Visualizations for Knowledge Building Assessment

Chris TEPLOVS and Marlene SCARDAMALIA
Institute for Knowledge Innovation and Technology,
Ontario Institute for Studies in Education of the University of Toronto

Abstract. In this paper we use visualizations of the contributions of students to a collective knowledge space to examine semantic near neighbors and suggest how visualization tools might help users gain insight and control over the processes and means for knowledge building.

Introduction

Knowledge building is a process of sustained idea improvement fostered by communities in which participants take responsibility for the advancement of community knowledge [1]. Its focus on the creation of new knowledge and on the provision of solutions to knowledge problems [2] requires more than the application of traditional assessment tools. Whereas we have demonstrated that engaging in knowledge building can yield conceptual advances and concomitant gains in traditional literacy scores [3] our goal is to extend this analytic framework to enable embedded and transformative assessment of the knowledge building process, not simply assessing its end point.

1. Knowledge Forum

Knowledge Forum, the original Knowledge Building Environment [4], stores discourse, knowledge objects and artifacts, time-stamped information about activities and interactions, and revision histories in an extensible database [5].

Users of Knowledge Forum have had access to powerful analytic tools for nearly a decade. Early analytic tools provided access to tabular representations of data. We experimented briefly with tools to graph results, but in advance of embedding tools in a comprehensive framework for assessment, and creating visualizations that allow results to be integrated into the ongoing work of the community, their use was limited [6]. The current assessment framework supports concurrent, embedded, and transformative assessment, thus represents an important step forward. However, it remains difficult for the end user to synthesize the vast amount of information provided by these assessment tools.

The aspects of knowledge advance that we are interested in measuring have historically been difficult to measure without engaging in time-consuming manual analyses: depth of understanding and explanation, idea diversity, curricular coverage. Additionally, the goal of transformative assessment requires that we present those results to users to inform their ongoing process, not simply to researchers or instructors to analyze results. Our goal is to employ complex yet comprehensible visualizations that convey the results of such assessments to participants as young as 5 and as old as 105.

Our assessment questions include: in what contexts has a particular individual worked? What are the dominant ideas in the discourse space? What does the evolution of the discourse space over time look like? How similar are the writings of participants to those of experts? Can we identify depth of understanding and conceptual change?

2. Latent Semantic Analysis

A promising approach to finding answers to these questions lies in the use of Latent Semantic Analysis (LSA). LSA is a statistical technique used to extract the deep meaning of patterns of words in specific contexts of use. The technique is performed by applying methods from linear algebra (matrix decomposition and dimension reduction) to matrices that represent usage patterns of terms in [7,8]. It is also a theory about knowledge acquisition in human beings [9,10]. Current applications of LSA include indexing and information retrieval [11], assessment of text coherence [12,13,14], automated grading of essays [15], and summary scoring and revision [16].

3. Current Work

We have developed a visualization tool for Knowledge Forum that projects the structural and semantic relations among notes from one or more databases in a two-dimensional space.

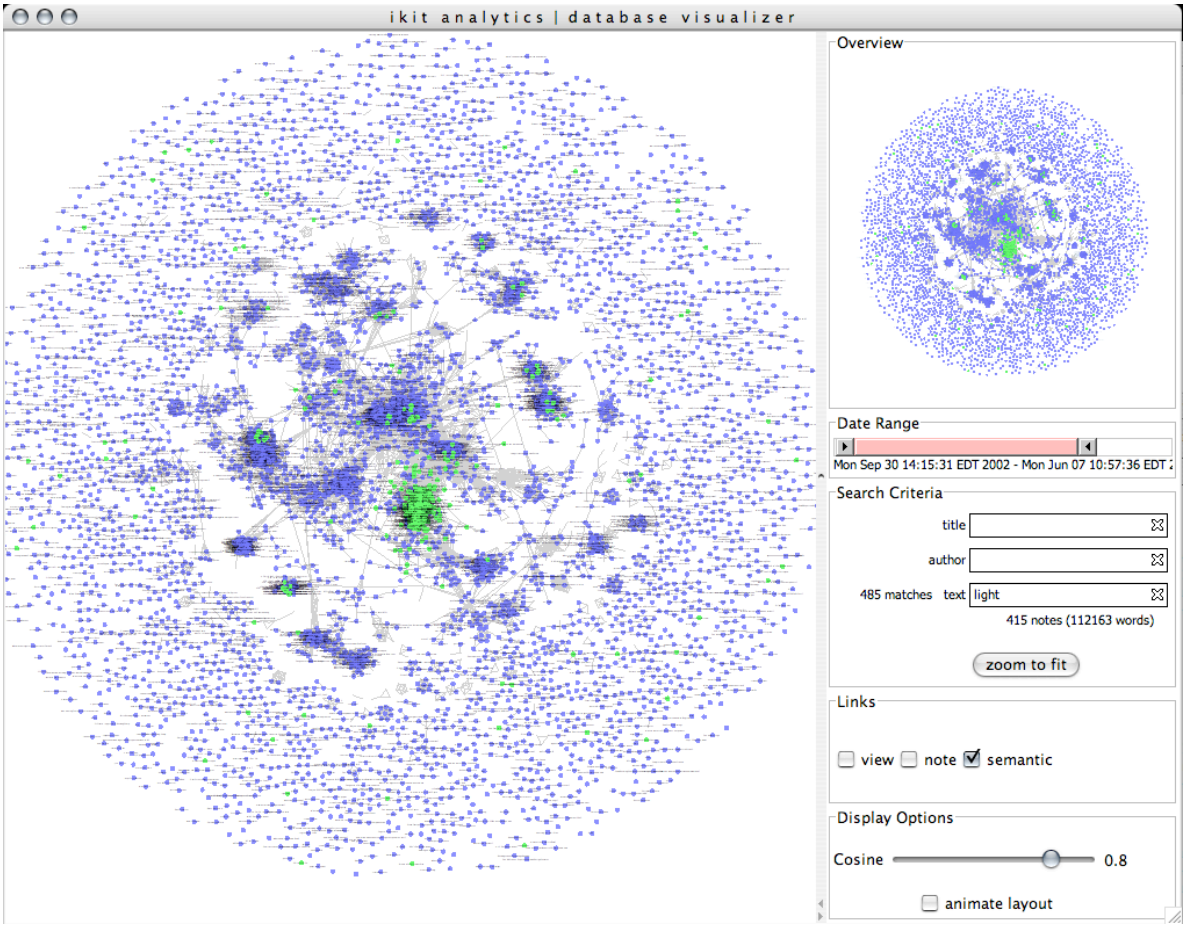

Figure 1a: Semantic layout of Year 1 + Year 2 notes. Notes containing the word "light" are highlighted.

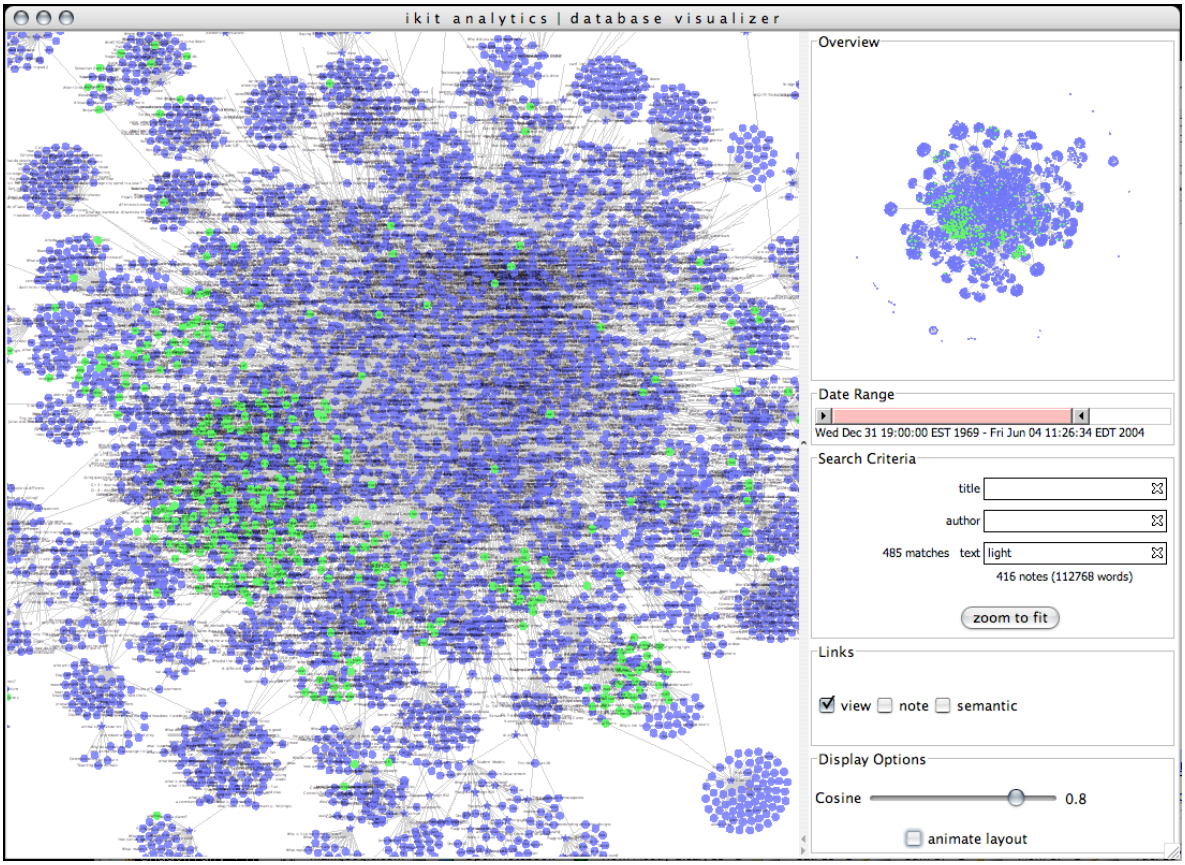

Figure 1b: Structural layout of Year 1 + Year 2 notes. Notes containing the word "light" are highlighted.

Figure 1a shows the extent to which the word "light" is found in notes in the students' database. Figure 1b shows the extent to which the same word is integrated into different sub-networks in the database. This simple "vocabulary overlay" function could be extended to a global network, so users are able to find their semantic near neighbors in a global network. Of course LSA allows for much more complex semantic analyses, using synonyms, frequently co-located terms, and other indicators of semantic overlap. Procedures can also be extended to locate student work in some benchmark corpus, with users (individually and collectively) located in the texts of a target group, For example, in one case a researchers compared early and late database entries, showing their respective overlap with a scholarly article. In another, student notes were mapped onto curriculum guidelines.

The visualization tool has several key features: an overview, the ability to zoom and filter, and details on demand [17]. The small overview panel in the upper right is automatically scaled to show all the notes in the analysis. The large pane on the left can be zoomed in or out and panned to allow the user to see relationships at a finer level of granularity. Notes can be highlighted (shown in green) by matching search criteria including authors' names, words in the object's title, or full-text searching. A slider is used to constrain the display of objects according to modification or creation dates. The lower and upper ranges of the date slider can be moved independently, thereby providing the user with an animation of the temporal growth of the database, or in unison, thereby providing temporal "slices" of database activity. The details of the notes, including authorship, revision dates, and full content, are shown as an overlay when the user hovers over an object. Selecting an object by clicking on it will highlight both it and its immediate neighbours.

Notes (nodes) are represented as circles and are labelled with the titles of the notes they represent. Links (edges) are derived from structure (i.e. the assignment of a note to one or more views, or by virtue of relationships such as build-on, annotation, or reference to another note) or semantics (derived from LSA). In the case of semantic links, the visibility of these links can be controlled by a slider that determines the threshold at which any two notes are considered sufficiently similar to merit a link.

Of note is the provision of a simple metric (word count) derived from the objects matching the search criteria. This is an important feature: it marks the shift from assessment as a stand-alone activity to assessment as an integrated feature of the user interface. The visualization integrates textual, dialogic, and graphical literacy, with potential to enhance all.

References

[1] M. Scardamalia, Technology for understanding, pp. 83--88 in: K. Leithwood, P. McAdie, N. Bascia, and A. Rodrigue (Eds.) *Teaching for deep understanding: towards the Ontario curriculum that we need*, Toronto, 2004.

[2] C. Bereiter, Understanding technology, pp 48--5188 in: K. Leithwood, P. McAdie, N. Bascia, and A. Rodrigue (Eds.) *Teaching for deep understanding: towards the Ontario curriculum that we need*, Toronto, 2004.

[3] Y. Sun, J. Zhang, and M. Scardamalia (in press). Developing deep understanding and literacy while addressing a gender-based literacy gap. *Canadian Journal of Learning and Technology*.

[4] M. Scardamalia, and C. Bereiter, Knowledge building environments: Extending the limits of the possible in education and knowledge work. In A. DiStefano, K.E. Rudestam, and R. Silverman (Eds.), *Encyclopedia of distributed learning*, Thousand Oaks, CA: Sage Publications, 2003.

[5] ZooLib. Available from http://zoolib.sourceforge.net/. Accessed 11 May, 2007.

[6] C. Teplovs, Z. Donoahue, M. Scardamalia, and D. Philip, *Tools for Concurrent, Embedded, and Transformative Assessment of Knowledge Building Processes and Progress*, CSCL2007.

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

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

[9] T. K. Landauer and S. T. Dumais (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.

[10] W. Kintsch (1998). *Comprehension: a paradigm for cognition*. New York: Cambridge University Press.

[11] S. Dumais. LSA and Information Retrieval. pp. 293-322 in: T. K. Landauer, D. S. McNamara, S. Dennis and W. Kintsch (Eds.) *Handbook of Latent Semantic Analysis*. Mahwah, NJ. (2007).

[12] P.W. Foltz (1996). Latent Semantic Analysis for text-based research. *Behavior Research Methods, Instruments and Computers* 28(2): 197-202.

[13] P.W. Foltz, W. Kintsch, and T.K. Landauer, (1998). Analysis of text coherence using Latent Semantic Analysis. *Discourse Processes* 25.

[14] P.W. Foltz. Discourse Coherence and LSA. pp. 167-184 in: T. K. Landauer, D. S. McNamara, S. Dennis and W. Kintsch (Eds.) *Handbook of Latent Semantic Analysis*. Mahwah, NJ. 2007.

[15] Landauer, T.K., Laham, D., & Foltz, P.W. (2003). Automatic essay assessment. Assessment in Education: Principles, Policy & Practice, 10(3), 295-308.

[16] E. Kintsch, D. Caccamise, M. Franzke, N. Johnson, and S. Dooley. Summary Street: Computer-Guided Summary Writing. pp. 263-278 in: T. K. Landauer, D. S. McNamara, S. Dennis and W. Kintsch (Eds.) *Handbook of Latent Semantic Analysis*. Mahwah, NJ. (2007).

[17] B. Shneiderman. The eyes have it: a task by data type taxonomy for information visualization. *Proceedings of IEEE Workshop on Visual Languages '96* (1996), 336--343.