

## Visualization and Elaboration of Students' Group Reading Processes

Sarah A. Sullivan, Sadhana Puntambekar,  
Department of Educational Psychology, University of Wisconsin, Madison, WI  
Email: sasullivan2@wisc.edu, puntambekar@education.wisc.edu

**Abstract:** This study discusses the contribution of data mining of log file data to video and interview data for studying students' collaborative engagement with digital texts presented online. Given that reading is generally an individual activity, understanding this group process is important. Data analyses at multiple levels can reveal group processes of knowledge construction, but can also give conflicting results, particularly related to students' perceptions of the process.

As part of learning through inquiry, students often work in collaborative groups to use multiple digital text resources to follow their own investigation paths while engaging with information. Data mining techniques, such as log file analyses, help to identify diverse navigational styles used to process and integrate multiple sources of information from digital text environments by allowing examination of navigation paths taken by different groups of learners (e.g., Puntambekar, Stylianou, & Hübscher, 2003). But, given that reading is generally an individual activity, understanding group processes of interacting with and making sense of multiple texts during group collaboration is important. Data mining techniques can be used to uncover patterns of navigation through digital systems that are not directly observable (Reimann, Yacef, & Kay, 2011). However, navigation profiles provide only one perspective on groups' processes of information selection and integration. Multiple scales of analysis are important to both understand phenomena at various levels and understand the relationships across scales (Suthers & Medina, 2011). This study discusses the contribution of data mining and visualization of log file data in the form of navigation paths to understanding group processes revealed by other measures, such as group video and interview data. The question addressed is: What processes do students use to construct knowledge from multiple digital texts and how can data mining techniques combined with other data sources increase understanding of these processes?

### Procedure

The log file analysis focused on 19 students in a 6<sup>th</sup> grade science classroom in which students read multiple digital texts using the CoMPASS system (Puntambekar, 2006) as part of a project-based inquiry physics curriculum. CoMPASS helps students engage in inquiry about the physics related to multiple topics, such as simple machines, and provides navigable concept maps designed to mirror the conceptual structure in the domain of physics and help students gain a rich understanding of concepts and relationships. CoMPASS provides two representations: a navigable concept map and textual descriptions. The concept that is being read about becomes the focal point of the map and the other concepts move accordingly based on the strength of their relationship to the concept of focus.

Navigation log files were used to look at the students' navigation behavior while using CoMPASS for approximately 15 minutes to learn about physics concepts to help them with the challenge of lifting a water bottle off of a table with the least amount of force. The topics and concepts that students visited and the time spent on each were recorded in chronological order. The log files were later used to develop navigation profiles to help understand students' meaning making processes while using CoMPASS by making students' navigation choices visible. In another iteration of the physics curriculum, video was collected of students' group interactions with CoMPASS in order to investigate their processes of interacting with the texts and making navigation decisions. Further, five student small groups (N=4 in each group) were interviewed at the end of the curriculum to examine students' reasoning behind their group's interactions with the texts and whether their perceptions of their processes for selecting concepts align with the analyses of strategies from the video data and the previous navigation profiles developed from the log file data.

### Analysis and Findings

The log files of students' interactions with CoMPASS were analyzed using Pathfinder (Schvaneveldt, 1990). Pathfinder is a graph theoretic technique used to create network representations consisting of nodes and links that characterize navigation patterns in CoMPASS, allowing us to look at similarities and differences in navigation paths. Navigation patterns (see Figure 1) show the number of times students "navigated to" a concept and the number of times students "navigated from" a concept and from which concepts students navigated to others.

The log file data reveal distinct ways in which students structured their paths through information presented in the CoPASS system. For example, two clusters that were found highlight the differences between students who focused on concepts within a single topic versus those who chose to explore concepts in multiple topics. Cluster 1 (N=15) reveals that students investigated how all of the concepts within pulley may apply to the challenge and also looked at the topics of other simple machines. However, students did not look at any concepts within these other machines. Cluster 2 (N=4) reveals that students explored all concepts within pulley with regard to their challenge as well as many of those same concepts and how they apply in a wedge.

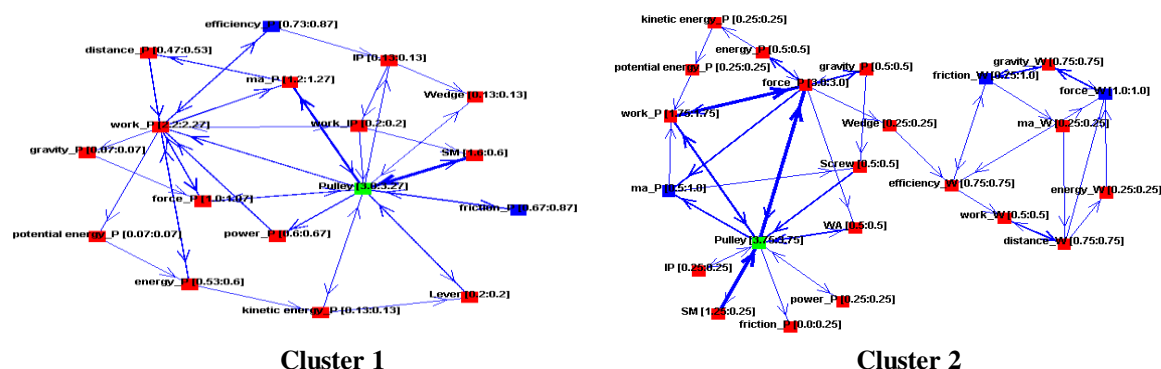


Figure 1. Navigation patterns of the two clusters. Each node has the name of the concept and the topic, so work\_P means students visited the concept of work in the topic pulleys.

From the log data, it appears that students in Cluster 1 were relatively goal-focused in their navigation. Log file data for Cluster 2 revealed search processes that appear to be less focused. However, the analysis of the video data revealed that log file data only captures the navigation preferences of a subset of students when working in collaborative groups. Students struggled at times to take up and consider all navigation suggestions, and some students tended to be deferred to more than others. Despite these collaboration issues revealed by the video data, in the interview data, students often perceived their group processes as conducive to conducting research related to their goal. Students also reported that they were able to use the concept maps to focus on content, but the navigation profiles indicate that students' navigation behaviors are not always goal-focused.

The multiple forms of data used for this study provide different perspectives on students' processes of interacting with multiple texts. The analyses also give insight into how the use of multiple data sources can shed light on the importance of different levels and scales of analysis and interpretation when studying engagement with multiple digital texts. Of particular concern is that, despite the issues revealed by the navigation profiles and video data, students perceived their processes for engaging with multiple texts as a collaborative group as productive and goal-focused overall. Kulikowich & Young (2001) have argued for making the collaboration process visible to learners in order to enhance their performance. This study underscores the need for investigating collaboration at multiple levels to not only inform understanding of the collaborative process for researchers and teachers but also to collect information for learners to help them understand their processes of navigation and collaboration. Perhaps, allowing collaborative reflection on navigation profiles may help students to make productive navigation choices.

## References

- Kanungo, T., Mount, D. M., Netanyahu, N., Piatko, C., Silverman, R. & Wu, A. (2000). *The analysis of a simple k-means clustering algorithm*. Symposium on Computational Geometry.
- Kulikowich, J. M., & Young, M. F. (2001). Locating an Ecological Psychology Methodology for Situated Action. *The Journal of the Learning Sciences*, 10(1&2), 165-202.
- Puntambekar, S. (2006). Learning from digital text in inquiry-based science classes: lessons learned in one program. *Proceedings of the 7<sup>th</sup> International Conference of the Learning Sciences* (pp. 564-570).
- Puntambekar, S., Stylianou, A., & Hübscher, R. (2003). Improving navigation and learning in hypertext environments with navigable concept maps. *Human Computer Interaction*, 18(4), 395-426.
- Reimann, P., Yacef, K., & Kay, J. (2011). Analyzing collaborative interactions with data mining methods for the benefit of learning. In S. Puntambekar, G. Erkens, & C. Hmelo-Silver (Eds.), *Analyzing Interactions in CSSL: Methods, Approaches & Issues*, New York: Springer.
- Schvaneveldt, R. W. (1990). *Pathfinder associative networks*. Norwood, NJ: Ablex.
- Suthers, D., & Medina, R. (2011). Tracing interaction in distributed collaborative learning. In S. Puntambekar, G. Erkens, & C. Hmelo-Silver (Eds.), *Analyzing Interactions in CSSL: Methods, Approaches & Issues*, New York: Springer.