Unraveling Idea Development in Discourse Trajectories

Abstract: With the present paper we want to shed light onto an issue that is central within the knowledge building theory but only little studied – the development of ideas in collaborative learning discourse. Starting from the construction of a network of explicit and implicit relations between ideas, we apply a scientometric method to tackle the temporality of collaborative processes based on the structure of successive ideas. The resulting discourse trajectories are shown to give a holistic and also a detailed view on how knowledge advances when their interpretation is combined with a qualitative analysis of the content of the ideas and their relations. The weighted relevance of relations between ideas enables the identification of sub-topics in the discourse, important ideas, and influence or uptake events.

Introduction

How is knowledge about the world created and advanced? Philosophers have spent enormous efforts to answer this question (e.g. Popper, 1972). Knowledge building is an approach from the learning sciences that attempts to build on contemporary answers from philosophical inquiries and research on expertise (Bereiter, 2002; Bereiter & Scardamalia, 1993; Scardamalia & Bereiter, 2006) to engage students in the kinds of knowledge work that are widely assumed important in the 21st century, including ability to collaborate, deal with novelty, and solve ill-structured problems. At the heart of knowledge building is a computer-mediated collaborative discourse that is oriented toward idea improvement. Following Popper's (1972) theory of objective knowledge, knowledge-building theory considers ideas as "real" objects that can be critiqued, tested and modified, much like how real objects like bicycles undergo these processes (Bereiter & Scardamalia, 2003). Ideas do not reside in the minds of participants but take on lives on their own in this discourse. Hong, Chen, Chang, Liao and Chan (2009) emphasized that the "idea-centered" educational design is enabled through the Knowledge Forum software by allowing interaction around ideas regardless of the discussion threading. Hence, the *process* of idea development is fundamental for understanding knowledge building. Moreover, focus on processes is widespread in the field of computer-supported collaborative learning (CSCL), reflecting the dynamic nature of discourse as an object of study.

However, despite this acknowledged role, there is a dearth of analytical approaches for investigating the dynamic development of ideas in knowledge-building discourse. Therefore, the main goal of this paper is to provide an example of studying what we call a *discourse trajectory*, i.e. the genuine process characteristics of a discourse based on idea development over time. In order to do so, we first outline briefly previous research and then present a new methodological technique and its application to knowledge building discourse.

Related Research

A starting point to studying a discourse process is the evaluation of surface indicators of participation and communication like number, length of contributions, etc. (Strijbos, Kirschner and Martens, 2004). Such results can automatically be evaluated from the log file data of the software but they are regarded as only very basic descriptors of a collaborative process.

Most studies of knowledge building have followed a content analysis approach (Chi, 1997; Gunawardena, Lowe, & Anderson, 1997; Henri, 1992), where qualitative data is segmented into idea units and these are coded for their cognitive, metacognitive, social, motivational and other aspects. The frequency of the assigned codes is then statistically regarded for comparisons of different students, discussion groups or phases (e.g. Hakkarainen, 2003; Lipponen, 2000; Oshima, Scardamalia, & Bereiter, 1996). The content analysis approach is fruitful for determining the epistemic level of discourse like complexity and elaboration of ideas, as well as its progress in time. Students' contributions can also be categorized as different discourse activities, on the basis of which a sequential analysis (Jeong, 2005) of their temporal ordering can reveal patterns and facilitate deeper insight into the collaborative discourse as a process. A review of the many different coding schemes in the CSCL research field is provided by de Wever, Schellens, Valcke and van Keer (2005).

However, as is becoming increasingly clear in CSCL research, the coding and counting technique neglects some important discourse qualities as it takes statements out of their context and generally addresses actions of individuals instead of the group as a whole (Strijbos & Stahl, 2007). Details on the interactional dynamics of the discourse are ignored, too (Stahl, 2002). However, these are important for CSCL researchers studying how knowledge is advanced collaboratively.

Students' contributions should not be regarded as isolates but in relation to the discourse, because content, context, preceding and subsequent contributions are all interdependent (Barnes & Todd, 1977). This view implies a perspective shift from the individual to the intersubjective level of analysis. Stahl, Koschmann and Suthers (2006) defined collaborative learning as an interactive process of shared meaning-making in a group. According to that they pleaded for exploratory and interpretative in-depth studies of it that account for the complexity of the process. Such is the conversation analysis of case study narratives (Stahl, 2002), where the activities of all contributors and the overall context are elaborated. The ephemeral nature of an intersubjective discourse process was corroborated by Reimann, Frerejean and Thompson (2009) who found that groups do not follow deterministic process models but actively shape the development of their discourse.

In-depth discourse analyses also capture a limited part of the process. Zemel, Xhafa and Chakir (2007) noted that the interactional qualities can be approached more appropriately starting with an in-depth conversation analysis to identify meaningful chunks and then enriching the analysis by applying coding and counting techniques. Hybrid methodology was also applied by Arvaja, Salovaara, Häkkinen and Järvela (2007) enhancing in-depth interpretations in order to analyze interactivity at the individual and the group level of collaborative discourse. The application of mixed methods to a complex phenomenon like the discourse process has been also generally discussed (Hmelo-Silver, 2003).

The interactive dimension of a discourse process was addressed by Henri (1992) distinguishing independent from implicit interactive and explicit interactive contributions depending on if and how they refer to other contributions in the discourse. Later, Gunawardena et al. (1997) extended this view noting that in a knowledge constructing discourse all contributions can be linked to one or more other contributions and to the discussed topic. This marks a change from understanding interactivity as reference between contributions to treating it as diffusion of knowledge and other more general forms of uptake or influence. Empirical analysis of interactivity according to Henri's definition was performed using a mapping technique (e.g. Hara, Bonk, & Angeli, 2000; Howell-Richardson & Mellar, 1996; Schrire, 2004) that allows visualization and classification of discussion threads based on their message structure. The holistic approach of Gunawardena et al. was also used (e.g. Fahy, Crawford, & Ally, 2001) for coding of contributions and general discourse characterization. In the knowledge building literature the semantic connection between contributions was analyzed to trace "inquiry threads" based on shared topics (Zhang, Scardamalia, Lamon, Messina, & Reeve, 2007). This was proposed by Lipponen, Rahikainen, Lallimo and Hakkarainen (2003) who noted that structural threads obscure the majority of semantic and conceptual relations between contributions. Manca, Delfino and

Mazzoni (2009) confirmed that by developing and using a semantic coding scheme that reveals much more dense communication flows. Suthers, Dwyer, Medina and Vatrapu (2010) introduced and employed contingency graphs for uptake analysis with the goal to identify and map also very subtle kinds of contribution uptakes that often remain unnoted.

Another approach to studying interactivity in discourse processes presents the social network analysis (SNA), which is a well established methodology in the CSCL research field (e.g. Aviv, Erlich, Ravid, & Geva 2003; Cho, Sefanone & Gay, 2002; de Laat, Lally, Lipponen, & Simons, 2007; Palonen & Hakkarainen, 2000; Reffay, & Chanier, 2002; Sha & van Aalst, 2003). It has been used for studying relations between *persons* embedded in a network. Networks are ususally defined based on the logged collaborative interaction between students, i.e. a network link between students is established when one reads or responds ot the contribution of another one. Analyses of such networks yield information on the cohesiveness of the learning group and on the position of single students relative to the others, at different points in time and overall.

Following the reasoning of Stahl (2006) on intersubjective meaning-making in CSCL in terms of networks of references the discourse process can also be approached from the perspective of a network of collaboratively created conceptual artifacts. There are accordingly a few notable examples of studies utilizing a network analysis of content relations, i.e. studying networks of related *contributions*. Both Sha, Teplovs, and van Aalst (2010) and Oshima and Oshima (2007) applied automatic algorithms for the identification of lexical relations between posted content in order to define a so-called semantic network of contributions and calculate general indices of the network and the position of single contributions. The authors performed additional steps like semantic clustering of the topics of the contributions. Irrespectively of the network type, persons or contributions, network analysis cannot deliver a complete picture of a discourse process, because it is based on surface indicators of relations that lack deeper meaning in the discourse. Therefore the application of network analysis relies on a combination with other in-depth and content related methods.

In sum, an appropriate methodology for studying discourse as a process has to be multi-faceted and address: the temporal dimension of development; the interactivity between contributions and between participants; the content of the discourse (Arvaja et al., 2007).

Main Path Analysis of Discourse Trajectory

The main goal of our study is to provide a primer for a novel methodology that tackles the trajectory of a discourse process analyzing the collaborative development of ideas over time. Our assumption is that time represents the main dimensionality of processes. Aggregations over the whole history of a discourse may yield biased results.

The approach presented here is rather simple and grounds on a network analysis of interrelated ideas. Based on a set of identified relations a main path analysis (Hummon & Doreian, 1989) assigns different weights to the relations according to their position in the network. This produces a differentiated visualization of the whole discourse trajectory that can be interpreted in order to identify the most influential ideas, idea paths, i.e. successions of related contributions, the developing structure of the paths over time, etc. The procedure of main path analysis stems from the scientometric research tradition that deals with citations between scholarly publications (e.g. Carley, Hummon, & Harty, 1993). In the present application of the analysis again relations between authored content are of interest. To the best of our knowledge, the present work is a pioneering attempt of applying scientometric methodology in the field of CSCL, and knowledge building in particular.

De Nooy, Mrvar and Batagelj (2005) reason that networks of scientific citations represent systems of knowledge flows. This perfectly fits our goal to analyze development of ideas in a knowledge building discourse. Moreover, Scardamalia and Bereiter (2006) made an explicit parallel between students in a knowledge building class and scientists conducting research and building on each others' achievements. Therefore, we conceptualize a network of contributed ideas based on content relations of manifest and latent building up actions. In accordance with the reviewed literature we regard explicit relations of thread links as equivalent to implicit relations between successive contributions that can be interpreted as diffusion, influence, reference, uptake, etc. A network constructed in that way represents the advancement of knowledge over time and can be analyzed as any scientific citation network. The resulting discourse trajectory is comparable to the evaluation of scientific history.

Main path analysis calculates a weight for each relation in a citation network according to number of times that relation is used when tracing all realized connections between all possible pairs of preceding and subsequent contributions (De Nooy et al., 2005). So, a relation on idea paths is more important, the more often it is used. Contributions that relate to many other preceding and subsequent contributions are easily identified as important for a discourse, as they synthesize old ideas, add new knowledge and represent a basis for developing new ideas. A main path in the discourse process can be identified by following the links with the highest weights. Special topics of pursued inquiry can be identified at the presence of separated main paths.

A discourse trajectory offers a dynamic view on the evolving knowledge and can be analyzed for instances of integration, fragmentation, specialization or paradigmatic changes (De Nooy et al., 2005). We therefore think that the present method is highly suited for the tracking of idea development and may highlight new aspects of online interaction among students. Moreover, we see very promising possibilities for combining it with in-depth and content analysis methods.

Method

Data Set

We reanalyzed Knowledge Forum data from a study of two classes of a Grade 10 Social Studies course — one regular and the other honors — who investigated aspects of environmental issues; see Niu & van Aalst (2009) for details. Despite some differences, the conclusion of the original study was that not only the honors but also the low-achieving students were able to profit from the course developing and sustaining a fruitful knowledge building discourse.

The students worked on Knowledge Forum in groups of approximately eight over a period of three weeks with daily synchronous sessions in the computer lab and occasional individual participation at home. Each group worked on a specific environmental issue. In the present study, we compare the discourses of one group in each class that worked on the handling of the nuclear accident at Chernobyl; the honors group wrote 60 notes and the regular group 164 notes on this issue. However, taken across all issues, students in the two classes wrote similar numbers of notes: 339 for the honors class, compared to 388 for the regular class. We analyzed the log data and the content of the computer notes; the log files contain writing (saving a note) and reading (opening a note) events, which were recorded with a timestamp while using Knowledge Forum.

Analysis

As we were interested in collaborative learning at the level of idea development in discourse it was important to take all ideas and all paths of development over time into account. We adopted Lipponen's (2000, p.185) definition that an "idea is a set of propositions that formed a coherent unit of meaning" and determined that a single note represents a suitable unit of analysis for our data that also simplifies the analysis (Rourke, Anderson, Garrison, & Archer, 2001). Direct linkages between notes as expressed by responses ("build-on" notes in Knowledge Forum) capture only a small part of all existing relations between notes, however. Thus, in order to capture relations between ideas, we made use of both — explicit links and implicit connections — by analyzing data from two different sources, the log files of the software as well as the content of the contributed notes. Based on the timestamps we were able to identify the notes each student had read before writing a new note. Building on an idea from Manca, Delfino and Mazzoni (2009), we then compared the content of each note the student had read and the content of her own new note. At this step we performed a qualitative analysis of the content relations between the notes. Even though a note may present new information taken from an external source, almost always it is also connected to some previously stated ideas, as put forward by Gunawardena et al. (1997). Hence, beyond including explicit references such as answering a question, citing a statement or agreeing with another student we identified more subtle but also important kinds of relations between ideas such as giving additional information or arguments, elaborating an idea, reasoning and summarizing based on a number of preceding ideas. In order to avoid an overestimation of relations between notes, we cautiously took the specific relations between note contents as well as the time interval between reading and writing into account. More precisely, a general relation of a note to the current focus of the discourse was not sufficient. Rather, our criterion for identifying a linkage between notes was the visible existence of a specific uptake or influence between ideas. With respect to the time interval we acted on the assumption that uptake rather occur between ideas created close in time. Both of these criteria, time interval and specificity of idea relations, were applied in consideration of the specific context of the whole discourse. As we were primarily interested in the paths of ideas development, the analyses reported here focus on the detection of links between posts. Beyond this, however, we also classified the links we detected following a grounded theory approach. The complete analysis of the coding is still in progress.

Having identified relations between notes based on a qualitative evaluation of the uptake of previous ideas, we employed a network approach to analyze the paths of idea development quantitatively. The goal was to explore the qualities of the discourse process in direct relation to its content. Similar to the analysis of citation networks for scientific publications, our networks visualized the relationships between a note and preceding notes. As we wanted to describe idea development over the course of time, we defined directed links from an earlier note to a subsequent related note. We then applied the main path analysis procedure (Hummon & Doreian, 1989; Carley, Hummon, & Harty, 1993) to the network, and obtained weights for each link based on its relative importance for all realized paths of idea development. This was done with the help of Pajek (Batagelj & Mrvar, 1998) network analysis software with standard procedures for the main path analysis.

Results

Figures 1 and 2 depict the results of the main path analysis of both group discourses. The vertically layered view illustrates idea development during the different days of activity in class and at home. Numbered points represent different notes. Arrows represent relatively important relations between notes with thicker arrows denoting more important relations regarding the whole discourse trajectory as calculated with the main path analysis method. The arrows are directed from an older to a newer note. It is important to mind that the isolated notes positioned on the left and on the right of the figures do not contain completely unrelated ideas. In fact, the initial coding process established idea relations between almost all notes for both discussions, so unweighted representations of the discussions would contain almost completely and much more densely connected networks. This is not surprising given the fact that our data was comprised of discussion posts that focused on a single topic, namely Chernobyl. We therefore focus on the most important relations and Figure 1 and 2 thus show only relations with weights greater than 0.05 for normalized weights between 0 and 1.

Another finding that can be seen in both figures is that the relatively important relations between the idea notes establish a single connected network of idea paths in both cases. This suggests that in both groups there were no disparate discussions around different topics. The main idea paths are interrelated as one would expect of a focused discussion. However, both figures show different discourse trajectories at a first glance that remain to be characterized with the help of the contributed content.

Honors Group (Figure 1) The notes from the first day (top of the figure) already generate different idea paths mostly addressing background information on causes and on effects of the accident. On April 4th, which was the second day of synchronous interaction in class, the honors group does not focus the development of ideas from the first day into one path. Instead, those ideas are interrelating and stimulating one another. The notes in this time

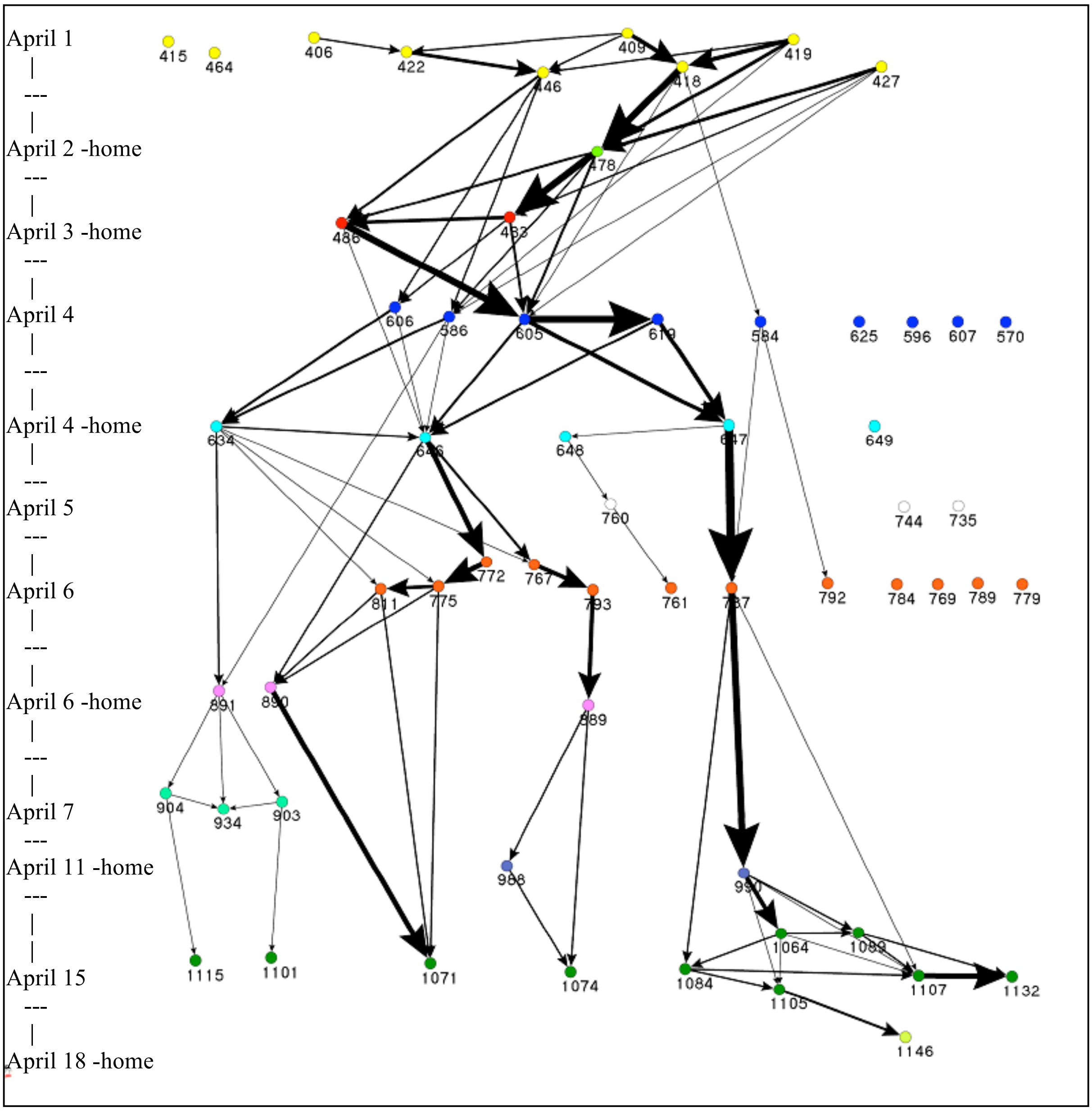

Figure 1. Discourse trajectory of the honors group: notes and idea relations with weights > 0.05

period deal with technological issues and safety measures. Hence, they are related to both the causes of the accident and potential future steps that might help to avoid future disasters. From April 4th and especially 6th on idea paths in the honors group are starting to separate resulting in four largely independent lines of inquiry that were pursued in the last days of the course. The thickest path deals with futuristic solutions of neutralizing radiation and inspires the largest amount of participation. The remaining three paths discuss the problem of the sarcophagus around the destroyed reactor, the use of nuclear energy in general and the politics behind the accident.

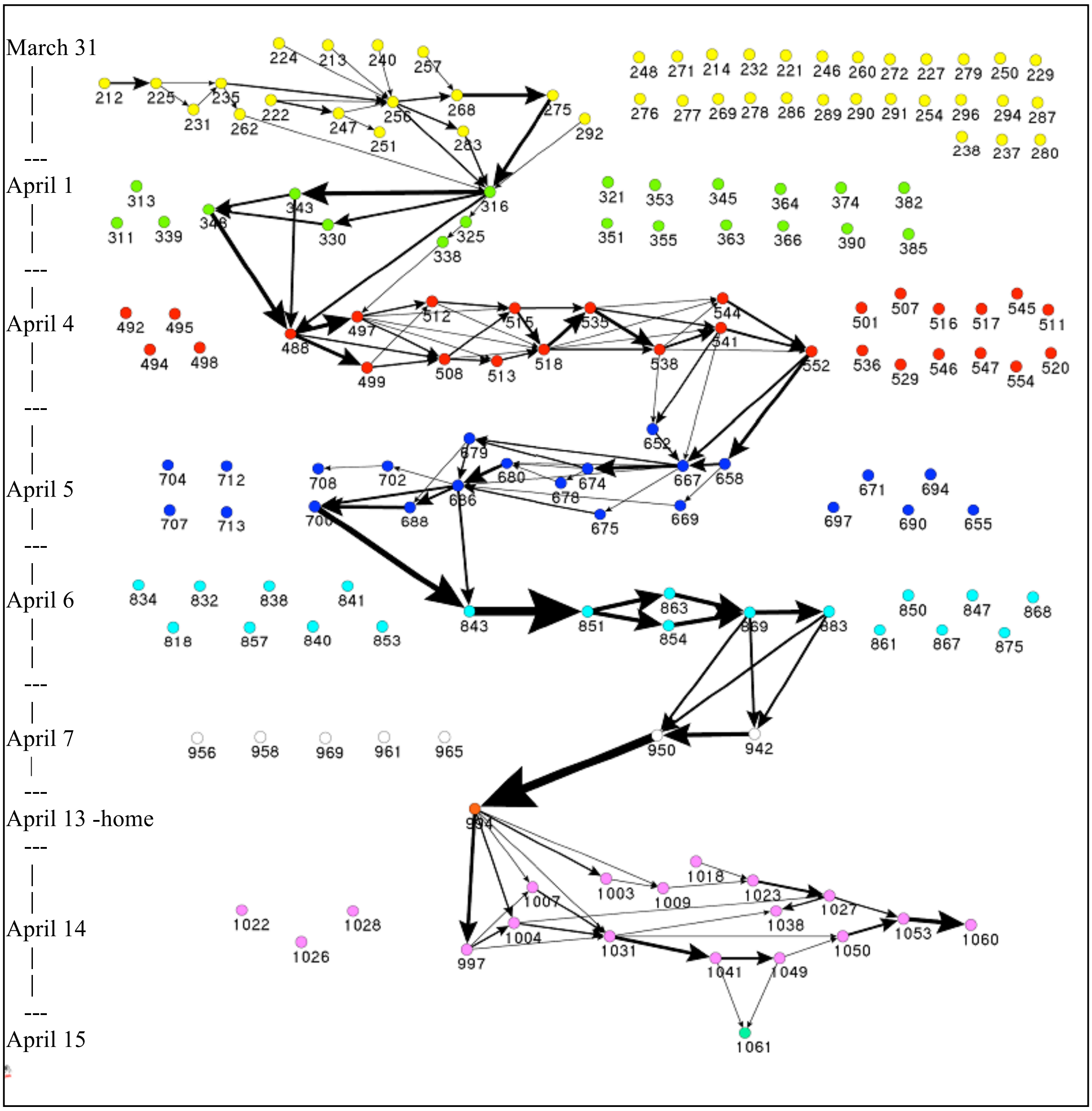

Figure 2. Discourse trajectory of the regular group: notes and idea relations with weights > 0.05

Regular Group (Figure 2) Initially there were different ideas, which, however, all merged into a single relevant note, 316, on April 1st. This may be understood as this note provided a summary of previous background information and also initiated a new topic by addressing a concrete problem and asking for solutions to it. Further tracking the process suggests that this pattern somehow reflects the discussion style within the regular class.

Repeatedly, one may identify overlapping idea developments that meld into a single note later that day or on the next day. For instance, the problem that was raised within the first central note, 316, was subsequently restated until a solution was proposed in the next central note, 488, that again unifies all previous paths. Then, again, this central idea gives rise to various paths, which involve notes that address different problems coming with the proposed solution (April 4th). During April 5th there are still more or less different paths, which contain elaborations regarding different solutions to the problems that emerged previously. A poll (note 843) on the preferences for the different solutions that have been worked out, ultimately joins the paths (April 5th and 6th). Then, interestingly, the elaboration of a previously unmentioned problem (April 6th and 7th) is brought to an end by note 934 that argues that the problem is irrelevant and again focuses the discussion. The discourse end partly with resignation and partly with new solutions and related problems, which is visible in the branching trajectory the discourse ends with.

Conclusion

The short comments on the results illustrate some initial ways of interpreting discourse trajectories obtained through main path analysis. With the present paper we pursued the goal to open up a field of possibilities for studying collaborative learning processes as we introduced a new method to the field. We showed that it handles the temporal perspective a process at the level of idea development very well. The application of the main path analysis to meaningfully defined relations provides an objective measure of the relevance of relations between notes and enables the identification of core ideas and central links. These can then be examined more closely with regard to their contents in a mixed methods approach (see also Carley et al., 1993). The obtained

discourse trajectories also provide a holistic view on the collaborative process. As in our examples groups may show different styles of collaboration regarding the convergence-divergence polarity and nevertheless be successful in knowledge building (Halatchliyski, Kimmerle, & Cress, 2011).

Our ideas for future research are connected to the present data set and methodology, as we plan to study the patterns of relations between ideas after finishing the coding procedure. It is interesting to compare what kind of contributions receive higher or lower weighting by the method, are more or less important for the process. Furthermore, we see possibilities to incorporate an individual view on students and their contribution to a discourse trajectory.

References

Arvaja, M., Salovaara, H., Häkkinen, P. & Järvela, S. (2007). Combining individual and group-level perspectives for studying collaborative knowledge construction in context. *Learning & Instruction*, 17, 448-459.

Aviv, R., Erlich, Z., Ravid, G., & Geva, A. (2003). Network Analysis of Knowledge Construction in Asynchronous Learning Networks. *Journal for Asynchron Learning Networks*, 7, 1-23.

Barnes, D., & Todd, F. (1977). *Communication and Learning in Small Groups*. London: Routledge and Kegan Paul.

Batagelj, V., & Mrvar, A. (1998). Pajek – Program for Large Network Analysis. *Connections*, 21, 47-57.

Bereiter, C. (2002). *Education and mind in the knowledge age*. Mahwah, NJ: Lawrence Erlbaum Associates.

Bereiter, C., & Scardamalia, M. (1993). *Surpassing ourselves: An inquiry into the nature and implications of expertise*. Chicago, IL: Open Court.

Carley, K. M., Hummon, N., & Harty, M. (1993). Scientific Influence: An Analysis of the Main Path Structure in the Journal of Conflict Resolution. *Knowledge: Creation, Diffusion, Utilization*, 14, 417-447.

Chi, M. T. H. (1997). Quantifying qualitative alalysis of verbal data: A practical guide. *The Journal of the Learning Sciences, 6*, 271-315.

Cho, H., Stefanone, M., & Gay, G., (2002). Social Network Analysis of Information Sharing Networks in a CSCL Community. In: G. Stahl (Ed.), *Proceedings of Computer Support for Collaborative Learning (CSCL) 2002 Conference, Jan. 7-11,* (pp. 43-50). Boulder, CO. Mahwah, NJ: Lawrence Erlbaum.

de Laat, M., Lally, V., Lipponen, L., & 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, 87-103.

de Nooy, W., Mrvar, A., & Batagelj, V. (2005). *Exploratory Social Network Analysis with Pajek*, Cambridge University Press, 2005.

de Wever, B., Schellens, T., Valcke, M., & van Keer, M. (2006). Content analysis schemes to analyze transcripts of online asynchronous discussion groups: A review. *Computers & Education*, 46, 6-28.

Fahy, P., Crawford, G., & Ally, M. (2001). Patterns of interaction in a computer conference transcript. International Review of Research in Open and Distance Learning, 2. http://www.irrodl.org/index.php/ irrodl/article/view/36/74 (Retrieved November 5, 2011).

Gunawardena, C. N., Lowe, C. A., & Anderson, T. (1997). Analysis of A Global Online Debate and The Development of an Interaction Analysis Model for Examining Social Construction of Knowledge in Computer Conferencing. *Journal of Educational Computing Research*, 17, 397-431.

Hakkarainen, K. (2003). Emergence of progressive-inquiry culture in computer-supported collaborative learning. *Learning Environments Research, 6*, 199-220. doi: 10.1023/a:1024995120180

Halatchliyski, I., Kimmerle, J., & Cress, U. (2011). Divergent and convergent knowledge processes on Wikipedia. In H. Spada, G. Stahl, N. Miyake, & N. Law (Eds.), *Connecting Computer-Supported Collaborative Learning to Policy and Practice: CSCL2011 Conference Proceedings* (Vol. II, pp. 566-570). Hong Kong: International Society of the Learning Sciences.

Hara, N., Bonk, C.J. & Angeli, C. (2000). Content analysis of online discussion in an applied educational psychology course. Instructional Science 28: 115–152.

Henri, F. (1992). Computer Conferencing and Content Analysis. In: A. Kaye (Ed.), Collaborative Learning through Computer Conferencing: The Najaden Papers (pp. 117-136). Berlin: Springer-Verlag.

Hmelo-Silver, C.E. (2003). Analyzing collaborative knowledge construction: multiple methods for integrated understanding. *Computers & Education*, 41, 397-420.

Howell-Richardson, C. & Mellar, H. (1996). A methodology for the analysis of patterns of participation within computer-mediated communication courses. Instructional Science 24: 24–69.

Hong, H.-Y., Chen, F. C., Chang, H. M., Liao, C. C. Y., & Chan, W. C. (2009). Exploring the effectiveness of an idea-centered design to foster a computer-supported knowledge building environment. In C. O'Malley, D. Suthers, P. Reimann, A. Dimitracopoulou (Eds.), Computer Supported Collaborative Learning Practices: CSCL2009 Conference Proceedings (pp. 142-150). Rhodes, Greece: International Society of the Learning Sciences, Inc.

Hummon, N. P., & Doreian, P. (1989). Connectivity in a Citation Network: The Development of DNA Theory. *Social Networks*, 11, 39-63.

Jeong, A. (2005). A guide to analyzing message-response sequences and group interaction patterns in computer-mediated communication. *Distance Education*, 26, 367-383.

Lipponen, L. (2000). Towards knowledge building: from facts to explanations in primary students' computer mediated discourse. *Learning Environments Research, 3*, 179-199.

Lipponen, L., Rahikainen, M., Lallimo, J., & Hakkarainen, K. (2003). Patterns of participation and discourse in elementary students computer-supported collaborative learning. *Learning and Instruction*, 13, 487-509.

Manca, S., Delfino, M., & Mazzoni, E. (2009). Coding procedures to analyse interaction patterns in educational web forums. *Journal of Computer Assisted Learning*, 25, 189–200.

Niu, H., & van Aalst, J. (2009). Participation in knowledge-building discourse: An analysis of online discussions in mainstream and honours social studies courses. *Canadian Journal of Learning and Technology 35*(1). http://www.cjlt.ca/index.php/cjlt/article (Retrieved November 5, 2011).

Oshima, J., Oshima, R., & Knowledge Forum ® Japan Research Group (2007). Complex network theory approach to the assessment on collective knowledge advancement through scientific discourse in CSCL. *Proceedings of CSCL2007*, pp. 563-565.

Oshima, J., Scardamalia, M., & Bereiter, C. (1996). Collaborative learning processes associated with high and low conceptual progress. *Instructional Science, 24*, 125-155.

Palnonen, T. & Hakkarainen, K. (2000). Patterns of interaction in computer-supported learning: A social network analysis. In: Fishman, B. und O'Conner-Divelbiss, S. (Eds.), *Proceedings of the Fourth International Conference on the Learning Sciences.* Mahwah, NJ: Lawrence Erlbaum.

Popper, K. R. (1972). *Objective knowledge: An evolutionary approach*. Oxford, UK: Clarendon Press.

Reffay, C. & Chanier, T. (2002) Social Network Analysis used for modelling collaboration in distance learning groups, In S. A. Cerri, G. Gouardères, and F. Paraguaçu (Eds.): *ITS 2002 Lecture Notes in Computer Science*, 2363, 31–40.

Reimann, P., Frerejean, J., & Thompson, K. (2009) Using process mining to identify models of group decision making processes in chat data. In C. O'Malley, et al. (Eds.), *Computer-supported collaborative learning practices: CSCL2009 conference proceedings*, (pp. 98-107). International Society for the Learning Sciences: Rhodes, Greece.

Rourke, L., Anderson, T., Garrison, D. R., & Archer, W. (2000) Methodological issues in the content analysis of computer conference transcripts. *International Journal of Artificial Intelligence in Education*, 11, 8-22.

Scardamalia, M., & Bereiter, C. (2006). Knowledge building: Theory, pedagogy, and technology. In R. K. Sawyer (Ed.), *The Cambridge handbook of the learning sciences* (pp. 97-115). New York, NY: Cambridge University Press.

Schrire, S. (2004). Interaction and cognition in asynchronous computer conferencing. *Instructional Science*, 32, 475-502.

Sha, L., & van Aalst, J. (2003). An application of social network analysis to knowledge building. Paper presented at the Annual Meeting of the American Educational Research Association, April 21-25, Chicago, IL.

Sha, L., van Aalst, J., & Teplovs, C. (2010). A visualization of group cognition: Semantic network
analysis of a CSCL community. In K. Gomez, L. Lyons, and J. Radinsky (Eds.)

Proceedings of the 9th International Conference of the Learning Sciences - Volume 1 (ICLS '10) (pp. 929-936).
Stahl, G. (2002). Rediscovering CSCL. In T. Koschmann, R. Hall & N. Miyake (Eds.), *CSCL 2: Carrying forward the conversation* (pp. 169-181). Mahwah, NJ: Lawrence Erlbaum Associates.
Stahl, G. (2006). *Group cognition: Computer support for building collaborative knowledge*. Cambridge, MA: MIT Press.
Stahl, G., Koschmann, T., & Suthers, D. (2006). Computer-supported collaborative learning: An historical perspective. In R. K. Sawyer (Ed.), Cambridge handbook of the learning sciences. Cambridge, UK: Cambridge University Press.
Strijbos, J. W., Kirschner, P. A., & Martens, R. L. (2004). What we know about CSCL: ... and we do not (but need to) know about CSCL. In P. Dillenbourg (Series Ed.) & J. W. Strijbos, P. A. Kirschner & R. L. Martens (Vol. Eds.), Computer-supported collaborative learning: Vol 3. What we know about CSCL:And implementing it in higher education (pp. 241-259). Boston, MA: Kluwer Academic Publishers.
Strijbos, J.-W., & Stahl, G. (2007). Methodological issues in developing a multi-dimensional coding procedure for small group chat communication. *Learning & Instruction*. Special issue on measurement challenges in collaborative learning research, 17, 394-404.
Suthers, D. D., Dwyer, N., Medina, R., & Vatrapu, R. (2010). A framework for conceptualizing, representing, and analyzing distributed interaction. *International Journal of Computer-Supported Collaborative Learning, 5*(1), 5-42.
Zemel, A., Xhafa, F., & Chakir, M. (2007). What's in the mix? Combining coding and conversation analysis to investigate chat-based problem solving. *Learning and Instruction*, 17, 405-415.
Zhang, J., Scardamalia, M., Lamon, M., Messina, R., & Reeve, R. (2007). Socio-cognitive dynamics of knowledge building in the work of 9- and 10-year-olds. *Educational Technology Research & Development, 55*(2), 117-145.

Acknowledgments

This study resulted from an exploration the authors began during the Summer School "Making Sense of Social Media", co-sponsored by the Knowledge Media Research Center and Tuebingen ScienceCampus, August 1-4, 2011. The data were originally collected as part of a New Economy Collaborative Research grant from the Social Sciences and Humanities Research Council of Canada to Marlene Scardamalia (OISE/University of Toronto). The authors thank Hui Niu (Simon Fraser University) for permission to reanalyze the Knowledge Forum database.