Ernest F. Martin, Jr., Ph.D. efmartin2@vcu.edu
KEYWORD – ASSESSMENT

Martin, E.F., Jr. 2004 “Assessment Outcome Coherence Using LSA Scoring” Academic Exchange Quarterly, Fall, 2004.

Assessment Outcome Coherence Using LSA Scoring

Ernest F. Martin, Jr., Virginia Commonwealth University

Ernie Martin, Ph.D. is Associate Professor of Public Relations in the School of Mass Communications.

Abstract
This paper introduces latent semantic analysis (LSA) as an automated, statistically reliable metric for comparing desired outcomes or objectives across educational programs. After a general introduction to LSA, demonstrating the ability of LSA to compare texts contexts, we sketch an LSA metric (assessment outcomes/objectives scoring tool) for measuring the relatedness of program outcomes and objectives across a university. This metric quickly identifies programs with outcome/objective sets that are very divergent from normative objective/outcome sets.

Introduction
A plethora of stakeholders perceives higher education as extremely important and increasingly costly. They are pushing universities for accountability data relative quality of university programs and the institution as a whole. Universities have responded with assessment programs as initial steps in meeting the demands of governing boards, state educational bodies, regional and professional accreditation organizations, politicians, students, parents, and many other stakeholders.

One mid-Atlantic university encourages all academic and support programs to use a web-based program assessment tool for reporting mission, outcomes/objectives, measures, findings and actions. The tool allows real-time review, consolidated reporting, tracking of actions, and other analysis of all program assessment efforts. The university is complex, enrolling 26,000 students in more than 170 programs in the arts, sciences and humanities and 112 support units on multiple campuses, including an international location.

Because of the variety of programs and the desires of faculty and administrators to write outcomes/objectives in the way that best fit their individual program, the web-based program assessment tool allows a great deal of flexibility for individual programs to construct their own narratives. At the same time, however, the institution must keep on top of the process.

This exploratory study adapts Latent Semantic Analysis (LSA) to create a metric to indicate quickly programs where additional review of activities is necessary.

Previous Research on LSA

Cognitive scientists and computational psycholinguistics have developed LSA as a statistical model for comparing semantic similarity of units of text to each other. The LSA statistic computations represent the contextual-usage meaning of words in a text (Landauer & Dumais, 1997). LSA is more completely described in Landauer, Foltz & Laham (1998) and the seminal article of Deerwester, Dumais, Furnas, Landauer and Harshman, (1990).

LSA provides an automatic method for comparing units of textual information to each other in order to determine their semantic relatedness. It was originally designed to improve information retrieval methods by performing retrieval based on derived "semantic" content of words in a query (e.g. "Googling") as opposed to performing direct word matching.

The development of LSA was at Bellcore (now Telcordia) in the late 1980s. The United State Patent 4,839,853 was granted on June 13, 1989 to inventors Scott C. Deerwester, Susan T. Dumais, George W. Furnas, Richard A. Harshman, Thomas K. Landauer, Karen E. Lochbaum, Lynn A. Streeter. A cross-language information retrieval patent followed (US Patent 5,301,109, April 5, 1994).

The specific application of LSA for informational retrieval is now called latent semantic indexing (LSI). What is the difference between LSA and LSI? Simply, LSI refers to using the approach for indexing or information retrieval; LSA refers to all other applications. The Telcordia LSI reference site provides excellent background materials. (http://lsi.argreenhouse.com) Additionally, a very comprehendible discussion of the information retrieval process using the technique is from the National Institute of Technology in Liberal Education, javelina.cet.middlebury.edu/lsa/out/lsa_intro.htm

The theoretical assumption of LSA is there is some underlying or "latent" structure in the pattern of word usage across documents. Meanings of a word modify as we use a word in different contexts. Similarly, different words sometimes have the same meaning depending on contexts. The idea of latent semantic analysis is that the aggregate of all the word contexts in which a particular word does, and does not, appear provides a set of mutual constraints that reflects the similarity of meaning of words to each other. (Foltz, 1996)

Since LSA is automatic, without human coders, there are no constraints on the size of the text analyzed. This permits efficient and economical analysis of the semantic structure of large bodies of text.

In short, this measure is based on a powerful statistical analysis of direct and indirect relations among words and passages in a large text corpus and can capture the extent to which two text units are discussing semantically related information.

Research Questions

Can LSA be adapted to help some aspects of program assessment construction and implementation? The beginning exploration asks:

1. Can LSA provide a metric to indicate relatedness of outcomes/objective sets across an institution?
2. Do LSA metrics provide a range to discriminate semantic relatedness of outcomes/objectives sets?
3. Can the LSA cosine ranges be evaluative -- indicating strong, medium, weak or minimal levels?

Method

In this study, we use latent semantic analysis to determine the relatedness of sample sets of outcomes/objectives for programs across the university. The university uses a web-based program assessment software that allows program flexibility for construction of outcomes/objects, with related measurements, findings and actions, while providing an overall picture of all units for internal and external reporting. Forty-two program outcomes/program sets are used in this exploratory study, including an equal number of bachelor's and master's degree programs and a variety of programs in the sciences, humanities and social sciences. The method is described in Foltz (1996) and more completely summarized in readings at http://lsa.colorado.edu/.

As previously mentioned, LSA is an automatic statistical technique for inferring relations for expected contextual usage of words in passages of text. This application of LSA takes as its input only raw text parsed into words, defined as unique character strings and separated into groupings (sets of outcomes/objectives.) It uses no humanly constructed dictionaries, semantic networks, grammars, syntactic parsers, etc. Inference of semantic relatedness is from contextual usage.

The first step is to select the programs for comparison.

Using LSA software, the second step is to generate a matrix of occurrences of each word in each document. Text is represented as a matrix with each row standing for a unique word and each column stands for a text passage. Each cell contains the frequency with which the word of its row appears in the passage denoted by its column.

Next, the cell entries undergo a preliminary transformation with each cell frequency weighted by a function that expresses both the word's importance in the particular passage and the degree to which the word type carries information in the discourse domain in general.

LSA applies singular value decomposition (SVD) to the matrix. This is similar to factor analysis. In SVD, the rectangular matrix is decomposed into the product of three other matrices. One component matrix describes the original row entities as vectors of derived orthogonal factor values, another describes the original column entities in the same way, and the third is a diagonal matrix containing scaling values. SVD decomposes the word-by-document matrix into a set of *k* orthogonal factors (generally 100 to 300) from which the original matrix can be approximated by a least-squares best fit.

The result of the SVD analysis is a *k*-dimensional vector space containing a vector for each term and each document. The vector location reflects the correlations in their use across documents. The location of document vectors reflects correlations in the terms used in the documents. In this space, the cosine between vectors corresponds to estimate semantic similarity. The theoretical cosine range can be from +1.00 to -1.00, although in practicality, even two completely different English language texts will, at most, fall into the teens. Our concern is generally at the other end of the spectrum. A higher cosine (closer to 1.00) means higher shared meanings. For a visual representation of the cosine and vector relationship, see http://www.ies.co.jp/math/java/trig/cosbox/cosbox.html.

In short, determining the cosine of vectors of two pieces of textual information allows us to determine the semantic similarity between them.

Findings

We select outcomes/objective listings from the academic programs in the university assessment database to conduct several LSA explorations into outcomes/objectives relatedness.

Research question 1 asks: Can LSA provide a metric to indicate relatedness of outcomes/objective sets?

Yes. The exploration indicates the LSA metric can show relatedness of outcomes/objectives sets across a variety of university academic programs.

The cosign of each program with every other program in the sample averaged .48. The range of the average cosigns is .58 to .15. The cosine represents the angles of the vectors and corresponds to estimate of semantic similarity. A higher cosine (closer to 1.00) means higher shared meanings.

In this exploration, LSA identifies relatedness of the outcomes sets to other programs. On the upper end of the spectrum, Computer Science has an average cosign of .58 with all other programs and a range of .78 to a low of .17. The cosign is above .60 with 21 of the 41 other programs; the cosign is above .50 with 32 of the programs. Only 3 programs have a cosign of below .40. Also on the upper end of the spectrum, Urban Studies has an average cosign of .56 with all other programs and a range of .77 to a low of .15. The cosign above .60 occurs with 18 of the 41 other programs. In a similar pattern, only 2 programs have a cosign of below .40.

On the lowest end of the spectrum, African American Studies (BA) and Biology (BA) have each have average cosigns of .15 -- little relatedness to other program outcomes/objectives. African American Studies ranged from .07 to .22; Biology ranged from .07 to .21.

Overall, 25 (59.5 %) of the programs have average cosign relatedness of .50 or above; 13 (30.5 %) in the range of .40 to .49; 2 (10%) in the range of .30 to .39; and 2 (10% in the range of .10 to .19. This exploration shows the LSA metric can show relatedness of outcomes/objectives sets across a variety of university academic programs.

Research question 2 asks: Do LSA metrics provide a wide enough range to discriminate semantic relatedness of outcomes/objectives sets?

Yes. As previously mentioned, the LSA metric can provide a range to discriminate semantic relatedness of outcomes/objectives sets of individual academic programs.

The ranges not only show comparisons among individual programs, but can also indicate varying levels of categories of programs. For example, when comparing the average mean cosigns of master’s and bachelor’s programs, the master’s level is .51 and the bachelor’s is .45. 67% of the master’s programs have relatedness measures to each other above .50 compared with 33% of the bachelor’s programs. A second example compares relatedness among science, humanities and social science programs. The average mean cosign of social sciences is .53 (with 73 % of the cosigns above .50), for humanities .49 (with 67 % above .50) and sciences .46 (44 % above .50).

Research question 3 asks: Can the LSA cosine ranges be evaluative -- indicating strong, medium, weak or minimal levels?

Yes. The exploration indicates the LSA cosign ranges can be evaluative – indicating strong, medium, weak or minimal levels.

Beyond using average cosigns, the evaluative ranges of semantic relatedness can be shown by looking at the cosigns comparing individual master’s and bachelor’s programs. Those are: History (.96); Math (.92); Accounting (.85); Criminal Justice (.83); Economics (.78); Art History (.77); Computer Science (.77); Information Science (.73); Interior Design (.68); English (.37); Biology (.16).

By looking at the narrative Math outcomes/objectives it is clear the semantic relatedness to each other is great.

> MATH (MS)
> 1. Students will be able to think creatively through conjecturing, problem solving and/or computer simulations both independently and in directed research
> 2. Students will be able to analyze and write mathematical arguments and proofs with professional competence and refine symbolic calculation skills beyond the basic undergraduate level
> 3. Students will be able to read and interpret mathematical literature, including technical articles within a particular mathematical sub-field and to write or orally present mathematics with professional competence
> 4. Students will be able to use technology, including specialized computational and graphics software, to test validity of certain conjectures, for solving problems and for mathematical experimentation and research

5. Graduates will be able to pursue goals or career in education or industry that are consistent with and benefit from one's mathematical education at the Masters level

MATH (BS)

1. Students will be able to improve their creative thinking through problem solving and/or computer simulations
2. Students will develop basic mathematical skills such as symbolic calculations and the ability to write proofs that can be effective in the analysis and solution of problems
3. Students will be able to read and interpret mathematical literature and to write/present mathematics commensurate with their level of technical sophistication
4. Students will learn to use technology effectively in solving problems and for mathematical experimentation where appropriate
5. Students will be able to pursue goals or career in secondary education or in industry commensurate with the preparation that comes with a BS degree in mathematics, or to proceed to graduate study for increased career opportunities in mathematical and scientific fields in academia, government or in the industry.

Similarly, the History outcomes/objectives, written in a very different style from Mathematics, also demonstrate high semantic relatedness to each other (.96)

HIST (MA)

1. Ability to comprehend Historical Arguments and Interpretations
2. Ability to Comprehend Scholarly Popular Works
3. Development of Research and Writing Skills
4. Enhanced Historical Consciousness
5. Familiarity with Research Methodologies

HIST (BA)

1. Ability to comprehend Historical Arguments and Interpretations
2. Ability to Comprehend, Summarize, Analyze, Evaluate Scholarly Popular Works
3. Methodology, Development of Research Skills, Interpretation of Evidence
4. Development of Writing Skills and Historical Consciousness

On the other hand, Biology has low relatedness. A review of the narratives indicate problems – previously identified with the LSA metric.

BIOL (MS)

1. A combination of coursework and thesis research will ensure that students are thoroughly trained within their chosen sub-discipline of Biology.
2. Upon completion of the graduate degree, students will be competitive for employment in laboratories, consulting firms, and local/state/federal agencies. Further, students will be competitive for admission to professional programs and advanced degree programs with graduates from our peer institutions and other Virginia universities.
3. Our MS program students will complete their MS degree in a timely manner and with a high degree of academic success.

4. Students will publish their theses in peer-reviewed scientific journals.
5. Students will present their thesis research at regional and national scientific conferences.

BIOL (BS)

1. Students will have a broad foundation in Biology and the other Life Sciences.
2. Students will be able to enter professional and graduate programs.
3. Students will be able to think critically and analytically from their exposure to research and scholarship.
4. Students` learning will be facilitated through the provision of flexible scheduling.
5. Students will be able to use the urban environment as a resource rich laboratory.
6. Students will have ready access to information on scheduling and curriculum planning.

If we look at the LSA literature, Martin (2004) found threshold levels of mid-40s and below for a known “unrelated” news release compared with news articles. At the higher end, the range of LSA cosines ran from strong impact (.86), to medium impact (.72) and to weak coherence (.63). Blackmon, Kitajima and Polson (2003) set a .60 cosine minimum level indicating relatedness of navigational tags on web sites.

The exploration initially demonstrates that although two outcomes/objectives sets are written very differently, LSA can provide a reliable metric measurement of relatedness of the sets.

Conclusion

The LSA metric can take a large number of items, such as all program outcomes/objectives, and quickly identify those which may be outlyers for focus and review. 59.5 % of the programs tested have outcome/objective sets with strong semantic similarity. Another 30.5 % are somewhat similar. Overall, the LSA metric identifies 10 % as very low in semantic relatedness. When reviewing narratives for the programs with stronger and weaker semantic relatedness, outcomes/objectives sets in need of revision are reflected with the cosign numbers.

It is clear from the exploratory study that LSA has the ability to differentiate levels of relatedness based on the context of words, regardless of words in charter strings. Because of the limitations of an exploratory study, however, additional research will be needed to verify and clarify LSA usage, especially the gradation of cosine levels.

References

Blackmon, Marilyn Hughes, Kitajima, Muneo and Polson, Peter G. (2003) “Repairing Usability Problems Identified by the Cognitive Walkthrough for the Web.” CHI letters

5:1 p. 498.
http://staff.aist.go.jp/kitajima.muneo/English/PAPERS(E)/CHI2002WSBKPL.html

Deerwester, Scott, Dumais, Susan T., Furnas, George W., Landauer, Thomas K., & Harshman, Richard (1990). Indexing by latent semantic analysis. *Journal of the American Society for Information Science*, 41(6), 391-407.
http://lsi.research.telcordia.com/lsi/papers/JASIS90.pdf

Foltz, Peter W. (1996). Latent semantic analysis for text-based research. Behavior research methods, instruments and computers. 28(2), 197-202.
http://psych.nmsu.edu/~pfoltz/reprints/BRMIC96.html

Foltz, Peter W., Kintsch, Walter & Landauer, Thomas K. (1998). The measurement of textual coherence with Latent Semantic Analysis. *Discourse Processes*, *25*, 2&3, 285-307. http://www.knowledge-technologies.com/papers/dp2.foltz.pdf

Landauer, Thomas K., & Dumais, Susan T. (1997) A solution to Plato's problem: The Latent Semantic Analysis theory of the acquisition, induction, and representation of knowledge. *Psychological Review* , 104, 211-240.
http://lsa.colorado.edu/papers/plato/plato.annote.html

Landauer, Thomas K., Foltz, Peter W., & Laham, Darrell (1998). Introduction to latent semantic analysis. *Discourse Processes,* 25, 259-284.
http://lsa.colorado.edu/papers/dp1.LSAintro.pdf

Martin, Jr., Ernest F. (2004). News Release Flow-Through: News Release/News Article LSA Metric. Paper presented at Association for Education in Journalism and Mass Communication, Public Relations Division, Toronto.
http://aejmc.org/convention/2004abstracts/PR.html

Turner, Althea A. and Edith Greene (1987). The propositional analysis system (Tech. Rep. No. 87-02). Boulder: University of Colorado, Institute of Cognitive Science.
http://psych.colorado.edu/ics/techpubs/pdf/87-02.pdf

IES Math Education and Technology
http://www.ies.co.jp/math/java/trig/cosbox/cosbox.html

National Institute of Technology in Liberal Education
http://javelina.cet.middlebury.edu/lsa/out/lsa_intro.htm

SALSA Lab, University of Colorado at Boulder http://www.lsa.colorado.edu/

Telcordia http://lsi.argreenhouse.com/lsi/LSI.html