Cohesion-Centered Analysis of CSCL Environments using ReaderBench

Mihai Dascalu, Gabriel S Gutu, Ionut S Cristian Paraschiv, Stefan S Ruseti, Philippe Dessus, Danielle A Mcnamara, Scott A Crossley, Stefan Trausan-Matu

Cohesion-Centered Analysis of CSCL Environments using *ReaderBench*

Mihai Dascalu[1,2], Gabriel Gutu[1], Ionut Cristian Paraschiv[1], Stefan Ruseti[1], Philippe Dessus[3], Danielle S. McNamara[4], Scott A. Crossley[5], Stefan Trausan-Matu[1,2]

[1] University Politehnica of Bucharest, Splaiul Independenței 313, 60042, Romania
{mihai.dascalu, gabriel.gutu, ionut.paraschiv, stefan.ruseti, stefan.trausan}@cs.pub.ro
[2] Academy of Romanian Scientists, Splaiul Independenţei 54, 050094, Bucharest, Romania
[3] Laboratoire des Sciences de l'Éducation, Univ. Grenoble Alpes, F38000 Grenoble, France
philippe.dessus@univ-grenoble-alpes.fr
[4] Institute for the Science of Teaching & Learning, Arizona State University, Tempe, USA
dsmcnama@asu.edu
[5] Georgia State University, Department of Applied Linguistics/ESL, Atlanta, 30303, USA
scrossley@gsu.edu

Abstract. Computer Supported Collaborative Learning (CSCL) environments have become a viable learning alternative from which valuable data can be extracted and used for advanced analyses centered on evaluating participants' involvement and their interactions. Such automated assessments are implemented within our *ReaderBench* framework, a Natural Language Processing platform that contains multiple advanced text analysis functionalities. The ReaderBench framework is based on Cohesion Network Analysis from which different sociograms, relying on semantic similarity, are generated in order to reflect interactions between participants. In this paper, we briefly describe the enforced mechanisms used to analyze three scenarios: individual chat conversations, virtual communities of practice and MOOCs.

Keywords: CSCL, Cohesion Network Analysis, Textual Cohesion, Natural Language Processing, Virtual Communities, Massive Open Online Course.

1 Introduction

Teachers and tutors have a limited amount of time to manually assess and grade student output. Moreover, monitoring and scoring student activities using indicators reflective of their performance in terms of participation or collaboration with peers is a cumbersome processes. Hence, there is necessity for automated analyses, which led us to develop our Cohesion Network Analysis (CNA) approach and integrate this approach within the *ReaderBench* framework available at http://readerbench.com. *ReaderBench* [1, 2] is a fully functional open-source framework centered on discourse analysis that consists of various Natural Language Processing (NLP) techniques designed to support students and teachers in their educational activities. This paper presents a brief overview of Computer Supported Collaborative Learning (CSCL) experiments performed within *ReaderBench* which help to validate the framework.

2 Performed Experiments

Three experiments in different CSCL environments have been conducted. These experiments focused on chat conversations, virtual communities, and a Massive Open Online Course (MOOC). In ReaderBench, CNA is defined by cohesive links spanning throughout the discourse and is tightly coupled with dialogism and polyphony which define the theoretical framing of CSCL. Moreover, CNA closely resembles Social Network Analysis by relying on equivalent indices to quantify participation within the generated sociograms [1, 2].

The first CNA evaluated the involvement of participants in CSCL ***chats*** [2]. The statistical analysis, run on a corpus of 10 multi-participant academic conversations, showed that cumulative utterance scores and outdegree were significant predictors of participation, while the cumulative utterance scores highlighted different participant interactions. The second CNA generated significant, moderate to strong, and positive correlations with human dialogue assessments while analyzing academic ***Virtual Communities of Practice*** [3]. This analysis showed that dialogue quality is higher for full time faculty members as they sustain a more elaborated dialogue and initiate longer discussion threads than part time members of the community (i.e., students). Third, CNA was used to predict students' completion rates in the context of the Big Data in Education ***MOOC*** [4]. Based on a cohesion graph and a longitudinal analysis, 56 indices were computed for each of the 300 students; twenty-six indices were statistically relevant and subsequently used in a Discriminant Function Analysis. Our method predicted the completion of the course with 76% accuracy, showing that students that are active on forums are more likely to finish the course.

These experiments demonstrate the capability of our *ReaderBench* framework to analyze a wide range of CSCL environments and to perform in-depth, cohesion-centered analyses of conversations' characteristics with a high accuracy.

Acknowledgments. This research was partially supported by the FP7 2008-212578 LTfLL project, by the 644187 EC H2020 RAGE project, as well as by the "Excellence Research Grants" Program UPB–GEX 12/26.09.2016.

References

1. Dascalu, M., Trausan-Matu, S., McNamara, D.S., Dessus, P.: ReaderBench – Automated Evaluation of Collaboration based on Cohesion and Dialogism. International Journal of Computer-Supported Collaborative Learning, 10(4), 395–423 (2015)
2. Dascalu, M., McNamara, D.S., Trausan-Matu, S., Allen, L.K.: Cohesion Network Analysis of CSCL Participation. Behavior Research Methods, 1–16 (2017)
3. Nistor, N., Trausan-Matu, S., Dascalu, M., Duttweiler, H., Chiru, C., Baltes, B., Smeaton, G.: Finding student-centered open learning environments on the internet: Automated dialogue assessment in academic virtual communities of practice. Computers in Human Behavior, 47(1), 119–127 (2015)
4. Crossley, S.A., Dascalu, M., Baker, M., McNamara, D.S., Trausan-Matu, S.: Predicting Success in Massive Open Online Courses (MOOC) Using Cohesion Network Analysis. In: CSCL 2017. ISLS, Philadelphia, PA (2017)

As a continuation of the interactive events at AIED 2013, LAK 2015, AIED 2015 and EC-TEL 2015, this demo is focused on the latest experiments centered on CSCL performed with the *ReaderBench* framework. All functionalities of *ReaderBench* (https://git.readerbench.com/ReaderBench/ReaderBench) are wrapped into dedicated packages and most of them can be tested freely online at http://readerbench.com. A demo developed for CSCL analysis is available at http://readerbench.com/demo/cscl.

Equipment provided by us: 1 Laptop (Mac Pro 15"), no additional hardware required.

Installed environment: Java JDK 1.8

Internet connection required to access *ReaderBench* website, in addition to the remote UPB server exposing the REST APIs. For demonstration purposes, serialized outputs will be used in order to speed up processing.