Strategies for Framing Argumentative Conclusion Generation

Philipp Heinisch[1], Anette Frank[2], Juri Opitz[2] and Philipp Cimiano[1]

[1]Bielefeld University, [2]Heidelberg University

Introduction& Task

Argument = premise → conclusion

Conclusions are often omitted in real-world online argumentation ⇒ need of reconstruction for downstream-tasks [2, 6]

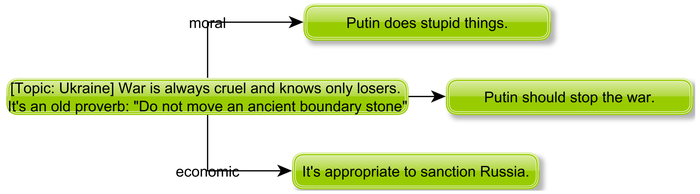


Fig. 1: Task: given a premise, generate an appropiate conclusion

Which conclusion of those three?

Conclusion generation is an underspecfied task: tailor the task with frame information (target aspect/ perspective) [8]

Datasets

Webis-Argument-Framing-19 [1]

- 12,326 arguments → 80% training, 10% dev, 5% conclusion-reranker-optimazation-split, 5% test
- each argument:
  - Topic+Premise text
  - conclusion text (target)
  - fine-grained issue-specific-frame (e.g. "women's rights" and "fetus rights" in topic *Abortion*; only 20% of these labels occur in at least two topics)

Media-Frames-Dataset [4]

- 17,826 annotated newspaper articles
- each article contains several:
  - marked text span
  - coarse-grained Media-Frame-class [3]: i) Economic, ii) Capacity and resources, iii) Morality, iv) Fairness and equality, v) Legality, constitutionality and jurisprudence, vi) Policy prescription and evaluation, vii) Crime and punishment, viii) Security and defense, ix) Health and safety, x) Quality of life, xi) Cultural identity, xii) Public opinion, xiii) Political, xiv) External regulation and reputation, as well as xv) Other
- use a subset with 21k text spans having an agreement by ≥ 2 annotators

Methods& Results...

Issue-specific-frames ↦ generic Media-Frame [5] (3 frame information pieces in total!)

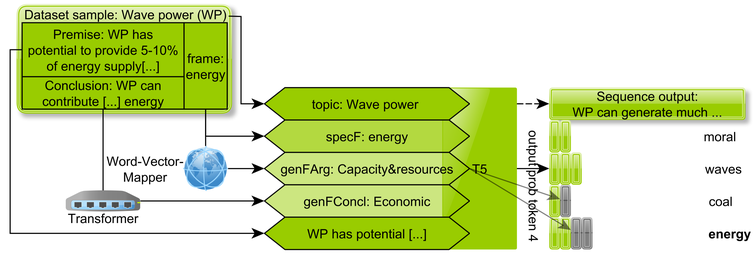
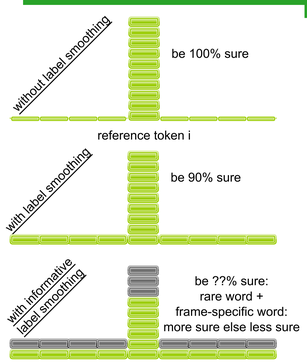


Fig. 2: Our inputs components with frame-sensitive decoding (frame-tailored term-frequencies by Media-Frames-Dataset)

Input	F1-BERTscore	Validity	Novelty	Both	frame-rel.
no frame	29.1	50%	50%	17%	67%
only specF	31.6	67%	37%	10%	**90%**
only genFArg	31.3	**73%**	37%	10%	87%
only genFConcl	31.6	67%	50%	13%	77%
specF+genFArg	30.4	40%	**63%**	7%	80%
specF+genFConcl	31.4	60%	47%	**20%**	77%
all 3 frames	33.8	70%	40%	10%	83%
reference	*100.0*	*73%*	*73%*	*47%*	*83%*

Fig. 3: Results with different frame inputs in addition to topic+premise. We performed a manual study to get the ratio of (generated) conclusions fulfilling validity, novelty, validity+novelty and relatedness to the issue-specific-frame (majority of 3 annotators/30 samples).

...adding informative label smoothing...



Fine-tuning: Regularization technique [7] advanced!

Preferring the conclusions with informative label smoothing to the ones without? preference-majorities - dispreference majorities

Input	BERTscore	Validity	Novelty	Both
no frame	29.4→30.6	+13%	+23%	+10%
only specF	31.6→33.8	+13%	+23%	+7%
only genFArg	31.3→31.9	-33%	+23%	tie
specF+genFArg	30.4→31.0	+3%	-27%	tie
specF+genFConcl	31.4→33.9	+30%	+13%	+13%

...adding conclusion reranking

Seq2Seq generates serval output sequences (beam search) ⇒ select best one ranked by reference-less metrics

(e.g.: BERTscore(premise, concl. candidate) scaled by factors trained on conclusion-reranker-optimazation-split

Input	BERTscore	Validity	Novelty	Both
no frame	30.6→34.6	+10%	-13%	-5%
only specF	33.8→36.5	+27%	-17%	+3%
only genFArg	31.9→34.4	+27%	-10%	tie
specF+genFArg	31.0→34.1	+20%	-20%	-3%
specF+genFConcl	33.9→**37.6**	+8%	-17%	-2%

Conclusion

Frame-guided conclusion generation results in better conclusions

✓ valid+novel: **from 17% to 40% with informative-label-smoothed specF+genFConcl**

Further takeaways:

1. better BERTscore (similarty to reference conclusion) ⇔ better conclusion (manual studies still needed)
2. trade-off between

   validity (tendency to copy premise parts)

   ⇔ novelty (risk something)

Acknowledgements

We are grateful to the anonymous reviewers for their valuable comments. This work has been funded by the DFG through the project ACCEPT as part of the Priority Program "Robust Argumentation Machines" (SPP1999).

References

[1] Yamen Ajjour et al. "Modeling Frames in Argumentation". In: *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*. Hong Kong, China: Association for Computational Linguistics, Nov. 2019, pp. 2915–2925. DOI: 10.18653/v1/D19-1290.

[2] Maria Becker, Siting Liang, and Anette Frank. "Reconstructing Implicit Knowledge with Language Models". In: *Proceedings of Deep Learning Inside Out (DeeLIO): The 2nd Workshop on Knowledge Extraction and Integration for Deep Learning Architectures*. Online: Association for Computational Linguistics, June 2021, pp. 11–24. DOI: 10.18653/v1/2021.deelio-1.2.

[3] Amber E. Boydstun et al. *Tracking the Development of Media Frames within and across Policy Issues*. Aug. 2014. DOI: 10.1184/R1/6473780.v1.

[4] Dallas Card et al. "The Media Frames Corpus: Annotations of Frames Across Issues". In: *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Volume 2: Short Papers)*. Beijing, China: Association for Computational Linguistics, July 2015, pp. 438–444. DOI: 10.3115/v1/P15-2072.

[5] Philipp Heinisch and Philipp Cimiano. "A multi-task approach to argument frame classification at variable granularity levels". In: *it - Information Technology* 63.1 (2021), pp. 59–72. DOI: 10.1515/itit-2020-0054.

[6] Shahbaz Syed et al. "Generating Informative Conclusions for Argumentative Texts". In: *Findings of the Association for Computational Linguistics: ACL-IJCNLP 2021*. Online: Association for Computational Linguistics, Aug. 2021, pp. 3482–3493. DOI: 10.18653/v1/2021.findings-acl.306.

[7] Christian Szegedy et al. "Rethinking the Inception Architecture for Computer Vision". In: *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 2016, pp. 2818–2826. DOI: 10.1109/CVPR.2016.308.

[8] Claes de Vreese. "News Framing: Theory and Typology". In: *Information Design Journal* 13 (Jan. 2005), pp. 51–62. DOI: 10.1075/idjdd.13.1.06vre.