AC 2010-1154: MOTIVATING EFFECTIVE PEER REVIEW WITH EXTRA CREDIT AND LEADERBOARDS

Edward Gehringer, North Carolina State University
Ed Gehringer, efg@ncsu.edu, is Associate Professor of Computer Science and Computer Engineering at North Carolina State University. His main research area is collaborative learning technology. He received his Ph.D. degree from Purdue University, and taught at Carnegie Mellon University, and Monash University in Australia.

Abhishek Gummadi, North Carolina State University
Abhishek Gummadi, agummad@ncsu.edu, is a graduate student in the Department of Computer Science at North Carolina State University. He is working on a masters thesis on the role of game mechanics in motivating effective peer reviewing.

Reejesh Kadanjoth, North Carolina State University
Reejesh Kadanjoth, rkadanj@ncsu.edu, is a graduate student in the Department of Computer Science at North Carolina State University. He is working on a masters thesis on supporting the creation of large documents such as wiki textbooks in an online peer-review system.

Yvonne Marie Andrés, Global Schoolnet.org
Yvonne Marie Andrés, yvonne@gsn.org, is the President & CEO of the Global SchoolNet Foundation, which partners with schools, communities and businesses to provide online collaborative learning programs that prepare students for the workforce and help them to become responsible and literate global citizens.
Motivating Effective Peer Review
with Extra Credit and Leaderboards

Abstract

Peer review is a pedagogically sound practice that has found its way into education in all fields, including engineering. Students are often skeptical of its value, and don’t give it the attention it merits. We have implemented several facilities to address this need in our Expertiza peer-review system. Rubric-based review guides each student through the review process. Metareviewing is performed to assign scores to each reviewer’s reviews. Scores for reviewing can be factored into the student’s grade. Leaderboards allow students to see, with permission, who are the class leaders in categories such as submitted work and review quality. We are currently working on implementing an approach to awarding extra credit to reviewers for interacting with their authors more frequently than required.

1. Motivation

Online gaming plays an important role in the lives of many of our students. They compete with each other to earn points and achieve status. The rewards motivate them to spend hours working on task. Suppose we could entice our students to contribute to each other’s learning experience by awarding points for providing formative feedback and answers to each other’s questions. This, it turns out, can be accomplished by a simple extension to student peer review.

Expertiza [1, 2] is a Web-based system for peer-reviewing student work. After students submit their work by uploading a file or writing on a wiki, other students are assigned to review their work. Author and reviewer communicate in double-blind fashion using rubrics. The reviewer fills out a rubric that includes several questions about the author’s work, as

![Figure 1. Filling out a review form](image-url)
shown in Figure 1. Each rubric question can be given a numeric score, and there is space for a prose comment on each question. The system computes a composite score based on the scores for the individual rubric questions.

While filling out a review rubric, the reviewer can be guided by rubric advice, which is a description of the characteristics of work meriting a particular score for a particular question. Clicking on “Show advice” will display the advice for the question. The reviewer can then click on the advice for a particular score, and that score will appear in the dropbox. Figure 2 shows how this works. We have begun to involve students in the process of creating rubrics, by having the class discuss first the rubric questions and then the rubric advice [3].

Similarly, authors give feedback to their reviewers, using a rubric that asks about the helpfulness, accuracy, and respectfulness of the review; and the scores from this feedback can be included in the reviewers’ grades. The form appears when the author views a review (Figure 3).

If we keep track of these scores for a whole class, we can display them—with the student’s permission—to others via a leaderboard. A leaderboard is simply a list of the leaders in a particular category, such as the students receiving the highest scores on reviews of their work, or the students whose feedback is rated most highly by the authors they have reviewed (Figure 4). Students will only appear on the leaderboard if they give permission (by editing their profile) for their scores to appear.

In Expertiza, grades can be based on a number of factors, including reviews by other students, metareviews, and feedback from the authors who are being assessed. The weightings are specified when an assignment is set up (Figure 5). The fact that their

Figure 2. Using rubric advice
reviewing counts in their grade gives the students an incentive to do careful reviews. If there are no serious disagreements between reviewers, or between reviewer and author, we normally accept student-assigned scores as the grade for the assignment. In case of controversy, however, the instructor or TA always has the ability to override student-assigned scores.

Figure 3. Giving feedback to reviewer

Figure 4. A leaderboard
2. Encouraging more frequent interaction

Author and reviewer are not limited to a single interaction. After receiving feedback from the reviewer, authors can revise their work, and reviewers can review it again. We are currently extending our system to track and give credit for these extra reviews. Providing formative feedback through extra reviews enhances learning, and those who give the most frequent reviews can be recognized with their own leaderboard category. To prevent credit for vacuous reviews, points are only awarded for reviews that are ranked above a certain threshold for helpfulness by the author.

Thus far, we have been talking about earning credit for assigned work—reviews, metareviews, teamwork, and the like. We are also interested in stimulating other kinds of interaction, such as asking and answering questions on a message board. We can allow readers to rate the questions and answers on a numeric scale, and award points for them, depending on how likely they are rated. However, rewarding unassigned work can lead to collusion. A group of students can get together and decide that they will rate each other’s questions and answers highly, regardless of merit. Preventing collusion is one goal of a reputation system.

Figure 5. How the instructor can count reviewing toward a student’s grade
3. Reputation systems

A reputation system is a way of measuring the reliability of ratings. Scores assigned by reviewers and metareviewers can be factored into a student's reputation. Several algorithms [4, 5, 6] have also been published for determining reviewer reliability, based only on the scores assigned by reviewers. These algorithms consider (i) consistency of scores assigned by this reviewer with scores assigned by others to the same work, and (ii) spread, how much the highest score the reviewer assigned differs from the lowest score (s)he assigned. Some algorithms also consider (iii) leniency, the tendency of a reviewer to give scores that are higher than other reviewers. Research [6] demonstrates that these algorithms provide effective quality control for reviews.

A reviewer's reputation can be used in two ways: as part of a reviewer's grade for an assignment, or to weight scores assigned by this reviewer in calculating other students' grades. A score assigned by a reviewer with a good reputation should count for more than a score assigned by a poor reviewer when computing the grade of the student being reviewed. Finally, students with the best reputations can be recognized on the leaderboard, just as they can be recognized for good reviewing or for performing a lot of extra reviews.

A student's reputation should be based both on assigned work and unassigned work. It is harder to compute reputation for unassigned work, because the strategy must deal with collusion. One simple strategy is to base reputation primarily on assigned work, and allow unassigned work to affect it only slightly, so that collusion is not worth the trouble. Another approach, in use by Slashdot, is to base reputation on the number of reviewers, not reviews, that rate a student's work as excellent. Then regardless how many "5"s a reviewer gives an author, it is no more valuable than giving a single "5". A third way is to reward a reviewer for being the first to give a particular post the same rating it later receives from other users; this is a takeoff on the SPEAR algorithm [7] used to combat spam on social Web sites. It would have the effect of damaging the reputation of colluders who rate a question or answer more highly than others rate it.

Students can also be awarded points for answering student questions on a message board. The questioner can rate the helpfulness of the response, and so can other students. Helpful answers to questions is another leaderboard category. To prevent collusion between students who agree to rate each other's answers highly, points are awarded for the number of students who highly rate a particular student's message-board answers.

4. Administration

Setting up an assignment entails several steps.

1. Creation of rubrics for review, feedback from authors, teammate review, and metareview.

2. Creation of the assignment, by specifying the rubrics, deadlines, and type of submission to be accepted.

3. Adding the participants to the assignment.
To make it easy to get started with the system, we have created several “public” rubrics of all four types. A new instructor can simply select from these rubrics, or copy and modify them, if desired.

Participants, teams, and review assignments can be imported from CSV files. The Expertiza project staff typically helps instructors set up their first assignment. After that, subsequent assignments can be created by copying and modifying an existing assignment.

Because the leaderboard system is new to Expertiza, we don’t have empirical data on how it affects engagement and student learning. However, we do have a survey on the peer-review process that has been administered to students over several previous semesters. This will allow us collect data on engagement and learning before and after leaderboards were incorporated into the system.

5. Conclusion

Peer review holds important advantages for students, both as assessors and as asessees. As assessors, the spend time reviewing, summarizing, diagnosing misconceived knowledge, and considering deviations from the ideal. As asessees, they write for an audience where they have the burden of making themselves understood, rather than writing for an expert grader, who is expected to be able to decipher their intentions. A large number of studies attest to the value of academic peer review in a wide variety of disciplines [8]. Students, however, are very skeptical about the value of peer review to their learning [9]. A clear need is to motivate students to “buy into” the peer-review process. In Expertiza, we have implemented several mechanisms for motivating good reviews: (1) authors can pose questions to reviewers, to understand the review, and how the reviewer thinks their work can be improved; (2) a student’s review can be metareviewed by the instructor or by another student, and a score assigned to it; (3) both feedback from the author and metareviews can be counted in determining a student’s final grade on the assignment, and (4) students can be recognized on a leaderboard as being among the best reviewers in the class. Future plans include incorporating a way to give extra credit for more frequent reviews, and implementing a reputation system that will bring additional evidence to bear on evaluation of reviewer competence.

Bibliography


