Multi-Modal Peer Discussion with RichReview on edX

Dongwook Yoon
Cornell University
Ithaca, NY 14850
dy252@cornell.edu

Piotr Mitros
edX
Cambridge, MA 02139
pmitros@edx.org

ABSTRACT
In this demo, we present RichReview, a multi-modal peer discussion system, implemented as an XBlock in the edX courseware platform. The system brings richness similar to face-to-face communication into online learning *at scale*. With this demonstration, we discuss the system's scalable back-end architecture, semantic voice editing user interface, and a future research plan for the profile based group-assignment scheme.

Author Keywords
Massive open online courses; peer discussion; multi-modal annotation; voice user interface; peer group assignment.

ACM Classification Keywords
H.5.3. Group and Organization Interfaces: Collaborative computing; H.5.2. User Interfaces: Interaction styles; H.5.1. Multimedia Information Systems: Audio input/output

INTRODUCTION
Peer discussion in classrooms can improve student learning outcomes with interactive engagement on the course concepts [1]. Massive open online course providers provide a large number of students with online peer discussion activities through discussion forums [2], video chats, and peer feedback on assignments [3]. These face a trade-off between the richness of synchronous interaction and the flexibility of asynchronous interactions.

We developed a multi-modal annotation system called RichReview [5] which brings a high level of expressivity into asynchronous document annotations through the use of student-recorded multimedia, including text, audio, and gestures. The system was originally built for the purpose of writing feedback, but was redesigned to support a range of modes of online peer discussion.

In this work, we demonstrate how integration of RichReview into a MOOC platform can potentially open an expressive discussion channel in asynchronous environments with large numbers of students. As the first step, we re-implemented the RichReview's front-end, integrating it into Open edX. We designed a scalable back-end architecture for transmitting and storing multimedia comments created by a number of students. We partner these with a novel peer group assignment scheme that maximizes overall diversity of group composition using student profile data.

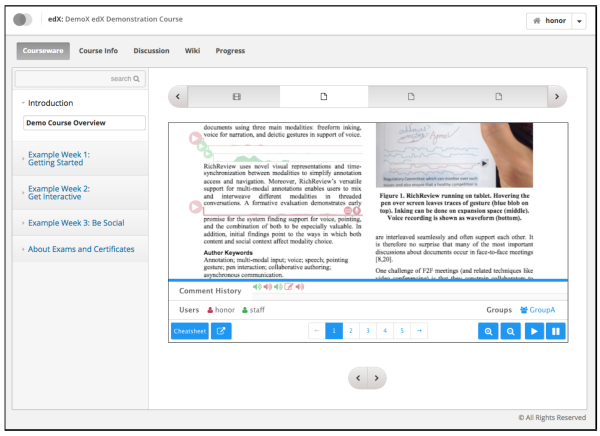

Figure 1. The RichReview XBlock running in the edX courseware.

RICHREVIEW MULTI-MODAL ANNOTATION SYSTEM
RichReview is a tablet based multi-modal annotation system for bringing richness of in-person conversation into document writing revision process. With RichReview, a commentator can record a combination of input modalities, such as pen writing and hovering, voice recording, as well as traditional modalities such as text. For example, RichReview users can verbally explain a math concept while pointing to a formula in a document and drawing a graph. Moreover, RichReview embeds the annotation thread within text lines of the annotated document, giving clear context for the comments. A prior lab study showed high potential of the system as a support tool for document-centric conversation [5].

INTEGRATION INTO EDX.ORG
Prior work demonstrated the value of RichReview in traditional classroom settings [6]. In order to more accurately measure the impact of RichReview, we would like to integrate it into a setting which supports massive numbers of learners. Open edX is a platform which support both MOOCs, with millions of learners on the edx.org web site, a large level of residential use (over 80% of MIT students use edX residentially), and over 100 open source deployments, including several deployments as part of national initiatives to improve training and education. Open edX was chosen due to the combination of support for research (including frameworks for

UIST'15 Adjunct, November 8–11, 2015, Charlotte, NC, USA.
ACM 978-1-4503-3780-9/15/11.
http://dx.doi.org/10.1145/2815585.2817809

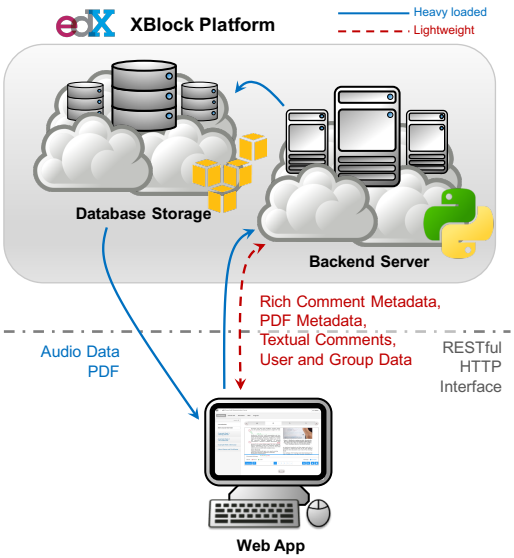

Figure 2. Scalable back-end architecture of the RichReview on edX.

randomized control trials, data collection, and frameworks for integration with novel pedagogical experiences like RichReview), anecdotal data from residential trials, and statistical evidence from MOOC trials.

A key to success is seamless integration into the edX UX, and robustly supporting tens of thousands of learners. As part of bringing RichReview to scale, we re-implemented the RichReview system as an XBlock — the components out of which Open edX courses are built. This allows the MOOC authors to include RichReview discussion sessions without managing multiple services, and for MOOC students to use RichReview seamlessly integrated into the edX interface. As can be seen in Fig. 1, the instructor can place the RichReview discussion session within the flow of the course contents. The overall architecture is shown in Fig. 2

Making the Back-end Scalable

RichReview is a media-heavy web application that exchanges large amounts of audio-visual data with the server. edX.org courses typically have tens of thousands of learners. In order to scale, we divided the data managed by RichReview into (1) high-bandwidth multimedia data (e.g. audio, PDF documents), and (2) low-bandwidth data (e.g. textual notes and metadata). The heavy-weight components are distributed through cloud file storage, which gives complete horizontal scalability. The metadata and other small data live in a conventional database. This allows queries for rapid simultaneous access to many pieces of metadata.

Semantic Voice Editing

From our prior RichReview deployment study [6], the discussant sometimes had to re-record an entire voice interaction to fix small mistakes in the middle of the recording, such as a stutter or a long pause. Recently, Rubin et al. presented an audio editing system that leverages speech transcriptions as semantic guidelines for editing voice as if it is text [4]. We are following this semantic editing approach, focusing on design and development of live editing features, such as partial deletion or insertion. Even after having the caption-based editing system, it is a tedious repetitive task to manually trim long-pauses and 'Um's [5]. We will solve this problem by providing a batch post-processing operation that can automatically delete several such unnecessary snippets at once.

Profile Based Group Formation

We would like to explore how different peer group formations influence performances of student discussion. In MOOCs, it is an open research question to assign a large pool of students into the group composition that maximizes overall group heterogeneity or homogeneity. We will approach this problem by modeling student characteristic based on a combination of (1) the readily available profile data, such as linguistic, cultural, geographic distance, and age, and (2) the profile data inferred from a student's past RichReview discussions, such as communication styles and topic of interests. We are developing a system which minimizes or maximizes overall distance between students on the axes of those characteristic dimensions.

ACKNOWLEDGMENT

We would like to gratefully acknowledge the support from the Kwanjeong educational foundation, the course staffs who have supported the development of RichReview, as well as the students who participate by providing each other with thoughtful, meaningful feedback.

REFERENCES

1. Chi, M. T., and Wylie, R. The icap framework: Linking cognitive engagement to active learning outcomes. *Educational Psychologist 49*, 4 (2014), 219–243.
2. Mitros, P., Affidi, K., Sussman, G., Terman, C., White, J., Fischer, L., and Agarwal, A. Teaching electronic circuits online: Lessons from MITx's 6.002x on edX. In *ISCAS*, IEEE (2013), 2763–2766.
3. Mitros, P., and Paruchuri, V. An integrated framework for the grading of freeform responses. *The Sixth Conference of MIT's Learning International Networks Consortium* (2013).
4. Rubin, S., Berthouzoz, F., Mysore, G. J., Li, W., and Agrawala, M. Content-based tools for editing audio stories. In *Proceedings of the 26th annual ACM symposium on User interface software and technology*, ACM (2013), 113–122.
5. Yoon, D., Chen, N., Guimbretière, F., and Sellen, A. Richreview: blending ink, speech, and gesture to support collaborative document review. In *Proceedings of the 27th annual ACM symposium on User interface software and technology*, ACM (2014), 481–490.
6. Yoon, D., Chen, N., Randles, B., Cheatle, A., Loeckenhoff, C. E., Jackson, S. J., Sellen, A., and Guimbretière, F. Deployment of a collaborative multi-modal annotation system for instructor feedback and peer discussion. In *Proceedings of the 19th ACM Conference on Computer Supported Cooperative Work & Social Computing*, ACM (2016).