UiO : University of Oslo

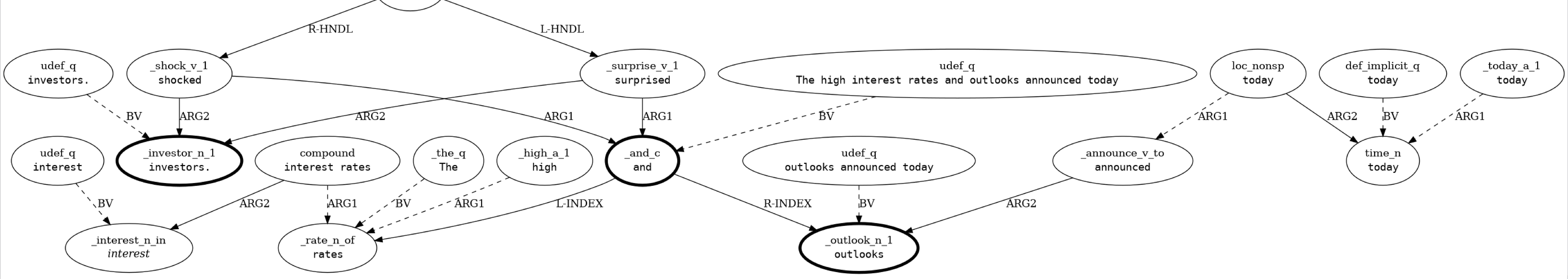
Argument Sharing in AMR, EDS & PTG

Maja Buljan **Lilja Øvrelid** **Stephan Oepen**

Language Technology Group, University of Oslo

EDS:

"The high interest rates and outlooks announced today surprised and shocked investors."



Introduction and motivation

In parsing into graph-based meaning representations, the widely accepted wisdom is that parsing accuracy deteriorates with growing structural complexity. However, the *why?* is still not clear.

Previous work shows a discrepancy in parser performance with growing graph complexity, between AMR & PTG vs. EDS.

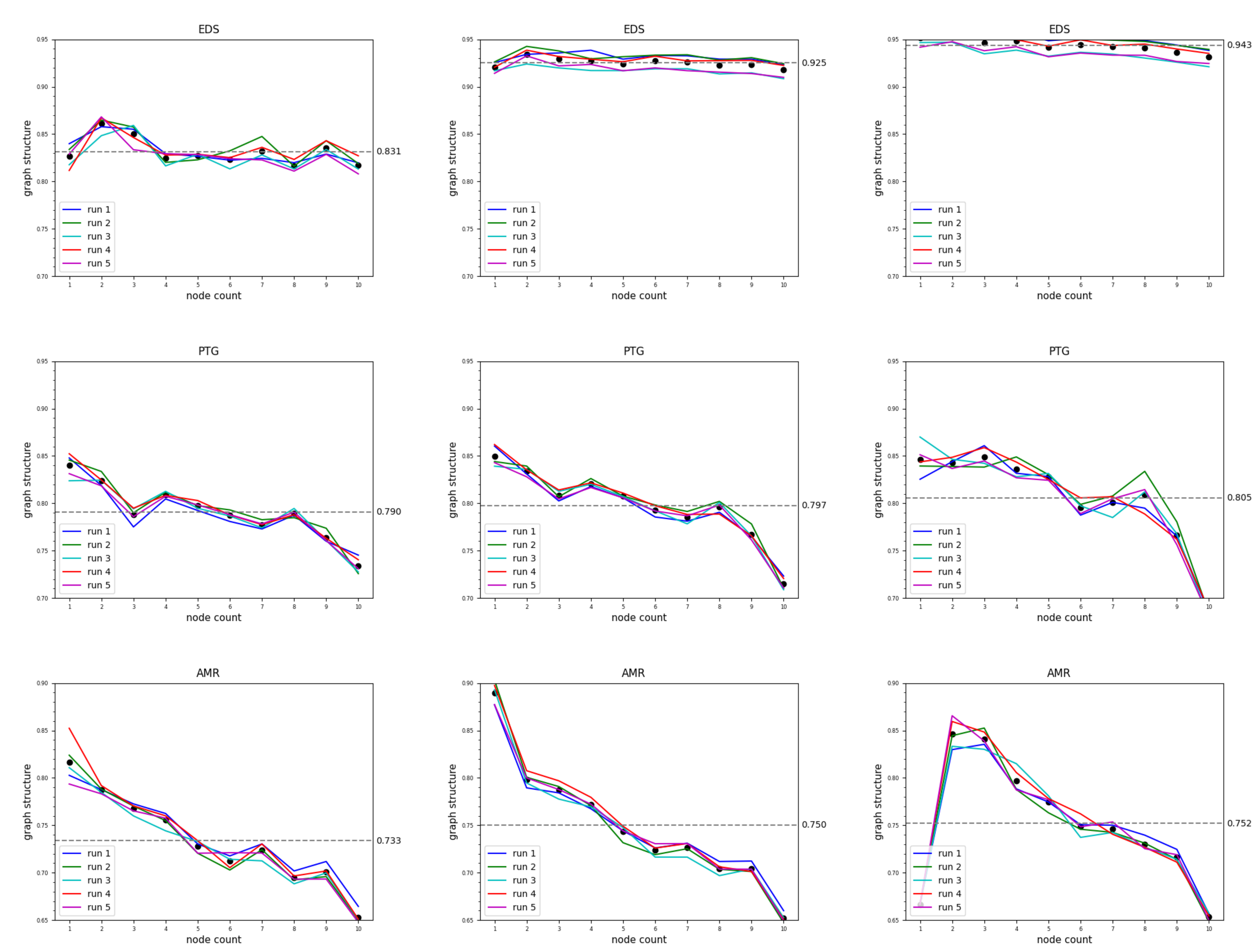
RQ1: Is this caused by shared arguments and reenrant nodes?

By expanding our understanding of framework design decisions, we can inform future annotation efforts, as well as parser development.

Graph structure and reentrancy

We approximate sentence-level graph complexity by number of nodes, group sentences into 10 decile bins, and compare F-scores for unlabelled graph structure contrasting by the reentrant status of edges.

tops + non-reentr. edges   tops + all edges   tops + reentr. edges



EDS parsing accuracy remains stable across complexity bins.

Reentrant edges actually seem easier to predict, in particular for EDS, and to an extent also for AMR.

Little difference for PTG, but slight improvement for less complex sentences when scoring reentrant edges only.

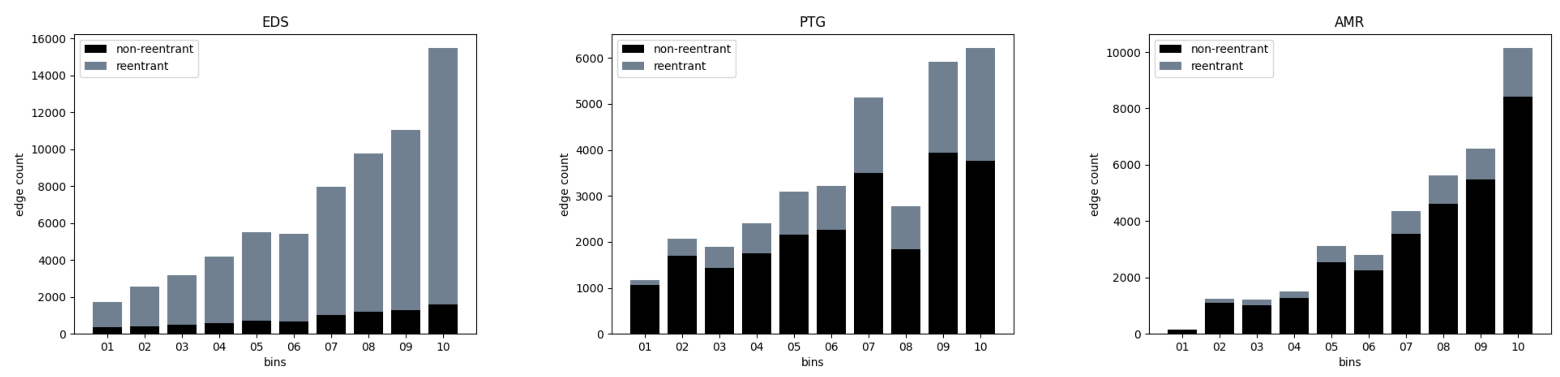
Annotation – 3 samples of 150 sentences

RQ2: Is the discrepancy due to the types of reentrancies?

Previous work on linguistic phenomena that give rise to reentrancies informs our pilot annotation and performance analysis.

Edges – reentrant and non-reentrant

There is a marked difference between the frameworks in proportions of non-reentrant vs. reentrant edges, including "predictable" ones.



Types of "non-predictable" reentrancies

	EDS		PTG		AMR	
Type	Freq.	Error	Freq.	Error	Freq.	Error
clause-like structures	.067	.058	-	-	.039	.428
comparative	.027	.000	-	-	-	-
control structures	.063	.250	**.136**	.291	**.106**	.342
coordination	**.127**	.125	**.165**	.258	**.134**	.166
coreference	.027	.142	**.401**	.382	**.243**	.436
modal	.011	.333	.014	.400	.008	.333
modification	**.167**	.071	**.176**	.451	-	-
named entity	-	-	.031	.818	-	-
partitive	-	-	-	-	.058	.285
possessive	.039	.500	.008	.666	.008	.666
relative clause	**.183**	.065	.022	.125	.008	.666
verbalisation	**.159**	.075	.002	.000	**.326**	.358
other	.122	.173	.039	.571	.067	.541

Findings and future work

AMR & PTG, coreference → frequent and hard to predict.
AMR, verbalisation→ frequent and non-uniform.
All three frameworks, possessives → high error rate.
EDS, lowest error rate → grammar-guided annotations.

Next steps: alternative approximations of graph complexity (reentrancies), expanded annotated dataset.

https://github.com/cfmrp/mtool

Maja Buljan, Joakim Nivre, Stephan Oepen, and Lilja Øvrelid. 2022. A tale of four parsers: methodological reflections on diagnostic evaluaion and in-depth error analysis for meaning representation parsing. *Language Resources and Evaluation*, 56(4):1075–1102.

Ryan McDonald and Joakim Nivre. 2011. Analyzing and integrating dependency parsers. *Computational Linguistics*, 37(1):197–230.

Dan Flickinger, Stephan Oepen, and Emily M. Bender. 2017. Sustainable development and refinement of complex linguistic annotations at scale. In *Handbook of Linguistic Annotation*, Springer.

Stephan Oepen, Omri Abend, Lasha Abzianidze, Johan Bos, Jan Hajic, Daniel Hershcovich, Bin Li, Tim O'Gorman, Nianwen Xue, and Daniel Zeman. 2020. MRP 2020: The second shared task on cross-framework and cross-lingual meaning representation parsing. *Proceedings of the CoNLL 2020 Shared Task*, ACL.

Ida Szubert, Marco Damonte, Shay B Cohen, and Mark Steedman. 2020. The role of reentrancies in abstract meaning representation parsing. In *Findings of the Association for Computational Linguistics: EMNLP 2020*

Rik Van Noord and Johannes Bos. 2017. Dealing with co-reference in neural semantic parsing. In *Proceedings of the 2nd Workshop on Semantic Deep Learning (SemDeep-2)*. ACL.