In P. Schuster and W. Singer (Eds.), Nonlinear Dynamics and Neuronal Networks.
New York: Springer-Verlag, 1991.

DISTRIBUTED HYPOTHESIS TESTING, ATTENTION SHIFTS, AND TRANSMITTER DYNAMICS DURING THE SELF-ORGANIZATION OF BRAIN RECOGNITION CODES

Gail A. Carpenter† and Stephen Grossberg‡
Center for Adaptive Systems
and
Graduate Program in Cognitive and Neural Systems
Boston University
111 Cummington Street
Boston, Massachusetts 02215 USA

1. ART: A Biological Theory of Autonomous Real-Time Learning about a Changing World

How the mammalian brain can rapidly but stably learn about a changing world filled with unexpected events is one of the most challenging scientific problems of our time. The brain's ability to autonomously discover and learn appropriate representations of the world in real-time, without the intervention of an external teacher to signal that external changes have occurred or the nature of these changes, lies at the heart of this problem. Adaptive Resonance Theory, or ART, was introduced in 1976 (Grossberg, 1976a, 1976b) in order to analyse how brain networks can autonomously learn about a changing world in a rapid but stable fashion. Popular alternative models, such as back propagation, can learn only slowly, in an off-line setting, about an essentially stationary environment that includes an external teacher whose explicitly coded answers drive learning using non-local operations that seem to have no biological analog (Carpenter, 1989; Grossberg, 1988b; Parker, 1982; Rumelhart, Hinton, and Williams, 1986; Werbos, 1974, 1982). The present chapter summarizes some recent results concerning how ART systems control distributed hypothesis testing and memory search in order to autonomously discover and learn predictive representations for recognition and recall.

---

† Supported in part by British Petroleum (89-A-1204), DARPA (AFOSR 90-0083), and the National Science Foundation (NSF IRI-90-00530).

‡ Supported in part by the Air Force Office of Scientific Research (AFOSR 90-0175 and AFOSR 90-0128), the Army Research Office (ARO DAAL-03-88-K0088), and DARPA (AFOSR 90-0083).

Acknowledgements: The authors wish to thank Cynthia E. Bradford, Carol Yanakakis Jefferson, and Diana Meyers for their valuable assistance in the preparation of the manuscript.

2. ART Predictions about Neural Information Processing and Learning

One classical test of a physical theory's promise is its ability to generate unifying explanations of paradoxical data and predictions of new data. In an interdisciplinary field such as the behavioral and brain sciences, where scientists use such different tools and have such different training, there does not exist any single intellectual community to rapidly assimilate and critically test interdisciplinary predictions. Notwithstanding these problems of communication, key ART predictions have received accumulating experimental support over the years.

Predictions of particular interest to the present chapter include the following: In 1972, it was predicted that norepinephrine and acetylcholine jointly control brain plasticity during learning of reinforcement and recognition codes (Grossberg, 1972b). In 1976, this analysis was extended to predict that norepinephrine and acetylcholine control cortical plasticity during the critical period in visual cortex (Grossberg, 1976b). Several experimental studies have described relevant data (Bear and Singer, 1986; Kasamatsu and Pettigrew, 1976; Pettigrew and Kasamatsu, 1978). A role for attention in the regulation of cortical plasticity was also predicted (Grossberg, 1976b) and subsequently reported (Singer, 1982). Standing waves of resonant cortical activity were predicted to subserve these cortical dynamics (Grossberg, 1976b; 1978). Several labs have recently explored the role of such resonant standing waves (Eckhorn *et al.*, 1988; and Gray *et al.*, 1989). These theoretical results built upon earlier predictions that synaptic plasticity is controlled by processes in which an inward Na^+ current and an outward K^+ current interact synergetically with an inward Ca^{++} current that competes with an Mg^{++} current (Grossberg, 1968; 1969a). Recent data about the role of NMDA receptors have refined contemporary understanding of such synergetic interactions (Kleinschmidt, Bear, and Singer, 1987). An associative learning law was introduced in which synaptic efficacy is gated by postsynaptic activity such that, with the learning gate open, synaptic strength can either increase or decrease (Grossberg, 1969b, 1976a, 1978). Such a gated learning law has since been reported in visual cortex and hippocampus (Levy, 1985; Levy, Brassel, and Moore, 1983; Levy and Desmond, 1985; Rauschecker and Singer, 1979; Singer, 1983). It is the basic learning law used in the ART models.

A top-down template matching event that regulates selective attention was derived in Grossberg (1976b, 1978). It has the properties of the Processing Negativity event-related potential that was reported by Näätänen, Gaillard, and Mäntysalo (1978). See also Näätänen (1982). A hippocampal generator of the P300 event-related potential, as distinct from possible neocortical generators, was predicted in Grossberg (1980, p.25) from an analysis of how short term memory is reset by novel events. A hippocampal P300 generator was experimentally reported in Halgren *et al.*, (1980). An analysis of how the emotional meaning of cues modulates attention led to a complementary prediction. Grossberg (1975) predicted that both negative and positive emotions generate positive attentional feedback. Experimental support for this prediction was found by Bower (1981) and Bower, Gilligan, and Monteiro (1981). The pathway subserving this attentional feedback process, called incentive motivation, was interpreted in Grossberg (1975) as a pathway from hippocampus to neocortex. Such a pathway was discovered by Rosene and Hoesen (1977).

A reciprocal pathway for associative learning of conditioned reinforcers was predicted to pass from neocortical sensory representations to hippocampal pyramidal cells (Grossberg, 1971, 1972a, 1972b, 1975). Experimental evidence was reported by Berger and Thompson (1978) who first interpreted their results as the discovery of a general neural "engram." Subsequent experiments considered the effects of selective ablations on learning both in hippocampus and in cerebellum, leading to the conclusion that hippocampal learning is indeed a variant of the predicted conditioned reinforcer learning, whereas the cerebellum carries out a type of motor learning (Thompson *et al.*, 1984). These experimental results concerning differences between reinforcement learning and motor learning are related to a prediction concerning the control of motivated instrumental behavior; namely, that hippocampal processes bifurcate into the aforementioned positive attentional feedback pathway to neocortex and into a motivationally signed motor mapping subsystem for control of approach and avoidance behavior (Grossberg, 1975). Experimental evidence for spatial mapping properties of the hippocampus was described by O'Keefe and Nadel (1978). A reciprocal cortico-hippocampal interaction between conditioned reinforcers and incentive motivational sources was also suggested in Grossberg (1975), along with the implication that the hippocampus mediates stimulus-reinforcement contingencies whose mismatch with sensory processing in the cortex prevents read-out of cortical commands. Gabriel, Foster, Orona, Saltwick, and Stanton (1980) have reported compatible data.

A study of how reinforcing cues are forgotten, or extinguished, led to a network design in which opponent processes are gated by slowly varying chemical transmitters (Grossberg, 1972b). These opponent processes were interpreted in terms of the dynamics of hypothalamus and medial forebrain bundle, and the chemical transmitters were interpreted to be catecholaminergic. A mathematical study of these opponent processes led to the discovery in Grossberg (1972b) of a formal behavioral syndrome wherein catecholaminergic underarousal could cause an elevated behavioral threshold to coexist with suprathreshold hypersensitivity. Moreover, an arousing drug could transform this underaroused syndrome

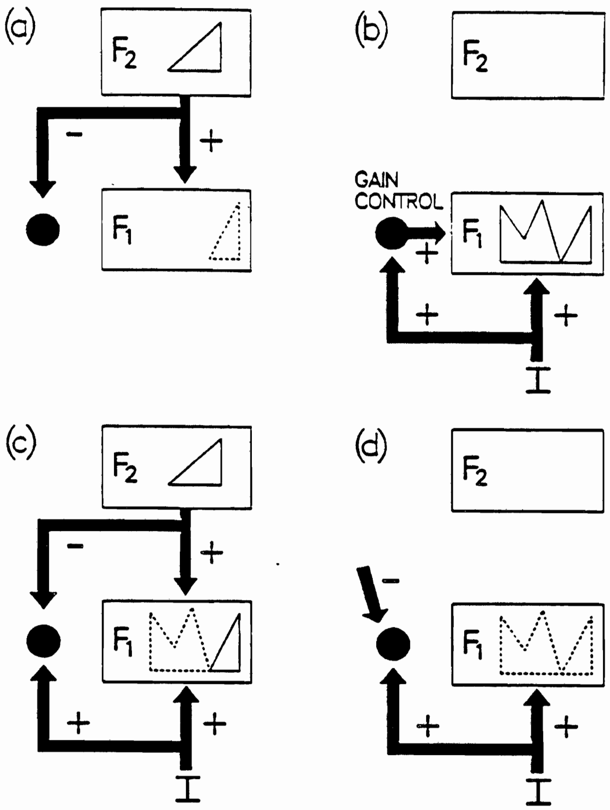


Figure 1. Matching by the 2/3 Rule in ART 1: (a) A top-down expectation from F_2 inhibits the attentional gain control source as it subliminally primes target F_1 cells. Dotted outline depicts primed activation pattern. (b) Only F_1 cells that receive bottom-up inputs and gain control signals can become supraliminally active. (c) When a bottom-up input pattern and a top-down template are simultaneously active, only those F_1 cells that receive inputs from both sources can become supraliminally active. (d) Intermodality inhibition can shut off the F_1 gain control source and thereby prevent a bottom-up input from supraliminally activating F_1, as when attention shifts to a different input channel. Similarly, disinhibition of the F_1 gain control source in (a) may cause a top-down prime to become supraliminal, as during an internally willed fantasy.

into an overaroused syndrome by moving the system over an inverted U whose peak corresponds to normal sensitivity. The overaroused syndrome has formal emotional properties symptomatic of certain schizophrenias. A similar underaroused syndrome in hyperactive children has been described by Shaywitz, Cohen, and Bowers (1977) and Shekim, DeKirmenjian, and Chapel (1977), notably the elevated threshold (Weber and Sulzbacher, 1975). Amphetamine is an arousing therapeutic drug which improves symptoms in small enough quantities (Swanson and Kinsbourne, 1976; Weiss and Hechtmann, 1979), but which is capable of causing schizophrenic syndromes in large enough doses (Ellinwood and Kilbey, 1980; MacLennan and Maier, 1983).

This partial list of predictions and supportive data illustrates the difficulty of understanding and evaluating interdisciplinary brain theories and their experimental implications from the perspective of traditional specialties. Simultaneous analysis of multiple levels of brain organization is characteristic of the theoretical development of neural architectures. Experimental evidence from multiple levels of brain organization is thus needed, from biochemical data about membrane channels and neurotransmitters, to emergent network properties about resonance and attention. Organizational principles and their mechanistic realizations need to simultaneously satisfy constraints on multiple levels of design in order to rule out the many hypotheses which seem plausible in the light of a limited set of data, but are incompatible with more global theoretical constraints. Fortunately, the sociological organization of the brain sciences is rapidly adapting to meet the intellectual challenge posed by the explanations and predictions of interdisciplinary brain theories, although experimental tests are still often reported either without knowledge or citation of their theoretical antecedents. As the infrastructure needed for cooperative work between theorists and experimentalists matures, more efficient and critical tests of theories can be anticipated.

Such a development is much to be desired, since many theoretical predictions have not yet been tested at all. One such neurophysiologically untested ART prediction is that cortical recognition codes are regulated by a matching law called the 2/3 Rule (Carpenter and Grossberg, 1987a). This prediction, as well as other neurobiologically testable predictions, are described below.

3. Control of Attention, Hypothesis Testing, and Memory Search by 2/3 Rule Matching

In its most general form, the 2/3 Rule summarizes a type of matching in which the effects of converging bottom-up and top-down excitatory signals are modulated by a third, inhibitory process. The 2/3 Rule was derived from one of the central ideas of ART; namely, that learning of top-down expectations, and focusing of attention upon particular featural groupings, are mechanisms that help to control rapid learning about novel behavioral events, without causing unselective forgetting of prior memories that are still behaviorally useful. A system with these properties resolves the *stability-plasticity dilemma* (Grossberg, 1982). The 2/3 Rule explains how a top-down expectation can attentionally *prime* a network to anticipate bottom-up data that may or may not occur (Figure 1). The priming state subliminally sensitizes the network to be ready for an expected event, without forcing the network to generate a full-blown, attentionally focussed resonance before the event actually occurs. Thus 2/3 Rule matching converts an ART system into an *intentional* system capable of anticipating events before they actually occur. As indicated in Figure 1c, this intentionality property implies that matching of a top-down expectation with bottom-up data carries out a type of analog spatial logic. Thus, within ART, aspects of

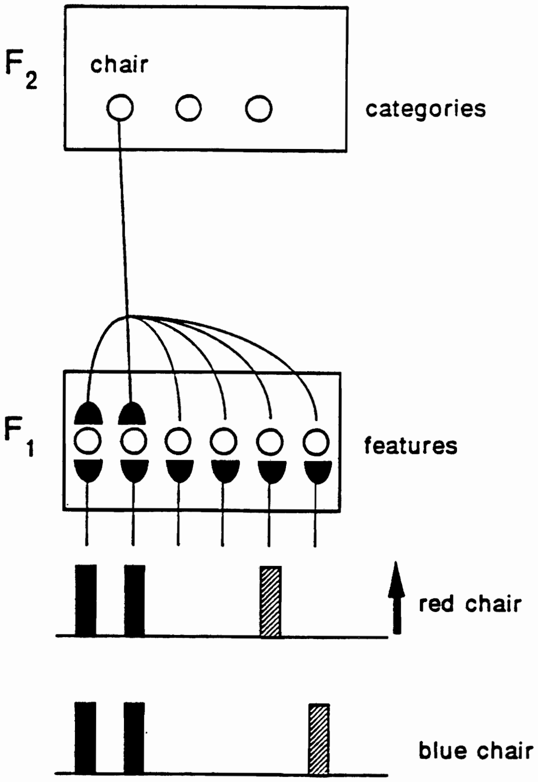


Figure 2. Due to prior experience the top-down expectation from the "chair" category reperesentation in level F_2 to the feature level F_1 encodes negligible LTM traces to represent color features.

intentionality imply aspects of logic.

The core mechanism that ensures code stability is matching of bottom-up inputs with their learned top-down expectations. In addition, the 2/3 Rule enables an ART system to rapidly initiate a cycle of hypothesis testing and memory search for more appropriate recognition codes when a selected recognition code is found to be unpredictive in a novel environment. Figures 2 and 3 illustrate how this works.

Figure 2 schematizes a two-level network in which level F_1 codes features and level F_2 codes recognition categories. In Figure 2, a recognition category at level F_2, labelled "chair," reads out a learned top-down expectation to level F_1. Due to prior learning, the adaptive weights, or Long Term Memory (LTM) traces, of the top-down expectation are large at combinations of features that are common to several types of chairs. During prior experience, however, the color of a chair, whether red or blue, has been irrelevant, so the bottom-up and top-down LTM traces corresponding to color features are very small or zero (Figures 2 and 3a).

As a result of 2/3 Rule matching, bottom-up activation at the level F_1 features corresponding to color features is suppressed when the chair category is activated (Figure 3b).

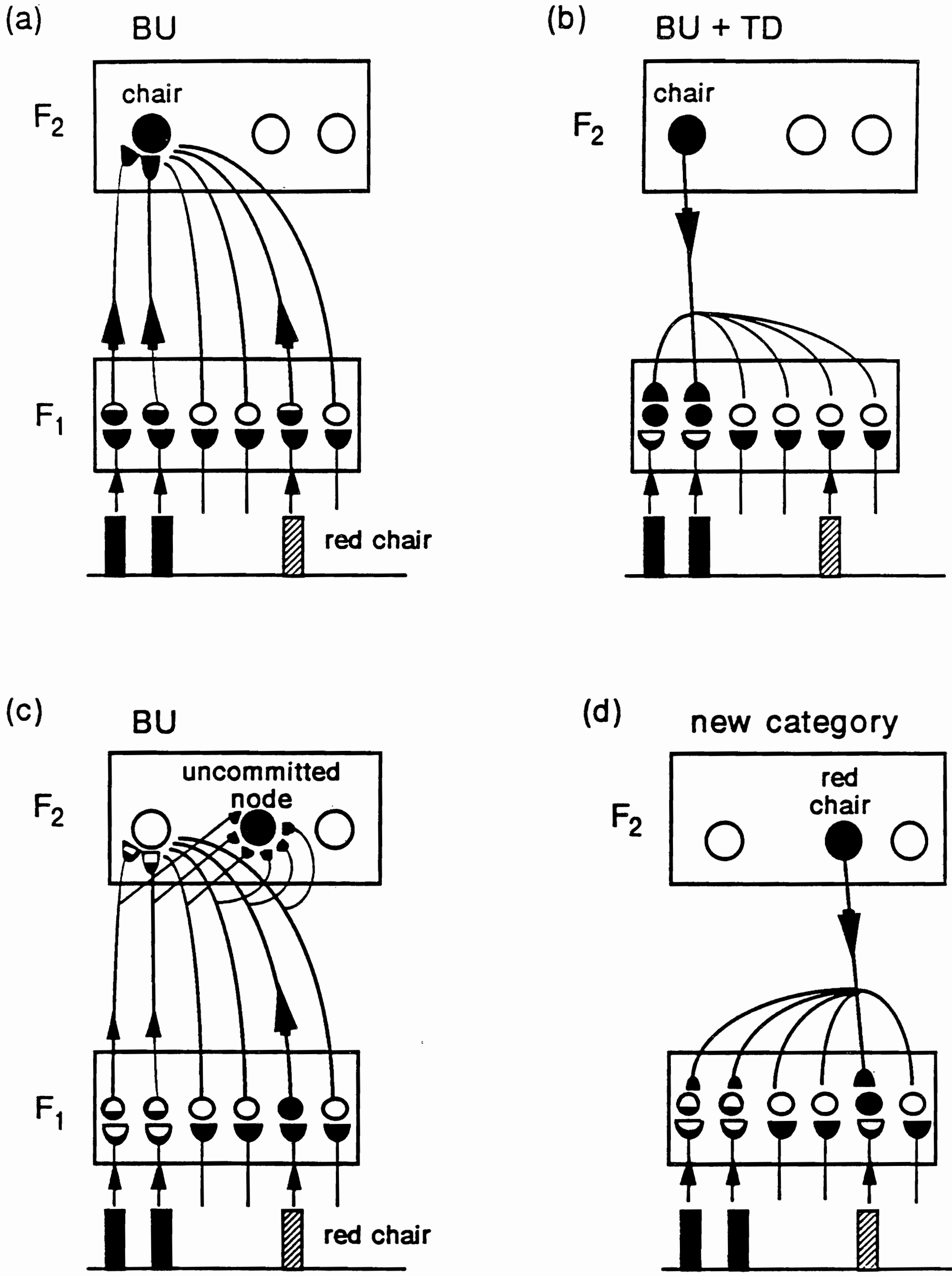


Figure 3. (a) When a red chair is presented, the color features corresponding to red are activated at level F_1, but their bottom-up LTM traces to the chair category in level F_2 are zero. Despite this fact, the chair category provides the best representation of red chair features, and is thus selected at F_2. (b) The top-down learned expectation from the "chair" category suppresses activation of red features at level F_1 via the 2/3 Rule. The chemical transmitters in the bottom-up pathways that abut the active F_1 features are habituated (see half-empty semi-circular synapses). (c) A reset event triggers a search by zeroing activity at all nodes. Since the transmitter in the pathways activating red features were not habituated, these features get more activity in the next processing cycle, leading to selection in (d) of a new category representation for "red chair" that will include color features when it learns its bottom-up and top-down LTM representations.

Color features are deemed irrelevant, based upon prior experience, even though they may receive large bottom-up activations. No "attention" is focussed upon the irrelevant color features, even though the color of the chair is processed by preattentive mechanisms.

Suppose, however, that environmental contingencies change, so that the red color of a chair becomes predictive in a new situation. Then the previously learned chair category leads to an action which causes disconfirming environmental feedback, such as negative reinforcement. This feedback triggers a nonspecific reset event. The reset event, in turn, initiates a cycle of parallel hypothesis testing, or memory search, within the network that automatically assigns a larger activation weight to features in the bottom-up input that have previously been suppressed by 2/3 Rule matching (Figure 3c). This search cycle leads to the discovery and learning of a new category in which both the color of the chair and its chair-like shape features are simultaneously incorporated (Figures 3d). The present chapter summarizes recent modelling results concerning how this search cycle is controlled.

4. A Role for a Neurotransmitter Medium Term Memory in Hypothesis Testing and Memory Search

This hypothesis testing model clarifies why ART predictions involve multiple levels of brain organization, from membrane channels and neurotransmitters to network properties like attention and resonance. On a formal level, the model implements parallel search of compressed or distributed recognition codes in a neural network hierarchy. The search process functions well with either fast learning or slow learning, and can robustly cope with sequences of asynchronous input patterns in real-time. Such a search process emerges when computational properties of neurotransmitters, such as transmitter accumulation, release, inactivation, and modulation, are embedded within an Adaptive Resonance Theory architecture called ART 3. A key part of the search process utilizes formal analogs of synergetic interactions between ions such as Na^+ and Ca^{2+}. These interactions control a nonlinear feedback process that enables the spatial pattern of presynaptic transmitter to model the spatial pattern of postsynaptic activation that represents a recognition code, as in Figure 3.

The spatial pattern of postsynaptic activation operates on a fast time scale, called Short Term Memory (STM). The spatial pattern of presynaptic transmitter operates on a slower time scale, called Medium Term Memory (MTM). Speaking intuitively, MTM encodes the state of transmitter *habituation* through time. Together, as in Figure 3, STM and MTM control a rapid search process which ends with a state of STM resonance that defines the network's focus of attention. The STM resonance also triggers selective learning within the adaptive weights whose pathways support the resonating STM activities. The adaptive weights operate on the network's slowest time scale, called Long Term Memory (LTM). The remainder of the chapter explains how STM, MTM, and LTM work together to control the search process.

5. Control of Hypothesis Testing by the Attentional and Orienting Subsystems

Adaptive Resonance Theory first emerged from an analysis of the coding properties and instabilities inherent in feedforward adaptive coding structures (Grossberg, 1976a). More recent work has led to the development of three classes of ART neural network architectures, specified as systems of differential equations. The first class, ART 1, self-organizes recognition categories for arbitrary sequences of binary input patterns (Carpenter and Grossberg, 1987a). A second class, ART 2, does the same for either binary or analog inputs (Carpenter and Grossberg, 1987b). The third class, ART 3, solves computational problems of ART systems embedded in network hierarchies where there can, in general, be either fast or slow learning and distributed or compressed code representations (Carpenter and Grossberg, 1990).

Both ART 1 and ART 2 use a maximally compressed, or winner-take-all, pattern recognition code. Such a code is a limiting case of the partially compressed recognition codes that are typically used in explanations by ART of biological data (Grossberg, 1982, 1987a, 1987b). Both winner-take-all and partially compressed recognition codes had previously been mathematically analysed in models for competitive learning, also called self-organizing feature maps. The basic equations and mathematical properties of competitive learning and self-organizing feature maps were described by Grossberg (1972c, 1976a, 1978), Malsburg (1973), and Willshaw and Malsburg (1976), and further developed by Kohonen (1984). The name "GKM models" may be used to summarize this historical development.

A GKM model forms part of the bottom-up dynamics of every ART model. The remaining ART mechanisms show how GKM learning can be self-stabilized in an arbitrary environment without slowing or terminating the learning rate. In particular, winner-take-all recognition codes were used in ART 1 and ART 2 to enable a rigorous analysis to be made of how the bottom-up and top-down dynamics of ART systems can be joined together in a real-time self-organizing system capable of learning a stable pattern recognition code in response to an arbitrary sequence of input patterns. These results have provided a computational foundation for designing ART systems capable of stably learning partially compressed recognition codes, as in the ART 3 systems.

The main elements of a typical ART 1 module are illustrated in Figure 4. F_1 and F_2

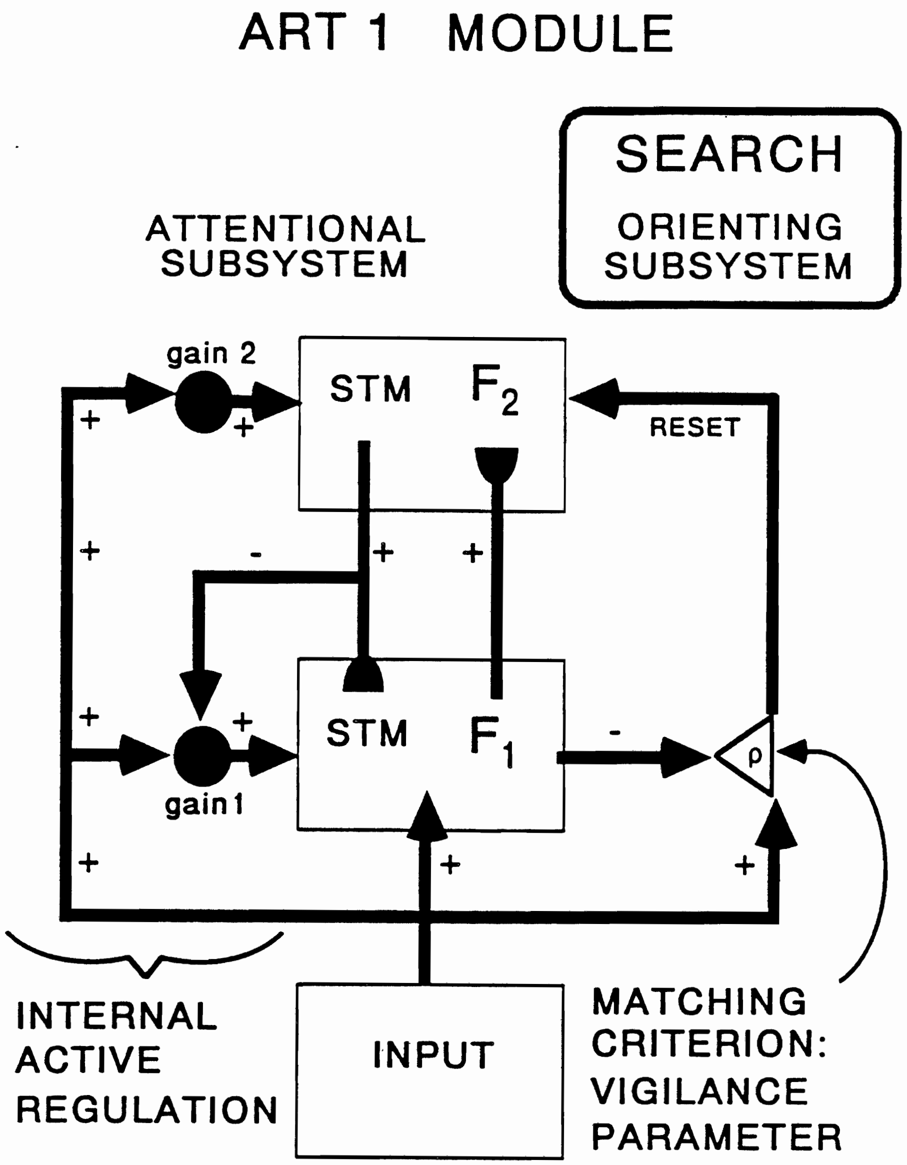


Figure 4. Typical ART 1 neural network module (Carpenter and Grossberg, 1987a).

are fields of network nodes. An input is initially represented as a pattern of activity across the nodes, or feature detectors, of field F_1. The pattern of activity across F_2 corresponds to the category representation. Because patterns of activity in both fields may persist after input offset yet may also be quickly inhibited, these patterns are called Short Term Memory, or STM, representations. The two fields, linked both bottom-up and top-down by adaptive filters, constitute the Attentional Subsystem. Because the connection weights defining the adaptive filters may be modified by inputs and may persist for very long times after input offset, these connection weights are called Long Term Memory, or LTM, variables. The Orienting Subsystem becomes active during search. It interacts with the Attentional Subsystem in order to enable new learning within the Attentional Subsystem to occur without causing unselective forgetting of previously learned categories to occur.

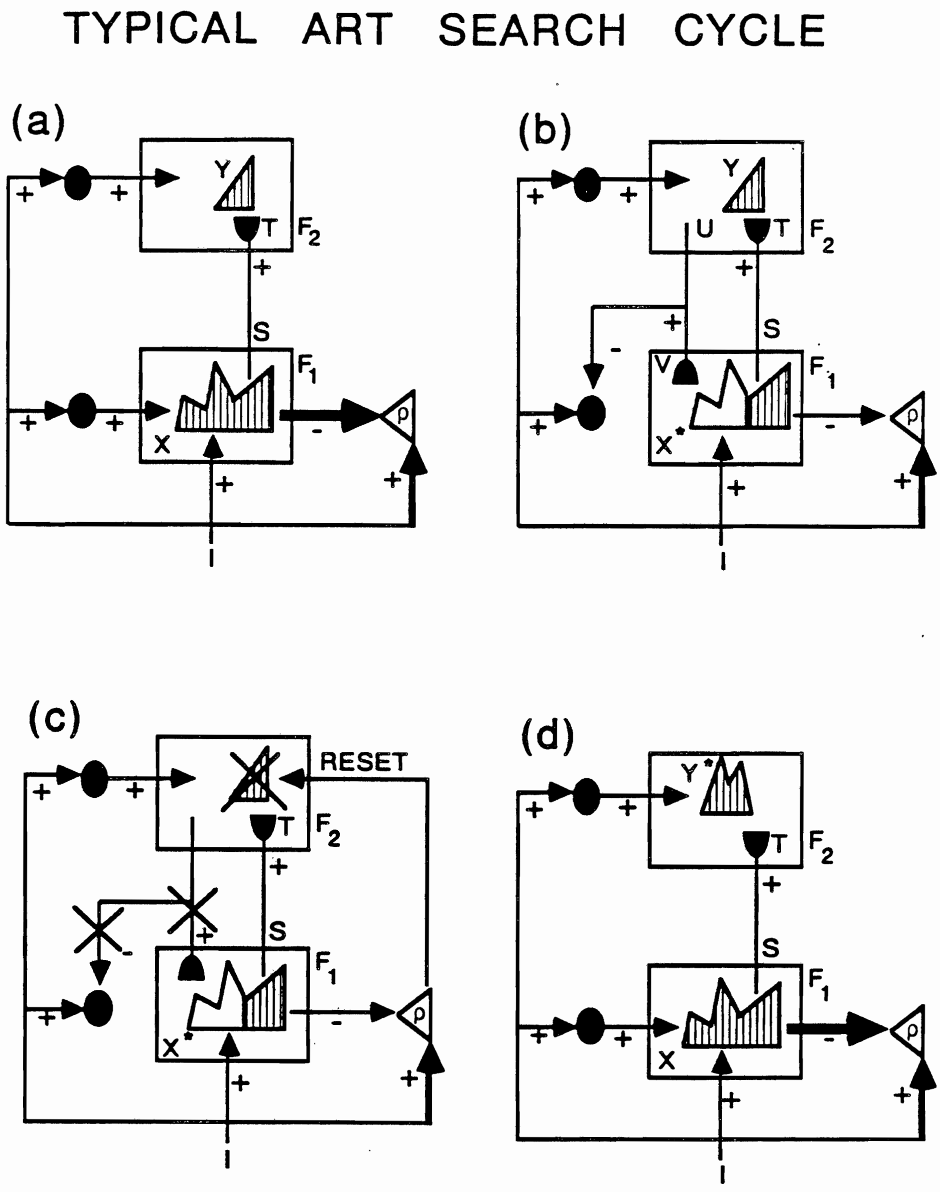

TYPICAL ART SEARCH CYCLE

Figure 5. ART search cycle (Carpenter and Grossberg, 1987a).

This search process is the subject of the present chapter.

6. An ART Search Cycle

Figure 5 illustrates a typical ART search cycle. An input pattern I registers itself as a pattern X of activity across F_1 (Figure 5a). The F_1 output signal vector S is then transmitted through the multiple converging and diverging weighted adaptive filter pathways emanating from F_1, sending a net input signal vector T to F_2. The internal competitive dynamics of F_2 contrast-enhance T. The F_2 activity vector Y therefore registers a compressed representation of the filtered $F_1 \rightarrow F_2$ input and corresponds to a category representation for the input active at F_1. These are the typical dynamics of competitive learning.

In addition, within ART, vector **Y** generates a signal vector **U** that is sent top-down through the second adaptive filter, giving rise to a net top-down signal vector **V** to F_1 (Figure 5b). F_1 now receives two input vectors, **I** and **V**. An ART system is designed to carry out a matching process whereby the original activity pattern **X** due to input pattern **I** may be modified by the *template pattern* **V** that is associated with the current active category. If **I** and **V** are not sufficiently similar according to a flexible matching criterion established by a dimensionless *vigilance parameter* ρ, a reset signal quickly and enduringly shuts off the active category representation (Figure 5c), allowing a new category to become active. Search ensues (Figure 5d) until either an adequate match is made or a new category is established.

Prior to the introduction of ART 3, we proposed that the enduring shut-off of erroneous category representations by a nonspecific reset signal could occur at F_2 if F_2 were organized as a gated dipole field, whose dynamics depend on habituative transmitter gates (Carpenter and Grossberg, 1987a; Grossberg, 1976b). Though the ART 3 search process does not use a gated dipole field, it does retain and extend the core idea that habituative transmitter dynamics can enable a robust search process when appropriately embedded in an ART system.

7. ART 2: Three-Layer Competitive Fields

Figure 6 shows the principal elements of a typical ART 2 module. In addition to incorporating many characteristics of the ART 1 module, ART 2 networks all have three processing layers within the F_1 field. These three processing layers allow ART 2 to stably categorize sequences of analog input patterns that can, in general, be arbitrarily close to one another. In Figure 6, one F_1 layer reads in the bottom-up input, one layer reads in the top-down filtered input from F_2, and a middle layer matches patterns from the top and bottom layers before sending a composite pattern back through the F_1 feedback loop. Both F_1 and F_2 are shunting competitive networks that contrast-enhance and normalize their activation patterns (Grossberg, 1982).

In applications, ART modules are often embedded in larger architectures that are hierarchically organized. When an ART module is embedded in a network hierarchy, it is no longer possible to make a sharp distinction between the characteristics of the input representation field F_1 and the category representation field F_2. In order for them to serve both functions, the basic structures of all the network fields in a hierarchical ART system should be homologous, in so far as possible. This constraint is satisfied if all fields of the

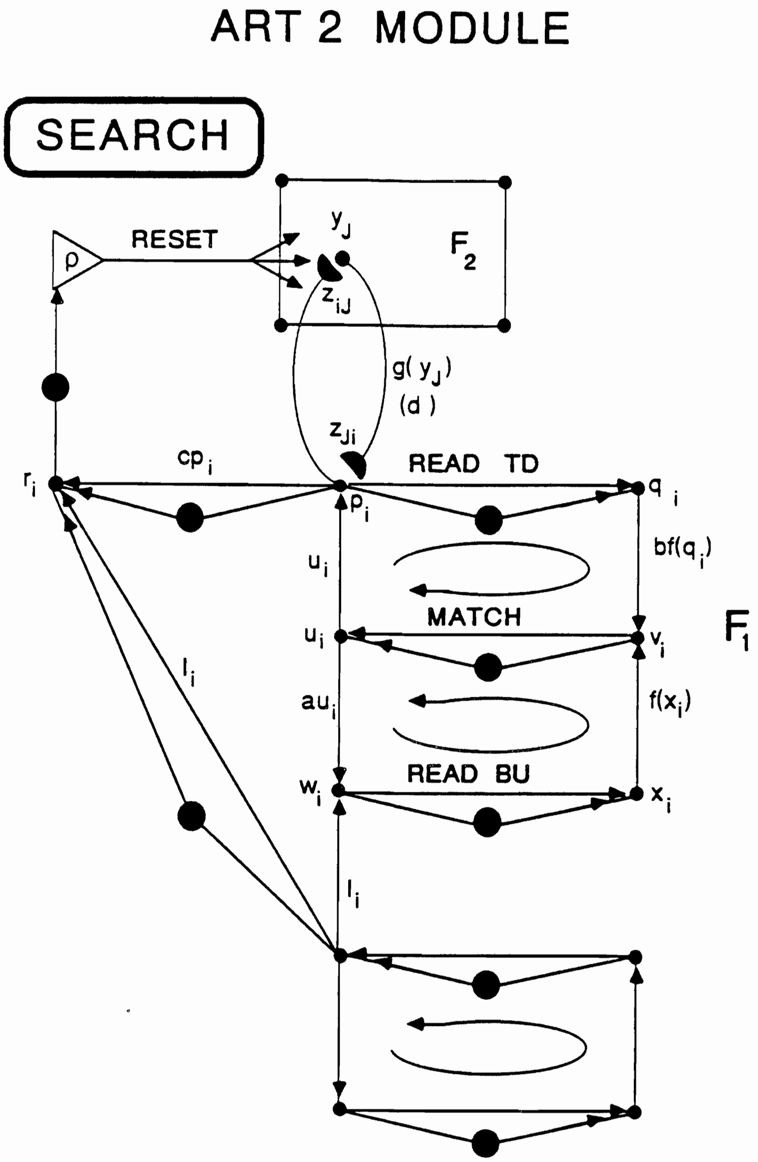


Figure 6. Typical ART 2 neural network module, with three-layer F_1 field (Carpenter and Grossberg, 1987b). Large filled circles are gain control nuclei that nonspecifically inhibit target nodes in proportion to the Euclidean norm of activity in their source fields.

hierarchy are endowed with the F_1 structure of an ART 2 module (Figure 6). Such a design is sufficient for the F_2 field as well as the F_1 field because the principal property required of a category representation field, namely that input patterns be contrast-enhanced and normalized, is a property of the three-layer F_1 structure.

8. Parallel Search of Distributed Codes in an ART 3 Hierarchy

We now consider the problem of implementing parallel search among the distributed codes of a hierarchical ART system all of whose fields are homologous. Assume that a top-down/bottom-up mismatch has occurred somewhere in the system. How can a reset signal search the hierarchy in such a way that an appropriate new category is selected?

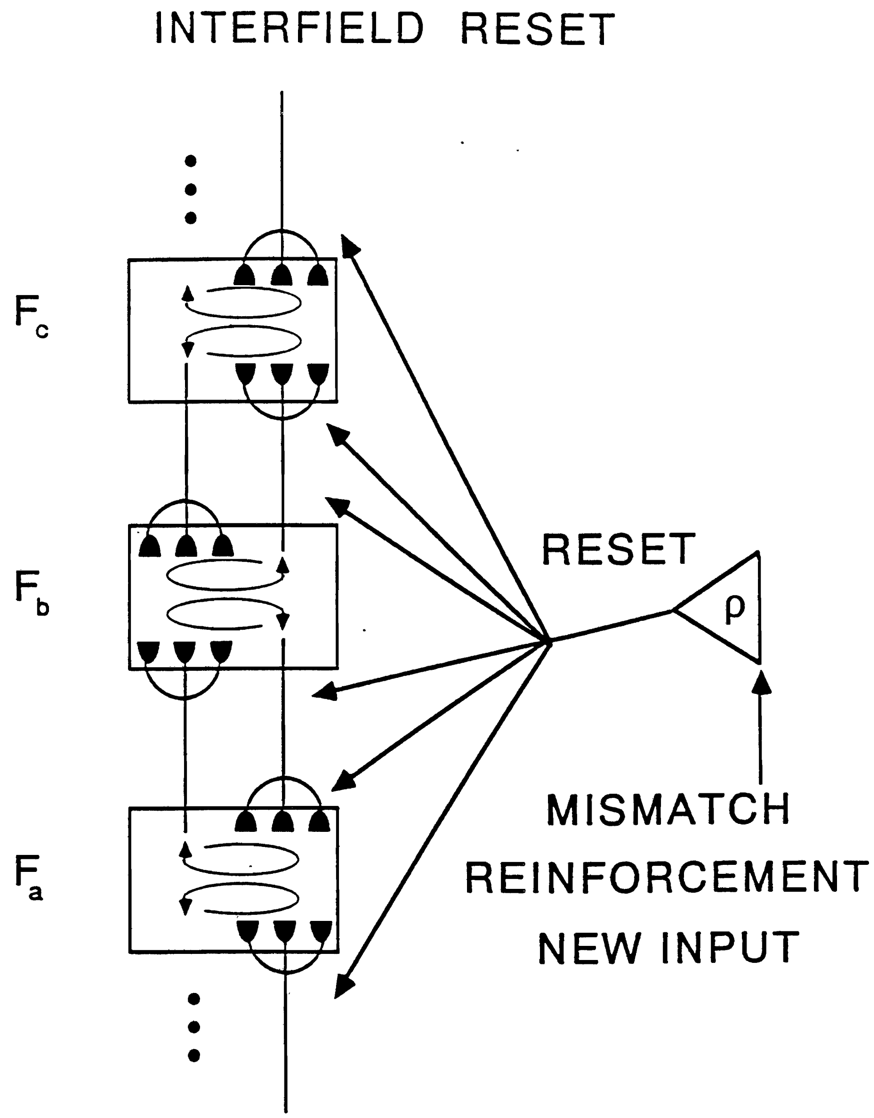


Figure 7. Interfield reset in an ART bidirectional hierarchy.

A key observation is that a reset signal can act upon an ART hierarchy *between* its fields F_a, F_b, $F_c \ldots$ (Figure 7). Locating the site of action of the reset signal between the fields allows each individual field to carry out its pattern processing function without introducing processing biases directly into a field's internal feedback loops.

9. Habituative Chemical Transmitters in ART Search

The computational requirements of the ART search process can be fulfilled by formal properties of neurotransmitters if these properties are appropriately embedded in the total architecture model. In particular, the ART 3 search equations incorporate the dynamics of production and release of a chemical transmitter substance; the inactivation of transmitter at postsynaptic binding sites; and the modulation of these processes via a nonspecific control signal. A conjoint interaction between a presynaptic Na^+ current and a postsynaptic Ca^{++} current is suggested to control the release of presynaptic transmitter (Ito, 1984,

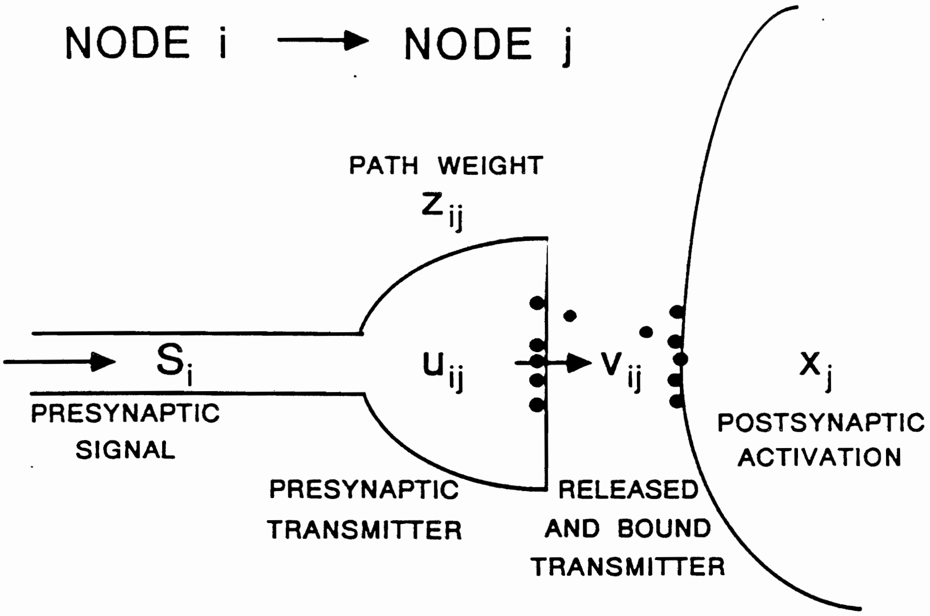


Figure 8. Notation for the ART chemical synapse.

p. 52; Zucker, 1989). The net effect of these transmitter processes is to alter the ionic permeability at the postsynaptic membrane site, thus effecting net excitation or inhibition of the postsynaptic cell. ART 3 dynamics hereby suggest a possible role for postsynaptic cell activation in the control of the presynaptic Ca^{++} current.

The notation to describe these transmitter properties is summarized in Figure 8, for a synapse between the ith presynaptic node and the jth postsynaptic node. The presynaptic signal, or action potential, S_i arrives at a synapse whose adaptive weight, or long term memory trace, is denoted z_{ij}. The variable z_{ij} is identified with the maximum amount of available transmitter. When the transmitter at this synapse is fully accumulated, the amount of transmitter u_{ij} available for release is equal to z_{ij}. When a signal S_i arrives, transmitter is typically released. Variables v_{ij} and w_{ij} denote the amount of transmitter released into the extracellular space from the bottom-up filter and intrafield feedback, respectively. A fraction of this total transmitter pool is assumed to be bound at the postsynaptic cell surface and the remainder rendered ineffective in the extracellular space. Finally, x_j denotes the activity, or membrane potential, of the postsynaptic cell.

The search mechanism can be realized using one of several closely related sets of equations, with corresponding differences in biophysical interpretation. An illustrative system of equations is described in terms of the variables z_{ij}, u_{ij}, v_{ij}, w_{ij}, and x_j at the ij^{th} pathway and j^{th} node of an ART 3 system.

Presynaptic Transmitter

$$\frac{du_{ij}}{dt} = (z_{ij} - u_{ij}) - u_{ij}[\text{release rate}] \tag{1}$$

Bound Transmitter in Bottom-Up Filter

$$\begin{aligned}\frac{dv_{ij}}{dt} &= -v_{ij} + u_{ij}[\text{release rate}] - v_{ij}[\text{inactivation rate}] \\ &= -v_{ij} + u_{ij}[\text{release rate}] - v_{ij}[\text{reset signal}]\end{aligned} \tag{2}$$

Bound Transmitter in Intrafield Feedback Pathways

$$\begin{aligned}\frac{d}{dt}w_j &= -w_j + [\text{intrafield feedback}] - w_j[\text{inactivation rate}] \\ &= -w_j + [\text{intrafield feedback}] - w_j[\text{reset signal}]\end{aligned} \tag{3}$$

Postsynaptic Activation

$$\begin{aligned}\epsilon\frac{dx_j}{dt} &= -x_j + (A - x_j)[\text{excitatory inputs}] - (B + x_j)[\text{inhibitory inputs}] \\ &= -x_j + (A - x_j)\Big[\sum_i v_{ij} + \ w_j\Big] - (B + x_j)[\text{internode competition}]\end{aligned} \tag{4}$$

Equation (1) says that presynaptic transmitter is produced and/or mobilized until the amount u_{ij} of transmitter available for release reaches the maximum level z_{ij}. The adaptive weight z_{ij} itself changes on the slower time scale of learning, but remains essentially constant on the time scale of a single reset event. Available presynaptic transmitter u_{ij} is released at a rate that is specified below.

A fraction of presynaptic transmitter becomes postsynaptic bound transmitter after being released. For simplicity, we ignore the fraction of released transmitter that is inactivated in the extracellular space. Equation (2) says that the bound transmitter is inactivated by the reset signal. Equation (3) posits a similar process for the transmitter released from intrafield feedback pathways.

Equation (4) for the postsynaptic activity x_j is a shunting membrane equation such that excitatory inputs drive x_j up toward a maximum depolarized level equal to A; inhibitory inputs drive x_j down toward a minimum hyperpolarized level equal to $-B$; and activity passively decays to a resting level equal to 0 in the absence of inputs. The net effect of bound transmitter at all synapses converging on the jth node is assumed to be excitatory, via the term

$$\sum_i v_{ij} \ + \ w_j. \tag{5}$$

Bottom-up signals and internal feedback from within the target field are excitatory (Figure 5), whereas the competitive interactions from other intrafield nodes are inhibitory. Parameter ϵ is small, corresponding to the assumption that activation dynamics are fast relative to the transmitter accumulation rate, equal to 1 in equation (1).

The ART 3 system can be simplified for purposes of simulation. Suppose that $\varepsilon << 1$ in (4); the reset signals in (2) and (3) are either 0 or $>> 1$; and net intrafield feedback is excitatory. Then equations (1), (6), (7), and (8) below approximate the main properties of ART 3 system dynamics.

Simplified ART 3 Equations

$$\frac{du_{ij}}{dt} = (z_{ij} - u_{ij}) - u_{ij}[\text{release rate}] \tag{1}$$

$$\begin{cases} \frac{dv_{ij}}{dt} = -v_{ij} + u_{ij}\ [\text{release rate}] & \text{if reset} = 0 \\ v_{ij}(t) = 0 & \text{if reset} >> 1 \end{cases} \tag{5}$$

$$\begin{cases} \frac{d}{dt} w_j = -w_j + [\text{intrafield feedback}] & \text{if reset} = 0 \\ w_{ij}(t) = 0 & \text{if reset} >> 1 \end{cases} \tag{7}$$

$$x_j(t) = \begin{cases} \sum_i v_{ij} + w_j & \text{if reset} = 0 \\ 0 & \text{if reset} >> 1. \end{cases} \tag{8}$$

10. Conjoint Influence of Presynaptic and Postsynaptic Activity on Transmitter Release

The transmitter release and inactivation rates in equations (1)–(3) will now be specified. Then we trace the dynamics of the system during a brief time interval after the input turns on ($t = 0^+$), when the signal S_i first arrives at the synapse; when subsequent internal feedback signals act from within the target field, following contrast-enhancement of the inputs; and when a reset signal implements a rapid and enduring inhibition of erroneously selected pattern features. We begin with the

ART Search Hypothesis 1: Presynaptic transmitter u_{ij} is released at a rate jointly proportional to the presynaptic signal S_i and a function $f(x_j)$ of the postsynaptic activity. That is, in equations (1), (2), and (6),

$$\text{release rate} = S_i f(x_j). \tag{9}$$

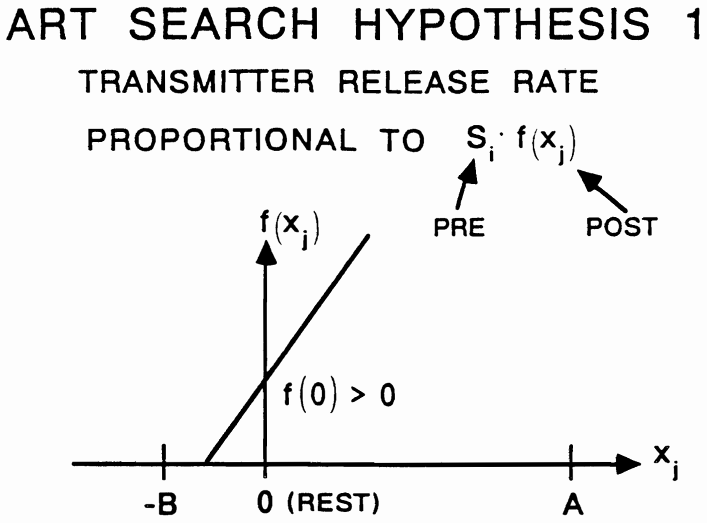


Figure 9. The ART Search Hypothesis 1 specifies the transmitter release rate.

The function $f(x_j)$ in equation (9) has the qualitative properties illustrated in Figure 9, where $f(x_j)$ has a positive value when x_j is at its 0 resting level, so that transmitter u_{ij} can be released when the signal S_i arrives at the synapse. In addition, $f(x_j)$ equals 0 when x_j is significantly hyperpolarized, but rises steeply when x_j is near 0. In our computer simulations, $f(x_j)$ is linear above a small negative threshold.

The form factor $S_i f(x_j)$ is a familiar one in the neuroscience and neural network literatures. In particular, such a product is often used to model associative learning, where it links the rate of learning in the ijth pathway to the presynaptic signal S_i and the postsynaptic activity x_j. Associative learning occurs, however, on a time scale that is much slower than the time scale of transmitter release. On the fast time scale of transmitter release, the form factor $S_i f(x_j)$ may be compared to interactions between voltages and ions, where the presynaptic signal depends on the Na^+ ion, the postsynaptic signal on the Ca^{2+} ion, and presynaptic transmitter release on the *joint* fluxes of these two ions (Ito, 1984). The ART Search Hypothesis 1 formalizes this type of synergetic relationship between presynaptic and postsynaptic processes in effecting transmitter release. Moreover, the rate of transmitter release is typically a function of the concentration of Ca^{2+} in the extracellular space, and this function has qualitative properties similar to the function $f(x_j)$ shown in Figure 9 (Kandel and Schwartz, 1981, p. 84; Kuffler, Nicholls, and Martin, 1984, p. 244).

11. System Dynamics at Input Onset: An Approximately Linear Filter

Some implications of the ART Search Hypothesis 1 will now be summarized. Assume that at time $t = 0$ transmitter u_{ij} has accumulated to its maximal level z_{ij} and that activity x_j and bound transmitter v_{ij} equal 0. Consider a time interval $t = 0^+$ immediately after a signal S_i arrives at the synapse. During this brief initial interval, the ART equations approximate the linear filter dynamics typical of many neural network models. In particular, equations (2) and (9) imply that the amount of bound transmitter is determined by equation

$$\frac{dv_{ij}}{dt} = -v_{ij} + u_{ij} S_i f(x_j) - v_{ij}[\text{inactivation rate}]. \tag{10}$$

Thus at times $t = 0^+$,

$$\frac{dv_{ij}}{dt} \approx z_{ij} S_i f(0) \tag{11}$$

and so

$$v_{ij}(t) \approx K(t) S_i z_{ij} \quad \text{for times } t = 0^+. \tag{12}$$

Because equation (12) holds at all the synapses adjacent to cell j, equation (6) implies that

$$x_j(t) \approx \sum_i K(t) S_i z_{ij} = K(t) \mathbf{S} \cdot \mathbf{z}_j \quad \text{for times } t = 0^+. \tag{13}$$

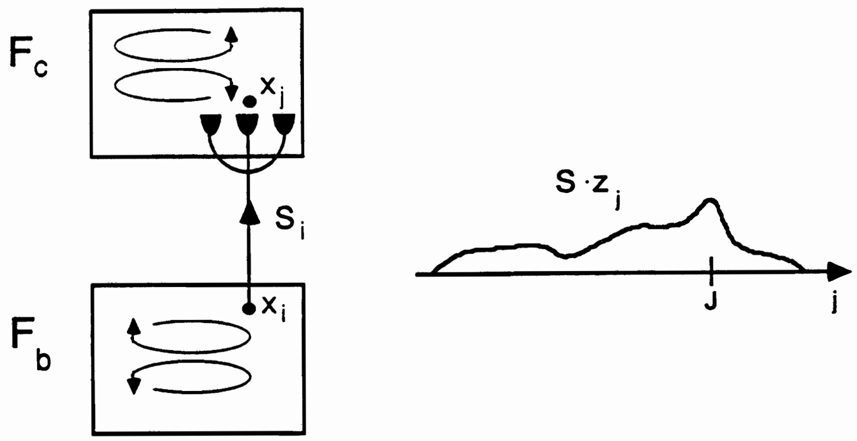
Here S denotes the vector $(S_1 \ldots S_n)$, $\mathbf{z}_j$ denotes the vector $(z_{1j} \ldots z_{nj})$, and $i = 1 \ldots n$. Thus in the initial moments after a signal arrives at the synapse, the small amplitude activity x_j at the postsynaptic cell grows in proportion to the dot product of the incoming signal vector S times the adaptive weight vector $\mathbf{z}_j$.

12. System Dynamics after Intrafield Feedback: Amplification of Presynaptic Transmitter Release by Postsynaptic Activation

In the next time interval, the intrafield feedback signal contrast-enhances the initial signal pattern (13) via equation (6) and amplifies the total activity across field F_c in Figure 10a, thereby starting to generate an attentive focus. Figure 10b shows typical contrast-enhanced activity profiles: partial compression of the initial signal pattern; or maximal compression, or choice, where only one postsynaptic node remains active due to the strong competition within the field F_c.

In summary, the model behaves initially like a linear filter. The resulting pattern of activity across postsynaptic cells is contrast-enhanced, as required in the ART 2 model as well as in the many other neural network models that incorporate competitive learning (Grossberg, 1982). These models implicitly assume that intracellular transmitter u_{ij} is

(a) $x_j \approx K(t)\, S\cdot z_j \quad (t=0^+)$



(b) FEEDBACK CONTRAST-ENHANCES x_j

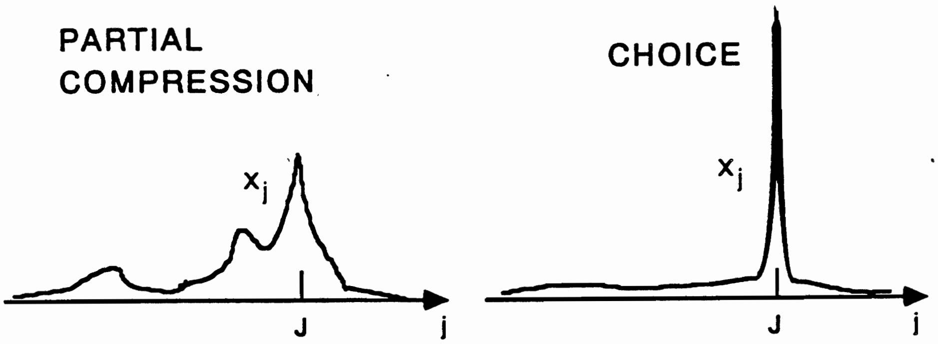


Figure 10. (a) If transmitter is fully accumulated at $t = 0$, low-amplitude postsynaptic STM activity x_j is initially proportional to the dot product of the signal vector S and the weight vector $\mathbf{z}_j$. (b) Intrafield feedback rapidly contrast-enhances the initial STM activity pattern. Large-amplitude activity is then concentrated at one or more nodes.

always accumulated up to its target level z_{ij} and that postsynaptic activity x_j does not alter the rate of transmitter release:

$$u_{ij} \approx z_{ij} \quad \text{and} \quad v_{ij} \approx z_{ij} S_i. \tag{14}$$

We now suggest how nonlinearities of synaptic transmission and neuromodulation can, when embedded in an ART circuit, help to correct coding errors by triggering a parallel search, allow the system to respond adaptively to reinforcement, and rapidly reset itself to changing input patterns. In equation (10), term

$$u_{ij} S_i f(x_j) \tag{15}$$

for the amount of transmitter released per unit time implies that the original incoming weighted signal $z_{ij}S_i$ is distorted both by depletion of the presynaptic transmitter u_{ij} and by the activity level x_j of the postsynaptic cell. If these two nonlinearities are significant, the net signal in the ijth pathway depends jointly on the maximal weighted signal $z_{ij}S_i$; the prior activity in the pathway, as reflected in the amount of depletion of the transmitter u_{ij}; and the immediate context in which the signal is sent, as reflected in the target cell activity x_j. In particular, once activity in a postsynaptic cell becomes large, this activity dominates the transmitter release rate, via the term $f(x_j)$ in (15). In other words, although linear filtering properties initially determine the small-amplitude activity pattern of the target field F_c, once intrafield feedback amplifies and contrast-enhances the postsynaptic activity x_j (Figure 10b), it plays a major role in determining the amount of released transmitter v_{ij} (Figure 11). In particular, the postsynaptic activity pattern across the field F_c that represents the recognition code (Figure 10b) is imparted to the pattern of released transmitter (Figure 11), which then also represents the recognition code, rather than the initial filtered pattern $\mathbf{S} \cdot \mathbf{z}_j$.

13. System Dynamics during Reset: Inactivation of Bound Transmitter Channels

The dynamics of transmitter release implied by the ART Search Hypothesis 1 can be used to implement the reset process, by postulating the

ART Search Hypothesis 2: The nonspecific reset signal quickly inactivates postsynaptic membrane channels at which transmitter is bound (Figure 12).

The reset signal in equations (5) and (6) may be interpreted as assignment of a large value to the inactivation rate in a manner analogous to the action of a neuromodulator. Inhibition of postsynaptic nodes breaks the strong intrafield feedback loops that implement ART 2 and ART 3 matching and contrast-enhancement (equation (3) or (6)).

The pattern of presynaptic transmitter provides a representation of the postsynaptic recognition code. The arrival of a reset signal implies that some part of the system has judged this code to be erroneous, according to some criterion. The ART Search Hypothesis 1 implies that the largest concentrations of bound extracellular transmitter are adjacent to the nodes which the system, based on its past experiences, deems the most likely representation of the data.. Correspondingly, the presynaptic transmitter stores of the most active postsynaptic cells are selectively inactivated, or habituated. If the hypothesis corresponding to this representation is supported, then the selective presynaptic transmitter habituation does not compromise the postsynaptic representation. The ART Search Hy-

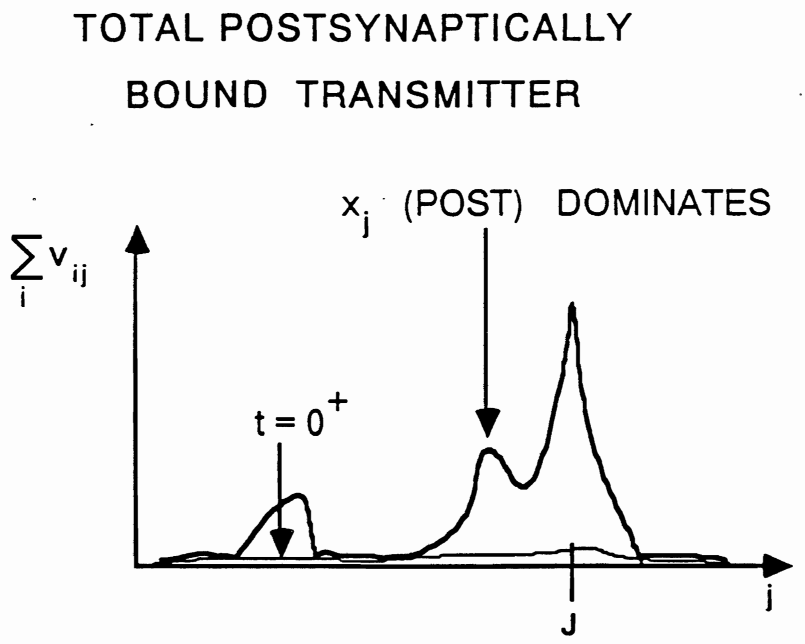


Figure 11. The ART Search Hypothesis 1 implies that large amounts of transmitter (v_{ij}) are released only adjacent to postsynaptic nodes with large-amplitude activity (x_j). Competition within the postsynaptic field therefore transforms the initial low-amplitude distributed pattern of released and bound transmitter into a large-amplitude contrast-enhanced pattern.

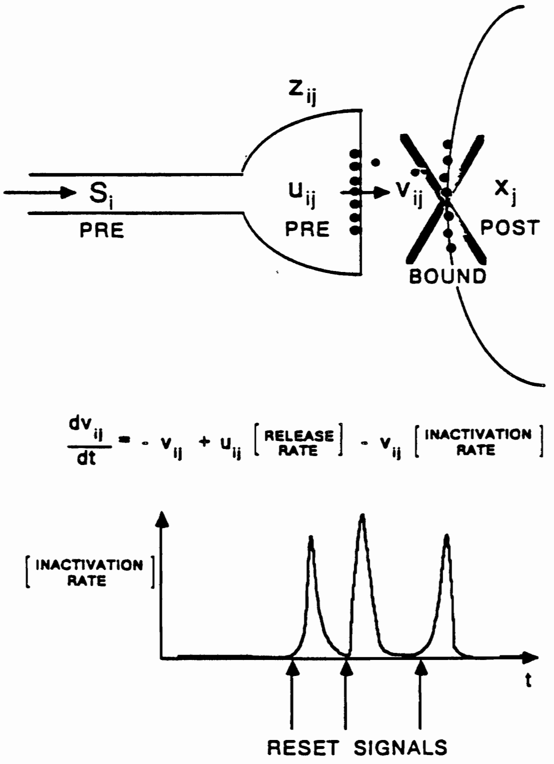


Figure 12. The ART Search Hypothesis 2 specifies a high rate of inactivation of bound transmitter following a reset signal. Postsynaptic action of the nonspecific reset signal is similar to that of a neuromodulator.

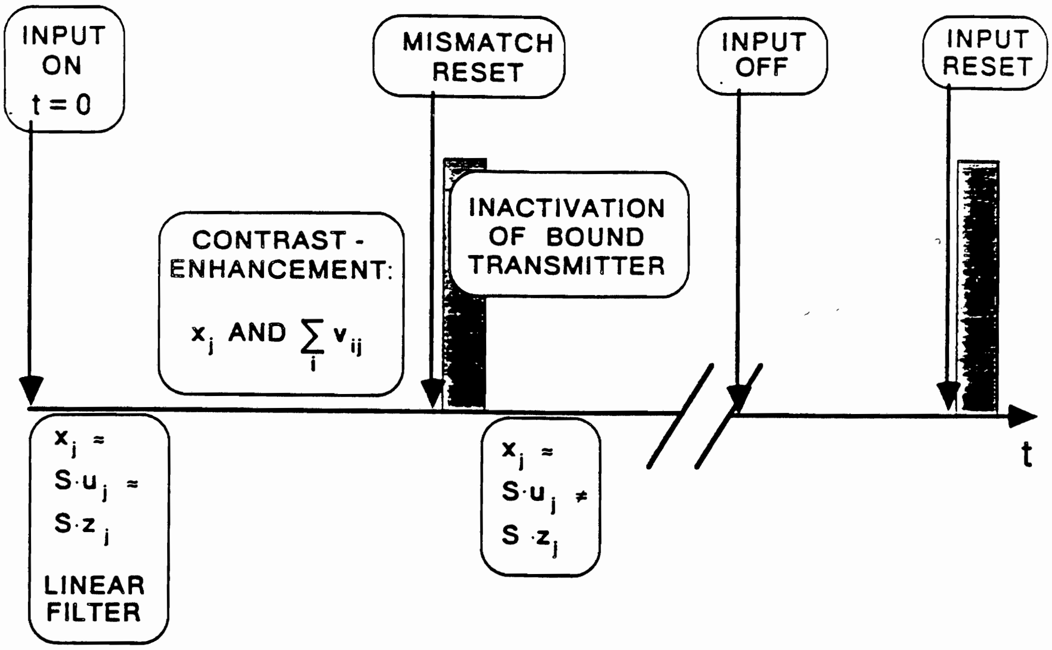


Figure 13. An ART 3 search cycle: Initial read-out of bottom-up signals behaves like a linear filter. Read-out leads to contrast-enhancement and amplification both of the postsynaptic activity pattern and the presynaptic transmitter pattern. A mismatch causes a reset event that inactivates the channels at which transmitter is postsynaptically bound. The presynaptic transmitter pattern biases the network against the previously most active feature detectors. After such a reset event, the adaptive filter delivers a smaller signal to the previously most active feature detectors, thereby testing a different hypothesis.

pothesis 2 implies that a reset event restores postsynaptic activation to an unbiased baseline of inactivation, and leaves the selective presynaptic transmitter habituation intact. This is the basis for calling the postsynaptic activation pattern Short Term Memory, or STM, and the presynaptic transmitter pattern Medium Term Memory.

In summary, after a reset event occurs, the system maintains a presynaptic MTM bias against postsynaptic activation of those nodes which were most responsible for the predictive failure that led to the reset event. Although the transmitter signal pattern $\mathbf{S} \cdot \mathbf{u}_j$ originally sent to target nodes at times $t = 0^+$ was proportional to $\mathbf{S} \cdot \mathbf{z}_j$, as in equation (12), the transmitter signal pattern $\mathbf{S} \cdot \mathbf{u}_j$ after the reset event is no longer proportional to $\mathbf{S} \cdot \mathbf{z}_j$. Instead, it is selectively biased against those features that were previously most active (Figure 13). The new signal pattern $\mathbf{S} \cdot \mathbf{u}_j$ will lead to selection of another contrast-enhanced representation, with more activation given to previously unattended features. This representation may or may not then be reset. This search process continues until an acceptable match is found, whence an attentive resonance is established and learning triggered in those adaptive weights, or LTM traces, which abut resonant activations.

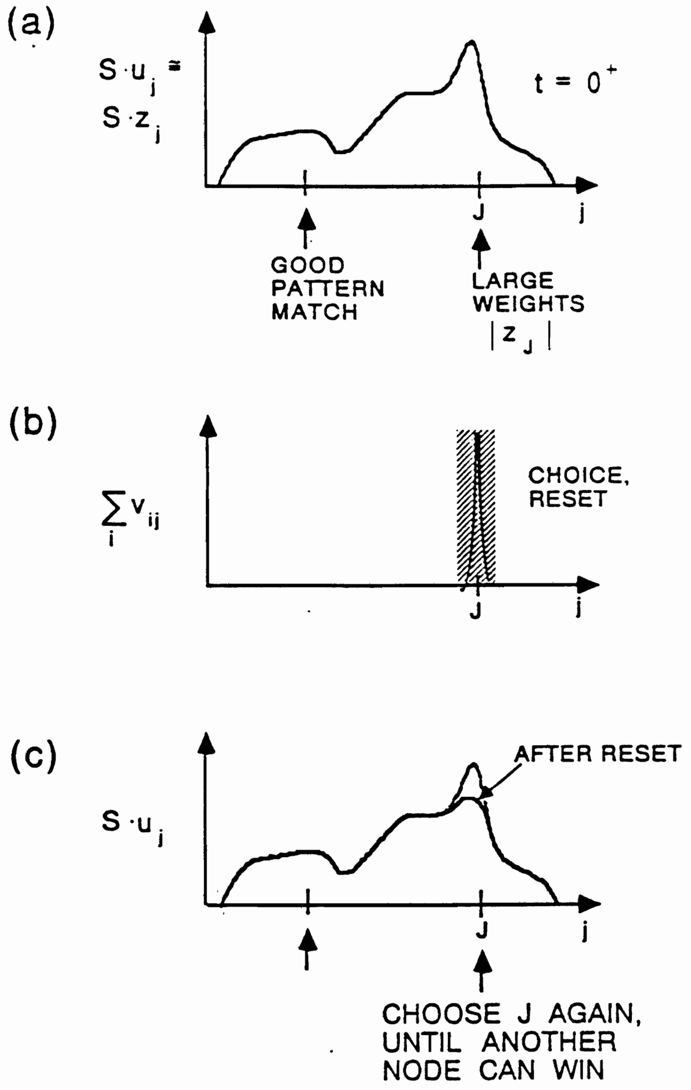


Figure 14. Change of transmitter and short term memory patterns due to search: (a) Initially the signal pattern $S\cdot u_j$ may deliver a bigger input to nodes whose adaptive weights z_J are larger than to nodes whose adaptive weights better match the input pattern; (b) contrast enhancement and amplification select the preferred nodes in short term memory; (c) a reset event biases the transmitted signal pattern $S \cdot u_j$ against these preferred nodes. More than one reset event may be needed to accumulate sufficient bias to select a new representation capable of supporting resonant attention and learning.

14. The Cycle of Search, Resonant Attention, and Learning: A New Round of Neurobiological Predictions

Figure 14 summarizes system dynamics of the ART search model during a single input presentation. Initially the transmitted signal pattern $\mathbf{S} \cdot \mathbf{u}_j$, as well as the postsynaptic activity x_j, are proportional to the weighted signal pattern $\mathbf{S} \cdot \mathbf{z}_j$ of the linear filter. The postsynaptic activity pattern is then contrast-enhanced, due to the internal competitive dynamics of the target field. The ART Search Hypothesis 1 implies that the transmitter release rate is greatly amplified in proportion to the level of postsynaptic activity. A subsequent reset signal selectively inactivates transmitter in those pathways that caused an error. Following the reset wave, the new signal $\mathbf{S} \cdot \mathbf{u}_j$ is no longer proportional to

$\mathbf{S} \cdot \mathbf{z}_j$ but is, rather, biased against the previously active representation, thereby causing a shift in the attentive focus to a potentially more predictive pattern of features in short term memory. A series of such reset events may ensue, until an adequate match or a new category is found, thereby allowing the network to go into resonance and to maintain its attentional focus. New code learning can then occur on a time scale that is long relative to that of the search process.

Mathematical development and computer simulations of these ART 3 interactions are described in Carpenter and Grossberg (1990). The ART 3 search cycle suggests another round of interdisciplinary predictions which functionally integrate and interpret interactions between microscopic cellular properties, such as the dynamics of ions and transmitters, with more macroscopic network properties, such as search, resonant attention, and learning.

REFERENCES

Bear, M.F. and Singer, W. (1986). Modulation of visual cortical plasticity by acetylcholine and noradrenaline. *Nature*, **320**, 172–176.

Berger, T.W. and Thompson, R.F. (1978). Neuronal plasticity in the limbic system during classical conditioning of the rabbit nictitating membrane response, I: The hippocampus. *Brain Research*, **145**, 323–346.

Bower, G.H. (1981). Mood and memory. *American Psychologist*, **36**, 129–148.

Bower, G.H., Gilligan, S.G., and Monteiro, K.P. (1981). Selectivity of learning caused by adaptive states. *Journal of Experimental Psychology: General*, **110**, 451–473.

Carpenter, G.A. (1989). Neural network models for pattern recognition and associative memory. *Neural Networks*, **2**, 243–257.

Carpenter, G.A. and Grossberg, S. (1987a). A massively parallel architecture for a self-organizing neural pattern recognition machine. *Computer Vision, Graphics, and Image Processing*, **37**, 54–115.

Carpenter, G.A. and Grossberg, S. (1987b). ART 2: Self-organization of stable category recognition codes for analog input patterns. *Applied Optics*, **26**, 4919–4930.

Carpenter, G.A. and Grossberg, S. (1988). The ART of adaptive pattern recognition by a self-organizing neural network. *Computer*: Special issue on Artificial Neural Systems, **21**, 77–88.

Carpenter, G.A. and Grossberg, S. (1990). ART 3: Hierarchical search using chemical transmitters in self-organizing pattern recognition architectures. *Neural Networks*, **3**, 129–152.

Cohen, M.A., Grossberg, S., and Stork, D. (1988). Speech perception and production by a self-organizing neural network. In Y. C. Lee (Ed.), **Evolution, learning, cognition, and advanced architectures**. Hong Kong: World Scientific Publishers, 217–231.

Dudek, S.M. and Bear, M.F. (1989). A biochemical correlate of the critical period for synaptic modification in the visual cortex. *Science*, **246**, 673–675.

Eckhorn, R., Bauer, R., Jordan, W., Brosch, M., Kruse, W., Munk, M., and Reitboeck, H.J. (1988). Coherent oscillations: A mechanism of future linking in the visual cortex? *Biological Cybernetics*, **60**, 121–130.

Ellinwood, E.H. and Kilbey, M.M. (1980). Fundamental mechanisms underlying altered behavior following chronic administration of psychomotor stimulants. *Biological Psychiatry*, **15**, 749–757.

Gabriel, M., Foster, K., Orona, E., Saltwick, S.E., and Stanton, M. (1980). Neuronal activity of cingulate cortex, anteroventral thalamus, and hippocampal formation in discriminative conditioning: Encoding and extraction of the significance of conditional stimuli. *Progress in Psychobiology and Physiological Psychology*, **9**, 125–231.

Gray, C.M., Konig, P., Engel, A.K., and Singer, W. (1989). Oscillatory responses in cat visual cortex exhibit inter-columnar synchronization which reflects global stimulus properties. *Nature*, **338**, 334–337.

Grossberg, S. (1968). Some physiological and biochemical consequences of psychological postulates. *Proceedings of the National Academy of Sciences*, **60**, 758–765.

Grossberg, S. (1969a). On the production and release of chemical transmitters and related topics in cellular control. *Journal of Theoretical Biology*, **22**, 325–364.

Grossberg, S. (1969b). On learning and energy-entropy dependence in recurrent and nonrecurrent signed networks. *Journal of Statistical Physics*, **1**, 319–350.

Grossberg, S. (1971). On the dynamics of operant conditioning. *Journal of Theoretical Biology*, **33**, 225–255.

Grossberg, S. (1972a). A neural theory of punishment and avoidance, I: Qualitative theory. *Mathematical Biosciences*, **15**, 39–67.

Grossberg, S. (1972b). A neural theory of punishment and avoidance, II: Quantitative theory. *Mathematical Biosciences*, **15**, 253–285.

Grossberg, S. (1972c). Neural expectation: Cerebellar and retinal analogs of cells fired by learnable or unlearned pattern classes. *Kybernetik*, **10**, 49–57.

Grossberg, S. (1975). A neural model of attention, reinforcement, and discrimination learning. *International Review of Neurobiology*, **18**, 263–327.

Grossberg, S. (1976a). Adaptive pattern classification and universal recoding, I: Parallel development and coding of neural feature detectors. *Biological Cybernetics*, **23**, 121–134.

Grossberg, S. (1976b). Adaptive pattern classification and universal recoding, II: Feedback, expectation, olfaction, and illusions. *Biological Cybernetics*, **23**, 187–202.

Grossberg, S. (1978). A theory of visual coding, memory, and development. In E. Leeuwenberg and H. Buffart (Eds.), **Formal theories of visual perception**. New York: Wiley and Sons.

Grossberg, S. (1980). How does a brain build a cognitive code? *Psychological Review*, **87**, 1–51.

Grossberg, S. (1982). **Studies of mind and brain: Neural principles of learning,**

perception, development, cognition, and motor control. Boston: Reidel Press.

Grossberg, S. (1984). Some psychophysiological and pharmacological correlates of a developmental, cognitive, and motivational theory. In R. Karrer, J. Cohen, and P. Tueting (Eds.), **Brain and information: Event related potentials.** New York: New York Academy of Sciences, 58–151.

Grossberg, S. (Ed.) (1987a). **The Adaptive Brain, I: Cognition, learning, reinforcement, and rhythm.** Amsterdam: North-Holland.

Grossberg, S. (Ed.) (1987b). **The Adaptive Brain, II: Vision, speech, language, and motor control.** Amsterdam: North-Holland.

Grossberg, S. (Ed.) (1988a). **Neural networks and natural intelligence.** Cambridge, MA: MIT Press.

Grossberg, S. (1988b). Nonlinear neural networks: Principles, mechanisms, and architectures. *Neural Networks*, **1**, 17–61.

Halgren, E., Squires, N.K., Wilson, C.L., Rohrbaugh, J.W., Babb, T.L., and Crandall, P.H. (1980). Endogenous potentials generated in the human hippocampal formation and amygdala by infrequent events. *Science*, **210**, 803–805.

Ito, M. (1984). **The cerebellum and neural control.** New York: Raven Press.

Kandel, E.R. and Schwartz, J.H. (1981). **Principles of neural science.** New York: Elsevier/North-Holland.

Kasamatsu, T. and Pettigrew, J.D. (1976). Depletion of brain catecholamines: Failure of ocular dominance shift after monocular occlusion in kittens. *Science*, **194**, 206–208.

Kleinschmidt, A., Bear, M.F., and Singer, W. (1987). Blockade of "NMDA" receptors disrupts experience-dependent plasticity of kitten striate cortex. *Science*, **238**, 355–358.

Kandel, E.R. and Schwartz, J.H. (1981). **Principles of neural science.** New York: Elsevier/North-Holland.

Kohonen, T. (1984). **Self-organization and associative memory.** New York: Springer-Verlag.

Kuffler, S.W., Nicholls, J.G., and Martin, A.R. (1984). **From neuron to brain, 2nd edition.** Sunderland, MA: Sinauer Associates.

Levy, W.B. (1985). Associative changes at the synapse: LTP in the hippocampus. In W.B. Levy, J. Anderson, and S. Lehmkuhle (Eds.), **Synaptic modification, neuron selectivity, and nervous system organization.** Hillsdale, NJ: Erlbaum, 5–33.

Levy, W.B., Brassel, S.E., and Moore, S.D. (1983). Partial quantification of the associative synaptic learning rule of the dentate gyrus. *Neuroscience*, **8**, 799–808.

Levy, W.B. and Desmond, N.L. (1985). The rules of elemental synaptic plasticity. In W.B. Levy, J. Anderson, and S. Lehmkuhle (Eds.), **Synaptic modification, neuron selectivity, and nervous system organization**. Hillsdale, NJ: Erlbaum, 105–121.

MacLennan, A.J. and Maier, S.F. (1983). Coping and the stress-induced potentiation of stimulant stereotypy in the rat. *Science*, **219**, 1091–1093.

Malsburg, C. von der (1973). Self-organization of orientation sensitive cells in the striate cortex. *Kybernetik*, **14**, 85–100.

Näätänen, R. (1982). Processing negativity: An evoked potential reflection of selective attention. *Psychological Bulletin*, **92**, 605–640.

Näätänen, R., Gaillard, A., and Mäntysalo, S. (1978). The N1 effect of selective attention reinterpreted. *Acta Psychologica*, **42**, 313–329.

O'Keefe, J. and Nadel, L. (1978). **The hippocampus as a cognitive map**. Oxford: Oxford University Press.

Parker, D.B. (1982). Learning-logic. Invention Report S81-64, File 1, Office of Technology Licensing, Stanford University.

Pettigrew, J.D. and Kasamatsu, T. (1978). Local perfusion of noradrenaline maintains visual cortical plasticity. *Nature*, **271**, 761–763.

Rauschecker, J.P. and Singer, W. (1979). Changes in the circuitry of the kitten's visual cortex are gated by postsynaptic activity. *Nature*, **280**, 58–60.

Rosene, D.L. and Hoesen, G.W. van (1977). Hippocampal efferents reach widespread areas of cerebral cortex and amygdala in the rhesus monkey. *Science*, **198**, 315–317.

Rumelhart, D.E., Hinton, G.E., and Williams, R.J. (1986). Learning internal representations by error propagation. In D.E. Rumelhart and J.L. McClelland (Eds.), **Parallel distributed processing**. Cambridge, MA: MIT Press.

Shaywitz, B.A., Cohen, D.J., and Bowers, M.B. Jr. (1977). CSF monoamine metabolites in children with minimal brain dysfunction: Evidence for alteration of brain dopamine. *Journal of Pediatrics*, **90**, 67–71.

Shekim, W.O., DeKirmenjian, H., and Chapel, J.L. (1977). Urinary catecholamine metabolites in hyperkinetic boys treated with d-amphetamine. *American Journal of Psychiatry*, **134**, 1276–1279.

Singer, W. (1982). The role of attention in developmental plasticity. *Human Neurobiology*, **1**, 41–43.

Singer, W. (1983). Neuronal activity as a shaping factor in the self-organization of neuron assemblies. In E. Basar, H. Flohr, H. Haken, and A.J. Mandell (Eds.), **Synergetics of the brain**. New York: Springer-Verlag.

Swanson, J.M. and Kinsbourne, M. (1976). Stimulant-related state-dependent learning in hyperactive children. *Science*, **192**, 1354–1356.

Thompson, R.F., Barchas, J.D., Clark, G.A., Donegan, N., Kettner, R.E., Lavond, D.G., Madden, J., Mauk, M.D., and McCormick, D.A. (1984). Neuronal substrates of associative learning in the mammalian brain. In D.L. Alkon and J. Farley (Eds.), **Primary neural substrates of learning and behavioral change**. New York: Cambridge University Press.

Weber, B.A. and Sulzbacher, S.I. (1975). Use of CNS stimulant medication in averaged electroencephalic audiometry with children with MBD. *Journal of Learning Disabilities*, **8**, 300–303.

Weiss, G. and Hechtman, L. (1979). The hyperactive child syndrome. *Science*, **205**, 1348–1354.

Werbos, P. (1974). Beyond regression: New tools for prediction and analysis in the behavioral sciences. Ph.D. Thesis, Harvard University, Cambridge, Massachusetts.

Werbos, P. (1982). Applications of advances in nonlinear sensitivity analysis. In A.V. Balakrishnan, M. Thoma, R.F. Drenick, and F. Kozin (Eds.), **Lecture notes in control and information sciences, Volume 38: System modeling and optimization**. New York: Springer-Verlag.

Willshaw, D.J. and Malsburg, C. von der (1976). How patterned neural connections can be set up by self-organization. *Proceedings of the Royal Society of London (B)*, **194**, 431–445.

Zucker, R.S. (1989). Short-term synaptic plasticity. *Annual Review of Neuroscience*, **12**, 13–31.