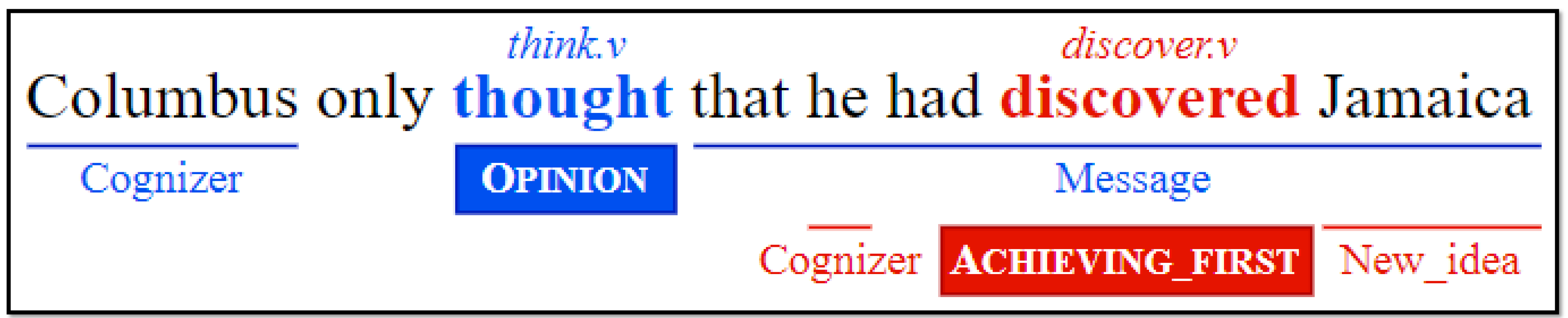
Robust Frame-Semantic Models with Lexical Unit Trees and Negative Samples

Jacob Devasier, Yogesh Gurjar, Chengkai Li

Department of Computer Science and Engineering
The University of Texas at Arlington
https://idir.uta.edu/

Frame-Semantic Parsing



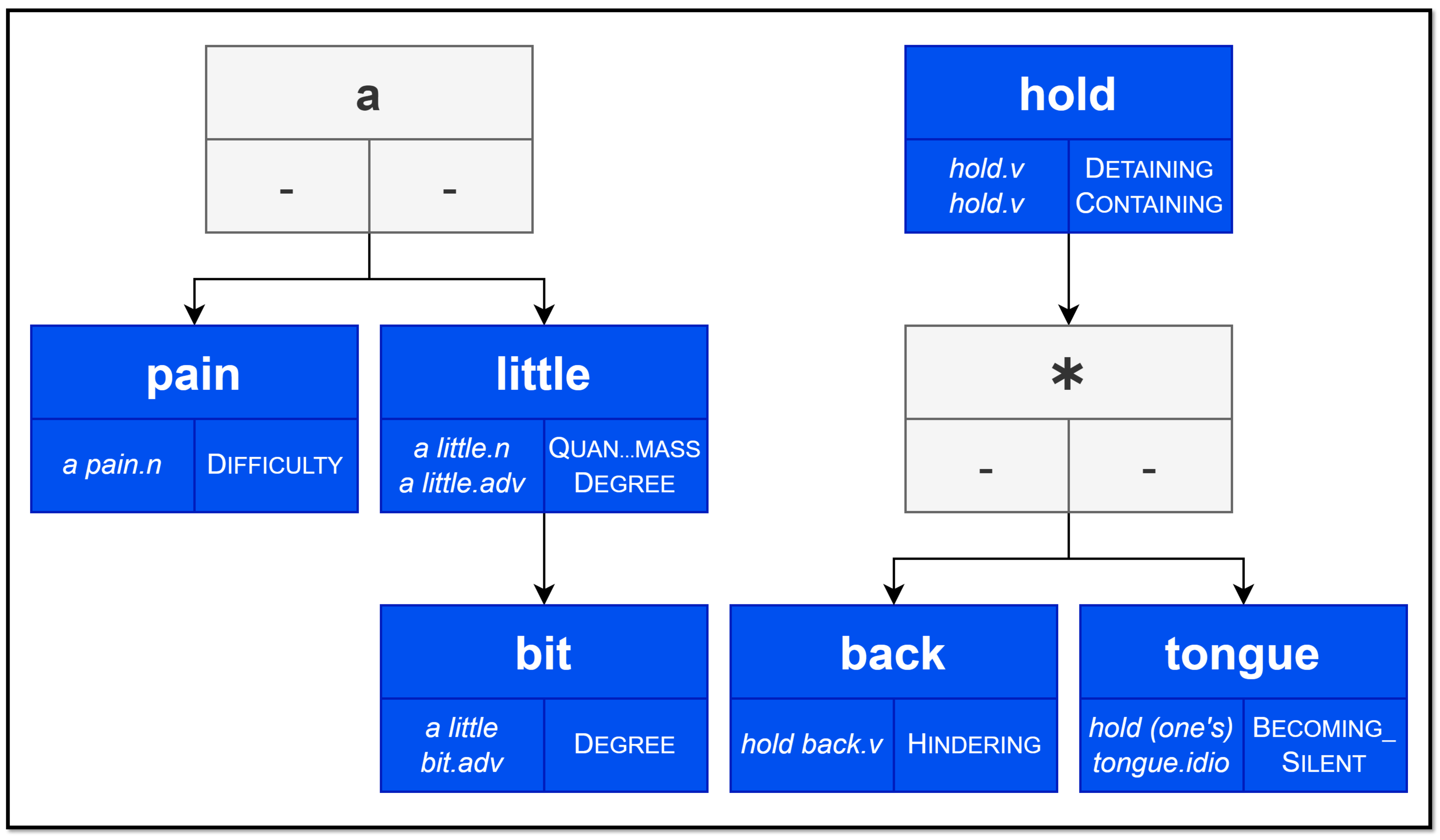
Terminology

- **Lexical Unit** – Pairing of a word and its meaning (*think.v*)
- **Frame** – Structures representing common situations (OPINION)
- **Target** – Words which evoke frames (**thought**)
- **Frame Element** – Key elements of a frame (Cognizer)

Dataset

- **FrameNet** – ~1,200 frames, ~6k sentences, ~19k targets
- **Test-1CF** – FrameNet samples with only 1 possible frame
- **Test-UU** – Test-1CF + 3 hard negative samples

Lexical Unit Tree



Candidate Target Generation

- Frames can only be evoked by certain words (Lexicon Filtering)
- Find frames for given word with lexical unit tree
- Wildcard supports disjoint lexical units using POS tags
- Covers **99.4%** of targets, **84.5%** false positive rate

Target Identification

Model	Acc	F1
Our model (candidate filter)	**0.788**	**0.775**
FIDO (Jiang and Riloff, 2021)	0.653	0.644
Our model	0.664	0.678

Model	FN1.5	FN1.7
Das et al. (2014)	0.454	-
Swayamdipta et al. (2017)	0.732	0.733
Bastianelli et al. (2020)	0.768	-
Lin et al. (2021)	0.769	0.763
Our model	**0.773**	**0.775**
Our model (manually filtered)	0.388	0.392

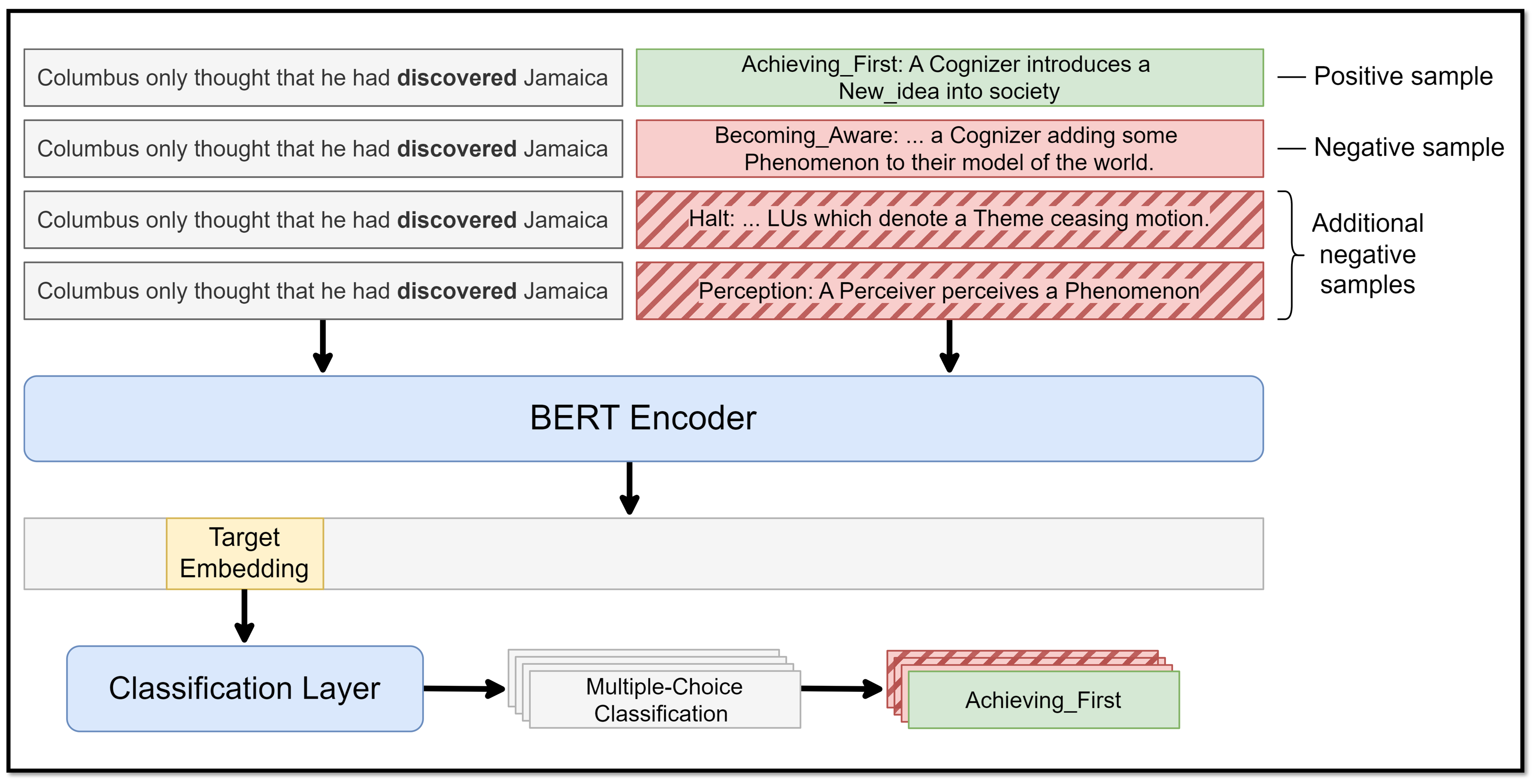
Frame Identification as Target Filtering (left)

- Can we directly replace target filtering with frame identification?
- **-9.7%** acc. vs target filter model, but **+3.4%** vs FIDO

Target Filtering (right)

- RoBERTa-based binary classification model
- **+1.2%** acc. vs SOTA model, **+38.3%** vs manual filtering

Frame Identification



K	# Frames	Our Model	FIDO	Δ
1	94	**0.781**	0.753	+0.028
3	235	**0.810**	0.778	+0.032
5	316	**0.853**	0.809	+0.044
10	426	**0.850**	0.826	+0.024

Model	Test-1CF	Test-UU
FIDO (Jiang and Riloff, 2021)	0.754	0.538
Our model	**0.893**	**0.603**

Additional Negative Sampling

- Improved performance on rare frames at all thresholds (left)
  - **+2.8%** acc. on 1-sample frames, **+4.4%** on 5-sample frames
- Enables learning on samples with only 1 possible frame
  - **+13.9%** acc. on Test-1CF, **+6.5%** on Test-UU
- Similar performance to best models while using less information
  - **+1.2%** acc. vs FIDO's frame-only model
  - **+0.2%** acc. vs full model
  - **-0.1%** acc. vs SOTA model

Model	FN1.5		FN1.7	
	All	Amb	All	Amb
Das et al. (2014)	0.836	0.692	-	-
Hermann et al. (2014)	0.887	0.737	-	-
Hartmann et al. (2017)	0.876	0.738	-	-
Yang and Mitchell (2017)	0.882	0.757	-	-
Swayamdipta et al. (2017)	0.864	-	0.866	-
Peng et al. (2018)	0.900	0.780	0.891	0.775
Bastianelli et al. (2020)	0.901	-	-	-
Lin et al. (2021)	0.906	-	0.906	-
Su et al. (2021)*	0.919	0.823	**0.924**	**0.844**
Tamburini (2022)*	**0.922**	**0.831**	0.922	0.843
Zheng et al. (2022a)	0.917	-	-	-
Jiang and Riloff (2021)	0.913	0.810	0.921	0.836
Jiang and Riloff (2021) (frame)	0.901	-	0.911	-
Our model (binary)	0.877	0.785	0.887	0.816
Our model	0.917	0.818	0.923	0.841

* Performance can not be verified due to private source code.

Contributions

- Novel lexical unit tree to enable support disjoint lexical units
- Developed bottom-up candidate target generation algorithm, leading to SOTA performance in target identification
- Evaluated effectiveness of language models for target filtering
- Enabled learning from single-frame targets in multiple-choice classification models
- Derived two new datasets from FrameNet for evaluating models on single-frame targets and similar lexical units

https://github.com/idirlab/frame

GitHub

Acknowledgements

This work is supported by the National Science Foundation under Grants 1719054, 1937143, and 2333834.

We would also like to extend our gratitude to the Texas Advanced Computing Center (TACC) for providing compute resources used in this work's experimentation.