

1993-01

# Context-sensitive spatio-temporal memory

Rinkus, Gerard J.

Boston University Center for Adaptive Systems and Department of Cognitive and  
Neural Systems

---

<https://hdl.handle.net/2144/2008>

*Boston University*

# CONTEXT-SENSITIVE SPATIO-TEMPORAL MEMORY

Gerard J. Rinkus

January 1993

Technical Report CAS/CNS-93-031

Permission to copy without fee all or part of this material is granted provided that: 1. the copies are not made or distributed for direct commercial advantage, 2. the report title, author, document number, and release date appear, and notice is given that copying is by permission of the BOSTON UNIVERSITY CENTER FOR ADAPTIVE SYSTEMS AND DEPARTMENT OF COGNITIVE AND NEURAL SYSTEMS. To copy otherwise, or to republish, requires a fee and/or special permission.

Copyright © 1993

Boston University Center for Adaptive Systems and  
Department of Cognitive and Neural Systems  
111 Cummington Street  
Boston, MA 02215

# Context-Sensitive Spatio-Temporal Memory

Gerard J. Rinkus

Department of Cognitive and Neural Systems  
Boston University, 111 Cummington Street, Boston, MA 02215

**ABSTRACT:** The proposed model, called the combinatorial and competitive spatio-temporal memory or CCSTM, provides an elegant solution to the general problem of having to store and recall spatio-temporal patterns in which states or sequences of states can recur in various contexts. For example, fig.1 shows two state sequences that have a common subsequence, C and D. The CCSTM assumes that any state has a distributed representation as a collection of features. Each feature has an associated *competitive module* (CM) containing  $K$  cells. On any given occurrence of a particular feature, A, exactly one of the cells in  $CM_A$  will be chosen to represent it. It is the particular set of cells active on the previous time step that determines which cells are chosen to represent instances of their associated features on the current time step. If we assume that typically  $S$  features are active in any state then any state has  $K^S$  different neural representations. This huge space of possible neural representations of any state is what underlies the model's ability to store and recall numerous context-sensitive state sequences. The purpose of this paper is simply to describe this mechanism.

**1. Introduction:** The type of solution to the context-sensitivity problem illustrated in fig. 1 that is used in the CCSTM is similar to the combinatorial memory scheme described in Lynch (1986) and in Miller (1991). The use of CMs in the model essentially implements an *orthogonalization* transformation over the input space. Kanerva's (1988) Sparse Distributed Memory (SDM) is similar in this regard although the SDM internal representation space is  $\{0,1\}^n$  (large  $n$ ) whereas the CCSTM internal representation space is  $\{0..K\}^n$ . The CCSTM's combinatorial approach differs significantly from models which utilize unitary representations of whole subsequences such as the Masking Field model of Cohen & Grossberg (1987).



Figure 1: Two state sequences,  $[ABCDEF]$  and  $[GHCDIJ]$ , have a common subsequence.

**2. Motivation of Model:** The following series of examples is intended to motivate the essential architectural feature of the CCSTM – the use of competitive modules (CMs). Initially (in fig. 2), assume there are no CMs. Instead there is just a fully-connected field of feature-detecting neurons. (Note that in this and subsequent figures, only synaptic connections pertinent to the example are explicitly shown). Suppose the model experiences the spatio-temporal pattern,  $\{\{K\},\{C, D, L\}\}$ , where this notation means that feature K is sensed on the first time step and features C, D, and L

are sensed on the following time step. Assume the model use the following *Hebbian* learning law.

**LR1:** each cell active on time step  $t$  increases its synaptic weight onto every cell active at  $t + 1$ , unless the synapse has already been increased.

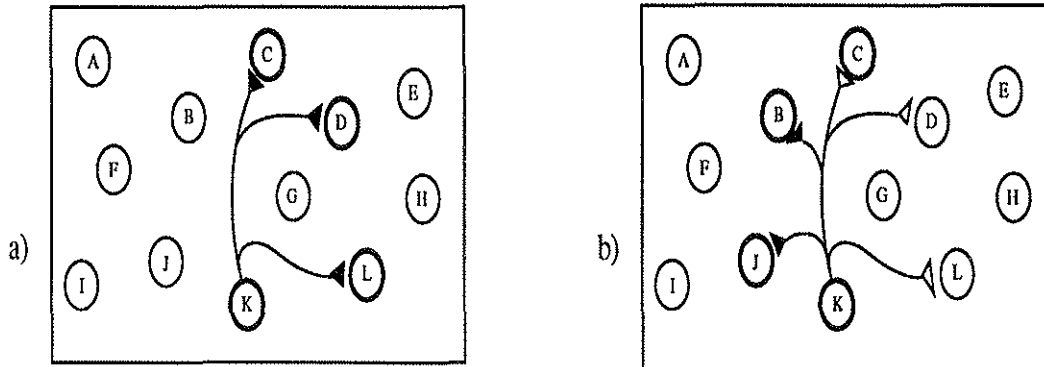


Figure 2: a) weights updated following presentation of pattern 1 =  $\{\{K\},\{C, D, L\}\}$ . b) weights increased due to pattern 2 =  $\{\{K\},\{C, B, J\}\}$  are highlighted in black.

Fig. 2a shows those weights that would be increased following presentation of pattern 1 =  $\{\{K\},\{C, D, L\}\}$ . We could then prompt for recall of pattern 1 by turning on cell K. Assume that during recall a cell will become active if it receives one or more signals via an increased synapse. Let the number of input signals via increased weights that is necessary to activate a cell be called the *recall threshold*,  $\Theta$ .

Now suppose we present pattern 2 =  $\{\{K\},\{C, B, J\}\}$  to the network. Fig. 2b shows synaptic increases due to pattern 2 highlighted in black. The problem is that if we now try to recall either of these patterns, they will interfere with each other. If we prompt the network by presenting feature K at time  $t$ , then the union of the two second time step patterns -  $\{C, D, L\} \cup \{C, B, J\} = \{C, D, L, B, J\}$  will become active at  $t + 1$ .

The problem in the previous example is the lack of specificity of effect that a single cell can generally have. This problem can be remedied by making the assumption that at least several cells will be active on any given time step. In this case we can rely on the *context* (i.e. set of other simultaneously active cells) within which a cell is active to *specify* the effect that cell has in turning on cells on the next time step. Fig. 3a depicts the learning that is due to a new pattern, pattern 3 =  $\{\{K, O, P\},\{C, D, L\}\}$ . Fig. 3b depicts the learning due to pattern 4 =  $\{\{K, M, N\},\{C, B, J\}\}$ . In this case, we can now set  $\Theta$  equal to either 2 or 3 and obtain perfect recall of either pattern even though cell K is common to both of their first time steps. If we set  $\Theta = 2$  and turn on cells K, O, and P, then cells C, D, and L will fire because they each have three large (i.e. increased) active synapses, but cells J and B will each have only one large active synapse and thus not fire. A similar analysis shows correct recall of pattern 4. Thus context differentiates or specifies the effect cell K has on subsequent neural activity.

Note also that this model is robust against cell death. For example, if cell P were to die, and assuming  $\theta = 2$  (but not 3), then both patterns would still be recalled completely correctly except for cell P itself.

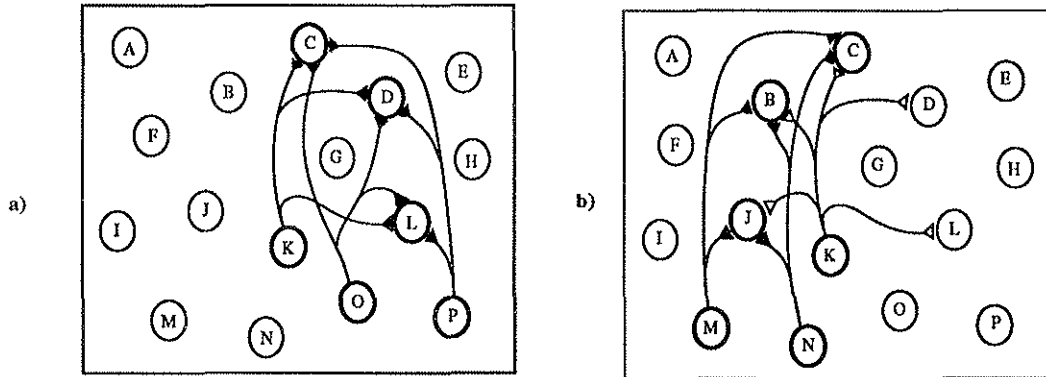


Figure 3: a) weights updated following presentation of pattern 3 =  $\{\{K, O, P\}, \{C, D, L\}\}$ . b) weights increased due to pattern 4 =  $\{\{K, M, N\}, \{C, B, J\}\}$  are highlighted in black.

But even with the assumption of distributed representations on each time step, problems can still occur. Specifically, what if the the network must store and recall state sequences in which states can recur in varying contexts (as shown in fig. 1). Fig 4. shows that this causes essentially the same problem we had originally (in fig. 2). Suppose that instead of presenting pattern 4, a pattern 5 =  $\{\{K, O, P\}, \{C, B, J\}\}$  were presented to the network. Figs. 4a,b show the synaptic situation following presentation of patterns 3 and 5, respectively. If we then reinstate the spatial activation pattern,  $\{K, O, P\}$ , the cells  $\{C, D, L, B, J\}$ , all become active. There is no way to set  $\Theta$  in order to enable correct recall.

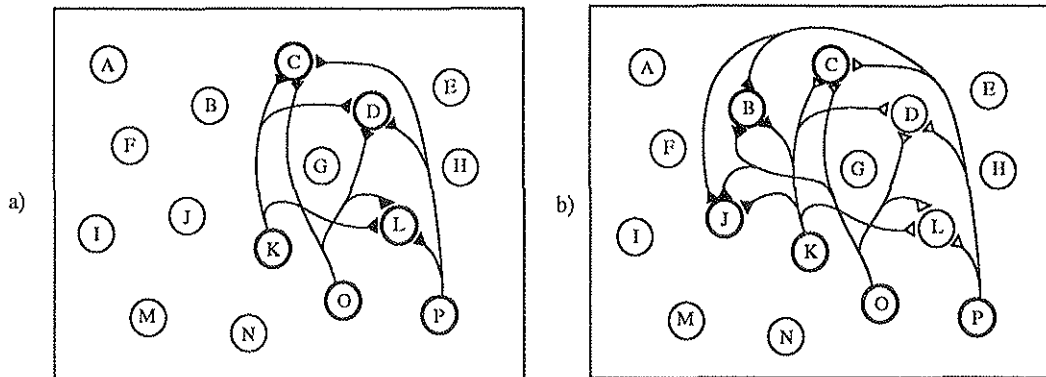


Figure 4: a) weights increased following presentation of pattern 3 =  $\{\{K, O, P\}, \{C, D, L\}\}$ . b) weights increased due to pattern 5 =  $\{\{K, O, P\}, \{C, B, J\}\}$  are highlighted in black.

The CCSTM solution to this general problem is the introduction of competitive modules (CMs) as shown in fig. 5. Now each feature can be represented by any of a number of cells in its associated CM rather than by just a single cell as before. On any given occurrence of a feature, exactly one of the cells in its associated CM will become active to represent it. If a particular feature is present on some time step  $t$ , then each cell in that feature's CM summates the signals arriving from other cells that were active on the previous time step. The one with the highest sum is the winner and becomes solely active in that CM on that time step. At that point the learning law is applied. In addition to the *Hebbian* LR1, a second *anti-Hebbian* learning law, LR2, is also in operation.

**LR2:** each synaptic weight from a cell that is *inactive* at time  $t$  onto a cell that is *active* at time  $t + 1$  is decreased, unless it has already been changed (either increased or decreased).

Figs. 5a,b show patterns 3 and 5 (respectively) again, but this time in the enriched representational format. Note that the two actual neural representations of the state  $\{K, O, P\}$  have only one cell,  $K_2$ , in common. Therefore, we can set  $\Theta$  equal to either 2 or 3 and get perfect recall of both patterns even though, at the featural level, they have the exact same initial state. The number of different neural representations of the state  $\{K, O, P\}$  is  $4^3$ . All of these representations differ in at least one CM. Thus, if we set  $\Theta = 3$ , this network could store a set of spatio-temporal patterns containing up to 64 occurrences of the state  $\{K, O, P\}$ , provided that all 64 unique representations were actually chosen by the model

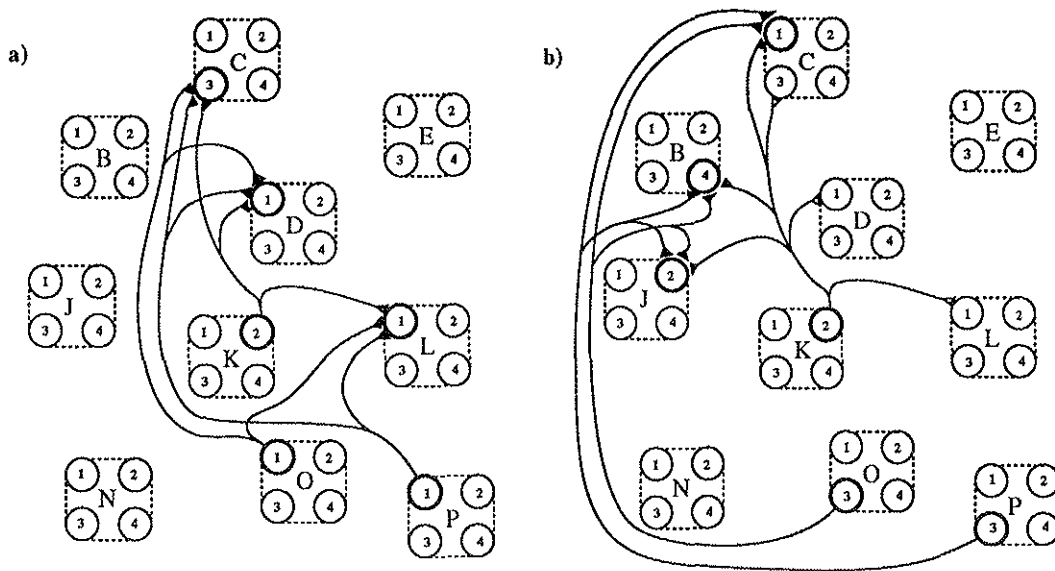


Figure 5: a) synaptic increases due to pattern 3 in the CM-based model b) the synaptic increases due to pattern 5 are highlighted in black.

Even for modest values of  $K$  and  $S$  like 100 and 20, respectively, the number of representations is huge. If the model chooses sparsely from that space, then any two chosen representations will have relatively few common cells. This allows us to set  $\Theta$  to some value much less than  $S$ . Two of the important benefits of being able to set  $\Theta$  low are a) the model can store spatio-temporal patterns in which some states can have less than  $S$  (but greater than  $\Theta$ ) features, and b) the model has significant robustness against cell death.

### 3. References:

1. Grossberg, S. and Cohen, M.I.. (1987) *Applied Optics* 26, 10
2. Kanerva, P. (1988) *Sparse Distributed Memory* Cambridge, MA, Bradford/MIT Press
3. Keeler, J.D. (1988) *Cognitive Science* 12:299-329
4. Lynch, G. (1986) *Synapses, Circuits, and the Beginnings of Memory* The MIT Press, Cambridge, MA
5. Miller, R. (1991) *Cortico-Hippocampal Interplay and the Representation of Contexts in the Brain* Springer-Verlag, Heidelberg, Germany