TESTING THE ROLE OF ASSOCIATIVE INTERFERENCE AND COMPOUND CUES IN SEQUENCE MEMORY

Frances S. Chance and Michael J. Kahana

Volen Center for Complex Systems
Brandeis University
Waltham, MA
E-mail: `chance@binah.cc.brandeis.edu`
E-mail: `kahana@fechner.ccs.brandeis.edu`

ABSTRACT

A major question facing computational models of human memory concerns the storage and retrieval of sequentially processed information. Many current models assume a chaining of associations. According to this view, each item, or memory pattern, is associated with the preceding item in the sequence. In reproducing the sequence, each recalled item serves as a retrieval cue for the next item. A serious problem facing these chaining models is their susceptibility to associative interference. This paper presents a novel experimental method designed to assess the effects of associative interference in the retrieval of ordered lists of items. Experimental findings presented here suggest that subjects use multiple prior items, as well as context, to overcome the effects of associative interference in list recall.

1. INTRODUCTION

The hypothesis that the pattern of neural activity at time t is associated with the pattern of neural activity at time $t + \tau$ is central to most neural network and abstract mathematical models of sequence memory (Abbott & Blum, 1996; Lewandowsky & Murdock, 1989; Sompolinsky & Kanter, 1986; Kleinfeld, 1986; Riedel, Kuhn & van Hemmen, 1988). This notion of chained associations, in which each recalled item in a sequential list facilitates recall of the next item, formed the cornerstone of Ebbinghaus's (1885) seminal research on human learning. During the 1960's experimental psychologists began to find holes in the notion of pure-chaining and proposed that chained, positional, and compound item associations were all factors in ordered

recall (see Young, 1968 for a review). These early studies, based on the "transfer of training" technique, failed to provide consistent evidence for or against a particular associative model.

A basic problem for all associative models is how subjects cope with interference from competing associations. Consider the challenge facing subjects when two lists containing a common element must be reproduced (e.g., A-B-C-D-E-F and P-Q-R-B-S-T). According to a chaining model the common element (B) is involved in two competing associations (B-C and B-S). Consequently, recall of the list-appropriate item following the shared element will be drastically impaired.

In this paper, a direct approach to studying the problem of associative interference in sequence memory is presented. In this new methodology, human subjects learn two sequences of unique items (common words) to a performance criterion. If the sequences share a single item, the two sequences may be thought of as "crossing paths". If the sequences share a sub-sequence of items, the sequences will have joined and then divided. Consider the sequences A-B-C-D-E-F-G-H and U-V-W-X-E-F-Y-Z. These two sequences contain a common subsequence E-F. In a pure-chaining model in which each recalled item facilitates recall of the following item, F is now associated with both G and Y. Thus a decrease in performance would be expected when the subject tries to recall the letter following F. However, in the case of a compound cue model, in which some combination of prior items facilitates recall of the following list item, there is no ambiguity between the cues for G and Y and the decrease in performance should be attenuated. Rather than measuring accuracy, we have chosen to examine a more sensitive measure of performance – namely, the inter-response time (IRT) at the exit transition out of the shared sub-sequence (i.e. F-G or F-Y).

2. METHODS

Twenty four paid volunteers from the Brandeis community participated in one practice session followed by four experimental sessions. In each experimental session, subjects learned two lists of fifteen words until they could recite them perfectly on three successive trials following a short distractor task. The words were randomly sampled from among the 300 highest frequency nouns in the Toronto Word Pool (Friendly, et al., 1982) and displayed, one at a time, on a computer monitor. Lists were chosen so as to "cross paths" for one, two, four, or eight words. To discourage positional coding, the overlapping items always began at a random point in the list such that there were at least three words following and three words preceding which were not included in the other list. At the end of each session subjects recalled both lists three times. IRTs were determined using a computer-assisted scoring program. IRTs deviating 3σ or more from the mean of each condition were excluded (σ represents the standard deviation for a particular position relative to the exit transition).

3. RESULTS

The results of an earlier unpublished study suggest that IRTs are not affected by list position, except possibly for those times at the very beginning and end of the list (results not shown). As discussed in the methods, the overlapping items of the two lists were placed so that the exit transition did not fall onto these positions.

Although all subjects were instructed to recite the list as quickly as possible without losing accuracy, IRT data was still extremely variable, as is typical in studies examining IRTs
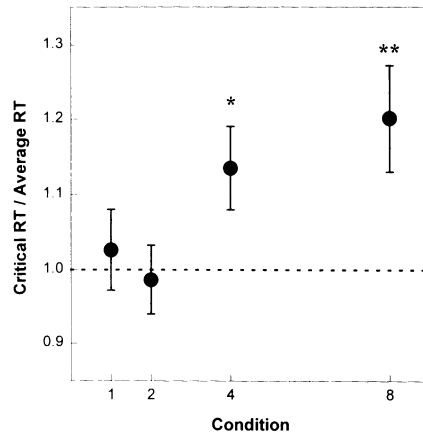
Figure 1. The effects of one, two, four, and eight word "crossed paths". On the vertical axis is the mean exit IRT (normalized). The four and eight word overlap conditions are significantly different from baseline ($p < 0.05$ in the four-word condition and $p < 0.01$ in the eight-word condition).

in recall tasks. To control for this, IRTs were normalized by dividing each IRT by the mean IRT for that list. Thus an unaffected transition should approximately equal the mean IRT and the ratio should equal one. If the IRT is significantly longer than the mean IRT, the ratio will be significantly greater than one.

Figure 1 shows the effect of increasing the number of overlapping words on the ratio of the critical (exit) IRT to the average list IRT. A significant increase in IRTs is observed for the four ($p < 0.05$) and eight ($p < 0.01$) word overlap conditions.

According to any simple associative chaining model of sequence memory, we would expect that even a single shared element should create substantial interference in recalling the item following the shared element in either list. The obtained IRT data fail to demonstrate associative interference when one or two consecutive elements are shared across two lists. These findings are consistent with data on intraserial repetition effects: if an item is repeated within a single list (e.g., A-B-C-D-E-F-C-G-H), IRTs to items following the repeated elements are not elevated (Kahana & Loftus, in press). Taken together with the results reported here, it seems clear that human subjects are relatively immune from associative interference in sequence learning tasks. This finding, coupled with the significant interference observed when large subsequences are shared across two lists, suggest the role of compound cues in sequence recall. The next section presents a preliminary mathematical analysis of the role of context and compound cues in associative models of sequence memory.

4. THEORETICAL ANALYSIS

In a chaining model, associations are formed between each item and its predecessor. Thus the effective stimulus is the prior item. If we represent item i in lists one and two as the vectors $\mathbf{a}_i$ and $\mathbf{b}_i$, the weight matrix which codes for the associations which make up the two lists is given by:

$$\mathbf{W} = \sum_{i=1}^{L_1} \mathbf{a}_i \mathbf{a}_{i-1}^T + \sum_{i=1}^{L_2} \mathbf{b}_i \mathbf{b}_{i-1}^T .$$

In this equation, vectors are columns, the superscript T above a vector denotes its transpose, and L_1 and L_2 are the lengths of lists one and two respectively.

When cued with an item which has two competing associates in the list (for instance, $\mathbf{a}_n\mathbf{a}_{n-1}^T$ and $\mathbf{b}_m\mathbf{a}_{n-1}^T$), the following information is retrieved:

$$\mathbf{W}\mathbf{a}_{n-1} = \mathbf{a}_n + \mathbf{b}_m.$$

In a non-linear model, both attractors will have equivalent strength and the model will not be able to disambiguate to the desired target. According to chaining models, the size of the overlapping region of the lists should not affect performance at the ambiguous exit point.

Adding context to a chaining model provides a powerful basis for list discrimination. In such a model, features representing item information may be added to features representing contextual information. If the vector features are sparse, addition should not produce many overlapping elements.

Two different patterns of activity, $\mathbf{x}$ and $\mathbf{y}$, represent context in lists one and two respectively:

$$\mathbf{a}' = \mathbf{a} + \mathbf{x}$$
$$\mathbf{b}' = \mathbf{b} + \mathbf{y}.$$

In this model, the effective stimulus is the prior item combined with the appropriate list context. The weight matrix for two lists sharing a common element ($\mathbf{a}_{n-1} = \mathbf{b}_{m-1}$, but $\mathbf{a}'_{n-1} \neq \mathbf{b}'_{m-1}$) is given by:

$$\mathbf{W} = \sum_{i=1}^{L_1} \mathbf{a}'_i\mathbf{a}'^T_{i-1} + \sum_{i=1}^{L_2} \mathbf{b}'_i\mathbf{b}'^T_{i-1}.$$

Consider the information which is retrieved when the memory is cued with $\mathbf{a}'_{n-1} = \mathbf{a}_{n-1} + \mathbf{x}$.

$$\mathbf{W}(\mathbf{a}_{n-1} + \mathbf{x}) = 2\mathbf{a}_n + \sum_{i \neq n}^{L_1} \mathbf{a}_i + \mathbf{b}_m + L_1\mathbf{x}.$$

In this model both associates are retrieved together with all of the $\mathbf{a}$ terms as well as a very strong contextual term ($\mathbf{x}$) list one. If the vectors are partitioned into context and content elements, we can retrieve the desired pattern $\mathbf{a}_n$.

In a compound cue model, associations are stored between the current item and an exponentially weighted sum of all prior items. The strength of the associations are greatest for the immediately preceding item and fall off exponentially:

$$\mathbf{W} = \sum_{i=1}^{L_1} \mathbf{a}'_i \sum_{j=0}^{i-2} e^{-kj}\mathbf{a}'^T_{i-j-1} + \sum_{i=1}^{L_2} \mathbf{b}'_i \sum_{j=0}^{i-2} e^{-kj}\mathbf{b}'^T_{i-j-1}$$

When the memory matrix is cued with $\mathbf{a}_{n-1}$:

$$\mathbf{W}(\sum_{j=0}^{n} e^{-kj}[\mathbf{a}'_{n-j-1}]) = 2\sum_{j=0}^{n} e^{-2kj}\mathbf{a}_n + \sum_{j=0}^{S} e^{-2kj}\mathbf{b}_m + \ldots$$

In this model, the size of the overlapping region, S, increases the interference produced by the major competitor, $\mathbf{b}_m$. Thus the size of the overlapping region affects performance of

this model, unlike the chaining model with context. The effect of context is to double the cue strength for the list-appropriate response a_n — making this model far more robust in the presence of noise. Note: the omission of numerous weakly matching terms in the equation above are indicated by the ellipses.

5. SUMMARY AND CONCLUSIONS

We have presented a novel experimental method for examining the effect of associative interference in recalling sequences of discrete items. Human subjects' performance during ordered recall of word lists is adversely affected when a sequence of several words is shared in two simultaneously learned lists. Contrary to the predictions of simple associative chaining models, the presence of one, or even two, shared items does not result in elevated inter-response times when exiting the shared sequence. However, significant interference is observed following shared sequences of four and eight words. These results support models in which multiple prior items combine to form compound cues used in retrieval of subsequent list items.

REFERENCES

[1] Blum, K. I., & Abbott, L. F. (1996) A Model of Spatial Map Formation in the Hippocampus of the Rat. *Neural Compuatation* 8: 85-93.

[2] Ebbinghaus, H. (1885/1913 [Reprinted by Dover, 1964]). *Memory: A Contribution to Experimental Psychology.* Teachers College, Columbia University, New York.

[3] Friendly, M., Franklin, P. E, Hoffman, D., Rubin, D. C. (1982) Nouns for the Toronto Word Pool. *Behavior Research Methods and Instrumentation* 14: 375-379.

[4] Kahana, M. J. & Loftus, G. (in press). Response Time and Accuracy in Human Cognition: Are They Two Sides of the Same Coin? In R. Sternberg (Ed.) *The Concept of Cognition.* MIT Press.

[5] Kleinfeld, D. (1986) Sequential State Generation by Model Neural Networks. *Proc. Natl. Acad. Sci. USA* 83: 9469-9473.

[6] Lewandowsky, S. & Murdock, B. B. (1989) Memory for Serial Order. *Psychological Review* 96: 25-57.

[7] Riedel, U., Kuhn, R., & van Hemmen, J. L. (1988) Temporal Sequences and Chaos in Neural Nets. *Physical Review A* 38: 1105-1108.

[8] Sompolinsky, H. & Kanter, I. (1986) Temporal Association in Asymmetric Neural Networks. *Physical Review Letters* 61: 259-262.

[9] Young, R. K. (1968) Serial Learning. In T. R. Dixon & D. L. Horton (Eds) *Verbal Behavior and General Behavior Theory.* Prentice Hall, England Cliffs, NJ. pp. 122-148.