

## Human-Centered Clustering and Visualizations of Student Code for Personalized Feedback at Scale

Elena Glassman, University of California Berkeley

**Abstract:** In massive programming classes, a single problem may yield hundreds or thousands of student submissions. For example, UC Berkeley's Structure and Interpretation of Computer Programs now enrolls ~1500 students each fall semester. To serve these students, the teaching staff now includes ~250 teaching assistants, tutors, and lab assistants. It is difficult for any one teacher to understand and provide feedback on more than a small slice of student work. Students primarily get feedback through simple test-based autograders, but this feedback alone is insufficient for helping students learn how to write great code. In this talk, I describe several systems which cluster and visualize hundreds or thousands of student code submissions in a human-readable way, exposing common and uncommon algorithmic choices, syntactic style, variable names, and mistakes. Using these systems, teachers can deliver personalized feedback to thousands of current and future students.

**Elena Glassman** is an EECS postdoctoral researcher at the Berkeley Institute of Design, advised by Bjoern Hartmann and Marti Hearst. She earned her EECS PhD at MIT CSAIL in August 2016, where she created scalable systems that analyze, visualize, and provide insight into the code of thousands of programming students. Prior to entering the field of human-computer interaction, she earned her M.Eng. in the MIT CSAIL Robot Locomotion Group. She has been a visiting researcher at the Stanford Biomimetics and Dextrous Manipulation Lab and a summer research intern at both Google and Microsoft Research, working on systems that help people teach and learn. She was awarded the Intel Foundation Young Scientist Award, both the NSF and NDSEG graduate fellowships, the MIT EECS Oral Master's Thesis Presentation Award, a Best of CHI Honorable Mention, the MIT Amar Bose Teaching Fellowship for innovation in teaching methods, and the Moore/Sloan Data Science Fellow at the Berkeley Institute for Data Science (BIDS).