Identifying Editor Roles in Argumentative Writing from Student Revision Histories

Tazin Afrin(✉) and Diane Litman

University of Pittsburgh, Pittsburgh, PA 15260, USA
{tazinafrin,litman}@cs.pitt.edu

Abstract. We present a method for identifying editor roles from students' revision behaviors during argumentative writing. We first develop a method for applying a topic modeling algorithm to identify a set of editor roles from a vocabulary capturing three aspects of student revision behaviors: operation, purpose, and position. We validate the identified roles by showing that modeling the editor roles that students take when revising a paper not only accounts for the variance in revision purposes in our data, but also relates to writing improvement.

Keywords: Editor role · Argumentative writing · Revision

1 Introduction

Knowing that experienced and successful writers revise differently than inexperienced writers [4], various intelligent writing tools have been developed that provide localized feedback on text characteristics [3,5,6,9]. These tools typically suggest edits to guide revision, rather than model the editing process after observing revisions. With the long term goal of developing an intelligent revision assistant, this paper presents an approach to modeling student editor roles.

Prior natural language processing (NLP) approaches to student revision analysis have focused on identifying revisions during argumentative writing and classifying their purposes and other properties [1,7,11,12]. In contrast, editor roles have generally been studied in NLP using online collaborative writing applications such as Wikipedia [10]. Inspired by the use of Wikipedia revision histories [10], in this paper we similarly use topic modeling applied to revision histories to identify editor roles in the domain of student argumentative writing. To model student revision histories, between-draft essay revisions are extracted at a sentence-level and represented in terms of the following three aspects: operation (add, delete, or modify a sentence), purpose (e.g., correct grammar versus improve fluency), and position (revise at the beginning, middle or the end of an essay). To identify editor roles, a Latent Dirichlet Allocation (LDA) [2] graphical model is then applied to these revision histories. Finally, we show that the identified roles capture the variability in our data as well as correlate with writing improvement.

S. Isotani et al. (Eds.): AIED 2019, LNAI 11626, pp. 9–13, 2019.
https://doi.org/10.1007/978-3-030-23207-8_2

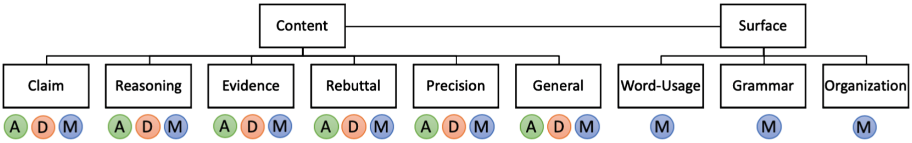

Fig. 1. The taxonomy of revision purposes [12] (A: Add, D: Delete, M: Modify).

Table 1. Example revision from aligned drafts of an essay from the Modeling Corpus.

Original draft	Revised draft	Operation	Purpose	Position
Self-driving vehicles pose many advantages and disadvantages	While self-driving vehicles pose many advantages and disadvantages, I am not on the bandwagon for them at this time	Modify	Claim	Beg.

2 Corpora

Our work takes advantage of several corpora of multiple drafts of argumentative essays written by both high-school and college students [11,12], where all data has been annotated for revision using the framework of [12]. We divide our data into a Modeling Corpus (185 paired drafts, 3245 revisions) and an Evaluation Corpus (107 paired drafts, 2045 revisions), based on whether expert grades are available before (Score1) and after (Score2) essay revision. Although the grading rubrics for the college and high-school essays in the Evaluation Corpus are different, both are based upon common criteria of argumentative writing, e.g., clear thesis, convincing evidence, clear wording without grammatical errors, etc. We apply linear scaling[1] to bring the scores within the same range of [0,100]. After scaling, the average Score1 and Score2 are 64.41 and 73.59, respectively.

For all essays and prior to this study, subsequent drafts were manually aligned at the sentence-level based on semantic similarity. Nonidentical aligned sentences were extracted as the **revisions**, resulting in three types of revision **operations** - *Add*, *Delete*, *Modify*. Each extracted revision was manually annotated with a **purpose** following the revision schema shown in Fig. 1 (modified compared to [12] by adding the Precision category). For this study, each revision's **position** was in addition automatically tagged using its paragraph position in the revised essay. To maintain consistency across essays, instead of using paragraph number, we identify whether a revision is in the first (*beg*), last (*end*), or a middle (*mid*) paragraph. Table 1 shows a modified claim at the beginning of an essay from the Modeling Corpus.

3 Identifying Editor Roles

To create a vocabulary for topic modeling and to understand the repeating patterns of student editors, we represent each revision utilizing the three aspects

[1] Formula used to scale the scores = 100 * (x-min)/(max-min).

Table 2. Derived editor roles with top 10 revisions. (Blue: Surface, Orange: Content)

Proofreader	Copy editor	Descriptive editor	Analytical editor	Persuasive editor
Grammar_mid	Word-Usage_mid	+General_mid	Word-Usage_beg	+Reasoning_mid
Grammar_beg	Word-Usage_beg	Word-Usage_mid	+General_end	-Reasoning_mid
Word-Usage_mid	+Reasoning_mid	-General_mid	+Reasoning_end	+Claims_mid
Grammar_end	Word-Usage_end	General_mid	Word-Usage_end	+Evidence_mid
Word-Usage_end	Organization_mid	Evidence_mid	Organization_beg	+General_mid
Word-Usage_beg	-General_end	Precision_mid	-Reasoning_end	-General_mid
Precision_beg	General_end	-General_beg	+Claims_end	Reasoning_mid
General_mid	-Reasoning_mid	+General_beg	+Evidence_mid	-General_beg
General_end	Claims_mid	Reasoning_mid	+Rebuttal_end	-Claims_mid
Reasoning_beg	-General_mid	+Claims_beg	Organization_mid	+General_beg

described earlier: operation, purpose, and position. This yields a rich and informative vocabulary for modeling our data, consisting of 63 revision "words" (54 content, 9 surface). This is in contrast to the 24 word revision vocabulary used in the prior Wikipedia editor role extraction method [10], formed using a Wiki-specific revision taxonomy of operation and purpose. When describing our revision "words", add and delete revisions are represented with '+' and '−' sign, and no sign for modification, e.g., *Claim_beg* in Table 1. Editors are then represented by their history of revisions in terms of this revision vocabulary.

We trained the LDA model on the Modeling Corpus and experimented with 2 to 10 topics. After an extensive evaluation for topic interpretation based on top 10 revisions under each topic, we ended up with 5 topics where the revisions under each topic intuitively correspond to one of a set of potentially relevant editor roles for academic writing. We drew upon roles previously identified for writing domains such as newspaper editing (e.g., proofreader, copy editor), Wikipedia (e.g., technical editor, substantive expert), and academic writing[2] (i.e., descriptive, analytical, persuasive, and critical).

The final topics are shown in Table 2, labeled by us with the best-matching editor role from the anticipated set of potential roles, based on the vocabulary items in each topic. The defining characteristic of a **Proofreader** are surface-level error corrections. **Copy** editors ensure that the article is clear and concise as they revise for word-usage, clarity, and organization. **Descriptive** editors provide details and enhance clarity, with widespread development of general content. **Analytical** editors revise by adding information and better organizing thoughts, with top revision purposes being word-usage, content, reasoning, and rebuttal. **Persuasive** editors discuss ideas and facts with relevant examples and develop arguments with added information.

[2] https://sydney.edu.au/students/writing/types-of-academic-writing.html.

Table 3. Variance across editors for each revision purpose ($p < .001$:***, $N = 107$).

Purpose	Grammar	Word-usage	Organization	Claims	Reasoning	General	Evidence	Rebuttal
R^2**-value**	0.573***	0.537***	0.043	0.240***	0.397***	0.459***	0.223***	0.025

Table 4. Partial correlations between role probabilities and Score2 controlling Score1.

Editor roles	Proofreader	Copy	Descriptive	Analytical	Persuasive
Corr(p-value)	−0.175(0.073)	−0.049(0.621)	−0.180(0.064)	−0.013(0.891)	0.205(**0.035**)

4 Validating Editor Roles

Using the trained topic model, we first calculate the probability of an editor belonging to each of the 5 roles, for each editor in the Evaluation Corpus. These probabilities represent each role's contribution to the essay revision. Motivated by Wikipedia role validation [10], we first validate our editor roles by similarly using editor roles to explain the variance in revision purposes. We create 8 linear regression models, one for each revision purpose[3]. The models take as input a five dimensional vector indicating an editor's contribution to each role and the output is the editor's edit frequency for each revision purpose. The R-squared values in Table 3 show that our topic model can best explain the variance of Grammar, Word-Usage, General content, Claim, Reasoning, and Evidence edits.

A corpus study in [12] showed that content changes are correlated with argumentative writing improvement, reaffirming the statement of [4]. Using a similar method, we investigate if our editor roles are related to writing improvement. We calculate partial Pearson correlations between editor roles and Score2 while controlling for Score1 to regress out the effect of the correlation between Score1 and Score2 (Corr. $= 0.692$, $p < 0.001$). Table 4 shows that the roles consisting of only surface edits or a mixture of edits are not correlated to writing improvement. However, Persuasive editor, which consists of content revisions, shows a positive significant correlation to writing improvement. Our results suggest that the Persuasive editor is the role of an experienced writer.

5 Conclusion and Future Work

Although editor roles have been studied for online collaborative writing [8,10], our research investigates student revisions of argumentative essays. While our model follows previous methods [10], we introduce a unique vocabulary to model each editor's revision history, with evaluation results suggesting that our identified roles capture salient features of writing. Future plans include using a Markov model to consider revision order, expanding the revision vocabulary, and using the predictions to provide feedback in an intelligent revision assistant.

Acknowledgements. This work is funded by NSF Award 1735752.

[3] The Evaluation Corpus does not have precision revisions.

References

1. Afrin, T., Litman, D.: Annotation and classification of sentence-level revision improvement. In: Proceedings of the Thirteenth Workshop on Innovative Use of NLP for Building Educational Applications, New Orleans, Louisiana, pp. 240–246, June 2018
2. Blei, D.M., Ng, A.Y., Jordan, M.I.: Latent dirichlet allocation. J. Mach. Learn. Res. **3**, 993–1022 (2003)
3. Eli Review, T.: (2014). https://elireview.com. Accessed 02 June 2019
4. Faigley, L., Witte, S.: Analyzing revision. Coll. Compos. Commun. **32**(4), 400–414 (1981)
5. Grammarly (2016). http://www.grammarly.com. Accessed 02 June 2019
6. Roscoe, R.D., McNamara, D.S.: Writing pal: feasibility of an intelligent writing strategy tutor in the high school classroom. J. Educ. Psychol. **105**(4), 1010–1025 (2013)
7. Tan, C., Lee, L.: A corpus of sentence-level revisions in academic writing: a step towards understanding statement strength in communication. In: Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics, vol. 2: Short Papers, Baltimore, MD, USA, pp. 403–408, June 2014
8. Welser, H.T., et al.: Finding social roles in Wikipedia. In: Proceedings of the 2011 iConference, iConference 2011, pp. 122–129. ACM, New York (2011)
9. The Writing Mentor: ETS writing mentor (2016). https://mentormywriting.org/. Accessed 02 June 2019
10. Yang, D., Halfaker, A., Kraut, R.E., Hovy, E.H.: Who did what: editor role identification in Wikipedia. In: Tenth International AAAI Conference on Web and Social Media, ICWSM, pp. 446–455. AAAI Press (2016)
11. Zhang, F., Hashemi, H., Hwa, R., Litman, D.: A corpus of annotated revisions for studying argumentative writing. In: Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics, vol. 1: Long Papers, pp. 1568–1578. Association for Computational Linguistics (2017)
12. Zhang, F., Litman, D.: Annotation and classification of argumentative writing revisions. In: Proceedings of the 10th Workshop on Innovative Use of NLP for Building Educational Applications, Denver, Colorado, pp. 133–143. Association for Computational Linguistics, June 2015