

Deep Multitask Learning for Semantic Dependency Parsing

Hao Peng¹, Sam Thomson², Noah A. Smith¹

¹University of Washington, ²Carnegie Mellon University
@Flair

Dec 20, 2017



Meaning Representations

Sentence

Meaning Representations

Truth conditions

$Take(Sh, Money) \wedge Run(Sh)$

Shareholders took their money and ran.

Meaning Representations

Sentence

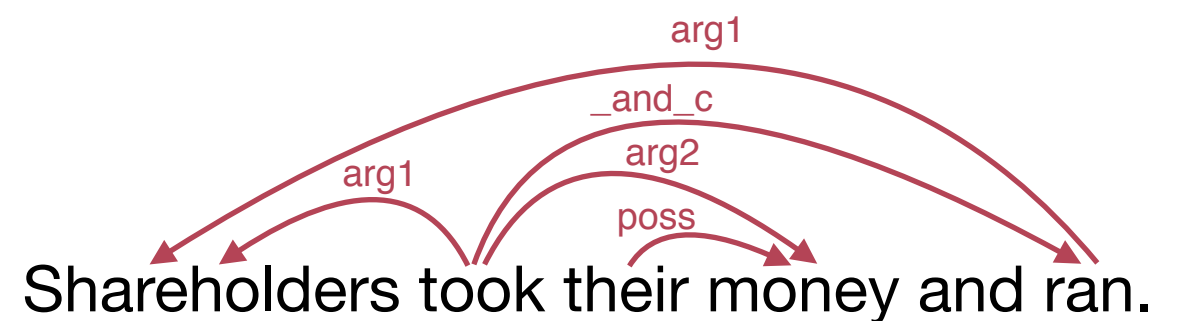
Meaning Representations

Truth conditions

$Take(Sh, Money) \wedge Run(Sh)$

Shareholders took their money and ran.

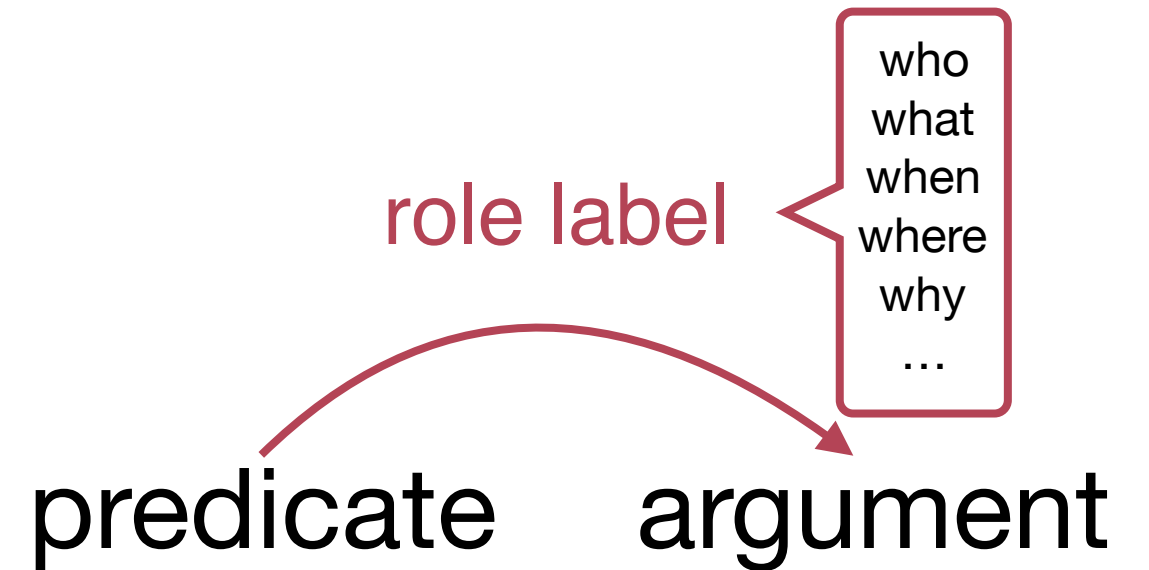
Predicate-argument



Semantic Parsing

Motivation

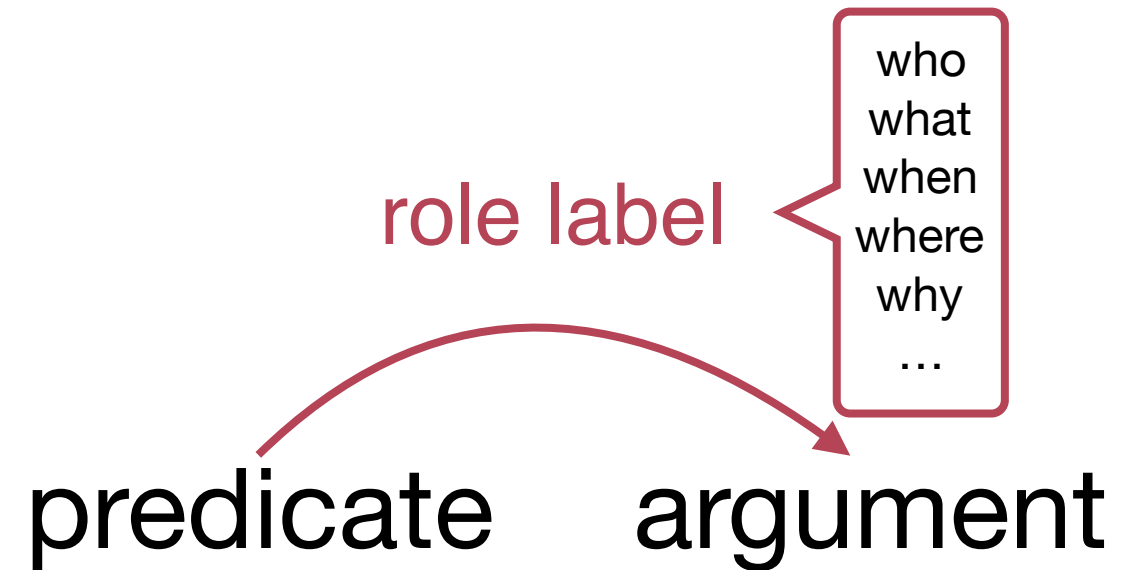
who did what to whom,
where and why...



Semantic Parsing

Motivation

who did what to whom,
where and why...

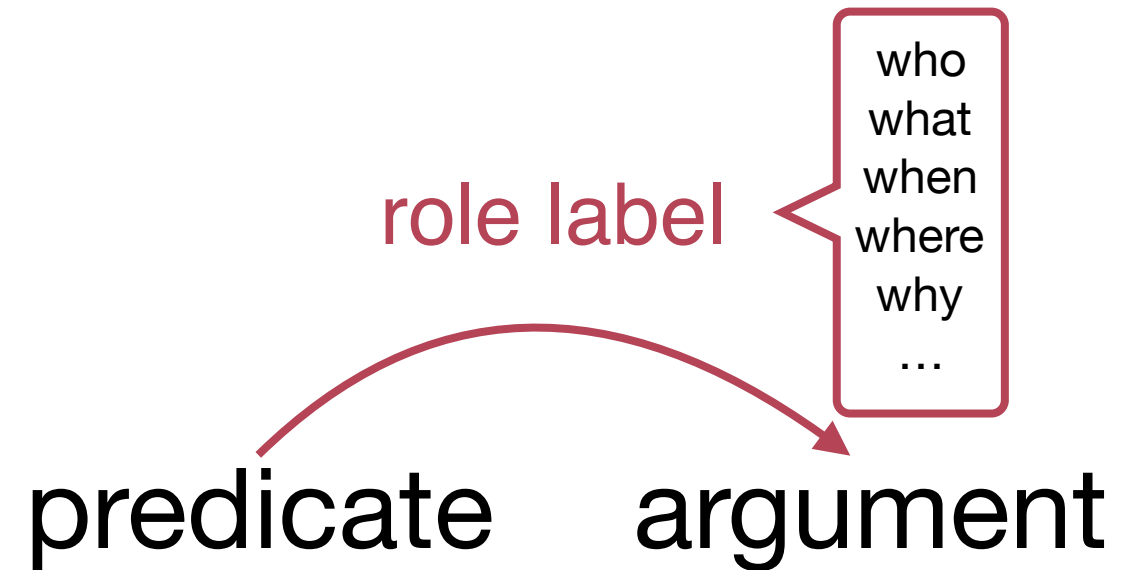


Shareholders took their money and ran .

Semantic Parsing

Motivation

who did what to whom,
where and why...

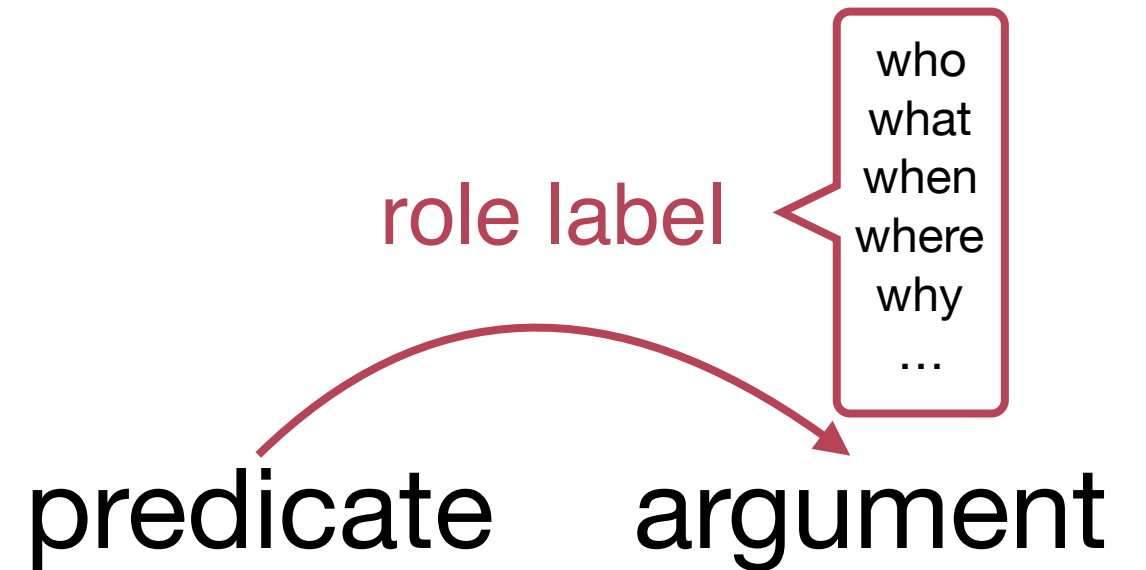


Shareholders took their money and ran .
did what

Semantic Parsing

Motivation

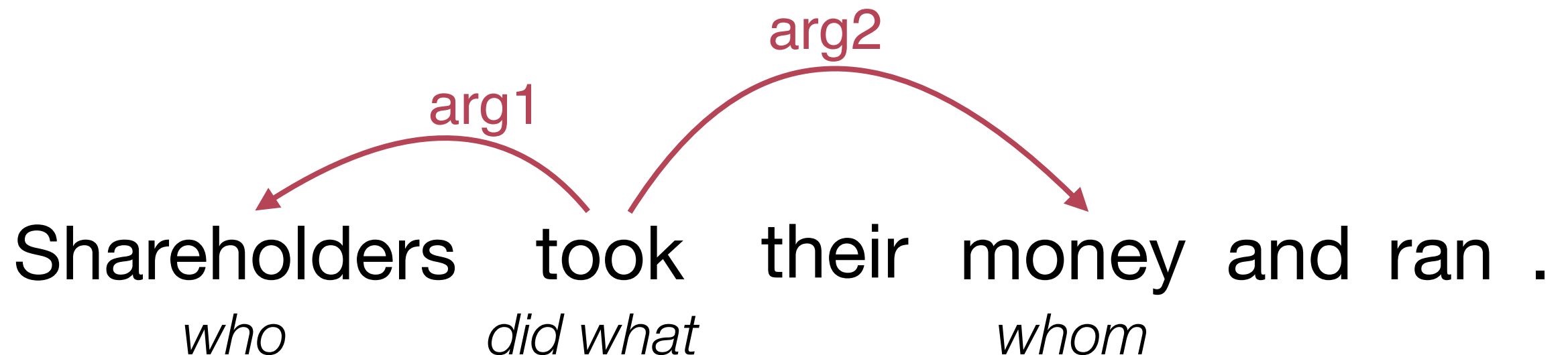
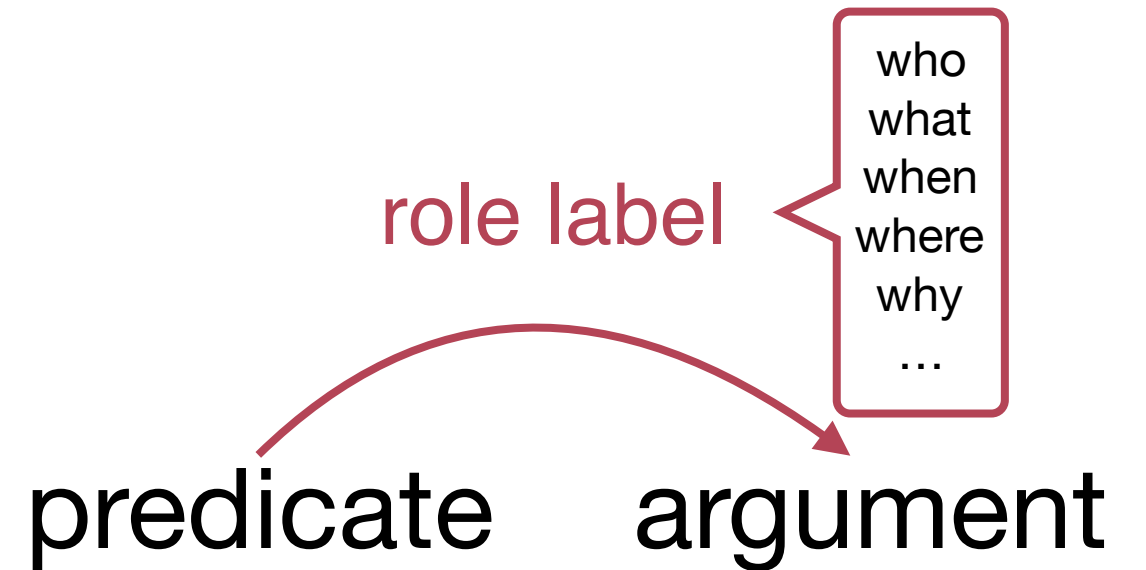
who did what to whom,
where and why...



Semantic Parsing

Motivation

who did what to whom,
where and why...



Semantic Parsing

Motivation

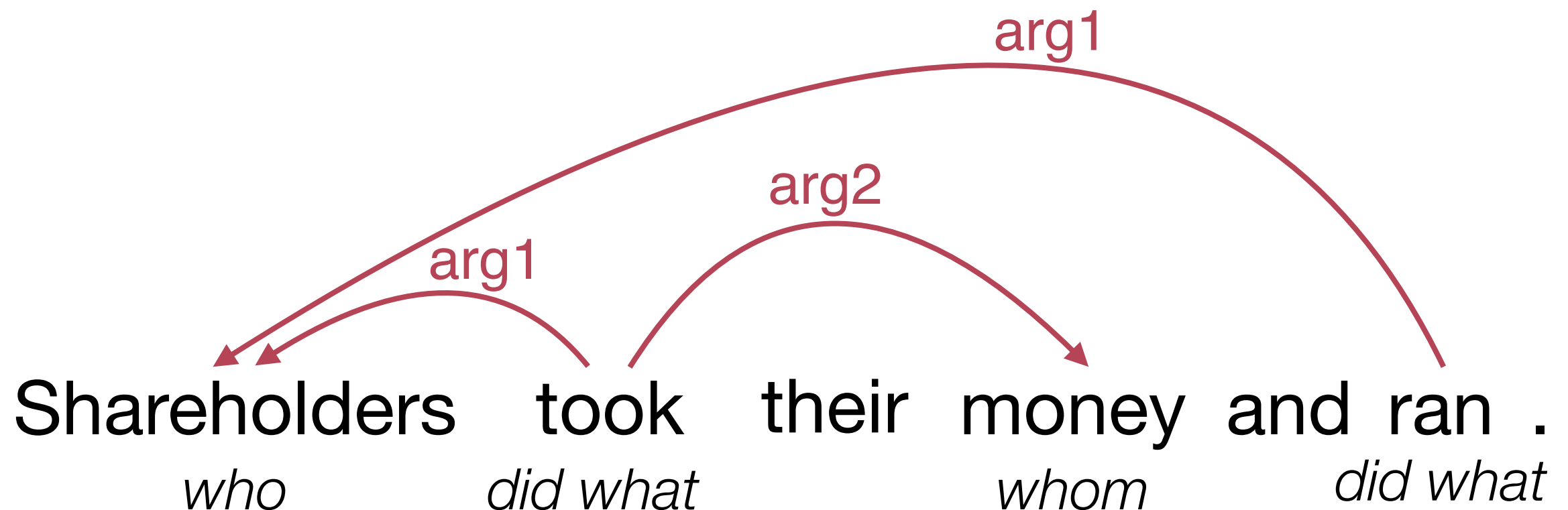
who did what to whom,
where and why...

role label

who
what
when
where
why
...

predicate

argument



Semantic Parsing

Motivation

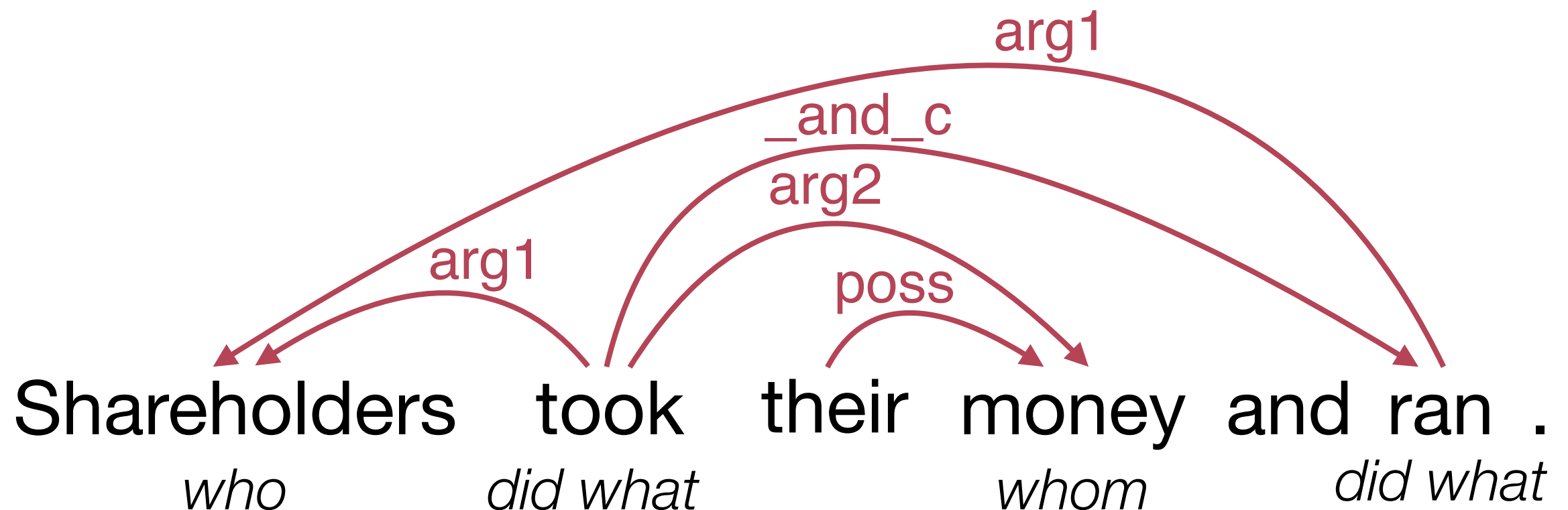
who did what to whom,
where and why...

role label

who
what
when
where
why
...

predicate

argument



DM (DELPH-IN MRS)



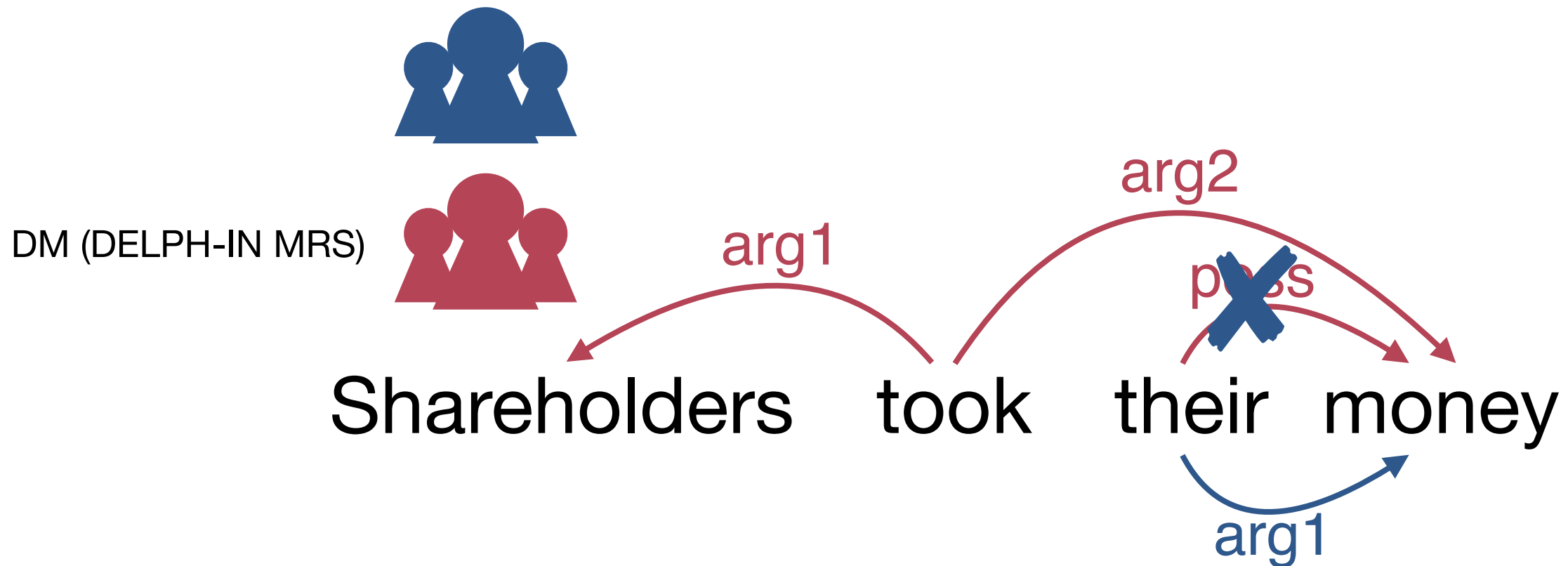
Shareholders took their money

arg1

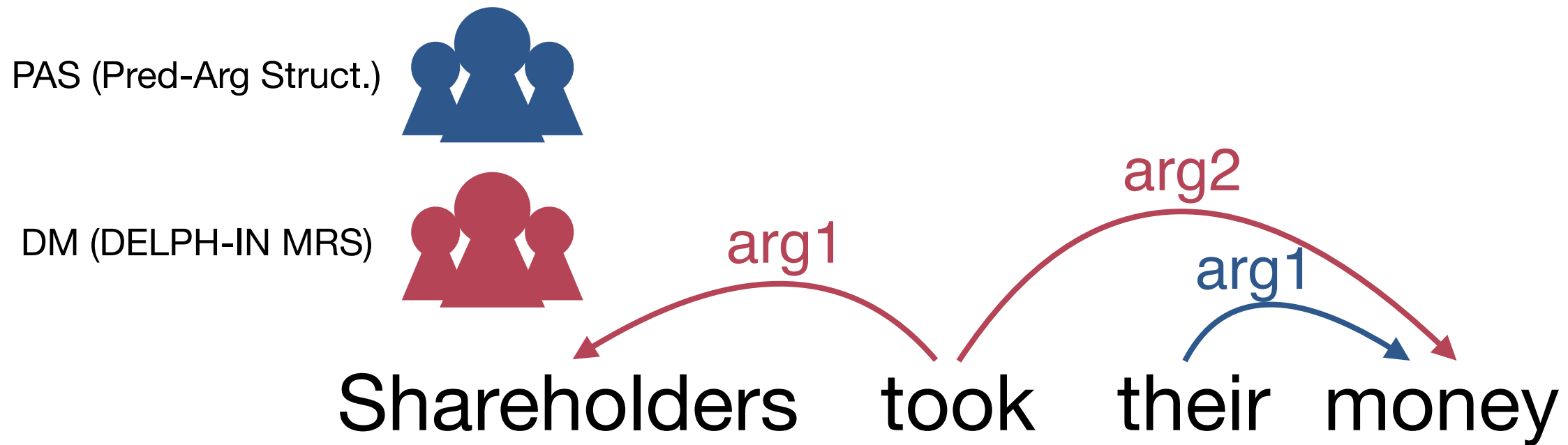
arg2

poss

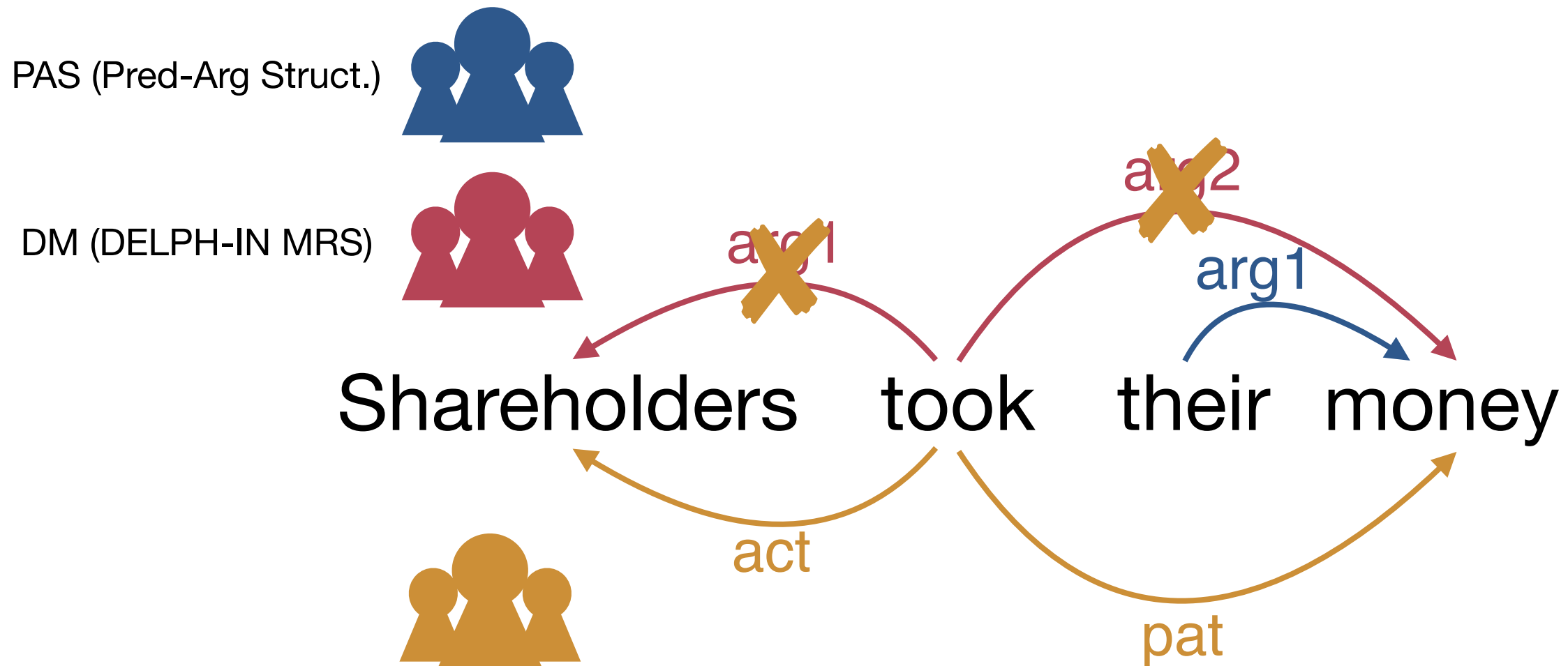
Fragmented Annotation Efforts



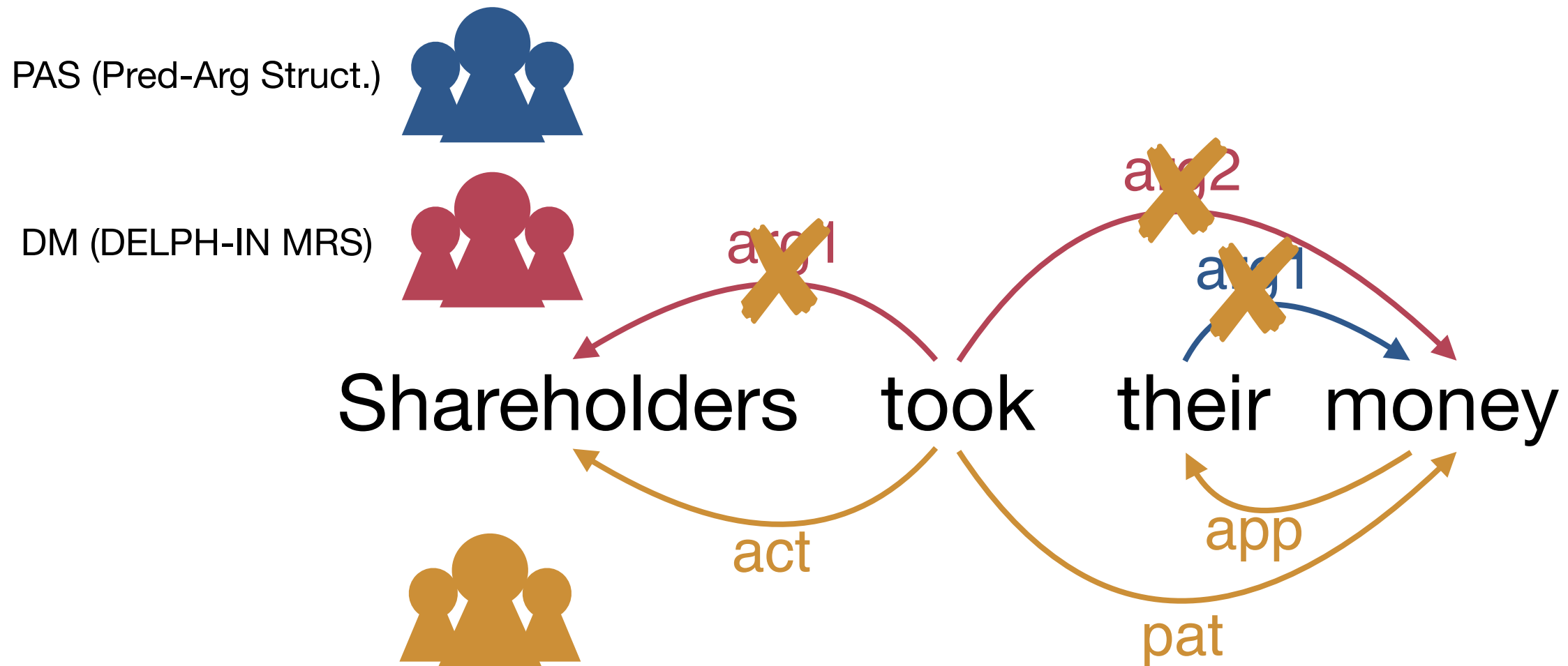
Fragmented Annotation Efforts



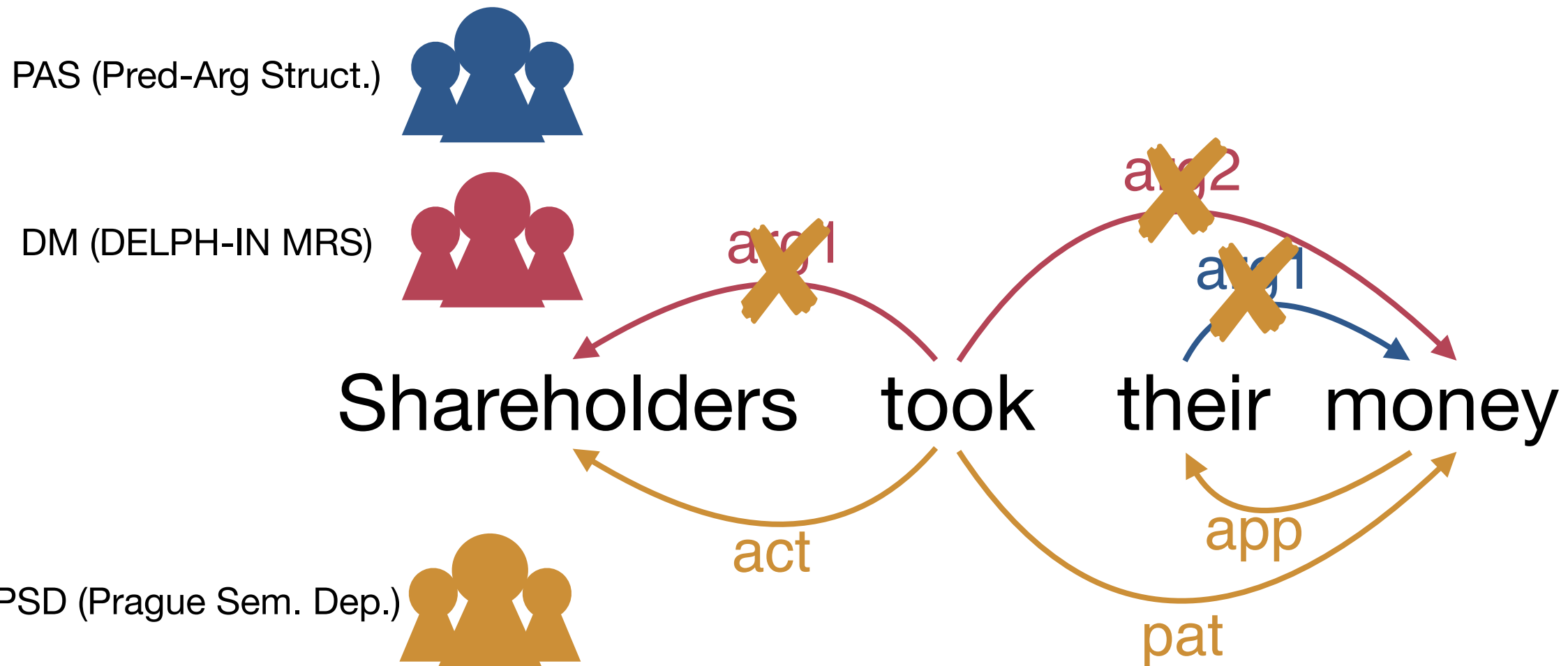
Fragmented Annotation Efforts



Fragmented Annotation Efforts



Fragmented Annotation Efforts



Hypothesis

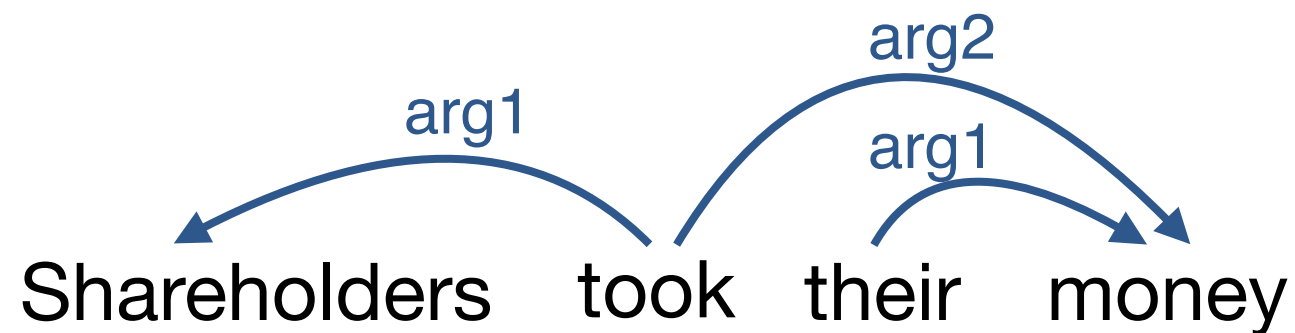
- Overlap among different semantic theories and their divergent implementations (annotations).
- Exploit with multitask learning.

Overview

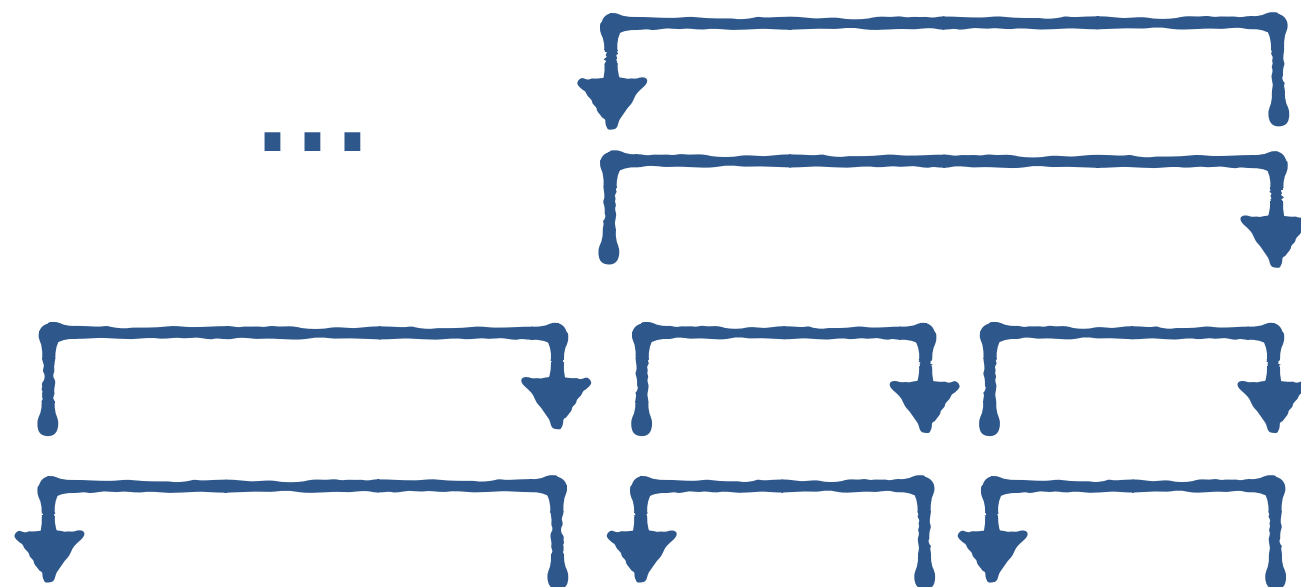
- Three parallel annotated formalisms.
- State-of-the art basic model.
- Two orthogonal joint learning approaches.

Parsing Semantic Graphs

objective: structured
hinge loss



inference: labeled
directed graphs



labeled bilexical
dependencies

Shareholders took their money

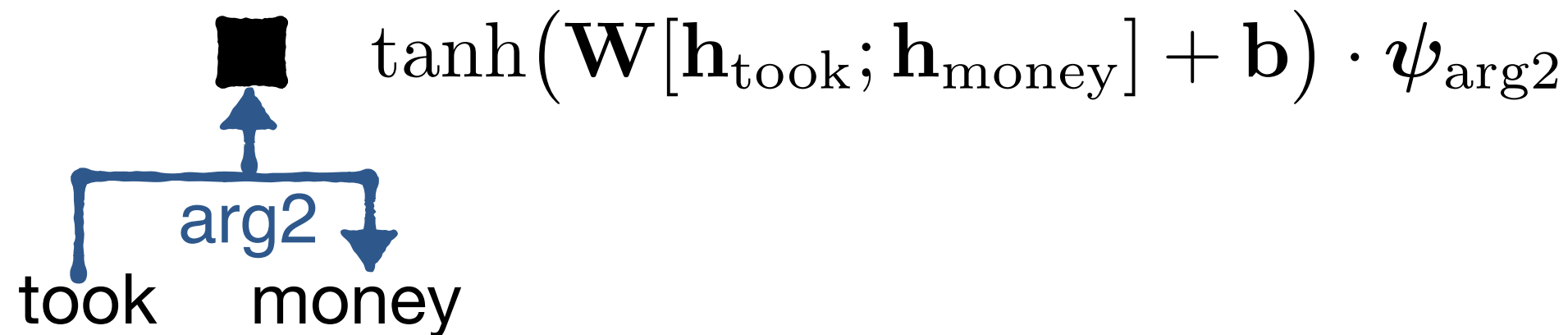
tokens

BiLSTMs



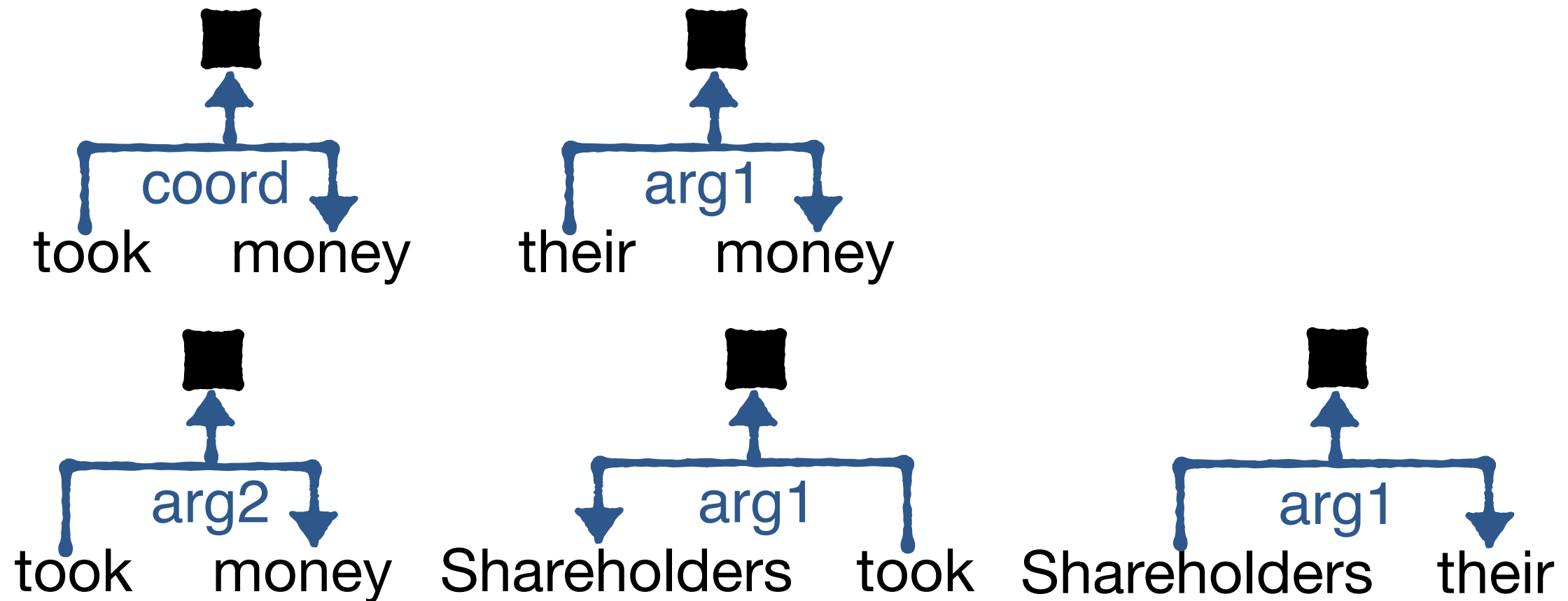
Shareholders took their money

Scoring with MLPs



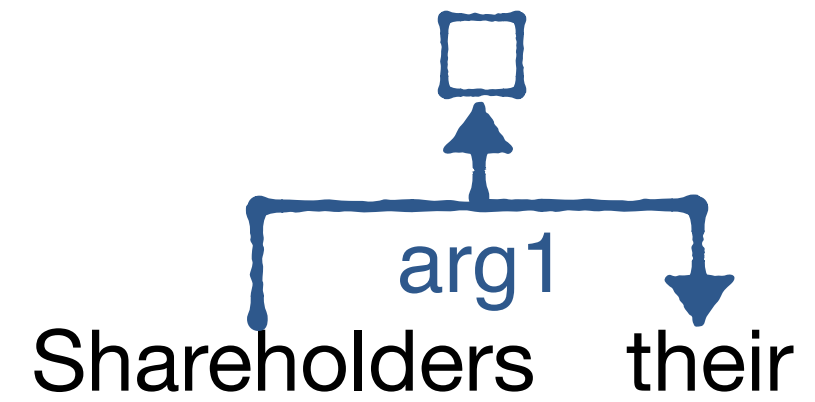
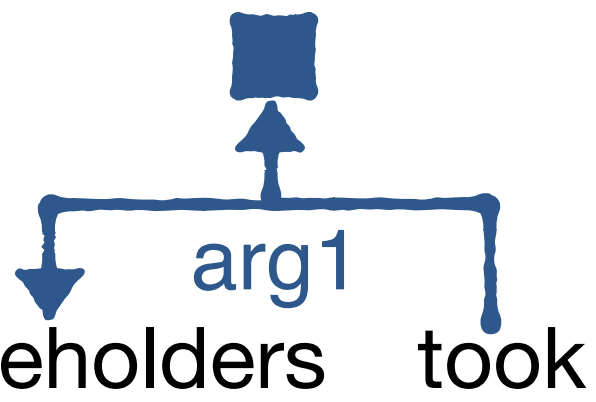
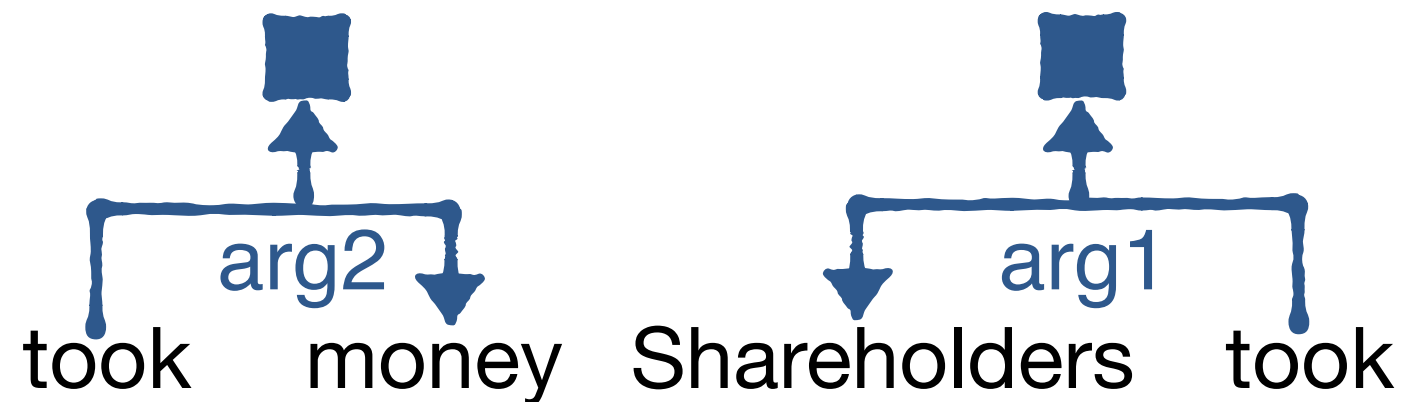
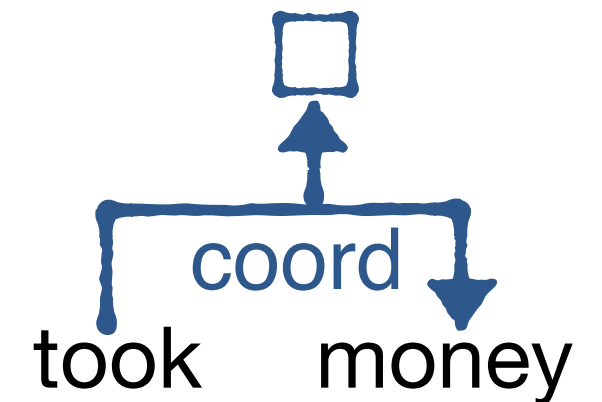
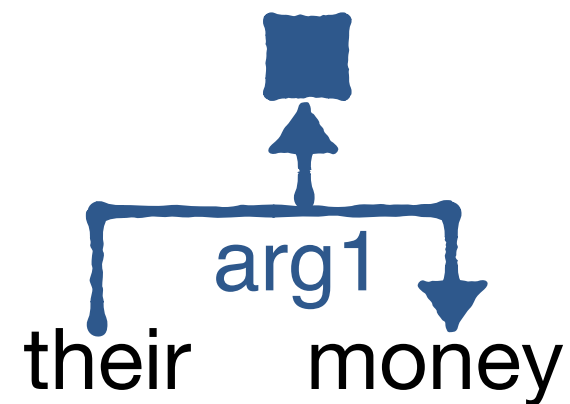
Shareholders took their money

Scoring with MLPs



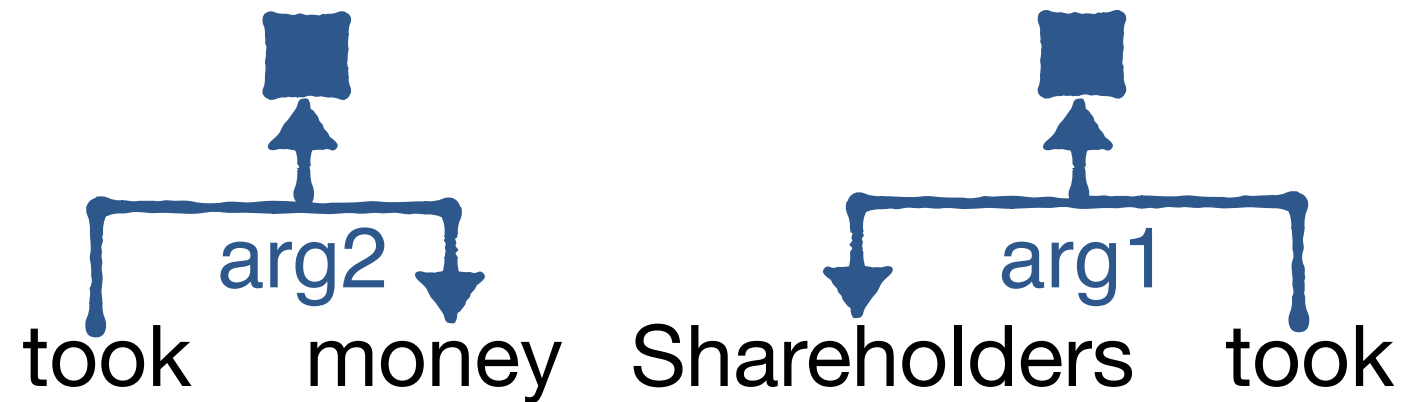
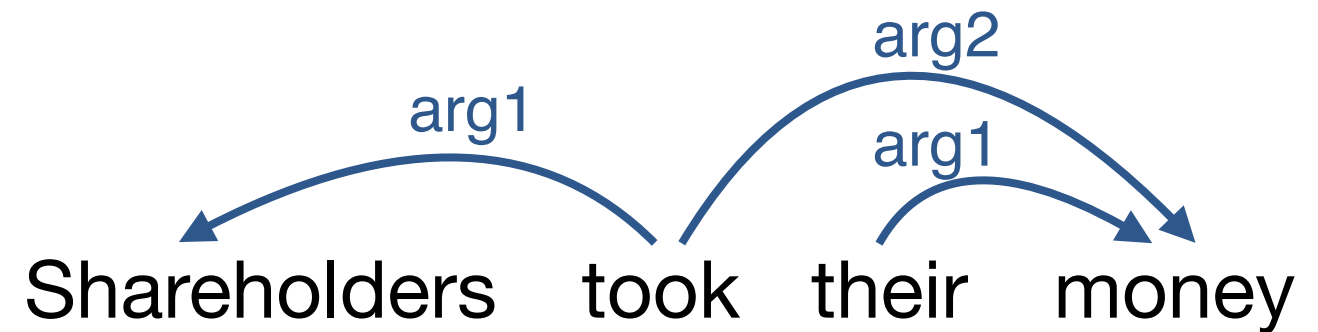
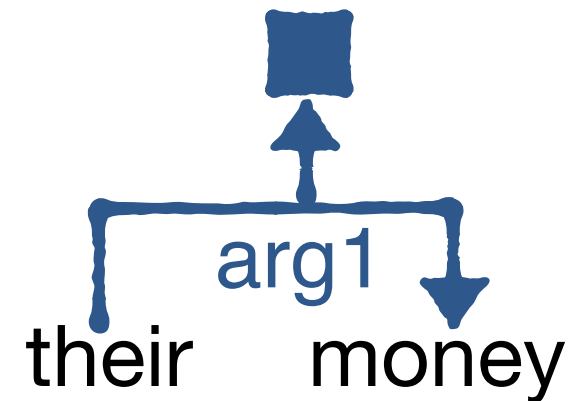
Shareholders took their money

Inference (Martins et al., 2014)



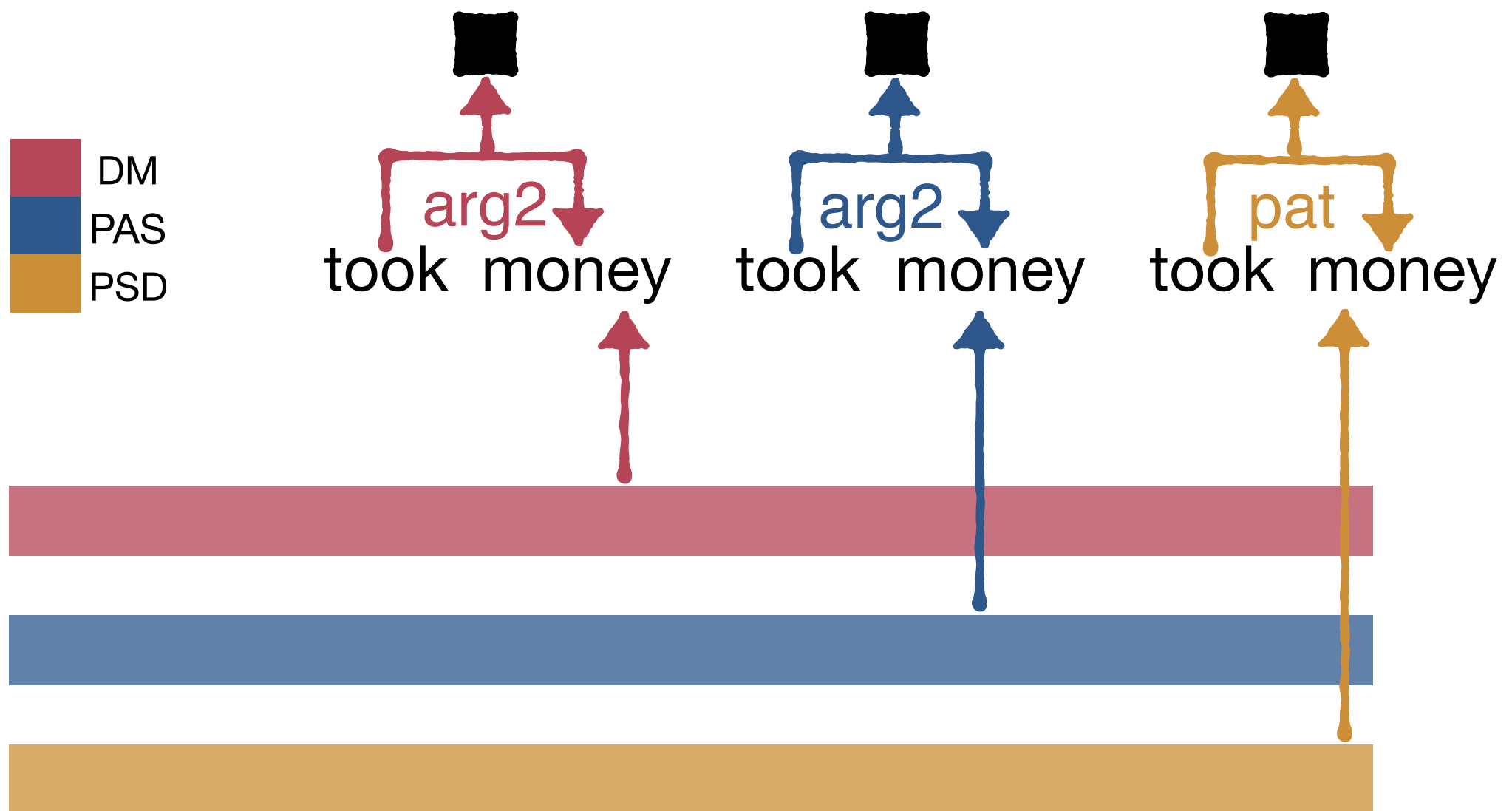
Shareholders took their money

Inference (Martins et al., 2014)



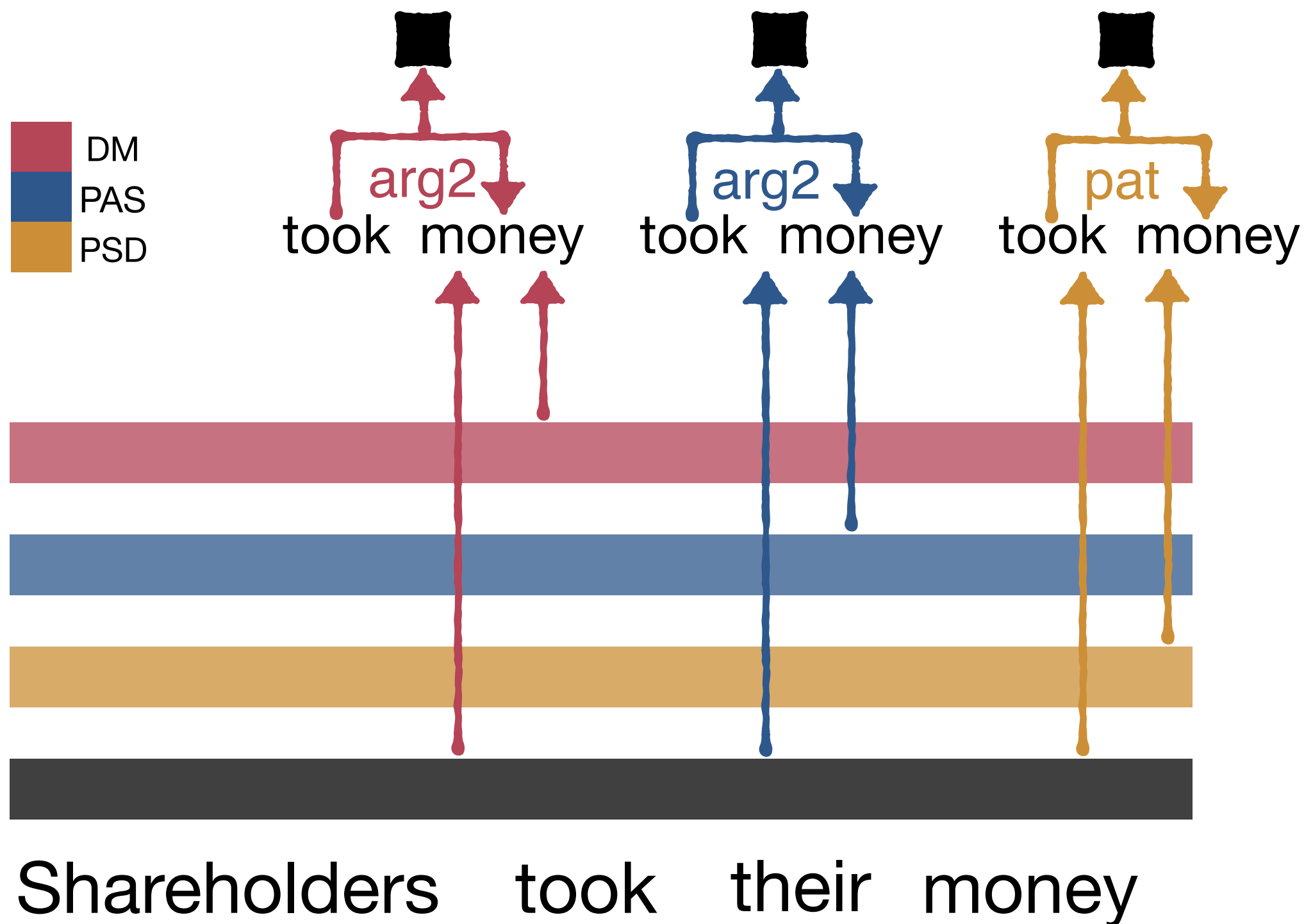
Shareholders took their money

Separate Parsers

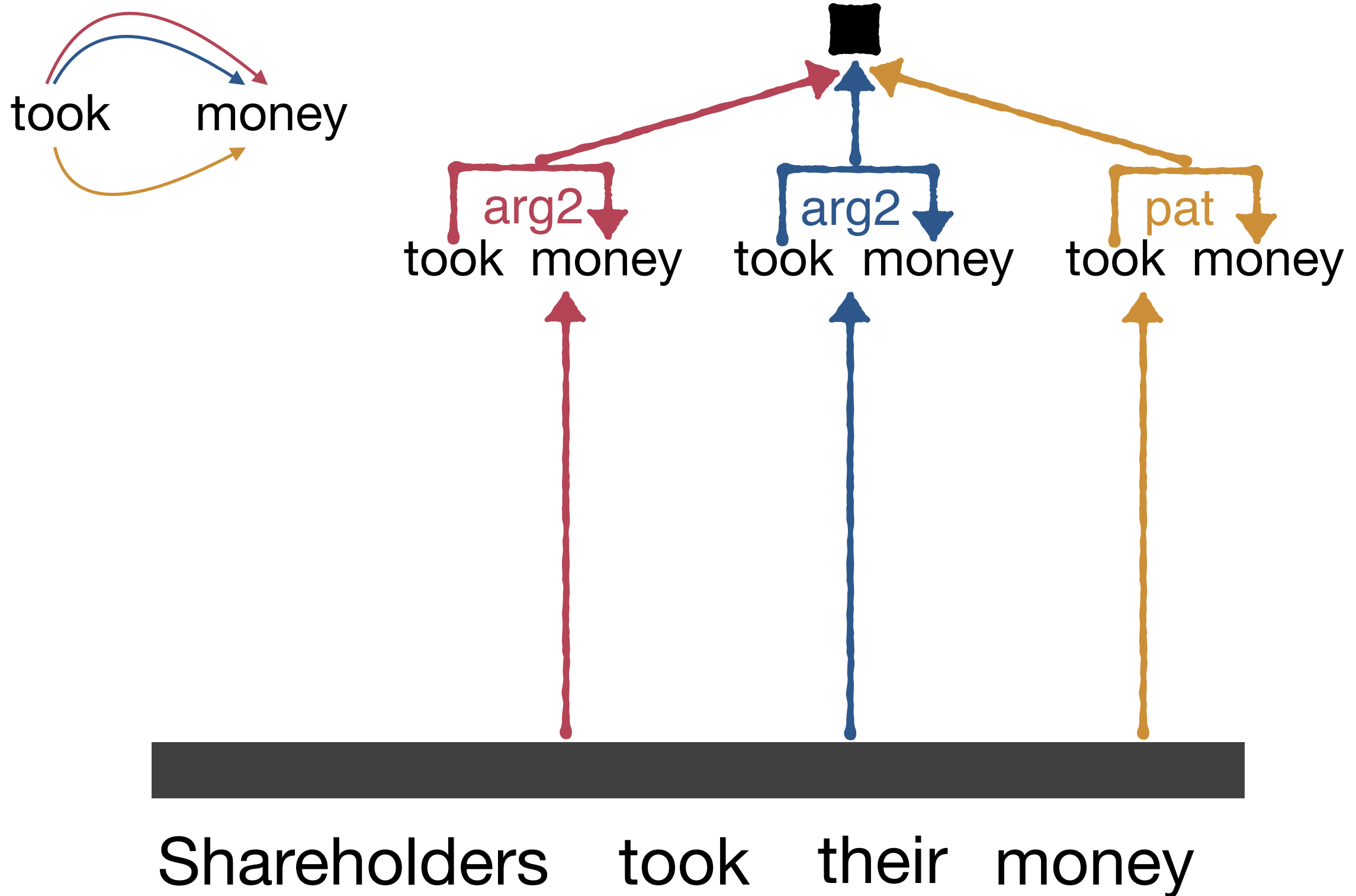


Shared Input Representation

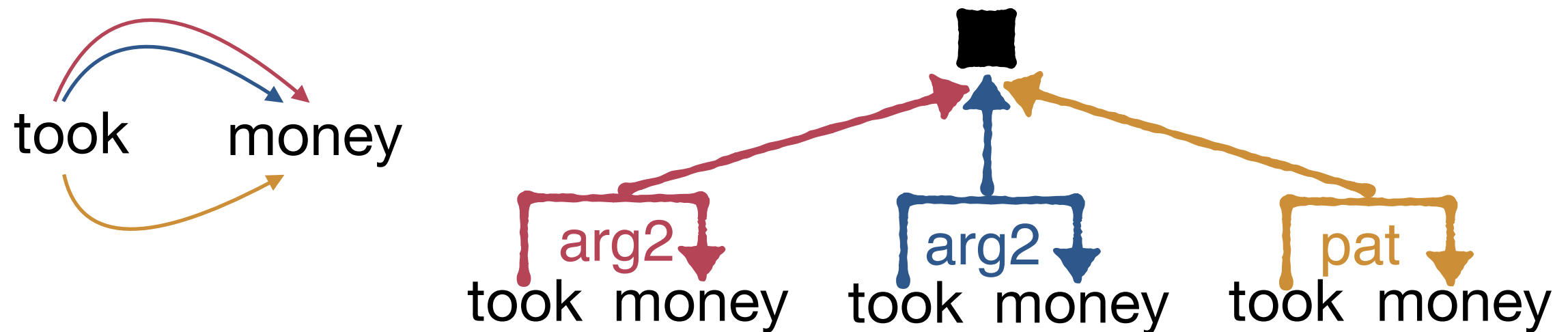
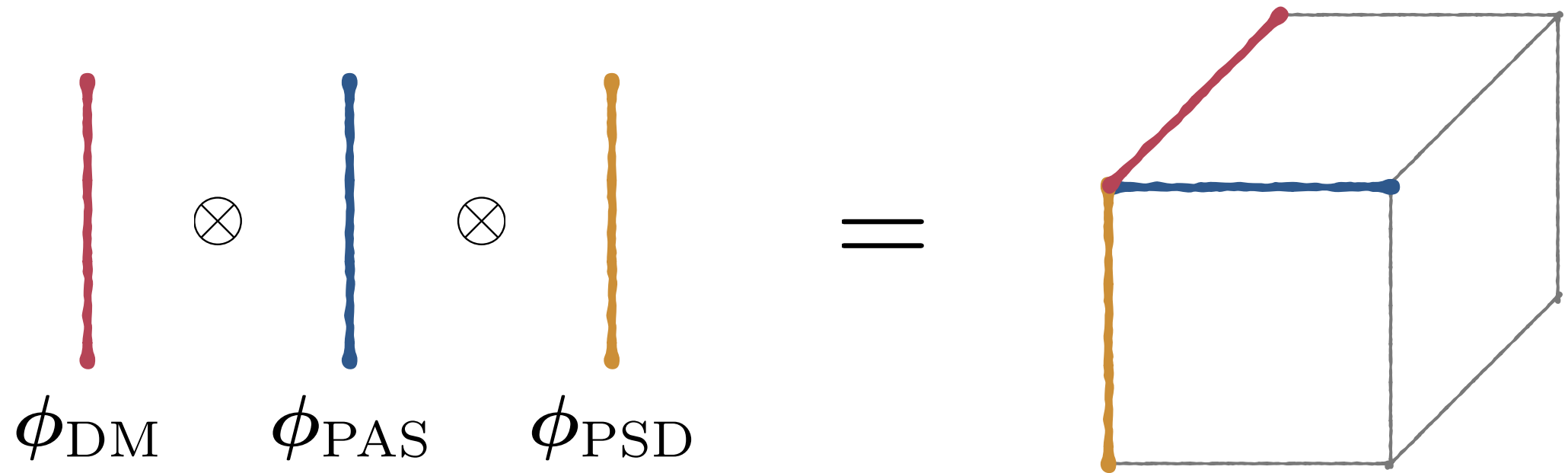
(Daume, 2007)



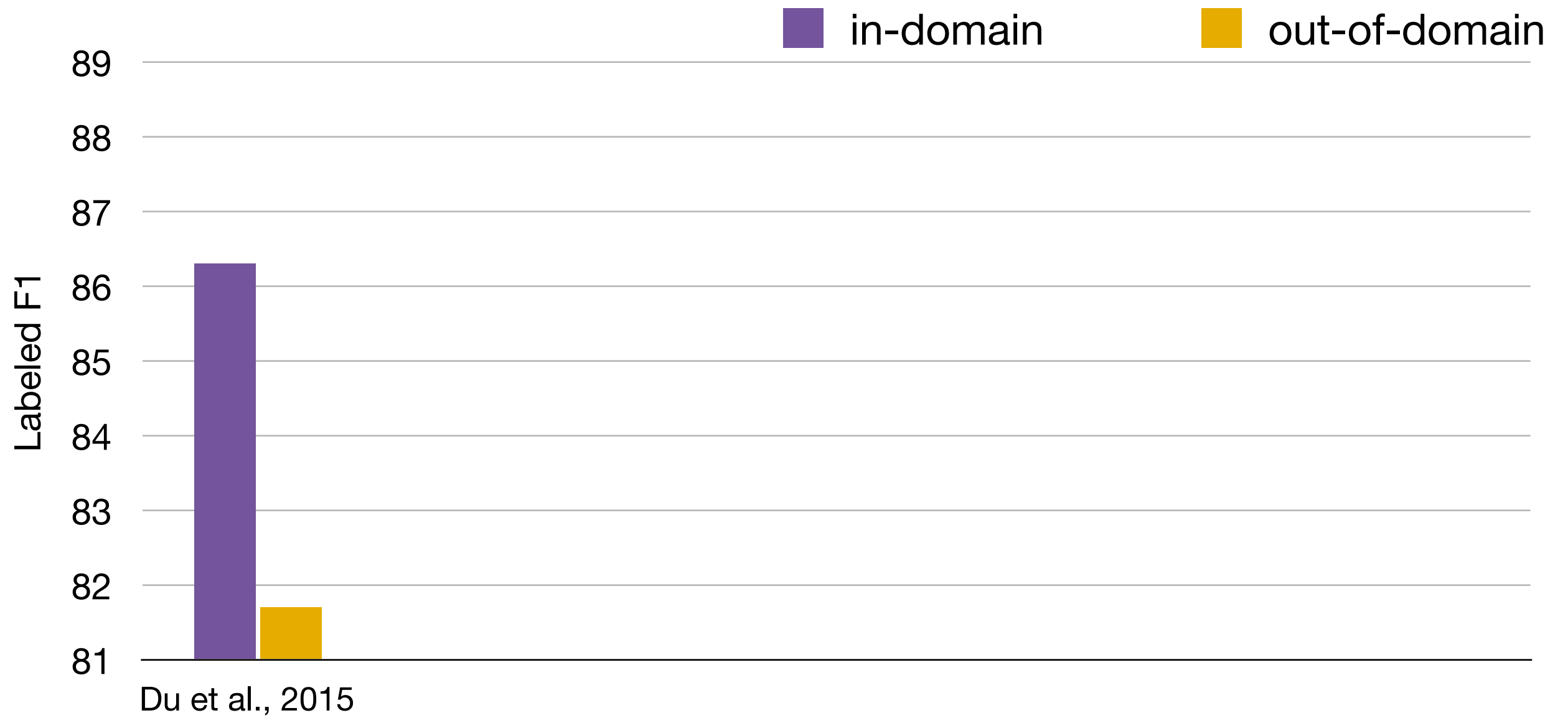
Jointly Predict Three Graphs



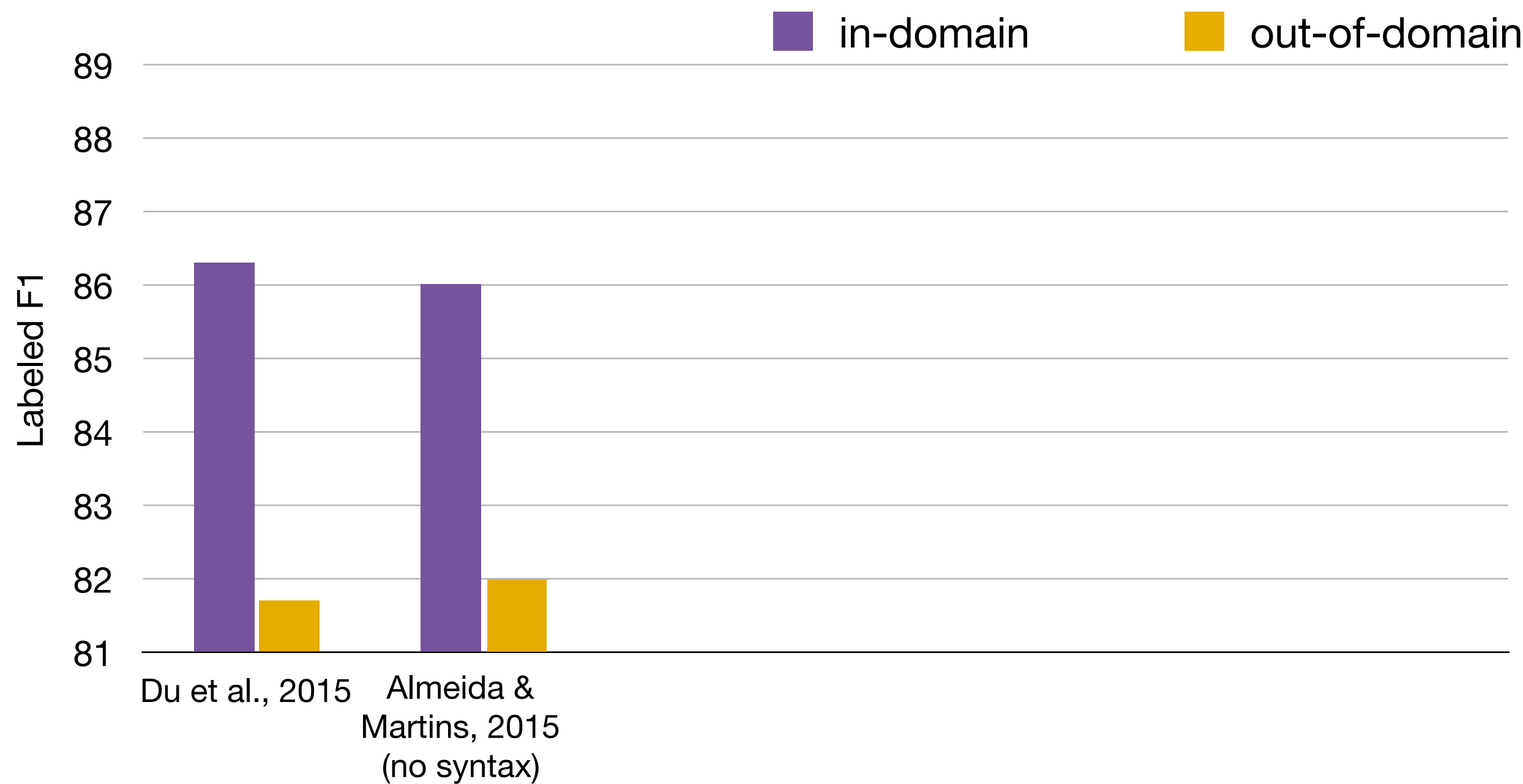
Jointly Predict Three Graphs



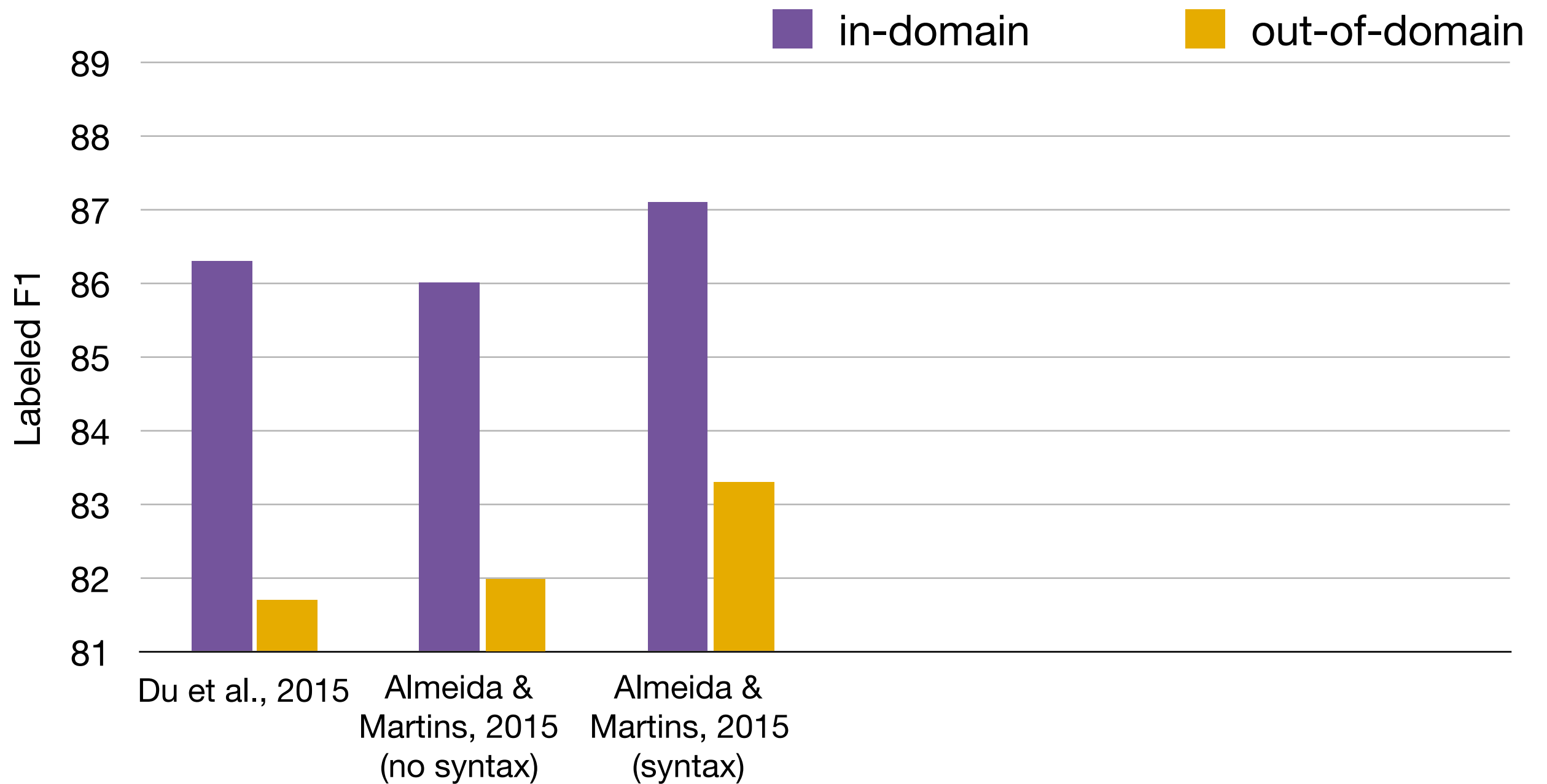
SemEval '15. Labeled F_1 micro-averaged over three formalisms.



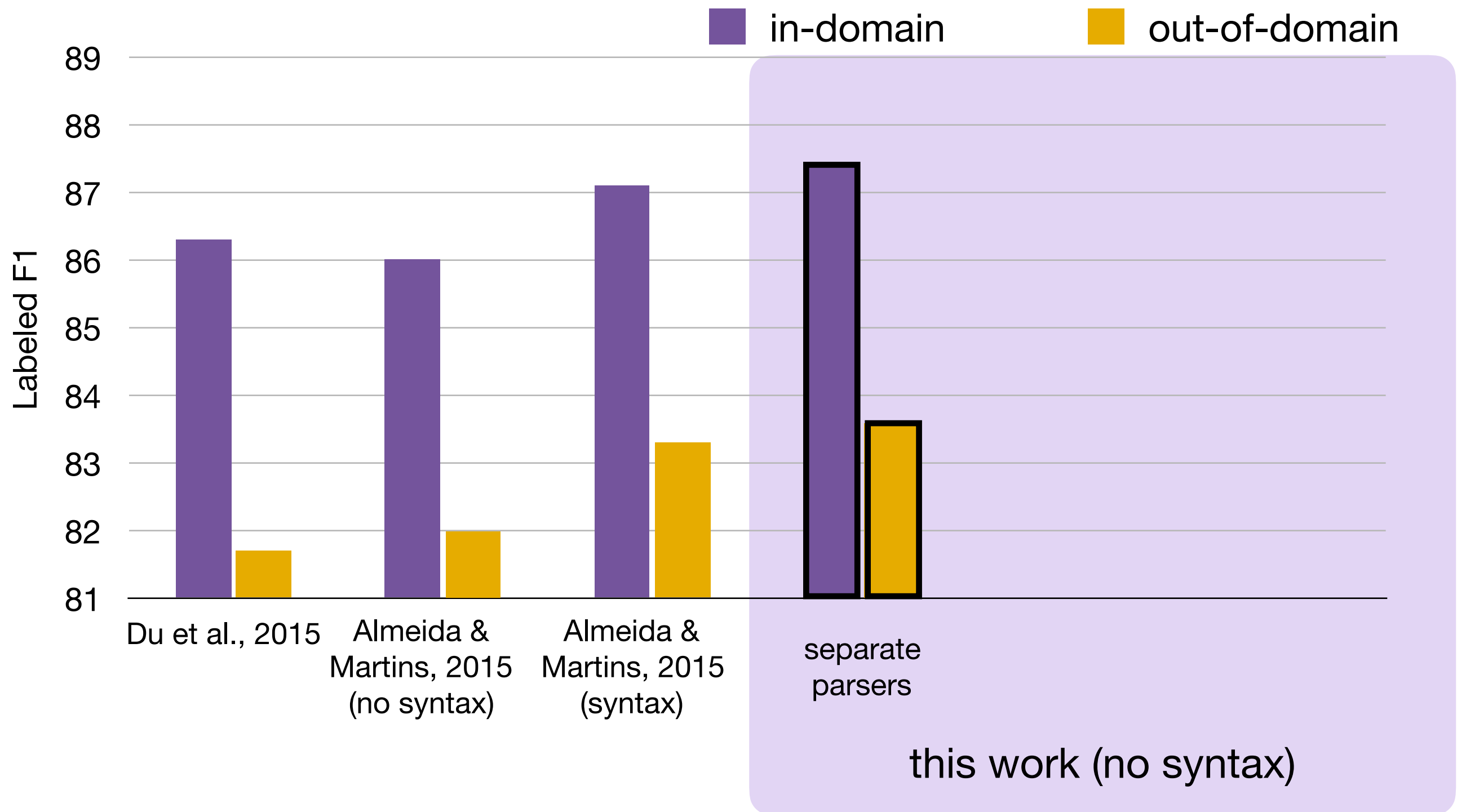
SemEval '15. Labeled F_1 micro-averaged over three formalisms.



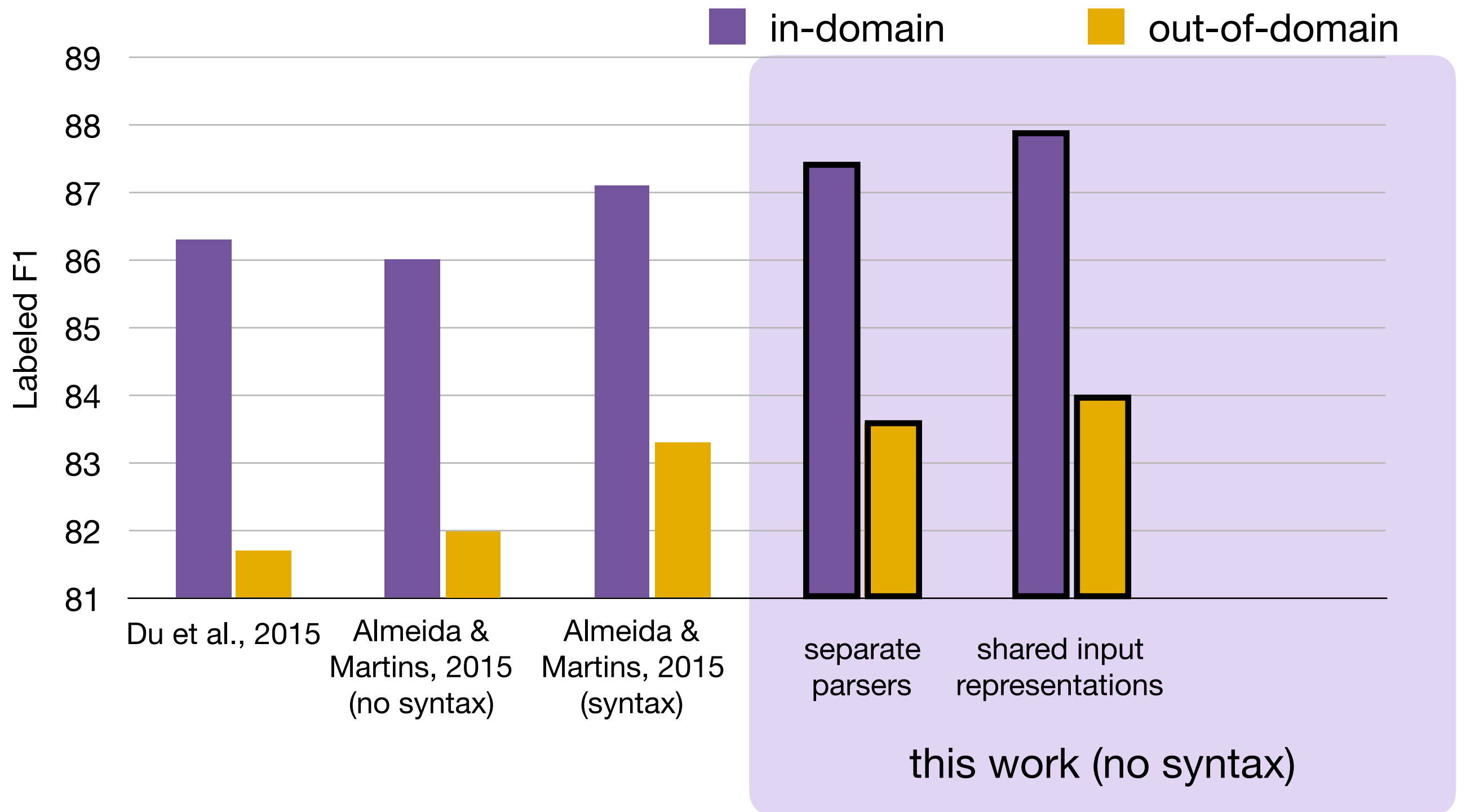
SemEval '15. Labeled F_1 micro-averaged over three formalisms.



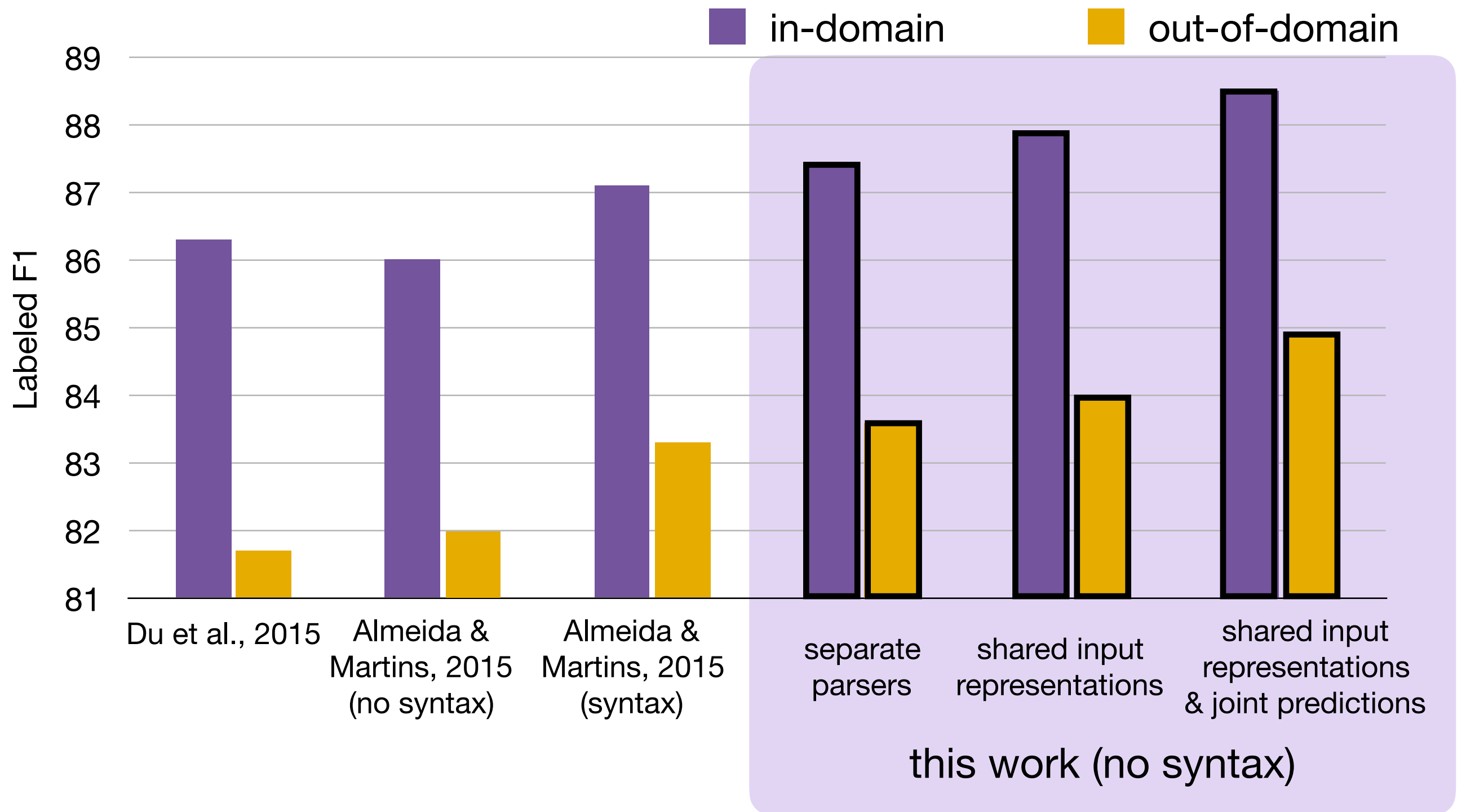
SemEval '15. Labeled F_1 micro-averaged over three formalisms.



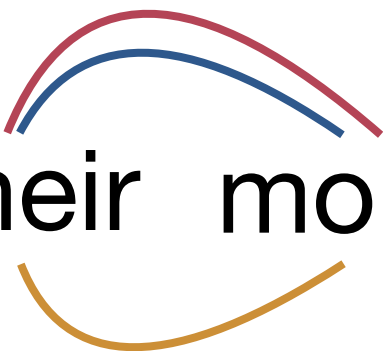
SemEval '15. Labeled F_1 micro-averaged over three formalisms.



SemEval '15. Labeled F_1 micro-averaged over three formalisms.



Structural Overlap



their money

undirected

	DM	PAS	PSD
DM	-	67.2	
PAS	70.0	-	
PSD			-

Structural Overlap


their money

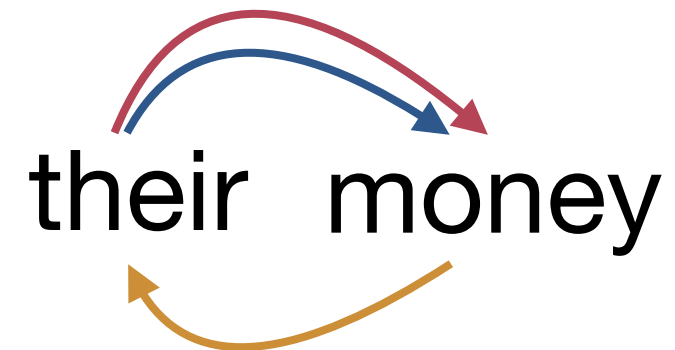
undirected

	DM	PAS	PSD
DM	-	67.2	56.8
PAS	70.0	-	54.9
PSD	57.4	56.3	-

Structural Overlap



undirected



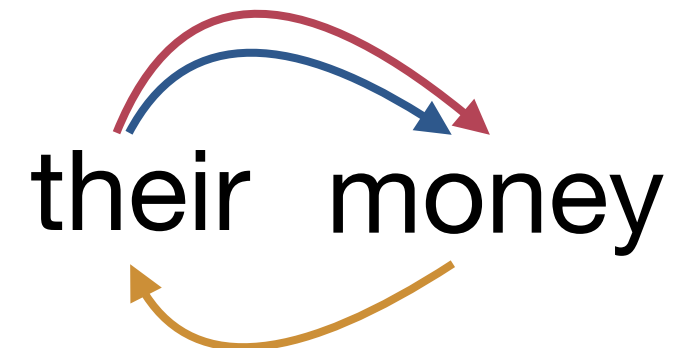
directed

	DM	PAS	PSD	DM	PAS	PSD
DM	-	67.2	56.8	-	64.2	
PAS	70.0	-	54.9	66.9	-	
PSD	57.4	56.3	-			-

Structural Overlap



undirected



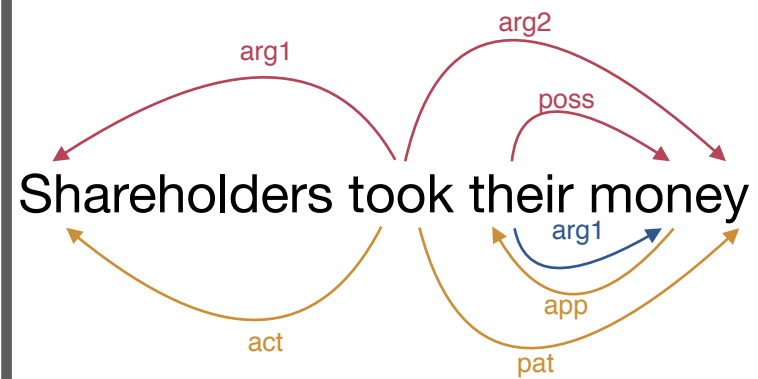
directed

	DM	PAS	PSD	DM	PAS	PSD
DM	-	67.2	56.8	-	64.2	26.1
PAS	70.0	-	54.9	66.9	-	26.1
PSD	57.4	56.3	-	26.4	29.6	-

PSD reverses the direction half of the time.

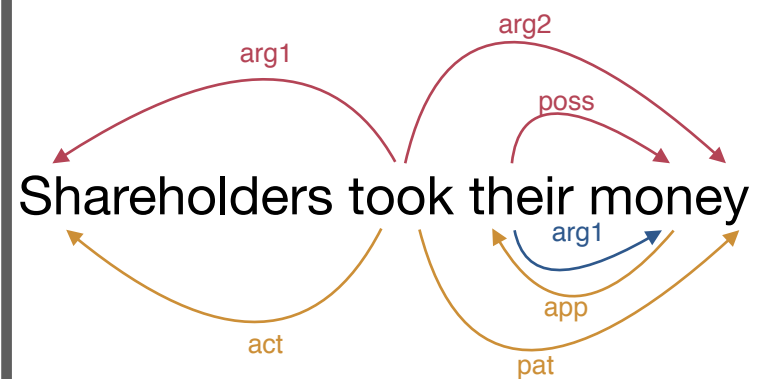
Conclusion

Problem

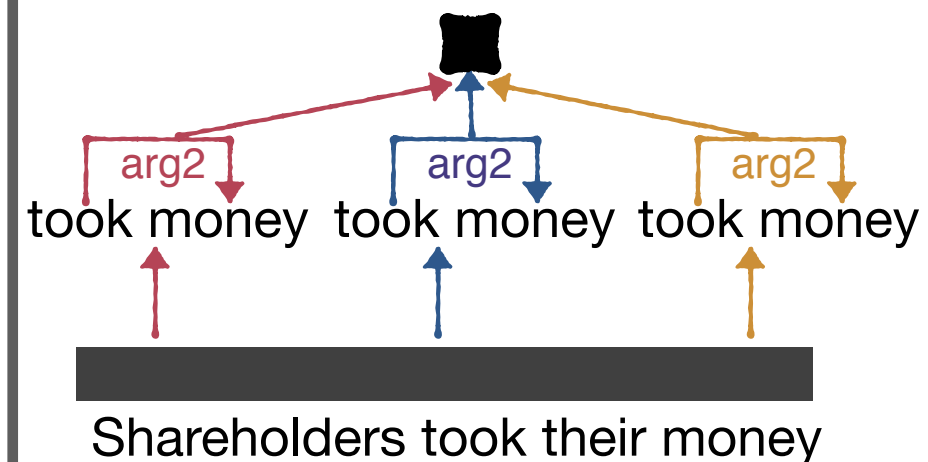
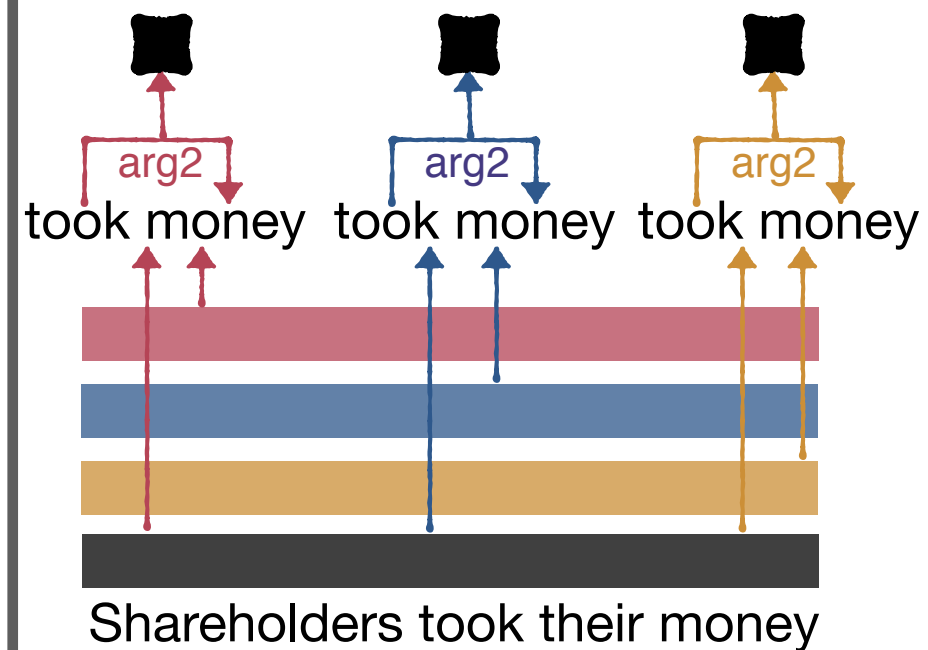


Conclusion

Problem

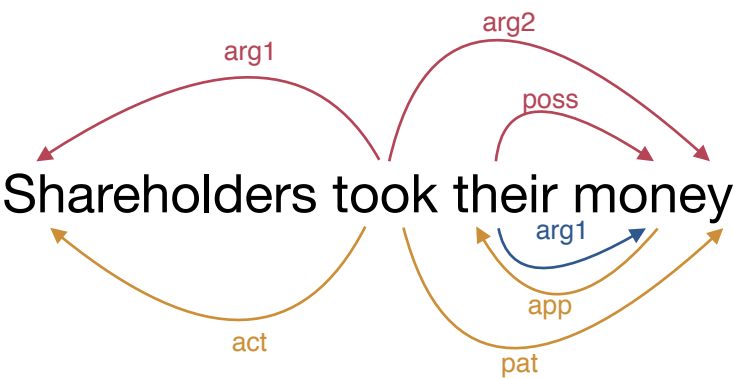


Method

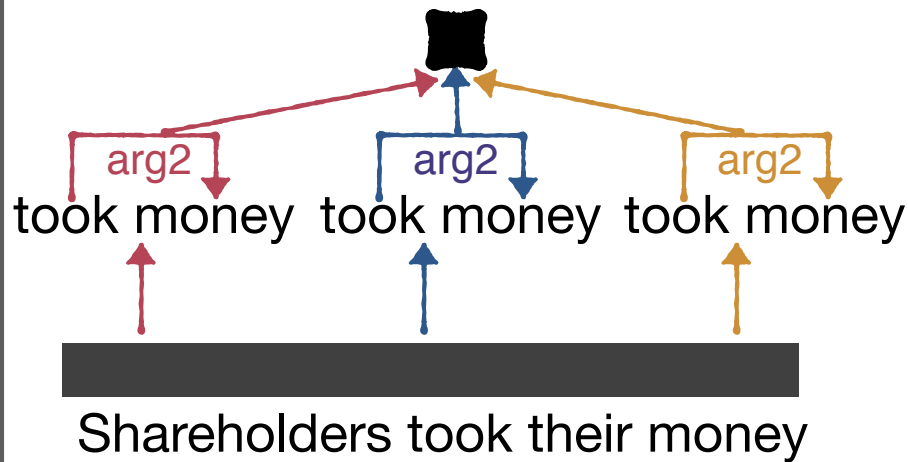
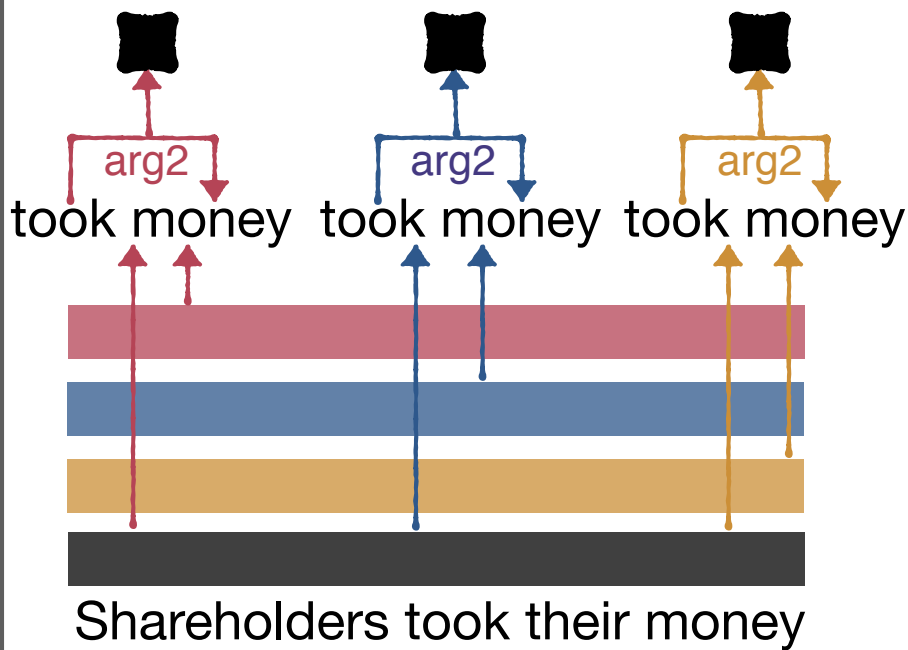


Conclusion

Problem



Method



Analysis



	DM	PAS	PSD
DM	-	67.2	56.8
PAS	70.0	-	54.9
PSD	57.4	56.3	-



	DM	PAS	PSD
DM	-	64.2	26.1
PAS	66.9	-	26.1
PSD	26.4	29.6	-

