

Department of Psychology
University of California
Los Angeles, California 90024

(net address: Jmiller @ Rand-Unix)

Dr. Feigenbaum,

This letter concerns my request for SUMEX-AIM access. I am currently a graduate student in cognitive psychology at UCLA (as well as a consultant at Rand-ISD), and am in the process of completing my dissertation [[1]]: a simulation model of semantic memory with the goal of predicting recent and future experiments on memory for the meaning and wording of sentences and texts. The work on my model, a UCI LISP program, was about one-third completed when the computer on which I had been working was removed to another section of UCLA. I cannot get time on the machine at this time, and its future plans are very unclear: Lisp may not even be provided should it later become available. Therefore, I would like to move my dissertation work to SUMEX, where a number of Lisp systems are of course available.

The current state of my work is outlined in the status report enclosed, but a short summary here might be appropriate. There have been a number of psychological experiments on the retention of sentences and texts that indicate that people generally have very good memory for the meaning of this material, but very poor memory for the exact words in which it was presented: discrimination of a presented sentence from a new, but synonymous, sentence is rarely much above chance. Correspondingly, the prevailing view in psychology is that incoming linguistic material is converted from its component words into an abstract, conceptual structure that represents the meaning of the passage. Wording information is available for a short period of time, but is soon lost. This position has been advocated in psychology by the Norman and Rumelhart group at UCSD, and in computer science by Roger Schank and his associates.

At the same time, some of the work here at Rand on text comprehension (by Barbara and Fredrick Hayes-Roth and Perry Thorndyke) has indicated that there are very real and observable effects of wording variation on the recall and recognition of sentences. For instance, subjects are much more successful at forming inferences between two sentences when the common components of those sentences are identical, rather than paraphrases of each other. If there were an underlying abstract representation, synonymous phrases should be represented identically, and no differences in inferencing ability should be observed. Hayes-Roth and Hayes-Roth have also found that, despite the near-chance recognition performance of sentence wording, subjects are reliably faster and more confident when verifying old than paraphrase sentences. This work suggests that

words are differentially represented in memory, but that some memory processes are occluding these differences.

My contention is that a semantic network controlled by a limited capacity spreading activation process could account for these results. This feature of the model encodes a sentence into memory by spreading activation throughout those parts of the network that are semantically related to the parts of the sentence. New words are incorporated into the context of the studied sentence as they are contacted by this spreading process. While such a process will build a strong representation of the sentence into memory, it will come at the cost of interference with the wording information, since both a word ("frightened") and one of its synonyms ("scared") may be encoded into the same context. A recognition test would then falsely recognize "John scared the babysitter" after having studied "John frightened the babysitter", presuming that "scared" had been activated by the model's study of "frightened".

The model can further predict the Hayes-Roth findings of higher confidence and faster reaction times to old sentences. Although both old and paraphrase sentences will be "recognized" as above, confidence ratings to old sentences should be higher as a result of the more highly activated state of such sentences. The words in old sentences would be directly activated by the full amount of activation which initially entered the sentence, while synonymous and other related words would receive only a portion of that original activation: the limited capacity feature of the model requires a division of activation entering a word node among all the links leaving that node. Confidence ratings will then vary as a function of the amount of activation at the node, being greater for older, more highly activated words. A similar manipulation could be expected to produce the finding of shorter reaction times to old sentences.

The current state of the model is this: a set of experimental sentences have been selected and a representation of the words in these sentences has been constructed from their dictionary definitions. The initial study and test routines have also been completed, so that the model is capable of "studying" a series of sentences, and then determining whether or not a particular sentence had been studied earlier. As previously mentioned, the model is falsely "recognizing" new sentences which are synonymous to previously studied sentences. The initial successes described above and in the enclosed status report indicate that the general direction of the research is an appropriate way to deal with the problem. What remains to be done is to develop the aspects of the model responsible for generation of correct test reaction times and confidence ratings, and expansion of the model to simulate a number of related memory phenomena (see the status report for instances of these).

As I mentioned before, I think that my work here is about one-third finished; ideally, I would like to finish my

dissertation by January. My needs at SUMEX actually seem quite moderate: the current model was running in about 40K in UCI LISP, and I don't foresee any drastic increase in that (although I will likely move the model to INTERLISP). Disk space should also be minimal: probably five to ten 30-40 block files. Given my other projects at UCLA and Rand, I would probably be using about 10 hours of terminal time per week; as much as one-half of that would likely be during evenings and weekends.

I hope to hear from you soon, and I hope that any affiliation between us will be productive. If you need any further information, please let me know. Thank you.

Sincerely,

James R. Miller

[[1]]

My dissertation committee is:

Robert A. Bjork, Department of Psychology, UCLA
Thomas D. Wickens, Department of Psychology, UCLA
Jackson L. Beatty, Department of Psychology, UCLA
Kenneth M. Colby, Department of Psychiatry, UCLA
Tyler Burge, Department of Philosophy, UCLA
Barbara Hayes-Roth, The Rand Corporation