Efficient Argument Structure Extraction with Transfer Learning and Active Learning

Xinyu Hua[1] and Lu Wang[2]
[1]Bloomberg, [2]University of Michigan

Motivation

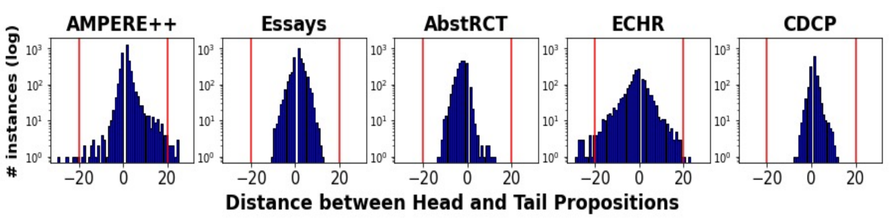
Argument Structure Extraction
- The goal is to identify the support/attack relations among propositions in a document.
- Annotation is costly, requires domain expertise (e.g., legal documents, scientific journals, policy making, etc.).
- Multiple datasets exist, but no unified framework that works across domains.

Data

AMPERE++ (new annotation)
- We label 400 reviews from `openreview.net`.
- Argument support/attack relations are labeled on top of our prior work's annotations [Hua+, 2019].
- More attack relations than existing datasets.

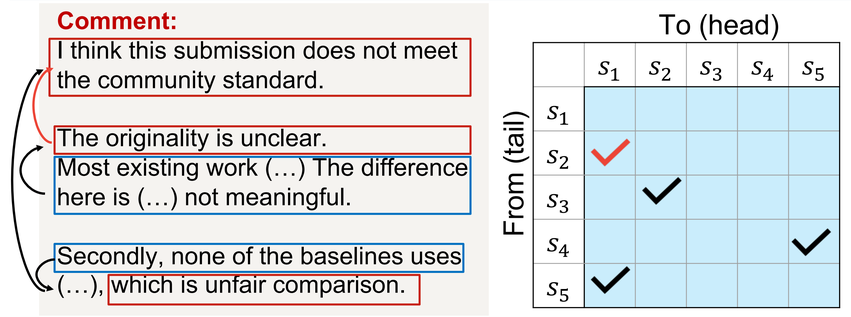
Other Datasets
- Essays [Stab & Gurevych, 2017]: 402 student essays
- AbstRCT [Mayer+, 2020]: 700 paper abstracts from PubMed
- ECHR [Poudyal+, 2020]: 42 case law documents from the European Court of Human Rights
- CDCP [Park & Cardie, 2018]: 731 online comments on the `eRulemaking` platform

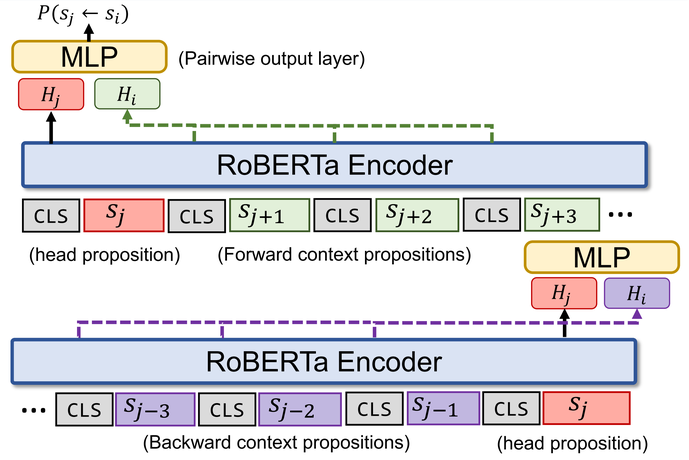
All datasets are annotated with support relations. Attack relations are labeled in AMPERE++, Essays, and AbstRCT.



- Most relations span less than 20 propositions.
- Review: verdicts first, support later.
- Paper abstract: evidence first, conclusion later.

Task and Model



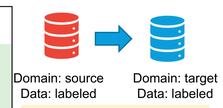


- Pairwise output with context-aware encoding
- Sliding window on forward/backward directions

	AMPERE++	Essays	AbstRCT	ECHR	CDCP
SVM-linear	24.82	28.69	33.60	21.18	29.01
SVM-RBF	26.38	31.68	32.65	21.36	30.34
SEQPAIR	23.40	38.37	**66.96**	13.76	35.23
OURS (head given)	**77.64**	**71.30**	63.62	**70.82**	**70.37**
OURS (end-to-end)	74.34	67.68	63.73	61.35	63.13

Transfer Learning

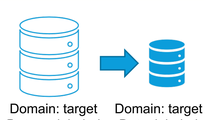
Source Domain \ Target Domain	AMPERE++	Essays	AbstRCT	ECHR	CDCP
AMPERE++		73.84	63.42	76.50	75.93
Essays	77.93		60.62	68.72	74.11
AbstRCT	76.29	71.17		73.31	69.17
ECHR	77.69	70.82	47.91		69.30
CDCP	77.87	68.37	62.38	72.03	



Transductive

Transfer learning can outperform standard supervised setting, with the same amount of target domain labels.

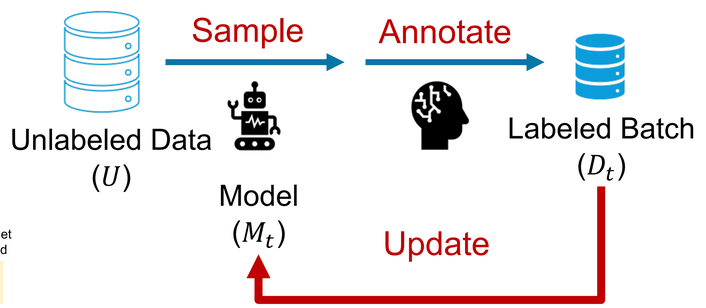
	AMPERE++	Essays	AbstRCT
MLM	78.10	74.21	64.48
Context-Pert	79.01	68.36	59.47



Inductive

MLM: masked language model with 15% masking probability. Context-Pert: context-aware sentence perturbation, (i) insert external sentence, (ii) shuffled, (iii) unchanged

Active Learning



Model-independent Acquisition Strategies
- NOVEL-VOCAB: encourages inclusion of unseen words
- DISC-MARKER: picks samples with discourse markers