Applications of Latent Semantic Analysis To Lessons Learned Systems

Michael J. Strait and Jacqueline A. Haynes
Intelligent Automation, Inc.
Rockville, MD, 20850
mstrait@i-a-i.com
jhaynes@i-a-i.com

Peter W. Foltz
Knowledge Analysis Technologies, LLC
Boulder, CO, 80303
pfoltz@knowledge-technologies.com

Abstract

This paper will present several examples of the application of Latent Semantic Analysis (LSA) to practical problems of knowledge acquisition, management and assessment. The purpose of this presentation is to make other knowledge management (KM) and artificial intelligence (AI) researchers aware of the value of LSA as an automated technique for improving the utility of Lessons Learned (LL) and similar knowledge and information management systems.

Introduction

What is Latent Semantic Analysis? Simply put, LSA is a set of automated statistical procedures for deriving a quantitative measure of the similarity in meaning between two words or groups of words. We say "group of words" rather than something like paragraph because it is completely immaterial whether the group of words possesses any syntactic structure. We will not take the time to describe LSA in great detail in this paper, since that information is readily available elsewhere in print (Landauer, Foltz, and Laham 1998) and online (http://lsa.colorado.edu).

LSA uses singular value decomposition (SVD), a general form of factor analysis, to condense a very large matrix of word-by-context data into a much smaller, but still large -- typically 100-500 dimensional – representation (Deerwester et al. 1990).

The similarity between resulting vectors for words and contexts, is measured by the cosine of their contained angle, and has been shown to closely mimic human judgments of meaning similarity and human performance based on such similarity in a variety of ways.

The focus our discussion of LSA here is on characteristics that bear most directly on its value with respect to LL systems.

There are many types of lessons learned and LL systems. The current discussion is intended to address a broad range of LL systems in military, corporate, and educational settings. In the military settings, for example, we are concerned with systems that collect and distribute observations regarding tactics, techniques and procedures critical to successful training and operations. In corporate settings, we are interested in LL systems that function primarily as corporate memories (van Heijst, van der Spek, and Kruizinga, E. 1996). In education settings, formalized LL systems are the least common, but functional equivalents are emerging in the forms of digital reference services and so-called "Ask-A" services using Internet technologies (Lankes 1999).

The applications of LSA reported here can improve the utility of LL systems in many forms, both with respect to the organization and dissemination of lessons learned. We will briefly describe each of the sample applications, and then summarize the characteristics of LSA manifest in these examples that seem of most value to improving the utility of LL systems.

Example Applications

The examples presented here are drawn from a series of research and development collaborations between Intelligent Automation, Inc. (IAI), of Rockville, Maryland, and Knowledge Analysis Technologies, LLP (KAT), of Boulder, Colorado. These collaborations were performed under contract with NASA, DARPA, and the US Navy.

Maintenance Manual Management System (MMMS)

The practical problem addressed by this project involved identifying the various portions of text in a complex maintenance manual that must be updated when any change is made in the equipment or procedures described (Haynes and Miller 1998).

We constructed a "semantic space" using the text from 182 "sub-procedures" representing the decomposition of an actual aircraft "Power Plant – Removal/Installation" procedure, which included about 200 pages of text and figures. The decomposition was accomplished by segmenting the Power Plant – Removal/Installation procedure using a small number of fully automated grammatical principles, similar to decomposing a novel into its constituent paragraphs. Each sub-procedure ranged in size from several words to several hundred words. LSA is typically trained on much larger bodies of text, but in this case we wanted to demonstrate that the limited lexicon of engine maintenance would promote positive results, and serve as proof-of-concept, of using a relatively small corpus.

This project demonstrated two previously untested capabilities of LSA that are important for future applications to Lessons Learned systems. In the past, LSA has been shown to work well as a search aid when applied to a large corpus of non-procedural (i.e., declarative) natural language text. Maintenance procedures differ from standard LSA training data in that the set of words and grammatical constructions used in procedures is more restricted than in unbounded declarative text. Maintenance procedures also contain a large number of figures and diagrams that cannot be directly evaluated by LSA techniques.

Our results demonstrated that LSA worked just as well with procedural text as it does with declarative text, and that we were able to successfully search out relevant figures much like text, by including a figure's labels, caption and other pieces of text as its own "sub-procedure" in our corpus.

LSA again proved its superiority over simple keyword matching, meaning that queries returned relevant matching results in many cases where there were no common terms between the query and the result. For example, the general query "protective caps" was used to find a variety of procedures concerned with the installation of protective caps. However, using our sample semantic space, none of the 14th-20th best matching sub-procedures mention protective caps at all, but are related to them in that they refer to items that you would put protective caps on (such as fuel lines).

E-mail Filtering System

The primary goal of this project was to establish the feasibility of using LSA to filter electronic mail messages and sort them into an appropriate set of folders, acting as an administrative assistant that separates topical mail from noise (Strait 1999).

We investigated the relative effectiveness of filtering using a semantic space comprised exclusively of email messages and one comprised exclusively of documents covering topics of interest (but including no email message text).

We constructed the email-only semantic space from the text messages from the Reagan-era National Security Council, that were made public under the Freedom of Information Act, and subsequently published in a book entitled White House E-Mail. The few hundred messages published in the book were supplemented with approximately 1,000 additional messages from the same source obtained from the National Security Archive. We then constructed a second, more generic space, from approximately 1,200 abstracts of scientific and technical documents concerning national security and related topics found at the Defense Technical Information Center web site (http://www.dtic.mil/stinet/).

Given these two semantic spaces, we tested their ability to filter sample messages from the published White House E-mail corpus. Our expectation was that filtering based upon the NSC e-mail corpus would be more accurate, but we were most interested in determining the relative performance of the two semantic spaces, since this would allow us to gauge the applicability of a general-purpose, and potentially customizable space.

As we expected, the email-specific semantic space was much more effective in categorizing test messages from the same source, while the general purpose semantic space performed just a little better than chance. Looking at the DTIC documents, we concluded that there just wasn't enough overlap of the critical terms between the DTIC corpus and the e-mail messages to achieve better results. Because of this, the computed similarity between messages was not always based on the terms that were likely most important for categorizing the messages (such as names and particular events within the White House.)

This finding raises the important point that while LSA can find and measure similarity of meaning without requiring the same words to be present in the compared texts, the words one wants to be able to compare must occur somewhere in the large corpus of text used to create the semantic space. [This is artificial intelligence, but not magic!] Nevertheless, we believe the results were sufficiently different from chance to indicate the feasibility of using the DTIC abstracts for a more generalized corpus.

Team Readiness Assessment Program (TRAP)

The objective of the TRAP project was to develop a technical framework to link Navy Mission Essential Tasks (NMTLs) to tactical training curricula for assessing training effectiveness (Strait and Haynes 1999). TRAP was proposed as an assessment tool that could be used to measure the effectiveness of training for both individuals and teams. TRAP will use LSA to measure similarities between textual descriptions of what is to be learned and textual descriptions of trainees' performance in training exercises.

The first challenge however was that no training tool exists today that links doctrine, tasks, conditions, and outcome measures found in the Universal Navy Task Lists (UNTLS) with tactics, techniques and procedures (TTP) for training use. TRAP was proposed to provide that critical missing performance and training effectiveness assessment methodological framework and capability for both individual and team training. A great deal of TTP information exists in the form of Lessons Learned data. We have proposed that LSA be used not only to assess training performance and the effectiveness of training, but also to help construct Navy Mission Essential Task Lists (NMETLs) and devise training scenarios based on Lessons Learned.

We were not able to work directly with data from the Navy's Lessons Learned Database, be we were able to demonstrate how LSA could be used to develop a model of team knowledge requirements contained in the Surface Ship Firefighting Manual, a basic component of Damage Control training. Surface Ship Firefighting is part of the Damage Control Task training area specified in Naval Tactical Task List (NTTL) 2.0. The textual representation of the domain was the full text of Naval Ships' Technical Manual (NSTM) 555, Volume 1.

The manual was broken up into separate sections representing individual components of knowledge that a team member would need in order to perform firefighting tasks. Based on this representation of knowledge requirements, we demonstrated how we can use measures of individuals' knowledge to provide an assessment of the extent to which a team of individuals have the right combination of knowledge required to perform the team task. We simulated the knowledge of trainees and the knowledge required by the team, and showed how LSA can be used to compute the similarity of a trainee's knowledge to that required by the task and how well that trainee's knowledge will add to the team's knowledge for successful task completion.

This demonstrated capability can easily be extended to search and include relevant knowledge from LL systems in development of NMETLs and for purposes of individual and team training assessment based on NMETLs.

Adult Literacy Tutor (LADDER)

Last but not least, IAI and KAT have recently received an SBIR Phase 2 award from the Navy (ONR, specifically) to build a reading comprehension tutoring system that relies upon LSA for multiple functions in lesson construction, assessment of student progress, and automated guidance through lesson content (Haynes and Strait 1999).

The tutoring system, called LADDER, builds on knowledge of the characteristics of adult learners and acquisition of literacy skills by using LSA to deliver reading passages selected to match relevant aspects of a reader's skills and knowledge. LADDER uses an instructional approach based on scaffolding, which moves readers up a 'Reading Ladder' in a step-wise fashion, based on an individual's growing capability in reading specific text. The steps of the ladder relate to levels of complexity in the structure and content of the text, so that readers gradually increase the difficulty of text they read as they continue interacting with the tutor.

An important feature of LADDER is that scaffolding is accomplished by using the individual reader's knowledge of the subject area before and after reading each passage to identify the next passage to be read, and to assess the reader's comprehension of text throughout the instructional process.

In our Phase 2 research, IAI and KAT will be exploring automated procedures for using LSA to identify Lessons Learned data available online that augments text-based reading material and incorporate it during the authoring process to enhance the learning experience and add motivation and interest for the student. This work is in progress.

Summary

The text components of Lessons Learned Systems are of indeterminate size, form, function and content. The characteristic of lessons learned most common across all examples is that it is a reported observation. Because the data they contain is often highly situational, the language used is frequently informal and cursory. Even though some LL systems may be more formal, there are no widely applied standards of form and content

The example applications described above show that LSA is capable of addressing some of the most difficult challenges of accessing, evaluating, and utilizing knowledge across multiple LL systems.

Applications of Latent Semantic Analysis as exemplified above and in related research show promise as automated techniques for improving the utility of Lessons Learned (LL) and similar knowledge and information management systems.

References

Landauer, T. K., Foltz, P., and Laham, D. 1998. An Introduction to Latent Semantic Analysis. *Discourse Processes*, 25: 259-284.

Deerwester, S., Dumais, S.T., Furnas, G.W., Landauer, T.K. and Harshman, R.A. 1990. Indexing by Latent Semantic Analysis. *Journal of the American Society for Information Science*, 41: 391-407.

van Heijst, G., van der Spek, R., and Kruizinga, E. 1996. Organizing Corporate Memories. In *Proceedings of the 10th Banff Workshop on Knowledge Acquisition for Knowledge-Based Systems*. Available online at http://ksi.cpsc.ucalgary.ca/KAW/KAW96/vanheijst/HTMLDOC.html.

Lankes, R. David. 1999. Building and Maintaining Internet Information Services: K-12 Digital Reference Services. ERIC Clearinghouse on Information & Technology, Syracuse University, Syracuse, NY.

Haynes, J. and Miller, M. 1998. The KCAP Maintenance Manual Management System, Technical Report, Intelligent Automation, Inc.

Strait, M. 1999. E-mail Filtering Project, Technical Report, Intelligent Automation, Inc.

Strait, M and Haynes, J. 1999. Team Readiness Assessment Program, Technical Report, Intelligent Automation, Inc.

Haynes, J and Strait, M. 1999. LADDER, Technical Report, Intelligent Automation, Inc.