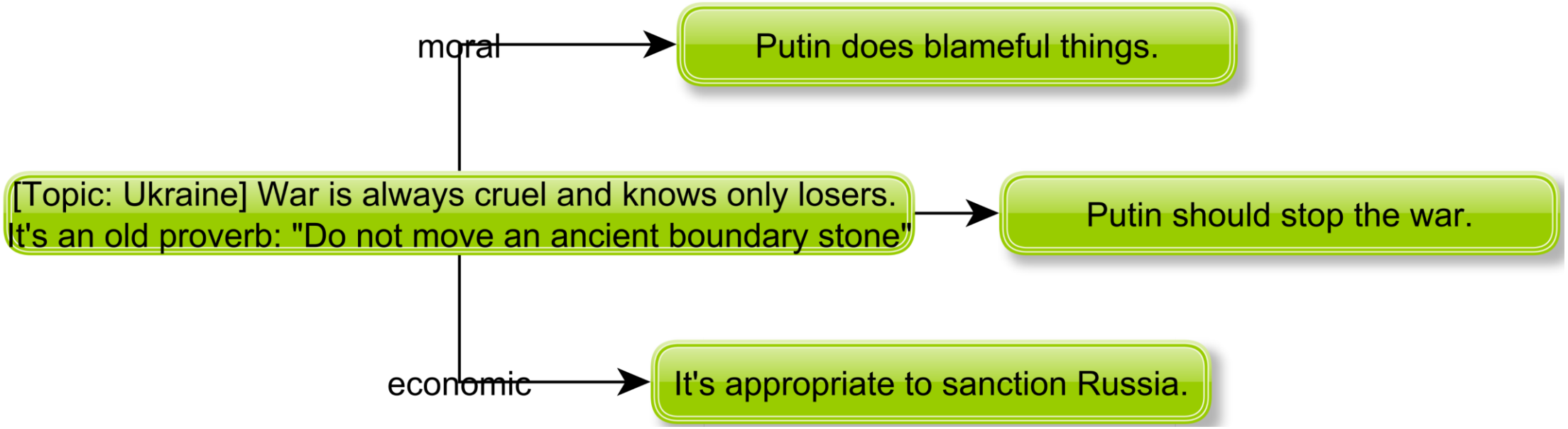
Philipp Heinisch, Anette Frank, Juri Opitz &
Philipp Cimiano

Strategies for Framing Argumentative Conclusion Generation

UNIVERSITÄT BIELEFELD

Conclusions are not unambiguous given a premise – target frame

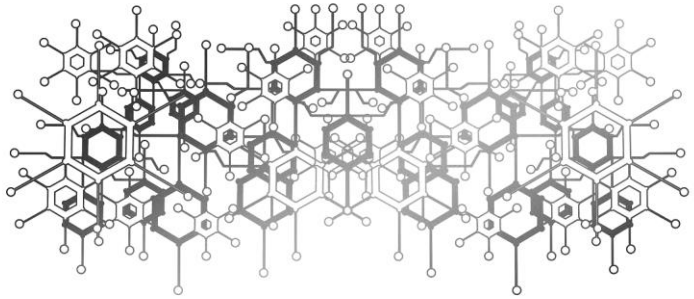


Task: generate conclusions with the help of frame information

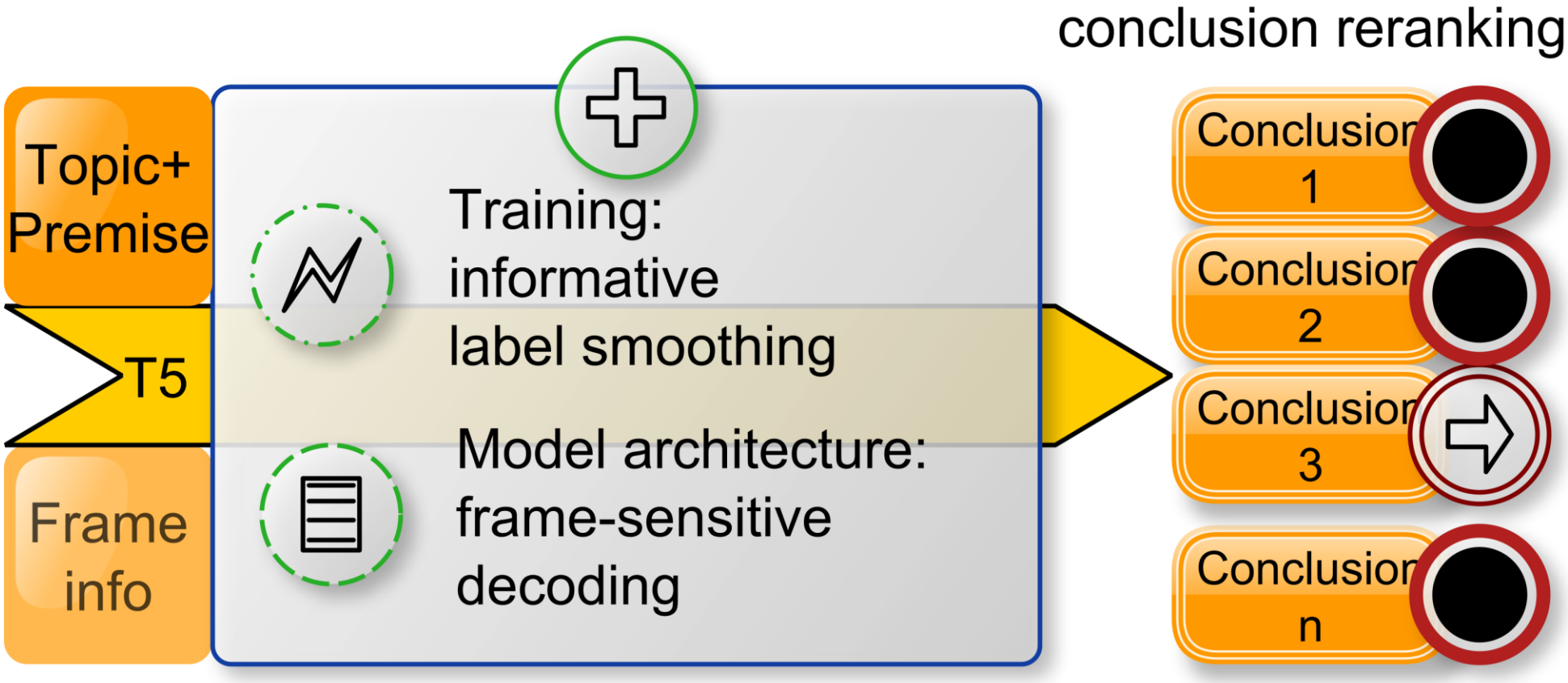
- Webis-Argument-Framing-19
  - 12k arguments (topic+ premise+ conclusion)
  - *Frame-info* for each argument
    - E.g.: Abortion -> *“woman's rights”*, *“fetus rights”*, *“safety”*
    - E.g.: Biofuel -> *“global warming”*, *“economics”*

Solving with

Seq2Seq-Language-Model



Overall strategies

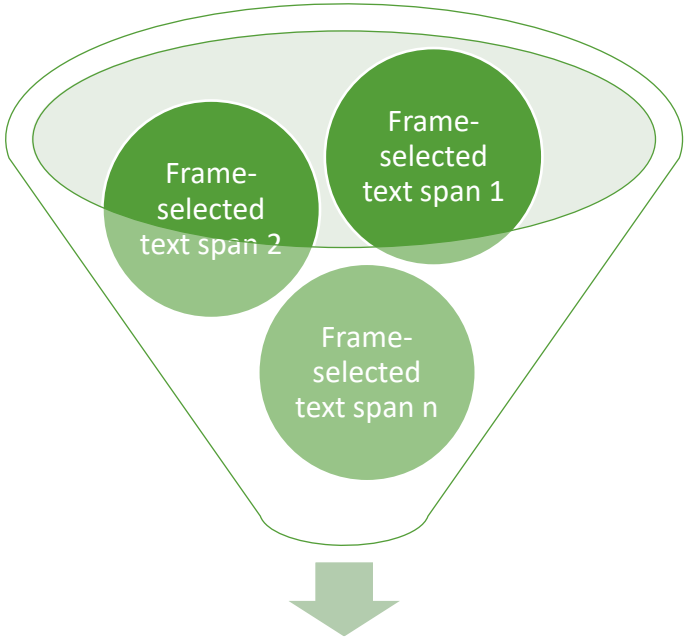


Generic frames (for informative label smoothing + frame-sensitive decoding)

Media-Frames

- Fixed set of 15 frames
  - Economic, Morality, Health and safety, …
- Media-Frames-Dataset
  - ~18k newspaper articles with span-frame-annotation

Frame-likely-words

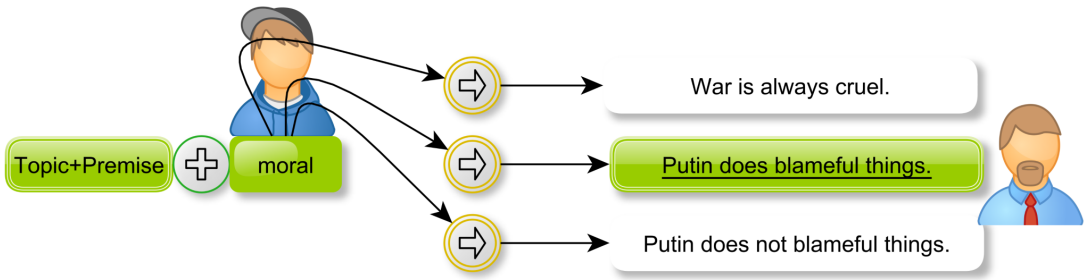


Term-frequency for each token/ frame
(e.g., $/economic = high correlation)

Conclusion reranking

Beam search in generation

=> several output sequences



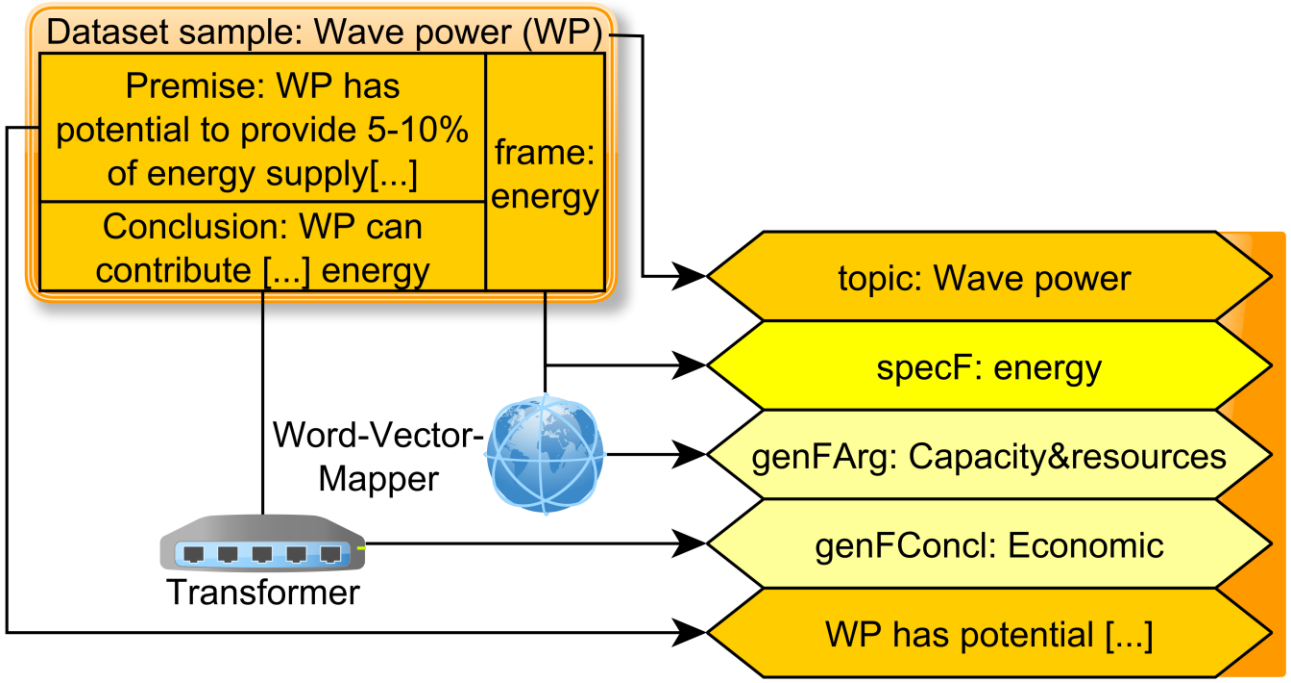
Conclusion reranking with reference-less metrics

1. BERTScore(prem, concl)
2. SupportAttackClassification (prem, concl)
3. Frame-likeliness(concl)
4. …

prem = premise/ concl = conclusion candidate

One problem to solve: different frame sets

- Fine- and course-grained frames
- Given in the Webis-ArgumentFraming-Dataset: sparse issue-specific-frames
- Needed for informative label smoothing + frame-sensitive decoding: fixed generic Media-Frames



Results – without informative label smoothing / conclusion reranking

Automatic (BERTscore(final selected conclusion, reference conclusion))

Input-Config	BERTscore
No frame	29.4
only specF	31.6
only genFArg	31.3
specF+genFArg	30.4
SpecF+genFConcl	31.4
All 3 frame infos	**33.8**
reference	100.0

Manual (majority of 3 annotators/ instance)

Valid	Novel	Both	Frame-rel.
50%	50%	17%	67%
67%	37%	10%	**90%**
73%	37%	10%	87%
40%	**63%**	7%	80%
60%	47%	**20%**	77%
70%	40%	10%	83%
73%	73%	47%	83%

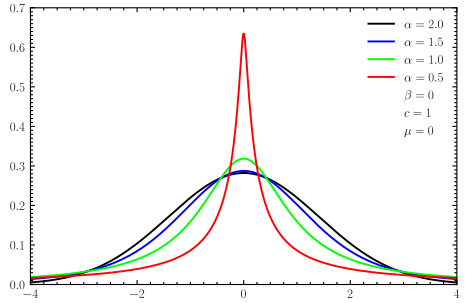
Results – with informative label smoothing

Automatic (BERTscore(final selected conclusion, reference conclusion))

Input-Config	BERTscore
No frame	29.4->30.6
only specF	31.6->33.8
only genFArg	31.3->31.9
specF+genFArg	30.4->31.0
SpecF+genFConcl	31.4->33.9

Do you favor the smoothed conclusion? (majority of 3 annotators/ instance)

Valid	Novel	Both
+13%	+23%	+10%
+13%	+23%	+7%
-33%	+23%	tie
+3%	-27%	tie
+30%	+13%	+13%



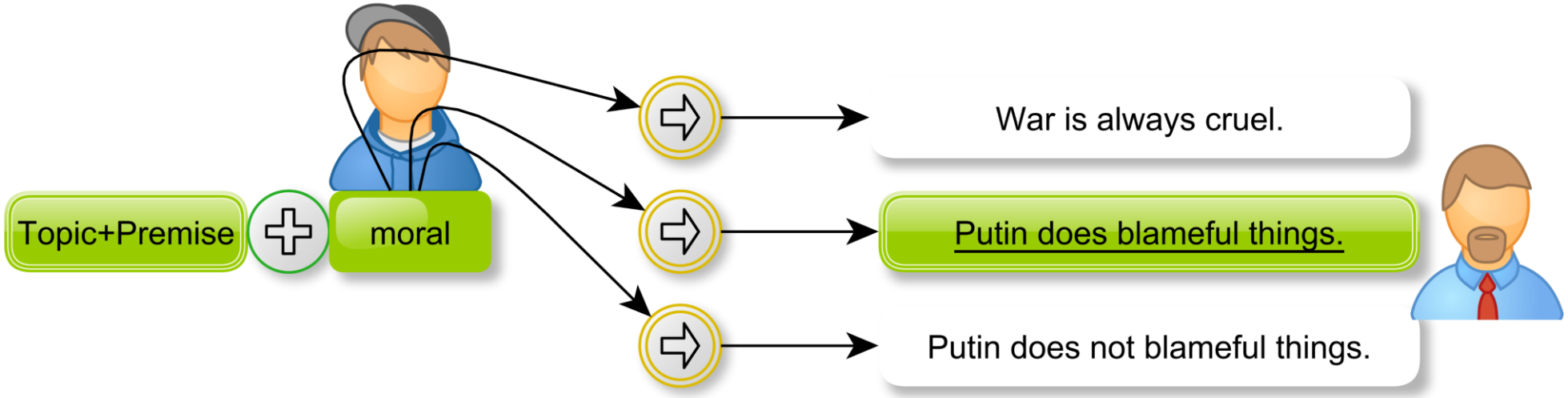
Results – with conclusion reranking

Automatic (BERTscore(final selected conclusion, reference conclusion))

Input-Config	BERTscore
No frame	30.6->34.6
only specF	33.8->36.5
only genFArg	31.9->34.4
specF+genFArg	31.0->34.1
SpecF+genFConcl	33.9->**37.6**

Do you favor the selected conclusion? (majority of 3 annotators/ instance)

Valid	Novel	Both
+10%	-13%	-5%
+27%	-17%	**+3%**
+27%	**-10%**	tie
+20%	-20%	-3%
+8%	-17%	-2%



Conclusion

- Real conclusion = valid + novel conclusion
- Frame information helps to fit better the reference conclusion (BERTscore: 29.4->37.6) and generate an appropriate conclusion as well (17% -> 40%)
  - Trade-off between validity and novelty
- Increasing BERTscore ≠ increasing conclusion quality (manual studies still necessary)