 + \alpha_2 I_2 + \alpha_3 I_3$ is elicited. A few cycles of this reverberation can quickly retune $\mathcal{F}_S^{(2)}$ until the STM pattern across $\mathcal{F}_S^{(2)}$ generates a feedback template to $\mathcal{F}_S^{(1)}$ that almost matches I . The resonant STM pattern at $\mathcal{F}_S^{(2)}$ then automatically generates a motor pattern at $\mathcal{F}_M^{(2)}$ that approximately equals O . With O active at $\mathcal{F}_M^{(2)}$ while I is active at $\mathcal{F}_S^{(1)}$, coding and associative learning can gradually close the $I \rightarrow O \rightarrow I$ loop as codes for I and O are synthesized.

Many interesting developmental questions are posed by the concept of a self-tuning filtered map. For example, what keeps the reset mechanism from totally inhibiting $v_{S_1}^{(2)}$ before $v_{S_2}^{(2)}$ is excited? Is this due to the fact that the partial match of I_1 to I creates a small arousal increment, and of $\alpha_1 I_1 + \alpha_2 I_2$ to I an even smaller arousal increment, etc.? Or is $v_{S_2}^{(2)}$ quickly excited by arousal because of its prior subliminal excitation? Or is the rebound mechanism weak at this developmental stage?

The self-tuning process describes a type of multidimensional inference in a real-time parallel processor (Anderson, 1958). The successive switching-in, or reset, of channels $v_{S_i}^{(2)}$ is analogous to principal component analysis, or discriminant analysis, of a space-time pattern (Donchin and Herning, 1975). It is of particular interest that multidimensional techniques have successfully been used to analyze the P300, which is interpreted herein as an index of the amount of reset.

57. No Sensory Feedback Implies No Map Formation

Hein and Held (1967) and Held and Bauer (1967) have shown that, when young kittens and monkeys reach for objects without being able to see their hands, then no positional map develops. By contrast, if the eye can see where the hand goes, then a map does develop. This occurs in the model for a basic reason.

Let $\mathcal{F}_M^{(1)}$ be the field of terminal maps for the hand-arm system, and

let $\mathcal{F}_S^{(1)}$ be stage (Ib) of the eye-head system (Section 48). If no sensory feedback reaches $\mathcal{F}_S^{(1)}$, then there is no way to build the associative bridges to $\mathcal{F}_M^{(2)}$ across which the coefficients α_i can be mapped by signals. Without these associative bridges, when the monkey looks at a new position, the position is not filtered into combinations of old positions that have the correct relative signal sizes. Hence no spatial map exists.

58. Does the Psychophysical Power Law Influence Imitation Errors?

Section 17 describes the simplest case of the adaptive coding model for purposes of exposition. Where a power law transformation (Mountcastle, 1967; Stevens, 1961) controls the filtering signals, the size of the power influences whether spatially localized or diffuse patterns will be preferentially weighted.

To illustrate this, suppose as in Section 17 that $\mathcal{F}^{(1)}$ normalizes its patterns. Let the input pattern I be transformed into a normalized pattern $\theta = (\theta_1, \theta_2, \dots, \theta_n)$. Then $\sum_{k=1}^n \theta_k = 1$. In Section 17, the signal from $\mathcal{F}^{(1)}$ to $v_j^{(2)}$ was defined by $S_j = \sum_{k=1}^n \theta_k z_{kj}$. More generally, $S_j = \sum_{k=1}^n f(\theta_k) z_{kj}$, and correspondingly (19) is replaced by

$$z_{ij} = [-z_{ij} + f(\theta_i)] x_j^{(2)} \quad (52)$$

Introduce the notation $f(\theta) = (f(\theta_1), f(\theta_2), \dots, f(\theta_n))$. Then

$$S_j = \|f(\theta)\| \cdot \|z_j\| \cos [f(\theta), z_j] \quad (53)$$

Often $f(w)$ is a sigmoid function of w , as in Section 15. The sigmoid can be approximated by a power law at small values of w , say $f(w) \cong f_p(w) \equiv w^p$. The size of p influences what patterns will be coded by each $v_j^{(2)}$ as follows.

Unbiased coding occurs if $p = \frac{1}{2}$. In this case, the normalization condition $\sum_{k=1}^n \theta_k = 1$ implies $\|f_{1/2}(\theta)\| = 1$. Furthermore (52) implies that $\|z_j\|$ approaches 1, owing to developmental tuning. Thus, after tuning takes place,

$$S_j \cong \cos [f_{1/2}(\theta), z_j]$$

so that S_j is maximized by the pattern $f_{1/2}(\theta) = z_j$. If $p > \frac{1}{2}$, there exists a tendency to code spatially localized patterns, because $\|f_p(\theta)\|$ is maximized by any point pattern: $\theta_i = 1$ and $\theta_j = 0, j \neq i$. If $p < \frac{1}{2}$, there exists a tendency to code spatially diffuse patterns, because $\|f_p(\theta)\|$ is maximized by the uniform pattern $\theta_i = 1/n, i = 1, 2, \dots, n$. Thus if $p \neq \frac{1}{2}$,

the signal law (53) mixes two maximizing tendencies: maximize $\cos [f_p(\theta), z_j]$ by choosing $f_p(\theta)$ parallel to z_j ; and maximize $\|f_p(\theta)\|$.

Given m normalized patterns $\phi_j, j = 1, 2, \dots, m$, suppose that z_j adapts to $f_p(\phi_j)$, and that pattern θ can be written as $\theta = \sum_{k=1}^m \alpha_k f_p(\phi_k)$. If the vectors $f_p(\phi_j)$ are mutually orthogonal, then the signal

$$S_j = \theta \cdot z_j = \alpha_j \|f_p(\phi_j)\| = \alpha_j \sum_{k=1}^m \phi_{jk}^{2p}$$

where $\phi_j = (\phi_{j1}, \phi_{j2}, \dots, \phi_{jn})$. If $p = \frac{1}{2}$, then $S_j = \alpha_j$; hence the coefficients α_j are mapped without bias. However, if $p > \frac{1}{2}$, then populations $v_j^{(2)}$ are favored whose patterns ϕ_j are spatially localized; if $p < \frac{1}{2}$, then populations are favored whose patterns ϕ_j are spatially diffuse.

To test these effects experimentally will require a correlative analysis of coding at $\mathcal{F}_S^{(1)}$, of the (approximate) power laws of $\mathcal{F}_S^{(1)} \rightarrow \mathcal{F}_S^{(2)}$ and $\mathcal{F}_S^{(2)} \rightarrow \mathcal{F}_M^{(2)}$ signals, and of trends in imitation errors.

59. Rhythm and Phrasing

When a musician plays a piece, how are the relative velocities of each note controlled? More generally, many sequential sensory-motor skills are performed with a fixed order and a characteristic rhythm. Yet the rhythm can be modified without destroying the correct order. How is this flexible relationship between order and rhythm established?

Section 9 notes an example of this phenomenon wherein varying the size of rehearsal arousal through time can alter performance velocity in an avalanche. This mechanism is instructive, but is insufficient in general. Consider a sensory chunk whose template encodes a sequence of motor commands. Then the chunk must also encode performance velocity. How is this accomplished? The avalanche example, along with the discussion of instrumental conditioning in Section 7, suggests that the chunk sends a conditionable pathway to the arousal source. Then the amount of arousal, and hence performance velocity of the entire sequence, can be changed by conditioning. This mechanism allows order and velocity information to be decoupled; any other input to the arousal source can change the rhythm without changing the order. However, the mechanism only alters the overall performance velocity *within* a sequence. How is arousal calibrated so that fixed changes in arousal determine prescribed velocity changes in different sequences? In partic-

ular, how is the amount of arousal calibrated to compensate for differences in sequence length?

This problem can be restated as follows. To heighten intuition, call the items controlled by the given command a *phrase*. Is there a tendency to quantize time so that each phrase fills a unit time interval of performance? If this were so, then items coded in longer phrases would be performed faster, other things being equal, and variations in (conditioned) arousal level could shrink or expand this quantized time unit, and thereby increase or decrease performance velocity by fixed amounts. In the special case that phrases are words, time quantization is compatible with the existence of a breath pulse, or syllable (Lenneberg, 1967, p. 115).

The time quantization problem can be restated as a technical question. Given an STM buffer that contains k active items in a monotone decreasing pattern $x_1 > x_2 > \dots > x_k$, how fast are the items performed as a function of k , other things being equal? Opposite answers can be derived if the design of the buffer is changed. Suppose for definiteness that total STM activity in the buffer is normalized. Then as k increases, each item has smaller activity, other things being equal. Smaller activities can imply slower or faster performance velocities, depending on other factors. For example, in the avalanche of Section 5, smaller activities imply slower velocities. This is because the time needed for activity to achieve suprathreshold values at $v_i^{(2)}$, given a fixed arousal base line, is a monotone decreasing function of signal size from $v_i^{(2)}$. In this situation, the rate at which excitation grows is the dominant effect. Whenever the rate of feedforward excitation growth is rate-limiting, longer phrases imply slower performance velocities.

Suppose by contrast that all STM activities are already actively reverberating at asymptotic levels before arousal is turned on. Suppose that these STM activities perturb the network's output cells, but that the output cells cannot fire until they are aroused. Let arousal act quickly when it is turned on, and when an output cell begins to fire, let it immediately begin to inhibit its STM source. In this situation, the rate at which feedback inhibition acts is rate-limiting, since as soon as one STM source is inhibited, the next fires, and so on. If feedback inhibition is rate-limiting, then smaller STM activities are more rapidly quenched. Smaller activities then imply faster velocities, and longer phrases imply faster performance. In summary, feedforward excitation and feedback inhibition have opposite effects on performance velocity.

A simple case of the feedback inhibition phenomenon is illustrated below. Let feedback inhibition grow at a rate proportional to suprathres-

hold STM activity, and let it inhibit STM activity at a rate proportional to its size. Denote by $x(t)$ the STM activity at time t , by $y(t)$ the amount of feedback inhibition, and by Γ the QT . Then $\dot{x} = -Ay$ and $\dot{y} = B[x - \Gamma]^+$ where $[w]^+ = \max(w, 0)$. Given initial STM activity I at time $t = 0$, it follows that STM reaches the QT at time $T(I) = \arccos [(I - \Gamma)^{-1}]$, which is a monotone increasing function of I .*

Many questions are raised by the above observations. If phrasing is an important factor in velocity control, then the number of phrases allowed in the performance buffer at any time must be carefully regulated. How is switching on of the next phrase accomplished in a way that prevents discontinuities in performance? Two different types of sensory feedback are probably important: sensory feedback that turns on new spaced sensory chunks which thereupon reset the motor buffer (Section 34), and reset of terminal maps when proprioceptive-terminal matching occurs (Section 52). A careful analysis of special cases is clearly needed, however.

For example, suppose that the buffer starts to renormalize its total activity after a population is quenched by feedback inhibition. If renormalization acts slowly relative to feedback inhibition, then items near the end of each phrase will have the smallest activities when they are performed, and performance rate will speed up as the phrase is executed. If renormalization is fast relative to feedback inhibition, then items near the end of the phrase will have the largest activities when they are performed, and performance rate will slow down as the phrase is performed. If the two effects are balanced, then a uniform performance rate occurs. Is there a mechanism that automatically balances the two inhibitory effects to guarantee uniform performance rates in all cases? Such a mechanism would reduce the QT , or amplify STM, if the total buffer activity decreases. This will happen, for example, if nonspecific excitatory interneurons that are driven by total STM activity recurrently excite the off-surrounds of the STM buffer.

An entirely different kind of performance will occur if arousal is turned on only when matching between proprioceptive and terminal maps occurs. In this case, if the buffers can renormalize themselves faster than a terminal map is executed, then a uniform rate of performing "syllables" can be achieved (cf. Lenneberg, 1967, p. 115). By changing the relative balance between buffer reset by sensory feedback, arousal onset by feedback due to map matching, arousal onset by descending commands, feedforward excitation, feedback inhibition, and field normalization, one can change performance from item perseveration, to uniform rates, to rhythmical speeding up or slowing down, or to a wide

*Erratum: $T(I) = \pi / 2\sqrt{AB}$ is independent of I , which makes uniform phrase performance velocity easier to achieve.

range of phrase velocities. Are certain pathologies in speech production due to such changes of balance?

60. Reciprocal Intermodality Feedback, Internal Hearing, and Naming

This section lists some important implications of network mechanisms. One of them will be used to provide a unified explanation of recent data on serial versus parallel visual information processing in Section 61.

If two or more modalities are associatively related, then their effects on each other can be reciprocal. Figure 44 schematizes two examples. In Fig. 44a, $\mathcal{F}_S^{(1)}$ denotes an auditory field, and $\mathcal{F}_M^{(1)}$ denotes a field of motor commands for speech-related musculature. The other fields establish codes for their base fields in the usual way. Field $\mathcal{F}_S^{(1)}$ can be excited via at least two routes. Activating $\mathcal{F}_M^{(1)}$ can elicit sounds that excite $\mathcal{F}_S^{(1)}$ via sensory feedback. Learned template signals from $\mathcal{F}_S^{(2)}$ to $\mathcal{F}_S^{(1)}$ can also activate $\mathcal{F}_S^{(1)}$. This activation serves as a subliminal sensory expectation if $\mathcal{F}_S^{(1)}$ is not aroused. If $\mathcal{F}_S^{(1)}$ is aroused, then the subliminal pattern becomes supraliminal and creates the impression of

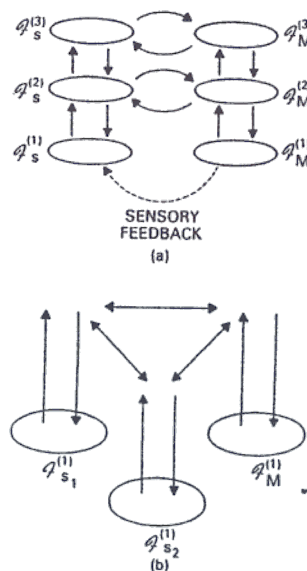


FIG. 44. Reciprocal intermodality associations

hearing an internally generated sound. Thus, motor activity at $\mathcal{F}_M^{(1)}$ is not needed to excite internal sounds at $\mathcal{F}_S^{(1)}$. As noted in Section 52, signals from $\mathcal{F}_S^{(3)}$ to $\mathcal{F}_M^{(3)}$ can generate anticipatory motor expectations, which ultimately elicit motor acts. Reciprocal signals from $\mathcal{F}_M^{(3)}$ to $\mathcal{F}_S^{(3)}$ can, in principle, generate anticipatory sensory expectations. In general, the several levels of associationally joined sensory and motor systems can mutually support each other's performance.

Figure 44b schematizes an important special case of reciprocal associations between the visual, auditory, and motor systems. $\mathcal{F}_{S_1}^{(1)}$ represents an auditory field, $\mathcal{F}_{S_2}^{(1)}$ represents a visual field, and $\mathcal{F}_M^{(1)}$ represents a motor field for speech-related musculature. Suppose that an object is visually presented to the network as its name is spoken aloud. The sequence of sounds at $\mathcal{F}_{S_1}^{(1)}$ can be imitated by using filtering properties from system S_1 to M . The sounds can hereby gradually generate a sequential motor code at $\mathcal{F}_M^{(3)}$ on successive trials. Simultaneously, the visual image of the object is coded at $\mathcal{F}_{S_2}^{(2)}$. The visual code can then sample the sequential motor code. Later, the visual image can elicit motor performance of the name. Furthermore, as the name is practiced, it can generate a sequential auditory code at $\mathcal{F}_{S_1}^{(3)}$. This code can sample the object's visual code, which in turn has learned a visual template at $\mathcal{F}_{S_2}^{(1)}$. Later, hearing the name can create a visual expectation of seeing the object. Similarly, seeing the object can create at $\mathcal{F}_{S_1}^{(1)}$ an auditory expectation of hearing its name. This expectation can be fulfilled either indirectly by arousing the motor commands of the name, or directly by arousing the auditory field.

The network constructions also admit hierarchical variations, as Section 24 implies. Figure 45 illustrates a hypothetical case in which a sensory field $\mathcal{F}_S^{(1)}$ excites several parallel hierarchies of adaptive codes and feedback templates. Each hierarchy discriminates ever-more-refined features of its base code. Field $\mathcal{F}_S^{(n+1)}$ interacts with all the hierarchies via adaptive codes and feedback templates. In such a system, a dominant feature in a sensory pattern at $\mathcal{F}_S^{(1)}$ can bias the entire hierarchy to expect a global ensemble of features that has often contained the dominant feature. This happens as follows. When a sensory event perturbs $\mathcal{F}_S^{(1)}$, it is coded by a pattern across $\mathcal{F}_S^{(2)}$. The features in this pattern are processed in two ways. They are projected directly onto $\mathcal{F}_S^{(n+1)}$. Here they excite the code that is closest to their pattern. This code, in turn, reciprocally excites the entire hierarchy via its template. Since the features computed at $\mathcal{F}_S^{(2)}$ are "simple," the feedback template can bias the field to ignore higher-order features that occur in an unfamiliar configuration. Signals between the hierarchy and $\mathcal{F}_S^{(n+1)}$ continue to reset each other until a consensus is reached.

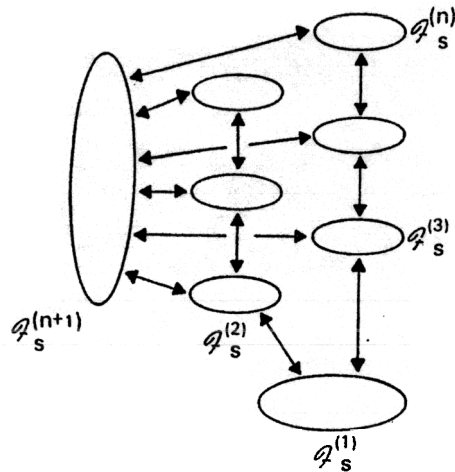


FIG. 45. The field $\mathcal{F}_s^{(n+1)}$ biases the entire hierarchy of fields to interpret patterns at $\mathcal{F}_s^{(n)}$ in terms of the "simple" features that are coded at $\mathcal{F}_s^{(1)}$.

These examples are set forth to emphasize a central problem for future theorizing: the classification of *which* features are computed by particular anatomical and physiological parameters.

61. Automatic versus Controlled Information Processing

Schneider and Shiffrin (1976) have described two complementary types of visual information processing in a series of interesting experiments. *Automatic* processing is said to be "a simultaneous, parallel, relatively independent detection process." *Controlled* processing is said to be a "serial terminating search" process. The authors argue that the two types of processing are associated with characteristic experimental paradigms, and that many earlier STM experiments about the serial or parallel nature of memory search can be classified into one or another of these paradigms. Below it is argued that both types of processing utilize common parallel operations, and that their apparent differences are due to shifts in the relative balance of these operations that are caused by experimental conditions. In particular, serial *properties* do not necessarily imply serial *operations*.

Both types of processing are studied by using a multiple frame search task. Four elements are presented simultaneously on a CRT screen. They are arranged in a square around a central fixation dot, and their

joint presentation over a brief time interval is called a *frame*. A trial consists of 20 frames presented in immediate succession at a prescribed rate. The subject's task is to detect one of several items in a *memory set* that is presented before each trial. Items that are not in the memory set are called *distractors*. Items that are neither distractors nor in the memory set are called *masks*.

Automatic processing occurs when the subject has practiced at giving a consistent detection response to memory set items that are never distractors—for example, detecting digits among letter distractors. This is also called a *consistent mapping* (CM) condition. Controlled processing occurs when memory set items and distractors are mixed from trial to trial—for example, detecting digits among digit distractors. This is also called a *varied mapping* (VM) condition. CM performance is much better than VM performance. During CM performance, there is almost no effect of varying the number of distractors in a frame, or of the memory set size; VM performance is monotonically related to each of these variables. Also, during CM performance, false alarms (detections when no target was present) increase substantially at fast frame speeds, but this does not occur during VM performance.

These data are used to conclude that during VM search serial comparisons are made by comparing all display items with a given memory item before a switch to the next memory item occurs. Also each switch to a new memory item takes some time. Data are reported to suggest that VM attentional deficits and search time are intimately related.

During CM search, it is assumed that "a mapping of stimuli to an internal detection or attention response can be learned in long-term memory. . . . Thus in long-term memory an automatic attention response to each target will be learned: the subject can simply wait for the occurrence of one of the learned attention responses . . . the target is always matched or compared first, before any distractors. . . ." Schneider and Shiffrin support this view by showing that automatic search can be learned, and that substantial negative transfer is produced if the memory set and distractor set are then interchanged. This negative transfer is attributed to the learning of an automatic attention response that continues to operate even after the memory set items are used as distractors. They also suggest that CM subjects carry out a controlled search after reversal, and that categorization may facilitate controlled search by reducing the effective memory set size. Results of LaBerge (1973) are also cited to indicate that the automatic response is analogous to the process whereby unknown and unexpected characters that are originally matched feature by feature are eventually matched as unitary

entities, or chunks, after they are repeatedly presented. More generally, their results support the dictum that "what is learned is what is attended," and that after automatic processing takes over at one level of behavioral organization, controlled processing can begin to organize the automatically processed behavioral units. This latter view is, in fact, the foundation on which the present theory of STM and LTM was originally constructed. In Grossberg (1969a), STM at previously coded behavioral units, which are represented by network vertices or nodes, drives associative LTM changes among these units until new units are thereby formed, whence the process repeats itself. Schneider and Shiffrin (1976) have a similar, albeit heuristic, model in mind when they write: "Suppose long-term store to consist of a collection of nodes that are associatively interrelated in a complex fashion. Each node may itself consist of a complex set of informational elements, including associative connections, programs for responses or actions, and directions for other types of information processing. The node is a distinguishable entity because it is unitized—when any of its elements are activated, all of them are activated."

The main effects found in the Shiffrin and Schneider data can be explained by the present theory. Consider VM search. Suppose that the memory set items are familiar. Then as they are read, they are recoded by their sequential auditory codes. This is an "automatic" process. If the memory set items are not familiar, they are filtered by those lower-order codes that are available. The explicit process of adaptive coding replaces the heuristic rubric of "unitization" in describing how these processes differ. Returning to the case of familiar items, an unfamiliar sequence of these items generates a spatial pattern of activity across the STM buffer of sequential auditory codes. This pattern stores order information in STM. As yet, there is no order information coded in LTM. A nonspecific rehearsal wave can read the items out of the STM buffer one at a time. This happens in a serial fashion (Section 28), and reset of the STM buffer takes some time, but the readout operation of nonspecific arousal is a *parallel* operation.

Suppose that a given item is read out of the auditory buffer into its visual code, where it forms a subliminal visual expectation. Then the items in a frame can be visually sampled until a match occurs. When this happens, a burst of activity from the visual code terminates search, as in Section 53. If no match occurs, a rehearsal wave can read the next item out of the auditory buffer, and so on.

Consider CM search. Repeated use of the same memory set gradually generates a higher-order auditory code that can sample the visual codes for all the items over successive trials. When the higher-order code is

activated, the visual codes of *all* memory set items can be subliminally activated. Matching with any one of these codes generates a resonant burst. The process therefore seems to be more parallel than VM search. I claim, however, that this is primarily because the higher-order code must be established before the visual codes of all memory set items can be sampled by a single internal representation. In other words, the auditory-to-visual *codes* and *templates* that are activated in VM and CM conditions are different, but the two conditions otherwise share common mechanisms.

Attention enters the search process in several ways. The simplest attentional reaction is amplification of network response to expected items. When a memory set item matches a subliminal visual expectation, a resonant visual response is generated. This type of attentional reaction occurs even under tachistoscopic conditions (cf. Berlyne, 1970). The amplification process can also move the eyes toward prescribed regions of the visual field by generating amplified feedback to the appropriate terminal eye coordinates if there is enough time to execute these motions. The attention theory in Grossberg (1975) suggests how other processes, such as incentive-motivational, CNV, and novelty-triggered feedback, can regulate the amount and pattern of STM amplification. The present theory supports, indeed refines, the dictum that "what is learned is what is attended." As Section 20 notes, the very stability of the STM code requires feedback processes that explicate attentional mechanisms.

Several other aspects of the Schneider and Shiffrin data are interesting. The "time for automatic search is at least as long as that for a very easy controlled search." This is paradoxical if CM search is a more efficient processing scheme. Is partial normalization of the visual template one reason for this? If more cues are subliminally active during CM than during VM search, then each cue will then have less subliminal activity. The reaction time for supraliminal signals to be generated during a match will then be greater during CM than during VM.

Also of interest are the data concerning performance accuracy when a memory set item occurs 0, 1, 2, or 3 frames away from an identical, or different, memory set item. During CM search, no performance decrement occurs if two distinct memory items are separated by any frame interval. This follows if the memory set visual templates are all subliminally active at once. Matching one item does not require reset to match a different item. However, if two identical items occur simultaneously, then the first match can interfere with registration of the second match by either of two mechanisms: if the match elicits performance and feedback inhibition of the item's STM activity, then the higher-order

code must be reactivated before the second match can occur; if the resonance due to the first match does not terminate before the second instance of the same item occurs, the resonant state cannot generate a distinguishable resonant burst. In summary, during CM search, two simultaneous distinct memory set items are recognized better than two simultaneous identical items.

The reverse is true during VM search. Here each display item is compared to the visual template of one memory set item at a time. Suppose that a display item matches the visual template. Then the item's visual code is amplified by resonance. This activates the automatic process of recoding the item auditorily. The visual resonance can excite the item's auditory code more strongly than would occur during casual reading of the item in a list of other items. Thus, a tendency exists to give the item maximal activity in the sequential auditory buffer. When the auditory buffer is then aroused, the item's visual template is reactivated. This reset effect makes it easier to detect two identical than two distinct memory set items in successive frames. Whether the items in successive frames are identical or distinct, the visual code is reset to detect the second item after a correct detection, and this takes time.

Similar reset effects can take place when an unexpected display item is processed. For example, suppose that an unexpected memory set item is visually scanned. Suppose that it is automatically recoded auditorily, and that its afferent auditory activity can summate with its activity in the sequential auditory buffer. Until this happens, the network has no way to tell that the display item is a memory set item. If summation does occur, then items with the largest prior STM activity should have the largest tendency to reset the visual code with their own template. This hypothesis should be tested experimentally. Analogously, during a VM search wherein CM items are used as distractors, a scanned CM item can activate its higher-order code, which then resets the visual template to expect CM items.

The above explanation of CM versus VM properties in terms of buffer reset can be tested by running the experiments again while also measuring P300. Also of interest is whether an expected memory set item can reset the auditory STM order more vigorously via visual resonance than an unexpected memory set item. This comparison might depend on subject strategies, since unexpected items can elicit a strong reset via the alarm system unless the alarm is already tuned by the search procedure to prevent this.

By explaining the Schneider and Shiffrin data in a unified way, we avoid several serious problems of their theory. They claim, and I agree, that automatic processing is used to rapidly code familiar behavioral

units so that controlled processing can then build these units into new unitized elements. I disagree that the "automatic attention response" in the CM condition is a mechanism that is qualitatively different from mechanisms operating in the VM condition. If the two types of conditions used serial versus parallel *operations*, as Shiffrin and Schneider claim, then how does the brain tirelessly alternate between serial and parallel mechanisms as it practices any new list of unitized elements? How do the serial and parallel processes compete when a visual scene contains both unitized and unfamiliar but relevant objects? How does the switchover from serial to parallel processing take place as an item is unitized? These problems evaporate in the present theoretical framework.

62. Visual versus Auditory Processing and Cerebral Dominance

A pattern of activity across a field of populations at a given time is inherently ambiguous. Does the pattern code a single event in time, such as the features in a visual picture, or does it code the order information in a series of events? Because of this fundamental ambiguity, two distinct types of STM reset mechanisms have been posited in the present theory: (1) deletion of population activity by feedback inhibition; (2) deletion of patterned activity across populations by mismatch with a comparison pattern. Type 1 has been used to explain various facts about order information. To use type 1, the cells in each population are allowed to be broadly distributed across the network, but their activities must be distributed close to a single average value. If the average activities of localized cell clusters in a given population are not approximately the same, there does not exist a simple index of order information between populations. Mechanisms of type 2 are then needed to delete intrapopulation activity.

Given a mechanism of type 1 in a field $\mathcal{F}^{(t)}$, the cells in the next field $\mathcal{F}^{(t+1)}$ automatically code order information in a manner that is sensitive to rehearsal strategies, as in Section 31. Thus, merely changing the inhibitory mechanism within $\mathcal{F}^{(t)}$ causes the adaptive coding mechanism from $\mathcal{F}^{(t)}$ to $\mathcal{F}^{(t+1)}$ to code data about time rather than space. To make this temporal code usable, individual populations in $\mathcal{F}^{(t)}$ must code the controlling features of entire behavioral events.

This latter constraint suggests a reason why visual data are often recoded auditorily to achieve an IMS of significant length. The data that are derived from a typical visual scene are of very high dimension, including colors, myriad shapes, distance information, etc. Coding all

the relevant dimensions of a typical visual object in a single population would require many stages of adaptive coding. Much simpler demands are placed on auditory coding, wherein many fewer dimensions covary in the sound spectrograms of simple sounds. Having coded such sounds in individual, albeit perhaps diffusely distributed, populations, order information among these codes is then readily learned by using STM reset by feedback inhibition. Thus the amount of data in a single perceptual frame (visual versus auditory) is traded off against the number of successive frames whose order can be coded.

A second distinction between visual and auditory coding emerges in the study of their circular reactions. The auditory-speech loop uses two modalities, each of which has low-dimensional codes—namely, sounds and speech motor acts. The visual system closes its loop with the motor systems of bodily position, but loop closure involves the motor system for moving the eyes as in Section 48, rather than scene analysis. The eye-neck-head system also has relatively simple codes. Order information among visual scenes can, however, be learned either if the codes for sequential eye movements sample visual representations, or if there does exist feedback inhibition of higher-order visual codes (“grand-mother cells”), say in the inferotemporal cortex (Rocha-Miranda *et al.*, 1975).

The above remarks note that an activity pattern per se across a field of populations has an inherently ambiguous interpretation. Does it code data about time or data about space? In order to unambiguously decode temporal versus spatial data, somehow the populations that code the different types of data must be spatially segregated. The patterns themselves do not suffice to make this distinction; rather, the nature of their reset mechanisms accomplishes this. The ambiguity problem therefore suggests the need to spatially segregate the processing of sequential, including language-like, codes from codes that concern themselves primarily with spatial integration. Perhaps this dichotomy is one reason for the emergence of cerebral dominance (Gazzaniga, 1970, Chapter 8). Since visual and auditory representations are bilateral, the trend toward segregation of temporal versus spatial processing in separate hemispheres would be superimposed on relatively localized spatial and temporal processors in each hemisphere; that is, there might exist a subtle symmetry-breaking due to a drift of visual-like processing into the nondominant hemisphere and auditory-like processing into the dominant hemisphere.

The theory thus contains a tantalizing question as it stands. Does the ambiguity problem necessarily lead to distinct type 1 and type 2 mechanisms? If so, can anatomical traces of a type 1 mechanism (for

example, negative feedback loops triggered by the output stage and feeding into the STM reverberation) be found in sequential STM buffers but not in spatial representations? If not, does there exist *in vivo* a unified mechanism—say a modified type 2 mechanism—that possesses both types of properties?

63. Concluding Remarks: Universal Adaptive Measurement

This section sketches a broader perspective within which to view the above results. Its central tenet is that the brain is a universal measurement device operating on the quantum level. By this is meant that data from all perceivable physical fields are translated into a common neural language, and that events on the quantum level, such as several photons, can be perceived. The universality of the neural language clarifies why results concerning the neural measurement process can have broad interdisciplinary implications.

A central result of the present theory is a description of an alternative to the probabilistic and computer memory models that have been used to explain cognitive data. In particular, probabilistic models are replaced by systems that undergo parallel interactions in real time. Why the formalism of probability theory works at all in describing physical processes is a nontrivial problem that is often overlooked because of the practical successes of probability models. In the present scheme, many probabilistic-like computations are described by competitive interactions among network codes. The universal problem of processing patterned data in noisy systems with finitely many sites requires the existence of such competitive interactions. Furthermore, the general problem of stabilizing adaptive codes in a fluctuating input environment requires that certain feedback relationships exist between the codes of individual events and the codes of various event combinations. Are such universal problems and their solutions by competitive systems one reason for the success of probability models? How generally can a more powerful alternative to probability theory be built up by using hierarchically organized competitive systems operating in real time? Especially in cases in which a system continually re-evaluates hypotheses based on disconfirming feedback does the present framework seem to be intrinsically richer than probability theory.

A related set of problems arises in the serial processing of lists. The spatial geometry of a list of events, represented as symbols on a serial tape, is not the same as the space-time geometry of the same list of events occurring in real time (Grossberg, 1969d). Indeed, suppose that a

list r_1, r_2, \dots, r_n of events is presented to a subject with one item occurring every w time units. It is not until at least w time units after r_n occurs that the subject can know that r_n is the last list item. Only then can this past event be reclassified, via a "backward effect in time," as being the list's end. This fact implies that the types of properties and paradoxes that can occur in formal systems, such as classical logic or model theory, and in real-time parallel systems can be quite different. The results on serial learning in Section 12, and the real-time probabilistic logic across a field of populations in Section 20, provide two examples of how the approaches differ. Problems concerning the field representations of mathematical versus empirical data, of infinite operations in networks with finite numbers of coded populations, and of plans for which no digital algorithm exists are among the many that are worth investigating.

The evolutionary properties of the brain's measurement process suggest another class of problems. One of the triumphs of modern physics was to geometrize the dynamics of physical laws, as in Einstein's general relativity theory. One of the important tasks of brain theory is to reverse this procedure; namely, to explain the four-dimensional geometry of the world in terms of a dynamical system operating in a non-Euclidean network of very high dimension. As noted in Section 1.F, the high dimension of unfamiliar behavioral data seems to be successively reduced as new codes and commands for organizing this data evolve. The sections on hierarchical coding and map formation begin to show how these lower-dimensional representations emerge.

For almost a century, the measurement problems that concerned physics and biology diverged. Before that time, distinguished physicists, such as Helmholtz, Mach, and Maxwell, were also distinguished psychologists or physiologists. This then ceased to be true if only because profound insights concerning the measurement processes of physics could still be expressed by using the available linear mathematics, whereas it became clear that psychophysiological measurement processes involved nonlinear systems whose laws, and underlying principles, were at best dimly understood. Recently, both physics and biology have been driven toward processes in which nonlinear collective effects have been implicated. Indeed, analogs of such currently interesting physical phenomena as phase transitions (Grossberg, 1969f, 1974), globally irreversible but locally reversible interactions (Grossberg, 1969f), and backward effects in time (Grossberg, 1969d, 1974) are found even in simple neural networks. As both physical and biological theory incorporate measurement concepts that are explicated by parallel, nonlinear, self-organizing, hierarchical, and feedback interactions, we

can anticipate a renewed flourishing of interdisciplinary studies and a deepening understanding of our interactions with the external world.

Appendix

Proof of Theorem 1. The proof is by induction. For $i = 2$, (29) and (30) imply

$$\mu(1 + \omega_2) = \mu\lambda + (1 - \lambda)M$$

or

$$\omega_2 = \lambda + (1 - \lambda)R - 1$$

as in (31). For $i > 2$, (29) and (30) imply

$$\mu[1 + \omega_i(1 + \omega_{i-1}(1 + \omega_{i-2}(\dots)))] = \mu\lambda^{i-1} + (1 - \lambda^{i-1})M$$

or

$$\omega_i(1 + \omega_{i-1}(1 + \omega_{i-2}(\dots))) = \lambda^{i-1} + (1 - \lambda^{i-1})R - 1 \quad (A1)$$

By the induction hypothesis, (31) can be used for all indices less than i , whence

$$1 + \omega_{i-1}(1 + \omega_{i-2}(\dots)) = \lambda^{i-2} + (1 - \lambda^{i-2})R \quad (A2)$$

which along with (A1) proves (31) for all i .

Equation (32) shows that, for $2 \leq k \leq j$, $x_{k-1,j} > x_{kj}$ if and only if $\omega_k > 1$. By (31), $\omega_k > 1$ if and only if

$$(R - 1)(1 - \lambda)\lambda^{k-2} > 1$$

Since $0 < \lambda < 1$, $x_{1j} > x_{2j} > x_{3j} > \dots > x_{jj}$ unless $j > J$, as (33) notes.

Proof of Theorem 2. By the Invariance Principle, the STM activities across $\mathcal{F}^{(1)}$ at successive times can be described by the rows

$$\begin{array}{ccc} \mu & & \\ \omega_2 \mu & \mu & \\ \omega_3 \omega_2 \mu & \omega_3 \mu & \mu \end{array}$$

and so on. Let the sampling signals from a given population in $\mathcal{F}^{(3)}$ to the successive rows be s_1, s_2, s_3, \dots . By hypothesis, $s_1 \geq s_2 \geq s_3 \geq \dots$. The product of sampling signal and STM trace determines LTM growth in each time frame, as in (2). These products are

$$\begin{array}{ccc} s_1 \mu & & \\ s_2 \omega_2 \mu & s_2 \mu & \\ s_3 \omega_3 \omega_2 \mu & s_3 \omega_3 \mu & s_3 \mu \end{array} \quad (A3)$$

Within each row, the entry in the i th column is at least as large as the

entry in the $(i + 1)$ st column, by the TMS hypothesis. Hence the $\mathcal{F}^{(2)}$ population samples a monotone decreasing pattern in every time frame, and then sums all the patterns to learn a net monotone decreasing pattern.

Inequalities (34) are discussed in the next proof.

Proof of Theorem 3. Deleting the superscripts (1) in the invariant parameters for simplicity, we find the chart

$$\begin{array}{ccc} s_1 \mu_1 & & \\ s_2 \omega_2 \mu_1 & s_2 \mu_2 & \\ s_3 \omega_3 \omega_2 \mu_1 & s_3 \omega_3 \mu_2 & s_3 \mu_3 \end{array} \quad (\text{A4})$$

By (35), each row is either monotone decreasing (if all ω_i in the row exceed 1), monotone increasing (if all ω_i in the row are less than 1), or unimodal (if some ω_i fall above and below 1). If a population in $\mathcal{F}^{(2)}$ starts sampling $\mathcal{F}^{(1)}$ when a given STM pattern is active, this pattern is encoded in its LTM traces. Because the STM pattern in the past field does not change, the same past STM pattern is encoded into LTM in every time frame. The total past field LTM pattern is the sum of these STM patterns, and hence has the same form as it had during its first sampling interval.

The future field LTM pattern is monotone decreasing because, by (35) and (36), $s_1 \mu_1 > s_2 \mu_2 > s_3 \mu_3 \dots$, $s_2 \omega_2 \mu_1 > s_3 \omega_3 \mu_2 > s_4 \omega_4 \mu_3 > \dots$, etc. That is, the inequalities (34) hold. The sum of STM values in the i th column of (A4) thus exceeds the sum of STM values in the $(i + 1)$ st column. Since the future field LTM pattern sums up column values for all columns that are first excited after it begins to sample, this LTM pattern is monotone decreasing.

REFERENCES

- Anderson, T. W. (1958). "An Introduction to Multivariate Statistical Analysis." Wiley, New York.
- Atkinson, R. C., and Shiffrin, R. M. (1968). *Adv. Psychol. Learning Motiv.* 2, 89.
- Atkinson, R. C., and Shiffrin, R. M. (1971). *Sci. Am.* p. 82, August.
- Berlyne, D. E. (1970). In "Attention: Contemporary Theory and Analysis" (D. E. Mostofsky, ed.), p. 25. Appleton, New York.
- Bizzi, E., Polit, A., and Morasso, P. (1975). *J. Neurophysiol.*
- Bjork, R. A. (1975). In "Cognitive Theory (F. Restle et al., eds.), Vol. 1, p. 151. L. Erlbaum Assoc., Hillsdale, New Jersey.
- Burns, B. D. (1958). "The Mammalian Cerebral Cortex." Arnold, London.
- Cant, B. R., and Bickford, R. G. (1967). *Electroencephalogr. Clin. Neurophysiol.* 23, 594.
- Chung, S.-H., Raymond, S. A., and Lettvin, J. Y. (1970). *Brain, Behav. Evol.* 3, 72.
- Cornsweet, T. N. (1970). "Visual Perception." Academic Press, New York.
- Craik, F. I. M., and Jacoby, L. L. (1975). In "Cognitive Theory" (F. Restle et al., eds.), Vol. 1, p. 173. L. Erlbaum Assoc., Hillsdale, New Jersey.
- Crosby, E. C., Humphrey, T., and Lauer, E. W. (1962). "Correlative Anatomy of the Nervous System." Macmillan, New York.
- Dethier, V. G. (1968). "Physiology of Insect Senses." Methuen, London.
- Donchin, E., and Hering, R. I. (1975). *Electroencephalogr. Clin. Neurophysiol.* 38, 51.
- Donchin, E., Gerbrandt, L. A., Leifer, L., and Tucker, L. (1972). *Psychophysiology* 9, 178.
- Donchin, E., Tueting, P., Ritter, W., Kutas, M., and Heffley, E. (1975). *Electroencephalogr. Clin. Neurophysiol.* 38, 1.
- Duda, R. O., and Hart, P. E. (1973). "Pattern Classification and Scene Analysis." Wiley, New York.
- Ellias, S. A., and Grossberg, S. (1975). *Biol. Cybernet.* 20, 69.
- Estes, W. K. (1972). In "Coding Processes in Human Memory" (A. W. Melton and E. Martin, eds.), p. 161. Holt, New York.
- Fry, D. B. (1966). In "The Genesis of Language" (F. Smith and G. A. Miller, eds.), p. 187. MIT Press, Cambridge, Massachusetts.
- Gazzaniga, M. S. (1970). "The Bisected Brain." Appleton, New York.
- Gibson, J. J. (1933). *J. Exp. Psychol.* 16, 1.
- Goldberg, M. E., and Wurtz, R. H. (1972a). *J. Neurophysiol.* 35, 542.
- Goldberg, M. E., and Wurtz, R. H. (1972b). *J. Neurophysiol.* 35, 560.
- Graham, C. H. (1966). "Vision and Visual Perception" (C. H. Graham et al., eds.), p. 548. Wiley, New York.
- Granit, R., ed. (1966). "Muscular Afferents and Motor Control." Wiley, New York.
- Grillner, S. (1969). *Acta Physiol. Scand., Suppl.* 327.
- Grossberg, S. (1967). *Proc. Natl. Acad. Sci. U.S.A.* 58, 1329.
- Grossberg, S. (1969a). *J. Math. Psychol.* 6, 209.
- Grossberg, S. (1969b). *J. Statist. Phys.* 1, 319.
- Grossberg, S. (1969c). *J. Math. Mech.* 19, 53.
- Grossberg, S. (1969d). *Math. Biosci.* 4, 201.
- Grossberg, S. (1969e). *J. Theor. Biol.* 22, 325.
- Grossberg, S. (1969f). *J. Differ. Eq.* 5, 531.
- Grossberg, S. (1970a). *Stud. Appl. Math.* 49, 135.
- Grossberg, S. (1970b). *J. Theor. Biol.* 27, 291.
- Grossberg, S. (1971a). *Proc. Natl. Acad. Sci. U.S.A.* 68, 828.
- Grossberg, S. (1971b). *J. Theor. Biol.* 33, 225.
- 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, 217.
- Grossberg, S. (1974). *Prog. Theor. Biol.* 3, 51.
- 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). In "Formal Theories of Visual Perception" (E. L. J. Leeuwenberg and H. Buffart, eds.), Wiley, New York.
- Grossberg, S. (1977b). *J. Math. Biol.* 4, 237.
- Grossberg, S. (1977c). *J. Math. Anal. Appl.* (in press).
- Grossberg, S. (1978a). This volume.

- Grossberg, S. (1978b). *J. Math. Psychol.* (in press).
- Grossberg, S., and Levine, D. S. (1975). *J. Theor. Biol.* 53, 341.
- Grossberg, S., and Pepe, J. (1971). *J. Statist. Phys.* 3, 95.
- Gustafson, T., and Wolpert, L. (1967). *Biol. Rev. (Cambridge Philos. Soc.)* 42, 442.
- Hein, A., and Held, R. (1967). *Science* 158, 390.
- Held, R., and Bauer, J. A., Jr. (1967). *Science* 155, 718.
- Held, R., and Hein, A. (1963). *J. Comp. Physiol. Psychol.* 56, 872.
- Hodgkin, A. L. (1964). "The Conduction of the Nervous Impulse." Thomas, Springfield, Illinois.
- Hogan, R. M. (1975). *Mem. Cognit.* 3, 197.
- Hogan, R. M., and Hogan, M. M. (1975). *Mem. Cognit.* 3, 210.
- Hubel, D. H., and Wiesel, T. N. (1962). *J. Physiol.* 160, 106.
- Hubel, D. H., and Wiesel, T. N. (1963). *J. Neurophysiol.* 26, 994.
- Irwin, D. A., Rebert, C. S., McAdam, D. W., and Knott, J. R. (1966). *Electroencephalogr. Clin. Neurophysiol.* 21, 412.
- John, E. R. (1966). In "Frontiers in Physiological Psychology" (R. W. Russell, ed.), p. 149. Academic Press, New York.
- John, E. R. (1967). In "The Neurosciences: A Study Program" (G. C. Quarten, T. Melnechuk and F. O. Schmitt, eds.), p. 690. Rockefeller Univ. Press, New York.
- Julesz, B. (1971). "Foundations of Cyclopean Perception." Univ. of Chicago Press, Chicago, Illinois.
- Kahneman, D., and Beatty, J. (1966). *Science* 154, 1583.
- Kennedy, D. (1968). In "Physiological and Biochemical Aspects of Nervous Integration" (F. O. Carlson, ed.), p. 285. Prentice-Hall, Englewood Cliffs, New Jersey.
- LaBerge, D. (1973). *Mem. Cognit.* 1, 268.
- Lashley, K. S. (1951). In "Cerebral Mechanisms in Behavior: The Hixon Symposium" (L. P. Jeffress, ed.), p. 112. Wiley, New York.
- Lenneberg, E. H. (1967). "Biological Foundations of Language." Wiley, New York.
- Levine, D. S., and Grossberg, S. (1976). *J. Theor. Biol.* 61, 477.
- Low, M. D., Borda, R. R., Frost, J. D., and Kellaway, R. (1966). *Neurology* 16, 771.
- McAdam, D. W. (1969). *Electroencephalogr. Clin. Neurophysiol.* 26, 216.
- McAdam, D. W., Irwin, D. A., Rebert, C. S., and Knott, J. R. (1966). *Electroencephalogr. Clin. Neurophysiol.* 21, 194.
- Miller, G. A. (1956). *Psychol. Rev.* 63, 81.
- Mountcastle, V. B. (1967). In "The Neurosciences: A Study Program" (G. C. Quarten, T. Melnechuk, and F. O. Schmitt, eds.), p. 393. Rockefeller Univ. Press, New York.
- Parzen, E. (1960). "Modern Probability Theory and Its Applications." Wiley, New York.
- Piaget, J. (1963). "The Origins of Intelligence in Children." Norton, New York.
- Remington, R. J. (1969). *J. Exp. Psychol.* 82, 250.
- Robinson, D. A. (1964). *J. Physiol. (London)* 174, 245.
- Rocha-Miranda, C. E., Bender, D. B., Gross, C. G., and Mishken, M. (1975). *J. Neurophysiol.* 38, 475.
- Rohrbaugh, J. W., Donchin, E., and Eriksen, C. W. (1974). *Percept. Psychophys.* 15, 368.
- Schneider, W., and Shiffrin, R. M. (1976). In "Basic Processes in Reading: Perception and Comprehension" (D. LaBerge and S. J. Samuels, eds.). L. Erlbaum Assoc., Hillsdale, New Jersey.
- Shiffrin, R. M. (1975). In "Cognitive Theory" (F. Restle et al., eds.), Vol. 1, p. 193. L. Erlbaum Assoc., Hillsdale, New Jersey.
- Squires, K. C., Wickens, C., Squires, N. K., and Donchin, E. (1976). *Science* 193, 1142.
- Stein, P. S. G. (1971). *J. Neurophysiol.* 34, 310.
- Sternberg, S. (1966). *Science* 153, 652.
- Stevens, S. S. (1961). In "Sensory Communication" (W. A. Rosenblith, ed.), p. 1. MIT Press, Cambridge, Massachusetts.
- Stryker, M. P., and Schiller, P. H. (1975). *Exp. Brain Res.* 23, 103.
- Stryker, M. P., and Sherk, H. (1975). *Science* 190, 904.
- Thomas, G. B., Jr. (1968). "Calculus and Analytic Geometry." 4th ed. Addison-Wesley, Reading Park, Massachusetts.
- Thompson, R. F. (1967). "Foundations of Physiological Psychology." Harper, New York.
- Townsend, J. T. (1974). In "Human Information Processing: Tutorials in Performance and Cognition" (B. H. Kantowitz, ed.), p. 133. Erlbaum, Potomac, Maryland.
- Walter, W. G. (1964). *Arch. Psychiatr. Nervenkr.* 206, 309.
- Werblin, F. S. (1971). *J. Neurophysiol.* 34, 228.
- Willows, A. O. D. (1968). In "Physiological and Biochemical Aspects of Nervous Integration" (F. O. Carlson, ed.), p. 217. Prentice-Hall, Englewood Cliffs, New Jersey.
- Wise, C. D., Berger, B. D., and Stein, L. (1973). *Biol. Psychiatry* 6, 1.
- Wurtz, R. H., and Goldberg, M. E. (1972a). *J. Neurophysiol.* 35, 575.
- Wurtz, R. H., and Goldberg, M. E. (1972b). *J. Neurophysiol.* 35, 587.
- Yarbus, A. L. (1967). "Eye Movements and Vision." Plenum, New York.