Computational Argumentation Analysis and Assessment on Essays in Dutch for Formative Feedback Generation



Marco Kalz

Howard Spoelstra

Welten-instituut Onderzoekscentrum voor leren, doceren en technologie Open Universiteit welten-instituut.ou.nl





- Feedback and automatic assessment tool.
- Survey of essay assessment and formative feedback.
- Computational argumentation analysis method.
- Discussion.

Welten-instituut Onderzoekscentrum voor leren, doceren en technologie



Essay, assessment, and formative feedback

- Essays are commonly used in the formal educational programs as an approach for assessing the learning outcome of the students (Miller et al., 2014).
- Essay assessment: grading and giving formative feedback.
- Formative feedback
 - Giving formative feedback on essay is one of the most effective approaches for student's development in different aspects (Smith & Gorard, 2005; Shute, 2008; Wingate, 2010).
 - Various types of formative feedback (content, develop student skills, etc.).
 - Various levels of formative feedback (acknowledgement, correction, explanation).

Welten-instituut



Automatic essay assessment

- Essay assessment brings heavy workload and it is timeconsuming.
- Related to natural language processing (NLP) technology.
- Previous works of automatic essay assessment tool.
 - Automated grading.
 - Keywords/key phrases.

Welten-instituut





Welten-instituut

Onderzoekscentrum \

Issues

- What do the teachers think of essays and formative feedback?
- What do the teachers need from automatic assessment tools?
- Are the current assessment tools suitable for the essays in Dutch?
- Is it possible to apply the state-of-the-art NLP technology for the automatic assessment tools?

Welten-instituut



The aim of this project

- The opinions of the teachers over essay, formative feedback, and automatic assessment tool.
 - Survey
- Developing a model of automatic assessment in the Dutch environment with Dutch corpus.
- Application of the latest NLP technology.
 - Argument component identification and its multilingual applicating (Stab &Gurevych, 2017; Eger et al., 2018).
 - Computational argument quality assessment (Wachsmuth et al., 2017).

Welten-instituut



Current work 1: Survey (still running)

- 20 responses.
- Most of the participants have background in education or computation science.
- 47% teaching, 30% research, 23% others.
- Almost everyone has experience in assessing essays in Wetenschappelijk Onderwijs (WO) level.

Welten-instituut



Some preliminary survey results: Importance level of various essay purposes

Essay purposes	Importance level
To recall and use knowledge	7.6
To develop an argument	8.6
To express idea	7.95
To improve writing skills	8.15

Welten-instituut



Some preliminary survey results: Formative feedback v.s. Assessment methods

	Consider the whole essay	Read sentence by sentence	Focus on specific sections	Search for keywords	Focus on coherence
To recall and use knowledge	6.15	5.45	6.75	6.3	5.5
To develop an argument	7.1	5.85	7.15	4.35	8.2
To express idea	7.2	5.7	6.65	4.45	8.05
To improve writing skills	6.9	5.65	6.65	4.1	8.05

Welten-instituut



Some preliminary survey results: Formative feedback v.s. Essay purpose

	To recall and use knowledge	To develop an argument	To express idea	To improve writing skills
Correct errors or misconception	8.8	8.65	7.85	6.95
Point to the omission of relevant material(s)	7.7	7.65	6.65	5.8
Point to the inclusion of irrelevant material(s)	7.25	8	6.5	6.05
Comment on the argumentation	7.55	8.85	8.1	8.05
Comment on the use of language	6.5	7.75	7.8	9.15
Comment on structure of the essay	6.4	7.95	7.6	8.8
Encourage dialogue	6.55	7.4	7.6	7.10
Point to additional resource materials	7.45	7.2	7.5	6.5
Praise the student	7.55	7.7	7.65	7.45

Welten-instituut

Onderzoekscentrum voor leren, doceren en technologie

Open Universiteit welten-instituut.ou.nl



Some preliminary survey results: Automatic assessment tool

Features provided by the automatic assessment tool	How helpful do you think?
Give insight into keywords/key sentences used	6.15
Provide statistic descriptions (essay length, word length, etc.)	5.7
Indicate the function of a sentence (background info, argument, summary, etc.) in an essay	6.4
Point to grammar mistakes	7
Indicate the quality of the argumentation	8.5

Welten-instituut



Summary of the current survey results

- Essays aiming to develop argumentation might be more important than other essay purposes.
- Teachers thinks it is more important to give formative feedback over argumentation, but it is also more difficult and timeconsuming.
- It'd be nice if an automatic assessment tool can evaluate the argumentation in an essay.





Current work 2: Developing argument component identification model

- Argument component identification for Dutch essay.
 - The argumentation function of a sentence in an essay: the stance of the author, the argument the stance, the support evidence of the argument, etc.
 - Stab & Gurevych (2017) present a new approach for parsing argumentation structure.
 - Eger et al. (2018) extent the model from Stab & Gurevych (2017) to multilingual application.
- Creating a Dutch essay corpus with argumentation structure annotated.

Welten-instituut



A standard machine learning pipeline



Argumentation component identification model for Dutch

- Data
 - Persuasive essay with argumentation structure annotation in English.
 - Machine translation from English to Dutch.
- Feature extraction
 - Hand-crafted linguistic features: specific characteristic of a word or a text, such as length of words, number of connectives in an essay.
 - **Automatic features**: word embedding, such as word2vec (Mikolov et al., 2013)
 - Each word in the corpus is represented by a vector space of hundreds dimension by training a huge corpus such as Wikipedia.

Welten-instituut





Argumentation component identification model for Dutch



- Training Model
 - Neural network. It is the state-of-the-art model for the task of argumentation component identification (Eger et al., 2017).

- Evaluation
 - Creating a Dutch corpus with argumentation structure annotation.
 - Evaluating the performance of the model with human-annotated data.





Future work

Automated argumentation quality assessment. (Wachsmuth et al., 2017)

- Generating formative feedback
 - Formative feedback based on the result of the argument components identification.
 - Comments over the argumentation quality.





Task

- Your opinions over formative feedback in terms of difficulty and time.
- Your opinions over formative feedback generated based on automated argument component identification tool.

Welten-instituut Onderzoekscentrum voor leren, doceren en technologie



Reference

- Eger, S., Daxenberger, J., Stab, C., & Gurevych, I. (2017). Cross-lingual Argumentation Mining: Machine Translation (and a bit of Projection) is All You Need!, (1).
- Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). Efficient estimation of word representations in vector space. arXiv preprint arXiv:1301.3781.
- Miller, A., Cox, K., & Imrie, B. W. (2014). Student assessment in higher education: a handbook
- Shute, V. J. (2008). Focus on Formative Feedback. Review of Educational Research, 78(1), 153–189.
- Smith, E., & Gorard, S. (2005). "They don"t give us our marks": The role of formative feedback in student progress. Assessment in Education: Principles, Policy & Practice, 12(1), 21–38.
- Stab, C., & Gurevych, I. (2017). Parsing argumentation structures in persuasive essays. Computational Linguistics, 43(3), 619-659.
- Wachsmuth, H., Naderi, N., Hou, Y., Bilu, Y., Prabhakaran, V., Thijmm, T. A., ... Stein, B. (2017). Computational Argumentation Quality Assessment in Natural Language. Eacl2017, 1, to appear. https://doi.org/10.18653/v1/E17-1017

Welten-instituut



Questions

Welten-instituut Onderzoekscentrum voor leren, doceren en technologie



Formative feedback

Formative feedback type	Difficulty	Time-consuming
Correct errors or misconception	5.5	5.8
Point to the omission of relevant material(s)	4.3	4.9
Point to the inclusion of irrelevant material(s)	4.2	3.75
Comment on the argumentation	6.15	6.65
Comment on the use of language	3.25	4.55
Comment on structure of the essay	4.75	5.35
Encourage dialogue	3.75	4.5
Point to additional resource materials	4.4	4.4
Praise the student	3.1	3.35

Welten-instituut

