Comparison of Network Heuristics for Understanding Small Groups in Synchronous Online Learning

XXXXX
XXXX
XXXXXX
XXXXXXXX
XXXXXXXXX

XXXXX
XXXX
XXXXXX
XXXXXXXX
XXXXXXXXX

XXXXX & XXXXX
XXXX
XXXXXX
XXXXXXXX
XXXXXXXXX

ABSTRACT
Social network analysis metrics can provide information about group structures, based on information about relationships and activities between group participants, which in turn can provide information to tutors, computational agents and students. In the case of chat discussions, however, the relationship involved in a participant replying to another is not explicit. In this poster, we compare four preliminary heuristics for estimating a reply structure from which individual and group network measures can be derived. Simple participation measures prove remarkably effective and a coding framework based on Systemic Functional Linguistics is shown the most effective heuristic for individual measures (r^2=0.96).

Categories and Subject Descriptors
K1.1. [**Computer Uses in Education**]: Collaborative Learning

General Terms
Measurement, Human Factors.

Keywords
Chat discussion, network heuristics, SNA.

1. INTRODUCTION
Social Network Analysis has been shown to be useful in contributing to understanding online discussions in Computer Supported Collaborative Learning (CSCL) situations [1][4], even in the case of small groups [18]. The insights gleaned from such techniques can be used in contributing to analysis, or in real-time to provide monitoring to group, to a human tutor or to a computational agent which is involved in scaffolding the situation.

Online chat discussion logs are the empirical manifestation of the complex group dynamic that unfolds during the discussion. Their simple sequential structure does not explicitly provide the links necessary for network analysis techniques to be applied. By manually or automatically adding this structure, it is possible to construct networks of sufficient quality for identification of major features of a discussion and the roles of actors [31]. Some authors have suggested using interfaces with explicit reply or even forums to obtain these structures [14], but these can be unreliably used (e.g. [30]). Using quotes [27] has been shown to be more reliable than relying on the reply structure, as has overlap in content [32]. In the case of asynchronous discussion without explicit structure, Goggins [10] shows that by considering the number of previous posts visible to a user, and factoring in time between messages, it is possible to construct an appropriate implicit reply structure. In the case of synchronous chat discussion, positive results have been achieved, combining temporal proximity, turn proximity and direct addressing of users [26].

In using automated techniques to expose implicit reply structures, we do not assume that the resulting reply structure will be exactly correct, but there is evidence that automated analysis at a turn by turn level may capture enough of underlying interaction that higher level indicators (such as network measures) might be expected to be valid. For example, Mayfield & Rosé [22] have achieved 68% accuracy in automatically applying a framework based on systemic functional linguistics that codes for negotiation of information, which resulted in an r^2 of .94 on an authoritativeness indicator (calculated at the participant level) derived from the automated coding. It is therefore reasonable to believe that both individual and group measures might be derived from an automated coding of reply structures.

In this poster, we present a comparison of four heuristics in their preliminary stages of elaboration. Our comparison is based on individual measures of centrality, and group measures of amount and equality of collaborative participation, with a gold standard derived from the manual coding of a reply structure. Our baseline heuristic is derived from counts of participation. A first heuristic from the perspective of adapting asynchronous Group Informatics heuristics [13] to synchronous discussion and two further heuristics from the perspective of Negotiation coding [22] are presented below. We find that the baseline performs surprisingly well, surpassed/equaled only by the Negotiation heuristic. In comparison, the Group Informatics heuristic performs poorly. We discuss why this may be the case and lay the grounds for further exploration of such heuristics.

2. CONTENTS OF THE POSTER
In the poster, we will present each of the heuristics in detail, showing with visual examples the meaning of each of the heuristics we use and describe the reasoning behind each of the heuristics being expected to work. The gold standard is manually coded. The group informatics heuristic extends a heuristic which was originally designed for asynchronous discussion and takes media affordances for reading and replying, as well as time into

account. The negotiation heuristics are based on a socio-linguistic coding which expect certain sequences to exist in interaction, allowing reply structure data to be inferred in various ways.

Based on these heuristics we create individual measures of participation/centrality and group measures of cohesiveness. For each heuristic, we calculate the r^2 coefficient between it and the gold standard, presented in Table 1. For the Degree Centrality, n=42, for Mean Degree Centrality and Coefficient of Variation, n=14.

Table 2. Correlation between gold standard and each heuristic

	Degree Centrality	Mean Degree Centrality	Coefficient of Variation
Turn Counts	0.85	**0.94**	**0.87**
Group Informatics	0.13	0.22	0.43
Liberal Negotiation	0.74	0.88	0.54
Conservative Negotiation	**0.96**	0.25	0.38

We conclude by discussing the meaning of network analytics in small groups and the potential value in predicting network structures, even in these extreme cases.

3. EXPECTED INTERACTIONS WITH PARTICIPANTS

We hope to connect with participants who have non-structured discourse data from which network measures might be derived via heuristics and to share methods with which this might be achieved. We are also interested in the fact that turn counts were more effective than might have been expected and wonder whether this result would carry over to larger groups. These interactions will help us shape our thinking in what is currently a side-track of our research, in order to decide in what direction we should go to bring maturity to our work.

4. POSTER FORMAT

While this was originally a short paper and we have not yet taken the time to produce a paper, we expect graphics and very short text to be able to support 5 of the 6 sections of the poster (basic idea, 3 network heuristics, results), while the discussion will be more bullet-pointed.

5. REFERENCES

[1] Aviv, R., Erlich, Z., Ravid, G., & Geva, A. 2003. Network Analysis of Knowledge Construction in Asynchronous Learning Networks. *Journal of Asynchronous Learning Networks, 7(3)*.

[2] Berry, M. 1981. Towards Layers of Exchange Structure for Directive Exchanges. In *Network* 2.

[3] Blincoe, K., Valetto, G., & Goggins, S. 2012. *Leveraging Task Contexts for Managing Developers' Coordination*. Proceedings from ACM Conference on Computer Supported Cooperative Work, 2012, 1351-1360, Seattle, WA.

[4] 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(1), 87-103.

[5] Dyke, G., Adamson, D., Howley, I., Rosé, C. 2012. Towards Academically Productive Talk Supported by Conversational Agents. *ITS 2012*.

[6] Erickson, T. 2009. 'Social'systems: designing digital systems that support social intelligence. *AI & Society*, *23*(2), 147-166.

[7] Fine, G. A., & Harrington, B. 2004. Tiny publics: small groups and civil society. *Sociological Theory*, *22*(3), 341-356. Retrieved from http://onlinelibrary.wiley.com/doi/10.1111/j.0735-2751.2004.00223.x/abstract

[8] Goggins, S., Laffey, J., & Amelung, C. 2011. *Context Aware CSCL: Moving Toward Contextualized Analysis*. Proceedings from CSCL 2011, 591-596, Hong Kong.

[9] Goggins, S., Mascaro, C., & Mascaro, S. 2012. *Relief after the 2010 Haiti Earthquake: Participation and Leadership in an Online Resource Coordination Network*. Proceedings from ACM Conference on Computer Supported Cooperative Work, 57-66, Seattle, WA.

[10] Goggins, S., Galyen, K., & Laffey, J. 2010. *Network Analysis of Trace Data for the Support of Group Work: Activity Patterns in a Completely Online Course*. Proceedings from ACM Group 2010, 107 - 116, Sanibel Island, FL.

[11] Goggins, S., Valetto, P., Mascaro, C., & Blincoe, K. 2012. Towards Group Awareness through the Analysis of Task Context Data. *User Modeling and User-Adapted Interaction: The Journal of Personalization Research*, *Accepted*.

[12] Goggins, S. P., Laffey, J., & Gallagher, M. 2011. Completely online group formation and development: small groups as socio-technical systems. *Information Technology & People*, *24*(2), 104-133. doi:10.1108/09593841111137322

[13] Goggins, S. P., Mascaro, C., & Valetto, G. 2012. Group Informatics: A Methodological Approach and Ontology for Understanding Socio-Technical Groups. *JASIS&T*, *Accepted with Revisions*.

[14] Harrer, A. Zeini, S., Pinkwart, N. 2006. Evaluation of communication in web-supported learning communities ? an analysis with triangulation research design. *International Journal of Web-Based Communities* 2(4): 428-446

[15] Howison, J., Wiggins, A., & Crowston, K. 2012. Validity Issues in the Use of Social Network Analysis with Digital Trace Data. *Journal of the Association of Information Systems*, *12(2)*(2).

[16] Howley, I., Mayfield, E. & Rosé, C. 2011. Missing Something? Authority in Collaborative Learning. *CSCL 2011*.

[17] Latour, B. 2007. *Reassembling the Social: An Introduction to Actor-Network Theory*. Oxford.

[18] Martin, J. 1992. *English Text: System and Structure.*

[19] Martin, J., & Rose, D. 2003. *Working with Discourse: Meaning Beyond the Clause*.

[20] Martinez, A., Dimitriadis, Y., Rubia-Avi, B., Gomez-Sanchez, E., & de la Fuente, P. (2003). Combining qualitative evaluation and social network analysis for the study of classroom social interactions. Computers & Education, 41, 353-368.

[21] Mayfield, E., Adamson, D., & Rosé, C.P. 2012. Hierarchical Conversation Structure Prediction in Multi-Party Chat. In *SIGDIAL Meeting on Discourse and Dialogue*.

[22] Mayfield, E., & Rosé, C.P. 2011. Recognizing Authority in Dialogue with an Integer Linear Programming Constrained Model. *ACL 2011*.

[23] McCulloh, I. 2009. *Detecting Changes in a Dynamic Social Network.* Carnegie Mellon University.

[24] McCulloh, I., & Carley, K. 2009. Longitudinal Dynamic Network Analysis Using the Over Time Viewer Feature in ORA. Retrieved from http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.161.3492

[25] Mead, M. 1958. Cultural determinants of behavior. *Behavior and evolution,* 480-503.

[26] Mutton, P. 2004. Inferring and Visualizing Social Networks on Internet Relay Chat. In Proceedings of the Information Visualisation, Eighth International Conference (IV '04). IEEE Computer Society, Washington, DC, USA, 35-43. DOI=10.1109/IV.2004.75 http://dx.doi.org/10.1109/IV.2004.75

[27] Sack, W., Detienne, F., Ducheneaut, N., Burkhardt, J.-M., Mahendran, D., & Barcellini, F. (2006). A methodological framework for socio-cognitive analyses of collaborative design of open source software. Computer Supported Cooperative Work, 15, 229-250.

[28] Schmidt, K., & Bannon, L. 1992. Taking CSCW seriously. *Computer Supported Cooperative Work (CSCW)*, *1*(1), 7-40.

[29] Stahl, G. 2006. *Group Cognition: Computer Support for Building Collaborative Knowledge.* Boston, MA: MIT Press.

[30] Suthers, D. D., Dwyer, N., Medina, R., & Vatrapu, R. (2010). A framework for conceptualizing, representing, and analyzing distributed interaction. I*nternational Journal of Computer Supported Collaborative Learning*, 5(1), 5-42.

[31] Suthers, D. D., & Desiato, C. (2012). Exposing chat features through analysis of uptake between contributions. Proceedings of the Hawaii International Conference on the System Sciences (HICSS-45), January 4-7, 2012

[32] Trausan-Matu, S, & Rebedea, T. (2010). A Polyphonic Model and System for Inter-animation Analysis in Chat Conversations with Multiple Participants. *Computational Linguistics and Intelligent Text Processing*. Lecture Notes in Computer Science, 2010, Volume 6008/2010, 354-363, DOI: 10.1007/978-3-642-12116-6_29

[33] Turner, W., Bowker, G. C., Gasser, L., & Zacklad, M. 2006. Information Infrastructures for Distributed Collective Practices. *Computer Supported Cooperative Work*, *15*, 93-110.

[34] Veel, R. 1999. Language, Knowledge, and Authority in School Mathematics. In *Pedagogy and the Shaping of Consciousness: Linguistics and Social Processes*