

## Essay-Grading Software Offers Professors a Break



Imagine taking a college exam, and, instead of handing in a blue book and getting a grade from a professor a few weeks later, clicking the "send" button when you are done and receiving a grade back instantly, your essay scored by a software program.

And then, instead of being done with that exam, imagine that the system would immediately let you rewrite the test to try to improve your grade.

EdX, the nonprofit enterprise founded by Harvard and the Massachusetts Institute of Technology to offer courses on the Internet, has just introduced such a system and will make its automated software available free on the Web to any institution that wants to use it. The software uses artificial intelligence to grade student essays and short written answers, freeing professors for other tasks.

The new service will bring the educational consortium into a growing conflict over the role of automation in education. Although automated grading systems for multiple-choice and true-false tests are now widespread, the use of artificial intelligence technology to grade essay answers has not yet received widespread endorsement by educators and has many critics.

Anant Agarwal, an electrical engineer who is president of EdX, predicted that the instant-grading software would be a useful pedagogical tool, enabling students to take tests and write essays over and over and improve the quality of their answers. He said the technology would offer distinct advantages over the traditional classroom system, where students often wait days or weeks for grades.

"There is a huge value in learning with instant feedback," Dr. Agarwal said. "Students are telling us they learn much better with instant feedback."

But skeptics say the automated system is no match for live teachers. One longtime critic, Les Perelman, has drawn national attention several times for putting together nonsense essays that have fooled software grading programs into giving high marks. He has also been highly critical of studies that purport to show that the software compares well to human graders.

"My first and greatest objection to the research is that they did not have any valid statistical test comparing the software directly to human graders," said Mr. Perelman, a retired director of writing and a current researcher at M.I.T.

He is among a group of educators who last month began circulating a petition opposing automated assessment software. The group, which calls itself Professionals Against Machine Scoring of Student Essays in High-Stakes Assessment, has collected nearly 2,000 signatures, including some from luminaries like Noam Chomsky.

"Let's face the realities of automatic essay scoring," the group's statement reads in part. "Computers cannot „read.“ They cannot measure the essentials of effective written communication: accuracy, reasoning, adequacy of evidence, good sense, ethical stance, convincing argument, meaningful organization, clarity, and veracity, among others."

But EdX expects its software to be adopted widely by schools and universities. EdX offers free online classes from Harvard, M.I.T. and the University of California, Berkeley; this fall, it will add classes from Wellesley, Georgetown and the University of Texas. In all, 12 universities participate in EdX, which offers certificates for course completion and has said that it plans to continue to expand next year, including adding international schools.

The EdX assessment tool requires human teachers, or graders, to first grade 100 essays or essay questions. The system then uses a variety of machine-learning techniques to train itself to be able to grade any number of essays or answers automatically and almost instantaneously.

The software will assign a grade depending on the scoring system created by the teacher, whether it is a letter grade or numerical rank. It will also provide general feedback, like telling a student whether an answer was on topic or not.

Dr. Agarwal said he believed that the software was nearing the capability of human grading.

"This is machine learning and there is a long way to go, but it's good enough and the upside is huge," he said. "We found that the quality of the grading is similar to the variation you find from instructor to instructor."

EdX is not the first to use automated assessment technology, which dates to early mainframe computers in the 1960s. There is now a range of companies offering commercial programs to grade written test answers, and four states — Louisiana, North Dakota, Utah and West Virginia — are using some form of the technology in secondary schools. A fifth, Indiana, has experimented with it. In some cases the software is used as a "second reader," to check the reliability of the human graders.

Source: [technews.acm.org](http://technews.acm.org)

Contributed by: Mrs. Kalpana Singh