Using Argument Diagramming to Improve Peer Grading of Writing Assignments

Mohammad H. Falakmasir [1], Kevin D. Ashley
Christian D. Schunn

Intelligent Systems Program, Learning Research and Development Center,
University of Pittsburgh
{mhf11, ashley, schunn}@pitt.edu

Abstract. One of the major components of MOOCs is the weekly assignment. Most of the assignments are multiple choice, short answer or programming assignments and can be graded automatically by the system. Since assignments that include argumentation or scientific writing are difficult to grade automatically, MOOCs often use a crowd-sourced evaluation of the writing assignments in the form of peer grading. Studies show that this peer-grading scheme faces some reliability issues due to widespread variation in the course participants' motivation and preparation. In this paper we present a process of computer-supported argumentation diagramming and essay writing that facilitates the peer grading of the writing assignments. The process has not been implemented in a MOOC context but all the supporting tools are web-based and can be easily applied to MOOC settings.

Keywords: Computer Supported Argumentation, Argument Diagramming, Peer Review and Grading

1 Introduction

MOOCs in general and Coursera, in particular, started with courses in the area of Computer Science. These courses offered a variety of homework including multiple choice, short answer, and programming assignments that can be graded automatically by the system. However, recently, many MOOCs have started offering courses in social sciences, humanities, and law subjects whose assignments naturally involve more writing and argumentation. Automatic grading of those kinds of assignments is more challenging given the current state of natural language processing technologies. Coursera and most of the other current systems use a peer-grading mechanism in order to address this issue. However, because of the open access nature of the MOOCs, a massive number of people with different educational backgrounds and language skills from all around the world participate in these courses and this heterogeneity in prior preparation negatively affects the validity and reliability of

[1] Corresponding Author

peer-grades. Researchers have investigated this issue (Duneier, 2012) and some steps have been taken to address it. Coursera, for example, flags students who give inaccurate grades and assigns their assessments less weight, but this method does not directly address the diversity of knowledge and writing skills among the students. In this paper, we recommend an approach to this issue that combines computer-supported argument diagramming and writing with scaffolded peer-review and grading. With support of the National Science Foundation,[2] our ArgumentPeer process combines two web-based tools (SWoRD and LASAD) that have been used in several university settings and courses, and applies them to support argumentation and writing assignments in science and law. The process enables the instructional team to carefully define and monitor the writing assignment and revision procedure and involves several machine learning and natural language processing components.

2 Background

Writing and argumentation are fundamental skills that support learning in many topics. Being able to understand the relationships among abstract ideas, to apply them in solving concrete problems, and to articulate the implications of different findings for studies and theories are essential for students in all areas of science, engineering, and social studies. However, inculcating these skills, or compensating for the lack of them, is especially difficult in MOOC setting where students have such diverse preparations and motivations.

Our approach to tackle this problem involves breaking down the process of writing into multiple measurable steps and guiding the student through the steps with careful support and feedback. The first step of the process, computer-supported argument planning, engages the students with a graphical representation for constructing arguments and provides them with feedback and intelligent support. We use LASAD[3] as our argument-diagramming tool (cf. Scheuer et al., 2010). LASAD is a web-based argumentation support system to help students learn argumentation in different domains. It supports flexible argument diagramming by enabling instructors to define a pre-structured palette of argumentation elements (Argument Ontology) along with a set of help system rules in order to give instant feedback to students while working on their diagrams.

The massive number of students in MOOC settings makes it impossible for the instructional team to provide reflective feedback on each individual student's argument. We handle this issue with computer-supported peer-review and grading using SWoRD[4] (Cho & Schunn, 2007). In general, peer review is consistent with learning theories that promote active learning. Furthermore, the peer-review of writing has some learning benefits for the reviewer, especially when the students provide constructive feedback (Wooley, Was, Schunn, & Dalton, 2008), and put effort into the process (Cho & Schunn, 2010). Moreover, studies have shown that

[2] Award No.1122504 DIP: Teaching Writing and Argumentation with AI-Supported Diagramming and Peer Review

[3] http://cscwlab.in.tu-clausthal.de/lasad/

[4] https://sites.google.com/site/swordlrdc/

feedback from a group of peers can be at least as useful as that of teachers (Cho & Schunn, 2007), especially when good rubrics and incentives for reviewing are included. Most relevant here, studies have shown that even students with lower levels of knowledge in the topic can provide feedback that is useful to the ones with higher levels (Patchan & Schunn, 2010; Patchan, 2011).

3 The Process

The ArgumentPeer process includes two main phases: 1) Argument Planning, and 2) Argument Writing. Fig. 1 shows an overview of the process and its underlying components and steps.

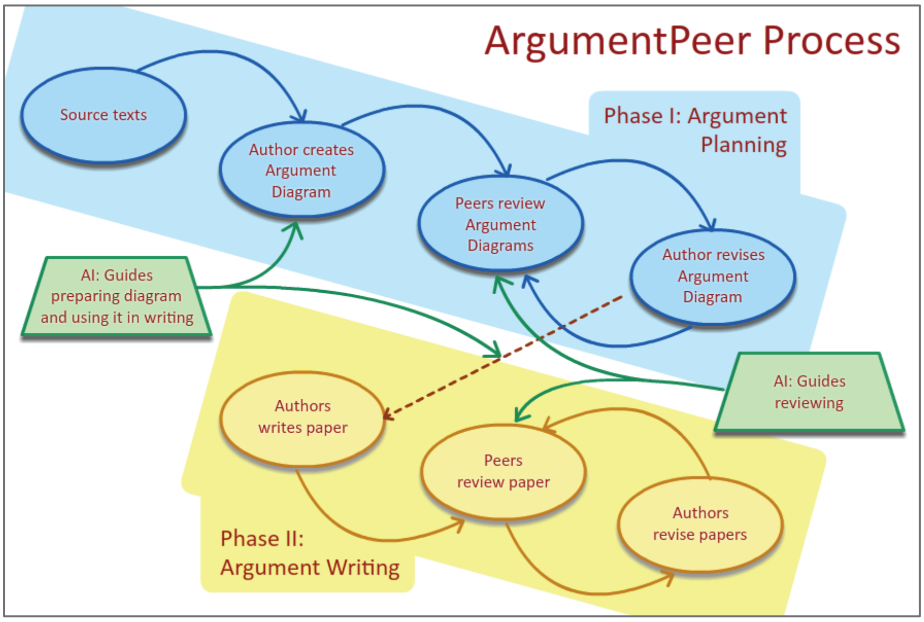

Fig. 1: ArgumentPeer Process

3.1 Phase I: Argument Diagramming

This phase includes studying the assigned resources and creating the argument diagram. As an example, students in a legal writing course used LASAD in order to prepare textual brief on appeal to the U.S. Supreme Court in the case of *United States v. Alvarez* (Lynch et al., 2012). The system had been introduced to them in a 45-minutes lecture session (that could easily be made a video) and students were directed toward a recommended stepwise format for written legal argumentation as set forth in a noted authority (Neumann 2005). Figure 2 shows an example diagram in this study.

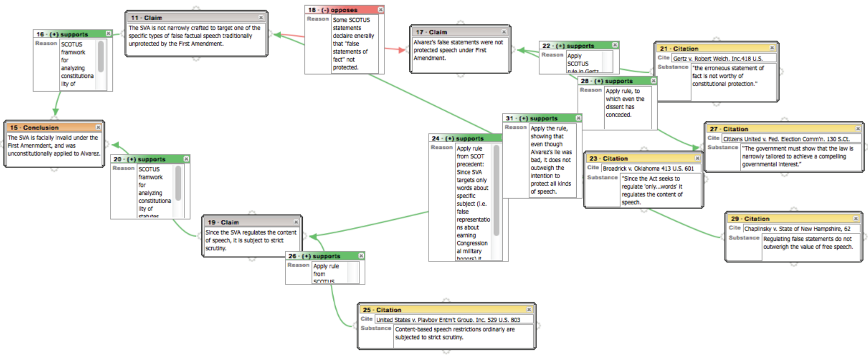

Fig. 2: Example Argument Diagram in Legal Writing Course

The instructional team tailored the argument ontology to support the recommended argumentation format; the nodes were basically legal "claim" and "conclusion" nodes that are connected together via "supporting" and "opposing" links providing reasons for and against. The development of a suitable ontology is a critical aspect in the design of an argumentation system and might involve iterative refinement based on observed problems and weaknesses (Buckingham et al., 2002). Specifically, ontologies affect the style of argumentation (Suthers, et al., 2001) and the level of details expected for students to provide. LASAD provides an authoring tool that enables the instructional team to carefully design the argumentation ontology.

After creating the argument diagrams, the students submit their diagrams to the SWoRD system for revision. As noted, SWoRD lets instructors provide a detailed rubric with which peers should assess the diagram. Moreover, it has a natural language processing (NLP) component that pushes reviewers to provide useful feedback that is not ambiguous or vague (more details in section 3.3). After receiving the reviews, the author will revise his/her argument diagram and get ready to write the first draft of the writing assignment in phase 2. To support this transition to a written argument, a system component creates a textual outline based on a depth-first traversal of the argumentation diagram and informed by the argument ontology. In this way, students are encouraged to create a well-annotated argumentation diagram because the diagram text is easily transferred directly to the written draft.

3.2 Phase II: Writing

In this phase, students write their first drafts using the outlines generated from the argument diagrams and submit them to SWoRD. After that, the system automatically assigns the draft to *n* reviewers based on the instructors' policy. The instructor can also assign the individual or groups of peers for the revision using various methods. For example, in the Legal Writing course, the instructor divided the students into two groups, one, writing for the majority and the other writing for the dissenting judge in the 9th Circuit U.S. Court of Appeals and assigned the peers in a way such that there is at least one peer from the other group among the reviewers.

In the next step, the instructor carefully designs the paper reviewing criteria (rubric) for the peers and then starts the reviewing process. The key feature of SWoRD is the ease with which instructors can define rubrics to guide peer reviewers in rating and commenting upon authors' work. The instructor-provided rubrics, which may include both general domain writing and content-specific criteria (Goldin & Ashley, 2012), should help to focus peer feedback and compensate for the wide diversity of peer-reviewers' preparation and motivation.

Reviewers, then, download the paper and evaluate them based on the defined rubric and submit their reviews and ratings to SWoRD. Again, the NLP component of the system, checks the reviews for usefulness and then the system deliverers the reviews back to the author. SWoRD automatically determines the accuracy of each reviewer's numerical ratings using a measure of consistency applied across all of the writing dimensions (Cho & Schunn, 2007). Finally, the author submits the second draft to the system and the final draft can either be grader by peers or the instructional team, although of course in a MOOC context peers would grade it again.

3.3 AI Guides Student Authors and Reviewers in Both Phases

As mentioned, the LASAD Authoring tool and its flexible ontology structure enable instructors to specify the level of detail on which they want the students to focus. Instructors can also use the Feedback Authoring tool to define help system rules that guide the students through the argumentation diagramming process. The instant feedback component of LASAD is an expert system that uses logical rules to analyze students' developing argument diagrams and to provide feedback on making more complete and correct diagrams. The hints can be as simple as telling the student to fill in a text field for an element, or as complex as telling the student to include opposing, as well as supporting, citations for a finding. Using this in-depth intervention, instructors can focus students on their intended pedagogical goals. For example, in the legal writing course, a help system rule asks students to include at least one opposing "citation" in their diagrams to anticipate possible important counterarguments that a court would expect an advocate to have addressed in his or her brief.

The NLP component of SWoRD helps the students improve their reviews by detecting the presence or absence of key feedback features like the location of the problem and the presence of an explicit solution. This feature has been implemented for review comments on both argumentation diagrams and the written drafts. The details of the computational linguistic algorithm that detects the feedback issues are described in (Xiong et al., 2012; Nguyen & Litman, in press). The interface provides reviewers with advice like: "Say where this issue happened." "Make sure that for every comment below, you explain where in the paper it applies." In addition, it provides examples of the kind of good feedback likely to result in an effective revision: "For example, on page [x] paragraph [y], …. Suggest how to fix this problem." "For example, when you talk about [x], you can go into more detail using quotes from the reading resource [y]." The system tries to be as helpful as possible, but in order to prevent frustration, it allows the reviewers to ignore the suggestions and submit the review as is. However, SWoRD considers these reviewers as less accurate and gives lower weight to their ratings.

4 Assessment and Grading

After submitting the final draft, the papers are assigned automatically or by the instructors to the same or another group of peers (or members of the instructional team in non-MOOC contexts) for grading. The same rubric can be used for the second round of review but it is also possible to define new criteria particularly for grading purposes.

According to (Cho, Schunn, & Wilson, 2006; Patchan, Charney, & Schunn, 2009) the aggregate ratings of at least 4 peers on a piece of writing in this setting are more highly reliable and just as valid as a single instructor's ratings. However, some studies (e.g., Chang et al., 2011) note that there can be systematic differences between peer and instructor assessment in a web-based portfolio setting. We believe that by breaking down the argument planning and writing process into multiple guided steps, each subject to review according to instructor-designed peer-review criteria, we move toward a more reliable peer-grading scheme that can be especially useful in a MOOC context.

5 Discussion

Grading writing assignments requires considerable effort, especially when the class size increases. Peer-review and grading is one way to deal with this problem but many instructors are hesitant to use it in their classrooms. The main concern is whether the students are actually capable of grading the papers accurately and responsively. Studies have shown that peer rating alone can be reliable and valid in a large-scale classroom under appropriate circumstances and well-chosen review criteria (Cho, Schunn, & Wilson, 2006; Patchan, Charney, & Schunn, 2009). The ArgumentPeer project not only enables the instructor to design the rubric but also makes it salient for the reviewer to see the deep structure of the argumentation by viewing the argumentation diagram. This positive synergy between diagramming and peer-review makes it easier for the reviewer to see the argument structure in the diagram and its reflection in the writing.

Regarding scalability and the possibility of being used in a MOOC setting, both SWoRD and LASAD are web-based projects developed using Java 2 Platform, Enterprise Edition (J2EE) architecture. LASAD uses automated load balancing in order to support a large number of students. The rich graphical interface of LASAD along with flexible structure of the ontologies helps students gain an understanding of the topic of argumentation (Loll, et al., 2010). Moreover, the collaborative nature of LASAD can be used in order to facilitate engagement, particularly in MOOC settings that face the problem of student retention.

SWoRD, which is the main platform for peer-review and grading, has also been successfully used in classrooms with a large number of students (Cho, Schunn, & Wilson, 2006). The basic review structure in SWoRD is quite similar to the journal publication process, which makes it a familiar process among academics. In addition, publicizing students' papers to their peers can make students put more effort into writing by increasing audience awareness (Cohen & Riel, 1989).

6 Conclusion

In this paper, we presented a process of argument diagramming and reciprocal peer-review in order to facilitate the grading of writing assignments. The ArgumentPeer process and its preexisting components, SWoRD and LASAD, have been applied across different university settings in different courses with large numbers of students. We have decomposed writing assignments into separate steps of planning an argument and then writing it, support students in each step with instructor- and AI-guided peer reviewing and grading. The results of our past studies show that high reliability and validity in the peer grading can be achieved with multiple reviewers per paper. The web-based nature of the components of the ArgumentPeer process makes it relatively easy to apply in MOOC settings. We believe that its fine-grained support for authoring and reviewing could help achieve higher levels of reliability and validity in MOOCs despite their massive numbers of highly diverse participants.

References

1. Buckingham Shum, S. J., Uren, V., Li, G., Domingue, J., Motta, E., & Mancini, C. (2002). Designing representational coherence into an infrastructure for collective sense-making. Invited discussion paper presented at the 2nd International Workshop on Infrastructures for Distributed Collective Practices.
2. Chang, C. C., Tseng, K. H., & Lou, S. J. (2011). A comparative analysis of the consistency and difference among teacher-assessment, student self-assessment and peer-assessment in a Web-based portfolio assessment environment for high school students. *Computers and Education*, 58(1), 303-320.
3. Cho, K., & Schunn, C. D. (2007). Scaffolded writing and rewriting in the discipline: A web-based reciprocal peer review system. *Computers & Education*, 48(3), 409–426.
4. Cho, K., & Schunn, C. D. (2010). Developing writing skills through students giving instructional explanations. In M. K. Stein & L. Kucan (Eds.), *Instructional Explanations in the Disciplines: Talk, Texts and Technology*. New York: Springer.
5. Cho, K., Schunn, C. D., & Wilson, R. W. (2006). Validity and reliability of scaffolded peer assessment of writing from instructor and student perspectives. *Journal of Educational Psychology*, 98(4), 891-901.
6. Cohen, M., & Riel, M. (1989). The effect of distant audiences on students' writing. *American Educational Research Journal*, 26, 143–159.
7. Duneier, M. (2012). Teaching to the world from central New Jersey. *Chronicle of Higher Education*, September 3.
8. Goldin, I. M. & Ashley, K. D. (2012) Eliciting Formative Assessment in Peer Review. *Journal of Writing Research* 4(2) pp. 203–237.
9. Loll, F., Scheuer, O., McLaren, B. M. & Pinkwart, N. (2010). Computer-Supported Argumentation Learning: A Survey of Teachers, Researchers, and System Developers. In M. Wolpers, P. A. Kirschner, M. Scheffel, S. Lindstaedt, & V. Dimitrova, Proceedings of the 5th European Conference on Technology Enhanced Learning (EC-TEL 2010), LNCS 6383, pp. 530-535. Springer.
10. Lynch, C., Ashley, K. D., Falakmassir, M. H., Comparing Argument Diagrams, in proceedings of The 25th Annual Conference on Legal Knowledge and Information Systems (JURIX), Amsterdam, Netherlands, December 2012, pp. 81-90.

11. Neumann, R. (2005) *Legal Reasoning and Legal Writing: Structure, Strategy, and Style*. (5th Ed.) Walters Kluwer.
12. Nguyen H., Litman D., (in press). Identifying Localization in Peer Reviews of Argument Diagrams. Accepted in the 16th International Conference on Artificial Intelligence in Education (AIED 2013), Memphis, TN.
13. Patchan, M. M., Charney, D., & Schunn, C. D. (2009). A validation study of students' end comments: Comparing comments by students, a writing instructor, and a content instructor. *Journal of Writing Research*, 1(2), 124-152.
14. Patchan, M. M., & Schunn, C. D. (2010). Impact of Diverse Abilities on Learning to Write through Peer-Review. Paper presented at the 32nd annual meeting of the Cognitive Science Society, Portland, OR.
15. Scheuer, O., Loll, F., Pinkwart, N. and McLaren, B. M. (2010). Computer-supported argumentation: A review of the state-ofthe-art. *International Journal on Computer Supported Collaborative Learning*, 5(1), 43-102. Springer.
16. Suthers, D. D., Connelly, J., Lesgold, A., Paolucci, M., Toth, E. E., Toth, J., & Weiner, A. (2001). Representational and advisory guidance for students learning scientific inquiry. In K. D. Forbus & P. J. Feltovich (Eds.), *Smart machines in education: The coming revolution in educational technology* (pp. 7–35). Menlo Park, CA: AAAI/MIT Press.
17. Wooley, R., Was, C., Schunn, C., & Dalton, D. (2008). The effects of feedback elaboration on the giver of feedback. Paper presented at the 30th Annual Meeting of the Cognitive Science Society.
18. Xiong, W., Litman, D., & Schunn, C. D. (2010). Natural Language Processing techniques for researching and improving peer feedback. *Journal of Writing Research*, 4(2), 155-176.