

Implementing Latent Semantic Analysis in Learning Environments with Conversational Agents and Tutorial Dialog

Arthur C. Graesser (a-graesser@memphis.edu)

Department of Psychology, 202 Psychology Building, University of Memphis, Memphis, TN 38152 USA

Xiangen Hu (xhu@memphis.edu)

Department of Psychology, 202 Psychology Building, University of Memphis, Memphis, TN 38152 USA

Brent A. Olde (baolde@memphis.edu)

Department of Psychology, 202 Psychology Building, University of Memphis, Memphis, TN 38152 USA

Matthew Ventura (mventura45@hotmail.com)

Department of Psychology, 202 Psychology Building, University of Memphis, Memphis, TN 38152 USA

Andrew Olney (aolney@hotmail.com)

Department of Psychology, 202 Psychology Building, University of Memphis, Memphis, TN 38152 USA

Max Louwerse (mlouwers@memphis.edu)

Department of Psychology, 202 Psychology Building, University of Memphis, Memphis, TN 38152 USA

Donald R. Franceschetti (dfrncsch@memphis.edu)

Department of Physics, University of Memphis, CAMPUS BOX 523390, Memphis, TN 38152 USA

Natalie Person (person@rhodes.edu)

Department of Psychology, Rhodes College, 2000 N. Parkway, Memphis, TN 38112 USA

We have been developing learning environments with animated conversational agents. The agents manage a mixed-initiative dialog between the learner and the computer system either by a direct conversational interaction or by serving as a navigational guide on a web site. Two of the systems simulate human tutors by (a) presenting difficult questions that require deep reasoning, (b) attempting to comprehend the learner's typed input, (c) formulating dialog acts that are sensitive to the learner's contributions, and (d) speaking to the student with the animated agent. AutoTutor teaches computer literacy whereas Why/AutoTutor teaches conceptual physics (Graesser, VanLehn, Rose, Jordan, & Harter, 2001). The Human Use Regulatory Affairs Advisor (HURAA) teaches officers in the military about the ethical use of human subjects on a web site with a search facility that accesses documents through questions posed in natural language.

All three systems have used Latent Semantic Analysis (LSA) as its primary representation of world knowledge. LSA is a statistical technique that compresses a large corpus texts into a space of 100-500 dimensions (Landauer, Foltz, & Laham, 1998). The K-dimensional space is used when evaluating the similarity between any two bags of words, with values ranging from 0 to 1. From the standpoint of AutoTutor and Why/AutoTutor, one bag of words is the set of assertions that a student expresses within a dialog turn; the other bag of words is the content of the curriculum script

for a particular topic. From the standpoint of HURAA, one bag of words is the learner's query in natural language and the other is a paragraph in the document space. LSA has generally been successful in evaluating the quality of student explanations, in evaluating the quality of student assertions in tutorial dialog, and in the retrieval of documents from natural language queries. Successes and failures of LSA are identified in these three learning environments.

Acknowledgements

This research was supported by the Office of Naval Research (N61339-01-C1006), the Institute for Defense Analyses (AK-2-1801), the National Science Foundation (REC 0106965), and the DoD Multidisciplinary University Research Initiative (MURI) administered by ONR under grant N00014-00-1-0600. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of ONR, IDA, DoD or NSF.

References

- Graesser, A.C., VanLehn, K., Rose, C., Jordan, P., & Harter, D. (2001). Intelligent tutoring systems with conversational dialogue. *AI Magazine*, 22, 39-51.
- Landauer, T.K., Foltz, P.W., Laham, D. (1998). An introduction to latent semantic analysis. *Discourse Processes*, 25, 259-284.