Latent Semantic Analysis: What is it and how can it improve and assess student learning?

Brent A. Olde, University of Memphis
Email: Baolde@memphis.edu

AutoTutor is a computer tutor that simulates natural discourse while executing pedagogically appropriate conversational turns. Currently there are two versions of AutoTutor, one designed for tutoring college students in an introductory computer course, the other designed for introductory physics. This paper describes a key component of AutoTutor's ability to "comprehend" student responses, latent semantic analysis (LSA). This natural language understanding technique has many practical applications for technologically sophisticated strategies of enhancing learning and aiding teaching. Our goal is to present an overview of LSA, briefly explore uses of LSA, and then discuss how it is implemented in our tutoring system.

Keywords: AutoTutor, Intelligent Tutoring Agent, Latent Semantic Analysis (LSA)

Links: **http://141.225.14.51/atclient/**
http://mnemosyne.csl.psyc.memphis.edu/trg/AutoTutor/trg/index.htm
http://lsa.colorado.edu/
Baolde@memphis.edu

Latent Semantic Analysis: What is it and how can it improve and assess student learning?

Brent A. Olde, Arthur C. Graesser, and the Tutoring Research Group
Department of Psychology, 202 Psychology Building
University of Memphis, Memphis, TN 38152 USA

Introduction

AutoTutor is an intelligent tutoring agent that interacts smoothly with a student using natural language dialogue. By smoothly it is meant that the tutor's interactions are not limited to single word answers or formulaic yes/no decision trees. Our research attempts to tackle the problem of understanding the ungrammatical, real word responses that flow naturally when a student and tutor interact, take those responses and reply in a pedagogically appropriate manner. It is this cooperative, constructive one-on-one dialogue that affords tutoring impressive learning gains (Chi, de Leeuw, Chiu, and LaVancher, 1994; Cohen, Kulik, & Kulik, 1982). To handle this formidable problem we use a recently emerged procedure using Latent Semantic Analysis (LSA). LSA is a statistical, corpus-based natural language understanding technique that computes similarity comparisons for terms and texts. Researchers have been exploring the strengths and limitations of LSA and have continually demonstrated convincing results of the technique's robustness and versatility in the domain of educational learning. The goal of this paper is to present an overview of LSA, briefly explore how it can be used in educational applications, and then discuss how it is used in our tutoring system.

What is Latent Semantic Analysis

LSA is a mathematical technique in which the information contained in the co-occurrences of words in a body of text is compressed into a set of vectors in N-dimensional space. The input to LSA is a set of texts related to the educational topic of interest. These texts need to be segmented into documents; a document is an arbitrarily defined unit, but normally is a sentence, paragraph, or section in a text (we typically used paragraphs as our document size since it provides a good approximation of an idea unit). The words in the document and the complete documents themselves are then subjected to mathematical transformations that take into account the frequency of the words and their use in the documents. Using the mathematical process of singular value decomposition, a word by document matrix (which would mainly be filled with zeros since the majority of the words do not occur in most paragraphs) is linearly decomposed into independent principal parts (broken into 3 separate matrices). All but the largest N singular values are set to zero, and the matrices are then re-multiplied to produce a matrix of the same dimensions as the original matrix (for a good review see Landauer & Dumais, 1997; Landauer, Foltz, & Laham, 1998)

By discarding some of the singular values when re-formulating the word by document matrix, a better representation of the text's meaning is captured. For instance, a positive relationship emerges between the coefficients in the rows corresponding to different words, if the words have similar or associated meanings. This new matrix can be used to evaluate the conceptual relatedness between any two "bags of words", an unordered set of one or more words. The match (i.e., similarity in meaning, conceptual relatedness) between two bags of words is computed as the geometric cosine between the two associated vectors, with values that normally range from 0 to 1. Using the cosines as an indicator of how similar two documents are, LSA has successfully predicted the coherence of successive sentences in text (Foltz, Kintsch, & Landauer, 1998), the similarity between student answers and ideal answers to questions (Graesser, Wiemer-Hastings, et al., 2000; Wiemer-Hastings et al., 1999), and the similarity between student essays and an ideal essay (Landauer, Foltz, & Laham, 1998)

What can been done with LSA

In this section we will briefly describe some of the research that indicates LSA has a real future in enhancing learning and aiding teaching. Let's first look at how LSA's performance compares with human performance. Landauer and Dumais (1997) compared LSA's performance to that of non-native English speakers on the Test of English as a Foreign Language (TOEFL). LSA was trained on a text of 5 million words derived from high-school level encyclopedia texts. LSA (just like the students) had to choose one of 4 alternative words/phrases most similar in meaning. LSA obtain 64% correct, which was the average for students from non-English countries. Not only did LSA perform on par with students, but the errors it committed were correlated (r = .44) with student errors. They also performed this same task using an introductory Psychology test. LSA was trained on an introductory Psychology textbook and given the same 4-alternative multiple choice test as students. LSA, received a 60%, lower than class average but still better than chance.

LSA not only takes tests, but also grades essays. It has been shown that LSA-based essay graders assign grades as reliably as experts in composition, even when the essays are not well-formed grammatically or semantically (Foltz, Gilliam, & Kendall, 2000). Landauer, Laham, & Foltz (1998) demonstrated in essays covering several diverse topics (The heart, introductory Psychology, the Panama Canal, and tolerance of diversity in America) that the correlation between LSA's score and an average human score was no different than the inter-rater reliability between human scorers (see Table 1). This result indicates that LSA can be a useful time saving tool. Foltz has taken this process one sep further by having his Introductory Psychology Class write essays using the web and having LSA grade the submitted essays. If the student is not satisfied with the grade, LSA can provide feedback as to the type of information that was lacking, and then the student can rewrite the essay for a better grade.

Table 1

	The Heart	Introductory Psychology	The Panama Canal	Tolerance of Diversity in America
LSA-human correlation	.80	.64	.68	.84
Human-human correlation	.83	.65	.54	.82

Another interesting application for LSA is in the field of second language acquisition. Landauer and Littman (1990) trained an LSA corpus with a list of translated text and the original text all in the same document. The resulting performance showed equivalent functioning independent of query language.

In our research LSA has been used to evaluate the quality of student contributions in interactive dialogs with a computer tutor. The computer system, called AutoTutor, has been designed to tutor in the domains of computer literacy and physics. The LSA module evaluates the quality of student answers to questions almost as reliably as graduate students (Graesser, Wiemer-Hastings, et al., 2000; Wiemer-Hastings, et al., 1999). Having established the utility of LSA we will now go into more detail on how we use LSA to evaluate student responses to our tutoring systems.

Overview of AutoTutor

AutoTutor's style of tutoring was modeled after actual human tutoring sessions (Graesser, Person, & Magliano, 1995; Wiemer-Hastings, Graesser, Harter, & the TRG, 1998). The tutor starts out by asking a question or posing a problem, then works with the student to cover the essential points that the tutor deems necessary to adequately understand the answer to the question. When one problem is understood, the process is repeated. AutoTutor's knowledge of its tutoring domain resides in a curriculum script. This is a list of the questions or problems that the tutor is prepared to handle in a tutoring situation. A major portion of the script is the LSA space; it gets created from an assortment of

texts collected from the domain of interest. For the physics version of AutoTutor, LSA was trained on two conceptual physics text. The script also contains very specific information about a set of questions or problems planned to be covered. This specific information is comprised of a relatively lengthy "ideal" answer. This answer is separated into a set of specific good answers, which address one aspect of the ideal answer. For each aspect of the ideal answer there are hints, prompts, and elaborations to help the student construct an appropriate answer. AutoTutor, using the LSA derived cosine matches, evaluates the student's response and compares it to an expected good response for the question that was just put forth. Based on the results, an appropriate dialog move is executed – such as asking a follow-up question, or summarizing the current problem.

Acknowledgments

This research was supported by grants or contracts from the Office of Naval Research (N00014-00-1-0600, N61339-01-C1006) and the National Science Foundation (SBR 9720314 and REC 0106965). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author and do not necessarily reflect the views of ONR or NSF.

References

Chi, M. T. H., de Leeuw, N., Chiu, M., and LaVancher, C. (1994). Eliciting self-explanations improves understanding. *Cognitive Science, 18,* 439-477.

Cohen, P. A., Kulik, J. A, and Kulik, C. C. (1982). Educational outcomes of tutoring: A meta-analysis of findings. *American Educational Research Journal, 19,* 237-248.

Foltz, W., Gilliam, S., & Kendall, S. (2000). Supporting content-based feedback in on-line writing evaluation with LSA. *Interactive Learning Environments, 8,* 111-128.

Foltz, P. W., Kintsch, W., & Landauer, T. K, (1998). The measurement of textual coherence with latent semantic analysis. *Discourse Processes 25,* 285-307.

Graesser, A. C., Person, N. K., & Magliano, J. P. (1995). Collaborative dialogue patterns in naturalistic one-to-one tutoring. *Applied Cognitive Psychology, 9,* 359.1-28.

Graesser, A.C., Wiemer-Hastings, P., Wiemer-Hastings, K., Harter, D., & Person, N., and the TRG (2000). Using latent semantic analysis to evaluate the contributions of students in AutoTutor. *Interactive Learning Environments, 8,* 128-148.

Landauer, T. K., & Dumais, S. T. (1997) A solution to Plato's problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge. *Psychological Review, 104,* 211-240.

Landauer, T. K., Foltz, P. W., & Laham, D. (1998). An introduction to latent semantic analysis. *Discourse Processes, 25,* 259-284.

Landauer, T. K., Laham, D., & Foltz, P. W., (1998). Learning human-like knowledge by Singular Value Decomposition: A progress report. In M. I. Jordan, M. J. Kearns & S. A. Solla (Eds.), *Advances in Neural Information Processing Systems 10,*(pp. 45-51). Cambridge: MIT Press.

Landauer, T. K., & Littman, M. L. (1990). Fully automatic cross-language document retrieval using latent semantic indexing. In *Proceedings of the 6th Annual Conferenceof the Centre for the New Oxford English Dictionaryand Text Research* (pp. 31-38). Waterloo, Ontario: UWCentre for the New OED.

Wiemer-Hastings, P., Graesser, A.C., Harter, D., and the Tutoring Research Group (1998). The foundations and architecture of AutoTutor. *Proceedings of the 4th International Conference on Intelligent Tutoring Systems* (pp. 334-343). Berlin, Germany: Springer-Verlag.

Wiemer-Hastings, P., Wiemer-Hastings, K., Graesser, A. & TRG (1999). Improving an intelligent tutor's comprehension of students with latent semantic analysis. In S. Lajoie & M. Vivet (Eds.), *Artificial intelligence in education* (pp. 535-542). Amsterdam: IOS Press.