The Impact of
Modeling Overall Argumentation with Tree Kernels

Henning Wachsmuth Bauhaus-Universität Weimar

Giovanni Da San Martino Qatar Computing Research Institute

Dora Kiesel Bauhaus-Universität Weimar

Benno Stein Bauhaus-Universität Weimar

Bauhaus-Universität Weimar

www.webis.de
henning.wachsmuth@uni-weimar.de

Overall argumentation of monological texts

| The death penalty is a legal means that as such is not practicable in Germany. | For one thing, inviolable human dignity is anchored in our constitution, | and further no one may have the right to adjudicate upon the death of another human being. | Even if many people think that a murderer has already decided on the life or death of another person, | this is precisely the crime that we should not repay with the same. |

(Peldszus and Stede, 2016)
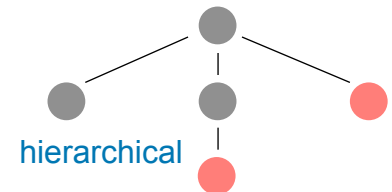
Why and how to model overall argumentation?

- **Hypothesis**
  The overall structure of argumentative texts is decisive for downstream analysis tasks of computational argumentation.
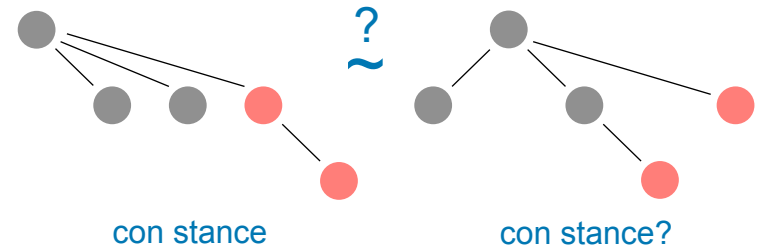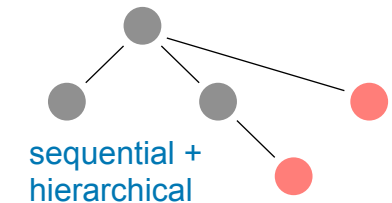

- **Existing work**
  - Some models for analyzing sequential overall structure
    (Persing et al., 2010; Wachsmuth and Stein, 2017)
 - No model for analyzing hierarchical overall structure

- **Research questions**
  1. How to jointly model sequential and hierarchical overall argumentation?
 2. What model has most impact in downstream tasks?

Downstream tasks on three argument corpora

- **Myside bias on AAE-v2** (Stab and Gurevych, 2016)
  - 402 persuasive student essays
  - 15.1 units per text, proprietary argument model
  - 251 one-sided, 151 two-sided
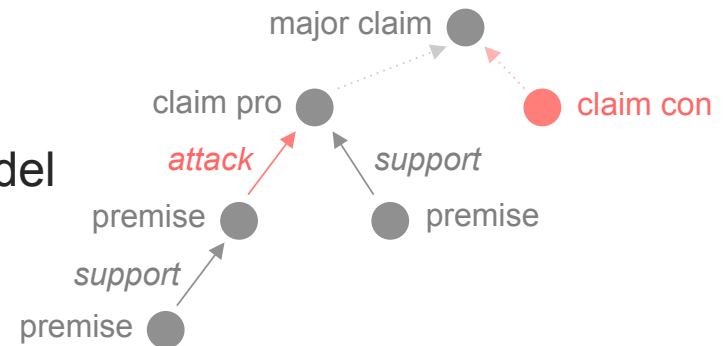
- **Stance on Arg-Microtexts** (Peldszus and Stede, 2016)
  - 112 short argumentative texts
  - 5.1 units per text, model of Freeman (2011)
  - 46 pro stance, 42 con stance, 24 unlabeled
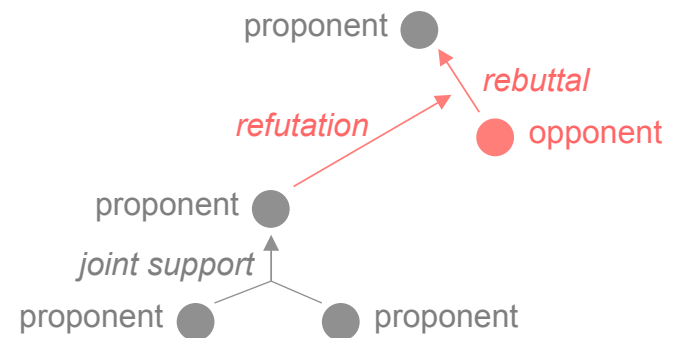
- **Genre on Web Discourse** (Habernal and Gurevych, 2015)
  - 340 argumentative web texts
  - 3.4 units per text, modified model of Toulmin (1958)
  - 216 comments, 46 blog posts, 73 forum posts, 5 articles
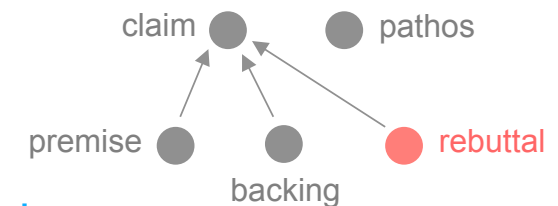
A general argument model

Map specific models to general model
- Nodes ordered according to position
- Node type encodes stance towards parent
- Relations between pairs of nodes only
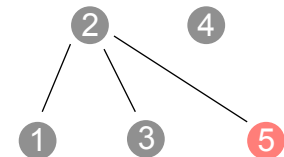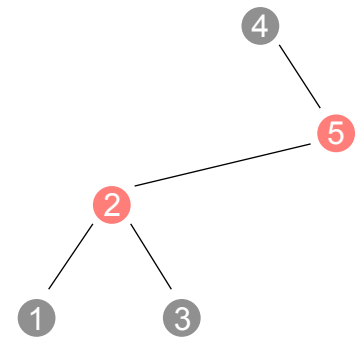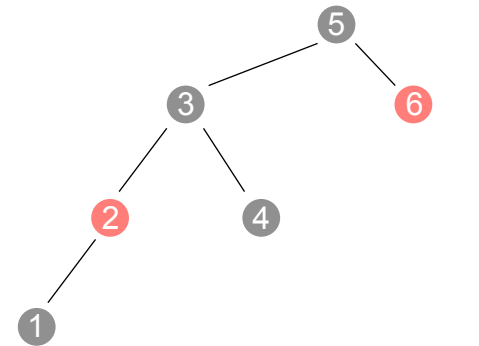- Root implicitly defines main claim

Pros and cons
- $+$ Sequential structure captured
- $+$ Same analyses on all corpora
- $+$ Comparisons across corpora
- $+$ Simpler argument mining (hypothesized)
- $-$ Partly less expressive

In this talk, only general model
- In the paper, also experiments for specific models
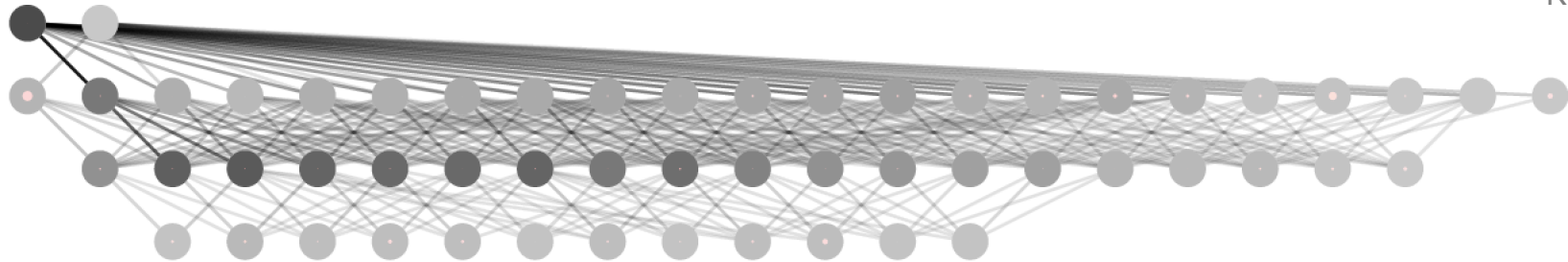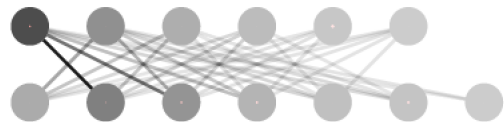
sequential structure

Dora
Kiesel

- **Myside bias on AAE-v2**   one-sided

- **Stance on Arg-Microtexts**   pro stance
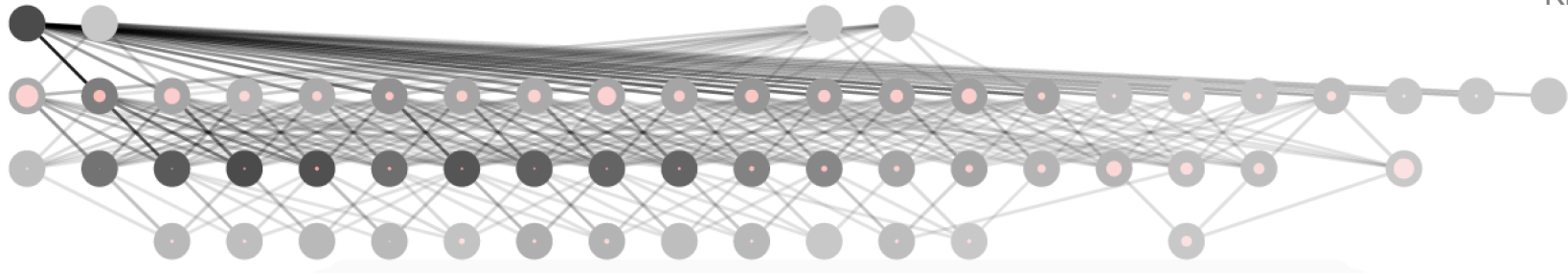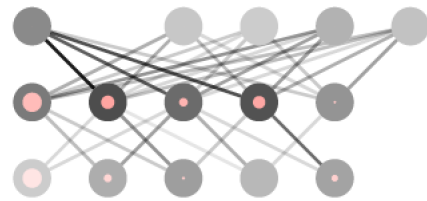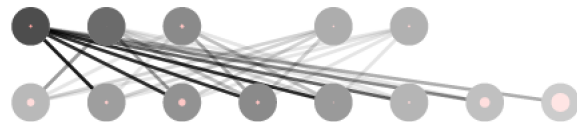

- **Genre on Web Discourse**   comments

Dora Kiesel

- **Myside bias on AAE-v2**   two-sided

- **Stance on Arg-Microtexts**   con stance

- **Genre on Web Discourse**   blog posts

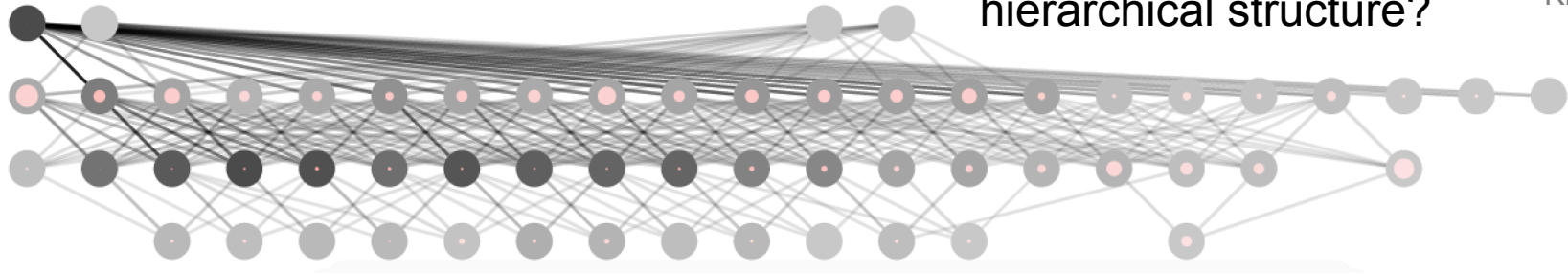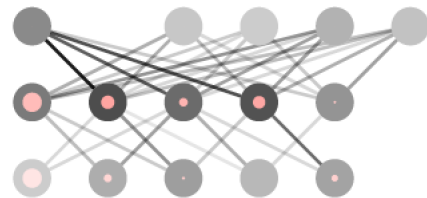Dora Kiesel

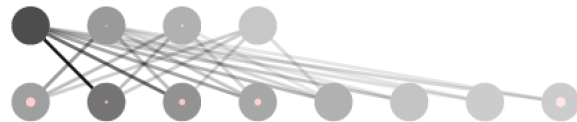- **Myside bias on AAE-v2** two-sided   → High impact of modeling hierarchical structure?

- **Stance on Arg-Microtexts** con stance   → Medium impact of modeling both types of structure?

- **Genre on Web Discourse** forum posts   → Low impact of modeling sequential structure if any?

Route kernels for modeling overall argumentation

Giovanni
Da San Martino

- **Kernel methods in machine learning**
  - Representation of instances in implicit feature space
  - Similarity function used by classifier (e.g., SVM)
  - Strong when good features unknown and data limited

- **Kernels for structured data**
  - Subsequence kernel for sequential structures
    (Mooney and Bunescu, 2006)
  - Tree kernels for hierarchical structures
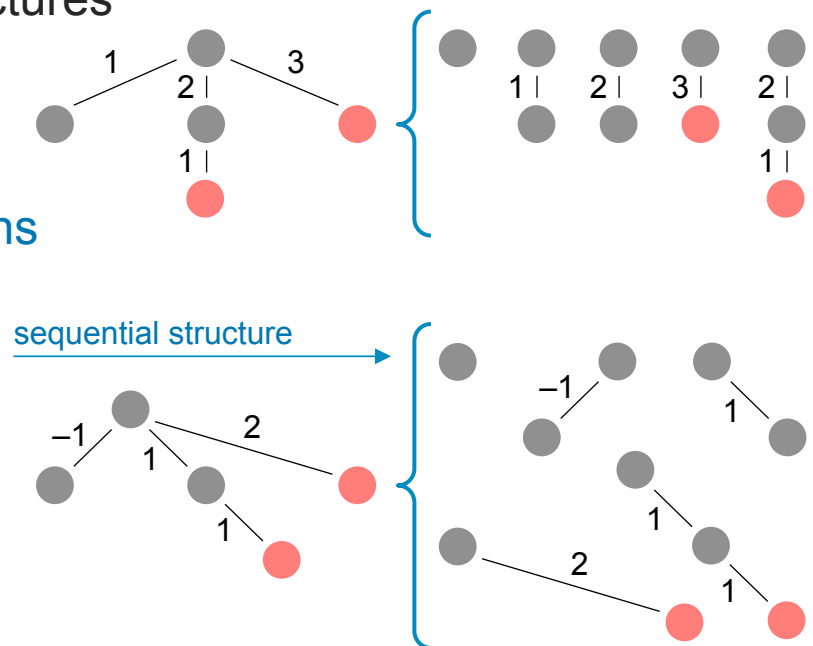   (Collins and Duffy, 2001; Moschitti, 2006)
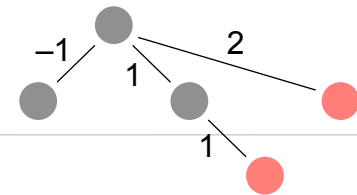 - Route kernels: Tree kernels with positions
    (Aiolli et al., 2009)

- **Route kernels for overall argumentation**
  - All paths starting from root
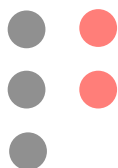 - Polynomial kernel "combines" paths
  - Positions relative to parent node

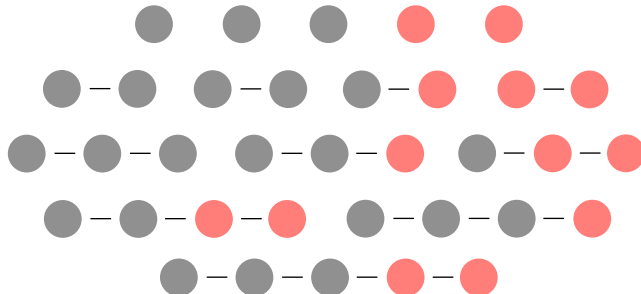# Experiments for each downstream task

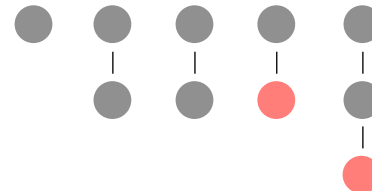- **Overall argumentation approaches**

frequencies	sequences	hierarchies	routes
linear kernel	subsequence kernel	tree path kernel	adapted route kernel
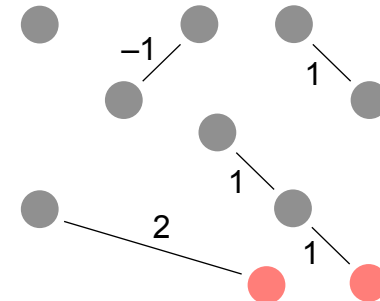
- **Baseline approaches**

majority	pos	tokens
always majority class	linear kernel	linear kernel

46 **pro stance**
42 ~~con stance~~

part-of-speech
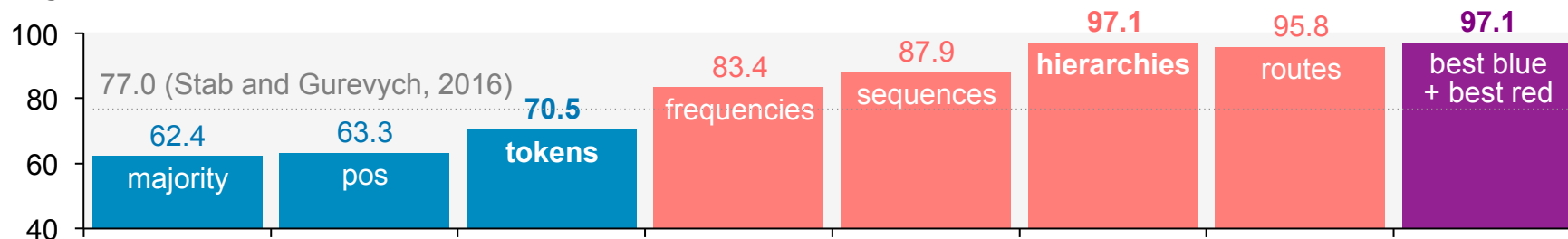1-, 2-, and 3-grams

token
1-, 2-, and 3-grams
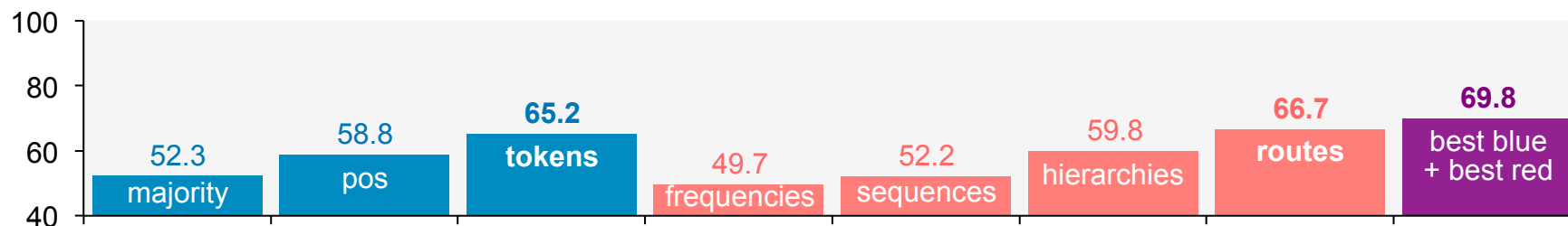
- **Experiments on ground-truth argument corpora**
  - SVM for each kernel in repeated 10-fold cross-validation
  - Hyperparameter tuning, fairness in training
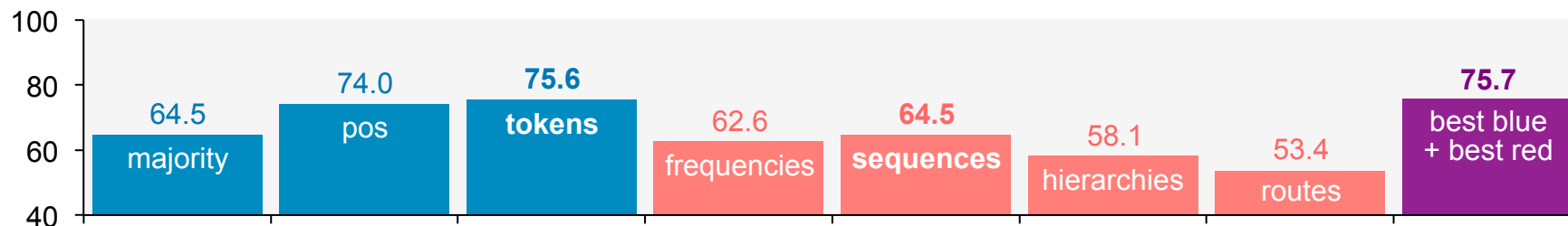
Accuracy results based on general model

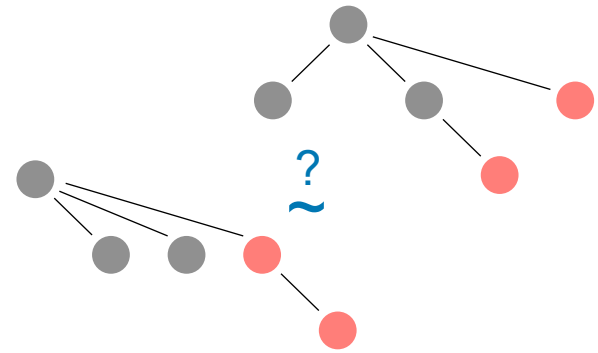**Myside bias on AAE-v2**

Stance on Arg-Microtexts

Genre on Web Discourse

Take aways

- **Route kernels** allow for analyzing sequential and hierarchical overall structure of argumentative texts jointly.

- **A general argument model** is almost competitive to Freeman's and Toulmin's model (see paper).

- **Modeling overall argumentation** benefits downstream tasks of computational argumentation, but the required type of structure varies.

- **Myside bias** is decided by hierarchical structure.

- **Stance** is affected by both types of structure.

- **The impact of mining errors** on the benefit of the modeling remains to be explored.

References

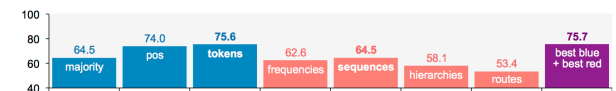- **Aiolli et al. (2009).** Fabio Aiolli, Giovanni Da San Martino, and Alessandro Sperduti. 2009. Route kernels for trees. In *Proceedings of the 26th Annual International Conference on Machine Learning*, pages 17–24.

- **Collins and Duffy (2001).** Michael Collins and Nigel Duffy. 2001. Convolution kernels for natural language. In *Advances in Neural Information Processing Systems 14*, pages 625–632.

- **Habernal and Gurevych (2015).** Exploiting debate portals for semi-supervised argumentation mining in user-generated web discourse. In *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing*, pages 2127– 2137.

- **Mooney and Bunescu (2006).** Raymond J. Mooney and Razvan C. Bunescu. 2006. Subsequence kernels for relation extraction. In *Advances in Neural Information Processing Systems 18*, pages 171–178. MIT Press.

- **Moschitti (2016).** Alessandro Moschitti. 2006. Efficient convolution kernels for dependency and constituent syntactic trees. In *Proceedings of the 17th European Conference on Machine Learning*, pages 318–329.

- **Peldszus and Stede (2016).** Andreas Peldszus and Manfred Stede. 2016. An annotated corpus of argumentative microtexts. In *Argumentation and Reasoned Action: 1st European Conference on Argumentation*.

- **Persing et. al. (2010).** Isaac Persing, Alan Davis, and Vincent Ng. 2010. Modeling organization in student essays. In *Proceedings of the 2010 Conference on Empirical Methods in Natural Language Processing*, pages 229–239.

- **Stab and Gurevych (2016).** Christian Stab and Iryna Gurevych. 2016. Recognizing the absence of opposing arguments in persuasive essays. In *Proceedings of the Third Workshop on Argument Mining (ArgMining2016)*, pages 113–118.

- **Wachsmuth and Stein (2017).** Henning Wachsmuth and Benno Stein. 2017. A universal model for discourse-level argumentation analysis. *Special Section of the ACM Transactions on Internet Technology: Argumentation in Social Media*, 17(3):28:1–28:24.