Title
Non-Hebbian properties of long-term potentiation enable high-capacity encoding of temporal sequences.

Journal
Proceedings of the National Academy of Sciences of the United States of America, 91(21)

ISSN
1091-6490

Authors
Granger, R
Whitson, J
Larson, J
et al.

Publication Date
1994-10-11

Peer reviewed

Proc. Natl. Acad. Sci. USA
Vol. 91, pp. 10104–10108, October 1994
Neurobiology

Non-Hebbian properties of long-term potentiation enable high-capacity encoding of temporal sequences

(non-Hebbian learning/hippocampus/long-term potentiation learning rules)

RICHARD GRANGER, JAMES WHITSON, JOHN LARSON, AND GARY LYNCH

Center for the Neurobiology of Learning and Memory, University of California, Irvine, CA 92717

Communicated by Richard F. Thompson, March 15, 1994

ABSTRACT **A hypothesis commonly found in biological and computational studies of synaptic plasticity embodies a version of the 1949 postulate of Hebb that coactivity of pre- and postsynaptic elements results in increased efficacy of their synaptic contacts. This general proposal presaged the identification of the first and still only known long-lasting synaptic plasticity mechanism, long-term potentiation (LTP). Yet the detailed physiology of LTP induction and expression differs in many specifics from Hebb's rule. Incorporation of these physiological LTP constraints into a simple non-Hebbian network model enabled development of "sequence detectors" that respond preferentially to the sequences on which they were trained. The network was found to have unexpected capacity (e.g., 50 × 10^6 random sequences in a network of 10^5 cells), which scales linearly with network size, thereby addressing the question of memory capacity in brain circuitry of realistic size.**

Physiological evidence has clearly shown that Hebb's postulate ". . . [when cell A] repeatedly or persistently takes part in firing [cell B], . . . A's efficiency, as one of the cells firing B, is increased" (1) is neither necessary nor sufficient for induction of synaptic long-term potentiation (LTP): the target cell (cell B) need not be fired, only depolarized, for LTP to occur (2, 3) (necessity), and yet even repeated firing will not induce LTP unless the depolarization exceeds the *N*-methyl-D-aspartate (NMDA) receptor channel voltage threshold (4) (sufficiency). Nonetheless, there has been a trend in the literature to embrace the generality of Hebb's proposal and to emphasize its similarities to the requirements for the induction of LTP; moreover, theoretical learning rules based on the Hebb or "correlation" postulate yield networks of considerable computational power (5–10). These discrepancies between the physiological constraints on LTP induction versus the Hebb correlation rule raise the question of whether non-Hebbian properties of LTP induction are largely extraneous to, or even impair, the behavioral and computational utility of synaptic plasticity, or whether such properties may yield learning rules that confer useful functional abilities to circuits that use them. Here we derive non-Hebbian LTP induction and expression rules from three physiological results [showing that the (simpler) Hebb rule emerges as a special case] and show that a network using these induction (learning) and expression (performance) rules acts as a high-capacity "sequence detector" that encodes and recognizes very large numbers of temporally patterned cue sequences.

Induction and Expression of LTP

Physiological Characteristics of LTP Induction and Expression. Studies have established that stimulation patterns based on the 4- to 7-Hz theta electroencephalogram rhythm, which appears throughout the olfactory–hippocampal circuit in learning animals (11), are ideally suited for producing robust and stable LTP (12–15). The unlikelihood of complete synchrony of afferents requires extension of LTP induction rules to account for somewhat asynchronous inputs occurring within the envelope of a single peak of the θ rhythm—i.e., brief (<100 ms) sequences of inputs.

A Hebbian coactivity rule would predict that as asynchronous afferents arrive, increased depolarization of the target neuron over the staggered arrival times will cause later inputs to be strengthened more than earlier inputs. However, experiments using asynchronous inputs have shown that this is not the case. Experiments were done using three small inputs stimulated in a staggered sequence (S1–S2–S3) over 70 ms (Fig. 1*a*), testing the induction of LTP in each of the three afferents. The synapses at S1 became potentiated the most, the synapses at S3 were potentiated the least, and the synapses at S2 attained an intermediate degree of potentiation (16).

LTP expression was found to exhibit order dependence corresponding to that found for LTP induction (Fig. 1*b*). Input S1 was potentiated and input S2 remained weak, as would occur in order-dependent LTP induction. Response to the sequence S1–S2, in which the strong (potentiated) input (S1) leads, is significantly larger than response to the opposite sequence S2–S1, in which the weaker (unpotentiated) input leads.

Moreover, the predominant effect of potentiation is on responses generated by the DL-α-amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid (AMPA) subclass of glutamate receptors, which are active during normal stimulation conditions; potentiation does not have a comparable effect on NMDA receptor-dependent potentials (17–19). Thus, subsequent LTP induction episodes are relatively unaffected by prior induction, another effect not predicted from Hebb's postulate and one not included in network models.

Derivation of LTP Learning and Performance Rules. These LTP induction and expression rules were abstracted and incorporated into a simple multicellular network based on the circuit in which the cited LTP rules were found, hippocampal field CA1. The resulting network learns brief simulated sequences lasting roughly 70–100 ms, the duration of the sequences used in the physiological LTP experiments on which the network is based. Intuitively, such brief sequences are commensurate with those that occur within a single syllabic unit of speech or during a single eye saccade.

The learning rule is abstracted from the order dependence of LTP induction (16) described above (Fig. 1*a*). In response to a trained input sequence, the performance rule selectively activates those cells whose synapses are potentiated in the appropriate order, as suggested by the LTP expression experiment in Fig. 1*b*. In particular, the network consists of *A* input lines innervating *M* "competitive" patches (5, 20–22) of *C* cells each, for a total of *CM* cells. The synaptic

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: LTP, long-term potentiation; NMDA, *N*-methyl-D-aspartate; AMPA, DL-α-amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid.

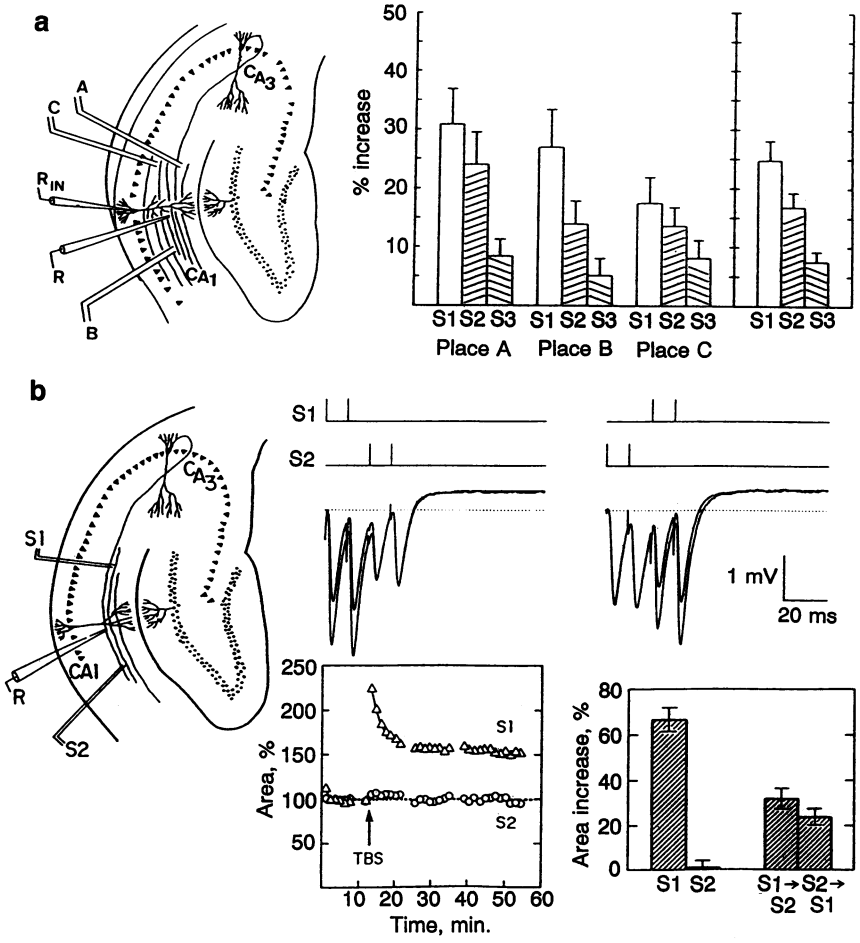

FIG. 1. Physiology of LTP induction via temporal sequence of inputs. (*a*) Order-dependent LTP induction. Sequential stimulation of multiple afferents to a target cell in field CA1 of *in vitro* slice preparations of hippocampus caused differential potentiation of synapses as a function of the order in which inputs were stimulated (16). In 22 experiments, when LTP was induced via stimulation of inputs S1, S2, and S3 in that order, regardless of the location of stimulation (A, B, or C), the degree of LTP induced (shown in the rightmost bar graph) was S1 > S2 > S3 ($P < 0.05$ each). [Reproduced with permission from ref. 16 (copyright 1989 Elsevier Science).] (*b*) Order-dependent LTP expression. The response of a target cell to the sequences S1–S1–S2–S2 and S2–S2–S1–S1 was measured both before and after induction of LTP in S1. (*Bottom Left*) A typical experiment lasting 50 min, with LTP induced in the S1 pathway at 12 min, resulting in stable potentiation of ≈66% ± 5%. (TBS, theta burst stimulation.) (*Top*) Excitatory postsynaptic potentials in response to the sequences S1–S1–S2–S2 and S2–S2–S1–S1, each measured 2 min before and 8 min after the potentiation episode (calibration = 1 mV, 20 ms). (*Bottom Right*) Bar graph shows percentage increases of the extracellular response area to S1 and S2 individually and to the two sequences due to potentiation of the S1 pathway. In 18 experiments, the response to the sequence with the potentiated pathway stimulated first (S1–S1–S2–S2) was significantly greater than the response to the reverse sequence ($P < 0.01$; one-tailed paired *t* test, 17 df = 2.75).

connection matrix W consists of random synaptic weights for the connection between each cell c and input line a. Each input or string X is a temporal sequence of S vectors, each of dimensionality A. Intuitively, then, the learning algorithm processes a temporal input sequence by performing order-dependent potentiation (Fig. 1*a*) of the synapses of those cell(s) that survive or "win" the lateral inhibition in the competitive patches. The performance algorithm is based on the order-dependent LTP expression found in Fig. 1*b*; cells are activated more strongly if the input sequence arrives at the cell in the order of decreasing synaptic strengths of those inputs at the cell. Thus, a target cell will respond selectively to sequences on which it has been trained via order-dependent LTP induction. Finally, LTP expression in the model incorporates the differential effects of LTP on AMPA and NMDA receptors (17, 19). The AMPA currents at a synapse are changed by learning and used during performance, corresponding to the typical connections used in virtually all artificial neural network models. But the NMDA channel, which is used during learning, remains relatively unchanged by prior learning. Thus, prior learning via LTP has little effect on subsequent learning episodes. This rule is a significant departure from the great majority of learning rules in neural network studies. To implement this distinction, the network uses potentiated weights, corresponding to AMPA receptor channels, during performance but uses naive (unchanged) weights, corresponding to NMDA receptor channel currents during learning.

Network Modeling

Network Characteristics. The derived functional LTP induction and expression rules are embedded in a network model whose design is based on prominent architectural features of hippocampal field CA1. Axons of primary excitatory (pyramidal) cells exit CA1 without recurrently contacting other excitatory CA1 cells. Inhibitory (basket) cells are outnumbered by pyramidal cells by ≈2 orders of magnitude, are densely contacted by pyramidal cell axons, and in return densely innervate pyramidal cells within a radius of ≈100 neighboring cells. Prior modeling results (22) have indicated that this patchlike arrangement of a few inhibitory cells densely contacting and contacted by many excitatory cells can yield lateral inhibitory activity of the type modeled by simpler competitive or "winner take all" networks (5, 9, 21).

Existing competitive-network results have focused on synchronously arriving inputs in such networks; the proposed extension to temporal sequences calls for a corresponding extension to the modeling of lateral inhibition in the network. In synchronous competitive or lateral-inhibition networks, target cells that are most strongly activated by the input win the competition and respond, whereas weakly activated target cells are suppressed via lateral inhibitory activity, losing the competition. In the temporal sequence network described here, each input in the temporal sequence gives rise to an increment of lateral inhibition, such that at each step the most weakly activated target cells become inhibited, whereas the rest survive. Thus, the target cell population is "honed" via lateral inhibitory responses to the successive inputs in the sequence, until the remaining cells are those most strongly responding to the overall input sequence.

Learning and Performance Formalisms. Table 1 contains a formal statement of the learning and performance rules to be incorporated into the temporal lateral inhibition network. The learning rule essentially specifies that for every time step s (assumed for simplicity to be divided into S discrete time "windows" each of fixed duration—e.g., 10 ms) in the input sequence X, the target cells ξ_s are subject to competitive lateral

Table 1. Learning and performance rules for the simplified network algorithm

Learning (induction)	Performance (expression)
Item L1. Do items L2–L3 for each input X to be learned.	Item P1. Do items P2–P3 for each input X to be tested.
Item L2. For each sequential step $s \in \{1, 2, \ldots, S\}$, retain active only winning subset of cells $\xi_s \leftarrow \text{win}(X_s, \xi_{s-1}, W^0)$.	Item P2. For each sequential step $s \in \{1, 2, \ldots, S\}$, retain active only winning subset of cells $\xi_s \leftarrow \text{win}(X_s, \xi_{s-1}, W)$.
Item L3. For final winning cell $c \in \xi_s$, perform order-dependent potentiation: train(W_c).	Item P3. Accept (recognize) if each patch responds or else reject.

Each input X is a temporal sequence of S orthogonal vectors each of dimensionality A; i.e., each input can be thought of as a word with each of the S steps corresponding to one letter in it. During learning, each competitive patch in the network performs lateral inhibition at each of the S time steps, at each step inhibiting insufficiently responsive cells. At the end of the temporal sequence, the synapses of the final winning cell(s) c become potentiated with the strongest potentiation at the first input and successively weaker potentiation at subsequent inputs. During performance, lateral inhibition patches behave as during learning; at the end of the input sequence, if each such patch responds with at least one cell surviving the lateral inhibition, then the sequence is recognized (accept or match); if any patch fails to respond, the sequence is not recognized (reject or mismatch).

inhibition $\xi_s \leftarrow \text{win}(X, \xi)$, and the final winning cell c surviving the honing process is trained according to the order-dependence rule described above; earliest inputs potentiate their synapses W the most, and later inputs potentiate successively less.

The performance rule assumes the same discrete time steps for the sequential input. At each step s, the same competitive lateral inhibition successively reduces the set of responsive target cells as described above. Whereas the learning rule used naive (unpotentiated) weights for all synapses, corresponding to activation of the NMDA receptors which induce LTP but exhibit little if any potentiation, the performance rule uses potentiated weights, corresponding to AMPA receptors. As a result of synaptic potentiation during the learning phase, different cells than when naive (unpotentiated) weights were used will now be the most strongly responding. Thus, the honing process now will find different cells as winners. In the network model, as the input sequence arrives, only synapses that were potentiated at the appropriate time step can survive the honing process and remain responsive. Thus, if the input was not previously learned by the network, the likelihood is low that any target cells will respond. Only if the network was trained on this or a related input sequence will the network yield a responsive winning cell at the end of the honing process. Multiple cell groups or patches of cells perform this same operation independently and in parallel.

Computation of the Network. The result is a network that performs an operation of "accept/reject" on its inputs. That is, having been trained on a set of input sequences via the learning rule, the resulting network will tend to recognize or accept those sequences on which it has been trained, and it will tend to reject, or not recognize, any sequences on which it has not been trained. Typical neural network algorithms successfully recognize inputs that are sufficiently similar to those on which training has occurred, even if the input itself has never been seen before; this constitutes "generalization" or "recognition of degraded cues"—i.e., a radius of similarity around learned cues, within which the network will respond positively. The present network performs no generalization; it perfectly remembers sequences on which it has been trained, but it does not generalize to other input sequences. The network exhibits another property, equally unusual for a network model: its capacity. The network uses synapses extremely efficiently, causing a network of given size to be capable of learning a very large number of input sequences with very low recognition error rates. More importantly, the network capacity (number of inputs learnable with fixed recognition error) grows linearly as the network grows, enabling very high capacity from large networks.

Computational Analysis

Formal Derivation of Network Capacity. The capacity of the network can be cast in terms of its errors as a function of the number of sequences stored in a network of a given size. Two types of errors can be distinguished: errors of collision, in which a particular set of target cells respond to more than one temporal string or "word" during training, and errors of commission, in which a target cell responds at testing (performance) time to a string on which it has not been trained. Approximations of the values of these two error rates are derived by the introduction of appropriate simplifying assumptions (Fig. 2). The number of collisions is shown to depend solely on the number n of inputs learned and the number of total possible output configurations, each consisting of the winning cell response from each competitive patch in the network. If Q is the number of distinct output patterns already assigned after n distinct random input patterns have been learned, then the number of collisions is $n - Q$. For a network of M competitive patches of C cells each, there are a total of $J = C^M$ possible distinct output patterns from the network. Let P_1 be the probability that a particular output pattern is activated by an input training pattern; $P_0 = 1 - P_1$ is the probability that a particular pattern is not activated by the input. Then the expected number of assigned output patterns $\overline{Q} = P_1 J = J(1 - P_0)$. If patterns are assigned randomly, then a particular output pattern is assigned with probability $1/J$; the probability of the pattern remaining unassigned after training on n inputs is $P_0 = (1 - 1/J)^n$. Combining these equations, the expected number of collisions is $n - J[1 - (1 - 1/J)^n]$, and, for n inputs, the rate E_L of collision errors is

$$E_L = 1 - \frac{J}{n}[1 - (1 - 1/J)^n]. \quad [1]$$

Unlike the collision rate, the probability of a commission error depends on the number of steps S in the input and the number of input lines A—i.e., the length and dimensionality of the input string. Input sequences are assumed to be relatively brief ($S < 50$), corresponding to asynchronous input stimulation occurring within a single peak of a synchronizing rhythm such as θ. (Such sequences can be thought of intuitively as coherent units such as, e.g., the brief sequence of phonemes comprising a spoken word.) Every learned sequence generates a memory trace consisting of S potentiated synapses or marks, for each entry in the input sequence. Spurious traces or paths through the set of marked synapses also arise with training, and an untrained input will be erroneously accepted or recognized by the net if it activates one of these spurious traces. (For instance, if a given cell is trained on the sequences ABC and DEF, then that cell if tested will falsely recognize sequences such as AEC.) For input vectors of dimensionality A (i.e., A letters in the input "alphabet"), there are A^S possible input sequences of length S. If learning n such strings produces a total of G traces per cell in each of M patches, n/C of which

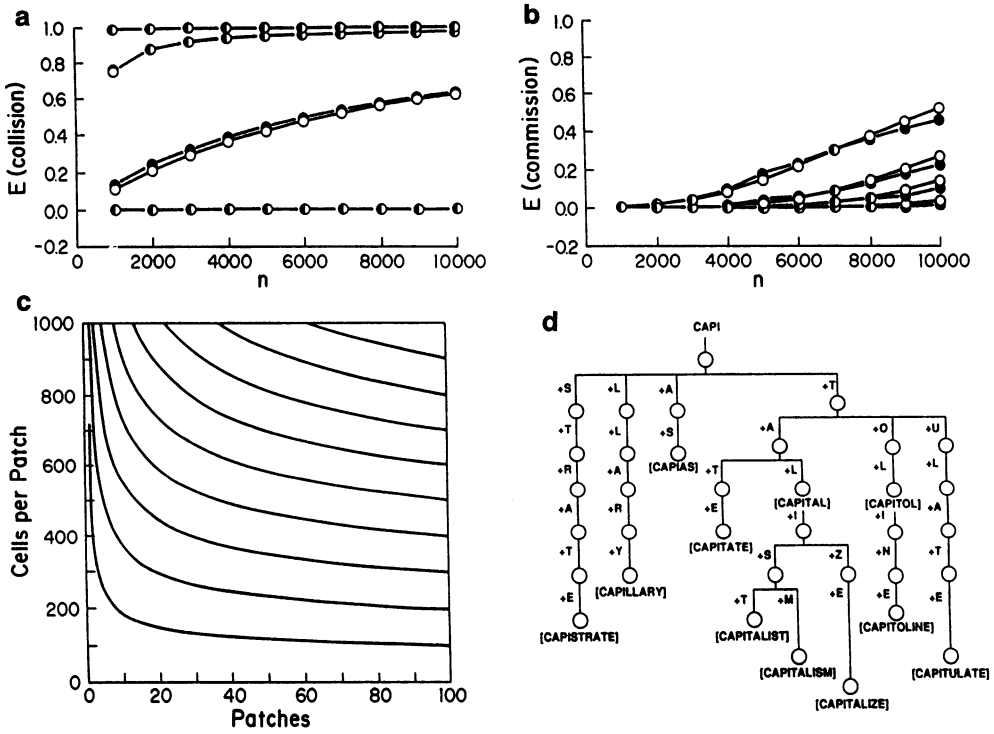

FIG. 2. Theoretical and empirical values for collision and commission errors by the derived algorithm. (*a*) Empirical (●) and theoretical (○) collision error rate (E), as a function of the number of sequences learned (n), plotted for four different networks of increasing size. Sequences learned are of length S = 4 randomly composed from an input alphabet of A = 500 orthogonal letters. Theoretical curves were generated from Eq. 1 (see text) for n = 1000 → 10,000 in steps of 1000. Empirical curves were generated from runs of a simulated network, trained incrementally over 10 training sessions of 10^3 new sequences per session. Training consisted of a single presentation per sequence, with learning proceeding according to the learning rule presented in the text and in Table 1. The four curves correspond to four different-sized networks. Top curve is the collision rate for a network with one patch (M = 1) of C = 16 cells; immediately below it is a network of two patches (32 cells total); next is a network of three such patches. Bottom curve corresponds to an 80-cell, five-patch network. Theoretical curves closely match empirical results, and collision rates shrink dramatically as the size of the network increases. (*b*) Empirical (●) and theoretical (○) percentage errors of commission are plotted for the same four network sizes as above (with 16, 32, 48, and 80 cells). Theoretical curves were generated via the equation for commission errors (Eq. 2), and empirical results were generated from the network as described. Again, the theoretical equation closely matches the empirical results, and, again, errors shrink with increasing network size, becoming vanishingly small for a network of just 80 cells learning 10^4 sequences. (*c*) Capacity and scaling properties of the network. Contour lines in the graph indicate number of sequences that can be learned by networks of various sizes. In particular, the number n of sequences of length S = 10 that can be learned without exceeding a fixed error probability E = 0.001, for networks with an input alphabet of A = 10,000, varying in size up to 100,000 cells. The x axis gives the number of patches of cells in the network (see text), and the y axis denotes number of cells per patch. Each contour curve corresponds to a specific number n of learned sequences. The lowest curve denotes 5×10^6 learned sequences, which can be learned by a network with 5 patches (x axis) of ≈250 cells per patch (y axis), or 40 patches of ≈125 cells per patch, etc. Each contour curve is 5×10^6 learned sequences higher than the one below it. The highest curve shown corresponds to 45×10^6 learned sequences (via networks with, e.g., 60 patches of 1000 cells each, or 100 patches of 900 cells each). The network corresponding to the upper right corner of the graph consists of 100,000 cells (100 patches of 1000 cells each) and has a capacity of 50×10^6 sequences, with error rate $E \leq 0.001$. (*d*) Recognition and sequence completion after training on 10,000 dictionary words. The simplified network algorithm was trained on 10^4 sequence strings of 6–10 letters drawn randomly from a dictionary (see text). The network consisted of 1000 target cells arranged into M = 5 patches of C = 200 cells each, with A = 702 input lines (corresponding to all possible letter pairs constructed from the 26 letters of the alphabet plus a space). The network exhibited 0.002 (0.2%) probability of collision error during training, and 0.01 (1%) probability of commission error during subsequent testing. The letter-order information in each learned word is stored in the network and is exhibited via a simple search algorithm which incrementally appends all possible letters to a given prefix and tests the network for recognition. Branches of the resulting search that led to recognized completions of the prefix CAPI by the 1000-cell network after being trained on the 10,000 words are shown.

are nonspurious (correct) traces, then the rate of commission errors will be $[(G - n/C)/A^S]^M$. If P_F is the probability that a particular synapse at a particular step is part of an accepted trace, then the expected number of synapses active per step is $P_F A$ and the expected number of traces that can be accepted for S steps is $G = (P_F A)^S$. For n randomly chosen strings, in each patch of C cells, a cell wins its patch competition n/C times, marking S synapses on that cell, each with a probability of $1/A$. The probability that a synapse is not marked for each cell that wins a competition is $[1 - (1/A)]^{n/C}$; thus $P_F = 1 - [1 - (1/A)]^{n/C}$, and the rate E_M of errors of commission is arrived at by combining these expressions:

$$E_M = \left(\frac{\{A[1 - (1 - (1/A))^{n/C}]\}^S - n/C}{A^S} \right)^M . \qquad [2]$$

Analytical and Empirical Findings. Fig. 2 *a* and *b* plots these theoretically derived values for collision and commission errors along with the empirical values from simulation runs. Four networks of increasing size were tested for their recognition capabilities after being trained on 10^4 randomly generated input sequences each of length four (S = 4) with a single training trial per input string (see Fig. 2 legend for details). It is seen that the number of both types of error decreases extremely rapidly with increasing network size; with an input alphabet of 500 letters, the 16-cell network (with 500 synapses per dendrite) exhibits nearly 100% errors of collision during training after learning only 10^3 inputs; there are somewhat fewer errors for the 32- and 48-cell networks. The probability of collision error for the 80-cell, five-patch network is only 0.007 (0.7%) over the course of learning the full set of 10^4 input strings, a rate dramatically lower than that of the smaller networks. Intuitively, this is due to the combinatorial nature of the network; with only one or two patches in the network, there are few possible patterns and thus relatively many collisions (identical patterns) generated by different inputs. However, with multiple patches in the network, the number of possible patterns grows factorially, and the probability of two patterns colliding (generating identical patterns) becomes vanishingly small.

Similarly, errors of commission decrease dramatically with increasing network size, from an ≈50% error rate after

training the 16-cell network on 10^4 strings to an error rate of 0.046 (4.6%) for the 80-cell network.

Fig. 2*c* illustrates the scaling properties of the network—i.e., how the capacity changes as a function of the size of the network. In the commissions equation (Eq. 2), for $n \ll A^S$, the quantity $(n/C)/A^S$ may be neglected, allowing Eq. 2 to be approximated by $E_M = \{1 - [1 - (1/A)]^{n/C}\}^{SM}$. Solving for the embedded exponent n yields

$$n = \frac{C\log(1 - E_M^{1/SM})}{\log[1 - (1/A)]}. \quad [3]$$

The figure is a contour graph of the relationship among the number of patches in the network, the number of cells in each patch, and the number of sequences that can be stored without exceeding a fixed recognition error of p (commission) = 0.001. If cells are assumed to receive 10^4 synaptic contacts on their dendrites, then a 100,000-cell network (100 patches of 1000 cells each) stores $\approx 5 \times 10^7$ random sequences of length 10. This corresponds to one learned every minute for 100 years (or, for 8-hr days, an input learned every 10 s for 50 years) with a 0.001 recognition error rate.

Fig. 2*d* gives an example of the ability of the network to both recognize and perform completion on real words rather than on randomly generated letter strings. A network was constructed by using 1000 target cells receiving input from an input alphabet A of 702 input lines corresponding to all possible letter pairs using just the 26 letters of the alphabet (see Fig. 2 legend). When this network was trained on 10^4 words randomly chosen from the dictionary, it exhibited 0.002 (0.2%) collision rate during training and 0.01 (1%) probability of commission error after training. Moreover, the trained network contains order information sufficient to enable a simple search algorithm to complete partial strings into learned words; the figure shows all validly identified completions of the string "CAPI" contained in the 1000-cell network after being trained on 10^4 words.

Discussion

Hebb's 1949 insight bears resemblance to the properties of LTP, yet the detailed physiology and biophysics of long-lasting synaptic plasticity is more complex than is captured by this simple correlational rule, raising the question of whether the learning models that ignore these details are as powerful as (or more or less powerful than) models that incorporate them. The present results begin with detailed physiological induction and expression characteristics of LTP that strictly obey neither Hebb's postulate nor a generalized Hebb-like correlational rule. Rather, the resulting rule is one that depends on temporal order of arrival of afferent activity to a target. It is interesting to note that a correlational or Hebbian rule emerges as the special case occurring only when all afferents arrive simultaneously.

Embedding the LTP induction and expression (learning and performance) rules in a simple network architecture results in a mechanism that stores temporal sequences with high capacity and scales linearly with network size. Most existing computational results on learning of temporal sequences address distinct issues such as time compression and variable length delays, without directly addressing the issue of capacity in terms of error rates, and no study addressing capacity of temporal sequence storage has found capacity greater than that reported here (23–26).

Artificial neural networks typically perform some type of generalization—i.e., respond similarly to similar inputs, even those on which they were not trained—and this capability is often taken as canonical for brain circuits, mapping intuitively onto the need for generalization in living organisms. The present analysis of LTP in field CA1 of hippocampus, however, yields an algorithm that learns temporal sequences without any significant generalization. This of course does not imply that generalization is not performed by organisms or by different brain networks. Rather, it raises the question of functional interpretation of single modular components (e.g., CA1) embedded in much larger systems (e.g., the corticohippocampal pathway). The primary inputs to field CA1 come from hippocampal field CA3, not a sensory field, and the primary outputs from CA1 project to subiculum and parahippocampal (entorhinal) cortex, not motor fields. The relationship of the function of CA1 to that of the overall organism is thus far from straightforward. We hypothesize that a high-capacity, sequence-dependent, nongeneralizing accept–reject or match–mismatch function such as that achieved by this circuit may, in combination with the functions of the other constituent circuits of the corticohippocampal pathway, be of considerable utility in internal memory functions such as recognition of recency, expectation, and changing views with movement (27). Previous work using only synchronous activity (i.e., not temporal sequences) showed that rhythmic θ-like activity during LTP induction in cortical networks did give rise to generalization by performing the computational task of hierarchical clustering (10). Combining those findings with the present results suggests the existence of computationally efficient special-purpose brain circuitry for processing temporal and spatial structure in the learning and recognition of environmental stimuli.

This work was supported in part by the Office of Naval Research (Grant N00014-92-J-1625).

1. Hebb, D. O. (1949) *The Organization of Behavior* (Wiley, New York).
2. McNaughton, B. L., Douglas, R. M. & Goddard, G. V. (1978) *Brain Res.* **157**, 277–293.
3. Kelso, S. R., Ganong, A. H. & Brown, T. H. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 5326–5330.
4. Collingridge, G. L., Kehl, S. L. & McLennan, H. (1983) *J. Physiol. (London)* **334**, 33–46.
5. von der Malsburg, C. (1973) *Kybernetik* **14**, 85–100.
6. Ballard, D. H., Hinton, G. E. & Sejnowski, T. J. (1983) *Nature (London)* **306**, 21–26.
7. Bear, M. F., Cooper, L. N. & Ebner, F. F. (1987) *Science* **237**, 42–48.
8. Kohonen, T. (1984) *Self-Organization and Associative Memory* (Springer, Berlin).
9. Grossberg, S. (1987) *The Adaptive Brain* (North–Holland, Amsterdam), Vol. 2.
10. Ambros-Ingerson, J., Granger, R. & Lynch, G. (1990) *Science* **247**, 1344–1348.
11. Macrides, F., Eichenbaum, H. B. & Forbes, W. B. (1982) *J. Neurosci.* **2**, 1705–1717.
12. Larson, J. & Lynch, G. (1986) *Science* **232**, 985–988.
13. Staubli, U. & Lynch, G. (1987) *Brain Res.* **435**, 227–234.
14. Diamond, D. M., Dunwiddie, T. V. & Rose, G. M. (1988) *J. Neurosci.* **8**, 4079–4088.
15. Pavlides, C., Greenstein, Y. J., Grudman, M. & Winson, J. (1988) *Brain Res.* **439**, 383–387.
16. Larson, J. & Lynch, G. (1989) *Brain Res.* **489**, 49–58.
17. Muller, D. & Lynch, G. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 9346–9350.
18. Muller, D., Joly, M. & Lynch, G. (1988) *Science* **242**, 1694–1697.
19. Kauer, J. A., Malenka, R. C. & Nicoll, R. A. (1988) *Neuron* **1**, 911–917.
20. Grossberg, S. (1976) *Biol. Cybernetics* **23**, 187–202.
21. Rumelhart, D. & Zipser, D. (1985) *Cognit. Sci.* **9**, 75–112.
22. Coultrip, R., Granger, R. & Lynch, G. (1992) *Neural Networks* **5**, 47–54.
23. Tank, D. W. & Hopfield, J. J. (1987) *Proc. Natl. Acad. Sci. USA* **84**, 1896–1900.
24. Gabor, D. (1968) *Nature (London)* **217**, 1288–1289.
25. Jordan, M. I. (1987) *Proceedings of the 8th Annual Conference on Cognitive Science in Society* (Erlbaum, Hillsdale).
26. Taylor, J. (1991) *Int. J. Neural Syst.* **2**, 47–54.
27. Lynch, G. & Granger, R. (1992) *J. Cognit. Neurosci.* **4**, 189–199.