

Social Network Analysis of Collective Knowledge Advancement

Jun OSHIMA^{a*b} & Ritsuko OSHIMA^a

^a*RECLS, Shizuoka University, Japan*

^b*CoREF, University of Tokyo, Japan*

**joshima@inf.shizuoka.ac.jp*

Abstract: This study aims at applying social network analysis (SNA) to establishing indicators for collective knowledge advancement. To do so, a SNA application, Knowledge Building Discourse Explorer (KBDeX), was developed and the effectiveness of several indicators was examined by applying them to an evaluation of students' discourse identified in another study as somewhat advanced in their state of collective knowledge. Results indicated that SNA with our proposed indicators quantitatively captured the collective knowledge advancement. We discussed the mixed-method approach with descriptive studies of cases and quantitative SNA.

Keywords: Social Network Analysis (SNA), KBDeX, Collective Knowledge Advancement

1. Introduction

With the development of new theories of learning that unify the acquisition and participation metaphors (Sfard, 1998), it has become clear that current assessment techniques are not adequate for examining new theories of learning in the knowledge-creation metaphor (Paavola, Lipponen, & Hakkarainen, 2004). Here, attention is focused on the example of Bereiter's model of a knowledge-building community (Bereiter, 2002) in order to discuss the techniques needed for such an examination. Scardamalia and Bereiter explain the nature of knowledge building by referring to two modes of learning: the belief mode and the design mode (Scardamalia & Bereiter, 2005). In the belief mode, learners are concerned with what they or others believe or ought to believe, namely, with the mental states of individuals. On the contrary, in the design mode, learners are concerned with the usefulness, adequacy, improvability, and developmental potential of ideas. Learners in the design mode should be aware of whether their ideas are good enough to solve the problems to be addressed, and how they should contribute to improving those ideas. Knowledge building is a social process that engages both modes of learning. The belief mode is used by learners to investigate the current state of their community knowledge level in order to highlight any problems. Learning in the design mode thus enables the creation of knowledge to solve problems. Exchange between learning modes is iterative, such that learners continuously participate in social practices of knowledge creation, and individuals generate knowledge that not only directly contributes to the advancement of community knowledge but also determines how best to contribute to this advancement.

How do we assess learning in knowledge building? The detailed description analysis of discourse can reveal how a group of learners engage in their joint activities to construct shared knowledge in their groups as communities. However, as suggested by Mercer and colleagues (Mercer, 2005), detailed description analysis of discourse cannot be used for large datasets because it is highly time-consuming. Consequently, Wegerif and Mercer developed a methodology to combine detailed description analysis of discourse and computerized discourse analysis to handle large datasets (Wegerif & Mercer, 1997). Although their approach is useful in qualitative and quantitative analyses, the focus of the methodology is still on the participation metaphor, rather than a combination of the acquisition and participation metaphors or the knowledge-creation metaphor.

In this paper, SNA of discourse is proposed as an assessment approach in the knowledge-creation metaphor. In the next section, a brief review is given of literature that investigates knowledge creation and on educational studies applying SNA to collaborative learning.

2. SNA of Collaborative Learning Discourse

In educational research on networked learning and computer-supported collaborative learning, there have been discussions on the advantages of using SNA to investigate community knowledge advancement and individual learners' engagement in this advancement (de Laat, Lally, Lipponen, & Simons, 2007; Martinez, Dimitriadis, Rubia, Gomez, & de la Fuente, 2003; Reffay, Teplovs, & Blondel, 2011). A limited number of studies have used SNA, especially in the knowledge-creation metaphor. Over a period of three years, Zhang, Scardamalia, Reeve, and Messina (2009) implemented a complementary approach that used SNA to visualize and compare classroom collaboration among fourth grade elementary school students through a CSCL environment designed to support them in knowledge building. An analysis of the students' online participatory patterns and knowledge advancement indicated that this learning process facilitated students' knowledge advancement effectively, and that this was the case through critical changes in organizations within the classroom: from fixed small groups in the first year of the study to appropriate collaboration through dynamic formation of small teams based on emergent goals.

Oshima, Oshima, and Knowledge Forum Japan Research Group (2007) further extended the potential of SNA as a core assessment technique by describing a different type of social network. Ordinary SNA illustrates the social patterns of learners, namely, the learners' social network. As de Laat et al. suggested, this approach is thus informative when examining developments or changes in the participatory structure of a community (de Laat et al., 2007). However, Oshima et al. (2007) argued that existing social network models are unable to examine how community knowledge advances through learners' collaboration. Instead, they used a similar procedure to that in ordinary SNA, but proposed a different type of social network, one based on the words learners use in their discourse on a CSCL environment. This social network, with words selected as nodes representing learners' knowledge or ideas during discourse on a study topic, was compared with a network of words from the discourse of a group of experts on the same topic. The results showed that there were remarkable differences in the community knowledge of elementary school students and of experts in terms of the words centered on the networks. Oshima et al. (2007) concluded that SNA could provide a new type of representation of community knowledge building by learners, enabling researchers to adopt a new complementary assessment technique for investigating knowledge building community models.

Although educational studies have proposed the application of SNA to learning analysis as a new assessment technique in the knowledge-creation metaphor, an exact methodology has to be established. Currently, researchers familiar with SNA conduct analysis with whatever software is available to them. For those interested in using SNA to analyze discourse data for examining participation patterns and collective knowledge advancement during collaborative learning, SNA software that can easily explore discourse data is needed. As a consequence, we have developed the Knowledge Building Discourse Explorer (KBDeX) software. In this paper, we introduce KBDeX and explain how three different types of discourse-based social networks are represented. Moreover, SNA conducted with KBDeX is demonstrated by using written discourse data in a CSCL environment. Finally, the potential contribution of KBDeX to a new complementary assessment approach in the knowledge-creation metaphor is discussed.

3. Knowledge Building Discourse Explorer: An Application for SNA of Discourse in Collaborative Learning

To establish a methodology for discourse analysis in collaborative learning from the perspective of complex network science, we are currently developing a platform application called Knowledge Building Discourse Explorer (KBDeX) (Matsuzawa, Oshima, Oshima, & Sakai, 2012; Oshima, Oshima, & Matsuzawa, 2012) is an analysis platform for visualizing network structures of discourse based on a bipartite graph of words $\times$ discourse units (e.g., conversation turns, BBS postings, or sentences) (Figure 1). Using discourse data (in .csv format) and a list of target words for bipartite graph creation (a text file) as its input, KBDeX can create visualizations of three different network structures: (1) learners' network structure (top right window in Figure 1), (2) unit network structure (bottom left), and (3) the network structure of the target words (bottom right).

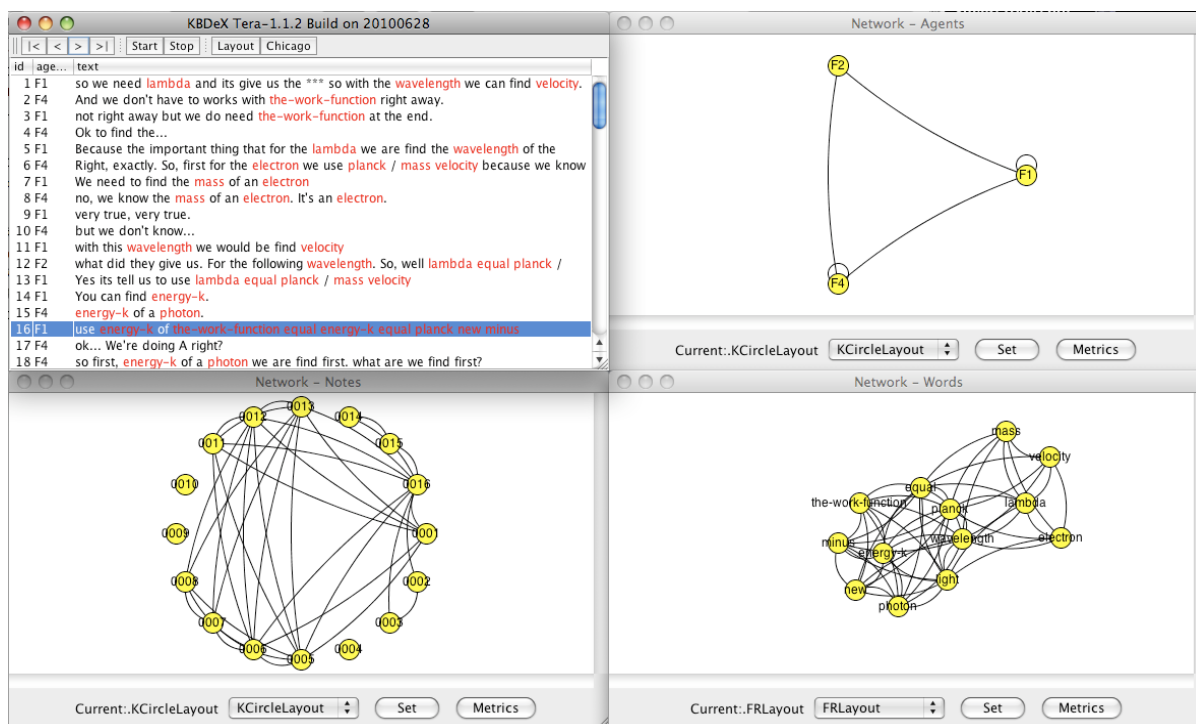


Figure 1. The Interface of KBDeX.

KBDeX automatically calculates typical measures for network structures used in network science, (1) the *betweenness centrality coefficient*, (2) the *degree centrality coefficient*, and (3) the *closeness centrality coefficient*. *Betweenness centrality* is a measure of the number of node pairs for which the shortest path between them passes through a selected node. High betweenness centrality suggests that the selected node works as a key mediator in linking other nodes. *Degree centrality* is a straightforward concept that indicates cumulative path lengths by which each node is linked to other nodes in the network. High degree centrality means that the node is at the center of the network as a whole, or near the center of a local cluster in the network. *Closeness centrality* is a more sophisticated measure of how close the node is to other nodes in a network, based on the geodesic distance.

4. An Exploratory Study to Establish Indicators for Collective Knowledge Advancement

For examining our SNA indicators for collective knowledge advancement, we designed a blind-test study. By using the same datasets as those that other researchers had already analyzed to draw their conclusion on the collective knowledge advancement in two different conditions of classroom discourse on a CSCL (Chan & Lam, 2010), we tested our SNA-based approach to evaluating the collective knowledge advancement and compared our conclusion with theirs.

Our SNA of students' discourse on CSCL in two different classrooms was conducted as follows. First, we proposed several indicators to capture collective knowledge advancement, creating indicators for *continuous improvement of ideas* and *idea diversity* based on the theoretical framework of knowledge building (Scardamalia, 2002). Second, we compared the two classroom discourses with the two kinds of indicators for detecting differences in collective knowledge advancement between the classrooms. Finally, we examined the differences by integrating our results with findings from the original study (Chan & Lam, 2010) that analyzed the same data with a different methodological approach.

4.1 Classroom Descriptions

Eighty 10th graders in two chemistry classes in Hong Kong participated in the original study. Lessons were conducted in English, and students took notes in English in the CSCL environment. Both classes engaged in a knowledge-building inquiry with reflective assessment. The first class is called

Reflective Assessment (RA, $n = 40$) and the second Reflective Assessment with Scaffolds (RAS, $n = 40$). Both classes were taught by the same teacher, who had taught high school chemistry for more than 12 years and had used knowledge-building pedagogy for over 6 years.

This classroom study was conducted in the second semester of 2008-09, lasting from Feb–June (16–18 weeks). There were five chemistry lessons each week, and each lesson was 35 minutes long. In both classes, students studied electrochemistry using knowledge-building inquiry approaches. Students had similar inquiry experiences and both used the same instructional topics, textbooks, and reference materials, and conducted the same experiments. As a key difference, however, students in the RAS class wrote reflective summaries in the CSCL environment using a set of predesigned conceptual-change scaffolds, typical topics being “My initial ideas,” “What we think,” and “What I think now.” Students in the RA class wrote reflective summaries without the scaffolds. Chan and Lam (2010) evaluated students’ discourse for clarifying whether conceptual change happened in their explanations, and showed that students in the RAS class were significantly better in their conceptual understanding. The goal of SNA in this study was to detect differences in collective knowledge advancement between the two instructional conditions, in particular which class engaged in more collective knowledge advancement through discourse.

4.2 Data Analysis with Indicators for Collective Knowledge Advancement

Referring to the socio-cognitive determinants of knowledge building communities (Scardamalia, 2002), two aspects were selected for SNA, *continuous improvement of ideas* and *idea diversity*, and were measured with using our hypothetical indicators for collective knowledge advancement. Based on the premise that learners’ ideas are represented as clusters of nodes, specifically sets of words as nodes with links among them, we established an indicator for the *improvement of ideas* as follows. We first generated a benchmark network based on appropriate issue discourse by asking the teacher to provide us with a list of study questions that he wanted his students to pursue. We next created answers to the study questions, and used those answers as input data (a bipartite graph of 65 sentences $\times$ 91 words) for the benchmark network structure. We also created networks of student discourse using the same set of detected words (2309 sentences $\times$ 91 words in the RA class, and 2631 sentences $\times$ 91 words in the RAS class). We used correlations of *betweenness centrality* and *degree centrality coefficients* of words between the benchmark network and the student network as indicators for students’ *continuous improvement of ideas*. To examine time-serial change in the state of collective knowledge, students’ word networks were constructed at three different times when students reflected on their progress. Figure 2 shows the transition of correlations between the two classroom networks and the benchmark network. Fisher’s z-transformation tests indicated significant differences between phases 1 and 2 in the RA class, and between phases 1 and 3 in the RAS class. The results suggested that *the improvement of ideas* in the RAS class was more stable and oriented towards the appropriate benchmark network. We found no significant differences in analysis of the degree centrality coefficient.

Idea diversity was explored from two different perspectives, *temporal* and *individual diversity*. *Temporal diversity* was examined by analyzing the extent of differences in network structures across different phases. In particular, we compared centrality coefficients of nodes appearing in the previous and current phases. If students were involved in more diverse ideas in the current stage than the previous stage, centralities of nodes appearing in the current stage would be equal to or higher than those of nodes appearing in the previous stage. Multiple *t*-tests on the mean *degree and betweenness centrality coefficients* showed that: (1) in phase 2 the *betweenness centrality coefficients* of nodes in the previous phase were significantly higher than those in the current phase, $t(71) = 3.05676$, $p < .05$ for the RA class, and $t(73) = 2.76447$, $p < .05$ for the RAS class, (2) in phase 3 there were no significant differences in the *betweenness centrality coefficients*, and (3) in phases 2 and 3 the *degree centrality coefficients* of nodes in the previous phase were significantly higher than those in the current phase, $t(71) = 9.4707$, $p < .05$ for the RA class in phase 2, $t(73) = 6.2865$, $p < .05$ for the RAS in phase 2, $t(74) = 2.3767$, $p < .05$ for the RA class in phase 3, and $t(77) = 2.8379$, $p < .05$ for the RAS in phase 3. Based on results that new nodes appeared in the current phase had less centrality coefficients than did existing ones, we concluded that students in the both classrooms were more likely to extend their ideas rather than radically changed ideas across phases.

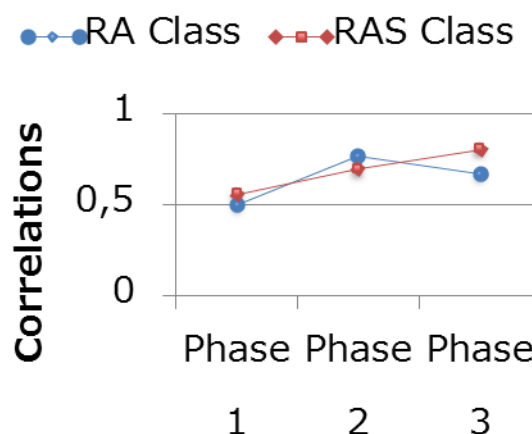


Figure 2. Transition of Correlation Coefficients of Betweenness Centrality between Students' Word Network and the Benchmark Network.

Individual diversity was examined by conducting stepwise analysis (Oshima et al., 2012), where each individual student's contribution to the network structure was evaluated by calculating the displacement of node coefficients after isolating target individuals' discourse from the total discourse. We examined correlation coefficient tables across individual contributions within each phase and found that in phase 3 there were fewer significant correlations in the RAS class than in the RA class. Based on results, we concluded that the RAS class was more diverse in individual ideas than was the RA class in the last moment of their learning.

5. Discussion

Our SNA of students' online discourse in classroom environments with and without scaffolds for conceptual change is summarized as follows. First, whereas there were significant correlations between students' networks and the benchmark network in the both classrooms, students in the RAS class engaged in more *stable and continuous improvement of ideas* toward the intended state of knowledge. Second, students' ideas in both the RA and RAS classes did not change radically across phases. They gradually constructed ideas based on what they had learned in previous phases. Finally, however, students' individual contributions in the RAS class were more diverse than those in the RA class, particularly in phase 3, indicating that students' contributions in the RAS class were more distributed in collective knowledge advancement than those in the RA class.

Combined with results of the original study analyzing the same discourse data (Chan & Lam, 2010), we conclude that the implementation of scaffolds facilitated students' conceptual change through their more productive engagement in collective knowledge advancement. Our SNA results complementarily support the original study by providing a process-oriented analysis of collective knowledge advancement. In particular, the difference in idea improvement between the two classrooms is a new finding. Students with conceptual change scaffolds were continuously engaged in improving ideas in comparison with those without the scaffolds. Why students without scaffolds discontinued their idea improvement in phase 2 and digressed from the expected state of knowledge in phase 3 is an interesting research question for us to further examine. One possibility to explain such a difference in idea improvement between the classrooms might be another finding in our SNA that students with conceptual change scaffolds were more diverse in their contributions to collective knowledge advancement. Scaffolding prompts for students to be *metacognitively* aware of their ideas from previous to current phase might facilitate them to elicit more variety of ideas through integrating different ideas out there. This is also an interesting question to examine in the future studies.

As a new methodological approach, SNA could provide the learning sciences researchers with new representations of knowledge at the collective level during collaborative learning. Collective level process-oriented analysis in collaborative learning supported by a new tool like KBDeX would strengthen our assessment of collaborative learning and understanding of how different types of

instructional scaffolds influence collective knowledge advancement and learners' individual accomplishments.

Acknowledgements

We are deeply grateful to Prof. Carol Chan and Prof. Jan van Aalst for providing their datasets for our exploratory study and comments on results. This study was supported by the Ministry of Education, Culture, Sports, Science, and Technology through Grants-in-Aid for Scientific Research (A) (No. 24240105; granted to Jun Oshima, Shizuoka University).

References

- Bereiter, C. (2002). *Education and mind in the knowledge age*. Hillsdale, New Jersey: Lawrence Erlbaum.
- Chan, C. K. K. and Lam, I. C. K. (2010). Conceptual Change and Epistemic Growth through Reflective Assessment in Computer-Supported Knowledge Building. In *Proceedings of ICLS2010* (Chicago, IL, June 29–July 02, 2010). 1071–1078.
- de Laat, M., Lally, V., Lipponen, L., and Simons, R. -J. (2007). Investigating patterns of interaction in networked learning and computer-supported collaborative learning: A role for social network analysis. *International Journal of Computer-Supported Collaborative Learning*, 2(1), 87–103.
- Martinez, A., Dimitriadis, Y., Rubia, B., Gomez, E., and de la Fuente, P. (2003). Combining qualitative evaluation and social network analysis for the study of classroom social interactions. *Computers & Education*, 41(4), 353–368.
- Matsuzawa, Y., Oshima, J., Oshima, R., & Sakai, S. (2012). Learners' Use of SNA-based Discourse Analysis as a Self-Assessment Tool for Collaboration. *Int. J. of Organisational Design and Engineering*, 2(4), 362–379.
- Mercer, N. (2005). Sociocultural discourse analysis: analysing classroom talk as a social mode of thinking. *Journal of Applied Linguistics*, 1(2), 137–168.
- Oshima, J., Oshima, R., and Knowledge Forum Japan Research Group. (2007). Complex network theory approach to the assessment on collective knowledge advancement through scientific discourse in CSCL. In *Proceedings of CSCL2007* (New Brunswick, NY, July 16–21, 2007). Lawrence Erlbaum, Mahwah, New Jersey, 563–565.
- Oshima, J., Oshima, R., & Matsuzawa, Y. (2012). Knowledge Building Discourse Explorer: A social network analysis application for knowledge building discourse. *Educational Technology Research & Development*, 60, 903–921. DOI: 10.1007/s11423-012-9265-2.
- Paavola, S., Lipponen, L., and Hakkarainen, K. (2004). Models of innovative knowledge communities and three metaphors of learning. *Review of Educational Research*, 74(4), 557–576.
- Reffay, C., Teplov, C., and Blondel, F. -M. (2011). Productive re-use of CSCL data and analytic tools to provide a new perspective on group cohesion. In *Proceedings of CSCL 2011* (Hong Kong, China, July 04 - 08, 2011). Lawrence Erlbaum, Mahwah, New Jersey, 846–850.
- Scardamalia, M. (2002). Collective cognitive responsibility for the advancement of knowledge. In B. Jones (Ed.), *Liberal education in a knowledge society* (pp. 67–98). Open Court, Chicago, IL.
- Scardamalia, M. and Bereiter, C. (2005). Does education for the knowledge age need a new science? *European Journal of School Psychology*, 3(1), 21–40.
- Sfard, A. (1998). On two metaphors for learning and the dangers of choosing just one. *Educational Researcher*, 27(2), 4–13.
- Wegerif, R. and Mercer, N. (1997). Using computer-based text analysis to integrate qualitative and quantitative methods in research on collaborative learning. *Language and Education*, 11, 271–286.
- Zhang, J., Scardamalia, M., Reeve, R., and Messina, R. (2009). Designs for collective cognitive responsibility in knowledge-building communities. *The Journal of the Learning Sciences*, 18, 7–44.