ReaderBench: An Integrated Cohesion-Centered Framework

Mihai Dascalu[1], Larise L. Stavarache[1], Philippe Dessus[2], Stefan Trausan-Matu[1], Danielle S. McNamara[3], Maryse Bianco[2]

[1] University Politehnica of Bucharest, Computer Science Department, Bucharest, Romania
{mihai.dascalu, stefan.trausan}@cs.pub.ro, larise.stavarache@ro.ibm.com
[2] LSE, Univ. Grenoble Alpes, Grenoble, France
{philippe.dessus, maryse.bianco}@upmf-grenoble.fr
[3] LSI, Arizona State University, Tempe, USA
dsmcnama@asu.edu

Abstract. *ReaderBench* is an automated software framework designed to support both students and tutors by making use of text mining techniques, advanced natural language processing, and social network analysis tools. *ReaderBench* is centered on comprehension prediction and assessment based on a cohesion-based representation of the discourse applied on different sources (e.g., textual materials, behavior tracks, metacognitive explanations, Computer Supported Collaborative Learning – CSCL – conversations). Therefore, *ReaderBench* can act as a Personal Learning Environment (PLE) which incorporates both individual and collaborative assessments. Besides the a priori evaluation of textual materials' complexity presented to learners, our system supports the identification of reading strategies evident within the learners' self-explanations or summaries. Moreover, *ReaderBench* integrates a dedicated cohesion-based module to assess participation and collaboration in CSCL conversations.

Keywords: Textual Complexity Assessment, Identification of Reading Strategies, Comprehension Prediction, Participation and Collaboration Evaluation.

1 ReaderBench's Purpose

Designed as support for both tutors and students, our implemented system, *ReaderBench* [1, 2], can be best described as an educational learning helper tool to enhance the quality of the learning process. *ReaderBench* is a fully functional framework that enhances learning using various techniques such as textual complexity assessment [1, 2], voice modeling for CSCL discourse analysis [3], topics modeling using Latent Semantic Analysis and Latent Dirichlet Allocation [2], and virtual communities of practice analysis [4]. Our system was developed building upon indices provided in renowned systems such as *E-rater*, *iSTART, and Coh-Metrix.*

However, ReaderBench provides an integration of these systems. ReaderBench includes multi-lingual comprehension-centered analyses focused on semantics, cohesion and dialogism [5]. For *tutors*, ReaderBench provides a) the evaluation of reading material's textual complexity, b) the measurement of social collaboration within a group endeavors, and c) the evaluation of learners' summaries and self-explanations. For *learners*, ReaderBench provides a) the improvement of learning capabilities through the use of reading strategies, and b) the evaluation of students' comprehension levels and performance with respect to other students. *ReaderBench* maps directly onto classroom education, combining individual learning methods with Computer Supported Collaborative Learning (CSCL) techniques.

2 Envisioned Educational Scenarios

ReaderBench (RB) targets both tutors and students by addressing individual and collaborative learning methods through a cohesion-based discourse analysis and dialogical discourse model [1]. Overall, its design is not meant to replace the tutor, but to act as support for both tutors and students by enabling continuous assessment. Learners can assess their self-explanations or collaborative contributions within chat forums. Tutors, on the other hand, have the opportunity to analyze the proposed reading materials in order to best match the student's reading level. They can also easily grade student summaries or evaluate students' participation and collaboration within CSCL conversations. In order to better grasp the potential implementation of our system, the generic learning flows behind *ReaderBench*, which are easily adaptable to a wide range of educational scenarios, are presented in Figures 1 and 2.

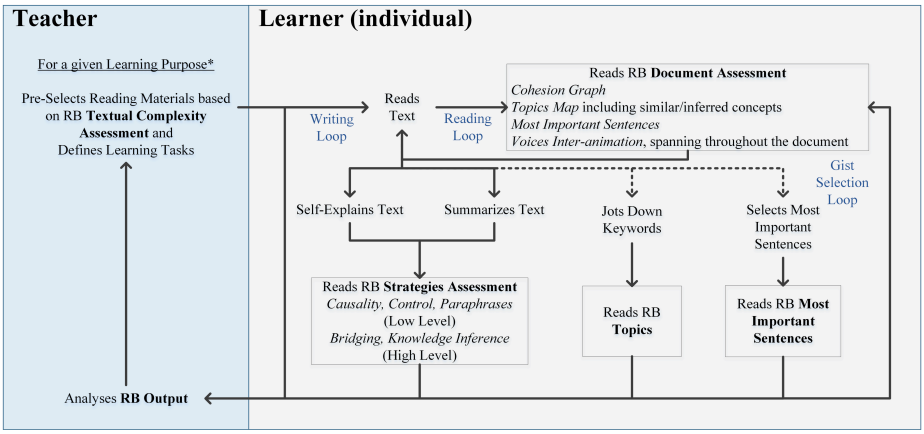

Fig. 1. Generic individual learning scenario integrating the use of *ReaderBench* (RB).

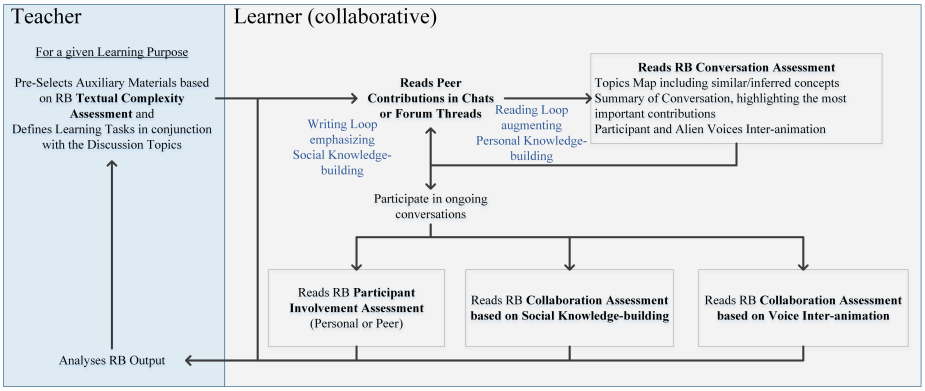

Fig. 2. Generic collaborative learning scenario integrating the use of *ReaderBench* (RB).

3 Validation Experiments

Multiple experiments have been performed, out of which only three are selected for brief presentation. Overall, various input sources were used for validating *ReaderBench* as a reliable educational software framework.

Experiment 1 [6] included 80 students between 8 and 11 years old (3^{rd}–5^{th} grade), uniformly distributed in terms of their age who were asked to explain what they understood from two French stories of about 450 words. The students' oral self-explanations and their summaries were recorded and transcribed. Additionally, the students completed a posttest to assess their comprehension of the reading materials. The results indicated that paraphrases and the frequency of rhetorical phrases related to metacognition and self-regulation (e.g., "*il me semble*", "*je ne sais*", "*je comprends*") and causality (e.g., "*puisque*", "*à cause de*") were easier to identify than information or events stemming from students' experiences. Furthermore, cohesion with the initial text, as well as specific textual complexity factors, increased accuracy for the prediction of learners' comprehension.

Experiment 2 [3] included 110 students who were each asked to manually annotate 3 chats out of 10 selected conversations. We opted to distribute the evaluation of each conversation due to the high amount of time it takes to manually assess a single discussion (on average, users reported 1.5 to 4 hours for a deep understanding). The results indicated a reliable automatic evaluation of both participation and collaboration. We validated the machine vs. human agreement by computing intra-class correlations between raters for each chat (avg $ICC_{participation}$ = .97; avg $ICC_{collaboration}$ = .90) and non-parametric correlations to the automatic scores (avg $Rho_{participation}$ = .84; avg $Rho_{collaboration}$ = .74). Overall, the validations supported the accuracy of the models built on cohesion and dialogism, whereas the proposed methods emphasized the dialogical perspective of collaboration in CSCL conversations.

Experiment 3 [7] consisted of building a textual complexity model that was distributed into five complexity classes and directly mapped onto five primary grade

classes of the French national education system. Multiclass Support Vector Machine (SVM) classifications were used to assess exact agreement (EA = .733) and adjacent agreement (AA = .933), indicating that the accuracy of classification was quite high. Starting from the previously trained textual complexity model, a specific corpus comprising of 16 documents was used to determine the alignment of each complexity factor to human comprehension scores. As expected, textual complexity cannot be reflected in a single factor, but through multiple categories. Although the 16 documents were classified within the same complexity class, significant differences for individual indices were observed.

In conclusion, we aim through *ReaderBench* to further explore and enhance the learning and instructional experiences for both students and tutors. Our goal is to provide more rapid assessment, encourage collaboration and expertise sharing, while tracking the learners' progress with the support of our integrated framework.

Acknowledgments. This research was partially supported by the 644187 RAGE H2020-ICT-2014 and the 2008-212578 LTfLL FP7 projects, by the NSF grants 1417997 and 1418378 to ASU, as well as by the POSDRU/159/1.5/S/132397 and 134398 projects. We are also grateful to Cecile Perret for her help in preparing this paper.

References

1. Dascalu, M.: Analyzing discourse and text complexity for learning and collaborating, Studies in Computational Intelligence, Vol. 534. Springer, Switzerland (2014)
2. Dascalu, M., Dessus, P., Bianco, M., Trausan-Matu, S., Nardy, A.: Mining texts, learners productions and strategies with ReaderBench. In: Peña-Ayala, A. (ed.) Educational Data Mining: Applications and Trends, pp. 335–377. Springer, Switzerland (2014)
3. Dascalu, M., Trausan-Matu, S., Dessus, P.: Validating the Automated Assessment of Participation and of Collaboration in Chat Conversations. In: ITS 2014, Vol. LNCS 8474, pp. 230–235. Springer, Honolulu, USA (2014)
4. Nistor, N., Trausan-Matu, S., Dascalu, M., Duttweiler, H., Chiru, C., Baltes, B., Smeaton, G.: Finding student-centered open learning environments on the internet. Computers in Human Behavior, 47(1), 119–127 (2015)
5. Dascalu, M., Trausan-Matu, S., Dessus, P., McNamara, D.S.: Discourse cohesion: A signature of collaboration. In: 5th Int. Learning Analytics & Knowledge Conf. (LAK'15), pp. 350–354. ACM, Poughkeepsie, NY (2015)
6. Dascalu, M., Dessus, P., Bianco, M., Trausan-Matu, S.: Are Automatically Identified Reading Strategies Reliable Predictors of Comprehension? In: ITS 2014, Vol. LNCS 8474, pp. 456–465. Springer, Honolulu, USA (2014)
7. Dascalu, M., Stavarache, L.L., Trausan-Matu, S., Dessus, P., Bianco, M.: Reflecting Comprehension through French Textual Complexity Factors. In: ICTAI 2014, pp. 615–619. IEEE, Limassol, Cyprus (2014)