Our Ongoing Memory Modeling Research: An Informal Overview

Robert Lindsey (robert.lindsey@colorado.edu)
Michael Mozer (mozer@colorado.edu)
Department of Computer Science
University of Colorado at Boulder

November 6, 2010

The way we study material influences how well we retain it. Psychologists have established that spaced practice leads to better long-term retention of declarative knowledge (facts such as foreign language vocabulary) than massed practice. However, determining the appropriate amount of time between practice sessions is complex and depends on the duration of time over which the material must be available for recall. We are exploring existing and novel computational models to explain a range of data on massed versus spaced practice.

We have developed a model that predicts the optimal spacing of study in a wide range of experiments, over retention intervals ranging from minutes to a year. The model makes surprising and counterintuitive predictions that we are currently testing.

We are also developing mechanistic accounts that explain other phenomena surrounding the conditions under which individuals are likely to learn and retain material. For example, when an individual is tested and then told the correct answer, their learning is better than when they merely study the material. We explain this finding with an error correcting learning account in which a better error signal is obtained when the individual produces a response, even if incorrect. We've also addressed the counterintuitive result that when individuals are willing to guess an answer, even when they are wrong, they will learn the material with less study than when they are unwilling to venture a guess. We have developed a model that links the strength of learning to willingness to guess (or confidence in a guess).

One of our most exciting projects involves the development of "intelligent" electronic tutoring systems. Though there are numerous tutoring systems out there that claim to exploit the spacing effect (e.g. SuperMemo), they suffer many problems stemming from a lack of scientific rigor. These existing systems are based on simple heuristics and intuitions, not on formal, quantitative models of human memory. They generally just increase the time to next study for items that are correctly recalled, and decrease the time for items that are not. Most critically, heuristic systems at best adapt only crudely to differences among individual learners and materials to be learned.

The large literature on the spacing effect shows that the "optimal" spacing of study depends heavily on the retention interval. The best way to space study depends not just on the nature of the material and student, but also on the time period over which the material will be needed. The relationship between the spacing of study and subsequent availability of the material at some later test is complicated, yet is something our memory model can predict quite well. With that in mind, we developed and are testing an online vocabulary tutoring system called the Colorado Optimized Language Tutor (COLT). COLT has three distinguishing features:

- It optimizes recall for a specific time window.[1] As opposed to existing systems, "optimize" is quantitatively defined and thus can be tackled using standard tools from Bayesian decision theory.
- It explicitly represents various sources of uncertainty and uses probabilistic inference to reason under that uncertainty. For example, we assume that we do not have any control over how often each student studies vocabulary on our system. COLT represents its belief of how often a student will study as a probability distribution. As time goes by and it observes how often they login and study, COLT updates its beliefs (in the sense of probabilistically conditioning its original belief on the observations). In this way, COLT can adapt to individual differences in material and students.
- Students can login to COLT when they want and study as long as they want. Despite this lack of control, COLT is opportunistic and makes use of each login session to optimize spacing.

We currently have a local upper-division Spanish class engaged in a semester-long study on COLT. COLT is trying to maximize each student's test performance on a final cumulative exam scheduled for December 13, 2010. Studying is done in a simple flashcard-like format: students are presented with a cue (i.e., a Spanish word), type in a response (i.e., the corresponding English word), receive feedback, and then repeat the process. The only thing COLT has control over is the decision of which vocabulary pair to present to a student next. Nonetheless, it can make selections so as to influence the spacing of study for an item. Because of our success at modeling experimental data, we believe we can make reasonably accurate predictions of the impact of this decision on the student's final exam score, despite significant uncertainty — uncertainty in how much and when students will study in the future, uncertainty in the difficulty of individual items, and uncertainty in the student's ability and willingness to study outside of the tutoring system. Taking these forms of uncertainty into consideration, COLT chooses vocabulary pairs whose further study it predicts will most benefit the student.

There are many exciting directions we could pursue with this research. Many, if not most, educational settings require the memorization of large sets of facts for use in a specific time window. For example, students studying large lists of obscure vocabulary for standardized tests could benefit greatly from our system. Such students presumably don't care about imprecise notions of long-term retention like what's promised by the hand-waving of current systems (cf SuperMemo, SpacedEd, smart.fm, etc..). They simply need to maximize their test performance for the date their exam is scheduled.

[1]More precisely, it maximizes expected recall given some probability distribution over retention intervals