

The Interplay of Emotions and Convincingness in Arguments

Lynn Greschner and Roman Klinger

Fundamentals of Natural Language Processing, University of Bamberg, Germany
{lynn.greschner, roman.klinger}@uni-bamberg.de

Argumentation, with the main goal of persuasion (Habernal and Gurevych, 2016a), is a central mechanism of human communication, where the argument quality depends on the logical structure (logos), speaker credibility (ethos), emotional appeal (pathos), and contextual relevance (kairos) (Aristotle, 1991; Schiappa and Nordin, 2013). While there is a substantial amount of research on the overall argument quality and convincingness (CVC) (Habernal and Gurevych, 2016b; Gleize et al., 2019; Wachsmuth et al., 2017; Lauscher et al., 2020), the emotional dimension of arguments remains underexplored in the natural language processing community, in parts due to the challenges emerging from the contextualized and implicit nature of capturing emotions in text (Casel et al., 2021; Klinger et al., 2018; Koga et al., 2024; Lee and Lau, 2020). This dissertation addresses this gap through interrelated studies that together establish a perspectivist, emotion-aware computational framework for subjective argument evaluation.

We introduce the contextualized argument appraisal framework, a novel theoretical framework that models how humans cognitively appraise arguments and simultaneously develop emotion and CVC assessments. We conduct a role-playing annotation study and obtain a corpus (CONTARGA) of 4000 individual argument evaluations for 800 arguments (5 annotations per argument). Analyzing this corpus reveals that positive emotions (e.g., *trust*, *joy*) correlate with high, negative emotions (e.g., *anger*, *fear*) with low CVC. The appraisal dimensions (adapted for argument evaluation) further reveal that perceiving arguments as familiar and pleasant correlates with high, perceiving them as norm-violating and cognitively demanding with low CVC. While the argument text is the primary driver of the emotion response of participants, emotion-specific patterns emerge regarding the relative contributions of sender, argument text, and receiver factors.

Building on CONTARGA, we investigate whether appraisal theories can computationally explain perceived CVC compared to categorical emotion models. Using zero-shot prompting experiments with 3 large language models, we find that incorporating emotion information consistently improves CVC predictions over a text-only baseline. Crucially, appraisal dimensions outperform categorical emotions. However, joint prediction of emotions, appraisals, and CVC consistently underperforms pipeline approaches, indicating that zero-shot prompting is insufficient to capture the complex bidirectional relationships in this task. This underscores the need for more capable appraisal and CVC predictors and motivates moving beyond zero-shot paradigms.

Following these results, we perform experiments to address improving the emotion and CVC predictions. We compare simple neural networks, encoder-only transformers, decoder-only models (both frozen and fine-tuned) on CVC and emotion classification. By plugging in individual gold-annotated features (emotions, appraisals, demographics, personality), we assess their respective contributions to CVC prediction. Preliminary results across all model types consistently show that *affective states* (emotions and appraisals) improve CVC prediction more than *traits* (demographics and personality). Fine-tuned models benefit more strongly from appraisal features than frozen ones, suggesting that learned representations can better integrate cognitive and emotion signals. These preliminary results reinforce the core thesis: fine-grained, contextualized emotion modeling is essential for robust and human-aligned CVC predictions. In the next step, it is crucial to further investigate CVC signals and how to recover individual appraisal assessments, for instance, by probing the models for CVC and appraisal representations from text and other features, to investigate subjective argument evaluation within context.

References

- Aristotle. 1991. *On Rhetoric: A Theory of Civic Discourse*, 2nd edition. Oxford University Press, New York.
- Felix Casel, Amelie Heindl, and Roman Klinger. 2021. [Emotion recognition under consideration of the emotion component process model](#). In *Proceedings of the 17th Conference on Natural Language Processing (KONVENS 2021)*, pages 49–61, Düsseldorf, Germany. KONVENS 2021 Organizers.
- Martin Gleize, Eyal Shnarch, Leshem Choshen, Lena Dankin, Guy Moshkovich, Ranit Aharonov, and Noam Slonim. 2019. [Are you convinced? choosing the more convincing evidence with a Siamese network](#). In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, pages 967–976, Florence, Italy. Association for Computational Linguistics.
- Ivan Habernal and Iryna Gurevych. 2016a. [What makes a convincing argument? empirical analysis and detecting attributes of convincingness in web argumentation](#). In *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing*, pages 1214–1223, Austin, Texas. Association for Computational Linguistics.
- Ivan Habernal and Iryna Gurevych. 2016b. [Which argument is more convincing? analyzing and predicting convincingness of web arguments using bidirectional LSTM](#). In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 1589–1599, Berlin, Germany. Association for Computational Linguistics.
- Roman Klinger, Orphée De Clercq, Saif Mohammad, and Alexandra Balahur. 2018. [IEST: WASSA-2018 implicit emotions shared task](#). In *Proceedings of the 9th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis*, pages 31–42, Brussels, Belgium. Association for Computational Linguistics.
- Yurie Koga, Shunsuke Kando, and Yusuke Miyao. 2024. [Forecasting implicit emotions elicited in conversations](#). In *Proceedings of the 17th International Natural Language Generation Conference*, pages 145–152, Tokyo, Japan. Association for Computational Linguistics.
- Anne Lauscher, Lily Ng, Courtney Napoles, and Joel Tetreault. 2020. [Rhetoric, logic, and dialectic: Advancing theory-based argument quality assessment in natural language processing](#). In *Proceedings of the 28th International Conference on Computational Linguistics*, pages 4563–4574, Barcelona, Spain (Online). International Committee on Computational Linguistics.
- Sophia Yat Mei Lee and Helena Yan Ping Lau. 2020. [An event-comment social media corpus for implicit emotion analysis](#). In *Proceedings of the Twelfth Language Resources and Evaluation Conference*, pages 1633–1642, Marseille, France. European Language Resources Association.
- Edward Schiappa and John Nordin. 2013. *Argumentation: Keeping Faith with Reason*, 1st edition. Pearson UK, Harlow.
- Henning Wachsmuth, Nona Naderi, Yufang Hou, Yonatan Bilu, Vinodkumar Prabhakaran, Tim Alberdingk Thijm, Graeme Hirst, and Benno Stein. 2017. [Computational argumentation quality assessment in natural language](#). In *Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics: Volume 1, Long Papers*, pages 176–187, Valencia, Spain. Association for Computational Linguistics.