Xu Wang

Research Statement

I conduct interdisciplinary research within the fields of Human-Computer Interaction, Cognitive Science, Artificial Intelligence and Education. My research goals are two-fold. 1) I contribute to scientific understanding of the nature of complex problem solving and learning. 2) Based on the insights, I develop content authoring techniques and intelligent tools to support teaching and learning.

In my doctoral work, I explored two threads of computational methods to build educational experience for scale and quality. Increasing numbers of people are seeking higher education. A challenge to meet the growing demand is to scale such education programs while maintaining their quality. My work tackles this challenge by leveraging the complementary strengths of peers, AI and teachers.

Traditionally, teachers take the most responsibility in providing learning opportunities to students, e.g., giving lectures, offering feedback. However, the efforts required from experts makes such learning opportunities less scalable. On the other hand, AI-based technologies have begun to tackle this scaling issue, such as automatic grading (Arrow #1 in Figure 1), intelligent tutoring systems that adaptively select questions based on students’ knowledge (Arrow #2 in Figure 1). But challenges remain. First, it requires lots of expertise and time for teachers to design high quality questions. I explore how we can take advantage of AI and the open-ended data past students produced to augment teachers’ capabilities on content creation (Arrow #3 in Figure 1). In this thread, I develop UpGrade [1], which leverages past students’ written solutions to semi-automatically create good quality multiple-choice questions. In contrast to experts’ beliefs, I find multiple-choice questions are at least equally beneficial for learning as open-ended questions in a variety of domains [2]. In another context, I develop a workflow recommender [3] to support end-users in using complex software applications. Second, online courses often only provide solitary learning experiences, whereas implementing collaborative team-based learning has been challenging. I explore ways AI could scaffold discussion online through signaling learning moments (Arrow #4 in Figure 1). I perform analytical studies of people’s discussion behaviors [5, 6] and developed interventions to facilitate in-depth conversations during team project discussions [4, 7].

On the one hand, my work demonstrates that advances in computer science have promise in improving education, e.g., natural language understanding techniques can be leveraged for automatic content creation and scalable teaching [1,2,4]. On the other hand, I seek to explore how advances in learning technologies could contribute to computer science. For instances, research on human intelligence could inform the design of machine intelligence; Human-AI hybrid methods that can make an impact in practice contribute knowledge around how to better structure and engineer training data as input to algorithms and how to leverage the complementary strengths of human and AI to tackle real world challenges.

Content Creation through Human-AI Partnership

Getting deliberate practice is critical in developing skill mastery. However, providing quality support to learners requires a tremendous amount of time and effort from instructors or other content providers. In college classrooms, for example, it is challenging for instructors to design sufficient high quality learning materials, including formative and summative assessments. In addition to the effort in designing learning materials, substantial...
grading effort is required when open-ended assignments are given, as they frequently are. From the students’ perspective, such open-ended assignments do not always provide sufficient support for novices, and tend not to facilitate deliberate practice with immediate feedback. Techniques that could both magnify the investment of instructor effort while also improving student learning are highly desirable for scaling quality education.

I develop UpGrade [1], a system to intelligently support instructors in creating scalable and quality learning opportunities. UpGrade takes past student open-ended solutions and past instructor feedback as input, and supports instructors in quickly creating quality multiple-choice questions. UpGrade first segments the written solutions from past students into components based on the assignment template. The instructor then specifies a question creation schema, identifying which components in the source will be used in the target question. The system reorganizes the source data components based on the schema to create multiple-choice questions. One critical step is selecting distractors. UpGrade leverages natural language understanding techniques (e.g., text similarity matching) to select plausible distractors from the data source. A final optional step is for the instructor to approve the questions. Two example questions produced by UpGrade on HCI research methods is shown in Figure 2.

We evaluated the quality of UpGrade-produced questions in a two-week classroom experiment in an HCI course. Students either learn about survey design and heuristic evaluation through UpGrade-produced questions or through traditionally used open-ended assignments. We demonstrate that 1) students learnt as much from UpGrade-produced multiple-choice questions as traditional open-ended assignments, 2) UpGrade reduces assignment completion time by ~30% and removes the need for instructor grading.

As I apply UpGrade in more courses, I observed an instructor preference to use open-ended assignments. A survey with 22 HCI professors revealed a belief that multiple-choice questions are less valuable than open-ended questions, and thus, using them entails making a quality sacrifice in order to achieve scalability. On the other hand, a study with 178 students in two college HCI research methods classes contradicted the instructor belief, suggesting that data-driven multiple-choice questions are equally hard as matched open-ended questions, and in “evaluation-heavy” domains, they assess, and exercise during practice, the same difficult skills that are exercised in open-ended questions [2].

UpGrade has been used in 7 modules of 8 classes at Carnegie Mellon University to help instructors create multiple-choice questions. So far, more than 400 students have learnt with UpGrade. In this process, I identified areas for improvement. The success of UpGrade relies on the data input from past peers to be structured. The existence of instructor feedback in the data source could largely improve question quality. With these insights, I’m exploring approaches to increasing the quality of data input to UpGrade, e.g., support instructors in designing open-ended assignments to collect better examples; cluster student solutions to enable detailed instructor feedback.
I also develop techniques to support automatic content creation in another context: learning to use complex software applications. Modern complex software applications often include hundreds or thousands of commands, which further form a much larger number of workflows a user can use. This raises two issues. First, pre-designed tutorials cannot cover all the different workflows. Second, users may get stuck in inefficient or suboptimal ways of completing tasks if they are not aware that better workflows exist. Knowing the tasks that a user is working on, and the workflows they are using, is the first step towards providing better personalized software learning support. I develop a technique [3] in collaboration with Autodesk Research, which models a user’s current workflow in a CAD software application, e.g., rendering, parametric design, through a hierarchical approach.

A diagram summarizing our approach is shown in Figure 4. The overall idea is to first use topic modeling to segment user logs into mutually exclusive sets based on high-level tasks (Layer 1), and then to apply frequent pattern mining to each set, to identify finer-grained patterns of command usage (Layer 2). This modeling approach provides insights into both what a user is currently working on in a CAD software application, and what tasks demonstration videos in online repositories are focusing on. We then designed and implemented recommender algorithms to recommend community-generated videos containing relevant workflows to end-users. With our workflow recommender, community user-uploaded demonstration videos or existing videos in online repositories can be leveraged as sources of reference for other users who are working on similar tasks.

Support In-depth Discussion in Online Learning

Online learning experiences are mostly solitary, despite the demonstrable benefits of collaboration and social interaction for learning. Making online learning to be more collaborative has been challenging, with reasons such as learners come to courses with a wide range of purposes and levels of motivation, they are typically distributed and do not have existing relationships, and interactions are often large-scaled and asynchronous in nature. In my work, I investigate two types of peer interactions in MOOCs to seek ways to support them. First, I run analytical studies on MOOC forums to understand how students are using them. I find that although the forums are designed to be interactive, on-topic interactive discussions are fairly scarce [5]. On the other hand, students who participate in the forum in a meaningful way, e.g., through displaying their reasoning and extending others’ reasoning, learn more in the course [6]. The two studies were the first ones that run fine-grained behavior analyses in MOOC discussion forums and have been well cited since 2015.

Second, I develop interventions to support synchronous team-based learning in MOOCs. Online learners often come to courses with very different backgrounds and perspectives, supporting them to share perspectives and acknowledge each other’s ideas is critical for successful teamwork. I develop a task-relevant conversational agent to support in-depth discussions between team members [4]. The agent is powered by natural language understanding models and makes two predictions from each contribution a student has typed in, 1) whether the student
is in support of a plan based on keywords matching. 2) whether the student shows reasoning, using a logistic
regression classifier built with pre-annotated data. The agent provides three types of prompts 1) When a student
does not show reasoning towards a certain plan, the agent asks the student to elaborate. 2) When two students
propose different plans, the agent asks one of the two to compare. 3) When there's nobody talking for 3 minutes,
the agent will encourage the least contributor to talk. An excerpt of a conversation transcript is shown in Fig-
ure 4. Through an evaluation with 252 learners on Amazon Mechanical Turk, we show that the BazaarAgent
supports multi-perspective learning [4] and shows promise to be deployed in future team-based online courses.
This paper has been used in the curriculum of three graduate-level courses at CMU.

In addition, I collaborated with colleagues to develop a group formation technique that automatically assigns
students to groups based on whom they had the most in-depth conversations with during the community delib-
eration. We show that the technique is effective in detecting pairs or groups of students who are interested in each
others' ideas and thus supports them to produce higher quality team products later in the course [7,8].

Research Agenda

The long-term purpose of my research is to support more people get access to high quality educational resources,
and help people navigate through their career with more flexibility and skill development opportunities. The
Human-AI hybrid methods I have described is a powerful way to achieve this goal. My common approach for
research follows 1) I start by understanding what are the critical learning processes students need to engage in
to develop a complex skill. This investigation is often done through theorized analyses, small-scale case studies,
and focused experiments on student thinking processes. I also conduct research to understand teachers’ needs
and beliefs through contextual inquiry, interviews, and surveys. 2) I then develop technologies to support teach-
ing and learning. The technology solutions are often evaluated through a mixed-methods approach, combining
randomized control experiments in classrooms and follow-up interviews with students and teachers. An iterative
design and development process is often involved. 3) In the end, I explore how such systems can be deployed.
Based on my past work, here are some future directions that I hope to continue pursuing.

A New Generation of Learning Management Platforms. Growing out of UpGrade, I see a new generation of
Learning Management Systems where student data is better structured, archived and reused. Instructor effort is
optimized, and repetitive work is reduced to the minimal. Future work may explore supporting instructors break
down open-ended assignment questions using past answers and reducing the repetitive efforts required in writing
feedback by pre-clustering student solutions. Additionally, I will continue to develop the theory behind genera-
tion and evaluation. With more nuanced categorization of skills embedded in Learning Management Systems, I
envision they will support instructors make more accurate and nuanced instructional decisions and designs.

Other Applications of Human-AI Hybrid Methods. I will continue building teacher support systems for use in
different contexts. UpGrade leverages past students’ written work as source data for question creation. I plan to
explore other data sources to support automatic content creation, e.g., leveraging teacher-student conversation
during lectures. In my undergraduate thesis, I developed an Augmented Reality game for 8th graders to visualize
and interact with the micro-structures of molecules when they learn chemistry [12]. Despite the benefits, we
observed that most teachers did not have the at-ready capability to create sophisticated AR applications. In my
future work, I plan to develop AR authoring toolkits to support multimedia content creation.

Support Informal Learning and Workplace Learning. In my prior work, I explored technology interventions to
support professional training in two domains. First, I collaborated with the MIT Teaching Systems Lab on devel-
opling ELK: a role-playing simulation for pre-service teachers to practice strategies on eliciting learner knowledge.
We show that the practice-based simulation helps teachers cultivate empathy with students and leads to improve-
ments in their questioning strategies [9]. This approach can be applied in other domains where perspective taking
is critical. For example, educating the public about the fairness issues behind AI algorithms.
Second, I collaborated with colleagues on adapting an industry standard “Mob Programming” into an online training paradigm to encourage teams of programmers to reflect on concepts and share code when working on team projects [10]. Consistent with other work, we observed that explicit scaffold [4, 9, 10] can be more helpful than implicit scaffold [4, 9, 11] for learning good practices in programming and strategies for communication and giving feedback. I plan to apply the methods and designs (e.g., explicit scaffolding, interactive examples) in developing novel professional training programs. Examples may include computer-based programs to support novice UberLyft drivers read maps and navigate between locations.

**AR Enhancement of Physical Task Performance** Many online tutorials are about physical tasks, e.g., cooking, setting up new devices, doing makeup, etc. I plan to explore methods to build alignment between 2D online tutorials and 3D physical task performance, e.g., using online tutorials and people’s learning and behavioral traces in physical spaces as an augmented layer of information to the existing world.

**REFERENCES**


[9] Xu Wang, Meredith Thompson, Kexin Yang, Dan Roy, Kenneth Koedinger, Carolyn Rose, Justin Reich. (Under Review) Practice-Based Teacher Education with ELK: A Role-Playing Simulation for Eliciting Learner Knowledge.

