UW MSR SUMMER INSTITUTE 2012

CROWDSOURCING PERSONALIZED ONLINE EDUCATION

ABSTRACTS

Sumit Basu and **Lucy Vanderwende**, *Teaching the Web: Learning How to Use Arbitrary Text as Teaching Material*

While there are a variety of efforts devoted to developing online courses and tutoring systems around existing curricula, the diverse set of possible learning topics is growing exponentially, and many of these do not have the advantage of a formal class or curriculum. This is particularly the case when one considers the educational needs of self-motivated learners who wish to master material for a new job, an unfamiliar task (e.g., buying a home), or a personal goal (e.g., learning about the history of the region one is traveling to). We propose to investigate how we might use the web and other unstructured sources as teaching material -- given a topic a user wants to learn about, can we collate appropriate material and deliver a learning experience that will help them master that topic? This involves a variety of research problems, from assembling an appropriate curriculum to helping them incrementally improve their coverage of the material. At the core is what we call the mastery loop, a cycle of material presentation, assessment, and adaptation, whose goal is to get the user to a place of mastery over the material. This is where the crux of our work has been thus far, and in this talk we will describe some of the technical problems and findings in this area, including our recently presented work on quiz generation and preliminary results on grading. Crowdsourcing is key to making this all work, not only in terms of data to help with learning how to do the tasks, but also for grading, usage data, and feedback that can lead to continuous improvement. We will close with an overview of the many open problems remaining in this space.

Jeff Bigham, *Accessible Education through Real-Time Crowd-Powered Systems*

11% of post-secondary students have a disability, yet accommodations remain expensive and limited in availability. A goal of crowdsourced education is to make education available to everyone, and leveraging the crowd to make education accessible is our best chance for ensuring that everyone can participate. In this talk, I'll overview three crowd-powered systems we have created that help make education more accessible: VizWiz provides real-time visual information for visually impaired students, Scribe provides real-time captions for deaf and hard of hearing students, and Legion/Multiverse provide interactive support to help overcome accessibility problems in existing interfaces. Collectively, these tools demonstrate the potential of the crowd to allow more people to participate in education.

Emma Brunskill, *Student Variability and Automated Instructional Policies*
Computer tutors can scale teaching that is customized to each individual student. But how should this customization work? What characteristics of a student are important for deciding how, what, and when the system teaches? Many existing tutoring systems adapt instruction to the stage of the student, but assume the rate of student progress is identical. Yet our results on a simple math domain suggest there is substantial variability across the dynamics of student learning, and, more importantly, that this variability has a significant impact on what is the best instructional policy for different students. Our work indicates that systems that explicitly learn and model student dynamics may lead to significantly more effective tutors. I'll also discuss how this work relates to our very preliminary work towards self-optimizing tutors that become better teachers as they interact with more students.

Stephanie Chang, *Personalized Learning at Khan Academy*
In this talk, I will share how Khan Academy began using machine learning to evaluate learner proficiency along with some lessons learned along the way. Measuring proficiency accurately is important for helping us advance learners appropriately. This way, proficient users spend less time on material they already know, while non-proficient users spend more time practicing unfamiliar material. We are able to do this by leveraging data from over 50,000 daily active users who answer randomly-generated problems on the Khan Academy site.

Micki Chi, *Two Approaches to Enhancing Online Learning: Dialogue Videos and Engagement Activities*
We know that having a teacher/instructor present a lecture in a monologue way (either in person or online) is not an effective form of instruction for learning. We present two alternative approaches that would enhance online learning far beyond a monolog lecture-style. One approach manipulates how content information is conveyed that is neither costly nor 1-1 adaptive, by using a video of tutorial dialogues. A second approach manipulates what students do in terms of engagement activities while a monologue lecture is given. Both approaches can also be combined.

Ed Dieterle and **Emily Dalton Smith**, *It is All About the Outputs*
A grand challenge for U.S. education is to support personalized learning for 26+ million secondary students with varying needs, skills, developmental levels, interests, and dispositions. Central to this challenge is our ability to continuously capture, derive meaning from, and act on data generated from activity within digital learning tools. We are pursuing this challenge by focusing intensely on the outputs generated from such activities and what they can explain about students' mastery of academic content and skill. In this interactive presentation, we discuss and seek feedback on three priority areas of investment: (a) building the field of education data scientists; (b) supporting the Shared Learning Collaborative and its development of the Shared Learning Infrastructure; and (c) supporting the development and investigation of digitally mediated social learning and crowdsourcing for learning and teaching.

Mira Dontcheva, *Learning how to use complex software with interactive tutorials and games*
I would like to talk about two projects: Tutorial Player and Level Up. Tutorial Player is an interactive iPad application that connects to your desktop computer and interacts with desktop software. It follows what you do in Photoshop and shows you how to accomplish steps when you are stuck. Tutorial Player has been available in the iPad app store since May 1st. We are working on extending the platform so that anyone can publish to the Tutorial Player platform and share interactive tutorial content. Level Up for Photoshop explores games as a mechanism for learning complex software. We have found that it is useful for both novices and advanced users. We are now

going through a second deployment and experimenting with turning Level Up into a game with a purpose.

Sumit Gulwani, *Intelligent Tutoring Systems: A Logical Perspective*
The need for use of cross-disciplinary technology in reinventing education cannot be understated. In this talk, I will describe the pivotal role that programming language and logical reasoning techniques can play in building intelligent tutoring systems that can support automated grading, solution/hint generation, and problem generation/recommendation. I will illustrate these capabilities for a variety of subject domains including geometry, algebra, programming, and language learning. I will also point out synergistic opportunities with use of statistical techniques that can leverage wisdom of the crowd.

Bjoern Hartmann, *Feedback and Snacks: Two Experiments in Crowd Education*
In this talk, I will present two approaches to educate the crowd, and to use the crowd for assessment in education. Our work on the Shepherd system demonstrates how timely feedback and self-assessment can lead to better work and higher worker perseverance. I will then discuss an alternative mechanism for crowdsourcing tasks that require specialized knowledge or skill from workers: communitysourcing – the use of physical kiosks to elicit work from specific populations. We apply communitysourcing to the problem of grading Computer Science exams.

Panos Ipeirotis, *What should I learn next? Discovering Economically-Rewarding Education Paths*
I will present some early experiments on suggesting types of projects that an online worker will find most economically rewarding to follow. Furthermore, based on an analysis of the skill sets of the workers, we describe how to recommend new skills to learn that will offer the highest "return on investment" in terms of effort to learn the skill and expected economic rewards in the marketplace.

David Karger, *The Textbook of the Future: Collaborative Annotation to Collaborative Authoring*
I'll start by talking about NB, our tool that situates class discussion forums in the margins of the documents---lecture notes and draft textbooks---that those forums are discussing. I'll argue that these "situated forums" preserve context, helping students to benefit from and contribute to the discussions about what they are reading at each moment, and that they provide valuable signals to the faculty about trouble spots in the text. I'll then broaden to ask some questions about the textbook of the future. How do we balance the goal of a perfectly clear teaching text against the desire to have students learn by struggling with and discussing the content? How do we combine the goal of wiki-style "peer production" of knowledge and content with the teacher's special authoritative role? What can we learn by instrumenting the textbooks that students are reading online? Why are there so few hypertextbooks? And most important, what can I do to ensure success when I make my class write its own textbook this fall?

Scott Klemmer, *Scaling studio critique: success, bruises, and future directions*
Creative fields like design are burgeoning. So are project-based learning approaches. However, providing feedback and assessment of design and other creative work is extremely time consuming -- this bottleneck is the major capacity constraint for scaling peer assessment. Given that assessment and feedback are central to learning, how might we address this?

This spring, we collaborated with Coursera to launch the first massive-scale class with self and peer assessment. We leveraged the self-assessment materials (like grading rubrics) that I've created and refined for my intro HCI class over the past several years. For peer assessment, we built on a technique

called calibrated peer assessment where students learn grading rubrics through training examples and then grade peers. In our first experimental class, http://hci-class.org, students anonymously graded 5 of their peers' assignments -- and then their own. My experience has been that students get a lot out of assessing their own work. And while there's a bit of gaming the system (which often usefully attune students), there's a lot of earnest and frank reflection and assessment that I think it tremendously valuable -- and increases students maturity.

Students also do remarkably well. For my in-person class, about 70% of students self-assessed within 3% of the staff assessment! Online, our preliminary analysis is that 40% of the 1800 students in the studio track self-assess within 5% of the median peer assessment. In this talk, I will describe the techniques we're using, our surprising successes (at least to me) and many bruises in deploying self and peer assessment, and what I see as exciting opportunities for future work.

Ken Koedinger, *Crowdsourcing Cognitive Models for Assessment, Tutoring, and In-Game Support*
Cognitive Tutors personalize online education by adapting problem-based instruction both within problems, by providing feedback and instruction specific to students' solution progress, and between problems, by selecting and pacing activities based on students' performance history. The approach relies on a cognitive model of the unobservable mental skills and concepts needed for success across problem-solving tasks. Widespread use of tutors (and educational games and simulations too) is producing vast data that can be used to create and improve cognitive models to drive better adaptive instruction. LearnLab's DataShop(http://learnlab.org/datashop) is the world's largest open repository of such data. I will discuss educational data mining algorithms to create, test, and improve cognitive models for better personalized assessment and tutoring within educational technologies. Large crowds of students and, more recently, small crowds of learning researchers are being leveraged in these approaches. In a recent study, we employed our Learning Factors Analysis algorithm on eleven different DataShop datasets to automatically discover better cognitive models in all cases.

Mausam, *Towards an AI Agent for Online Education*
Recent years has seen a splurge of research activity in enabling crowdsourcing platforms, such as Mechanical Turk and GalaxyZoo. We have recently proposed Clowder, an integrated AI agent, which performs worker tracking, task-level quality control and task allocation on Mechanical Turk and other micro-crowdsourcing platforms. In this talk, I argue that crowdsourcing and online education require analogous techniques to be successful at large scales. For example, tracking a worker's ability is similar to tracking a student's progress. Allocating a task to a worker is akin to allocating reading material to a student and so on. Moreover, AI-based control is extremely important in online education, be it for assignment of tasks to students for peer-evaluation or for automatically personalizing education to each student. Along the lines of Clowder, I lay out the design of an AI agent that can enormously boost the success of online education platforms.

Rob Miller, *Crowdsourced Code Review in Programming Classes*
Caesar is a distributed, social code reviewing tool specifically designed for use in a programming course. Caesar is capable of scaling to a large and diverse reviewer crowd, provides automated tools for increasing reviewer efficiency, and implements a social web interface for reviewing that encourages discussion and participation. Our system is implemented in three loosely-coupled components: a language-specific code preprocessor that partitions code into small pieces, filters out uninteresting ones, runs static analysis, and detects clusters of similar code; an incremental task router that dynamically assigns reviewers to code; and a language-agnostic web interface for reviewing code. We will discuss the experience of using Caesar in two semesters of a software

engineering course at MIT, drawing on a mixed crowd of reviewers that included peer students, course alumni, and teaching staff.

Andrew Ng, *The Online Revolution: High-Quality Education for Everyone*
Last year, Stanford University offered three online courses, which anyone in the world could enroll in and take for free. Students were expected to submit homeworks, meet deadlines, and were awarded a "Statement of Accomplishment" only if they met our high grading bar. Together, these three courses had enrollments of around 350,000 students, making this one of the largest experiments in online education ever performed. In the past few months, we have transitioned this effort into a new venture, Coursera, a social entrepeneurship company that partners with top universities to provide high-quality content to everyone around the world for free. In this talk, I'll report on this new experiment in education, and why we believe this model can provide both an improved classroom experience for our on-campus students, via a flipped classroom model, as well as a meaningful learning experience for the millions of students around the world who would otherwise never have access to education of this quality. I'll describe the pedagogical foundations for this type of teaching, and the key technological ideas that support them, including easy-to-create video chunks, a scalable online Q&A forum where students can get their questions answered quickly, sophisticated autograded homeworks, and a carefully designed peer grading pipeline that supports the at-scale grading of more open-ended homeworks, such as essay questions, derivations, or business plans.

Whereas technology and automation have made almost all segments of our economy – such as agriculture, energy, manufacturing, transportation---vastly more efficient, education today isn't much different than it was 300 years ago. Given also the rising costs of higher education, the hyper-competitive nature of college admissions, and the lack of access to a high quality education, we think there is a huge opportunity to use modern internet and AI technology to inexpensively offer a high quality education online. Through such technology, we envision millions of people gaining access to the world-leading education that has so far been available only to a tiny few, and using this education to improve their lives, the lives of their families, and the communities they live in.

Zoran Popović, *Engaged Learning: Discovering New Learning Pathways*
This talk will focus on the application of the most powerful question that arose from the Foldit experiment: is possible to achieve general world-class expertise in a specific domain through deliberate prolonged game play. I will discuss our efforts try to generalize this finding and apply it to education in general. Specifically, we focus on turning education enterprise into a data driven science through use of interactive artifacts, including but not limited to games, that optimally evolve towards each individual student while optimizing a dual objective function of prolonged engagement, and transferable learning. I will describe some of the machine learning, visual data mining, procedural content-generation problems that arise from this framework and the impact this technology could have on millions of student worldwide.

Luis Van Ahn, *DuoLingo: Learn a Language for Free while Helping to Translate the Web*
I want to translate the Web into every major language: every webpage, every video, and, yes, even Justin Bieber's tweets.

With its content split up into hundreds of languages – and with over 50% of it in English – most of the Web is inaccessible to most people in the world. This problem is pressing, now more than ever, with millions of people from China, Russia, Latin America and other quickly developing regions entering the Web. In this talk, I introduce my new project, called Duolingo, which aims at breaking this language barrier, and thus making the Web truly "world wide."

We have all seen how systems such as Google Translate are improving every day at translating the gist of things written in other languages. Unfortunately, they are not yet accurate enough for my purpose: Even when what they spit out is intelligible, it's so badly written that I can't read more than a few lines before getting a headache.

With Duolingo, our goal is to encourage people, like you and me, to translate the Web into their native languages.

Kurt Van Lehn, *Toward socially intelligent tutoring systems: Of the crowd, for the crowd*
After a brief orientation that distinguishes tutoring systems from other educational software, and the community of tutoring system researchers from other communities, we will dive into the major types of tutoring behavior and emerge with two reliable sources of effectiveness, namely step-based hinting and adaptive task selection, plus some new promising sources of effectiveness. Next is a peek under the hood at the three major technologies of tutoring, which leads to identification of a third reliable source of effectiveness: content re-engineering. All this good news leads naturally to the question, "If tutoring systems really are as effective as human tutors and will probably soon be even more effective than human tutors, then why isn't my kid's school using them?" At this point, we leave data behind and begin to speculate. I will claim that the bottleneck is *not* the publishing industry, nor the school districts, nor the IT departments of schools, nor the robustness of technology nor even the technophobia of teachers which is rapidly disappearing anyway. The bottleneck is that current tutoring systems disrupt an ordinary classroom's social structures and processes. I will sketch a vision of changes to both classrooms and tutoring systems (hint: it involves crowdsourcing) that will allow our kids to reap the benefits of the three+ sources of effectiveness identified earlier.

Jennifer Widom, *From 100 Students to 100,000*
I'll describe my recent experience teaching introductory databases to 60,000 students. Admittedly only 25,000 of them submitted their homework, and a mere 6500 achieved a strong final score. But even with 6500 students, I more than quadrupled the total number of students I've taught in my entire 18-year academic career. I began by "flipping" the way I teach my Stanford course and, as a side-effect, making all components of the course freely available online. But the big inflection point came when I offered the online course in a structured fashion with a schedule, automatically-graded assignments and exams, and most importantly a worldwide community of students. I'll cover a variety of topics related to the massive online course, both logistical and social, while avoiding speculation on the future of higher education.

Haoqi Zhang, *Towards Active Learning in Online Education: Ideas from the Classroom and Crowdsourcing Complex Tasks*
Active learning refers to instructional strategies that aim to improve student learning by engaging students in 'doing.' In this talk, I ask 'how can we construct online education environments in which active learning is a core component of the learning experience?' Focusing on collaborative interactions, I explore examples in which small and large groups of people work together to solve problems. I start in the classroom, and describe our experience designing a linear optimization course that promotes doing via breakout modeling sessions, active sections, and extreme optimization assignments. Moving online, I describe Mobi, a system for crowdsourcing itinerary planning. Mobi illustrates a novel *crowdware* concept by using a shared, collaborative workspace to tackle complex tasks that are difficult to decompose. I synthesize how ideas from these examples can help shape new online learning experiences, and note the challenges that remain.