Broad-Coverage Semantic Parsing

Sheng Zhang

Joint work w/ Xutai Ma, Kevin Duh, and Benjamin Van Durme.

Center for Language and Speech Processing (CLSP)

JOHNS HOPKINS
UNIVERSITY

Semantic Parsing

Natural language text ⇒ Meaning representation (**MR**)

E.g.

Caesar was stabbed by Brutus.

↓

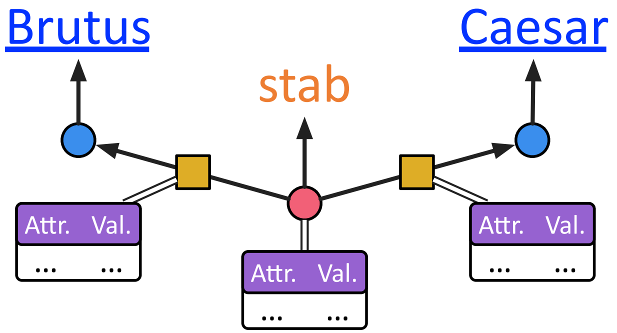
stab(Brutus, Caesar)

JOHNS HOPKINS UNIVERSITY

Shallow-to-Deep Semantic Processing

Shallow

Deep

	Task	**MR**
	Open Information Extraction	(Caesar, **was stabbed**, by Brutus)
	Semantic Role labeling	**stab(Brutus, Caesar)**
	Semantic Parsing	

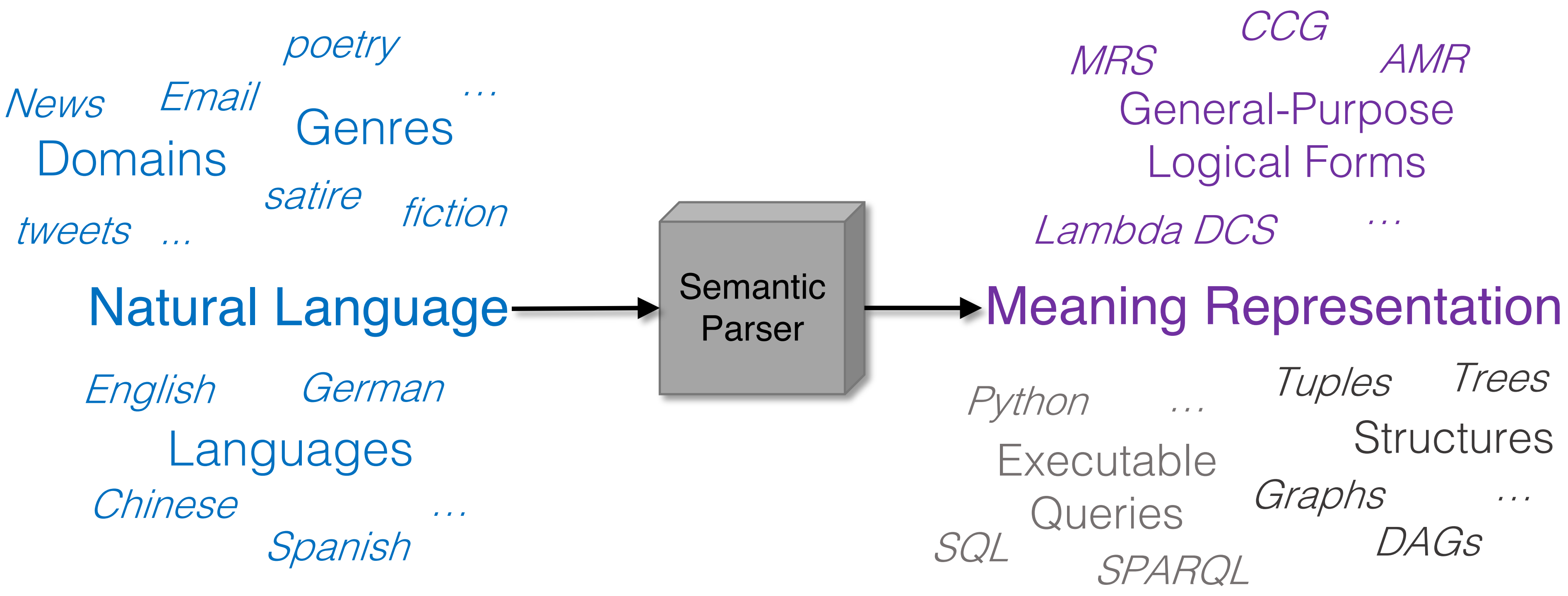
JOHNS HOPKINS UNIVERSITY

Semantic Parsing

Natural language text ⇒ Meaning representation (**MR**)

Broad-Coverage Semantic Parsing

Natural language text ⇒ Meaning representation (**MR**)



Broad-Coverage Semantic Parsing

(A long-standing topic of interest in CL)

Discourse Representation Theory
(**DRT**; Kamp, 1984, 1993; Bos, 2008)

Head-driven Phrase Structure Grammar
(**HPSG**; Pollard and Sag, 1994)

Minimal Recursion Semantics
(**MRS**; Copestake et al., 2005)

Combinatory Categorical Grammar
(**CCG**; Steedman 1996)

Abstract Meaning Representation
(**AMR**; Banarescu et al., 2013)

Universal Conceptual Cognitive Annotation
(**UCCA**; Abend and Rappoport, 2013)

Universal Decompositional Semantics
(**UDS**; Reisinger et al., 2015; White et al., 2016)

(Not a full enumeration.)

Lenhart Schubert. 2015. Semantic Representation. *AAAI*
Omri Abend and Ari Rappoport. 2017. The State of the Art in Semantic Representation. *ACL*

AMR Bibliography

https://nert-nlp.github.io/AMR-Bibliography/

Explore the research on AMR

The table below is sortable by column.
You can highlight rows by topic (click on a topic TAG). The main topics are:

Annotation: Research on and methods for AMR annotators
Parsing: Produce an AMR from natural language text
(See state of the art on NLP-progress)
Generation: Produce natural language text from an AMR
Applications: Summarization, Information Extraction, Biomedical, etc.
Alignment: Find the AMR subgraph corresponding to a word or phrase
AMR Extensions: Research which adds features to the AMR annotation scheme
Multilingual: Extensions of AMR from its original language (English) to more languages

```
(j / join-01
      :ARG0 (p / person :wiki -
            :name (p2 / name :op1 "Pierre" :op2 "Vinken")
            :age (t / temporal-quantity :quant 61
                  :unit (y / year)))
      :ARG1 (b / board
            :ARG1-of (h / have-org-role-91
                  :ARG0 p
                  :ARG2 (d2 / director
                        :mod (e / executive :polarity -))))
      :time (d / date-entity :month 11 :day 29))
```

The AMR for the sentence *Pierre Vinken, 61 years old, will join the board as a nonexecutive director Nov. 29.*

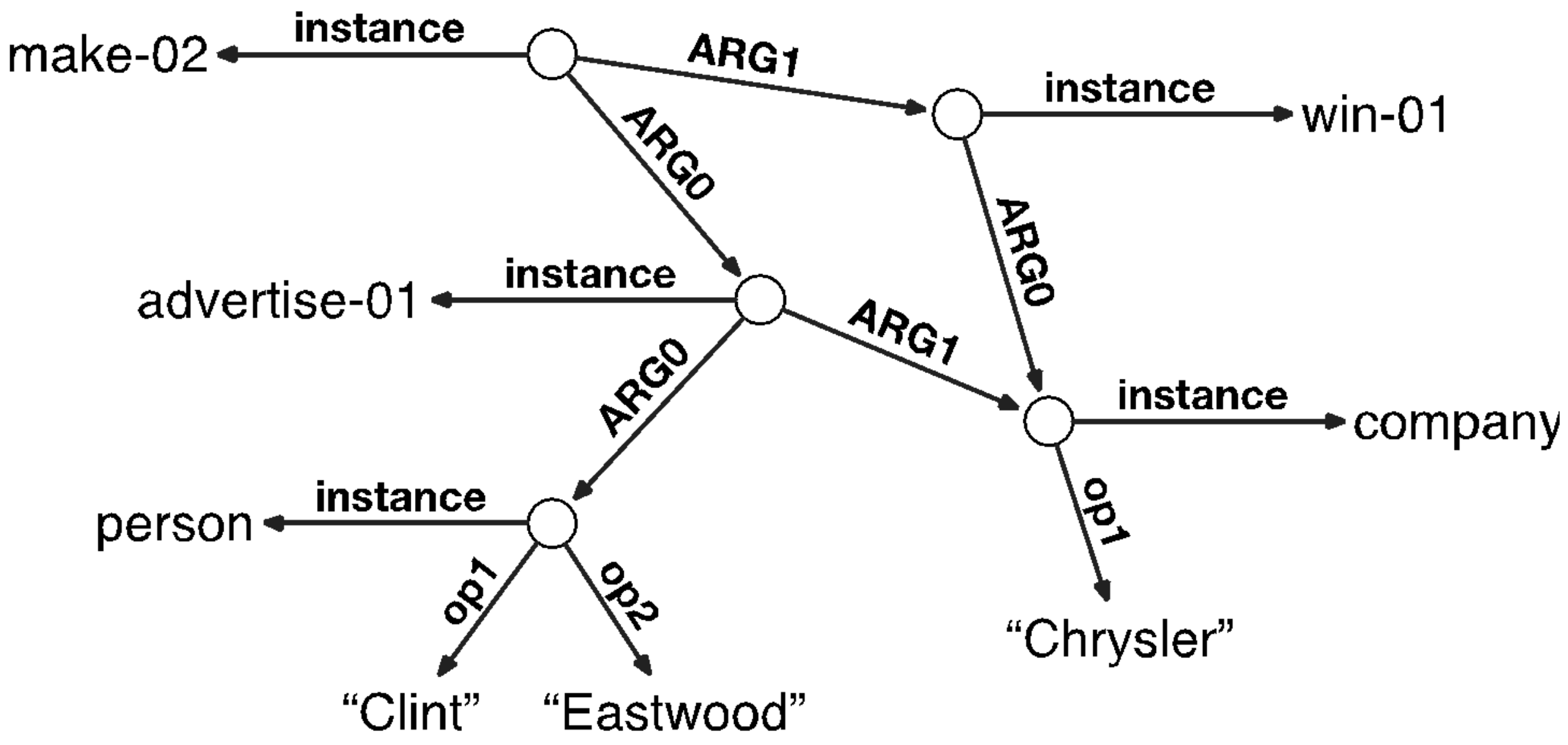
Annotation, Parsing, Generation, Applications, Alignment, AMR Extensions, Multilingual, etc.

Title	Authors	Venue	Year	Link(s)	Arxiv	Tags
Augmenting Abstract Meaning Representation for Human-Robot Dialogue	Claire Bonial, Lucia Donatelli, Stephanie M. Lukin, Stephen Tratz, Ron Artstein, David Traum, Clare R. Voss	Designing Meaning Representations Workshop	2019	pdf		AMR Extensions, Applications
Separating Argument Structure from Logical Structure in AMR	Johan Bos	preprint	2019		arXiv	AMR Extensions
Towards a General Abstract Meaning Representation Corpus for Brazilian Portuguese	Marco Antonio Sobrevilla Cabezudo, Thiago Alexandre Salgueiro Pardo	Linguistic Annotation Workshop	2019	pdf		Annotation, Multilingual
Core Semantic First: A Top-down Approach for AMR Parsing	Deng Cai, Wai Lam	EMNLP	2019		arXiv	Parsing
Factorising AMR generation through syntax	Kris Cao, Stephen Clark	NAACL	2019	pdf	arXiv	Generation

JOHNS HOP
UNIVERSI

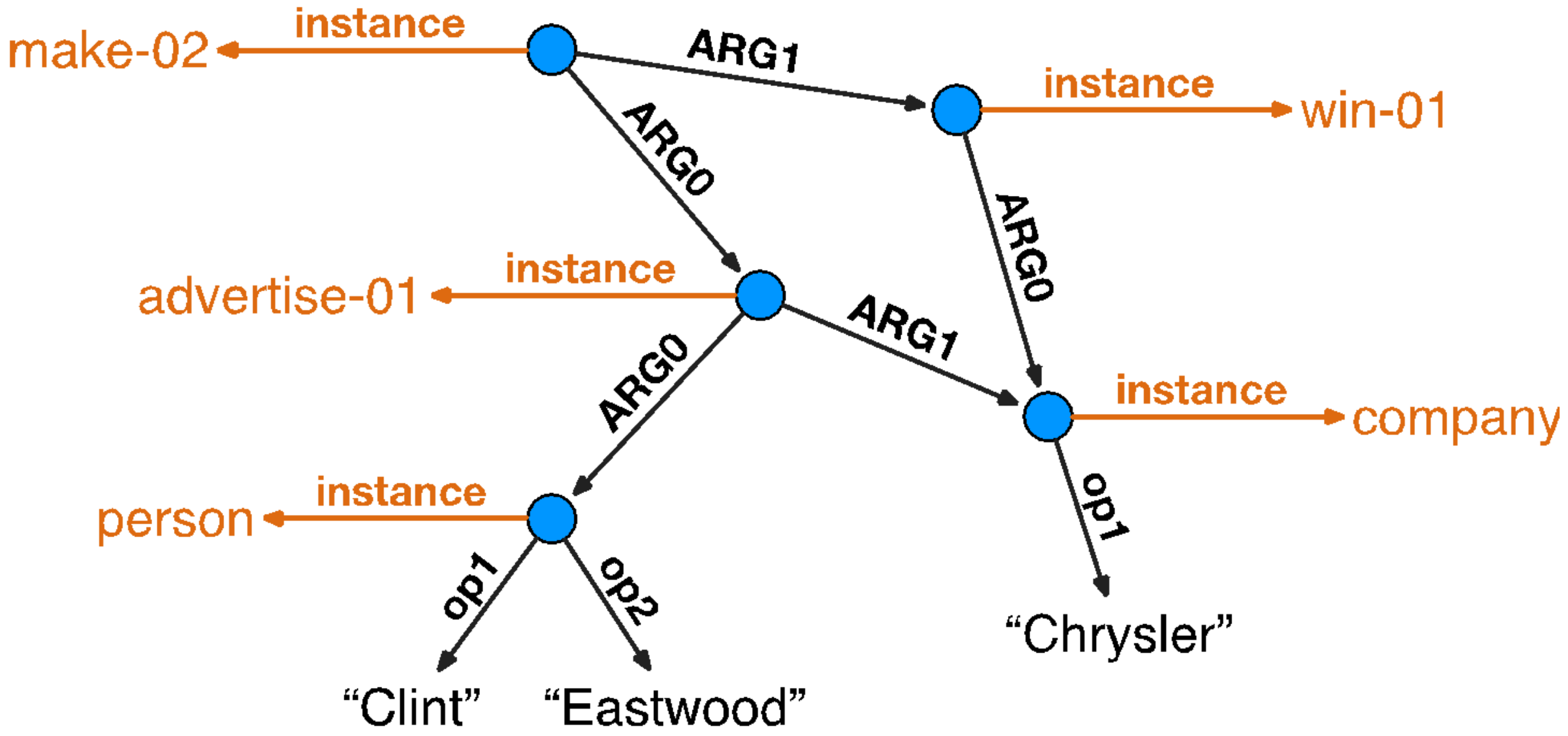
Abstract Meaning Representation (AMR)

Clint Eastwood's ad for Chrysler makes them the winner.



AMR: Instances

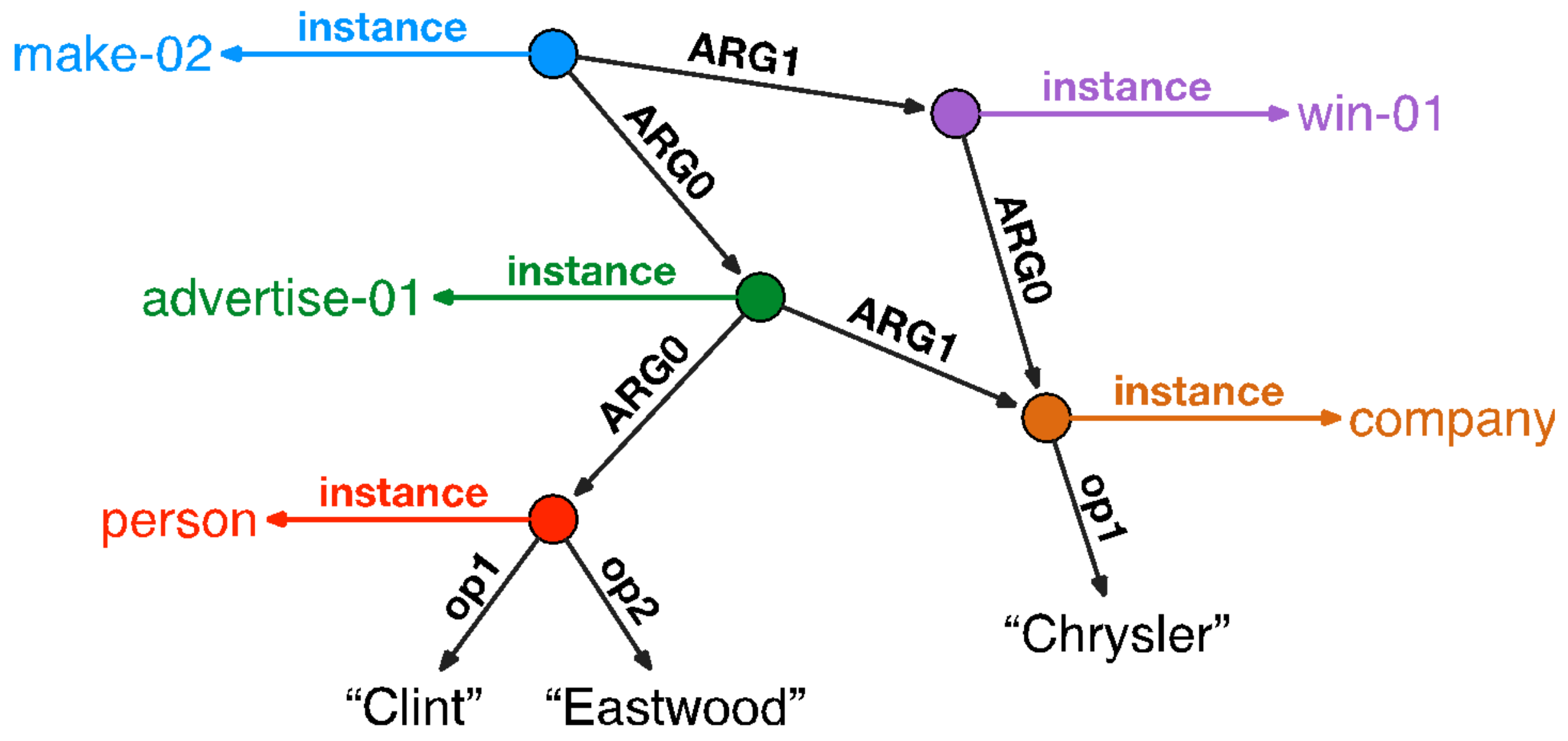
Clint Eastwood's ad for Chrysler makes them the winner.



JOHNS HOPKINS UNIVERSITY

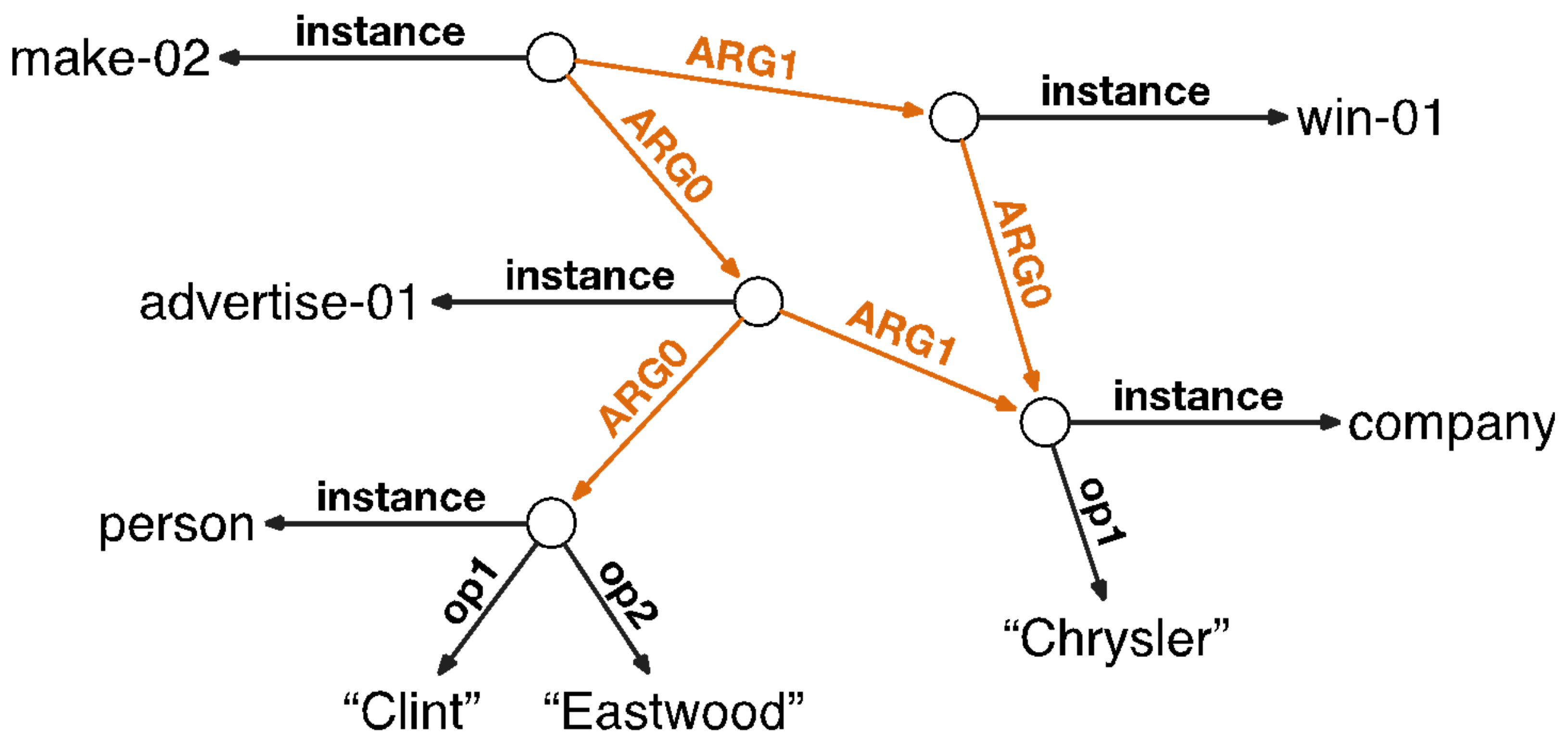
AMR: Instances

Clint Eastwood's ad for Chrysler makes them the winner.



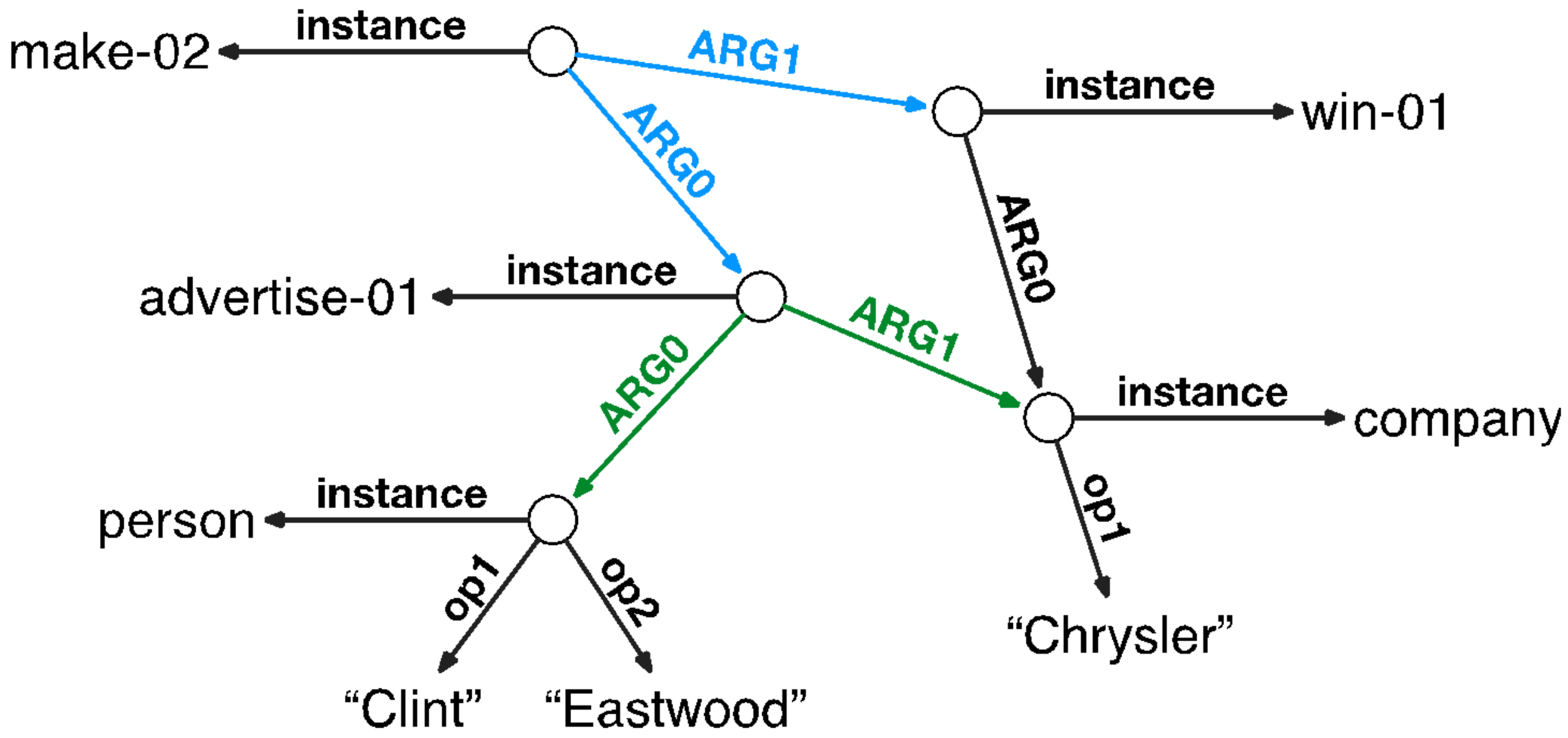
AMR: Relations

Clint Eastwood's ad for Chrysler makes them the winner.



AMR: Relations

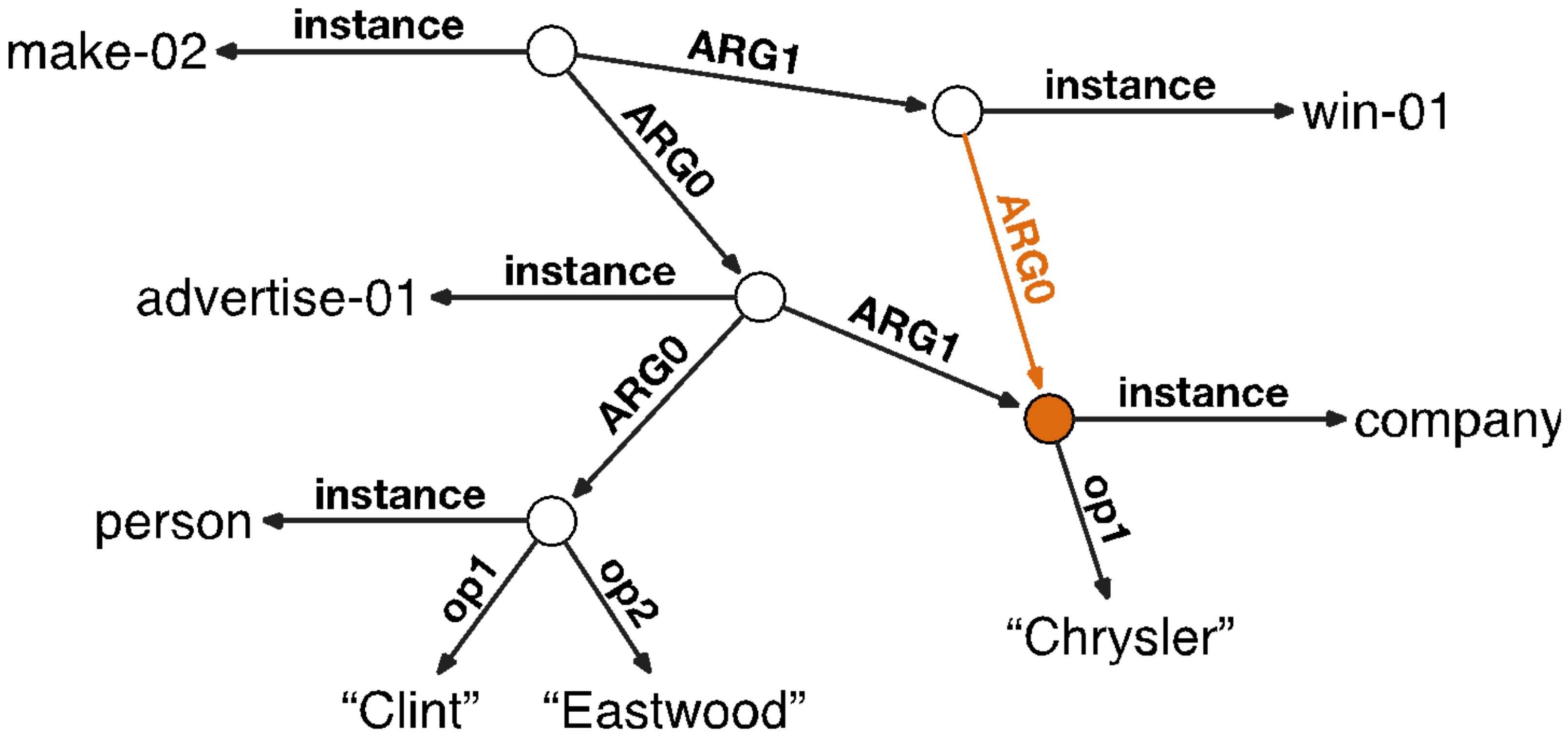
Clint Eastwood's ad for Chrysler makes them the winner.



JOHNS HOPKINS
UNIVERSITY

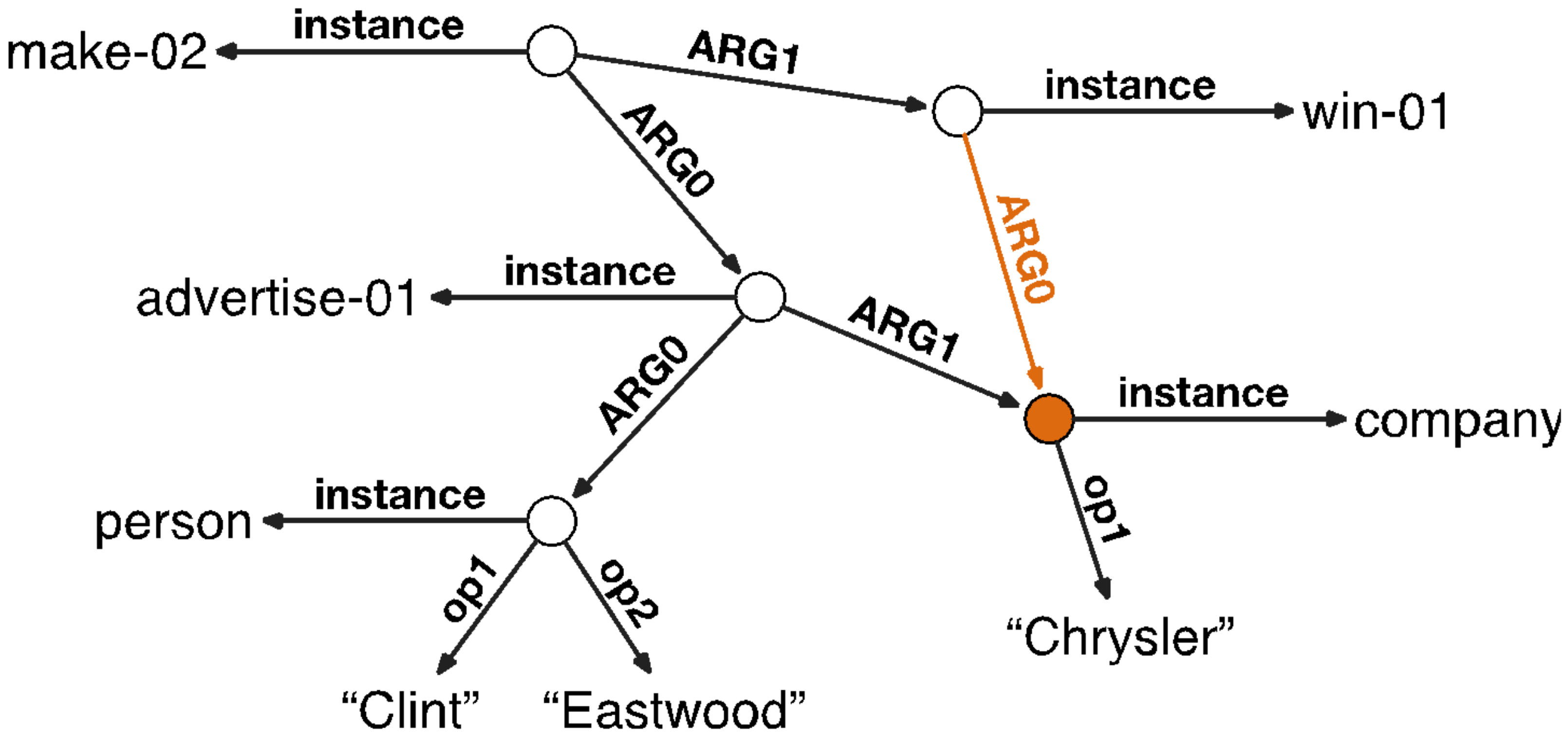
AMR: Reentrancy

Clint Eastwood's ad for Chrysler makes them the winner.

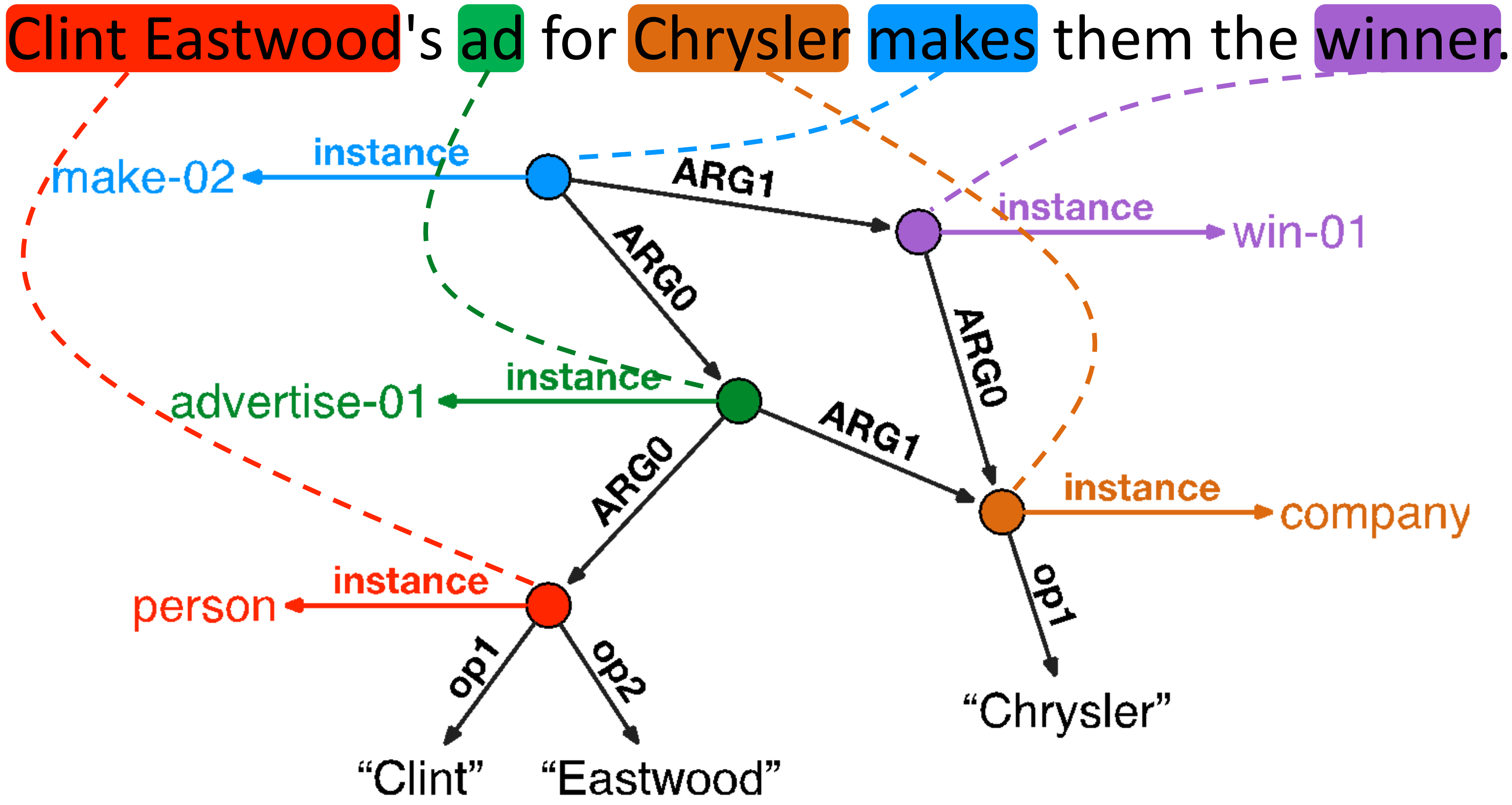


Challenge #1: Reentrancy

Clint Eastwood's ad for Chrysler makes them the winner.



Challenge #2: Lack of Alignment



Challenge #3: Limited Labeled Data

AMR 1.0 (LDC2014T12)

- ~10k **training** / 1k **development** / 1k **test** pairs

AMR 2.0 (LDC2017T10)

- ~37k **training** / 1k **development** / 1k **test** pairs

WMT18 en-zh

- ~22M **training** / 10k **development** / 2k **test** pairs

Three Challenges in AMR Parsing

1) **Reentrancy** 2) **Lack of Alignment** 3) **Limited Data**

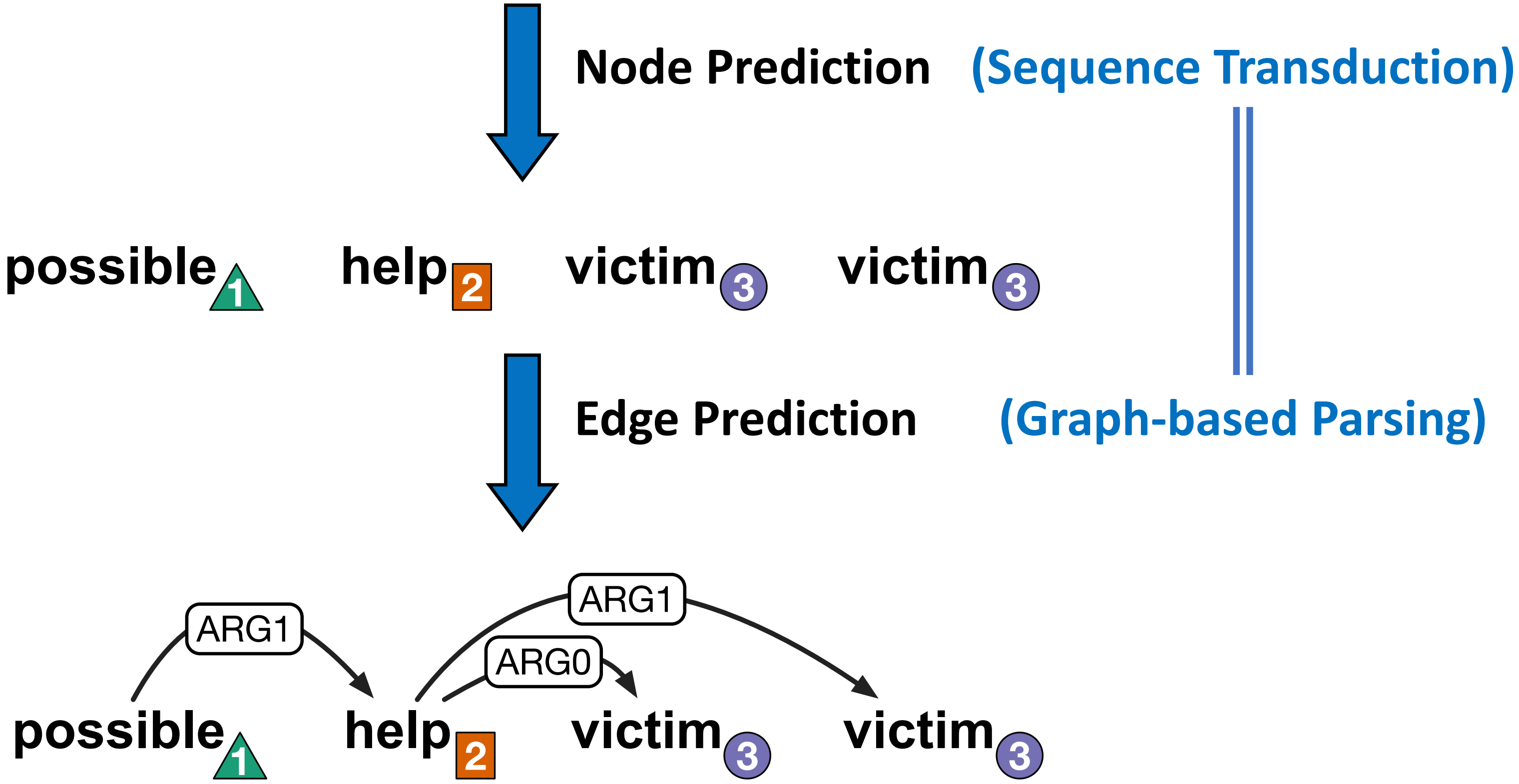
Is there a way to 1) **explicitly resolve reentrancy**,
2) **eliminate requirement for alignment**, and
3) **efficiently achieve SOTA w/ limited data**?

AMR Parsing as Sequence-to-Graph Transduction

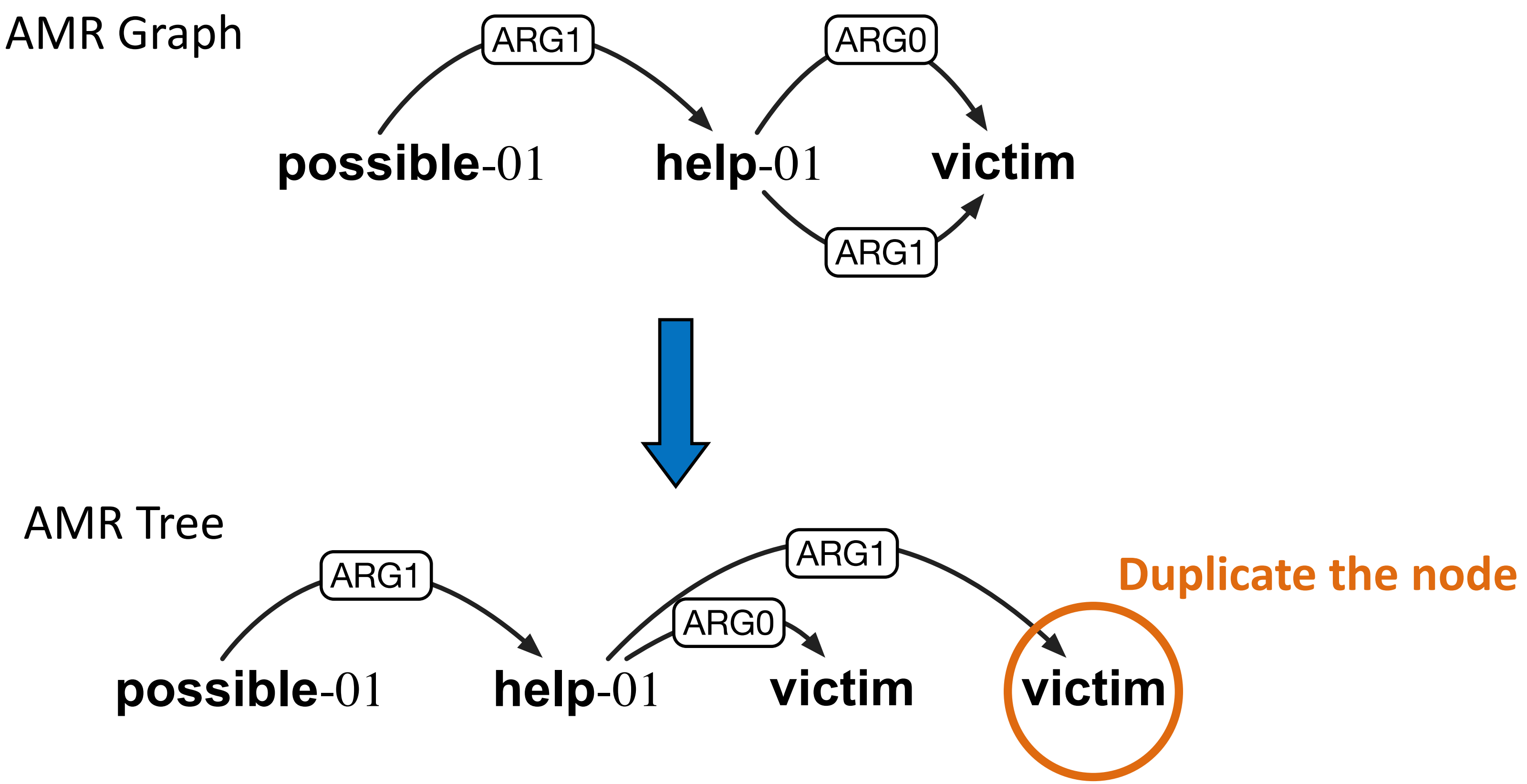
The victim could help himself.

Challenges

1) **Reentrancy**
2) ~~**Alignment**~~
3) ~~**Limited Data**~~

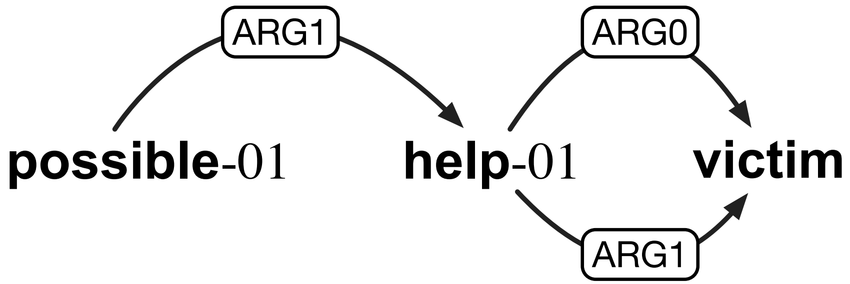


Another View of Reentrancy

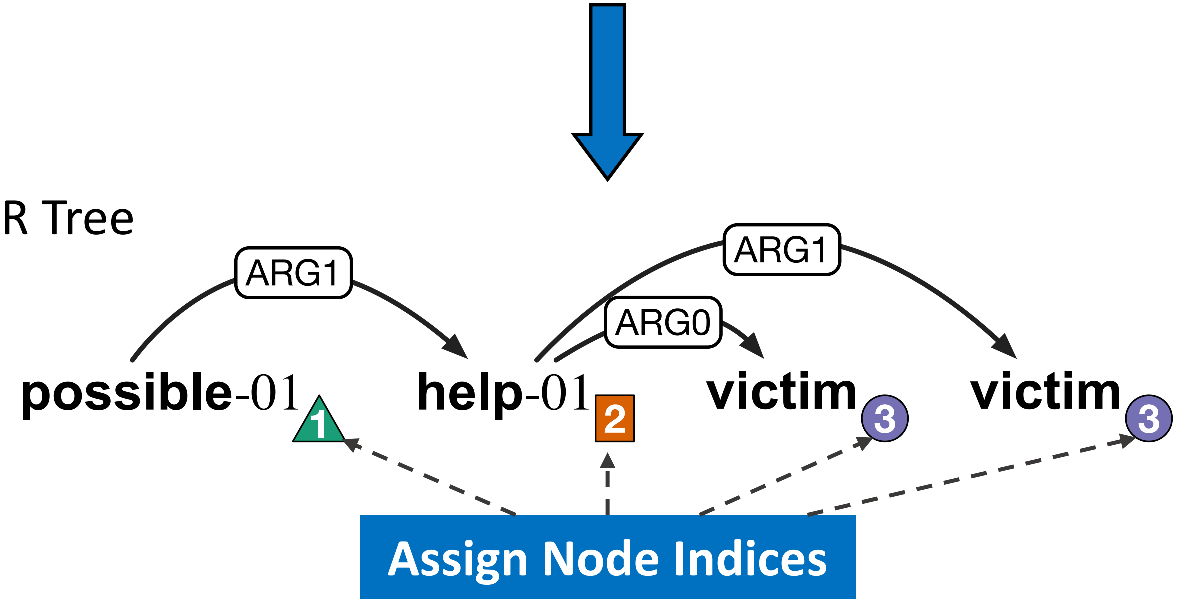


Another View of Reentrancy

AMR Graph



AMR Tree



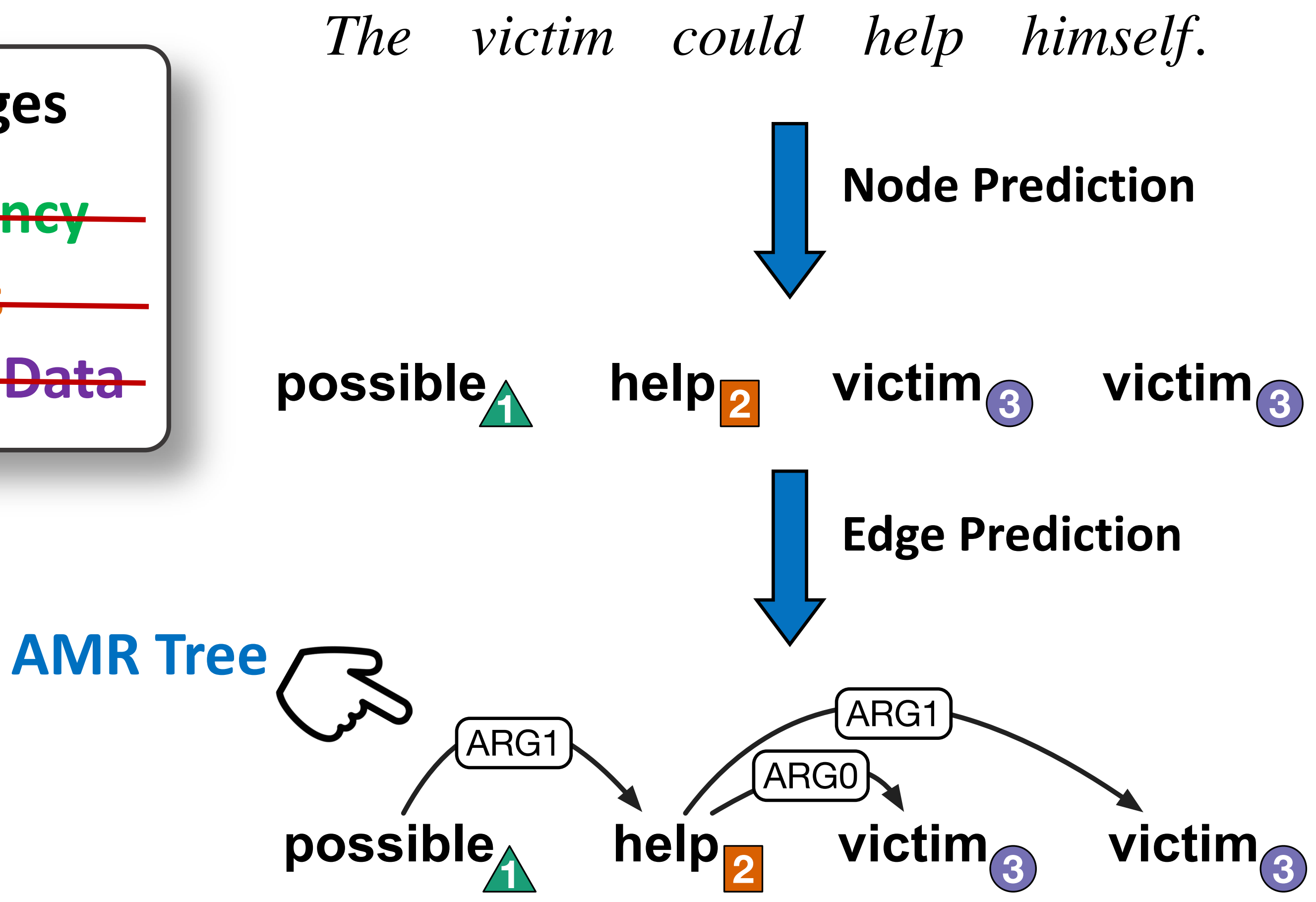
Our Approach in a Nutshell

Challenges

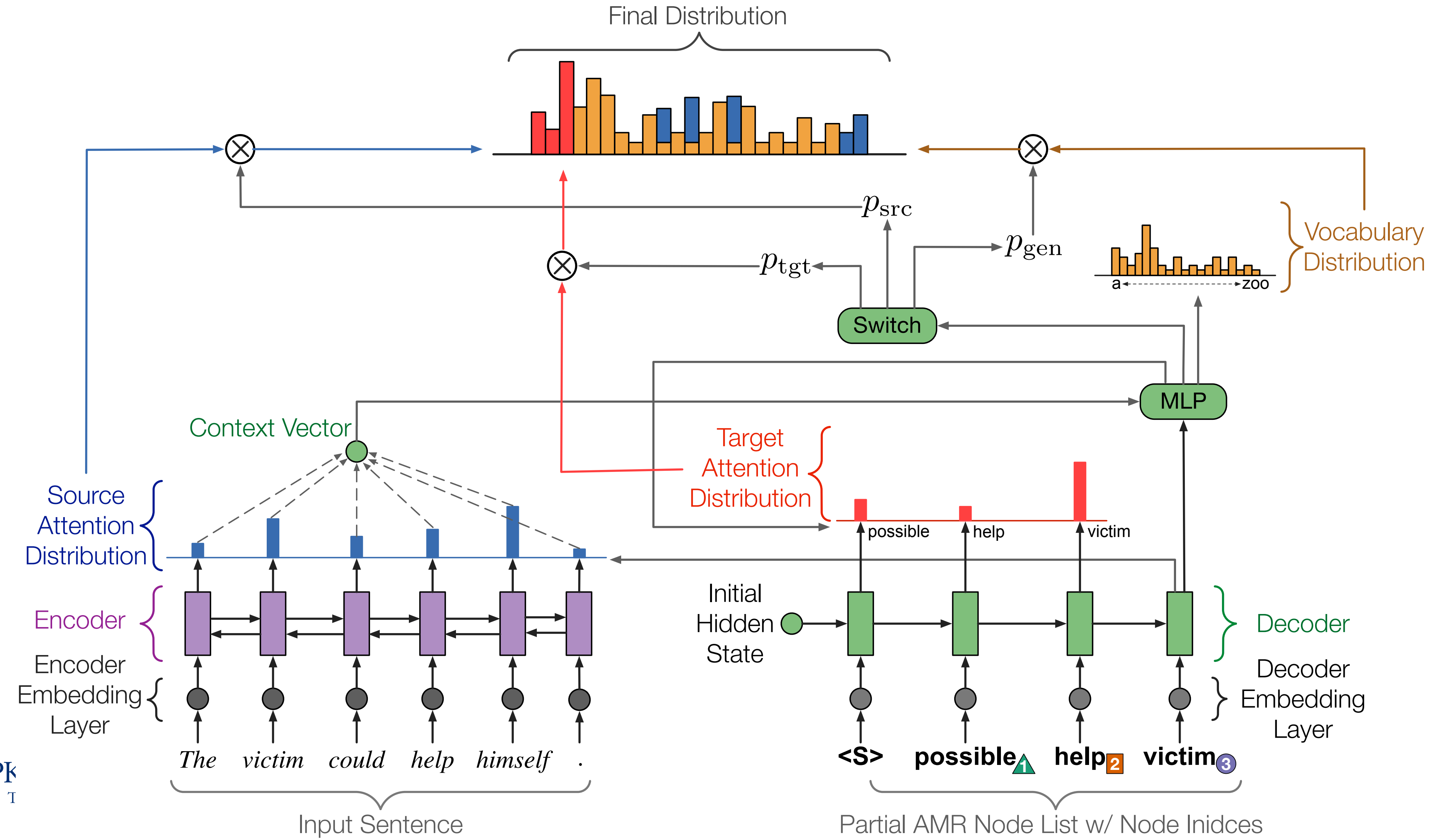
~~1) Reentrancy~~

~~2) Aligners~~

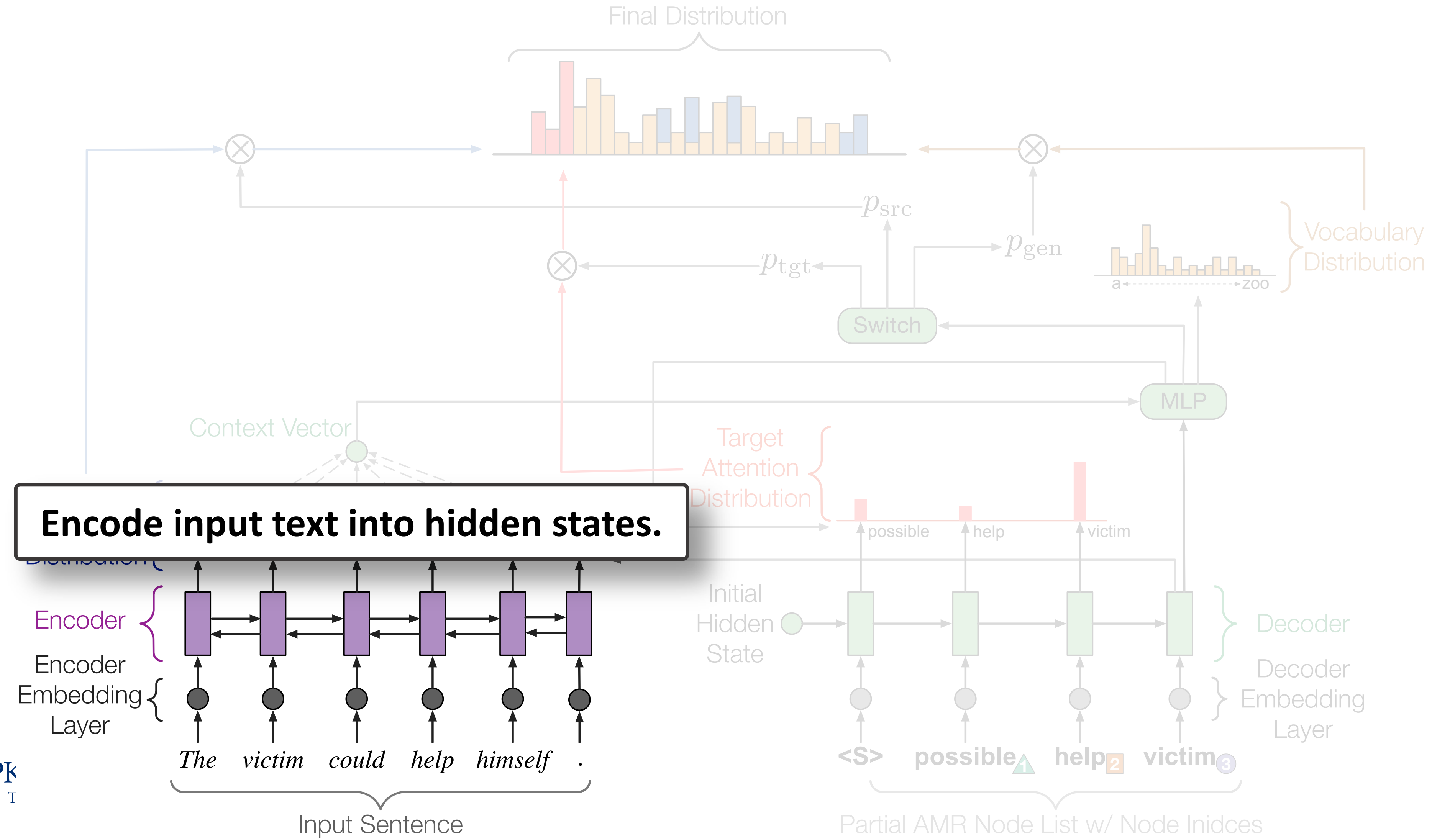
~~3) Limited Data~~



Extended Pointer-Generator Net
(Node Prediction)

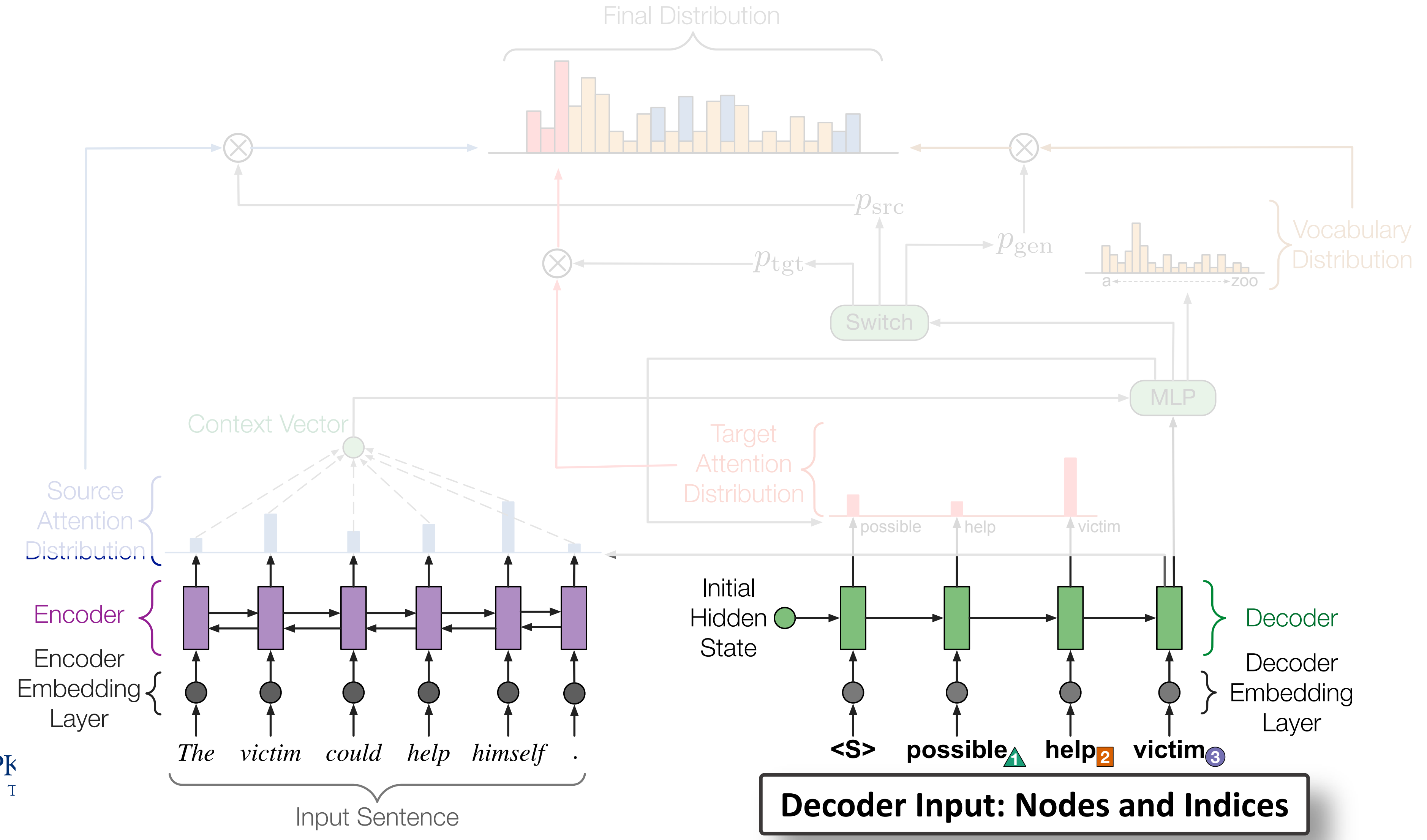


Extended Pointer-Generator Net
(Encoding)



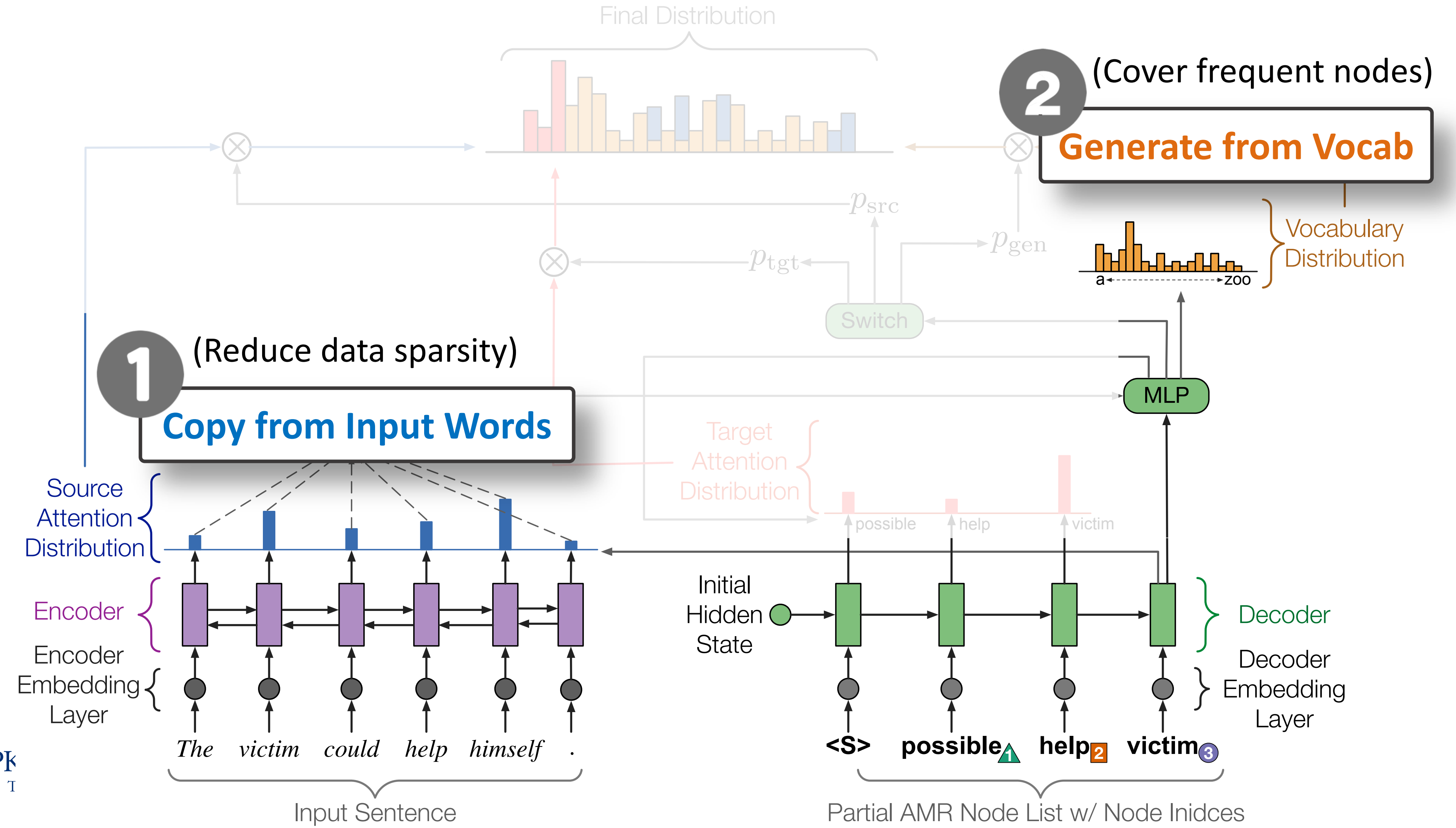
JOHNS HOPK UNIVERSIT

Extended Pointer-Generator Net
(Decoding)

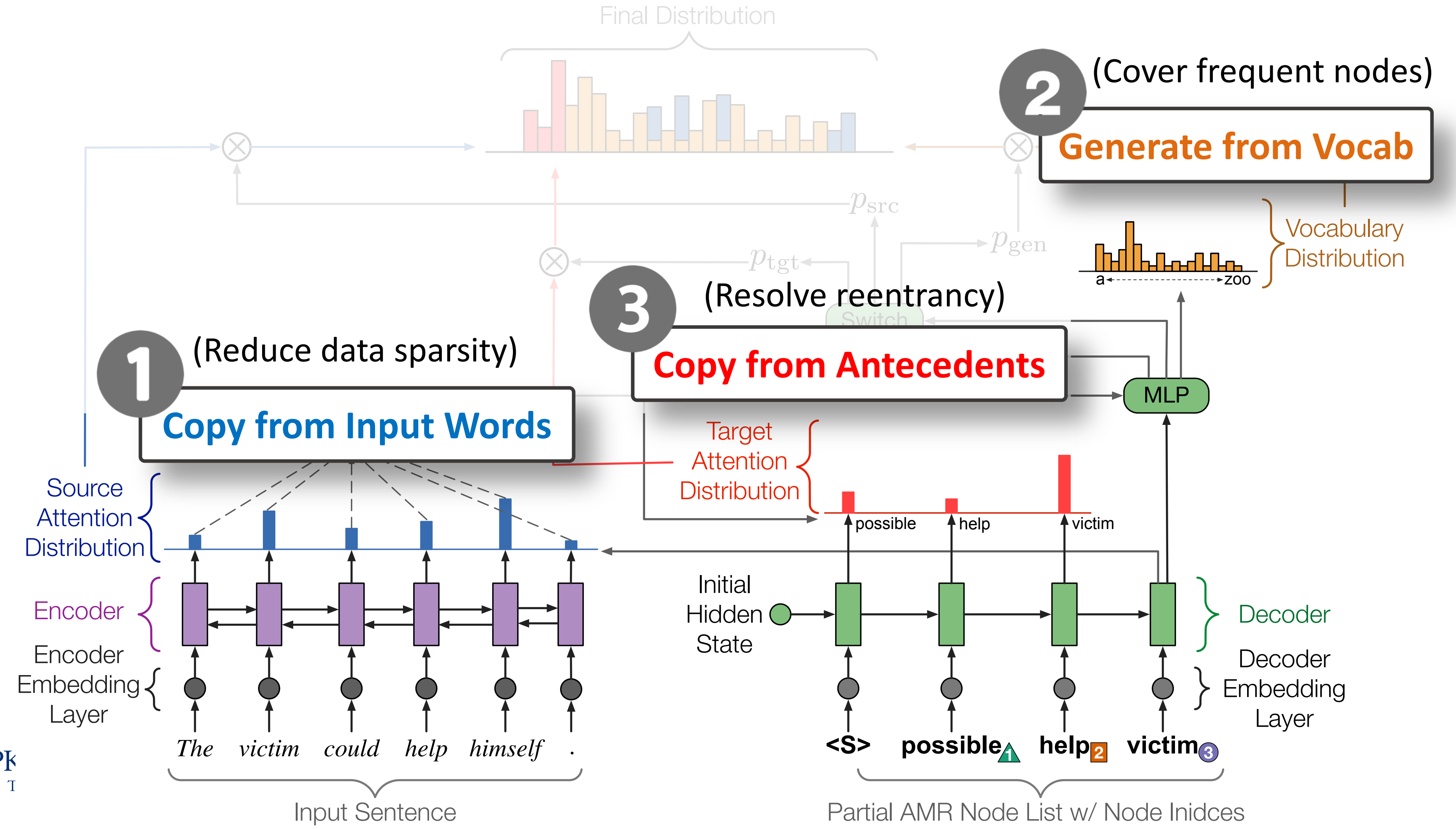


JOHNS HOPK UNIVERSIT

Extended Pointer-Generator Net
(Node Prediction)

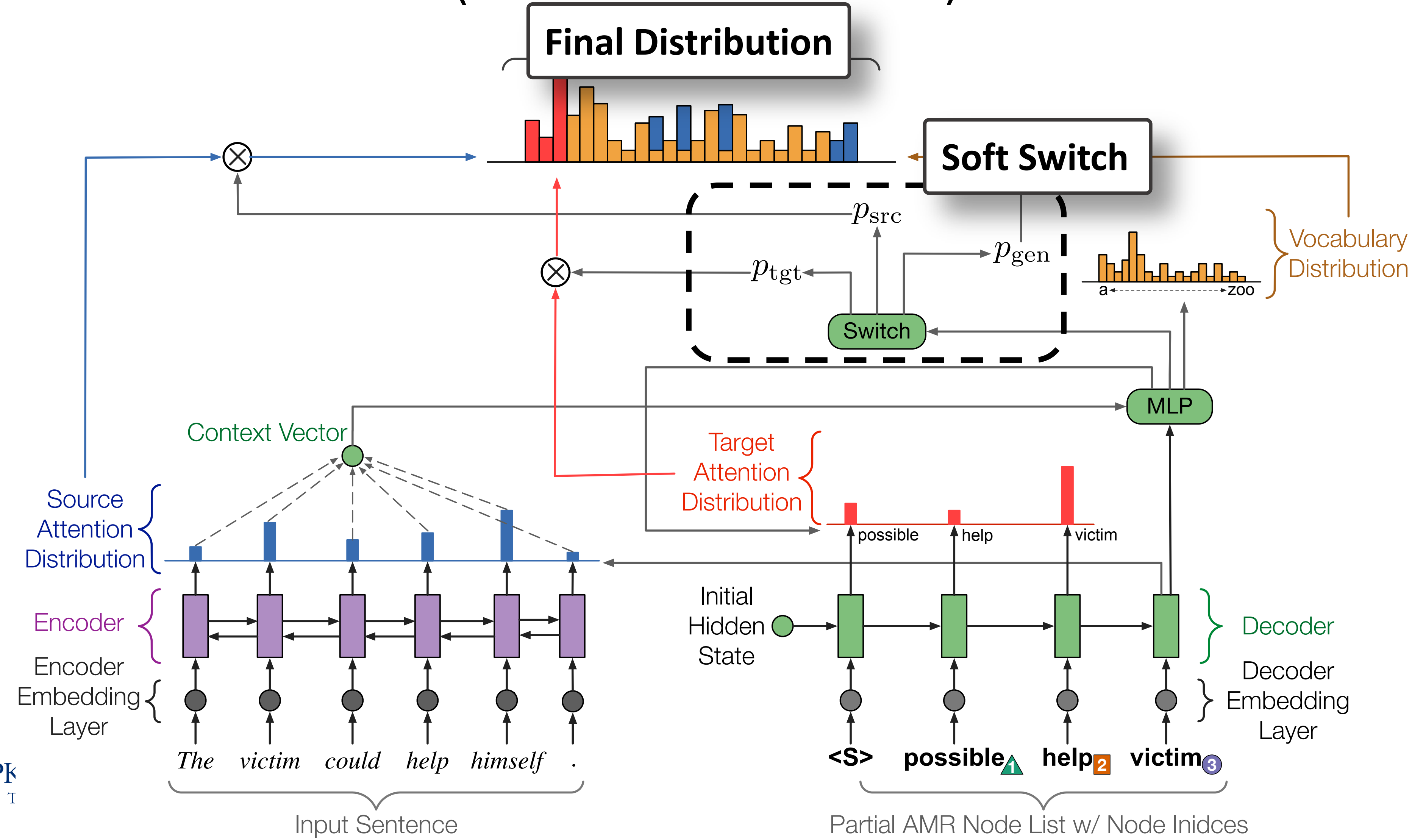


Extended Pointer-Generator Net
(Node Prediction)



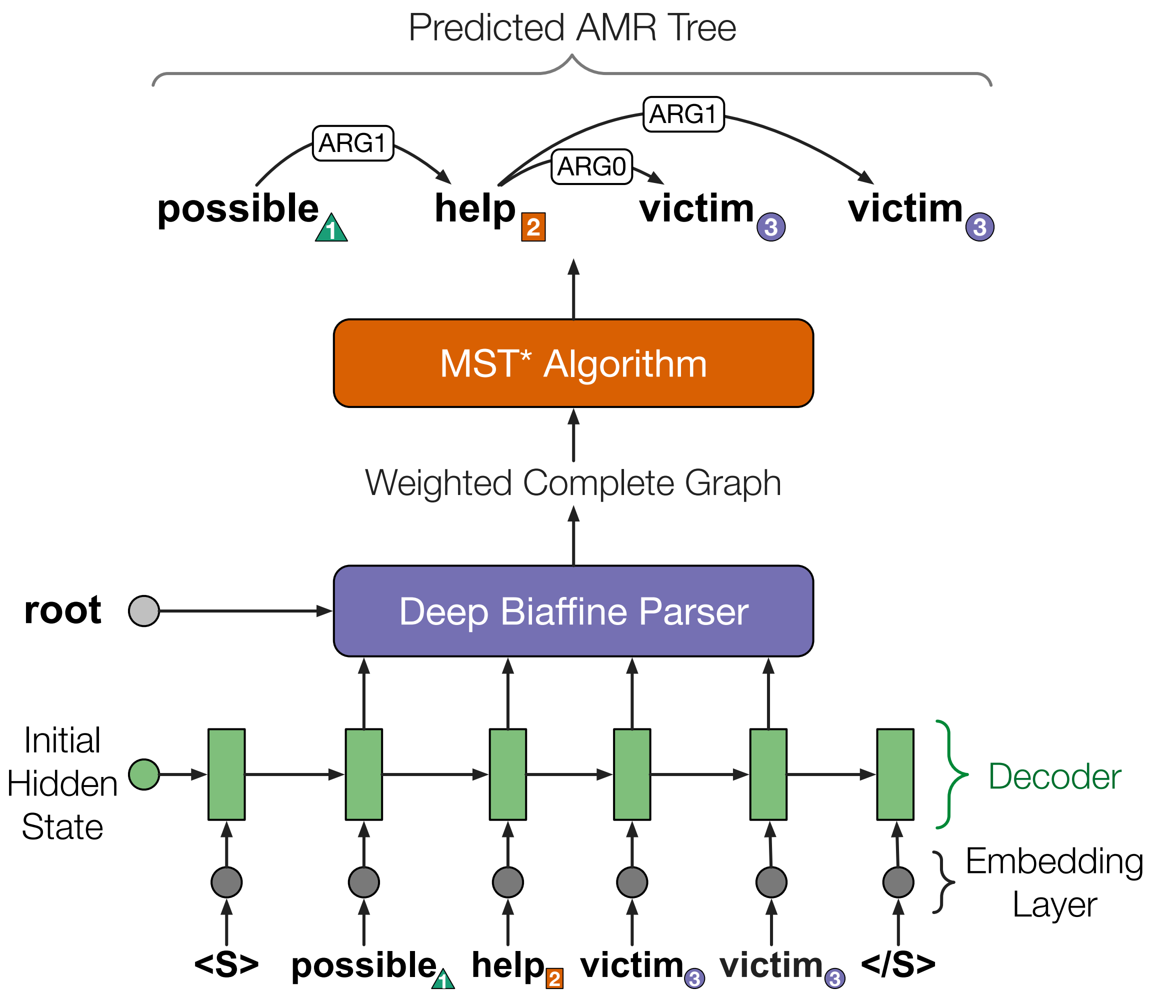
JOHNS HOPKINS UNIVERSITY

Extended Pointer-Generator Net
(Node Prediction)



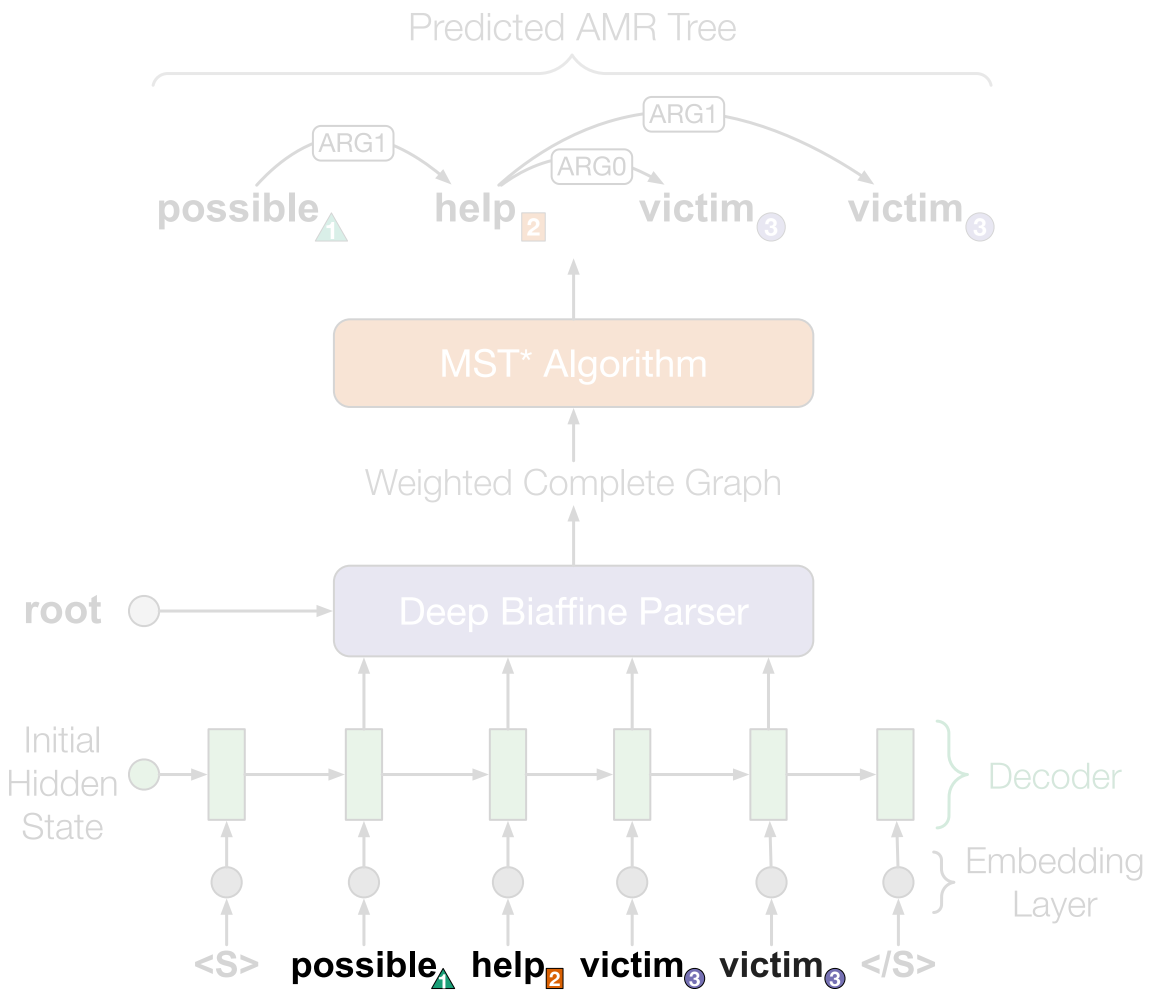
JOHNS HOPK
UNIVERSIT

Deep Biaffine Parser
(Edge Prediction)



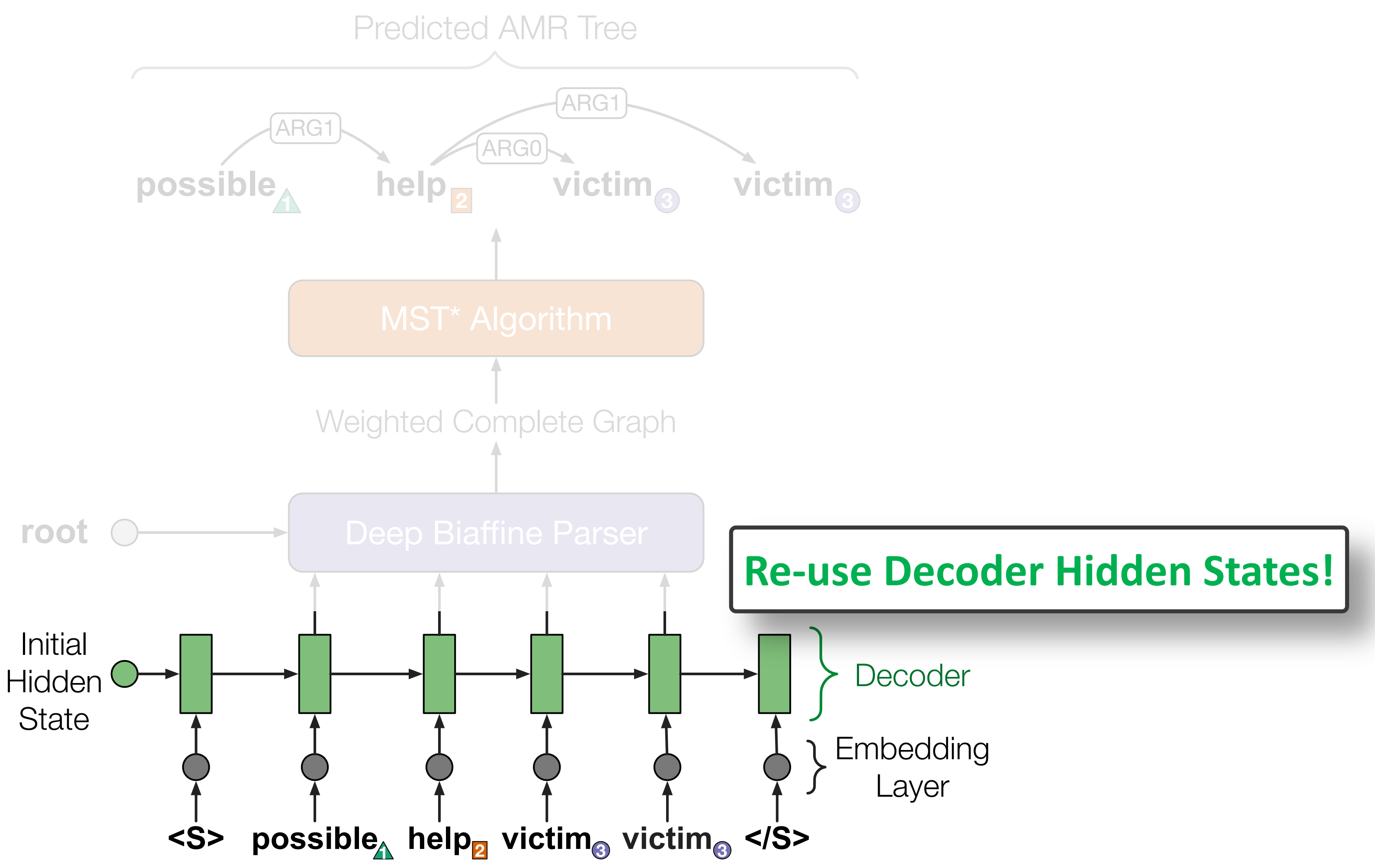
JOHNS HOPKINS UNIVERSITY

Deep Biaffine Parser
(Edge Prediction)



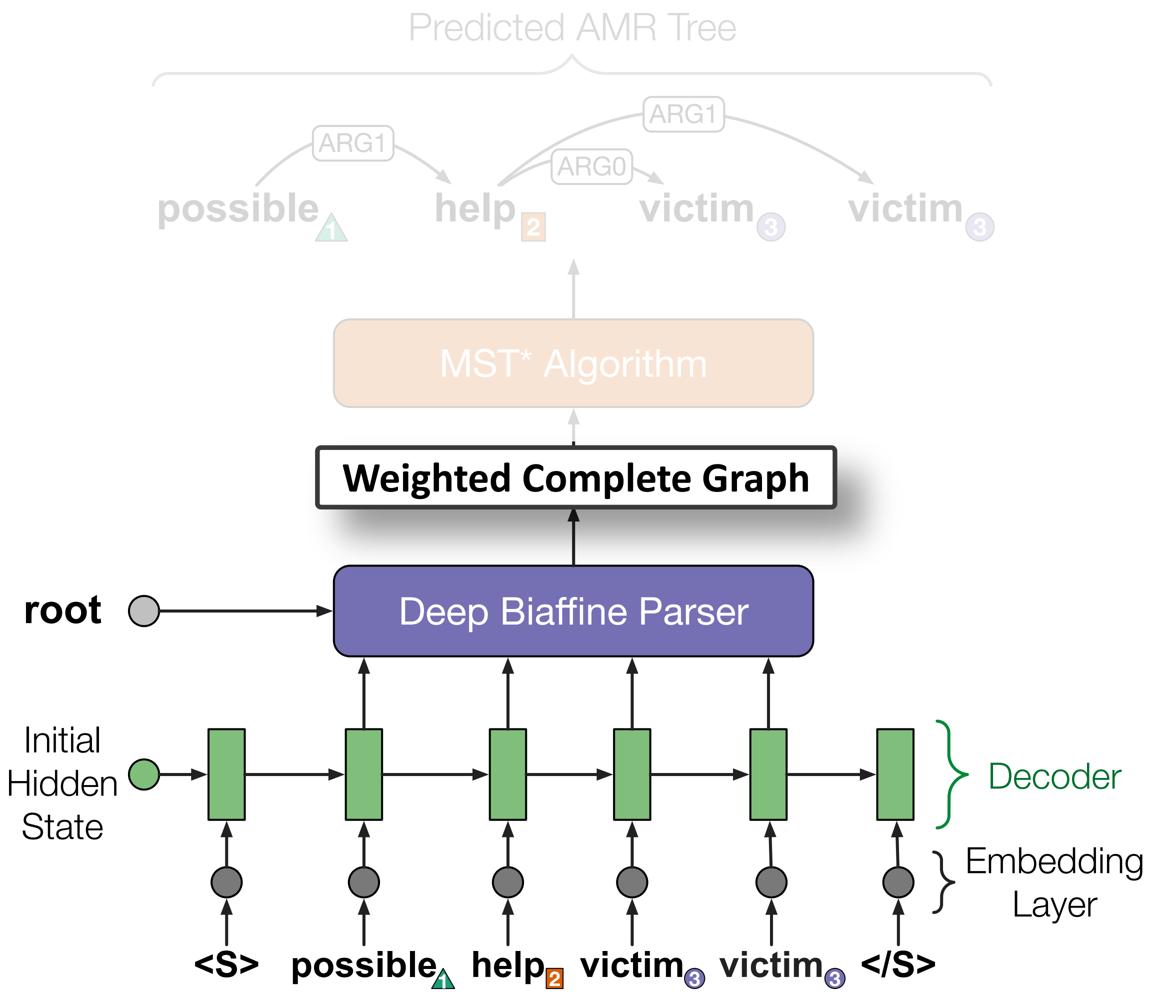
JOHNS HOPKINS
UNIVERSITY

Deep Biaffine Parser
(Edge Prediction)

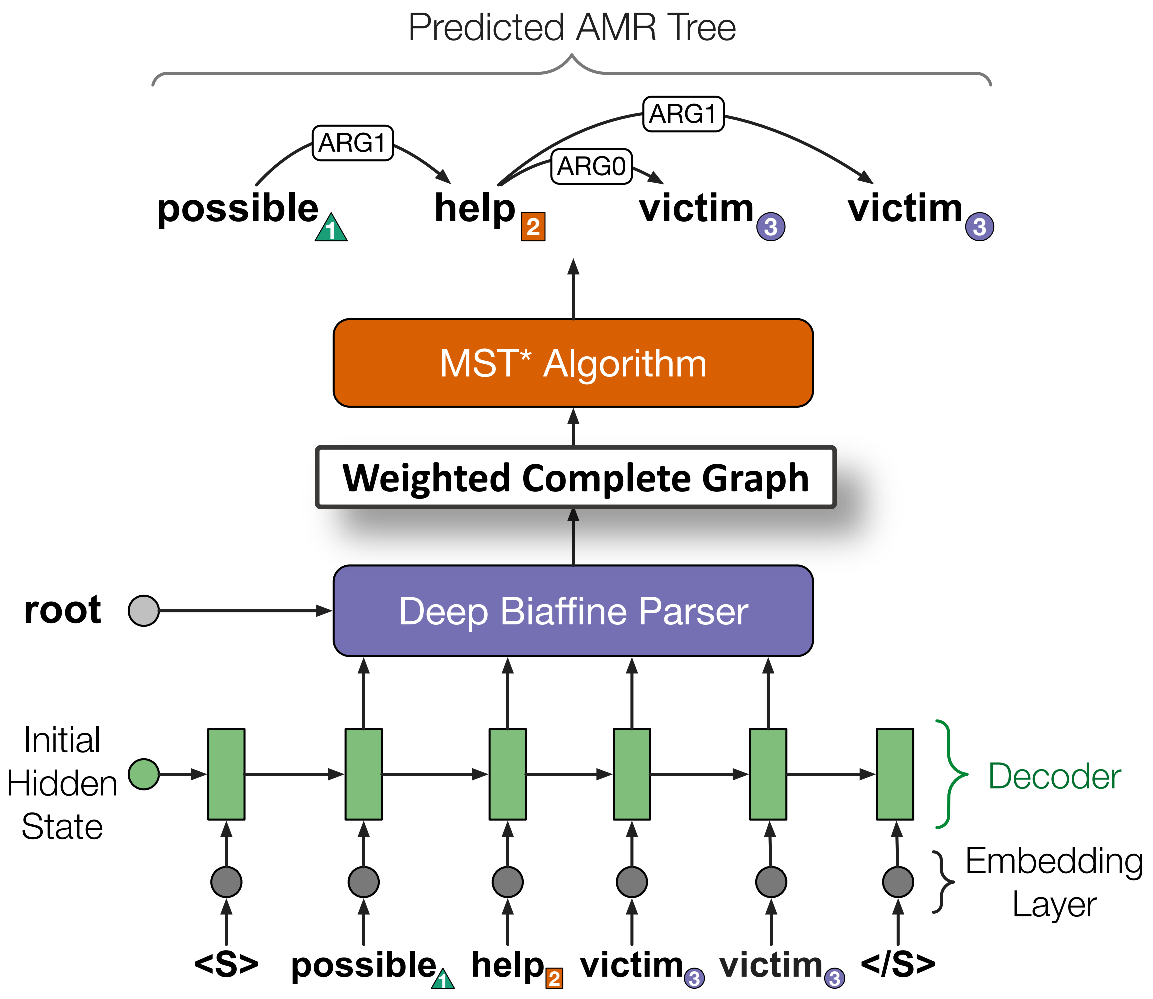


JOHNS HOPKINS UNIVERSITY

Deep Biaffine Parser
(Edge Prediction)



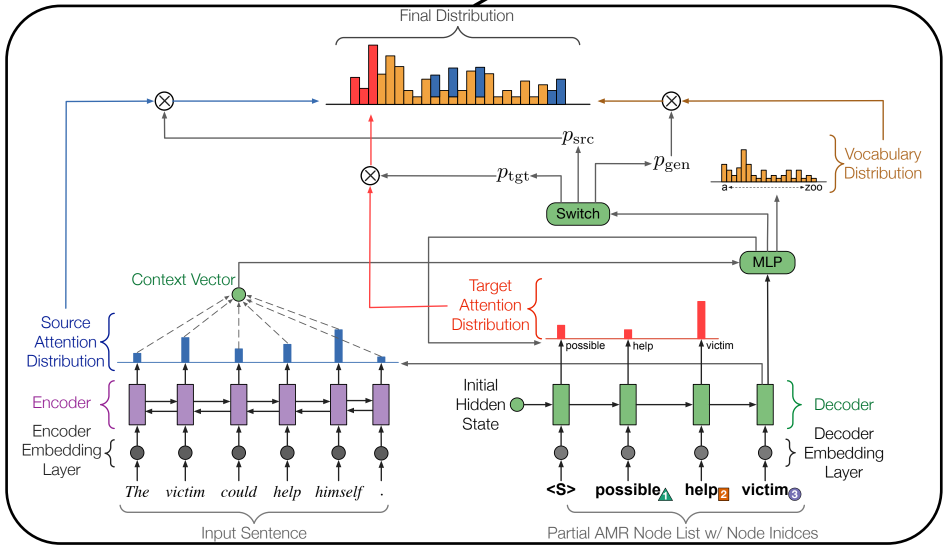
Deep Biaffine Parser
(Edge Prediction)



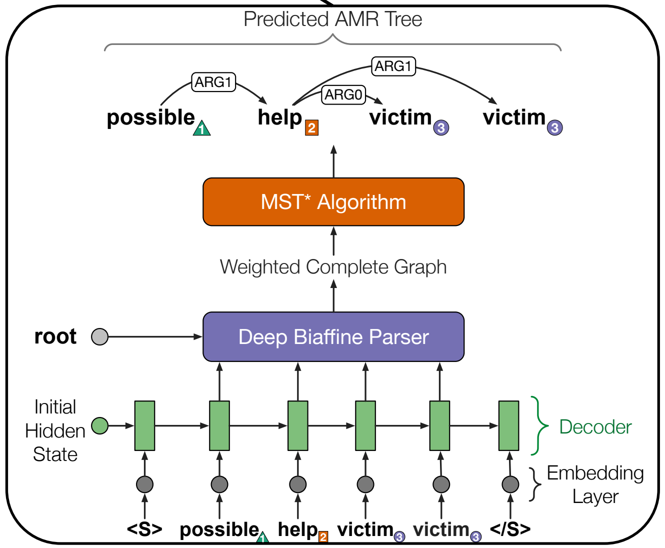
Johns Hopkins University

Multi-task Learning

$$\mathcal{L} = \mathcal{L}_{\text{node}} + \lambda \mathcal{L}_{\text{edge}}$$



Node Prediction



Edge Prediction

JOHNS HOPKINS UNIVERSITY

Experiments

AMR 1.0 (LDC2014T12)

- ~10k **training** / 1k **development** / 1k **test** pairs

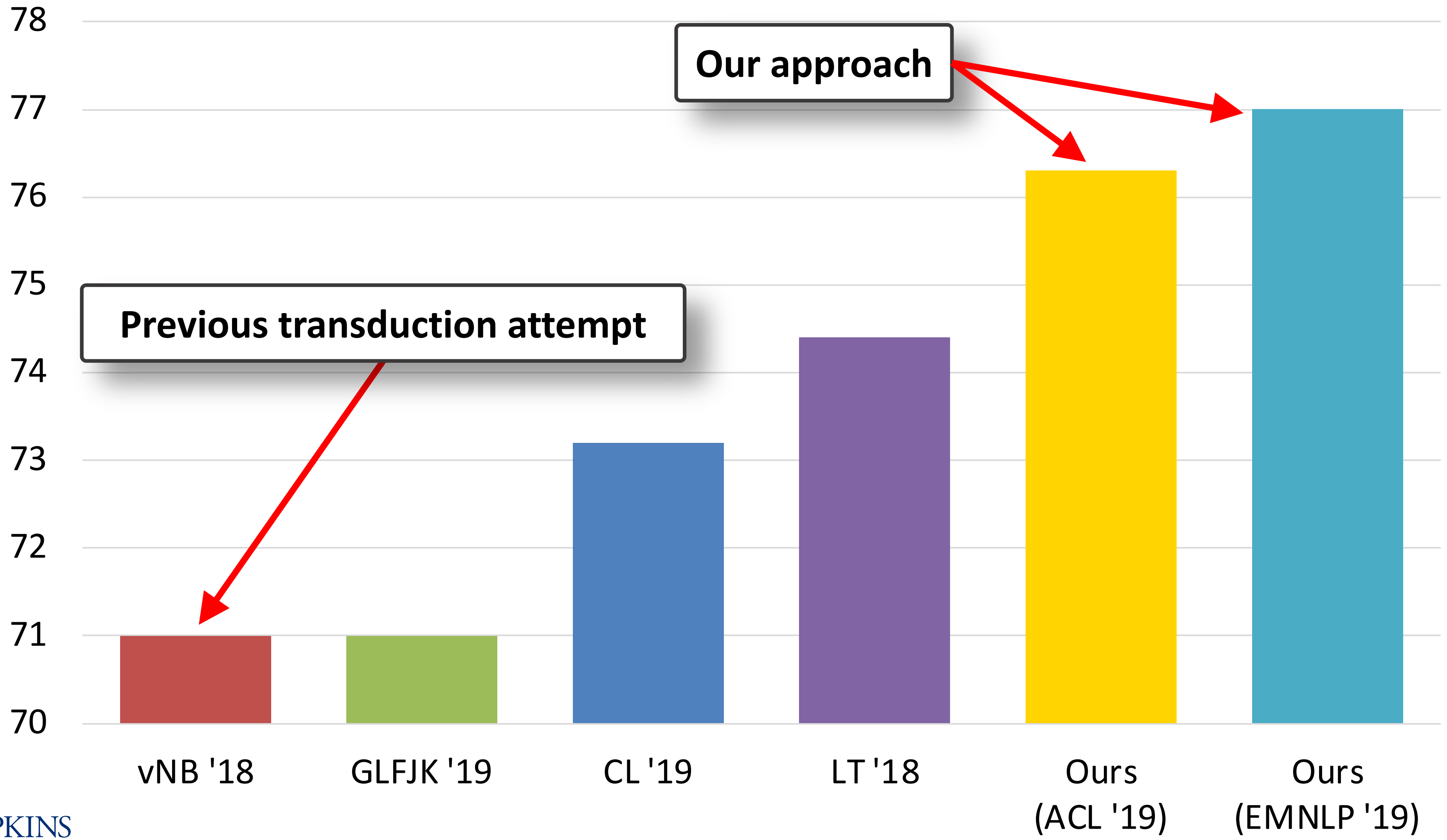
AMR 2.0 (LDC2017T10)

- ~37k **training** / 1k **development** / 1k **test** pairs

Metrics

- Smatch F1 (Cai and Knight, 2013)
- Fine-grained F-score (Damonte et al., 2017)

AMR Parsing %F1



Ablation Study

- Extended Pointer-Generator Modules

	AMR 1.0	AMR 2.0
Full model	70.2	76.3
no source-side copy	62.7	70.9
no target-side copy	66.2	71.6
no coverage loss	68.5	74.5
no BERT embeddings	68.8	74.6
no index embeddings	68.5	75.5
no anonym. indicator embed.	68.9	75.6
no beam search	69.2	75.3
no POS tag embeddings	69.2	75.7
no CharCNN features	70.0	75.8
only edge prediction	88.4	90.9

Ablation Study

- Extended Pointer-Generator Modules
- Distributed Language Representations

	AMR 1.0	AMR 2.0
Full model	70.2	76.3
no source-side copy	62.7	70.9
no target-side copy	66.2	71.6
no coverage loss	68.5	74.5
no BERT embeddings	68.8	74.6
no index embeddings	68.5	75.5
no anonym. indicator embed.	68.9	75.6
no beam search	69.2	75.3
no POS tag embeddings	69.2	75.7
no CharCNN features	70.0	75.8
only edge prediction	88.4	90.9

Ablation Study

- Extended Pointer-Generator Modules
- Distributed Language Representations
- Beam Search

	AMR 1.0	AMR 2.0
Full model	70.2	76.3
no source-side copy	62.7	70.9
no target-side copy	66.2	71.6
no coverage loss	68.5	74.5
no BERT embeddings	68.8	74.6
no index embeddings	68.5	75.5
no anonym. indicator embed.	68.9	75.6
no beam search	69.2	75.3
no POS tag embeddings	69.2	75.7
no CharCNN features	70.0	75.8
only edge prediction	88.4	90.9

Ablation Study

- Extended Pointer-Generator Modules
- Distributed Language Representations
- Beam Search
- Linguistic Features

	AMR 1.0	AMR 2.0
Full model	70.2	76.3
no source-side copy	62.7	70.9
no target-side copy	66.2	71.6
no coverage loss	68.5	74.5
no BERT embeddings	68.8	74.6
no index embeddings	68.5	75.5
no anonym. indicator embed.	68.9	75.6
no beam search	69.2	75.3
no POS tag embeddings	69.2	75.7
no CharCNN features	70.0	75.8
only edge prediction	88.4	90.9

Ablation Study

- Extended Pointer-Generator Modules
- Distributed Language Representations
- Beam Search
- Linguistic Features
- Node Prediction is the key!

	AMR 1.0	AMR 2.0
Full model	70.2	76.3
no source-side copy	62.7	70.9
no target-side copy	66.2	71.6
no coverage loss	68.5	74.5
no BERT embeddings	68.8	74.6
no index embeddings	68.5	75.5
no anonym. indicator embed.	68.9	75.6
no beam search	69.2	75.3
no POS tag embeddings	69.2	75.7
no CharCNN features	70.0	75.8
only edge prediction	88.4	90.9

AMR Sub-tasks %F1

Notable increase on sub-tasks:

- **Unlabeled**
- **Reentrancies**
- **Wikification**
- **Negation**

Metric	vN'17	G'18	L'18	Ours
SMATCH	71	71	74	**76.3**±0.1
Unlabeled	74	74	77	**79.0**±0.1
No WSD	72	72	76	**76.8**±0.1
Reentrancies	52	49	52	**60.0**±0.1
Concepts	82	84	**86**	84.8±0.1
Named Ent.	79	78	**86**	77.9±0.2
Wikification	65	71	76	**85.8**±0.3
Negation	62	57	58	**75.2**±0.2
SRL	66	64	**70**	69.7±0.2

Table 2: Fine-grained F1 scores on the AMR 2.0 test set. vN'17 is van Noord and Bos (2017); G'18 is Groschwitz et al. (2018); L'18 is Lyu and Titov (2018).

AMR Sub-tasks %F1

Notable increase on sub-tasks:

- Unlabeled
- Reentrancies
- Wikification
- Negation

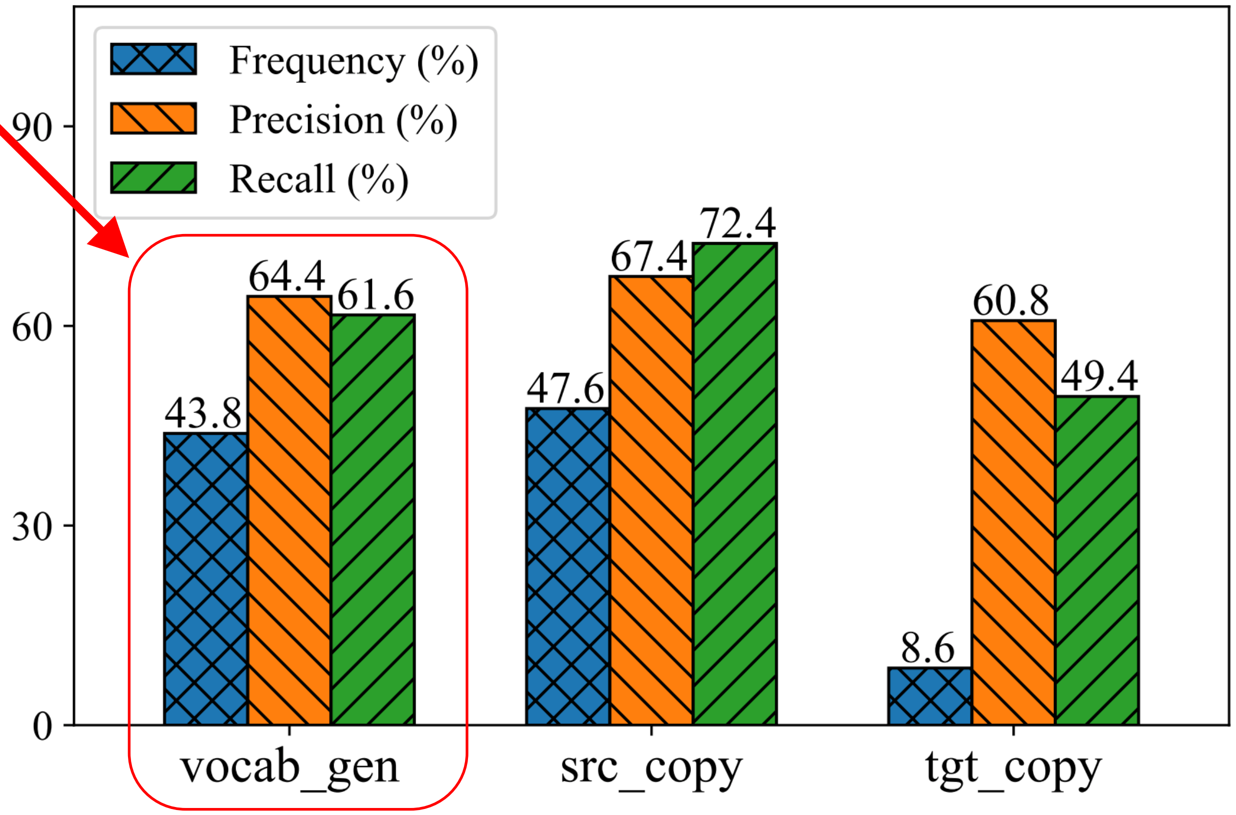
Metric	vN'17	G'18	L'18	Ours
SMATCH	71	71	74	**76.3**±0.1
Unlabeled	74	74	77	**79.0**±0.1
No WSD	72	72	76	**76.8**±0.1
Reentrancies	52	49	52	**60.0**±0.1
Concepts	82	84	**86**	84.8±0.1
Named Ent.	79	78	**86**	77.9±0.2
Wikification	65	71	76	**85.8**±0.3
Negation	62	57	58	**75.2**±0.2
SRL	66	64	**70**	69.7±0.2

Table 2: Fine-grained F1 scores on the AMR 2.0 test set. vN'17 is van Noord and Bos (2017); G'18 is Groschwitz et al. (2018); L'18 is Lyu and Titov (2018).

JOHNS HOPKINS UNIVERSITY

Efficiency of Source and Target Copy

43.8% are novel nodes.

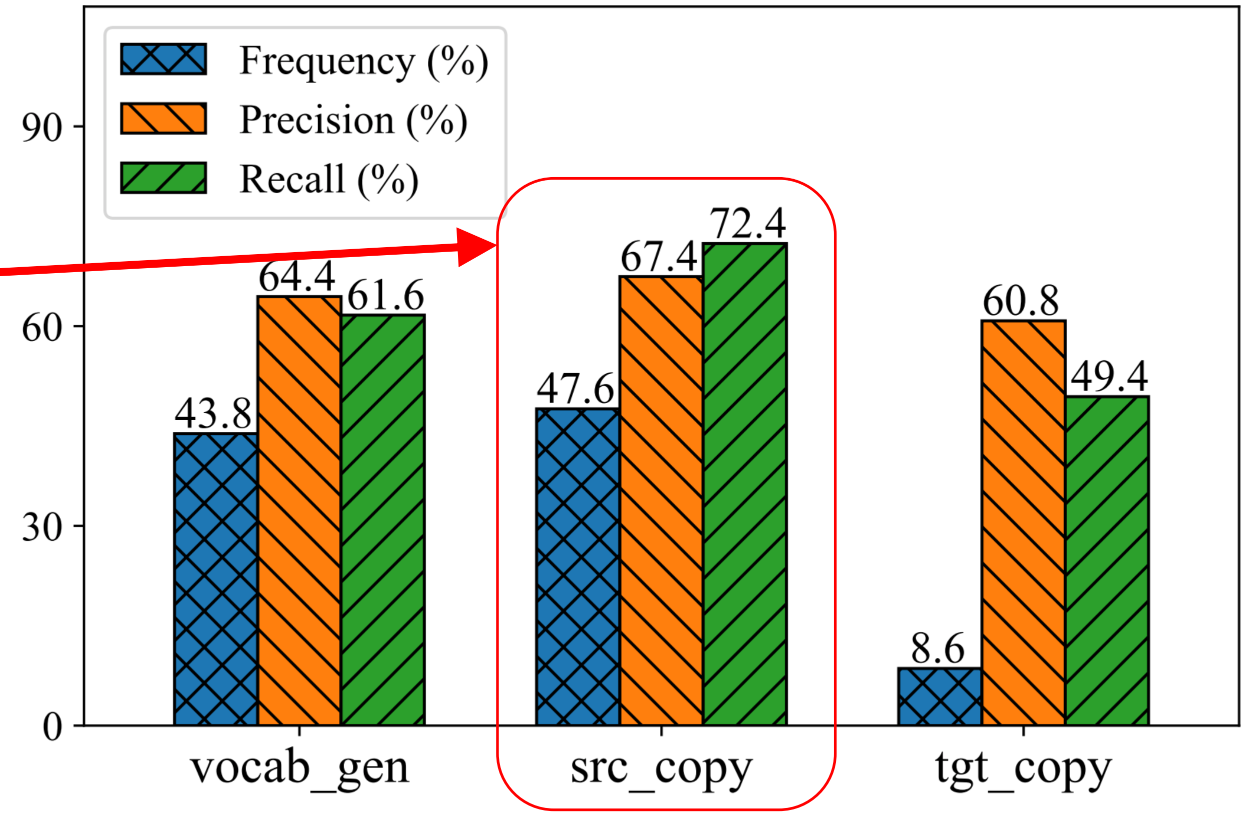


Johns Hopkins University

Efficiency of Source and Target Copy

43.8% are novel nodes.

47.6% are from source-side copy.



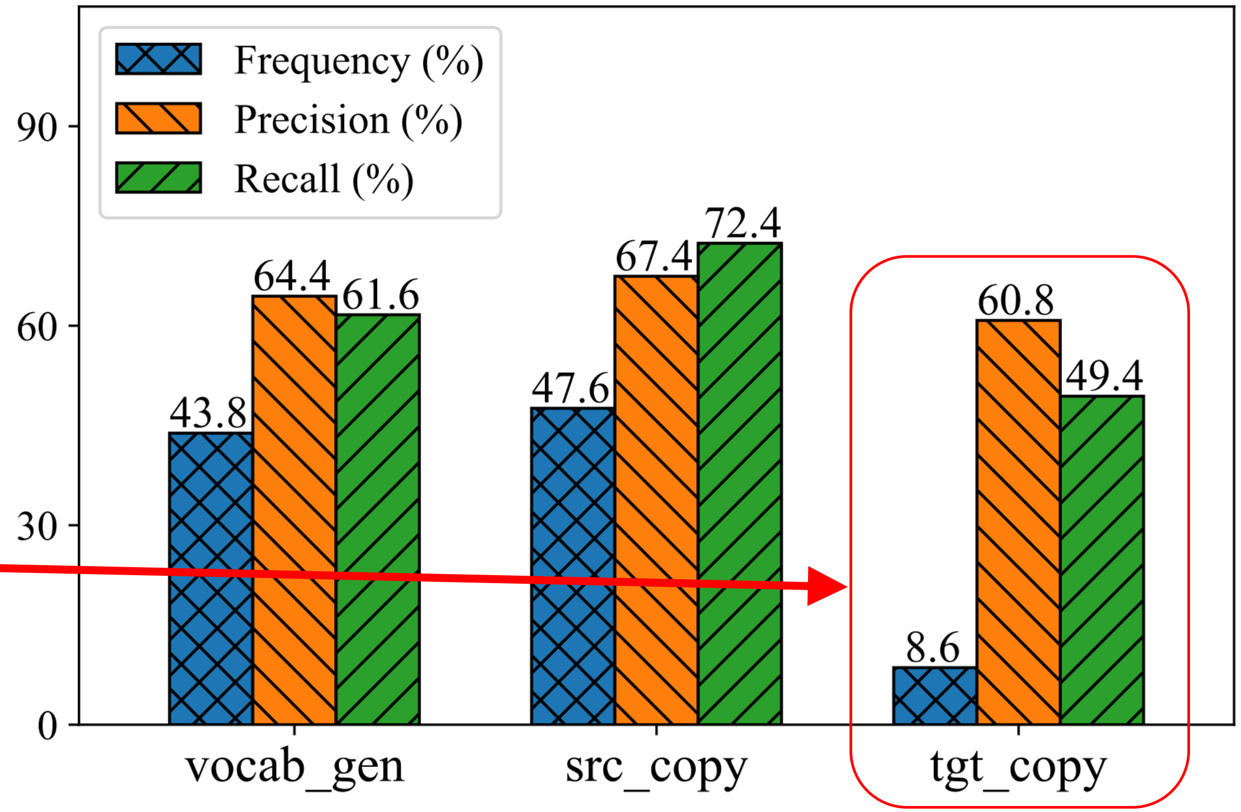
JOHNS HOPKINS UNIVERSITY

Efficiency of Source and Target Copy

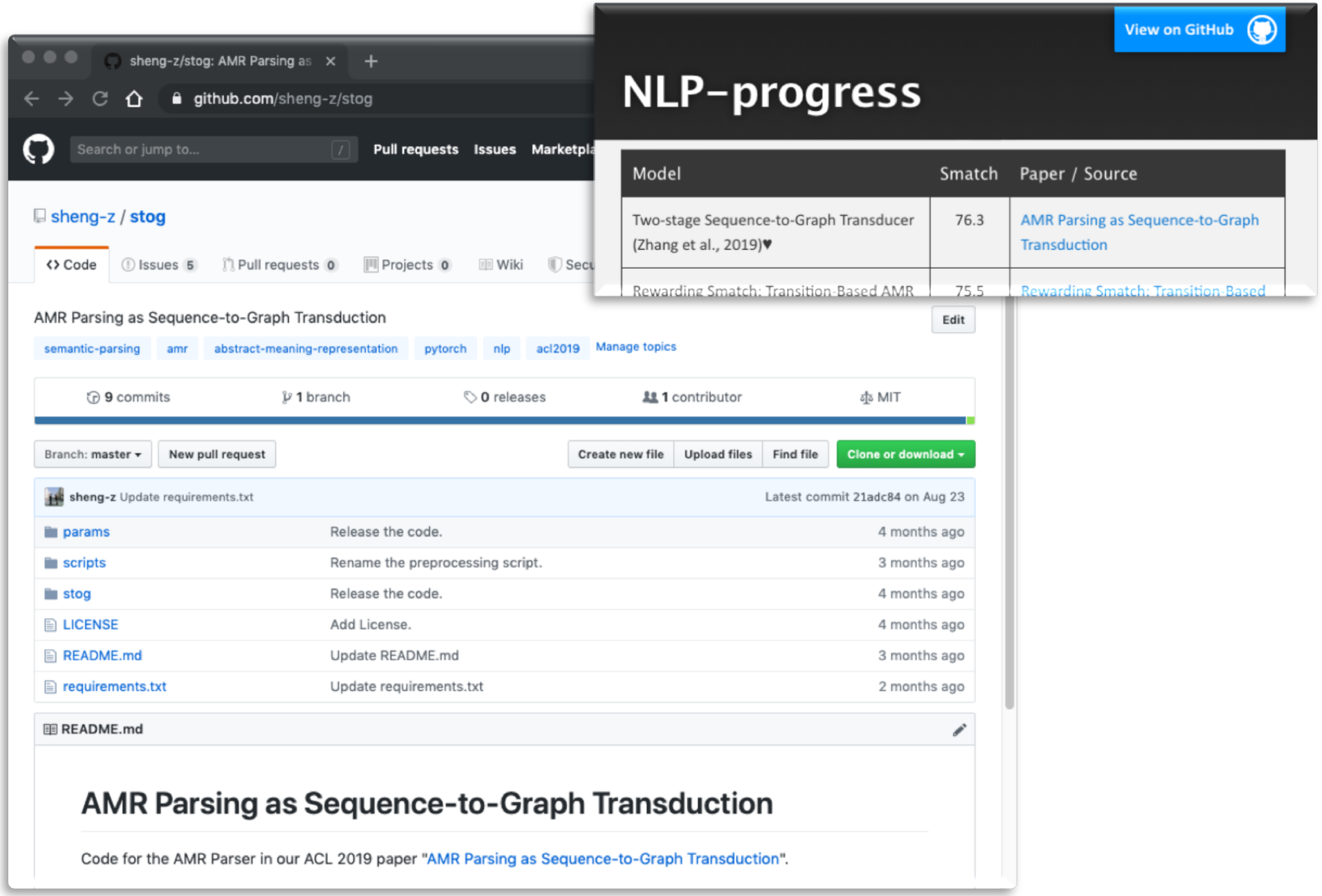
43.8% are novel nodes.

47.6% are from source-side copy.

Only 8.6% are from target-side copy.



Code Released!



JOHNS HOPKINS UNIVERSITY