2014 IEEE 14th International Conference on Advanced Learning Technologies

Creating Educational Resources at Scale

Piotr Mitros
edX
piotr@mitros.org

Felix Sun
MIT
felixsun@mit.edu

***Abstract*—Rising college tuitions are putting increasing financial pressure on families, while at the same time, teachers are repeating tasks like grading and problem creation thousands of times across campuses. Open educational resources (OERs) have the promise of reducing this inefficiency, offering higher quality education at lower cost. However, the adoption of OERs remains limited in part due to issues with organization, quality, search, and fit to the curriculum of individual classrooms. This paper presents a new format for OERs based on a common course skeleton shared among a community of students and instructors. That community jointly creates and improves educational resources around that skeleton. This model resolves many of the classical problems in finding and organizing OERs. We present ways to use such communities to create richer educational resources, such as intelligent tutoring systems (ITS).**

I. Introduction and Prior Work

There are 4,706 degree-granting institutions in the United States. An introductory course such as physics is taken by over one million students annually. This gives 3-4 orders of magnitude repetition and inefficiency on tasks such as creation of lectures and assessments, and 6 orders of magnitude on per-student tasks such as grading or tutoring. International numbers are approximately an order of magnitude greater. Open educational resources (OER) and at-scale learning organizations such as edX attempt to improve the quality of education by reducing those inefficiencies, providing higher-quality resources, and allowing more time for student-instructor interaction.

The OER approach has been limited in impact due to lack of coherence. With a few exceptions, OERs are spread among many repositories, and built on different conventions, syllabi, and technologies. As a result, finding and adapting OERs to specific courses is time-prohibitive for most faculty [5].

In contrast, institutional Massive Open On-line Courses (xMOOCs) [14] use a centralized approach where an institution creates a complete, coherent course. Those courses are used in blended classrooms across many campuses. xMOOCs are typically taught by top instructors, and many employ research-based pedagogies such as active learning, constructive learning, and mastery-learning [10]. They are further enhanced with data-driven techniques. The centralization allows for a greater investment of resources per course than traditional courses. Initial evidence suggests that well-designed xMOOCs can lead to high levels of student learning and satisfaction in both on-line and blended settings [8][10]. In contrast to OERs, xMOOCs are traditionally not open, and do not substantially leverage the creative input of external contributors.

In this paper, we present a hybrid model: the distributed course. In this model, a community of students and instructors collaborates around the creation of a common, shared course, and uses variants of that course across many classrooms. Since the resources are organized around a common curriculum, the barriers to use and contribution are much lower than traditional OERs. This model is analogous to crowdsourced models such as Wikipedia [6] or the Linux kernel, and has analogues to the LON-CAPA Shared Resource Pool [7]. It builds closely on the design goals of the MITx and the Berkeley CourseSharing platforms, both of which had a stated goal of becoming a "github for courses," as well as on cMOOCs[9].

We present evidence that the student experience can be enhanced by such a model, that communities are willing to contribute to a distributed course, explore ways to leverage community to create rich educational experiences, and look at potential limitations. In particular, we focus on case studies for building point-of-need help systems.

II. Prerequisites for crowdsourced content

For the distributed course to work, students and instructors must be willing to contribute. We have data from several experiments which show that with minimal encouragement, a single-digit percent of students contribute. In the 6.002x wiki, 586 students edited the wiki – roughly 2% of the active students. Those students generated 270 articles through 4645 edits. A distribution of how those contributions were spread among students is shown in fig. 1. The CDF of contributors is approximately log-linear, with substantial benefit both from a few major contributors, and from many minor contributors. In a hinting experiment, 8% of students contributed hints. While these are small portions of the students, this percentage still gives hundreds of contributors per course – sufficient to create very rich resources. In addition, qualitatively, several instructors who used MOOCs residentially expressed an interest in contributing back to those courses.

More active contributors developed complex content, such as problem solutions at quality levels beyond what course staffs can create, major platform technology enhancements, and similar. The long tail gives diversity, which allows, for example, interventions for less common student issues.

For crowdsourced resources to be effective, the benefits of contributed content must outweigh the costs of inconsistency. To begin exploring this issue, in 6.002x, Circuits and Electronics – the first edX/MITx MOOC – 30% of the students received a version of the main learning sequences which mixed two styles. The main videos were created by Agarwal, and used Khan-style tablet capture. The test group was shown additional set of videos consisting of dialogues between Mitros and Sussman, with camera capture of writing on paper, sometimes supported by experiments. In many cases, there was unintentional overlap between the Agarwal videos

978-1-4799-4038-7/14 $31.00 © 2014 IEEE
DOI 10.1109/ICALT.2014.15

IEEE computer society

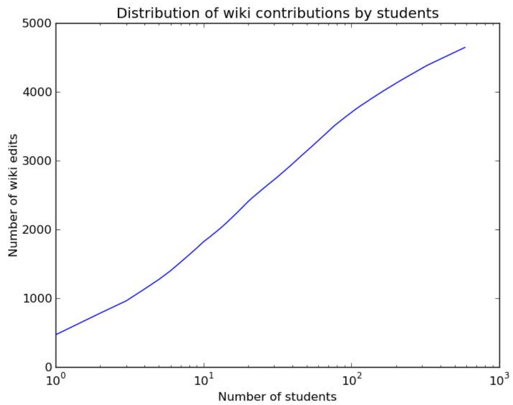

Fig. 1. A distribution of number of contributions made by students to the 6.002x wiki. Students are sorted from highest-contributing to lowest-contributing on the X axis. The Y axis shows the cumulative number of edits.

	Consistent	Inconsistent	Significant
Completion Rate	5.528%	5.558%	No
Grade (out of 4)	3.510	3.522	No
Rating (overall)	6.286	6.316	Yes
Rating (sequences)	6.220	6.375	Yes

TABLE I. 6.002X STUDENTS PREFERRED LEARNING SEQUENCES WITH MIXED STYLES. THERE WAS NO STATISTICALLY SIGNIFICANT EFFECT ON GRADES OR ON COURSE COMPLETION.

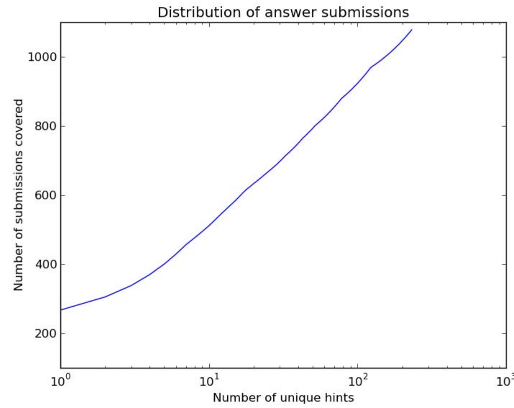

Fig. 2. A distribution of attempts on an exercise in mechanical engineering. The attempts are sorted on the X axis from most popular to least popular (submissions within 1% treated as equivalent). The Y axis shows the cumulative number of submissions. Practically, the X axis shows the number of hints needed to cover the number of submissions on the Y axis.

Make sure to use h=0.02 m, not h=0.2 m
Be carefully with units.
Check substitution of values.
units
Check your integrations for beam curvature.

Fig. 3. Five examplar hints chosen by a random number generator from the hints contributed by 2.01x MOOC students.

and the Mitros/Sussman videos, especially towards the end of the course. As shown in table I, students *preferred* learning sequences with additional content with mixed styles. This experiment gives preliminary evidence that conflicting styles may not detract from the learning experience. Whether this extrapolates to courses with many styles remains to be seen.

III. Personalized interventions

Providing immediate, specific, personalized feedback to students has been shown to significantly improve learning in both in-person [1] and technology-enhanced instruction [15], but is cost-prohibitive to create for thousands of university courses. Systems provide help either before the student reaches an answer ("How do I get started?") or after an incorrect answer ("What did I do wrong?"). It is easier to diagnose a student's misunderstanding after an attempt, so many ITS target incorrect answers. Such systems typically have hints for around half of student submissions [12]. Generating hints for the remaining half is extremely expensive. Fig. 2 shows the answer distribution to a typical numerical problem from 2.01x, an edX course in mechanical engineering. The CDF is approximately log-linear. While the first 50% of most common submissions can be covered with just a dozen hints, covering the top 95% of submissions would require hundreds of hints. To generate those hints, an expert would need to reverse-engineer the exact error the student made from each response.

We developed a system which permits students who answer a question incorrectly and later correctly to contribute a hint for that wrong answer. If a hint already exists, students have the option to either contribute a new hint, or vote on an existing hint. In a test problem in 2.01x, 619 students viewed the problem. Ignoring blank and junk submissions, there were 1078 attempts of which 223 were unique (numbers binned within 1%). Students contributed 51 hints. Virtually all hints were high quality, as evaluated by the course staff, with the exception of spelling and grammar errors (and one hint in Spanish), likely due to international demographics. A random subset is shown in fig. 3.

For help prior to answer submission, following the example of past MOOCs [13], we provided support through an on-line question-and-answer forums. Properly designed and managed, the effectiveness of such a forums can be very high. In the first run of 6.002x, 92% of questions received replies with a median time of 12 minutes. The staff felt the quality of the responses typically exceeded that of staff-provided help.

The primary role of the forum, however, was not for the benefit of the student asking the question, but for future students with the same question. The 13,000 threads formed a rich body of crowd-sourced content addressing a majority of the common student misconceptions, errors, and mental states. Only a minority of students asked questions, and students were 290 times as likely to read a thread as to create one.

We are investigating ways in which expert instructors can generate interventions for more common errors. Substantial research suggests that captures of expert-novice interactions may be an effective tool for addressing student misconceptions [3]. We ran a series of experiments in capturing expert-novice interactions in classrooms. We found that instructors were comfortable using a custom interface for helping students where writing and audio were captured on Microsoft Surface tablets. Preliminary experiments suggest that, if adopted across a number of classrooms, this may be an effective way to create a set of in-depth help resources.

IV. Limitations on Student Contributions

There are limitations to novices' ability to contribute. We tried a range of experiments where 6.002x students would tag content by learning objectives. Unsurprisingly [2], the quality of the tagging on the first pass was poor enough to be useless. Students tagged content based on superficial features, such as "circuit with a diode," rather than on the fundamental objectives, such as "analysis of nonlinear memoryless networks".

The original run of 6.002x included an experiment to see how well student crowd-sourcing could be used to generate a set of course notes, with the intention of eventually leading to an open-access textbook. We modified our wiki to allow for the creation, editing, and display of course-specific content such as in-line circuit schematics. We seeded the platform with structure as a repository for course notes, as well as some preliminary content. Students generated very high quality content, including course notes containing key equations, tables, summaries, derivation, and other information-at-a-glance. However, it was unclear whether these could evolve into a textbook – as novices, the students appeared to lack the ability to structure the text around key abstract concepts. The text was also not always at a level appropriate for novice learners.

This experience has parallels in peer grading. The edX platform includes a system which integrates AI, peer grading, self-assessment, and instructor grading for the grading of free-form responses such as essays and short answers [11]. As the system is used across increasing numbers of classrooms, grading from many instructors and potentially students may improve the quality of the machine learning. Experience [16] suggests that peer grading and self assessment can be quite accurate given a clear, well-defined rubric and instructions.

V. Conclusion

The distributed course offer an alternative model for creation and improvement of open educational resources. Since the resources are organized around a curriculum, the cost of finding resources is dramatically reduced. Placing all resources within a common course further guarantees that the resources follow common conventions (units, i vs j, etc.), use a common technology, and fit to the course.

By focusing a greater number of people on a smaller resource pool, the investment per resource can be higher. Since the resources are used by a large number of students in a common platform, data can be used to analyze and improve the quality of the resources, either explicitly with techniques from educational data mining, or implicitly, such as contributing help based on wrong answers and questions from students.

This approach allows us to bring techniques for improving the quality which would otherwise be prohibitively resource-intensive. ITS can lead to substantial learning gains [15], but require more complex content, such as hints for common student errors, problem banks [4], and individualized pathways. We have demonstrated that a community of students is able to construct such resources.

While the community of contributors was small as an percentage of participants, it was very large as an absolute number. The quality of the effort was exceptional. Since many of the contributors were novices, the efforts required appropriate structural and technological support, including clear instructions and structure, and a reasoned division of which types of contributions come from students, instructors, and other sources.

Even at this early stage, several of the frameworks described allow us to provide students with experiences surpassing those of traditional residential education on several axes. We were able to provide students with rich course notes, superior homework solutions, and multiple means of 24/7 point-of-need help. The next step will be to organize a distributed group of instructors around a course, and try to bring in a larger set of expert contributions.

References

[1] B Bloom. The 2 sigma problem: The search for methods of group instruction as effective as one-to-one tutoring. *Educational Researcher*, 13(6):4–16, 1984.

[2] M Chi, P Feltovich, and R Glaser. Categorization and representation of physics problems by experts and novices. *Cognitive science*, 5(2):121–152, 1981.

[3] S Craig, M Chi, and K VanLehn. Improving classroom learning by collaboratively observing human tutoring videos while problem solving. *Journal of Educational Psychology*, 101(4):779–789, 2009.

[4] S Embretson and S Reise. *Item response theory*. Psychology Press, 2000.

[5] D Harley, A Head, D Nasatir, S Paletz, J Henke, I Miller, X Sheng, and J Guo. Understanding the use of digital collections in undergraduate humanities and social science education: Status report, preliminary analyses, and next steps. *CSHE Digital Resource Study, Center for Studies in Higher Education, University of California at Berkeley*, 2004.

[6] A Kittur, E Chi, B Pendleton, B Suh, and T Mytkowicz. Power of the few vs. wisdom of the crowd: Wikipedia and the rise of the bourgeoisie. *World Wide Web*, 1(2):19, 2007.

[7] G Kortemeyer and W Bauer. Multimedia collaborative content creation (mc3) – the MSU lecture online system. *Frontiers in Education Conference*, 1998.

[8] T Lewin. Colleges adapt online courses to ease burden. *New York Times*, April 2013.

[9] A McAuley, B Stewart, G Siemens, and D Cormier. The MOOC model for digital practice. *Online:¡ http://davecormier.com/edblog/wp-content/uploads/MOOC_Final.pdf¿[accessed 17/04/11]*, 2010.

[10] P Mitros, K Affidi, G Sussman, C Terman, J White, L Fischer, and A Agarwal. Teaching electronic circuits online: Lessons from MITx's 6.002x on edX. In *ISCAS*, pages 2763–2766. IEEE, 2013.

[11] P Mitros and V Paruchuri. An integrated framework for the grading of freeform responses. *The Sixth Conference of MIT's Learning International Networks Consortium*, 2013.

[12] D Pritchard and E Morote. Reliable assessment with CyberTutor, a web-based homework tutor. In *World Conference on E-learning in Corporate, Government, HealthCare, and Higher Education*, 2002.

[13] The AI and CS Learning Community. "http://www.aiqus.com/". Accessed: 2013-10-07.

[14] Introduction to artificial intelligence. "http://www.ai-class.org/". Accessed: 2011-11-15.

[15] K VanLehn. The relative effectiveness of human tutoring, intelligent tutoring systems, and other tutoring systems. *Educational Psychologist*, 46(4):197–221, 2011.

[16] M Walvoord, M Hoefnagels, D Gaffin, M Chumchal, and D Long. An analysis of calibrated peer review (cpr) in a science lecture classroom. *Journal of College Science Teaching*, 37(4):66, 2008.