UNIVERSITY of WASHINGTON

ALLEN INSTITUTE for ARTIFICIAL INTELLIGENCE

UWNLP

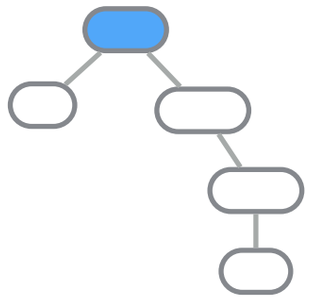
Neural AMR: Sequence-to-Sequence Models for **Parsing** and **Generation**

Ioannis Konstas

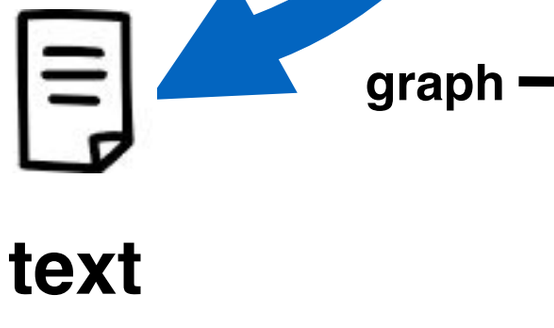
HERIOT WATT UNIVERSITY

joint work with **Srinivasan Iyer, Mark Yatskar, Yejin Choi, Luke Zettlemoyer**

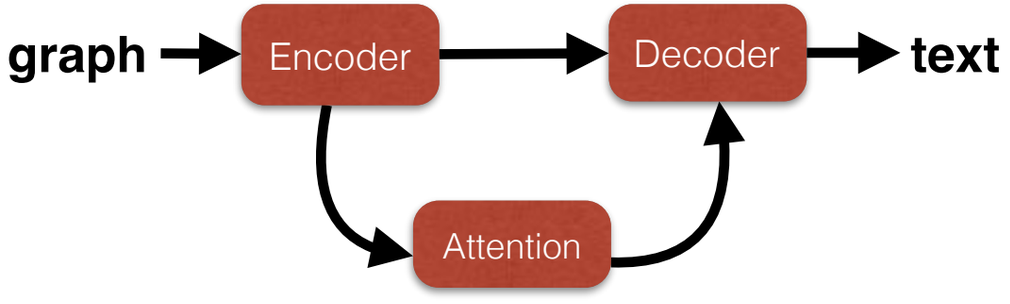
AMR graph



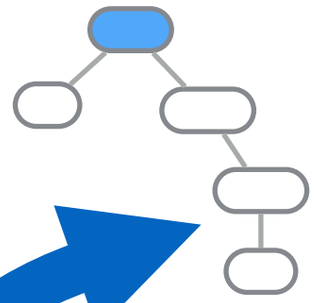
Generate from AMR



text

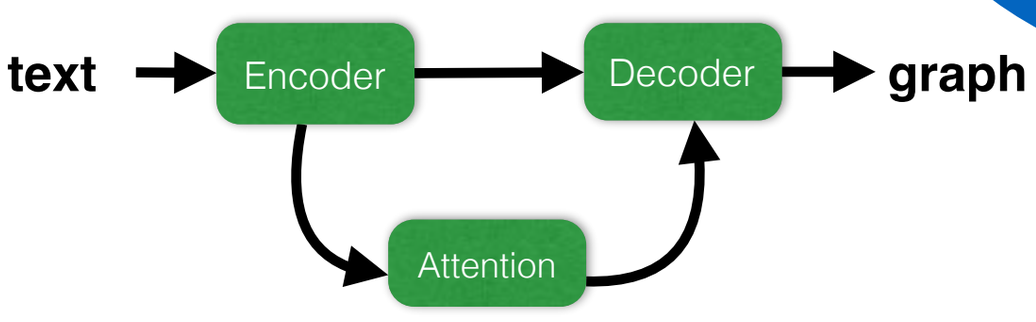


AMR graph

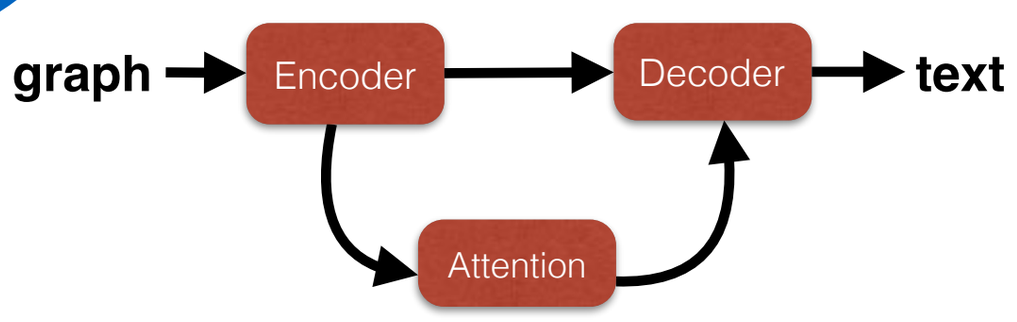


Parse to AMR

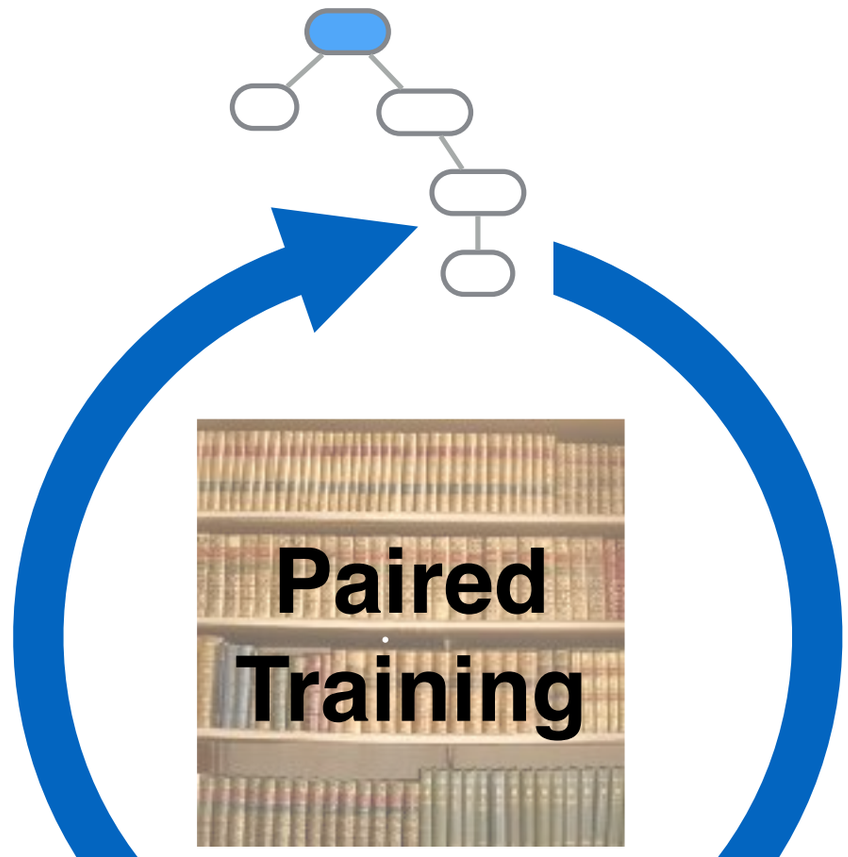
Generate from AMR



text

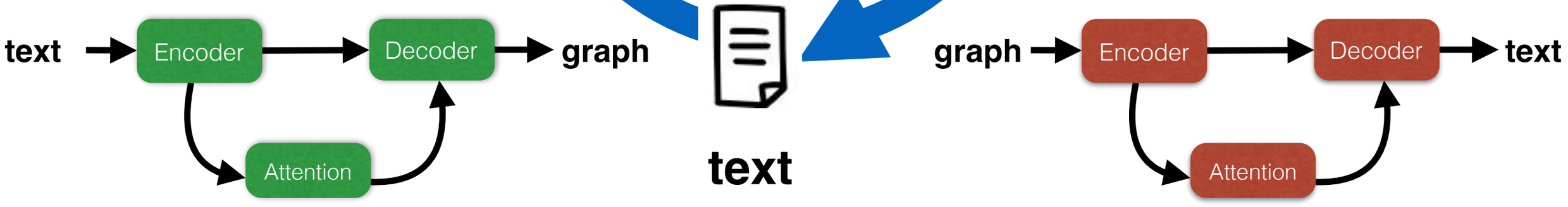


AMR graph



Parse to AMR

Generate from AMR

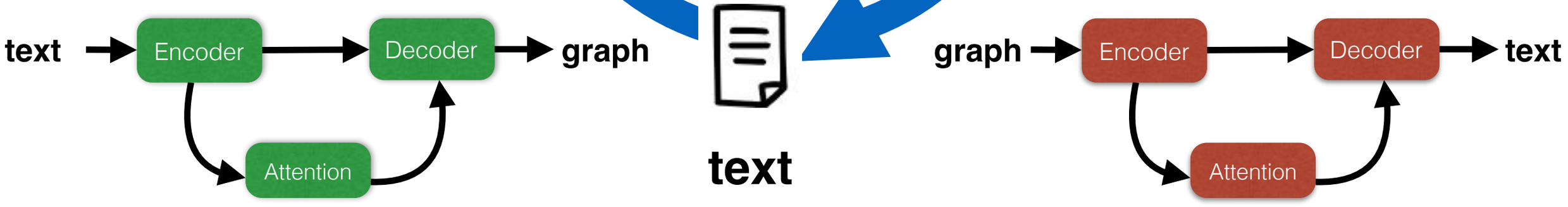


AMR graph

SOTA

Parse to AMR

Generate from AMR

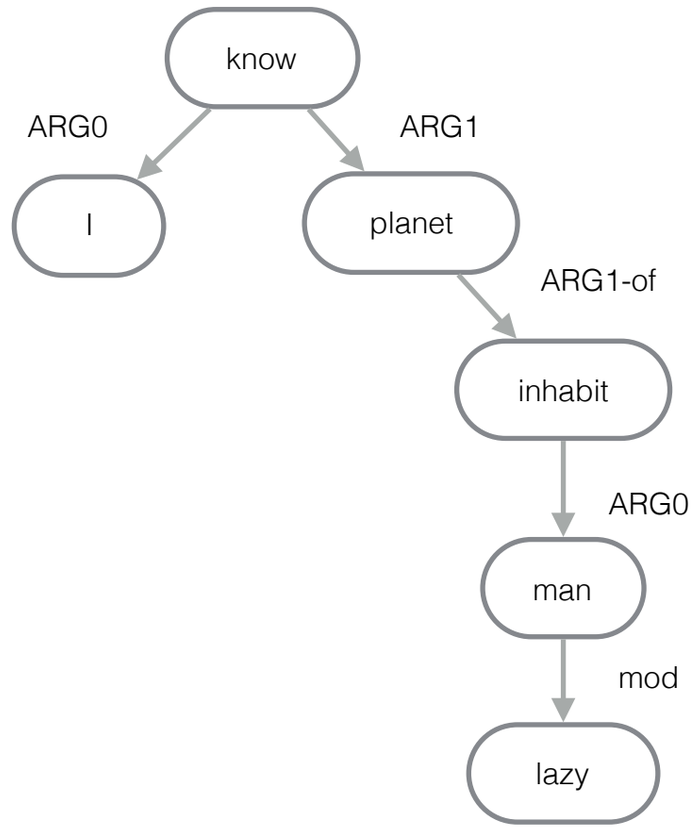


Abstract Meaning Representation

(Banarescu et al., 2013)

- Rooted Directed Acyclic Graph
- Nodes: concepts (nouns, verbs, named entities, etc)
- Edges: Semantic Role Labels

I have **known** a **planet** that was **inhabited** by a **lazy man**.

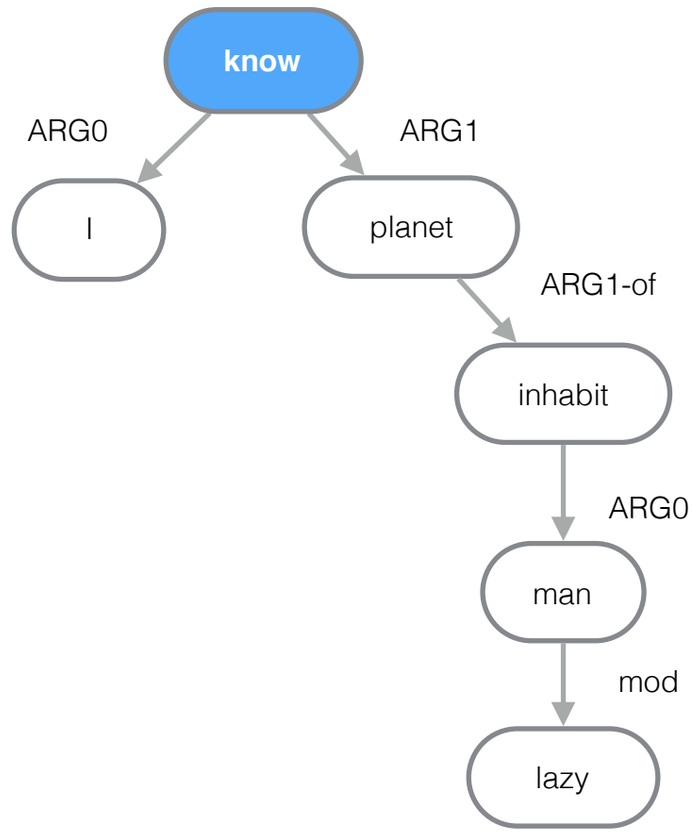


Abstract Meaning Representation

(Banarescu et al., 2013)

- Rooted Directed Acyclic Graph
- Nodes: concepts (nouns, verbs, named entities, etc)
- Edges: Semantic Role Labels

I have **known** a **planet** that was **inhabited** by a **lazy man**.

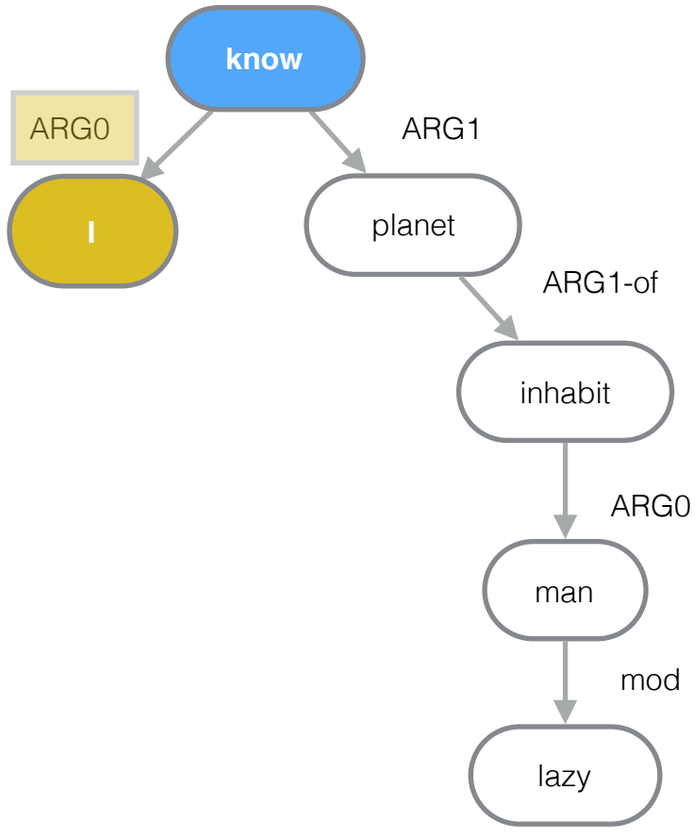


Abstract Meaning Representation

(Banarescu et al., 2013)

- Rooted Directed Acyclic Graph
- Nodes: concepts (nouns, verbs, named entities, etc)
- Edges: Semantic Role Labels

I have **known** a **planet** that was **inhabited** by a **lazy man**.

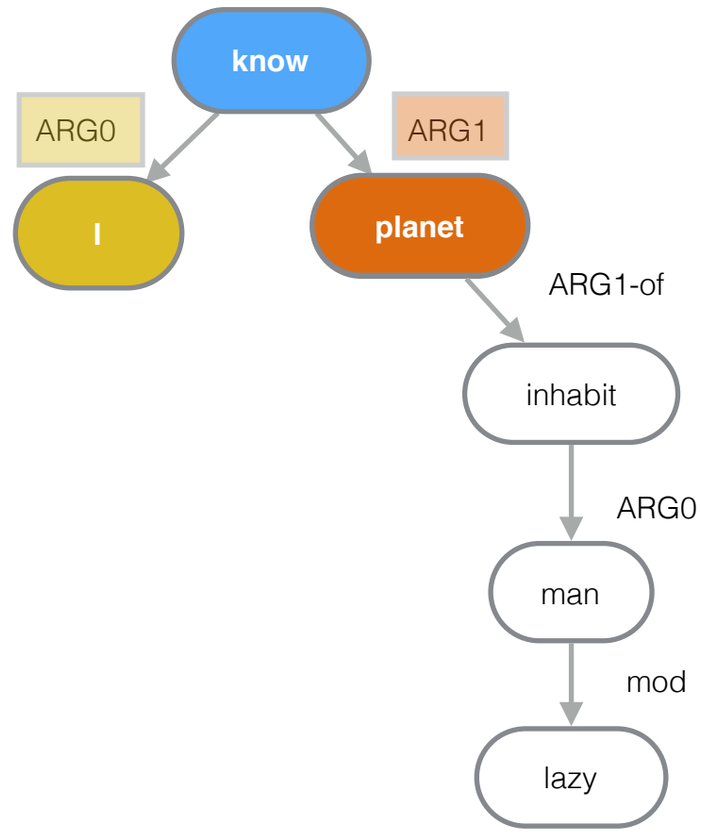


Abstract Meaning Representation

(Banarescu et al., 2013)

- Rooted Directed Acyclic Graph
- Nodes: concepts (nouns, verbs, named entities, etc)
- Edges: Semantic Role Labels

I have **known** a **planet** that was **inhabited** by a **lazy man**.



Abstract Meaning Representation

(Banarescu et al., 2013)

- Rooted Directed Acyclic Graph
- Nodes: concepts (nouns, verbs, named entities, etc)
- Edges: Semantic Role Labels

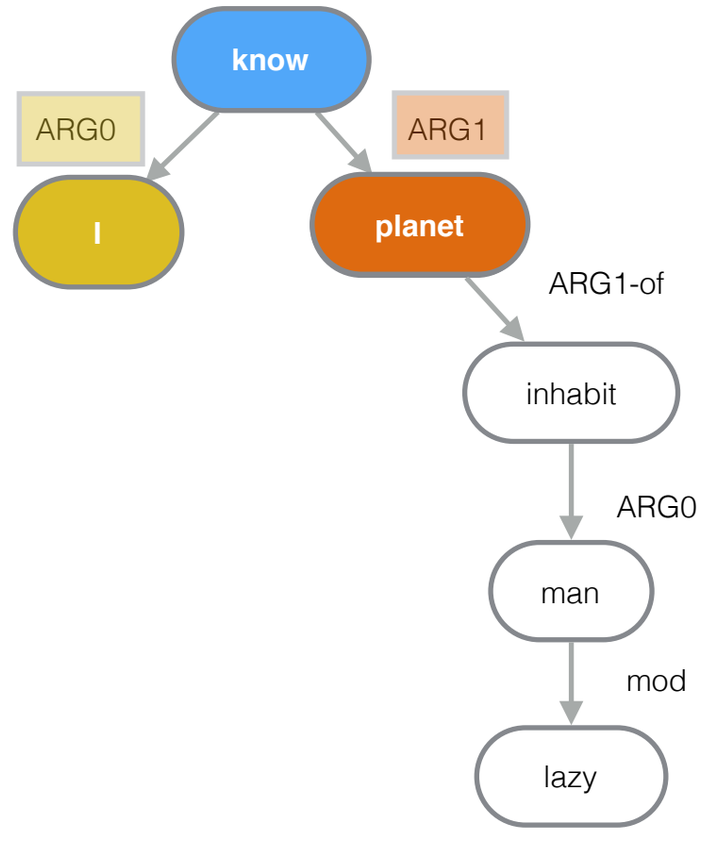
I knew a **planet** that was **inhabited** by a **lazy man**.

I have **known** a **planet** that was **inhabited** by a **lazy man**.

I know a **planet**. It is **inhabited** by a **lazy man**.

Input: AMR Graph

Generate from AMR



Abstract Meaning Representation

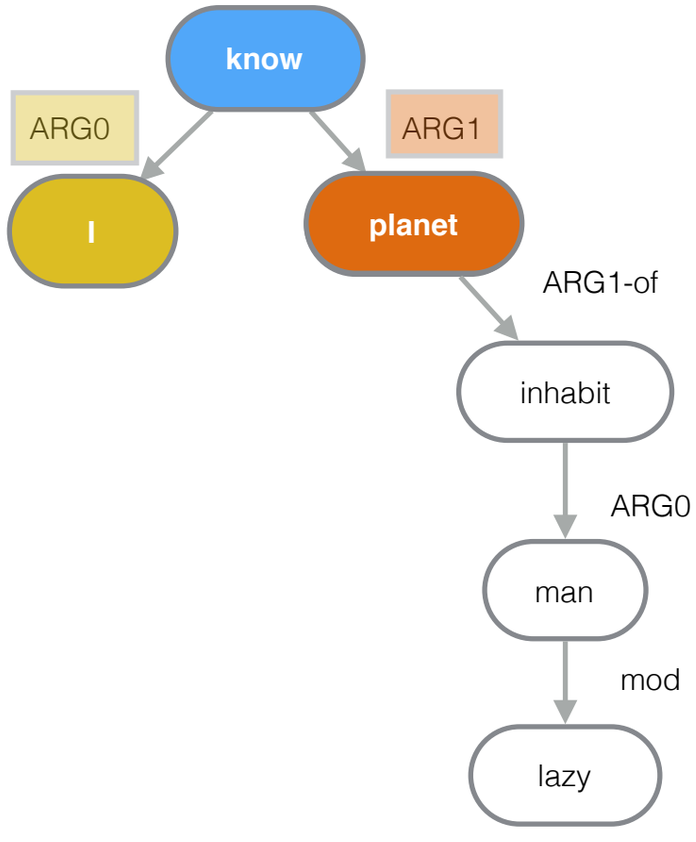
(Banarescu et al., 2013)

- Rooted Directed Acyclic Graph
- Nodes: concepts (nouns, verbs, named entities, etc)
- Edges: Semantic Role Labels

Input: Text

I have **known** a **planet** that was **inhabited** by a **lazy man**.

Parse to AMR

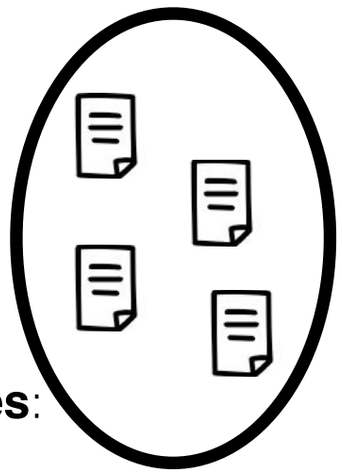


Applications

- **Text Summarization** (Liu et al., 2015)

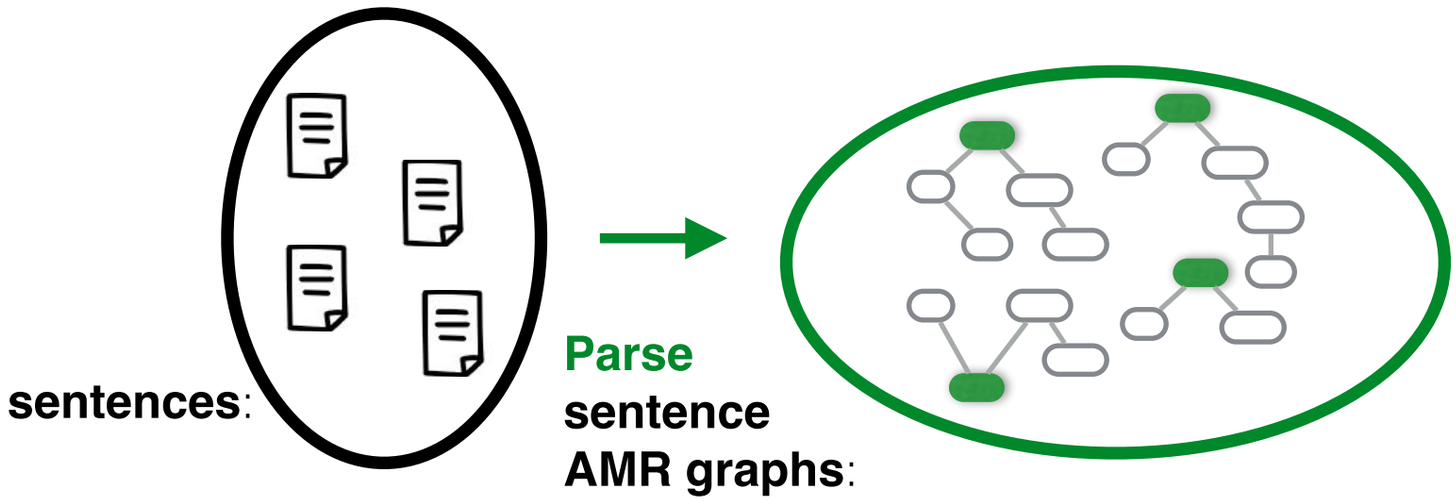
Applications

sentences:



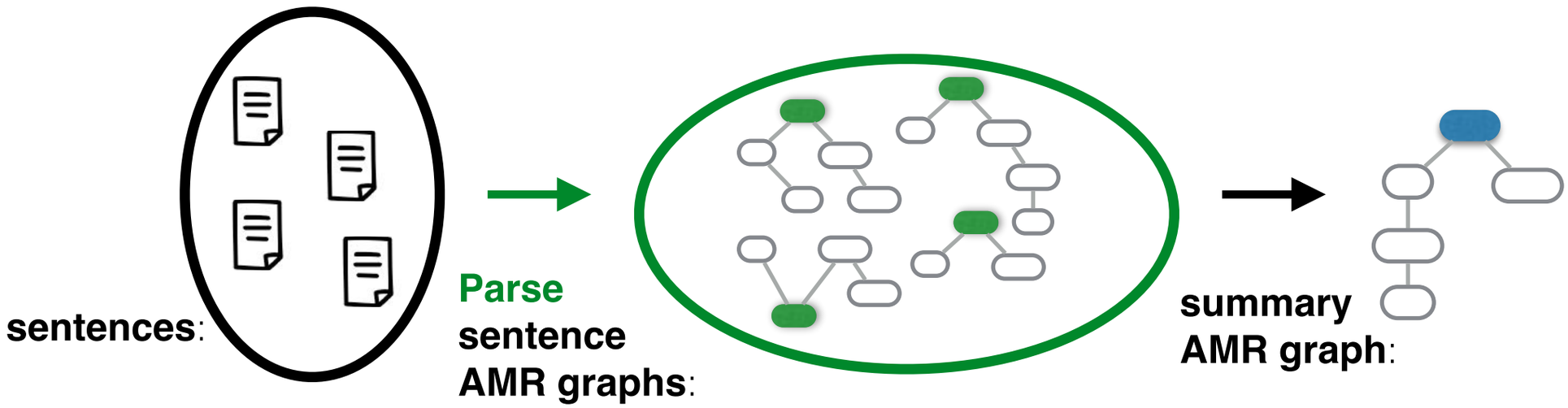
- **Text Summarization** (Liu et al., 2015)

Applications



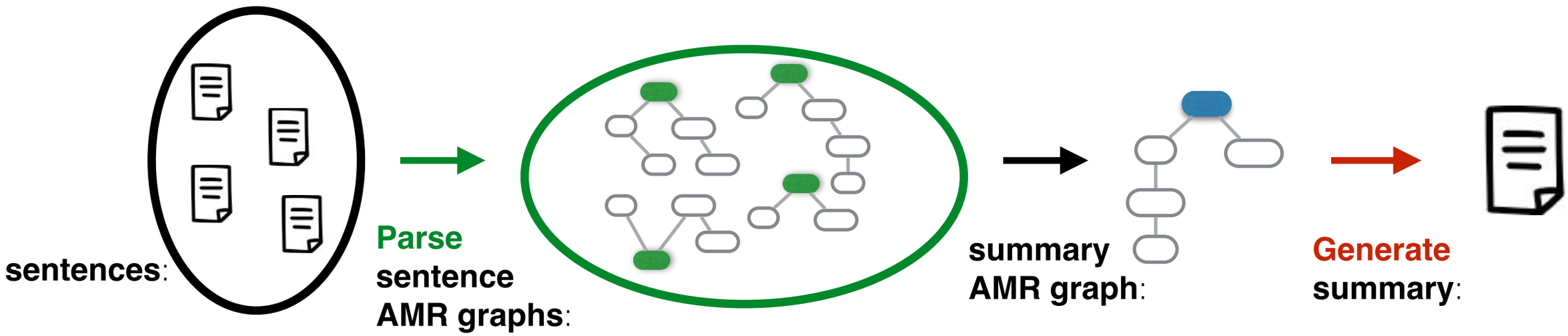
- **Text Summarization** (Liu et al., 2015)

Applications



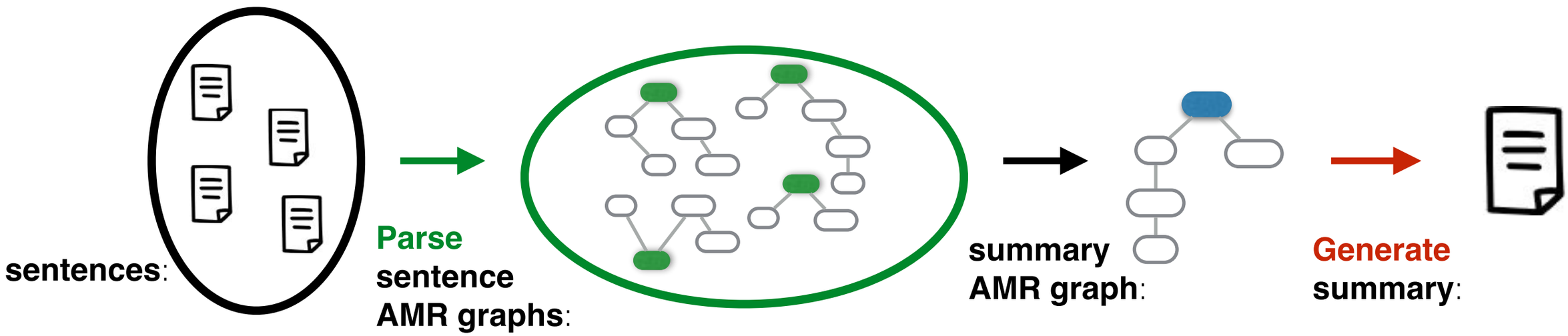
- **Text Summarization** (Liu et al., 2015)

Applications

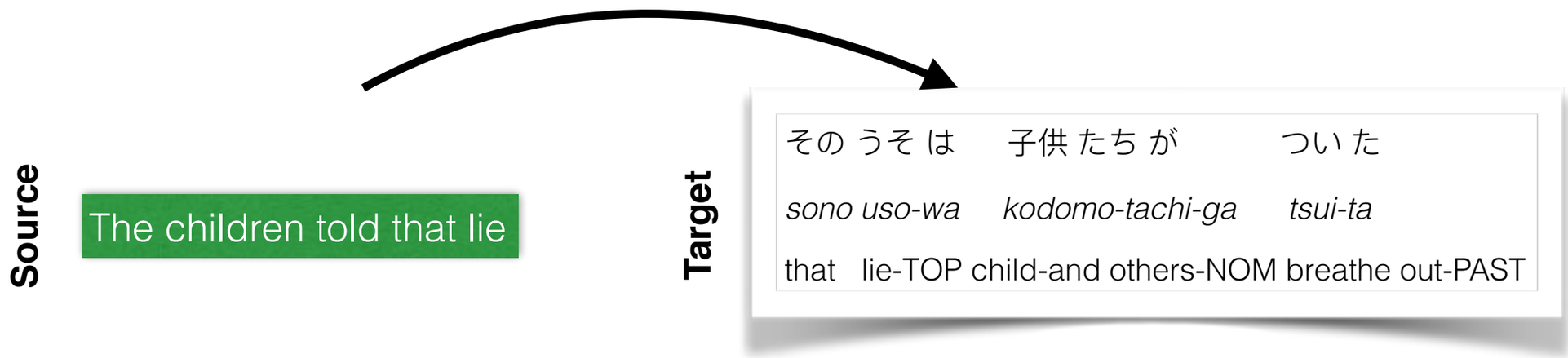


- **Text Summarization** (Liu et al., 2015)

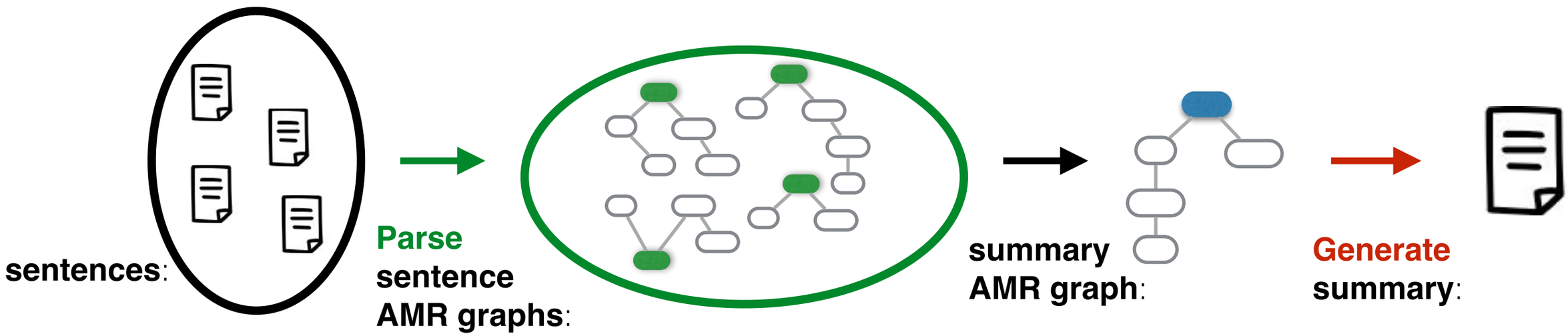
Applications



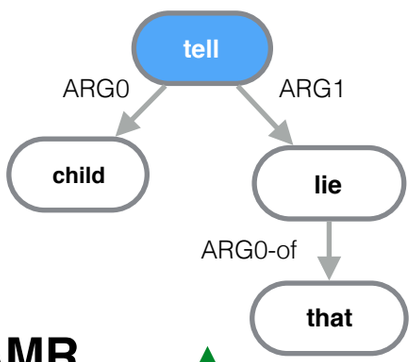
- **Text Summarization** (Liu et al., 2015)



Applications



- **Text Summarization** (Liu et al., 2015)



Parse AMR graph:

Source

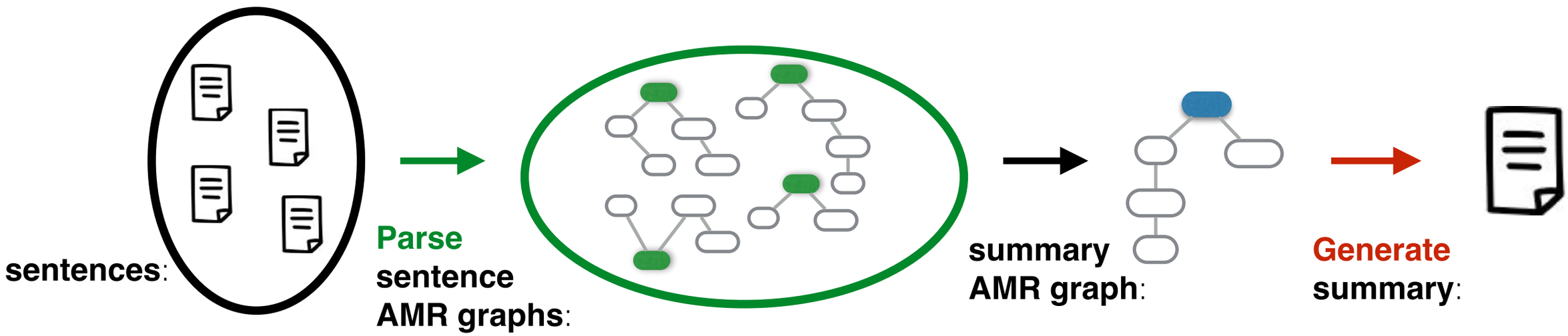
The children told that lie

Target

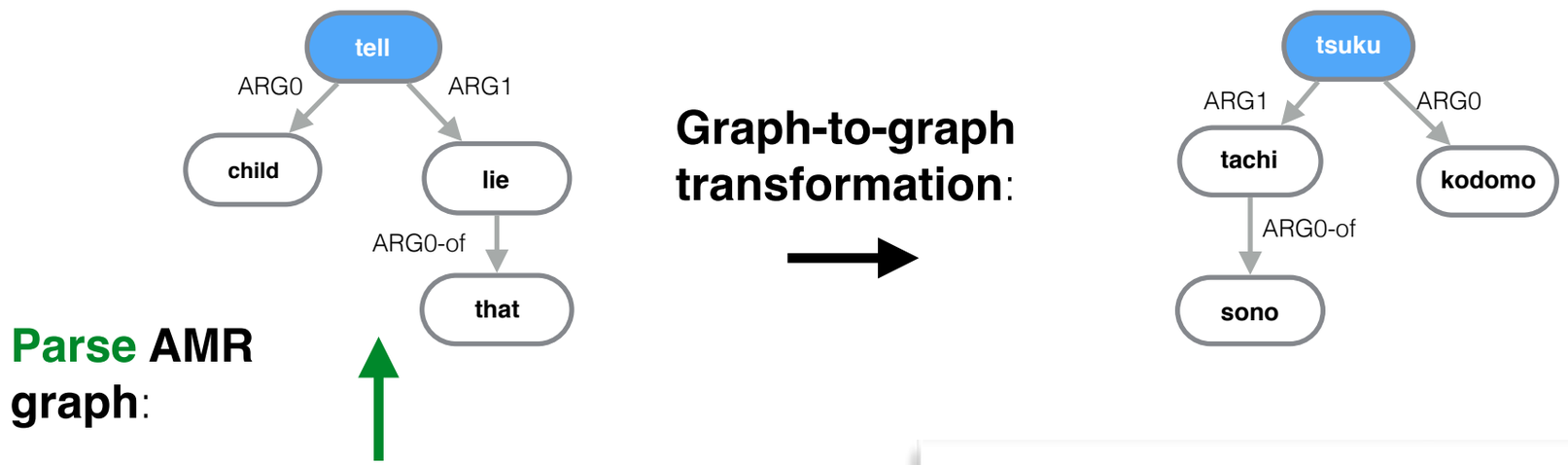
その	うそ	は	子供 たち が	ついた
sono	*uso-wa*		*kodomo-tachi-ga*	*tsui-ta*
that	lie-TOP		child-and others-NOM	breathe out-PAST

- **Machine Translation** (Jones et al., 2012)

Applications

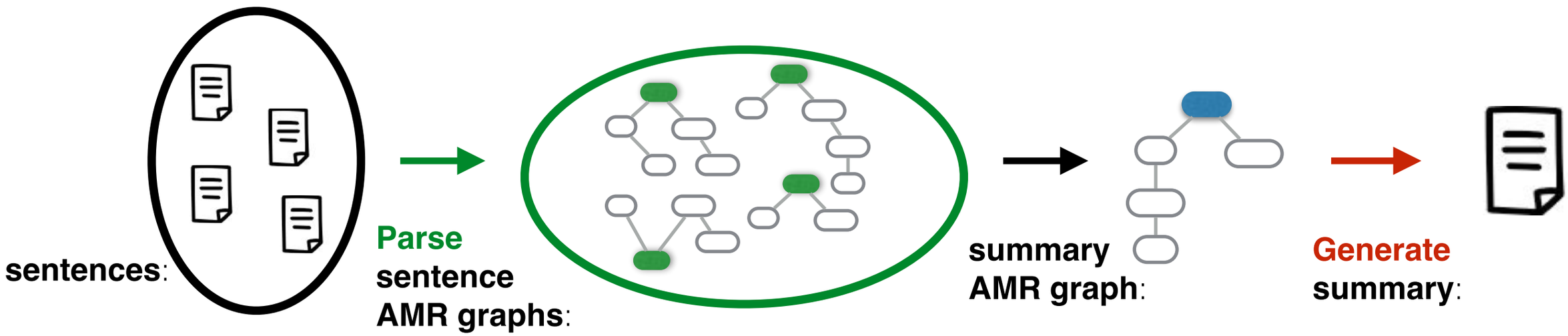


- **Text Summarization** (Liu et al., 2015)

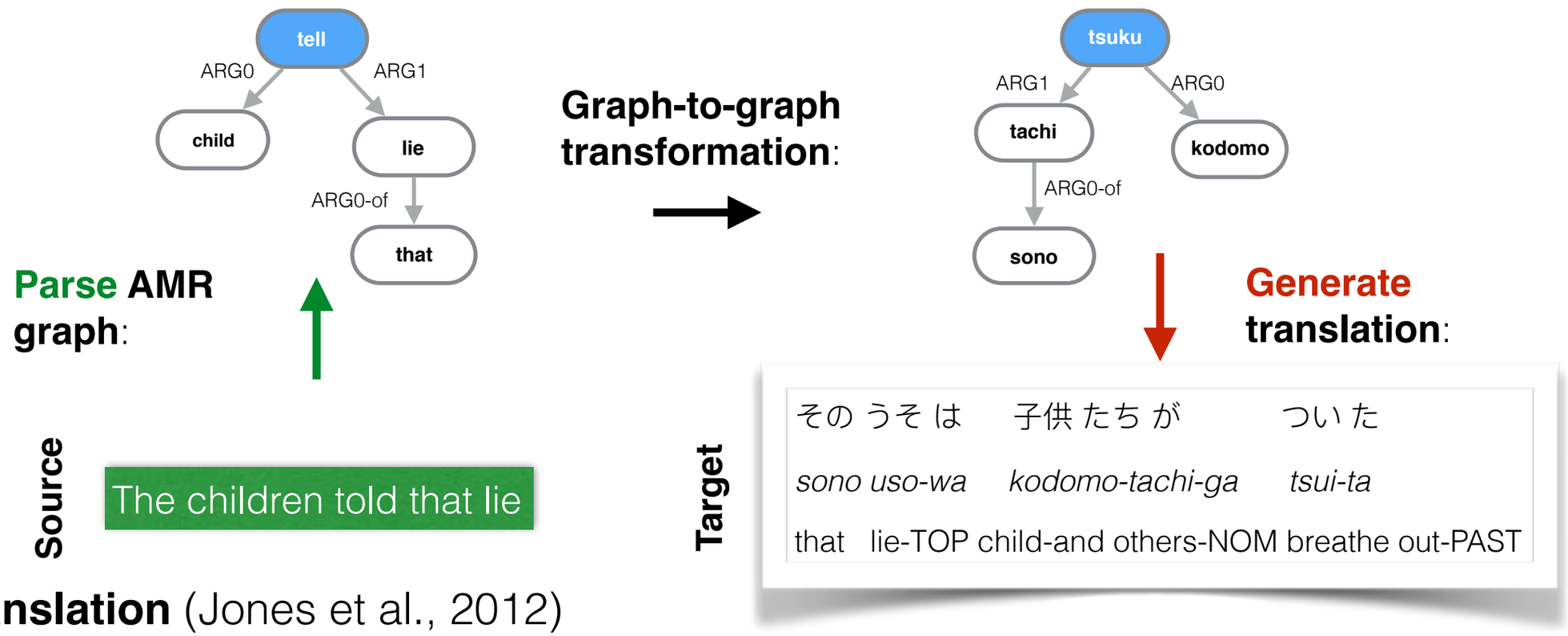


- **Machine Translation** (Jones et al., 2012)

Applications



- **Text Summarization** (Liu et al., 2015)



- **Machine Translation** (Jones et al., 2012)

Existing Approaches

Generate from AMR

- **MT-based**
  - Flanigan et al. 2016, Pourdamaghani and Knight 2016, Song et al. 2016
- **Grammar-based**
  - Lampouras and Vlachos 2017, Mille et al. 2017

Existing Approaches

Generate from AMR

- **MT-based**
  - Flanigan et al. 2016, Pourdamaghani and Knight 2016, Song et al. 2016
- **Grammar-based**
  - Lampouras and Vlachos 2017, Mille et al. 2017

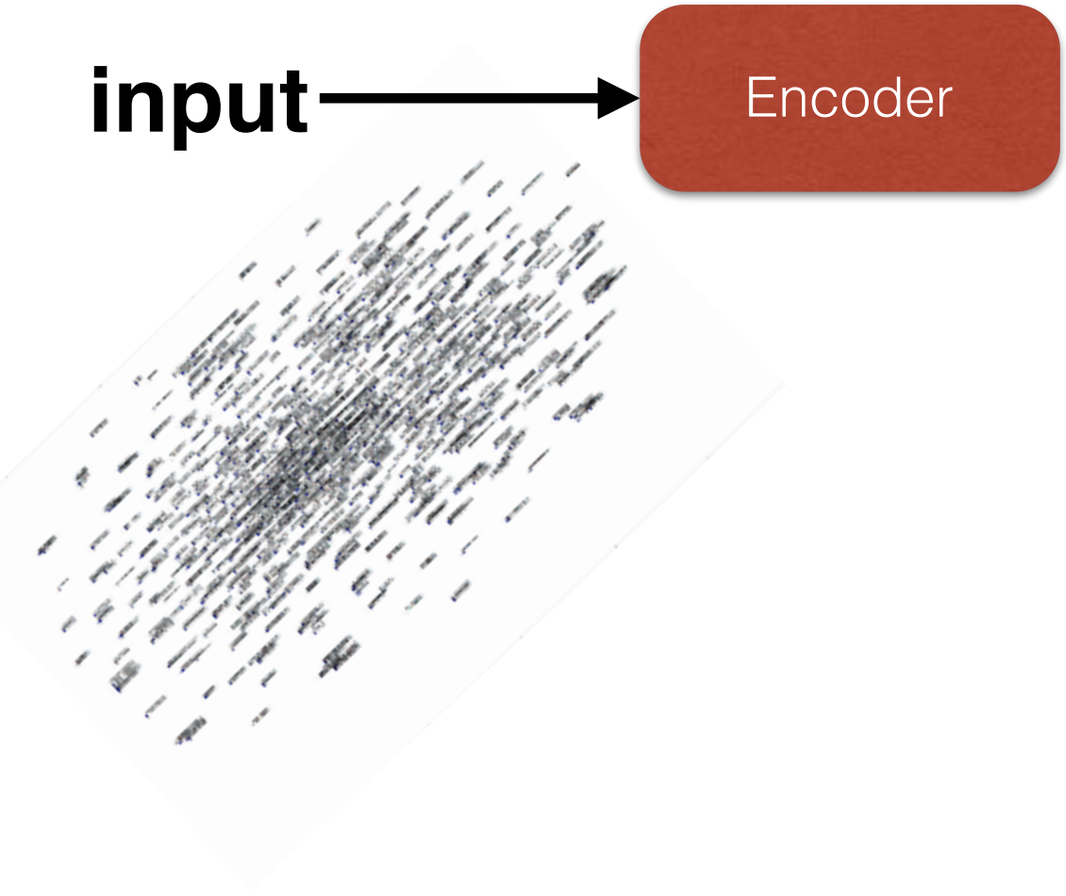
Parse to AMR

- **Alignment-based**
  - Flanigan et al. 2014, 2017 (JAMR)
- **Grammar-based**
  - Wang et al. 2016 (CAMR), Pust et al. 2015, Artzi et al. 2015, Damonte et al. 2017, Goodman et al. 2016, Puzikov et al. 2016, Brandt et al. 2017, Nguyen et al. 2017
- **Neural-based**
- Barzdins and Gosko 2016, Peng et al. 2017, Noord and Bos 2017, Buys and Blunsom 2017

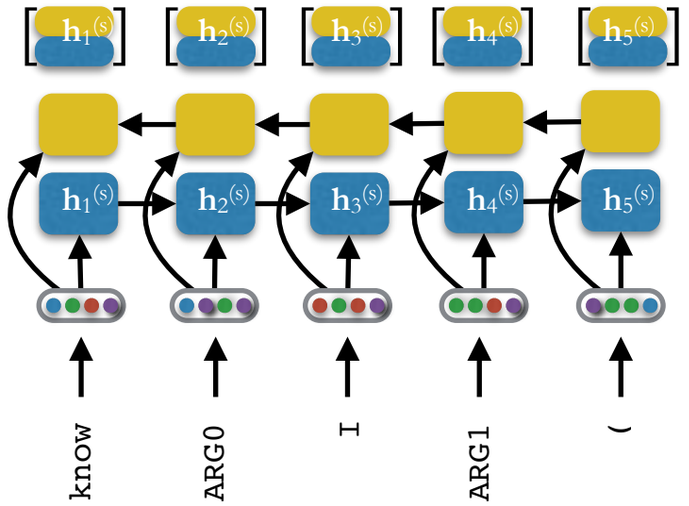
Overview

- Sequence-to-sequence architecture
  - End-to-end model w/o intermediate representations
  - Linearisation of AMR graph to string
  - Pre-process
- Paired Training
  - Scalable data augmentation algorithm

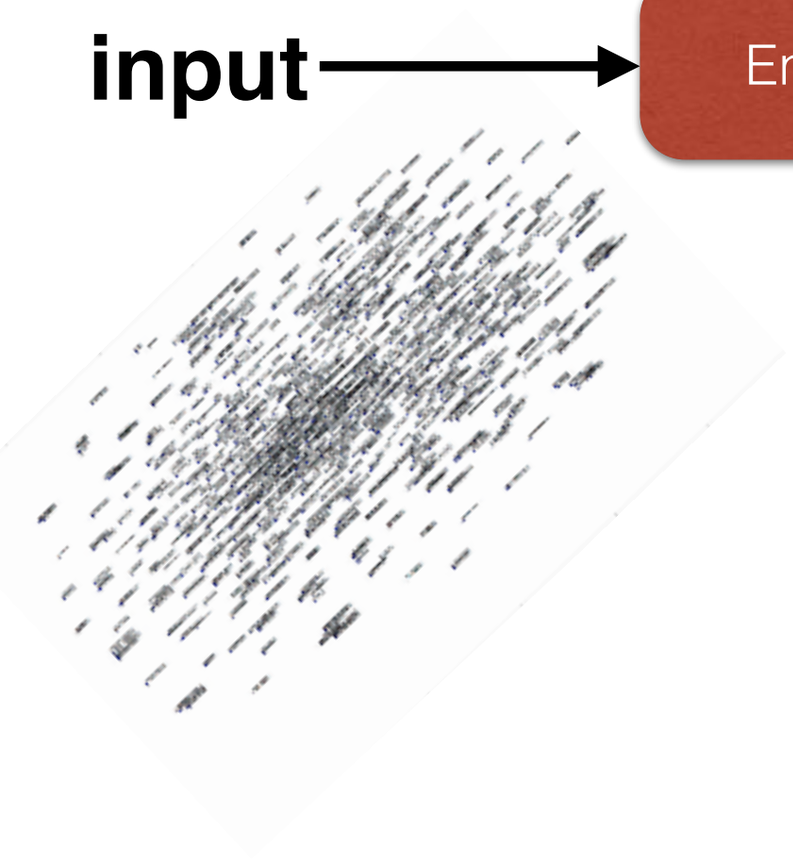
Sequence-to-sequence model



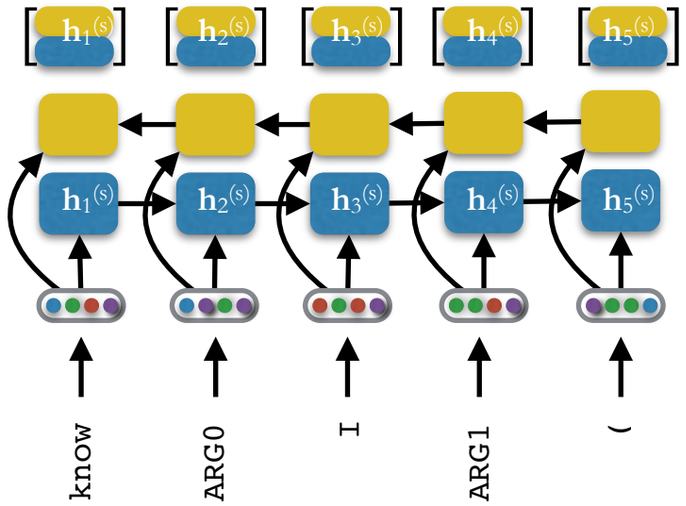
Sequence-to-sequence model

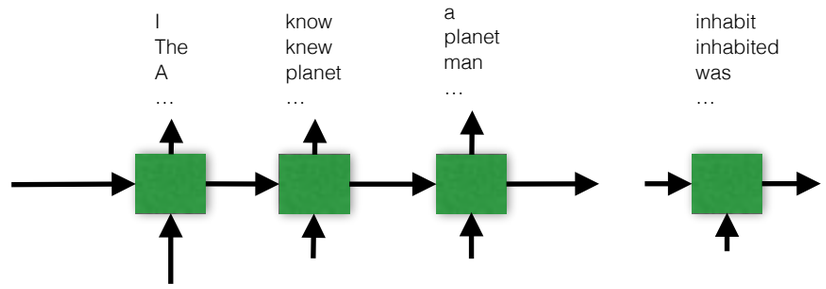


input → Encoder



Sequence-to-sequence model

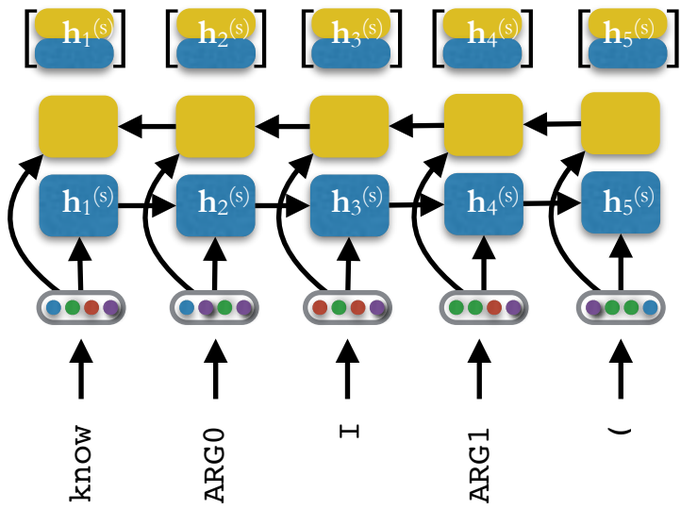


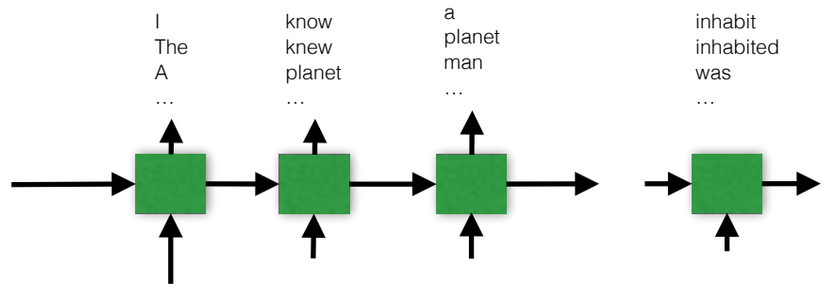


$$\hat{w} = \underset{w}{\operatorname{argmax}} \prod_i p\left(w_i | w_{<i}, \mathbf{h}^{(s)}\right)$$

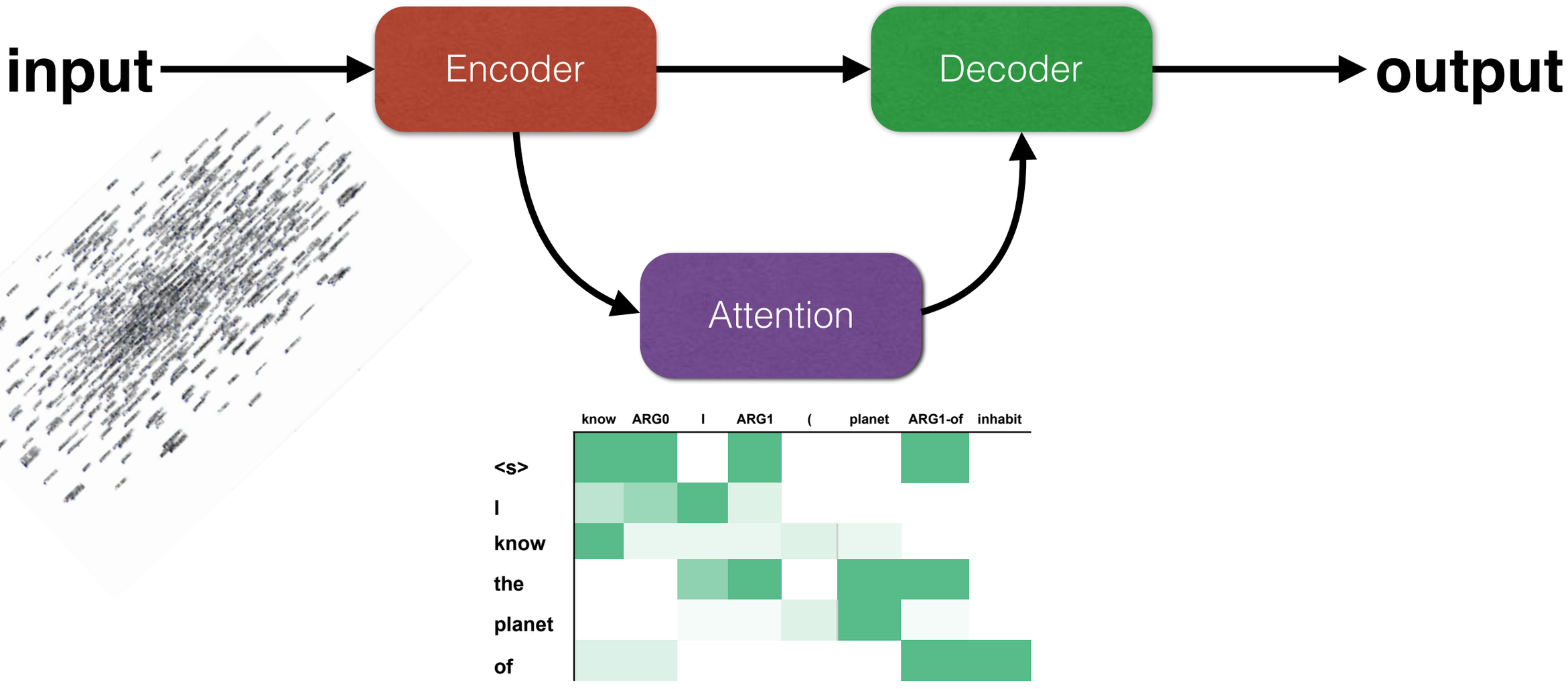
input → Encoder → Decoder → **output**

Sequence-to-sequence model



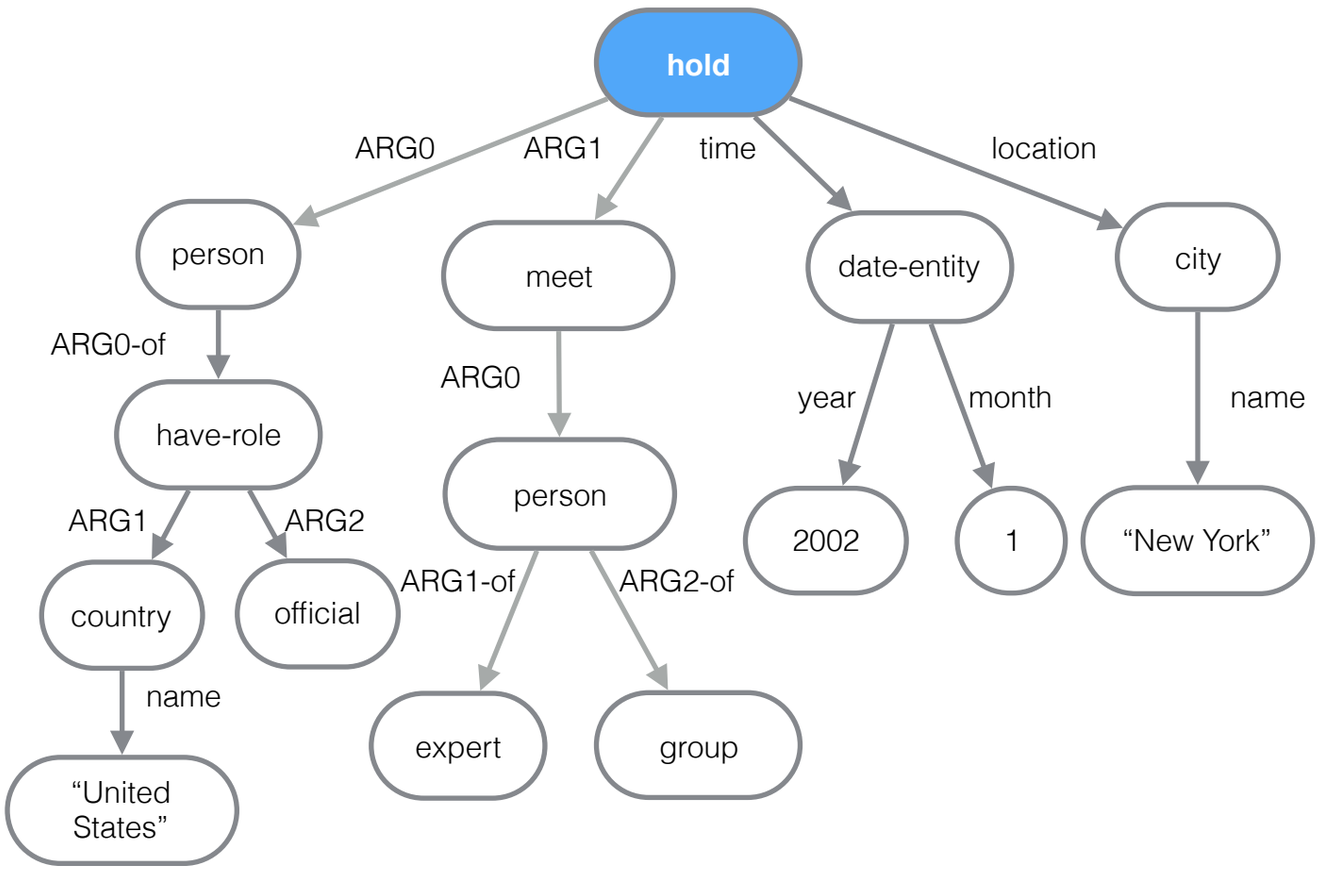


$$\hat{w} = \underset{w}{\operatorname{argmax}} \prod_i p\left(w_i | w_{<i}, \mathbf{h}^{(s)}\right)$$



Linearization

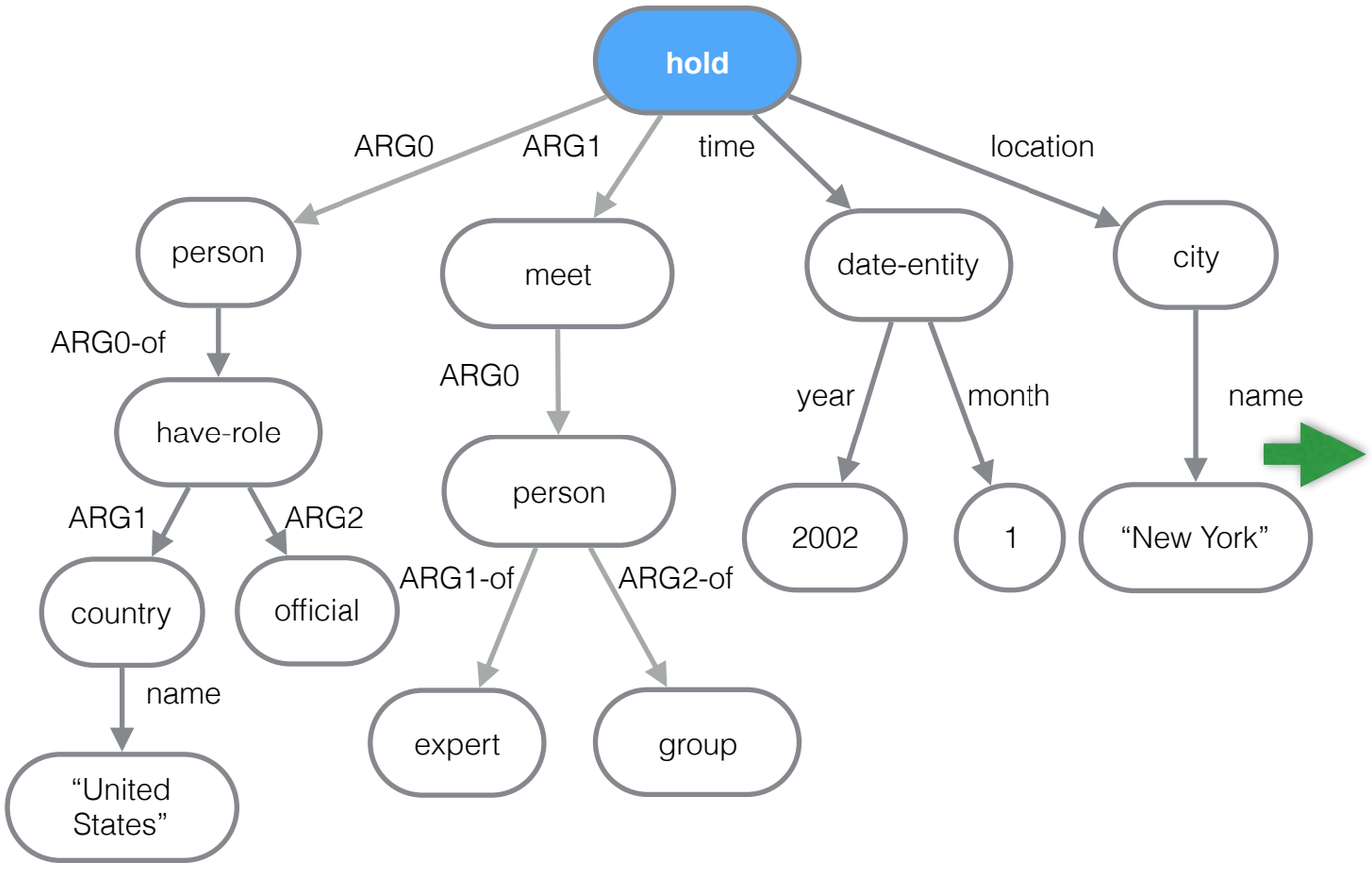
Graph —> Depth First Search (Human-authored annotation)



US officials held an expert group meeting in January 2002 in New York .

Linearization

Graph —> Depth First Search (Human-authored annotation)

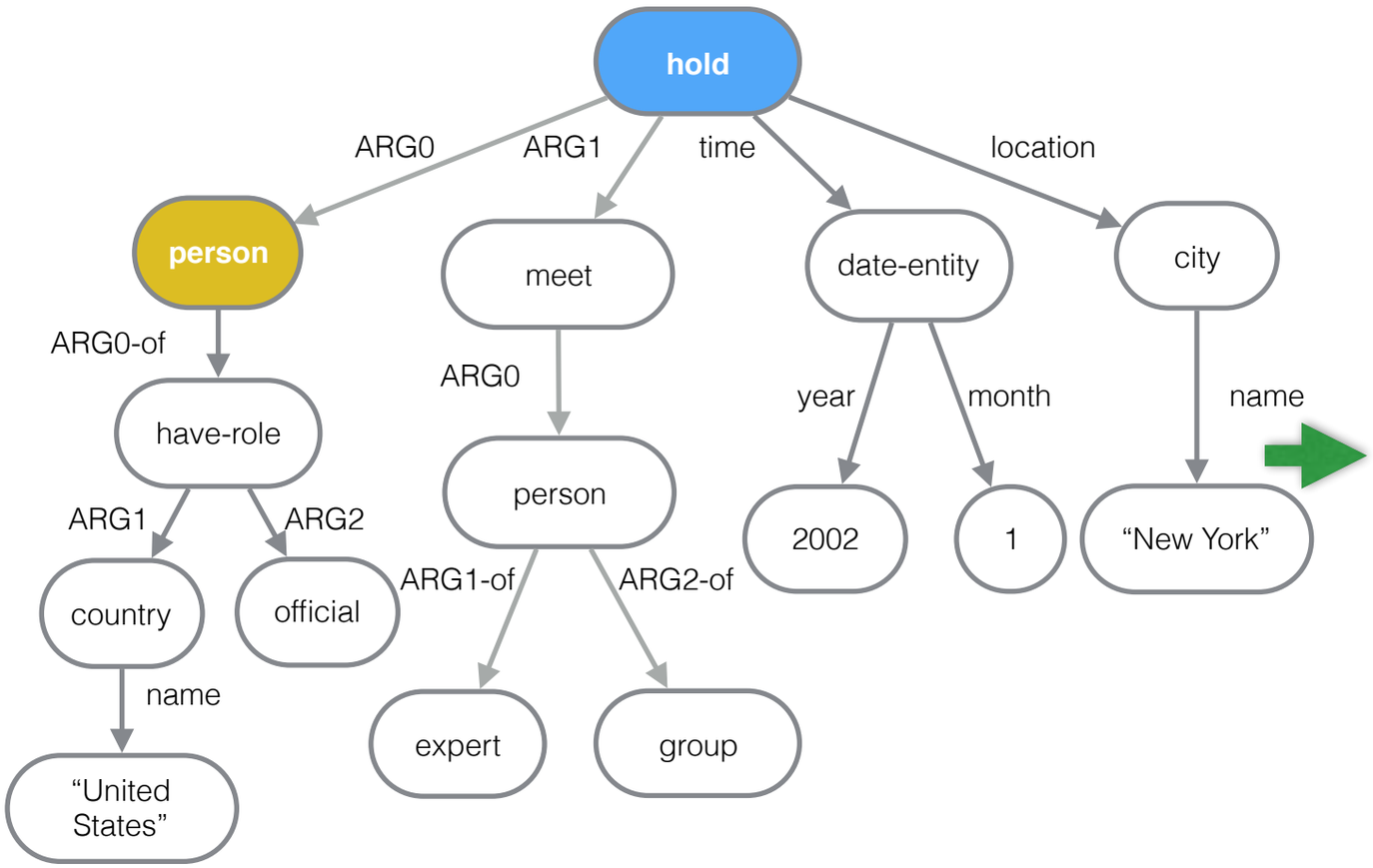



```
hold
   :ARG0 (person
            :ARG0-of (have-role
               :ARG1 United_States
               :ARG2 official)
         )
   :ARG1 (meet
            :ARG0 (person
               :ARG1-of expert
               :ARG2-of group)
         )
   :time (date-entity 2002 1)
   :location New_York
```

US officials held an expert group meeting in January 2002 in New York .

Linearization

Graph —> Depth First Search (Human-authored annotation)

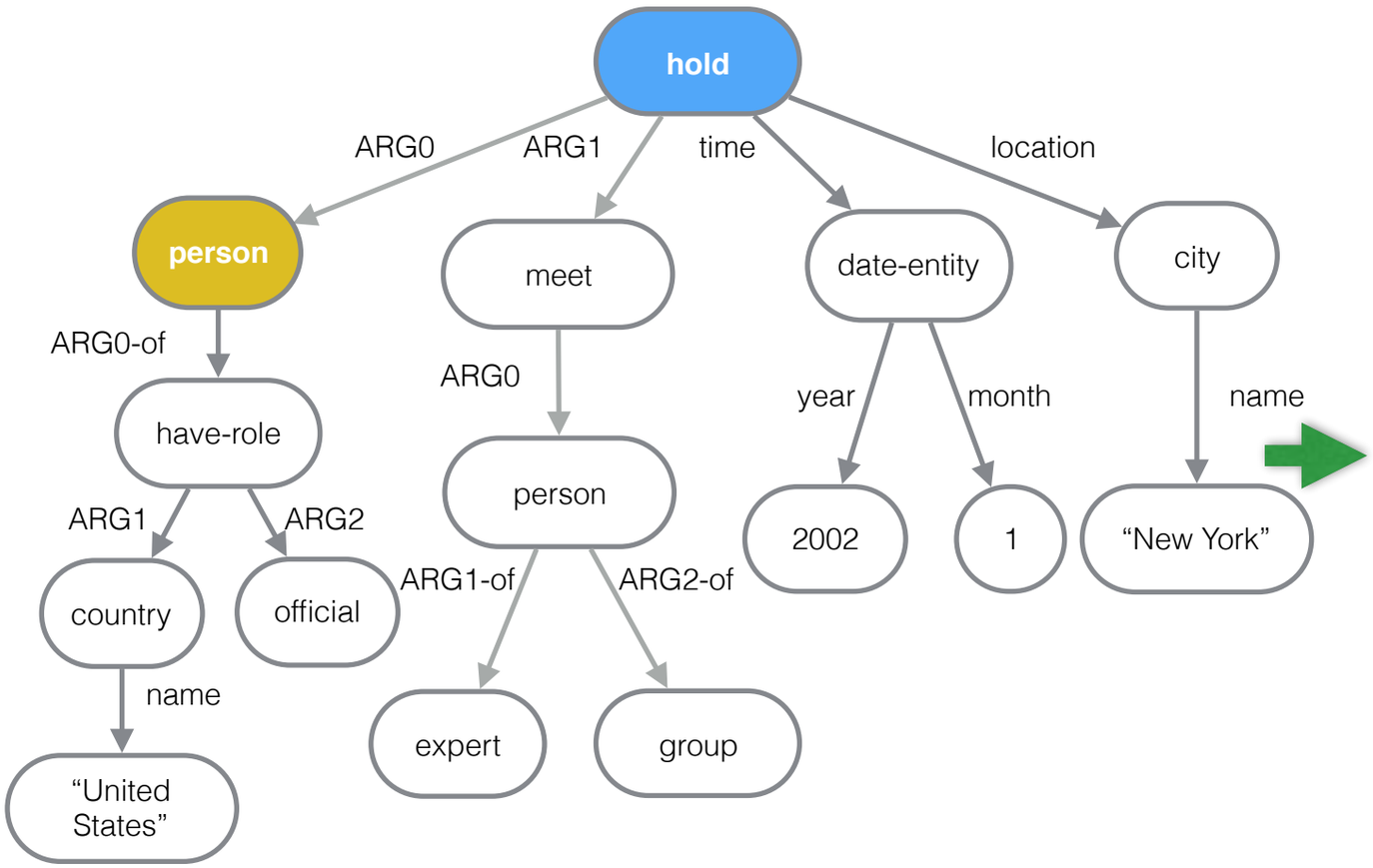



```
hold
   :ARG0 (person
            :ARG0-of (have-role
               :ARG1 United_States
               :ARG2 official)
         )
   :ARG1 (meet
            :ARG0 (person
               :ARG1-of expert
               :ARG2-of group)
         )
   :time (date-entity 2002 1)
   :location New_York
```

US officials held an expert group meeting in January 2002 in New York .

Linearization

Graph —> Depth First Search (Human-authored annotation)

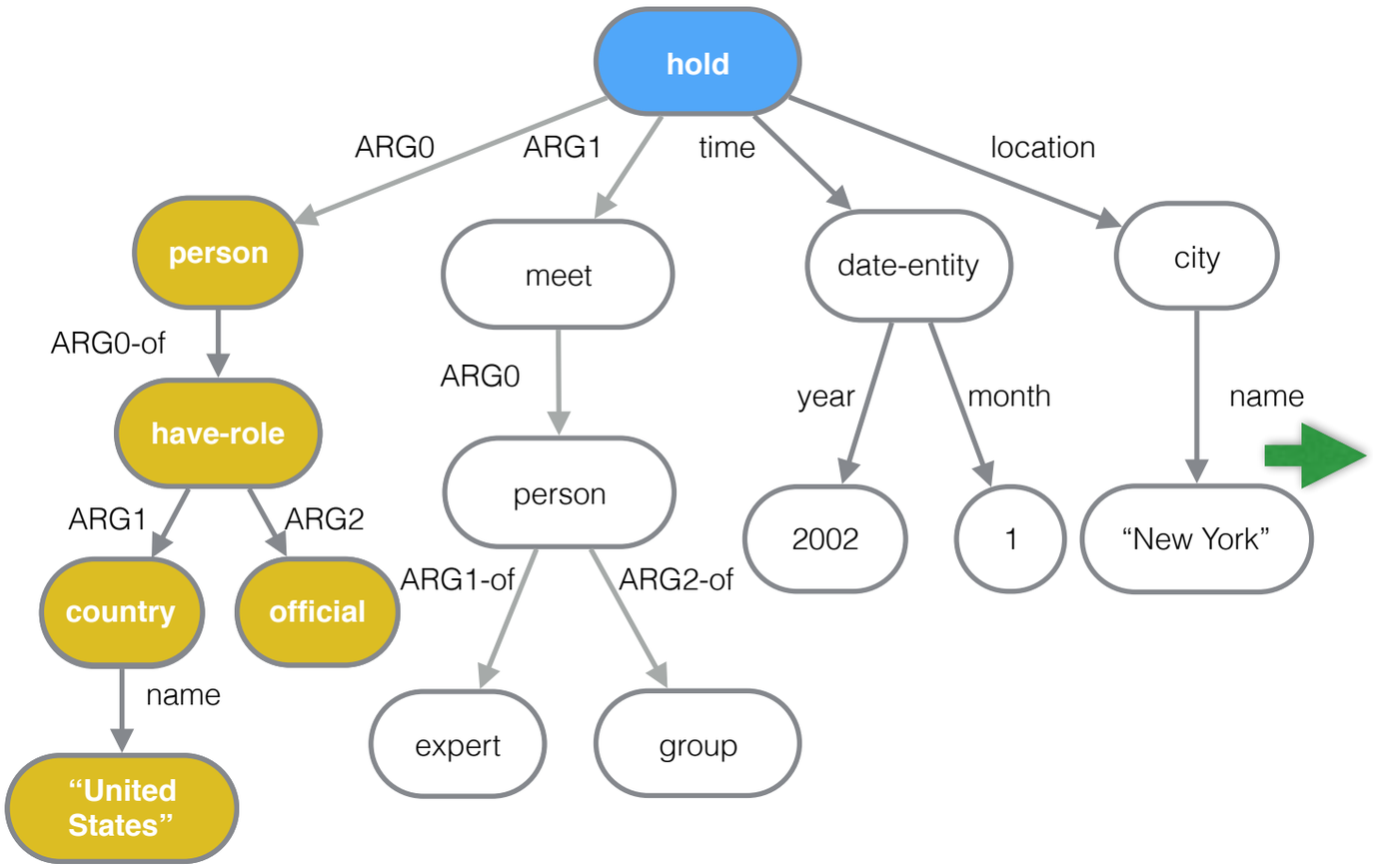



```
hold
    :ARG0 (person
                :ARG0-of (have-role
                    :ARG1 United_States
                    :ARG2 official)
          )
    :ARG1 (meet
                :ARG0 (person
                    :ARG1-of expert
                    :ARG2-of group)
          )
    :time (date-entity 2002 1)
    :location New_York
```

US officials held an expert group meeting in January 2002 in New York .

Linearization

Graph —> Depth First Search (Human-authored annotation)

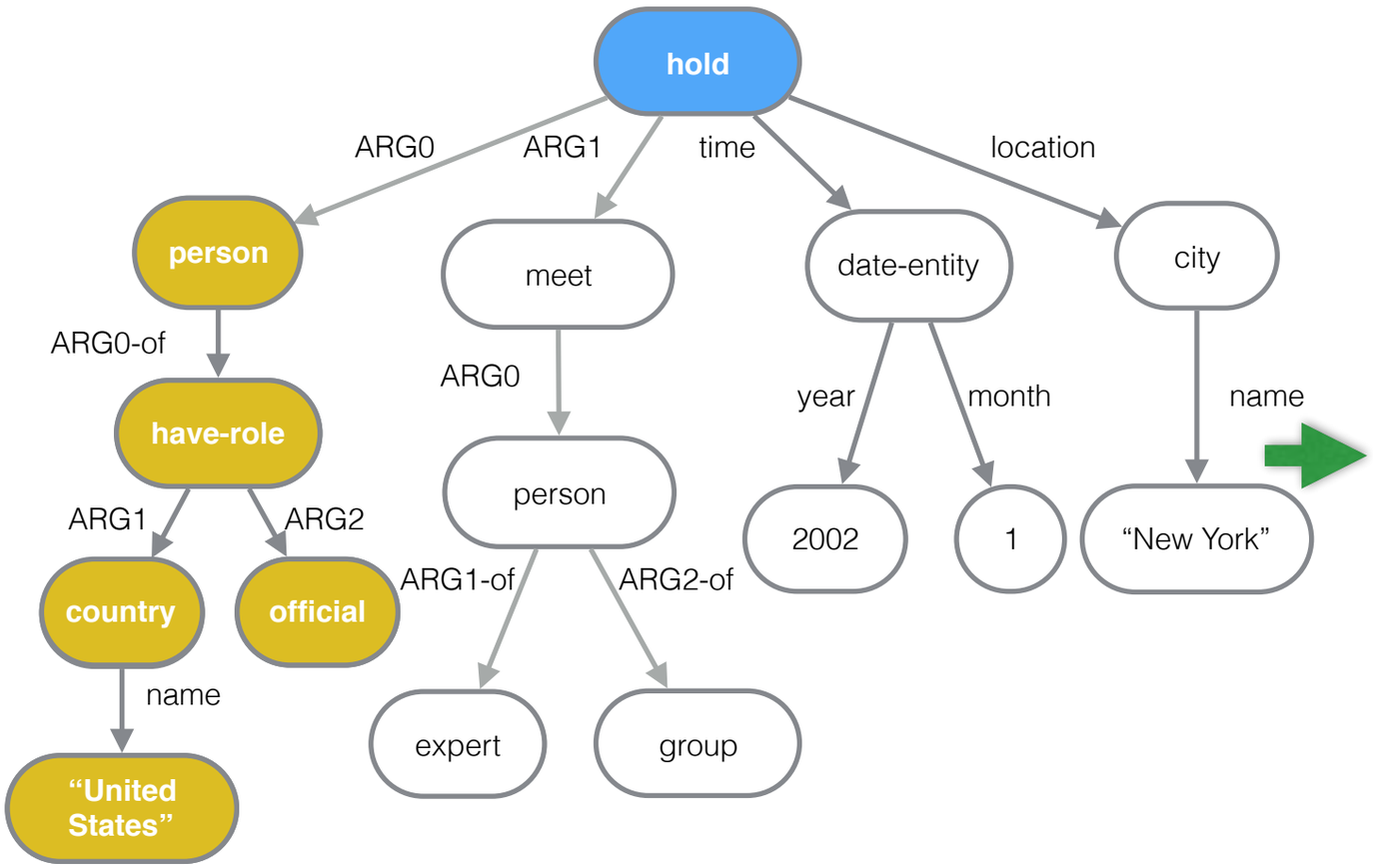



```
hold
    :ARG0 (person
                :ARG0-of (have-role
                    :ARG1 United_States
                    :ARG2 official)
          )
    :ARG1 (meet
                :ARG0 (person
                    :ARG1-of expert
                    :ARG2-of group)
          )
    :time (date-entity 2002 1)
    :location New_York
```

US officials held an expert group meeting in January 2002 in New York .

Linearization

Graph —> Depth First Search (Human-authored annotation)

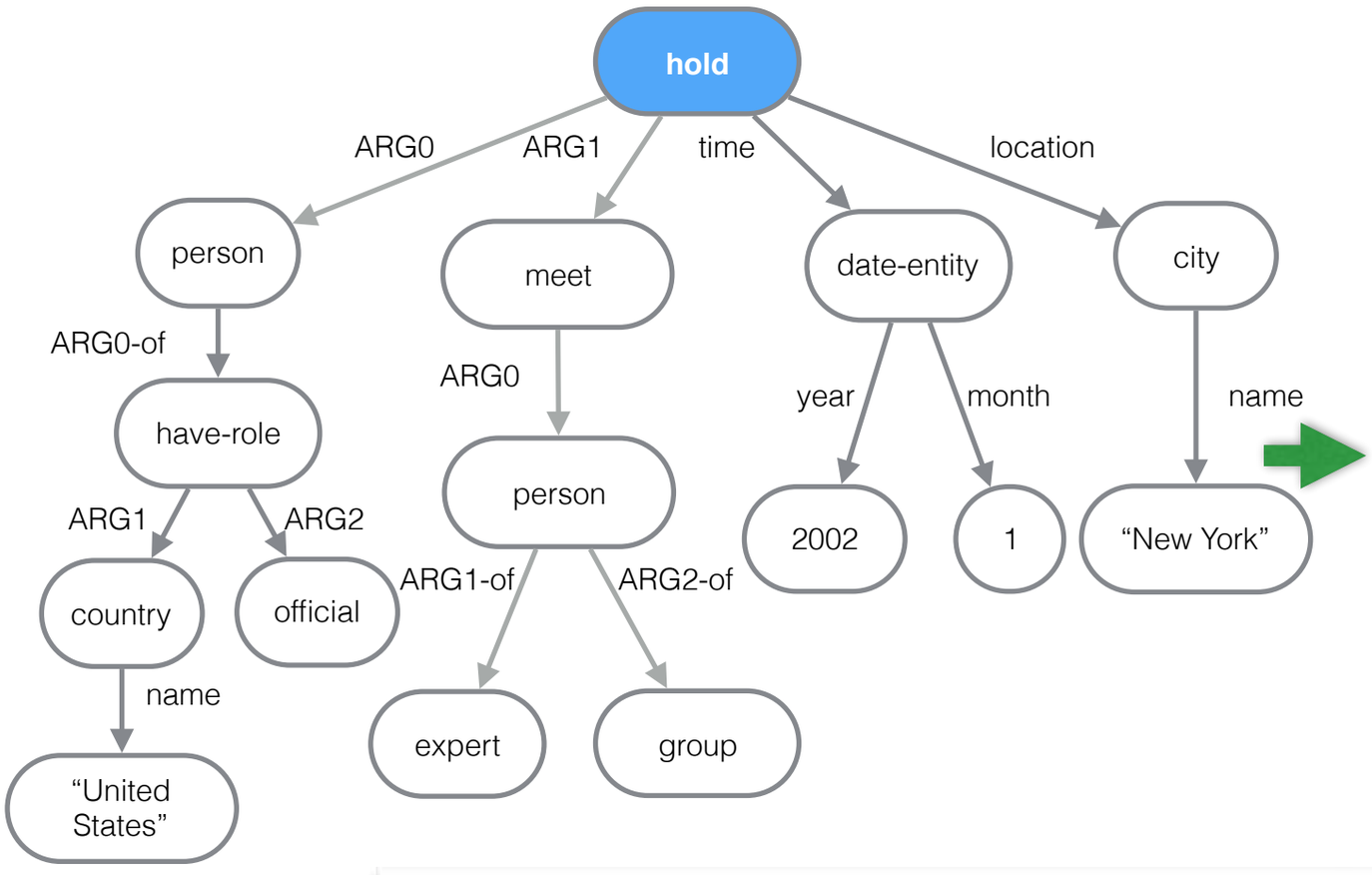



```
hold
    :ARG0 (person
                :ARG0-of (have-role
                    :ARG1 United_States
                    :ARG2 official)
           )
   :ARG1 (meet
               :ARG0 (person
                   :ARG1-of expert
                   :ARG2-of group)
          )
   :time (date-entity 2002 1)
   :location New_York
```

US officials held an expert group meeting in January 2002 in New York .

Pre-processing

Linearization —> Anonymization

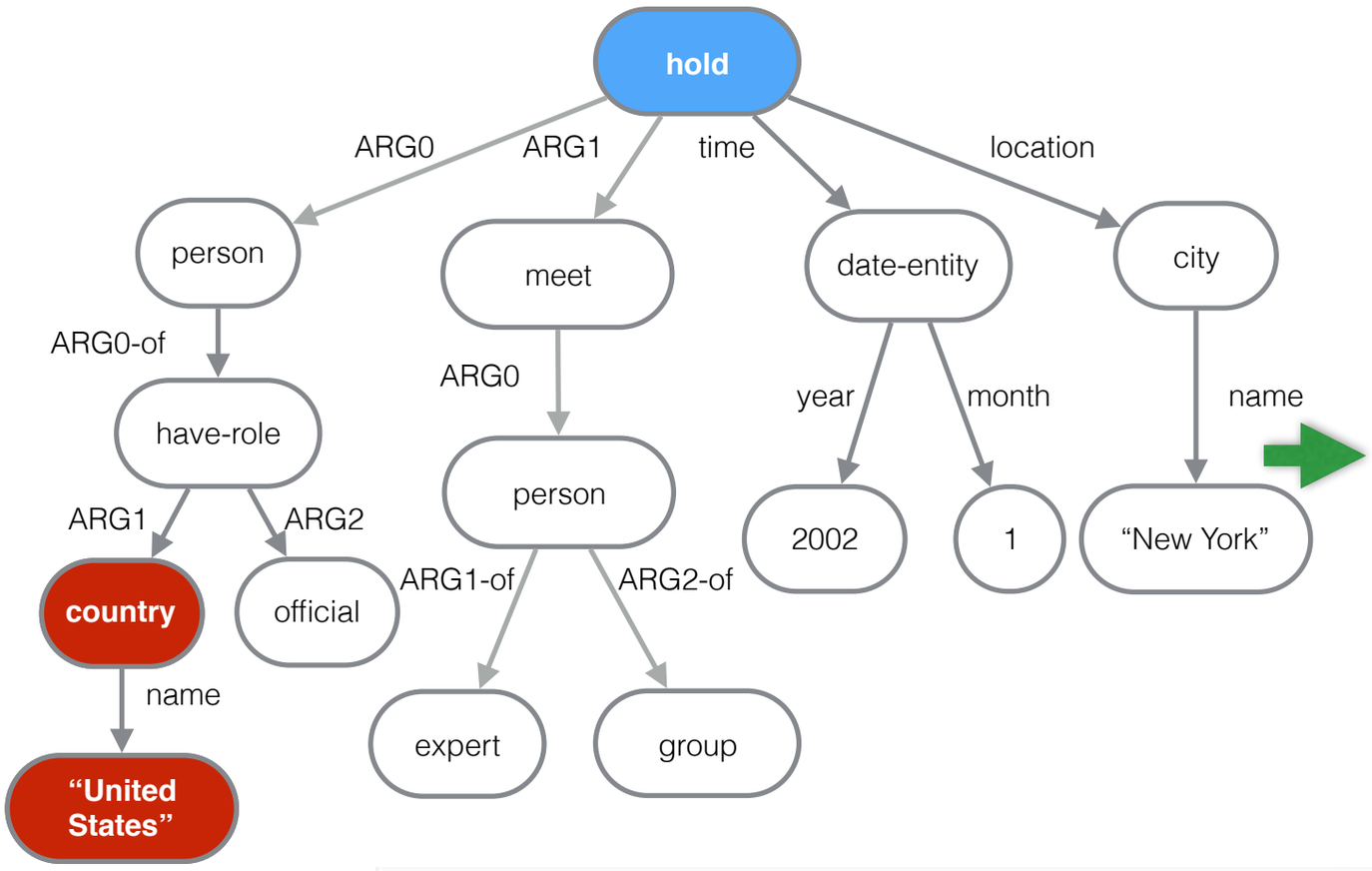



```
hold
   :ARG0 (person
            :ARG0-of (have-role
               :ARG1 loc_0
               :ARG2 official)
         )
   :ARG1 (meet
            :ARG0 (person
               :ARG1-of expert
               :ARG2-of group)
         )
   :time (date-entity year_0 month_0)
   :location loc_1
```

US officials held an expert group meeting in January 2002 in New York .

Pre-processing

Linearization —> Anonymization

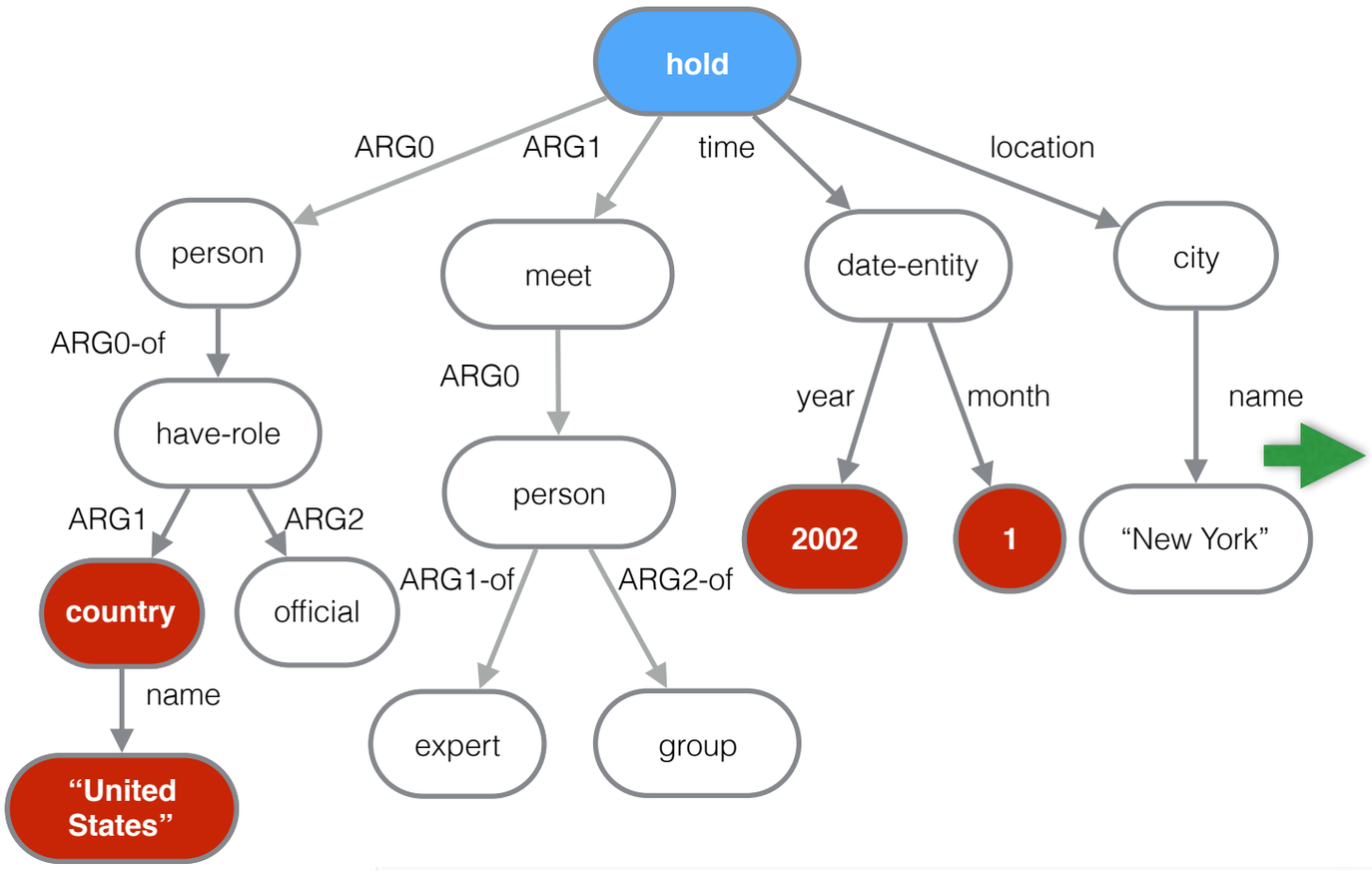



```
hold
   :ARG0 (person
            :ARG0-of (have-role
               :ARG1 loc_0
               :ARG2 official)
         )
   :ARG1 (meet
            :ARG0 (person
               :ARG1-of expert
               :ARG2-of group)
         )
   :time (date-entity year_0 month_0)
   :location loc_1
```

US officials held an expert group meeting in January 2002 in New York .

Pre-processing

Linearization —> Anonymization

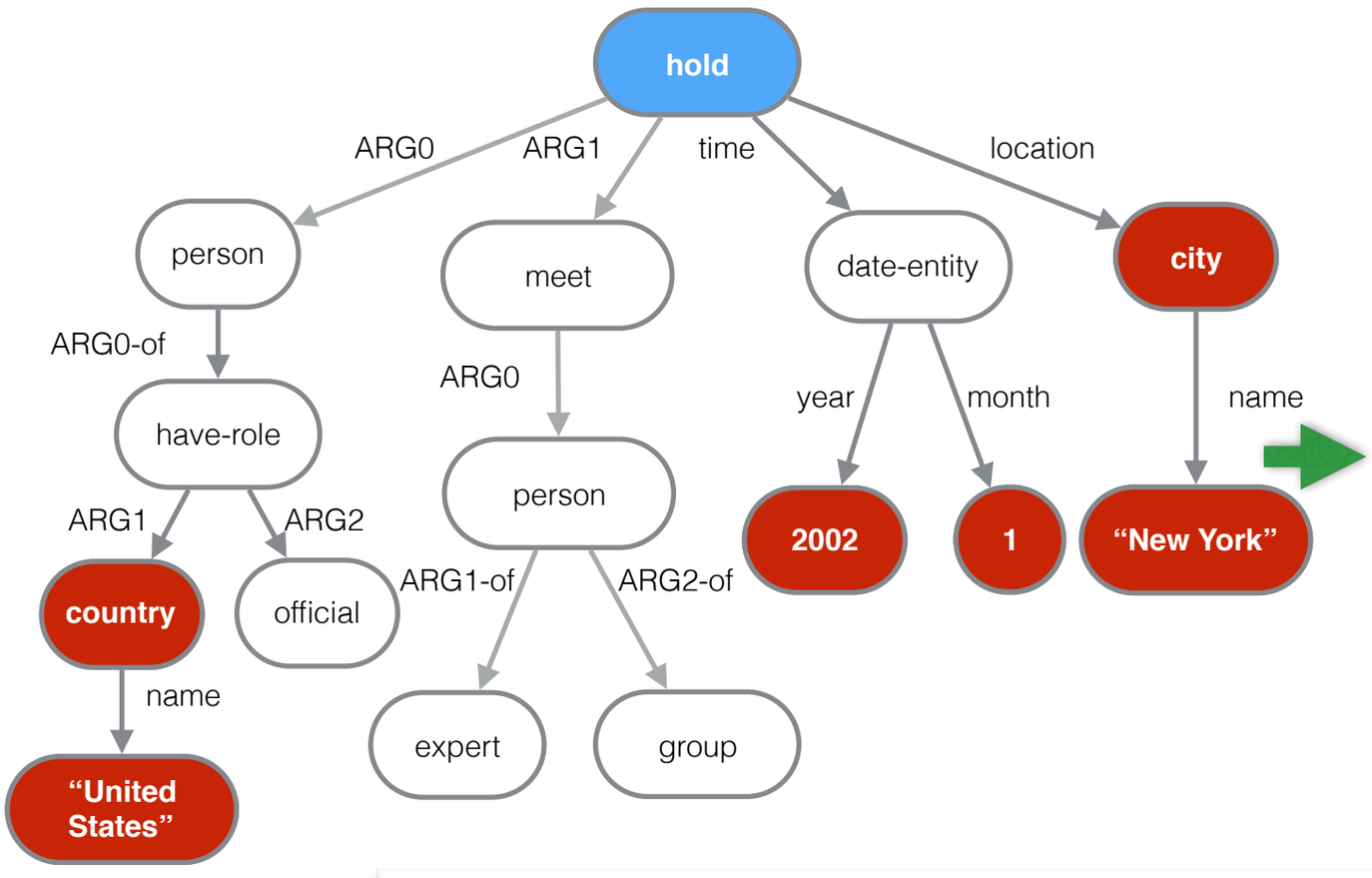



```
hold
   :ARG0 (person
            :ARG0-of (have-role
               :ARG1 loc_0
               :ARG2 official)
         )
   :ARG1 (meet
            :ARG0 (person
               :ARG1-of expert
               :ARG2-of group)
         )
   :time (date-entity year_0 month_0)
   :location loc_1
```

US officials held an expert group meeting in January 2002 in New York .

Pre-processing

Linearization —> Anonymization

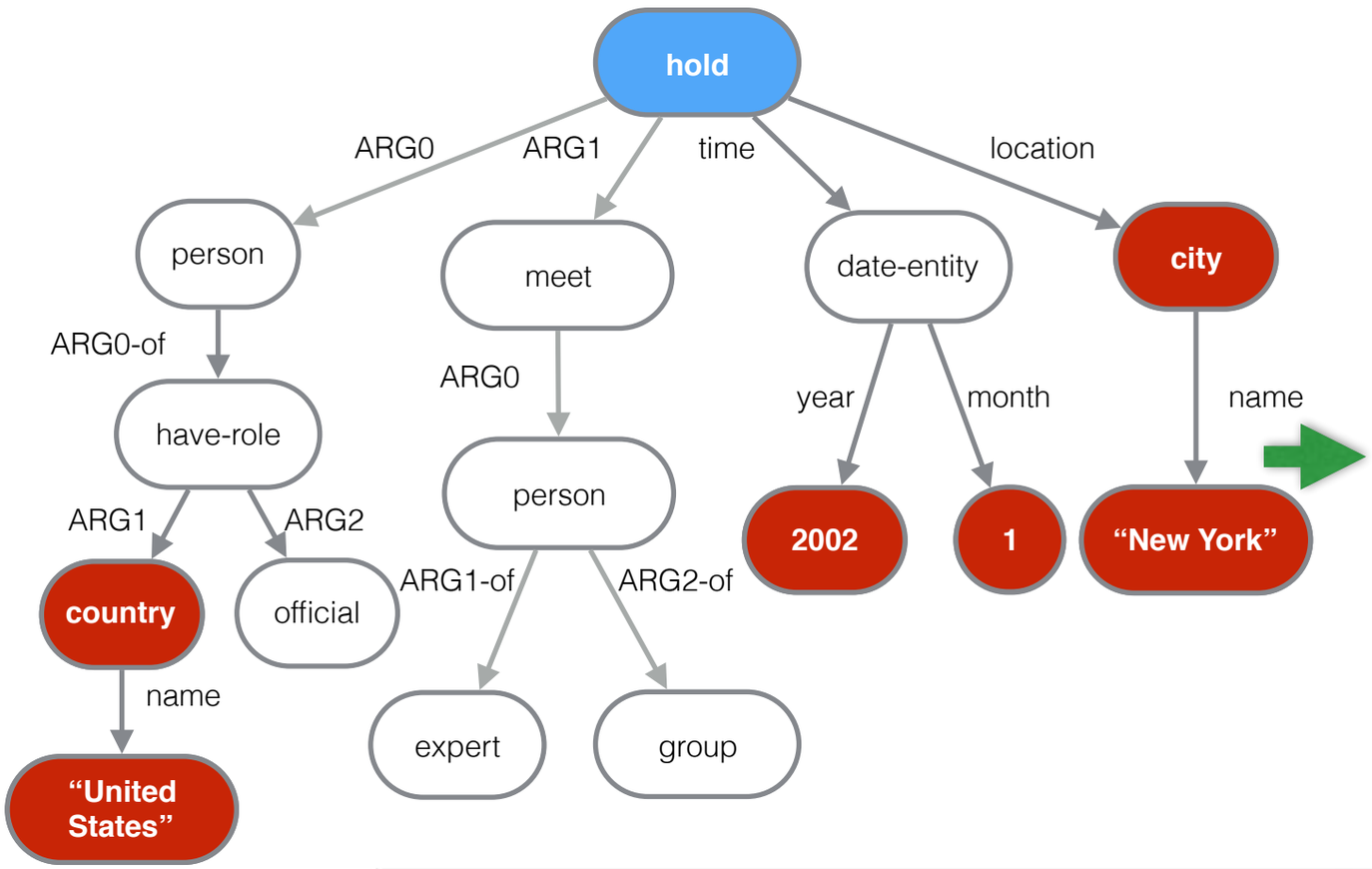



```
hold
    :ARG0 (person
                :ARG0-of (have-role
                    :ARG1 loc_0
                    :ARG2 official)
          )
    :ARG1 (meet
                :ARG0 (person
                    :ARG1-of expert
                    :ARG2-of group)
          )
    :time (date-entity year_0 month_0)
    :location loc_1
```

US officials held an expert group meeting in January 2002 in New York .

Pre-processing

Linearization —> Anonymization




```
hold
    :ARG0 (person
                :ARG0-of (have-role
                    :ARG1 loc_0
                    :ARG2 official)
          )
    :ARG1 (meet
                :ARG0 (person
                    :ARG1-of expert
                    :ARG2-of group)
          )
    :time (date-entity year_0 month_0)
    :location loc_1
```

US officials held an expert group meeting in January 2002 in New York .

loc_0 officials held an expert group meeting in **month_0 year_0** in **loc_1** .

Experimental Setup

AMR LDC2015E86 (SemEval-2016 Task 8)

- Hand annotated MR graphs: newswire, forums
- ~16k **training** / 1k **development** / 1k **test** pairs

Train

- Optimize cross-entropy loss

Evaluation

- BLEU n-gram precision (**Generation**)
  (Papineni et al., 2002)
- SMATCH score (**Parsing**)
  (Cai and Knight, 2013)

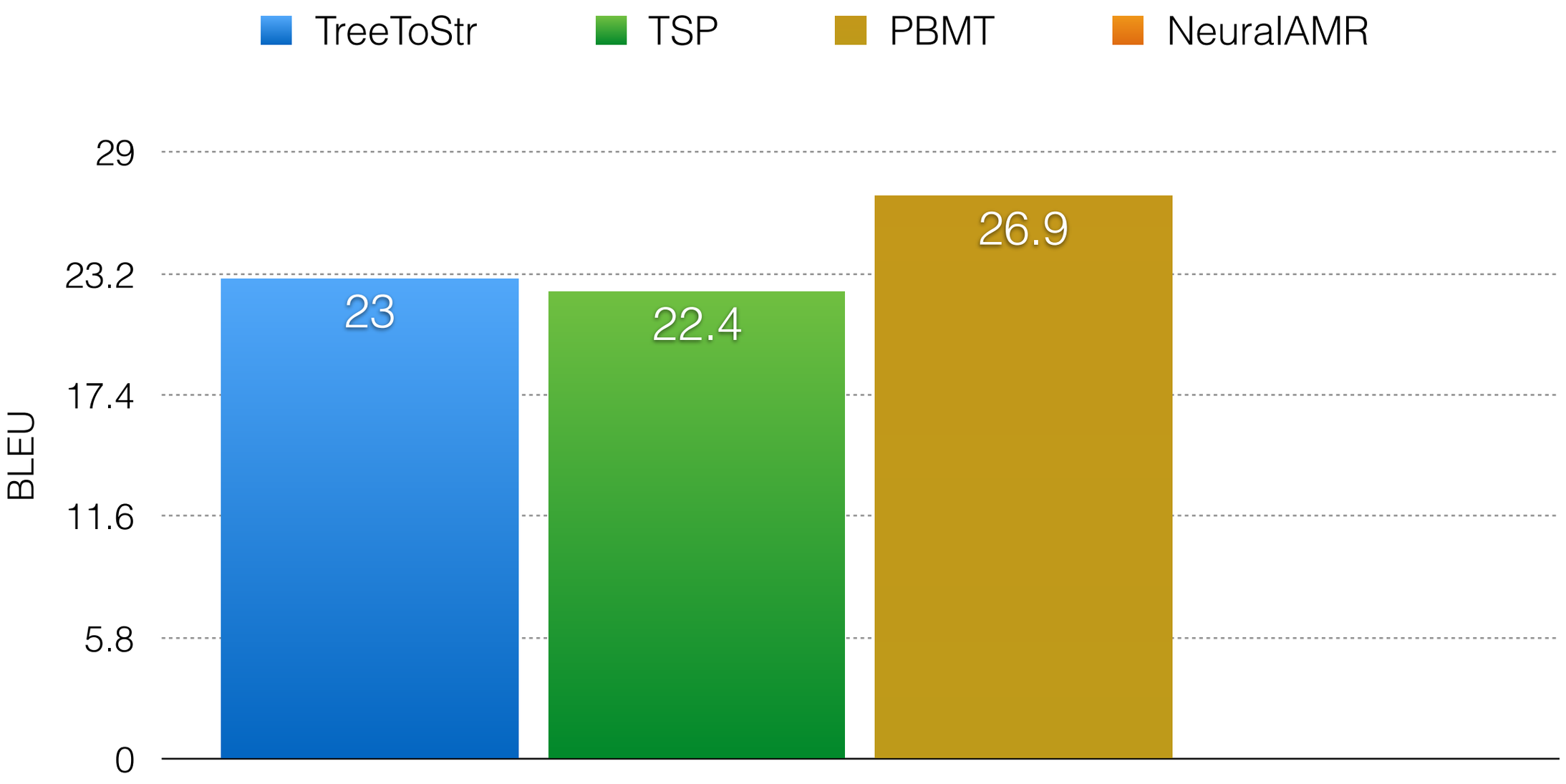
Experiments

- Vanilla experiment
  - Limited Language Model Capacity
- Paired Training
  - Data augmentation algorithm

First Attempt (**Generation**)

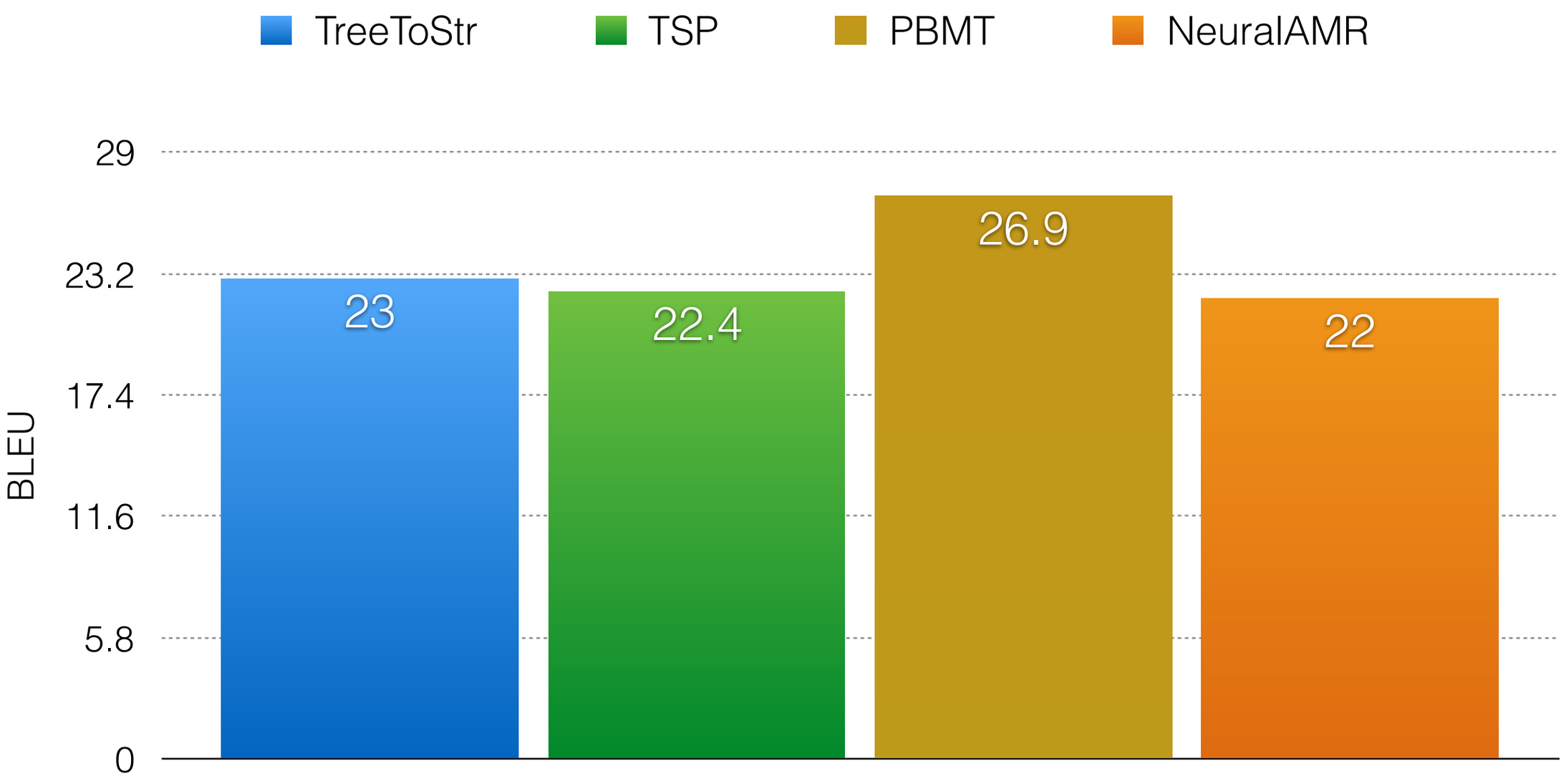
TreeToStr: Flanigan et al, NAACL 2016
TSP: Song et al, EMNLP 2016
PBMT: Pourdamaghani and Knight, INLG 2016

First Attempt (**Generation**)



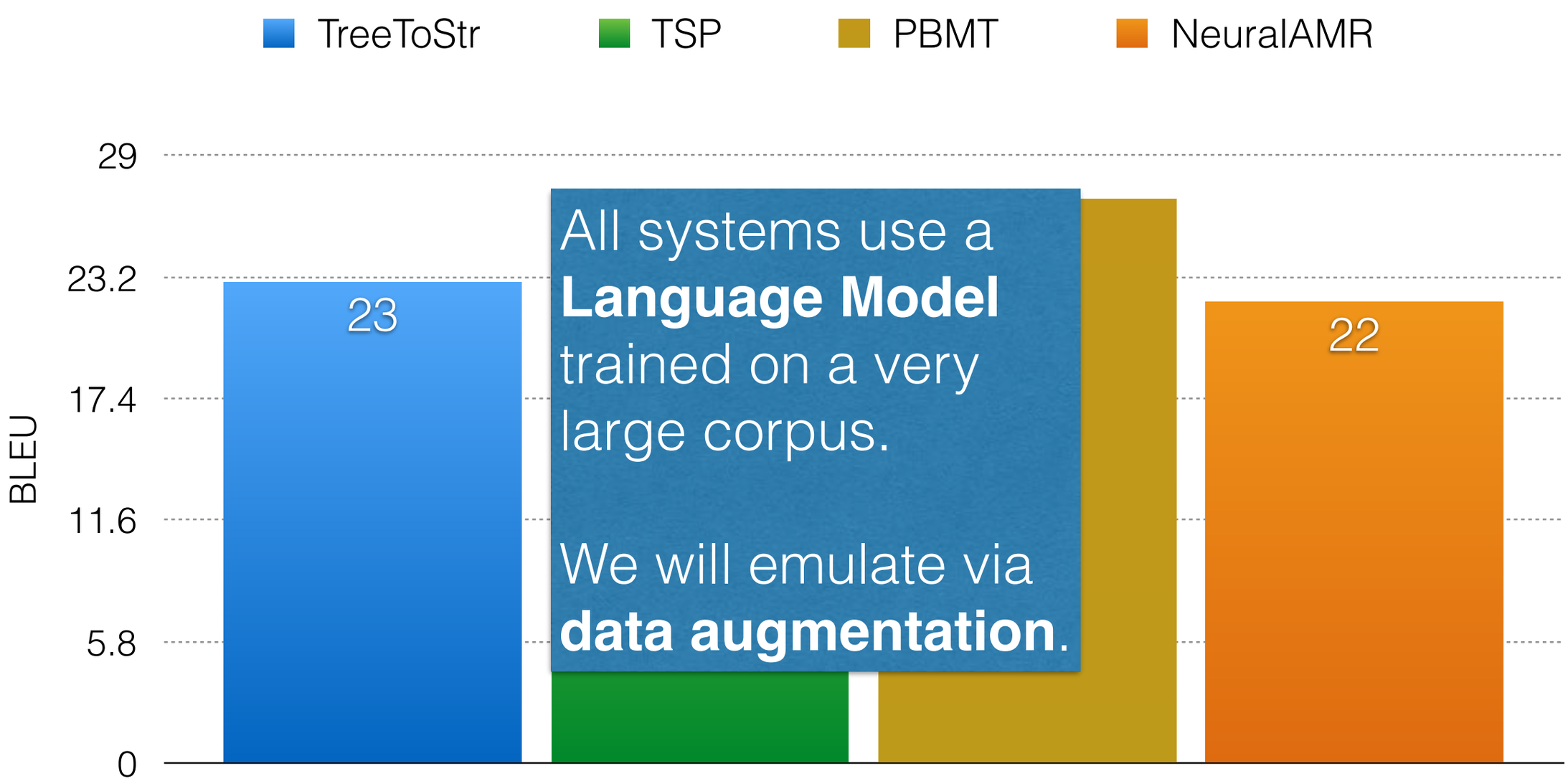
TreeToStr: Flanigan et al, NAACL 2016
TSP: Song et al, EMNLP 2016
PBMT: Pourdamaghani and Knight, INLG 2016

First Attempt (**Generation**)



TreeToStr: Flanigan et al, NAACL 2016
TSP: Song et al, EMNLP 2016
PBMT: Pourdamaghani and Knight, INLG 2016

First Attempt (**Generation**)



TreeToStr: Flanigan et al, NAACL 2016
TSP: Song et al, EMNLP 2016
PBMT: Pourdamaghani and Knight, INLG 2016

(Sennrich et al., ACL 2016)

What went wrong?

```
hold
   :ARG0 (person
            :ARG0-of (have-role
               :ARG1 loc_0
               :ARG2 official)
         )
   :ARG1 (meet
            :ARG0 (person
               :ARG1-of expert
               :ARG2-of group)
         )
   :time (date-entity year_0 month_0)
   :location loc_1
```

Reference

US officials held an expert group meeting in January 2002 in New York .

Prediction

United States officials held held a meeting in January 2002 .

What went wrong?

```
hold
   :ARG0 (person
            :ARG0-of (have-role
               :ARG1 loc_0
               :ARG2 official)
         )
   :ARG1 (meet
            :ARG0 (person
               :ARG1-of expert
               :ARG2-of group)
         )
   :time (date-entity year_0 month_0)
   :location loc_1
```

Reference

US officials held an expert group meeting in January 2002 in New York .

Prediction

United States officials held held a meeting in January 2002 .

- Repetition

What went wrong?

```
hold
   :ARG0 (person
            :ARG0-of (have-role
               :ARG1 loc_0
               :ARG2 official)
         )
   :ARG1 (meet
            :ARG0 (person
               :ARG1-of expert
               :ARG2-of group)
         )
   :time (date-entity year_0 month_0)
   :location loc_1
```

Reference

US officials held an expert group meeting in January 2002 in New York .

Prediction

United States officials held held a meeting in January 2002 .

- Repetition
- Coverage

What went wrong?

```
hold
   :ARG0 (person
            :ARG0-of (have-role
               :ARG1 loc_0
               :ARG2 official)
         )
   :ARG1 (meet
            :ARG0 (person
               :ARG1-of expert
               :ARG2-of group)
         )
   :time (date-entity year_0 month_0)
   :location loc_1
```

- Repetition
- Coverage

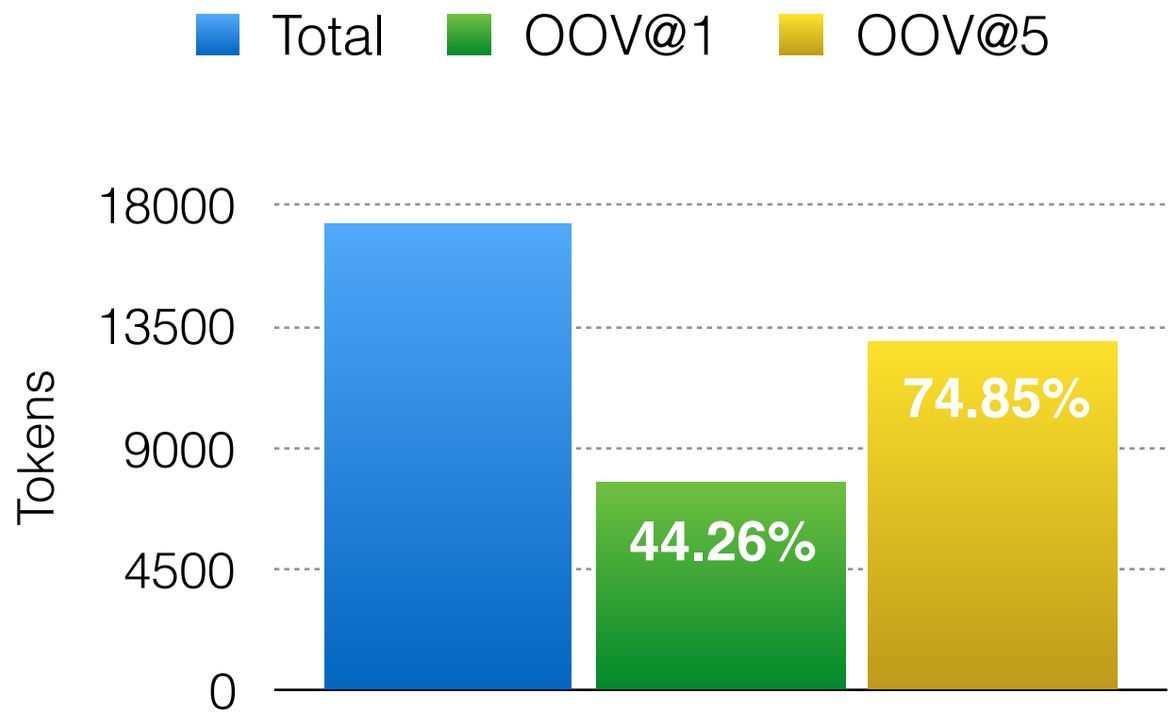
  a) Sparsity

Reference

US officials held an expert group meeting in January 2002 in New York .

Prediction

United States officials held held a meeting in January 2002 .



What went wrong?

```
hold
   :ARG0 (person
            :ARG0-of (have-role
               :ARG1 loc_0
               :ARG2 official)
         )
   :ARG1 (meet
            :ARG0 (person
               :ARG1-of expert
               :ARG2-of group)
         )
   :time (date-entity year_0 month_0)
   :location loc_1
```

Reference

US officials held an expert group meeting in January 2002 in New York .

Prediction

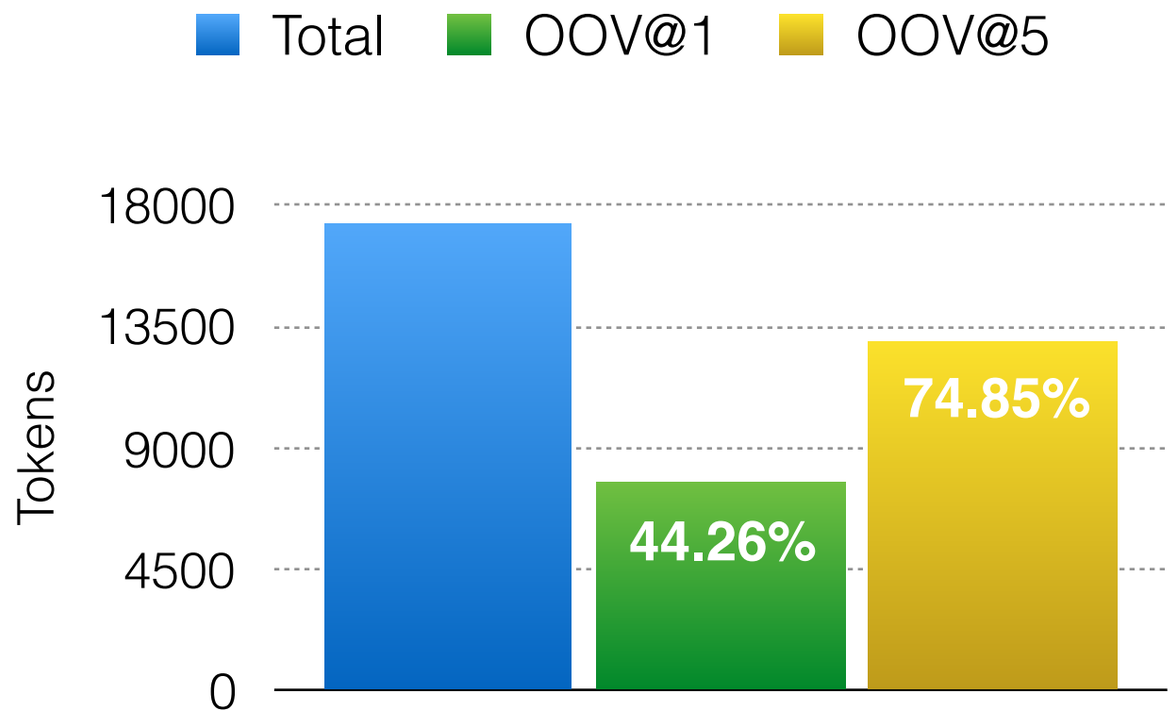
United States officials held held a meeting in January 2002 .

- Repetition
- Coverage

a) Sparsity

b) Avg sent length: 20 words

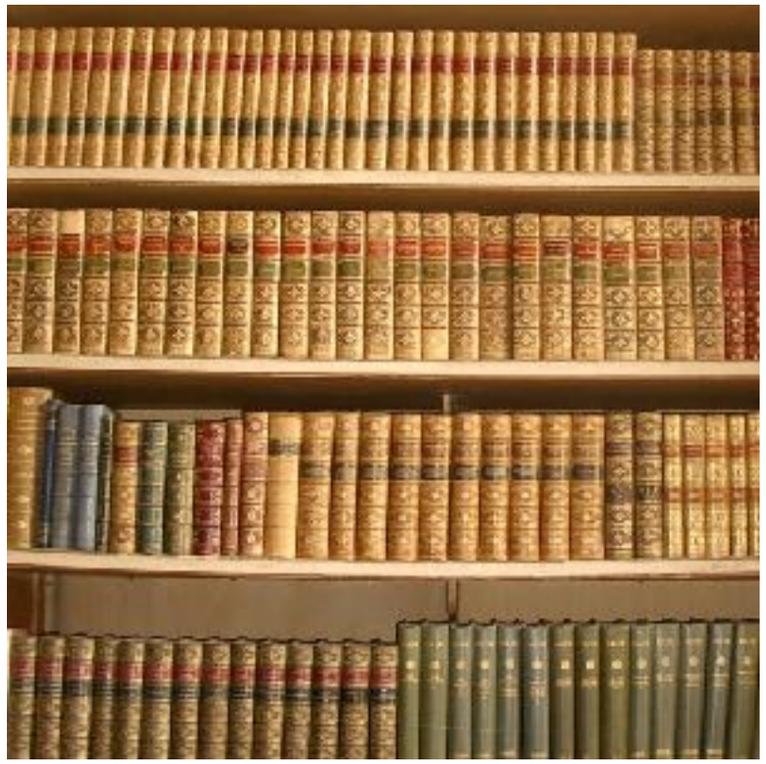
c) Limited Language Modeling capacity



Data Augmentation

Original Dataset: ~16k graph-sentence **pairs**

Data Augmentation



Original Dataset: ~16k graph-sentence **pairs**

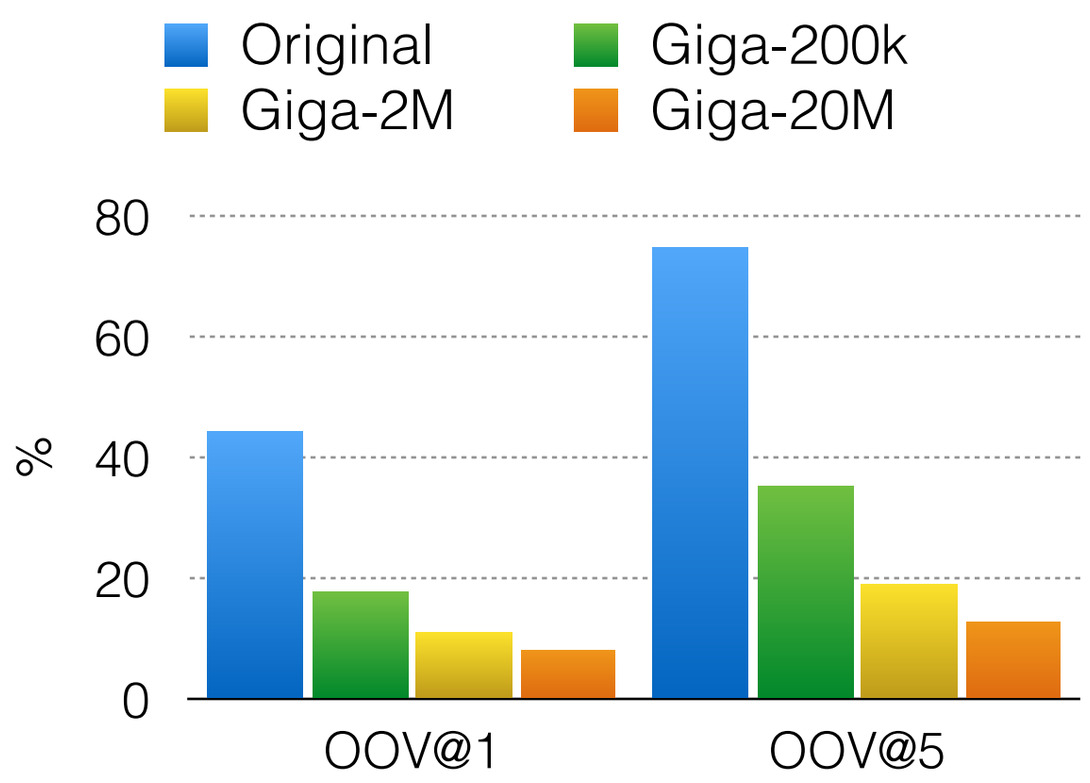
Gigaword: ~183M sentences ***only***

Data Augmentation

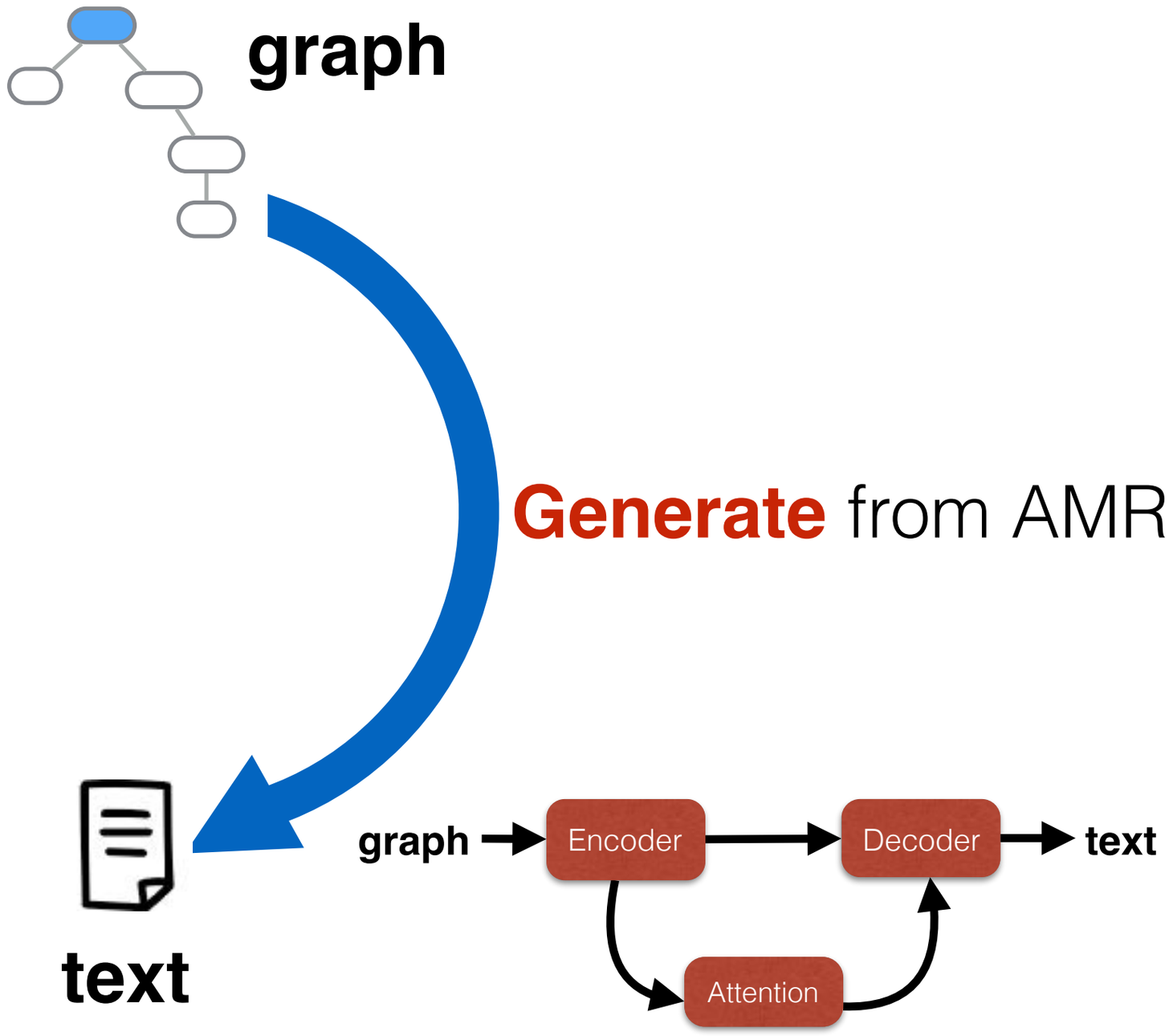
Original Dataset: ~16k graph-sentence **pairs**

Gigaword: ~183M sentences ***only***

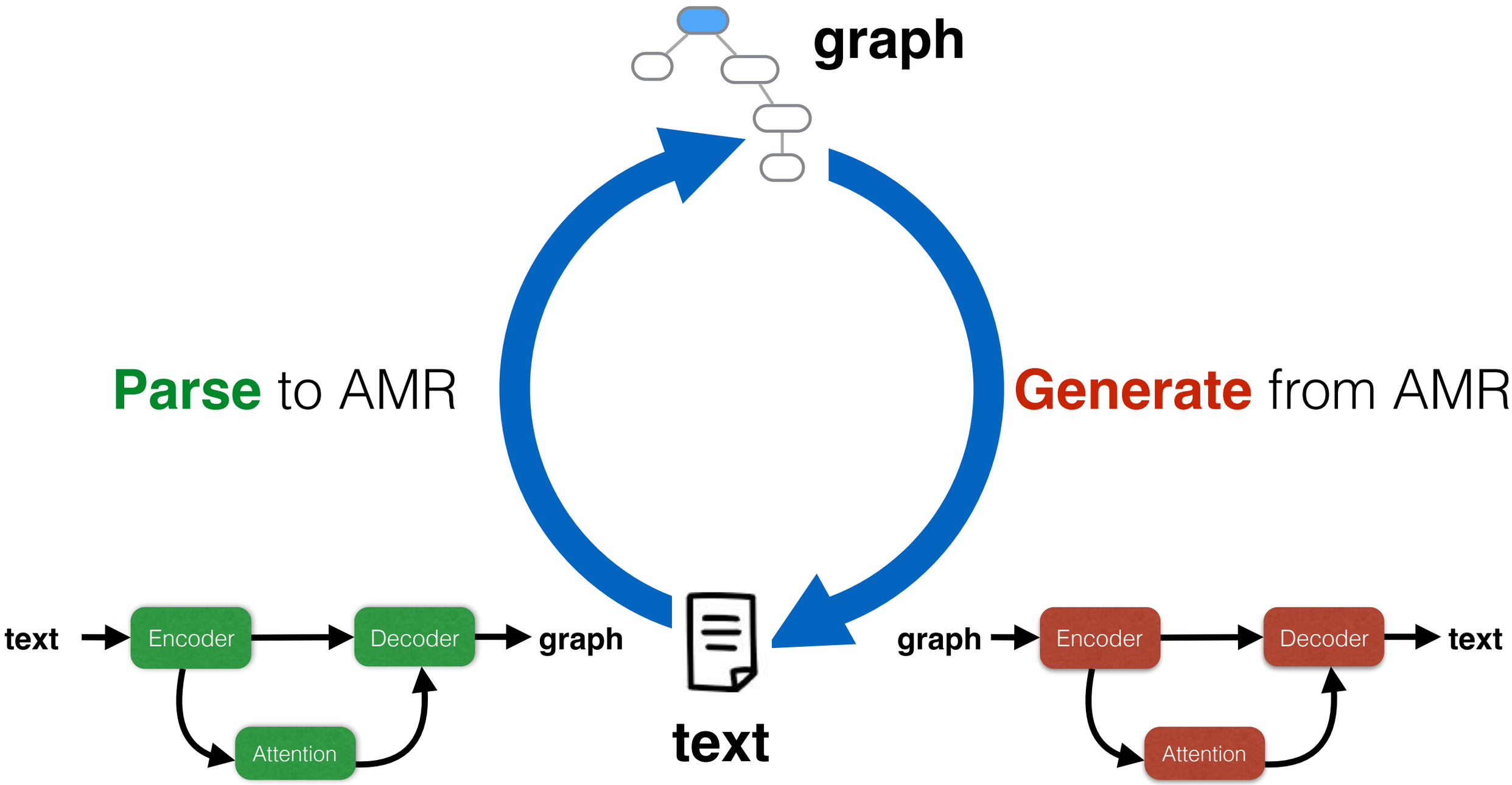
Sample sentences with vocabulary overlap



Data Augmentation

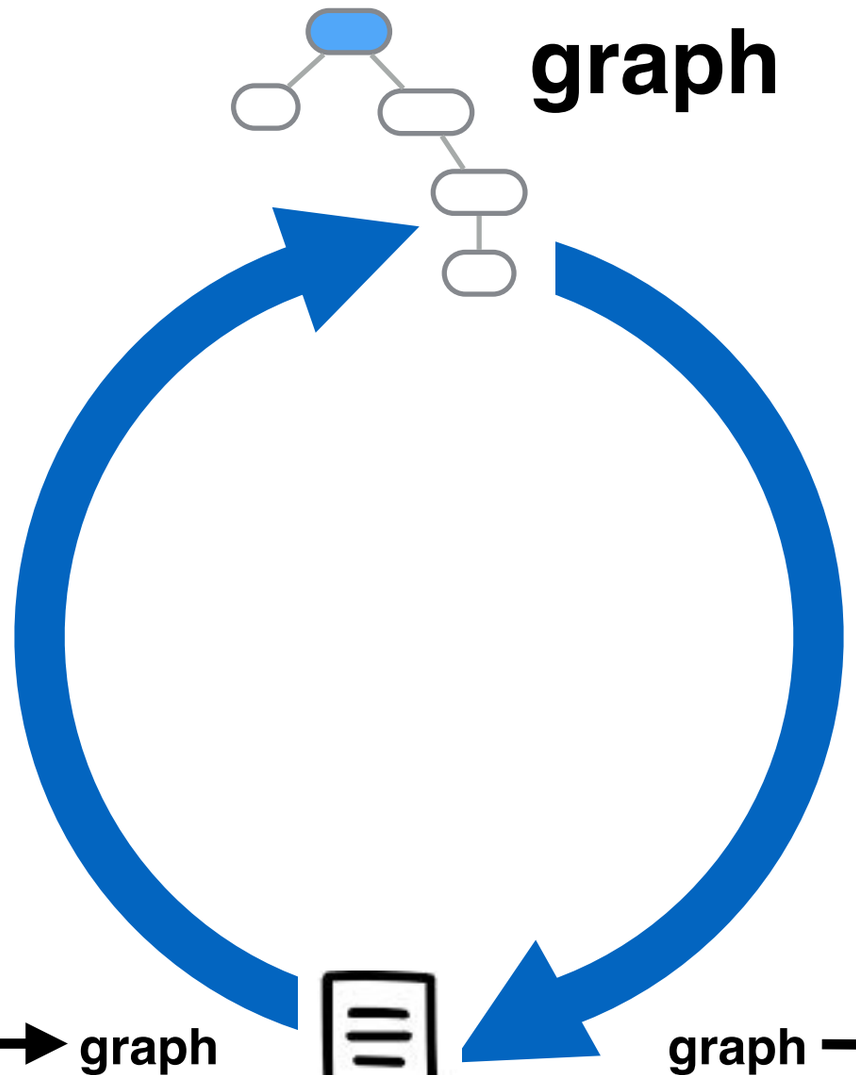


Data Augmentation

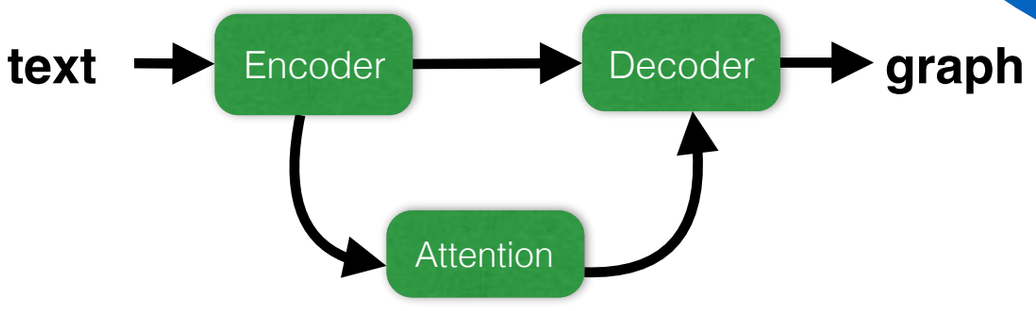


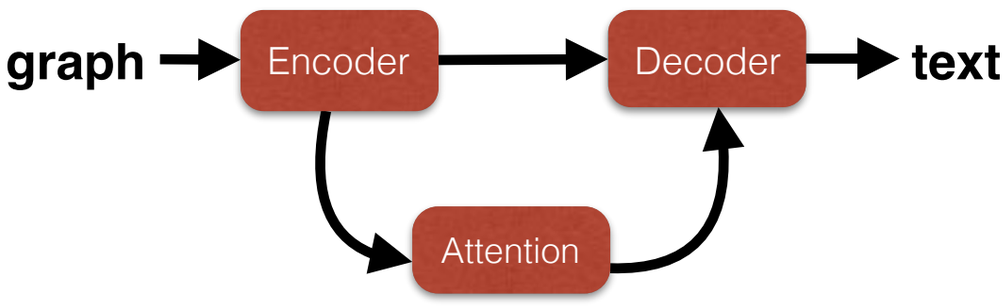
Data Augmentation

Parse to AMR



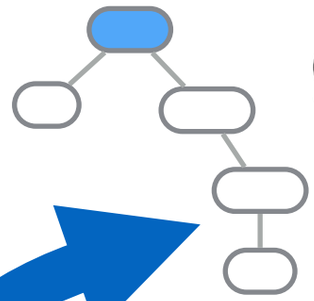
Generate from AMR





Data Augmentation

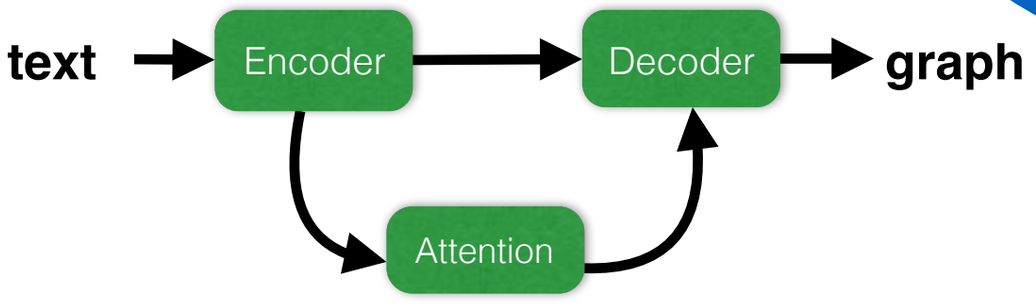
Parse to AMR



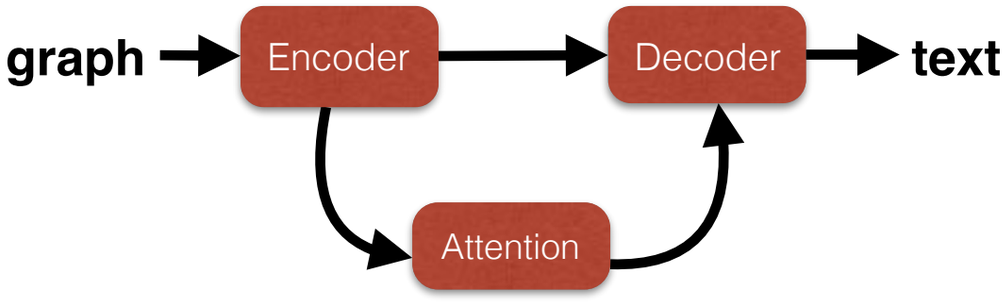
graph

Re-train

Generate from AMR



text



Semi-supervised Learning

- **Self-training**
  - McClosky et al. 2006

- **Co-training**
  - Yarowski 1995, Blum and Mitchell 1998, Sarkar 2001
  - Sogaard and Rishoj, 2010

Paired Training

Paired Training

Train AMR Parser **P** on Original Dataset

( , )

Paired Training

Train AMR Parser **P** on Original Dataset

( , )

for i = 0 … N

Paired Training

Train AMR Parser **P** on Original Dataset

for i = 0 … N

S_i =Sample **k** 10^i sentences from Gigaword

Paired Training

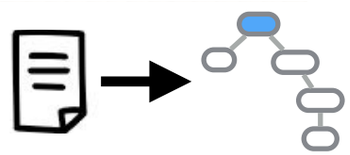
Train AMR Parser **P** on Original Dataset

( , )

for i = 0 … N

S_i =Sample **k 10^i** sentences from Gigaword

Parse **S_i** sentences with **P**



Paired Training

Train AMR Parser **P** on Original Dataset

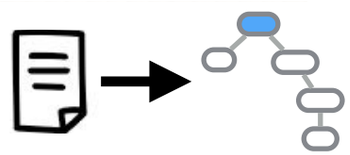
(,)

for i = 0 … N

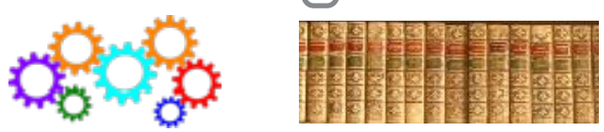
Self-train P

S_i =Sample **k 10^i** sentences from Gigaword

Parse **S_i** sentences with **P**



Re-train AMR Parser **P** on **S_i**



Paired Training

Train AMR Parser **P** on Original Dataset

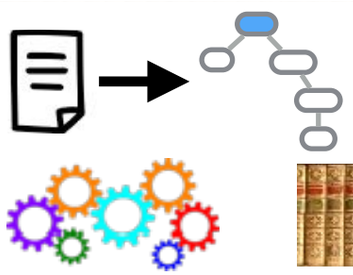
(,)

for i = 0 … N

Self-train P

S_i =Sample **k** 10^i sentences from Gigaword

Parse S_i sentences with **P**



Re-train AMR Parser **P** on S_i

Paired Training

Train AMR Parser **P** on Original Dataset

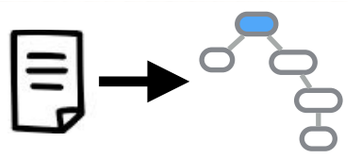
(,)

for i = 0 … N

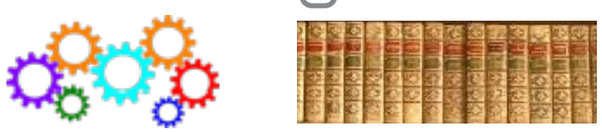
Self-train P

S_i =Sample **k** 10^i sentences from Gigaword

Parse S_i sentences with **P**



Re-train AMR Parser **P** on S_i

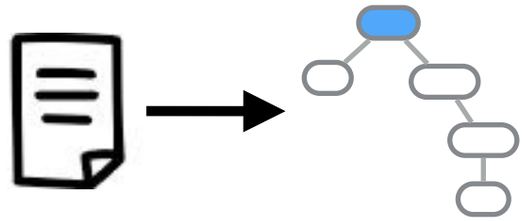


Train Generator **G** on S_N

(,)

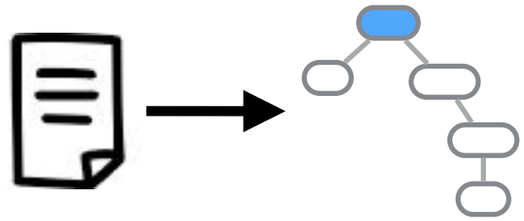
Training AMR **Parser**

Train **P** on
Original Dataset



Training AMR **Parser**

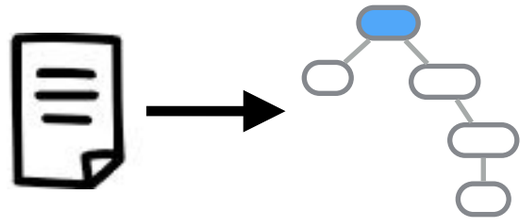
Train **P** on
Original Dataset



Training AMR **Parser**

Train **P** on Original Dataset

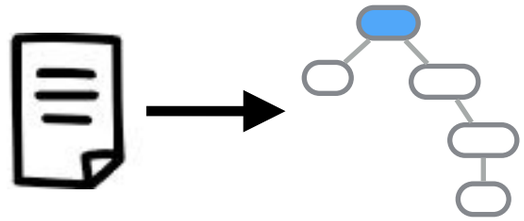
Sample $\mathbf{S_1}$=**200k** sentences from Gigaword



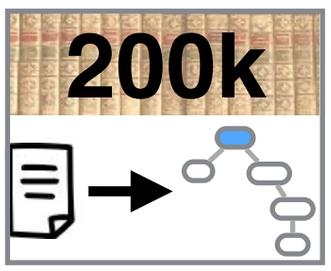
Training AMR **Parser**

Train **P** on Original Dataset

Sample $\mathbf{S_1}$=**200k** sentences from Gigaword



Parse $\mathbf{S_1}$ with **P** (,)



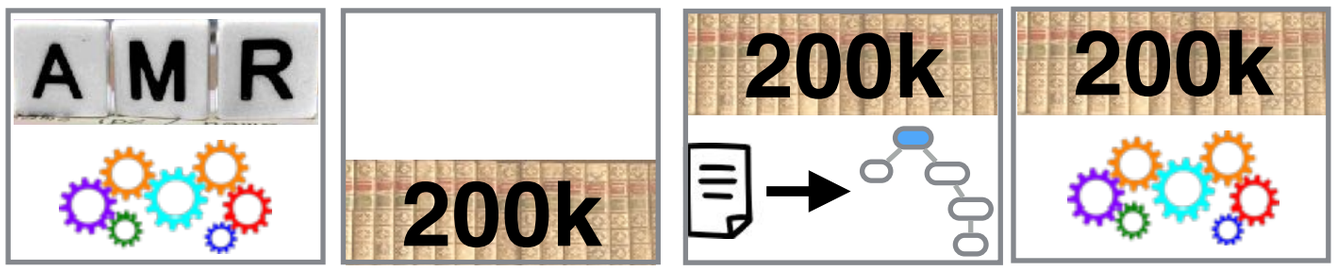
Training AMR **Parser**

Train **P** on Original Dataset

Sample $\mathbf{S_1}$=**200k** sentences from Gigaword

Parse $\mathbf{S_1}$ with **P** (,)

Train **P** on $\mathbf{S_1}$=**200k**



Fine-tune: init parameters from previous step and train on Original Dataset

Training AMR **Parser**

Train **P** on Original Dataset

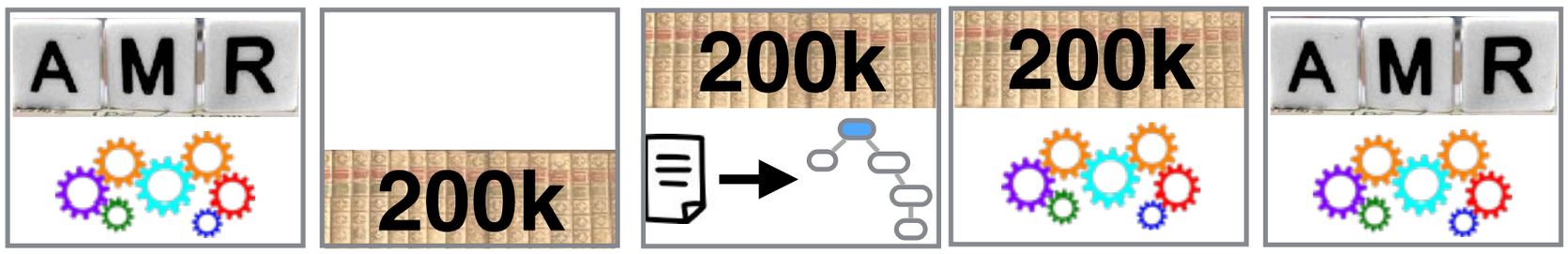
Sample $\mathbf{S_1}$=**200k** sentences from Gigaword

Parse $\mathbf{S_1}$ with **P**

(,)

Train **P** on $\mathbf{S_1}$=**200k**

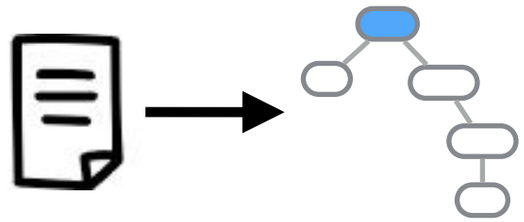
Fine-tune **P** on Original Dataset



Fine-tune: init parameters from previous step and train on Original Dataset

Training AMR **Parser**

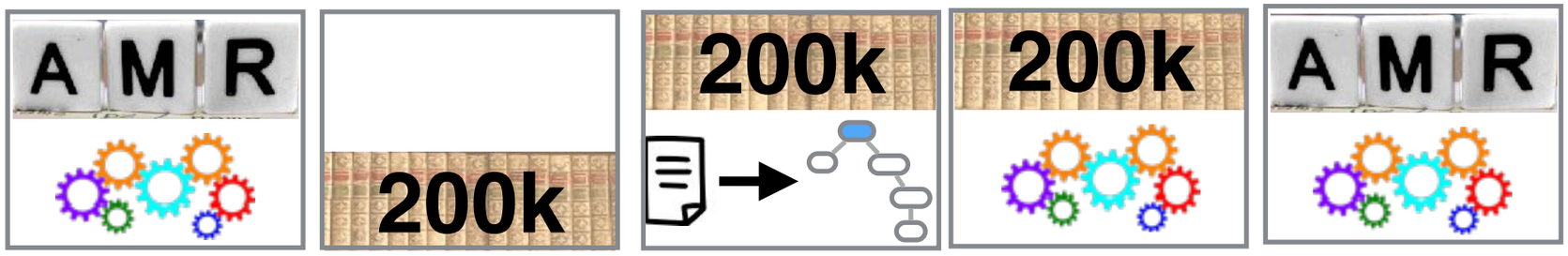
Sample $\mathbf{S_2}$=**2M** sentences from Gigaword



Parse $\mathbf{S_2}$ with **P** (,)

Train **P** on $\mathbf{S_2}$=**2M**

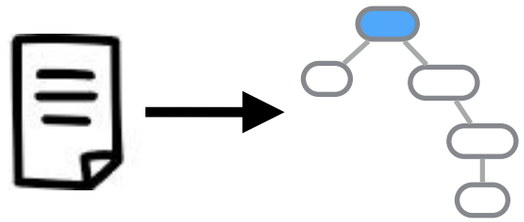
Fine-tune **P** on Original Dataset



Fine-tune: init parameters from previous step and train on Original Dataset

Training AMR **Parser**

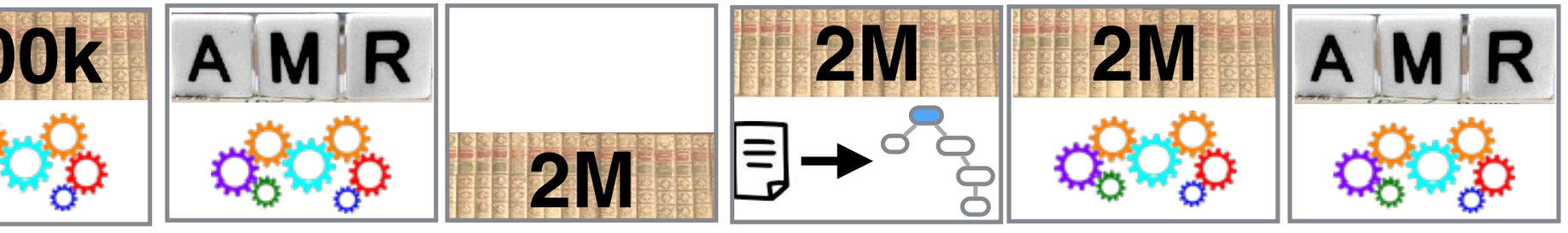
Sample $\mathbf{S_2}$=**2M** sentences from Gigaword



Parse $\mathbf{S_2}$ with **P** (,)

Train **P** on $\mathbf{S_2}$=**2M**

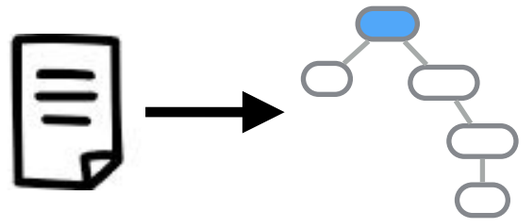
Fine-tune **P** on Original Dataset



Fine-tune: init parameters from previous step and train on Original Dataset

Training AMR **Parser**

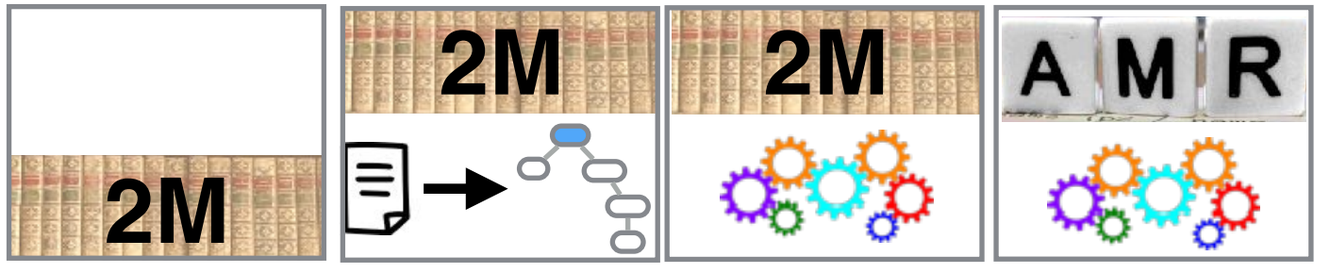
Sample $\mathbf{S_3}$=**20M** sentences from Gigaword



Parse $\mathbf{S_3}$ with **P** (,)

Train **P** on $\mathbf{S_3}$=**20M**

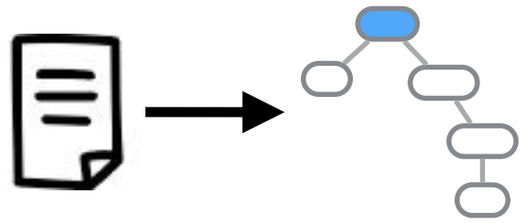
Fine-tune **P** on Original Dataset



Fine-tune: init parameters from previous step and train on Original Dataset

Training AMR **Parser**

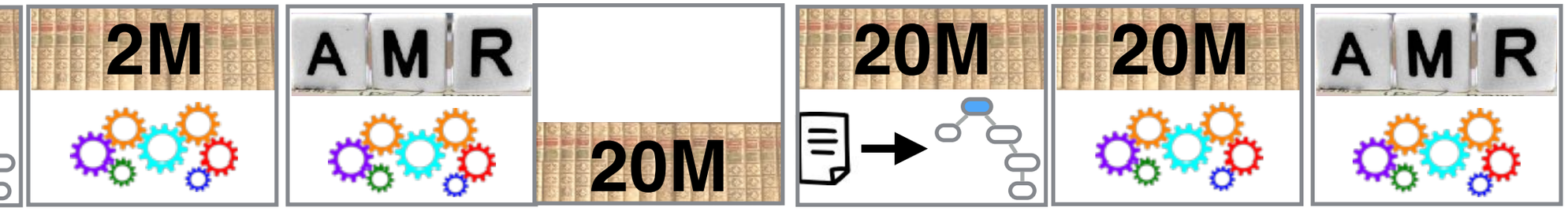
Sample $\mathbf{S_3}$=**20M** sentences from Gigaword



Parse $\mathbf{S_3}$ with **P** (,)

Train **P** on $\mathbf{S_3}$=**20M**

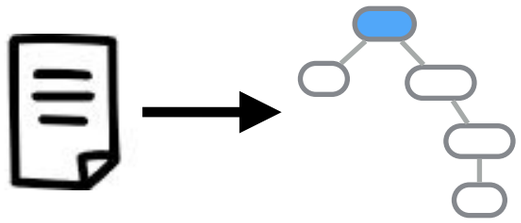
Fine-tune **P** on Original Dataset



Fine-tune: init parameters from previous step and train on Original Dataset

Training AMR **Parser**

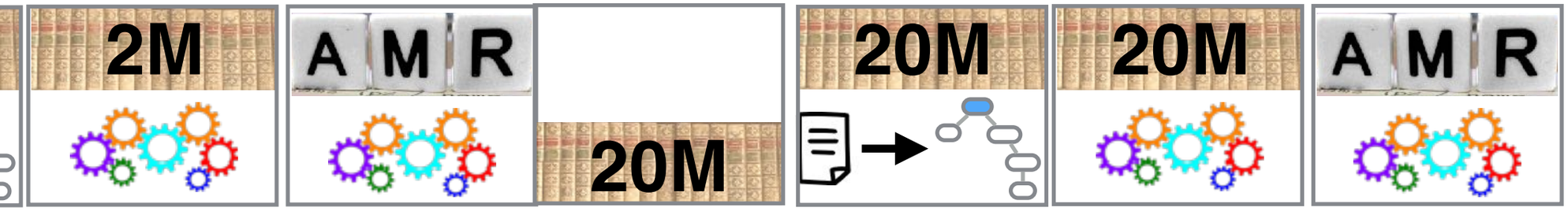
Sample **S_3=20M** sentences from Gigaword



Parse **S_3** with **P** (,)

Train **P** on **S_3=20M**

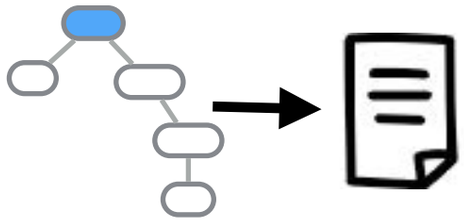
Fine-tune **P** on Original Dataset



Fine-tune: init parameters from previous step and train on Original Dataset

Training AMR **Generator**

Sample **S_4=20M** sentences from Gigaword

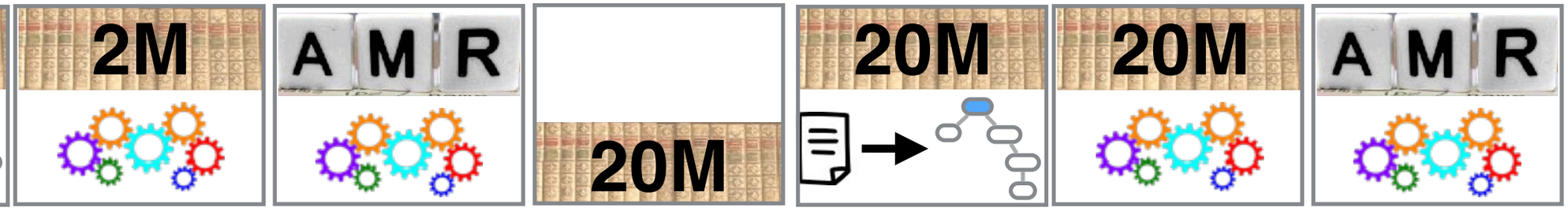


Parse **S_4** with **P**

( , )

Train **G** on **S_4=20M**

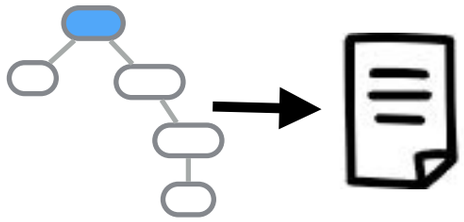
Fine-tune **G** on Original Dataset



Fine-tune: init parameters from previous step and train on Original Dataset

Training AMR **Generator**

Sample $\mathbf{S_4}$=**20M** sentences from Gigaword

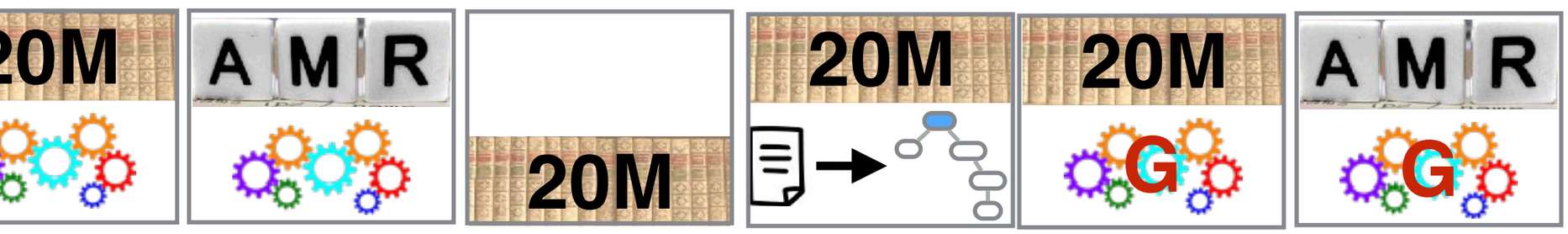


Parse $\mathbf{S_4}$ with **P**

( , )

Train **G** on $\mathbf{S_4}$=**20M**

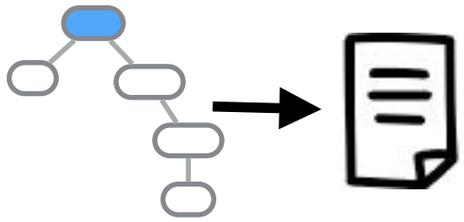
Fine-tune **G** on Original Dataset



Fine-tune: init parameters from previous step and train on Original Dataset

Training AMR **Generator**

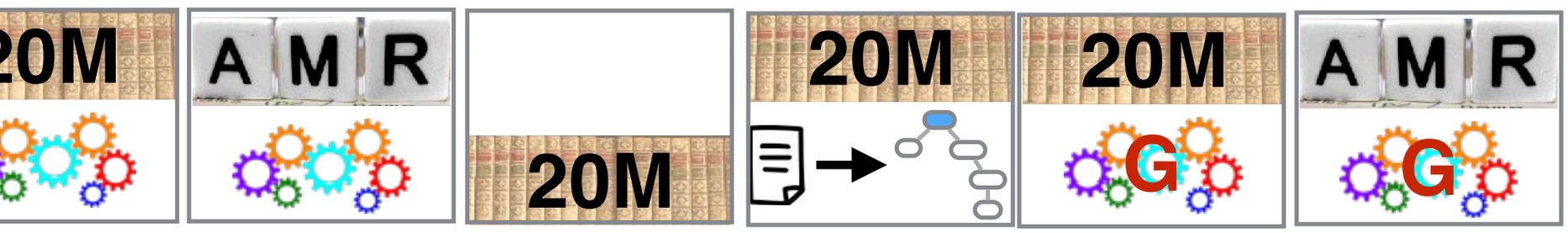
Sample $\mathbf{S_4}$=**20M** sentences from Gigaword



Parse $\mathbf{S_4}$ with **P** (,)

Train **G** on $\mathbf{S_4}$=**20M**

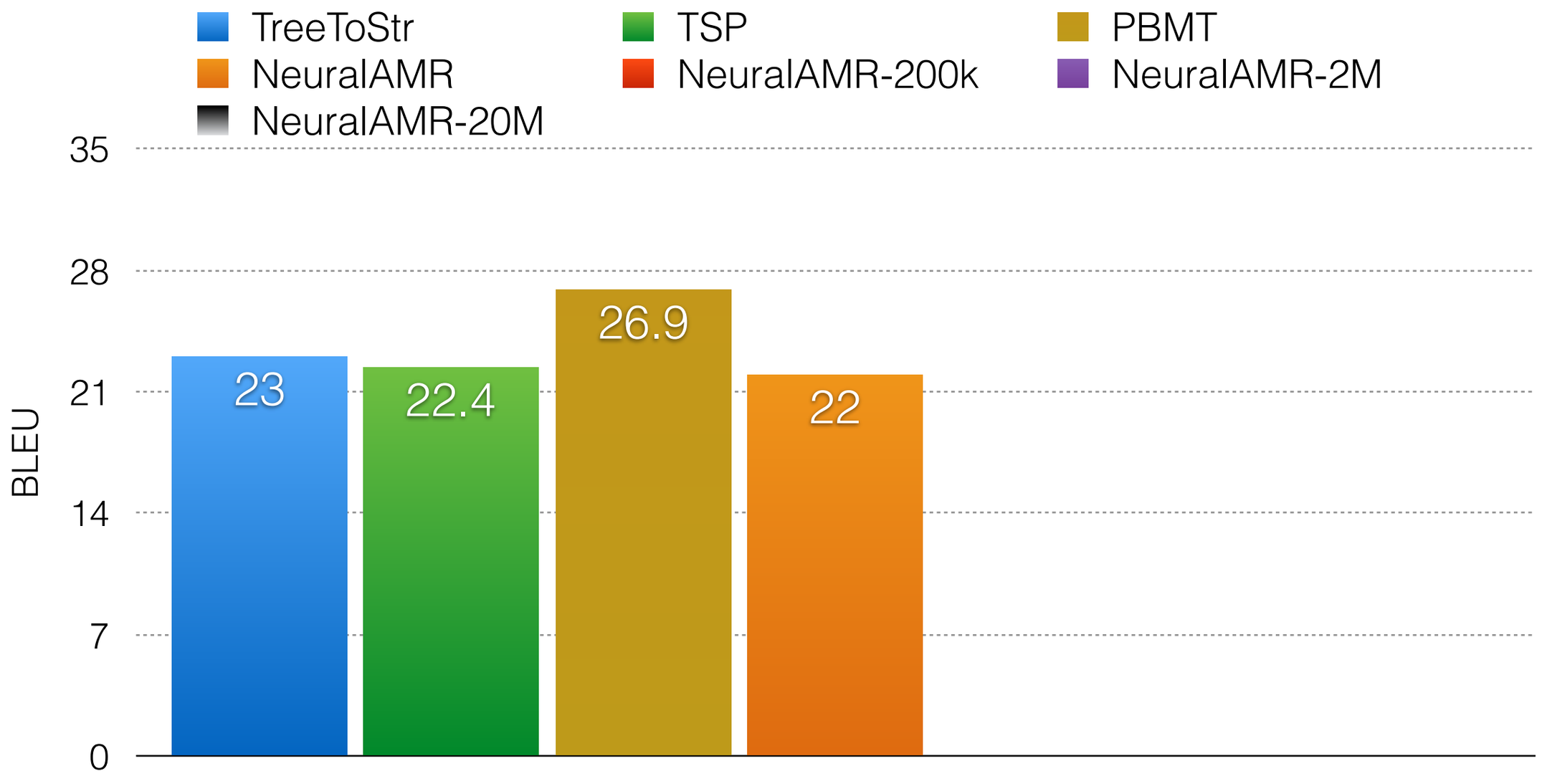
Fine-tune **G** on Original Dataset



Final Results (**Generation**)

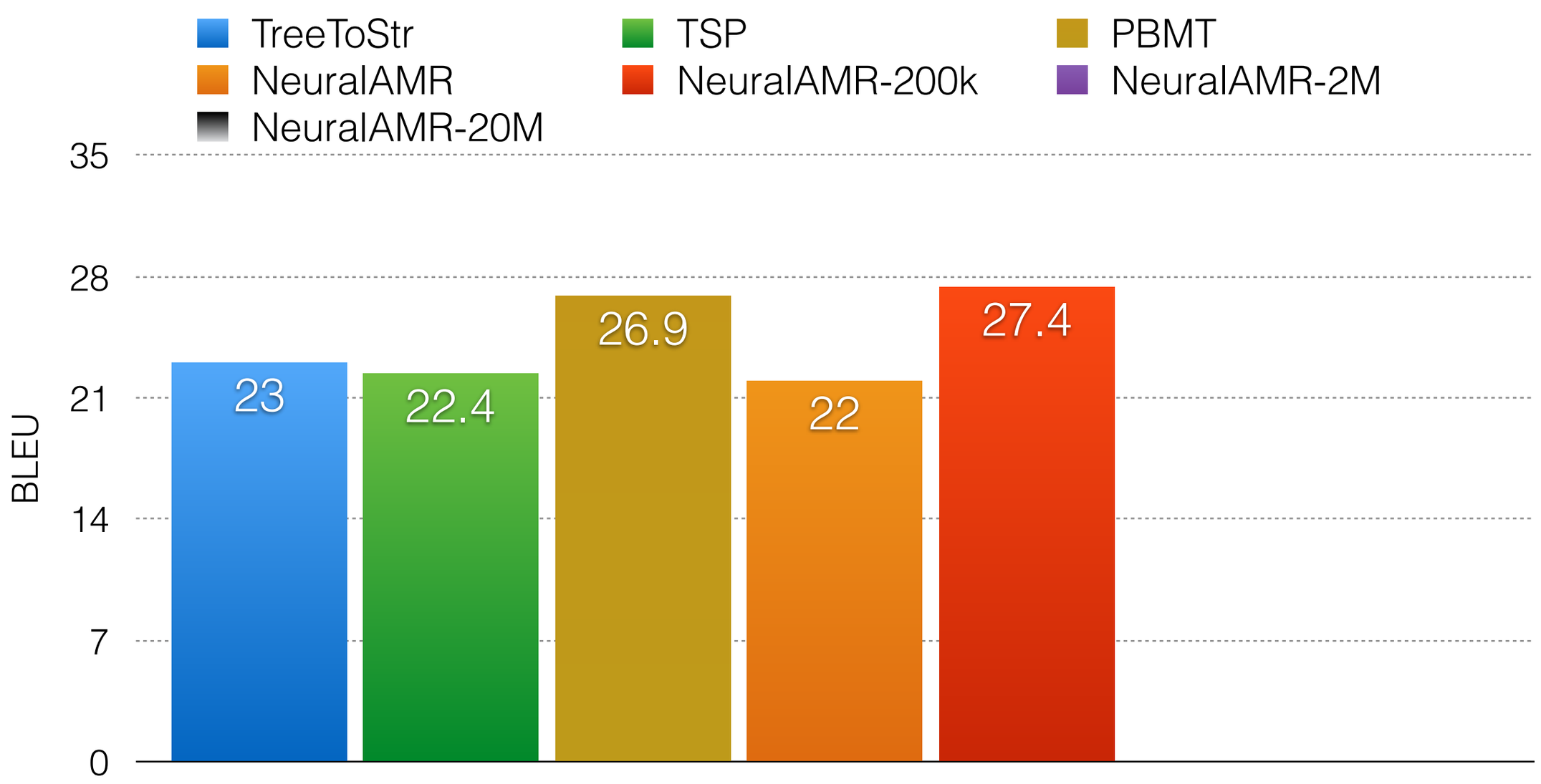
TreeToStr: Flanigan et al, NAACL 2016
TSP: Song et al, EMNLP 2016
PBMT: Pourdamaghani and Knight, INLG 2016

Final Results (**Generation**)



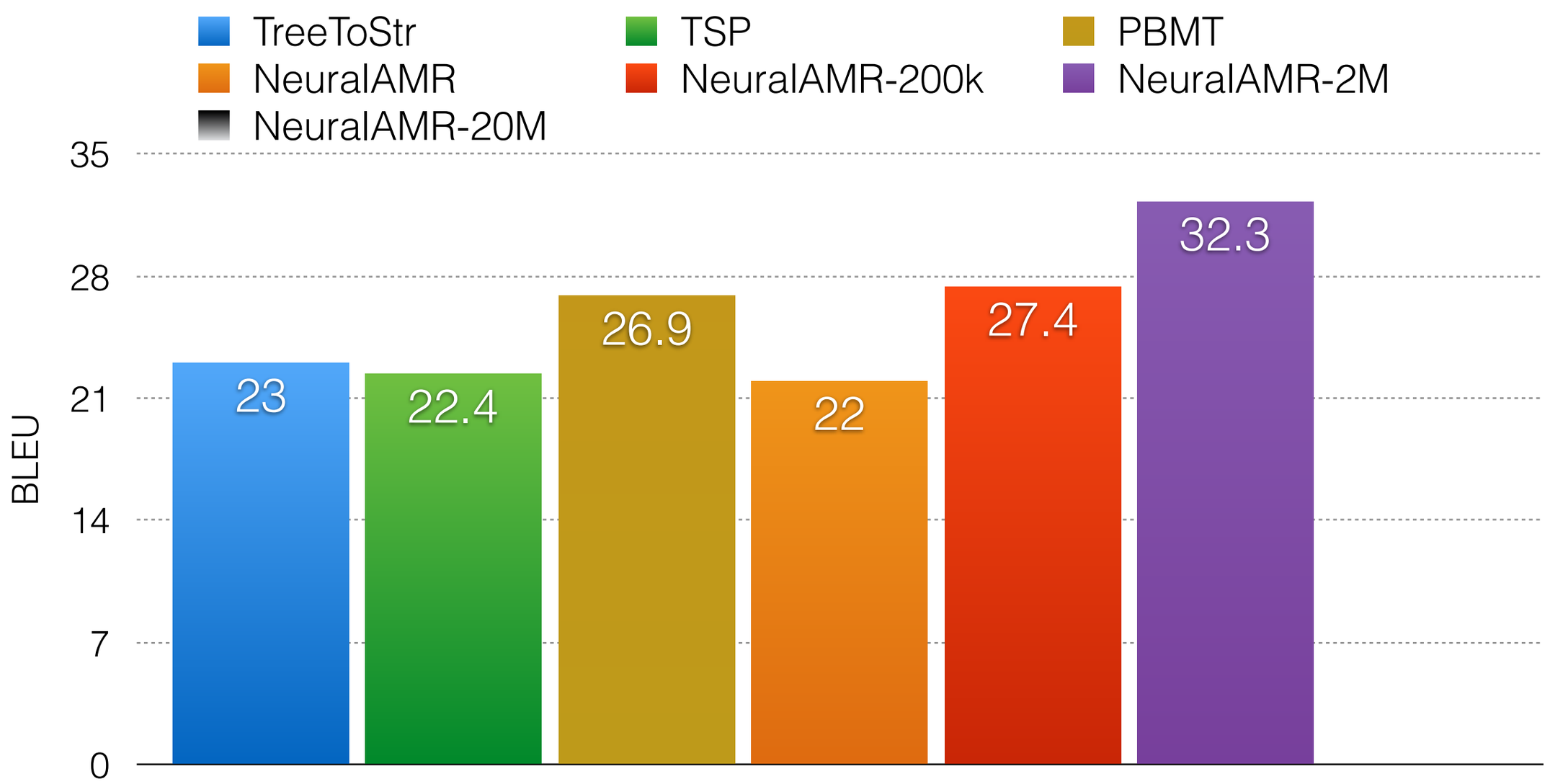
TreeToStr: Flanigan et al, NAACL 2016
TSP: Song et al, EMNLP 2016
PBMT: Pourdamaghani and Knight, INLG 2016

Final Results (**Generation**)



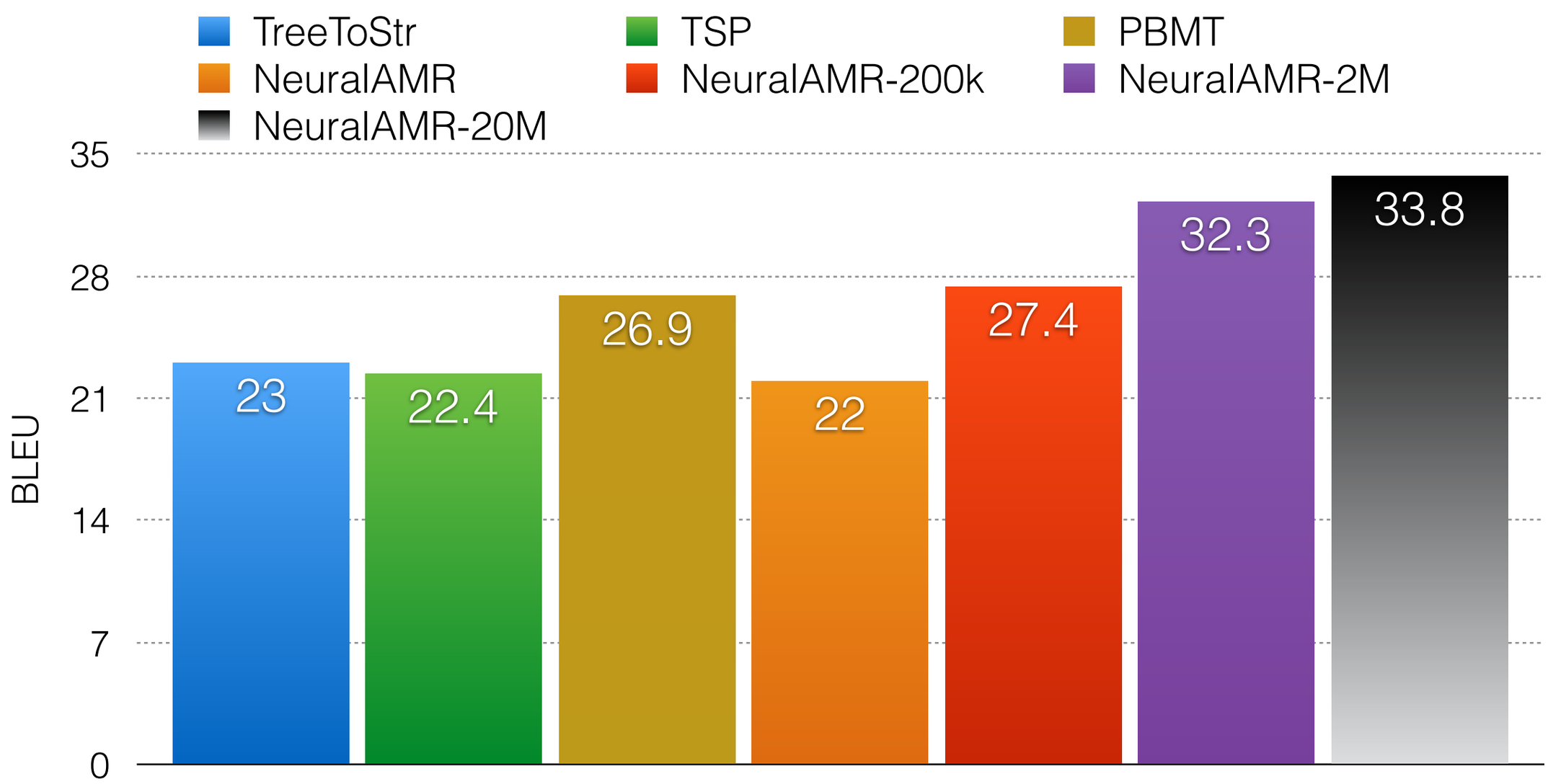
TreeToStr: Flanigan et al, NAACL 2016
TSP: Song et al, EMNLP 2016
PBMT: Pourdamaghani and Knight, INLG 2016

Final Results (**Generation**)



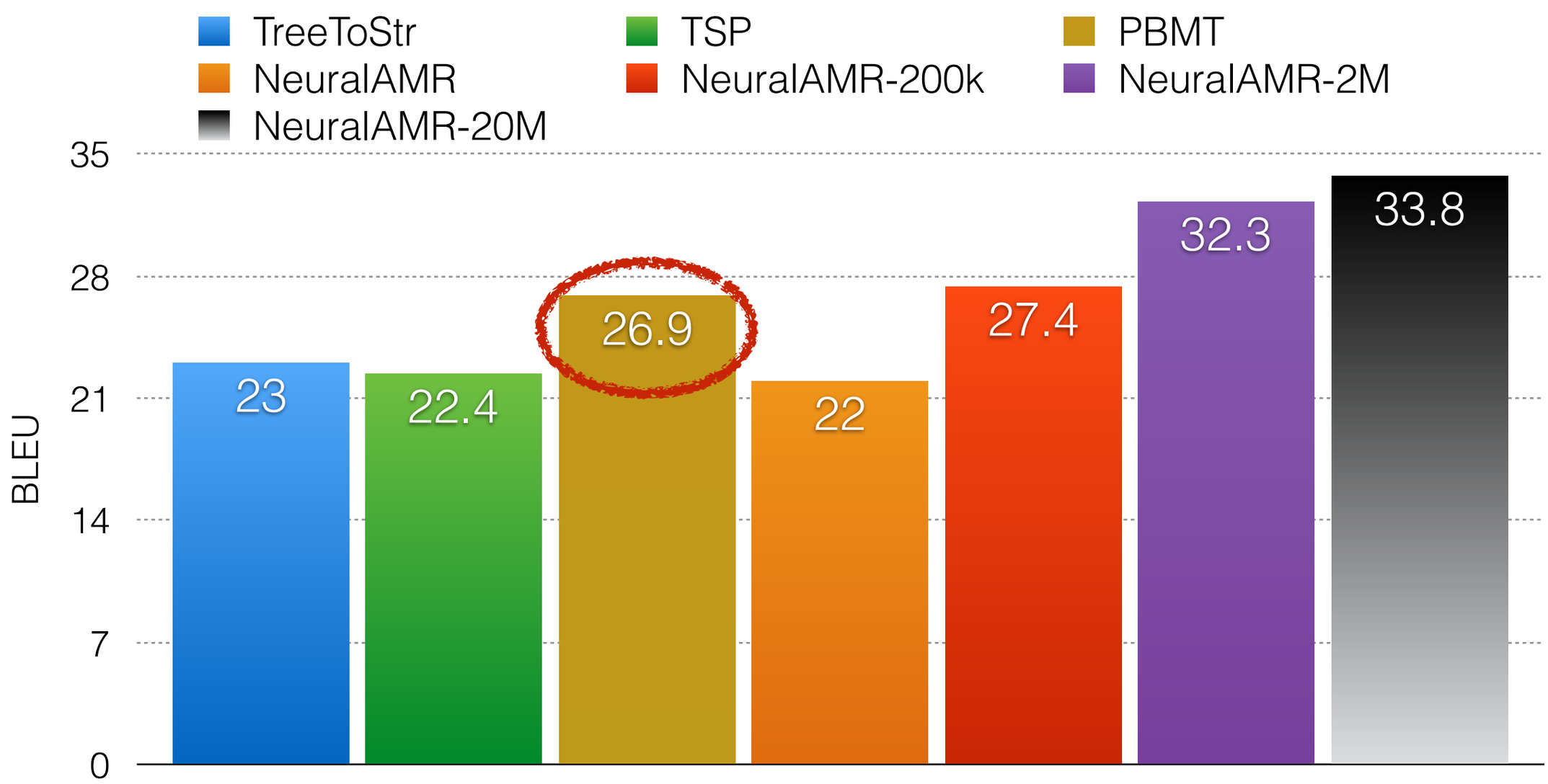
TreeToStr: Flanigan et al, NAACL 2016
TSP: Song et al, EMNLP 2016
PBMT: Pourdamaghani and Knight, INLG 2016

Final Results (**Generation**)



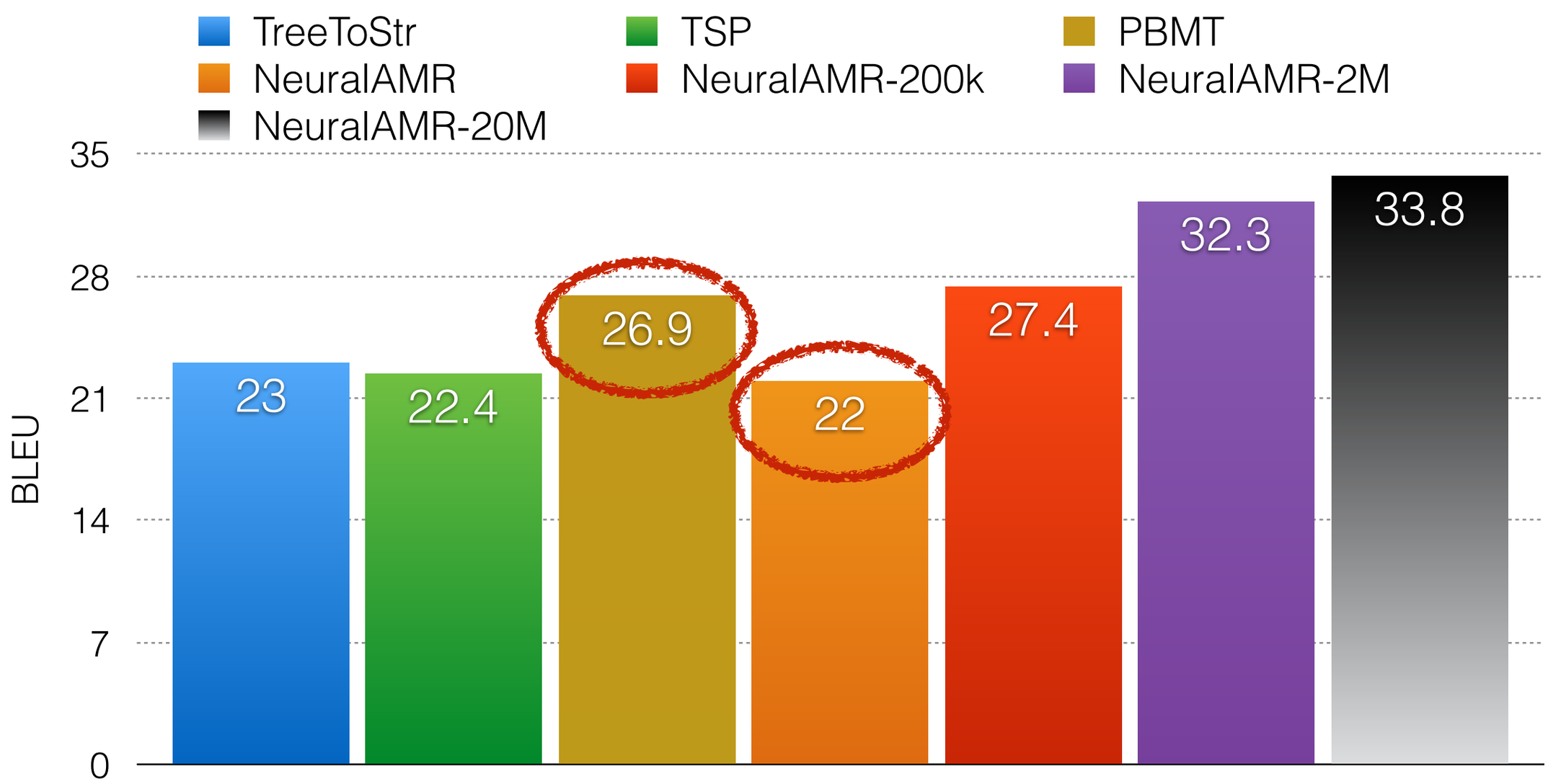
TreeToStr: Flanigan et al, NAACL 2016
TSP: Song et al, EMNLP 2016
PBMT: Pourdamaghani and Knight, INLG 2016

Final Results (**Generation**)



TreeToStr: Flanigan et al, NAACL 2016
TSP: Song et al, EMNLP 2016
PBMT: Pourdamaghani and Knight, INLG 2016

Final Results (**Generation**)

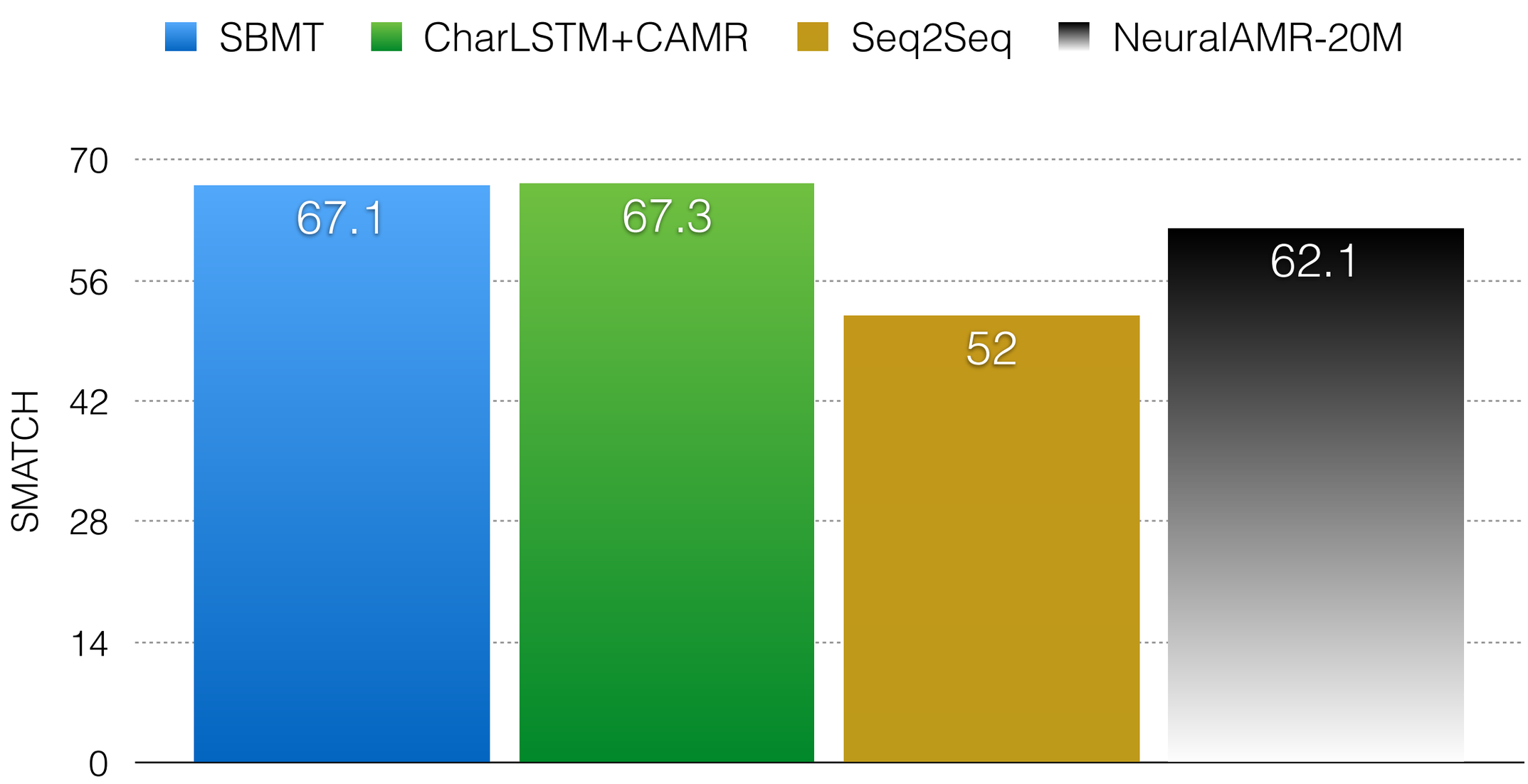


TreeToStr: Flanigan et al, NAACL 2016
TSP: Song et al, EMNLP 2016
PBMT: Pourdamaghani and Knight, INLG 2016

Final Results (**Parsing**)

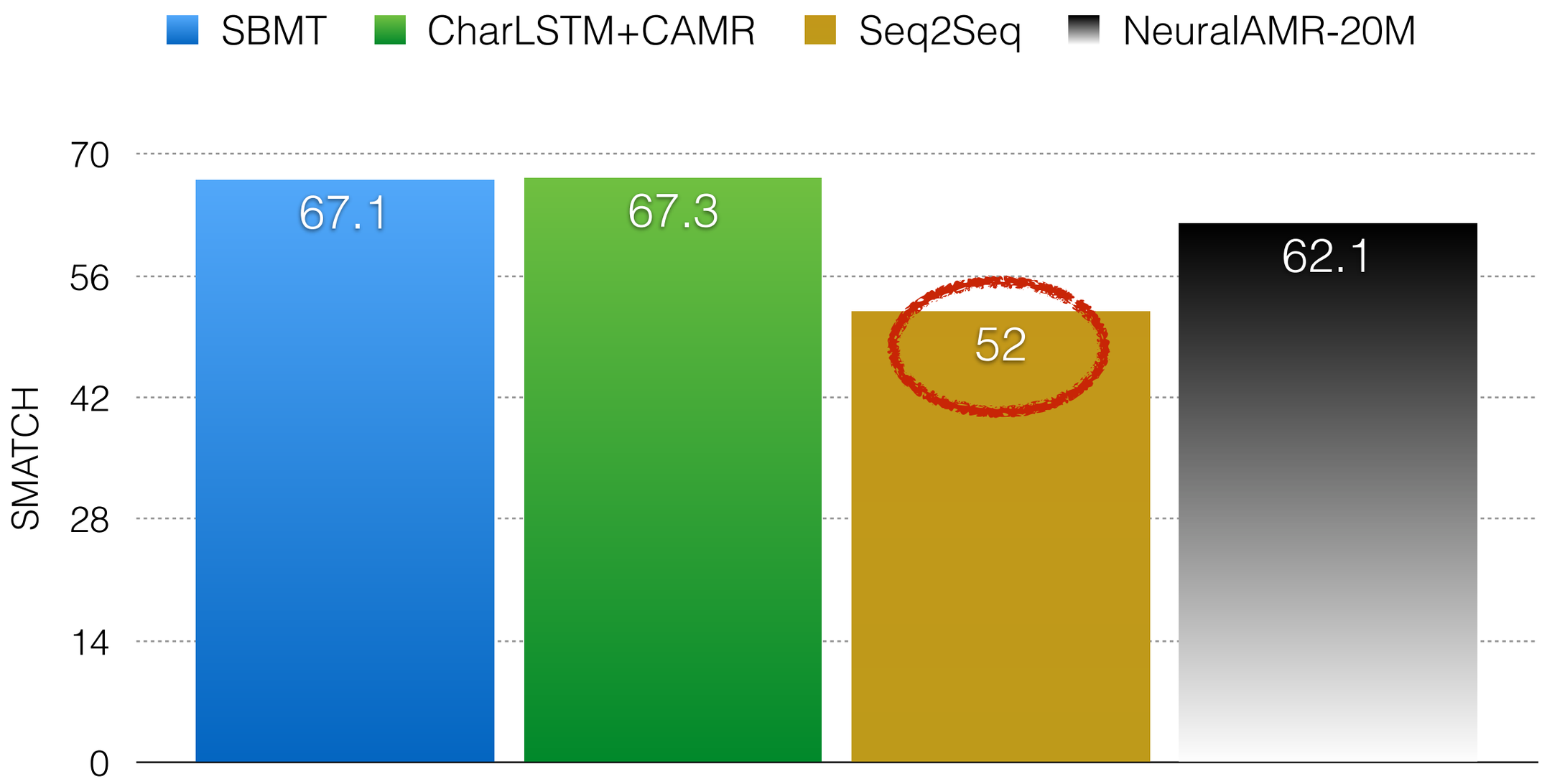
SBMT: Pust et al, 2015
CharLSTM+CAMR: Noord and Bos, 2017
Seq2Seq: Peng et al., 2017

Final Results (**Parsing**)



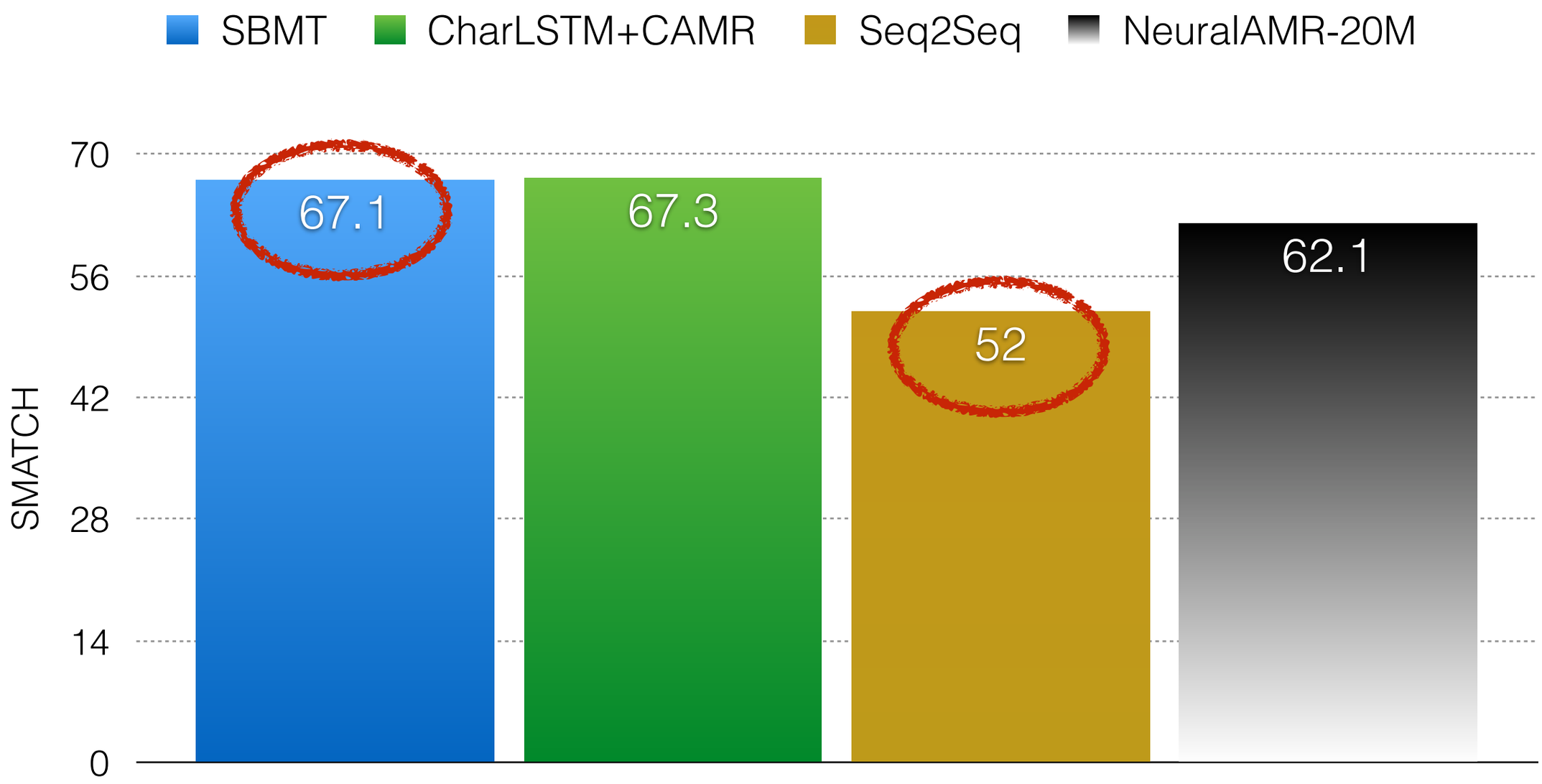
SBMT: Pust et al, 2015
CharLSTM+CAMR: Noord and Bos, 2017
Seq2Seq: Peng et al., 2017

Final Results (**Parsing**)



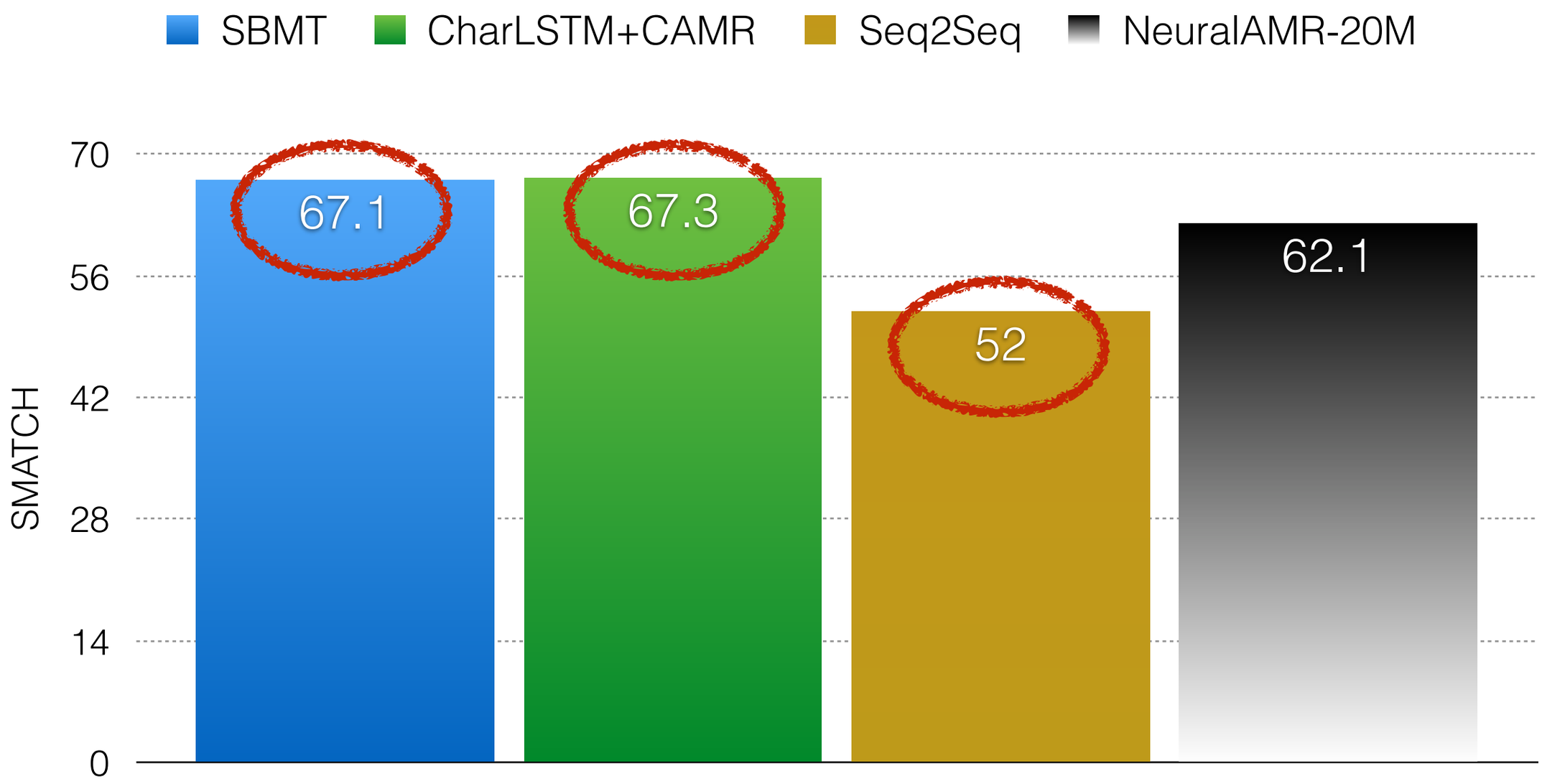
SBMT: Pust et al, 2015
CharLSTM+CAMR: Noord and Bos, 2017
Seq2Seq: Peng et al., 2017

Final Results (**Parsing**)



SBMT: Pust et al, 2015
CharLSTM+CAMR: Noord and Bos, 2017
Seq2Seq: Peng et al., 2017

Final Results (**Parsing**)



SBMT: Pust et al, 2015
CharLSTM+CAMR: Noord and Bos, 2017
Seq2Seq: Peng et al., 2017

How did we do? (**Generation**)

```
hold
   :ARG0 (person
            :ARG0-of (have-role
               :ARG1 loc_0
               :ARG2 official)
         )
   :ARG1 (meet
            :ARG0 (person
               :ARG1-of expert
               :ARG2-of group)
         )
   :time (date-entity year_0 month_0)
   :location loc_1
```

Reference

US officials held an expert group meeting **in January 2002** in New York .

Prediction

In January 2002 United States officials held a meeting of the group experts in New York .

Errors: Disfluency Coverage

How did we do? (**Generation**)

```
hold
   :ARG0 (person
            :ARG0-of (have-role
               :ARG1 loc_0
               :ARG2 official)
         )
   :ARG1 (meet
            :ARG0 (person
               :ARG1-of expert
               :ARG2-of group)
         )
   :time (date-entity year_0 month_0)
   :location loc_1
```

Reference

US officials held an expert group meeting **in January 2002** in New York .

Prediction

In January 2002 United States officials held a meeting of the group experts in New York .

Reference

The report stated **British government** must help to stabilize **weak states** and push for international regulations that would stop **terrorists** using freely available information to create and unleash new forms of biological warfare such as **a modified** version of the influenza **virus**.

Prediction

The report stated that the **Britain government** must help stabilize **the weak states** and push international regulations to stop the use of freely available information to create a form of new biological warfare such as **the modified** version of the influenza .

Errors: **Disfluency** **Coverage**

Summary

- Sequence-to-sequence models for **Parsing** and **Generation**
- **Paired Training**: scalable data augmentation algorithm
- Achieve **state-of-the-art** performance on **generating** from AMR
- Best-performing **Neural** AMR **Parser**
- Demo, Code and Pre-trained Models: http://ikonstas.net

Summary

- Sequence-to-sequence models for **Parsing** and **Generation**
- **Paired Training**: scalable data augmentation algorithm
- Achieve **state-of-the-art** performance on **generating** from AMR
- Best-performing **Neural** AMR **Parser**
- Demo, Code and Pre-trained Models: http://ikonstas.net

Thank You

Bonus Slides

Encoding

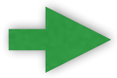
Linearize —> RNN encoding

```
hold
   :ARG0 (person
            :ARG0-of (have-role
               :ARG1 United_States
               :ARG2 official)
         )
   :ARG1 (meet
            :ARG0 (person
               :ARG1-of expert
               :ARG2-of group)
         )
   :time (date-entity 2002 1)
   :location New_York
```

Encoding

Linearize —> RNN encoding

```
hold
   :ARG0 (person
            :ARG0-of (have-role
               :ARG1 United_States
               :ARG2 official)
         )
   :ARG1 (meet
            :ARG0 (person
               :ARG1-of expert
               :ARG2-of group)
         )
   :time (date-entity 2002 1)
   :location New_York
```

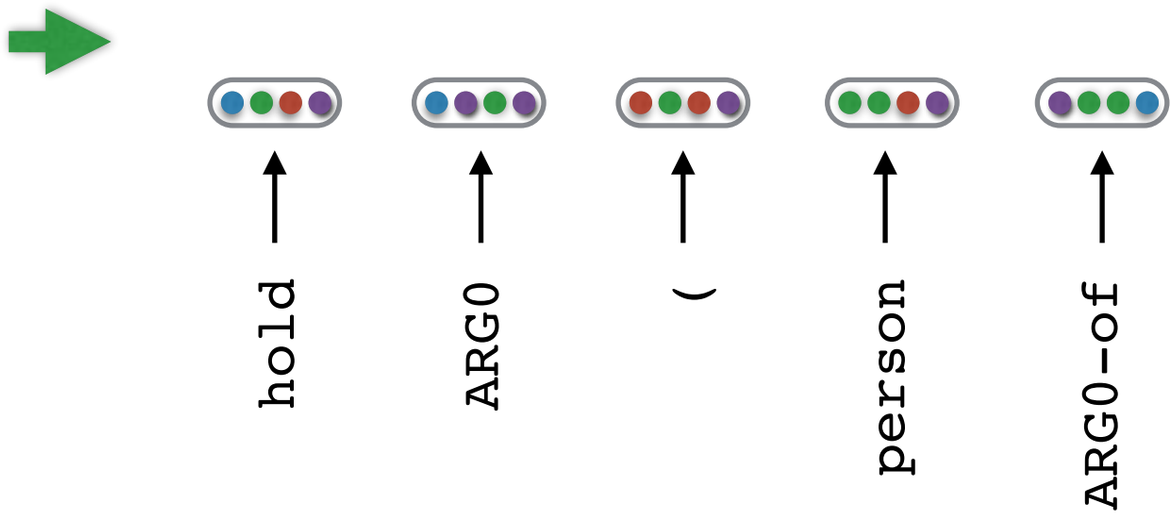


Encoding

Linearize —> RNN encoding

- Token embeddings

```
hold
   :ARG0 (person
            :ARG0-of (have-role
               :ARG1 United_States
               :ARG2 official)
         )
   :ARG1 (meet
            :ARG0 (person
               :ARG1-of expert
               :ARG2-of group)
         )
   :time (date-entity 2002 1)
   :location New_York
```

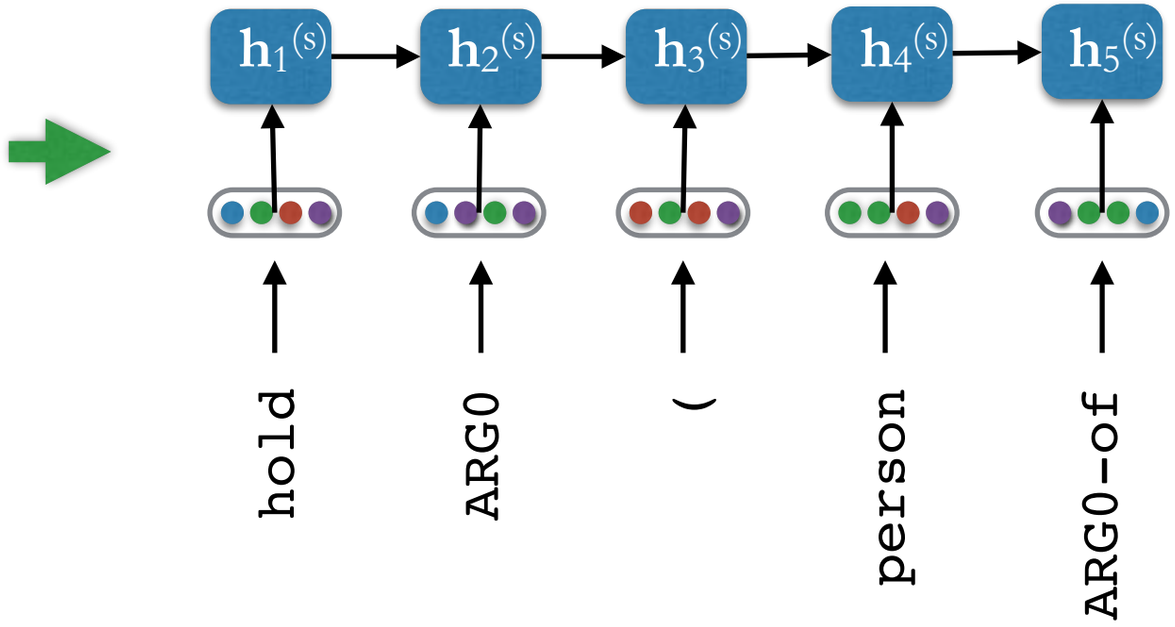


Encoding

Linearize —> RNN encoding

- Token embeddings
- Recurrent Neural Network (RNN)

```
hold
   :ARG0 (person
            :ARG0-of (have-role
               :ARG1 United_States
               :ARG2 official)
         )
   :ARG1 (meet
            :ARG0 (person
               :ARG1-of expert
               :ARG2-of group)
         )
   :time (date-entity 2002 1)
   :location New_York
```

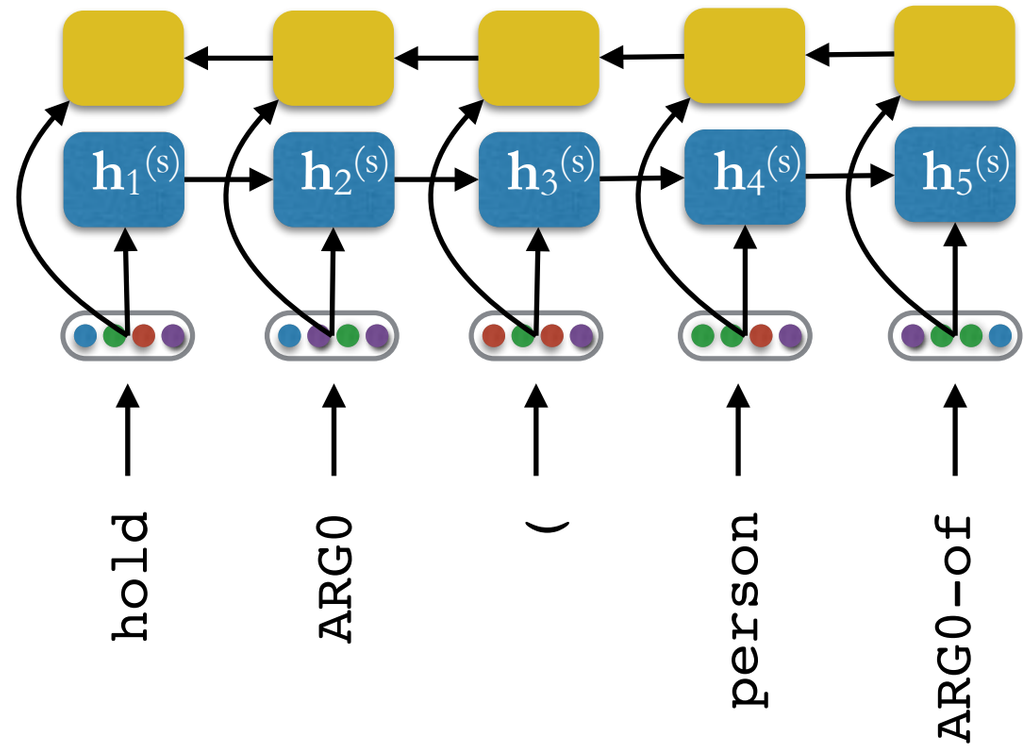


Encoding

Linearize —> RNN encoding

- Token embeddings
- Recurrent Neural Network (RNN)
- Bi-directional RNN

```
hold
   :ARG0 (person
            :ARG0-of (have-role
               :ARG1 United_States
               :ARG2 official)
         )
   :ARG1 (meet
            :ARG0 (person
               :ARG1-of expert
               :ARG2-of group)
         )
   :time (date-entity 2002 1)
   :location New_York
```

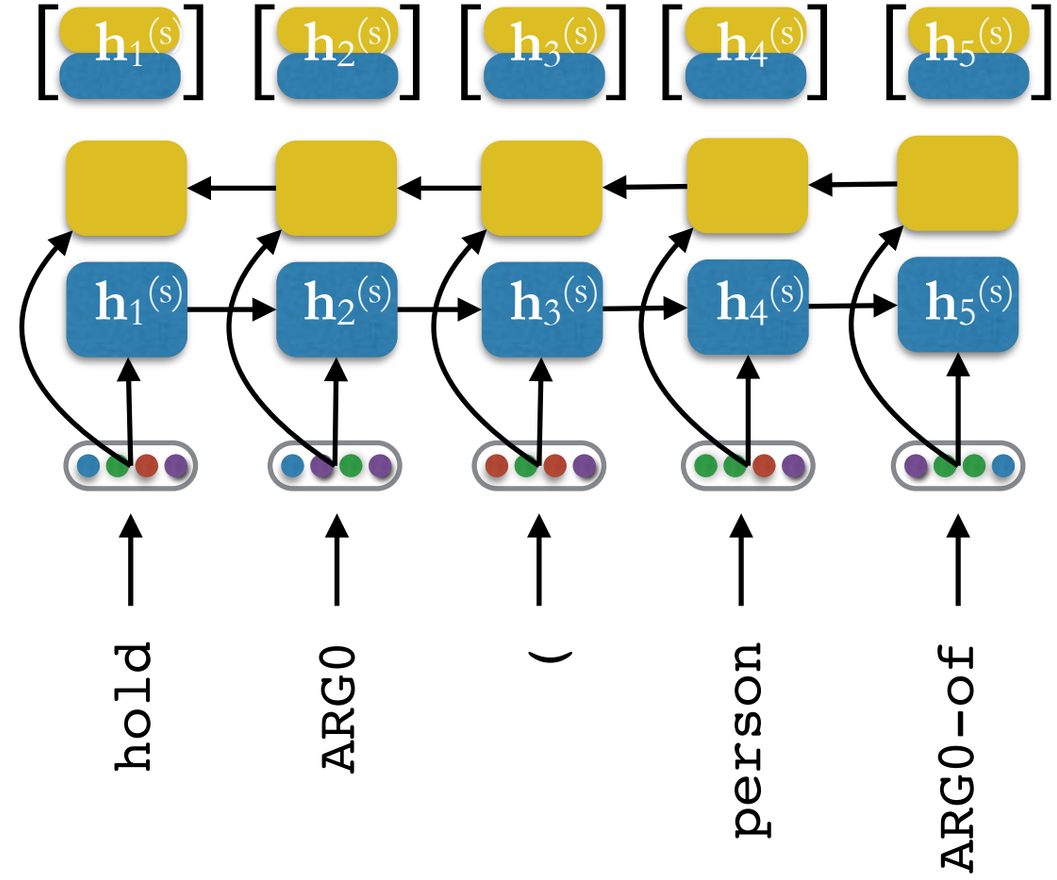


Encoding

Linearize —> RNN encoding

- Token embeddings
- Recurrent Neural Network (RNN)
- Bi-directional RNN

```
hold
   :ARG0 (person
            :ARG0-of (have-role
               :ARG1 United_States
               :ARG2 official)
         )
   :ARG1 (meet
            :ARG0 (person
               :ARG1-of expert
               :ARG2-of group)
         )
   :time (date-entity 2002 1)
   :location New_York
```

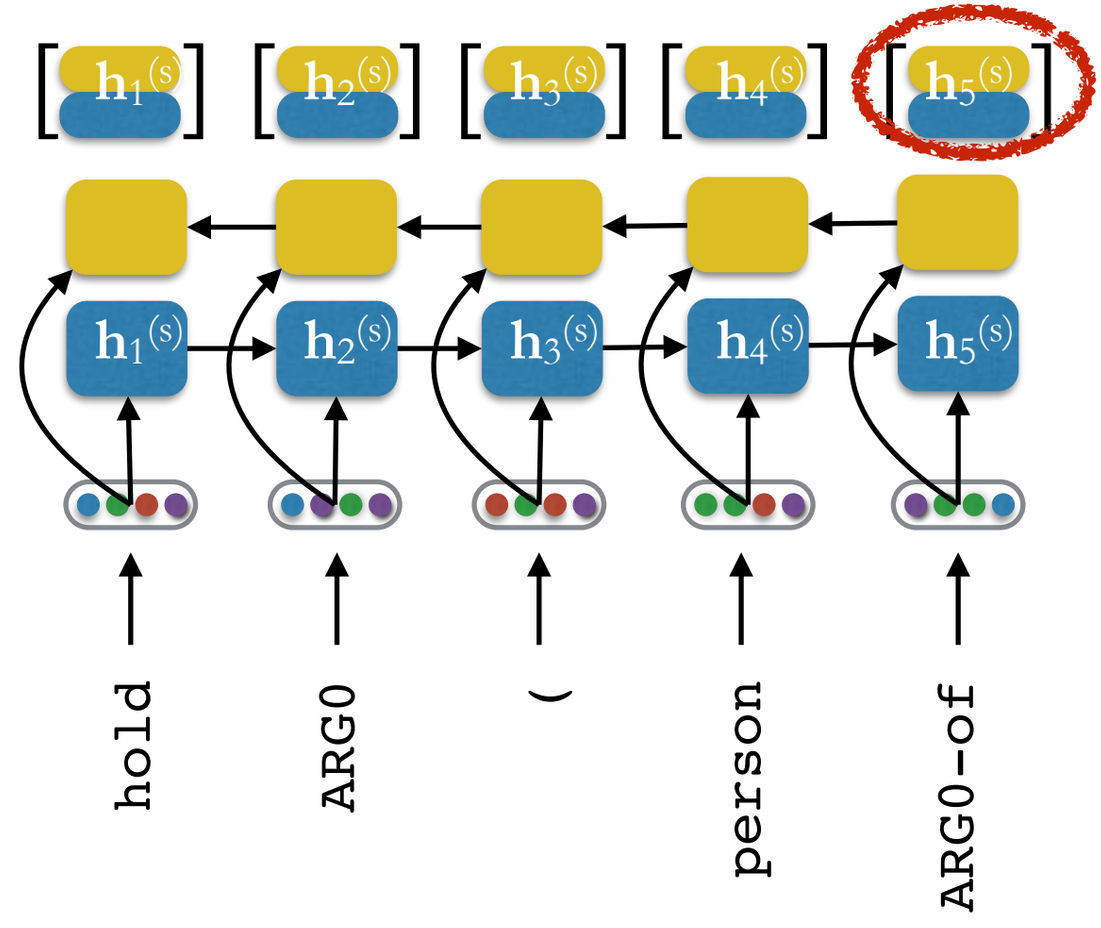


Encoding

Linearize —> RNN encoding

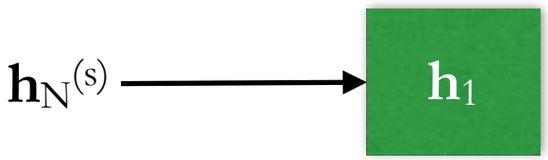
- Token embeddings
- Recurrent Neural Network (RNN)
- Bi-directional RNN

```
hold
    :ARG0 (person
              :ARG0-of (have-role
                  :ARG1 United_States
                  :ARG2 official)
          )
    :ARG1 (meet
              :ARG0 (person
                  :ARG1-of expert
                  :ARG2-of group)
          )
    :time (date-entity 2002 1)
    :location New_York
```



Decoding

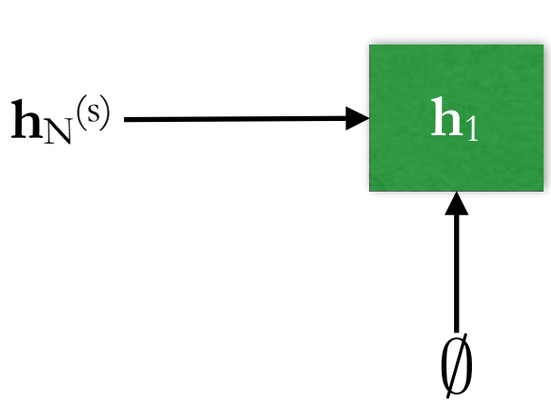
RNN Encoding —> RNN Decoding (Beam search)



Decoding

RNN Encoding —> RNN Decoding (Beam search)

- init $\mathbf{h}^{(s)}$

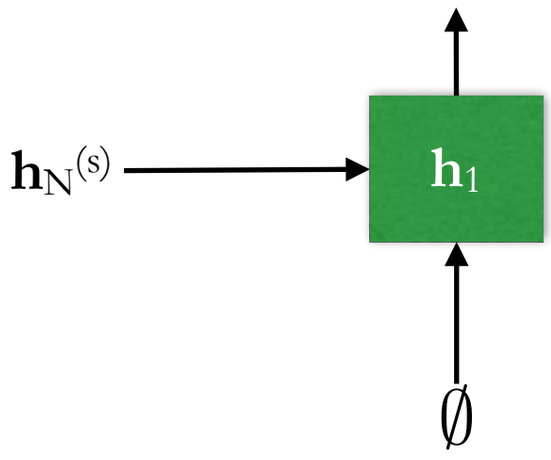


Decoding

RNN Encoding —> RNN Decoding (Beam search)

- init $\mathbf{h}^{(s)}$
- softmax

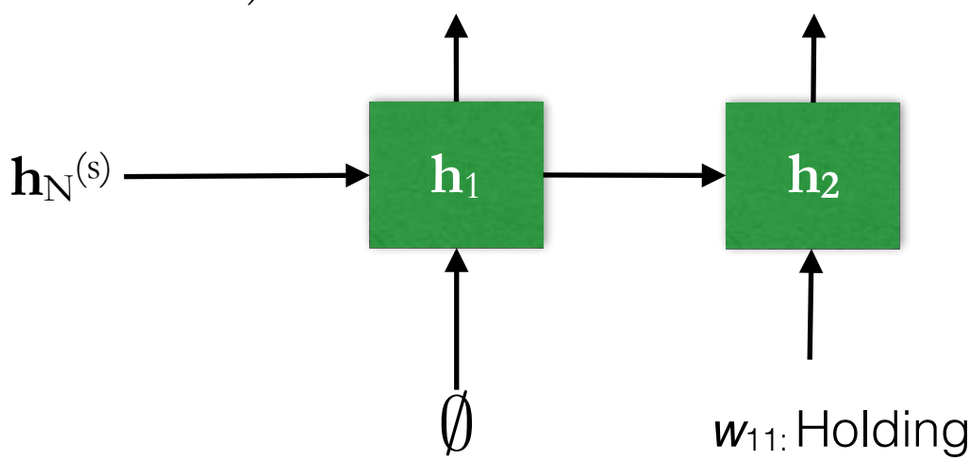
Holding
Held
US
…



Decoding

RNN Encoding —> RNN Decoding (Beam search)

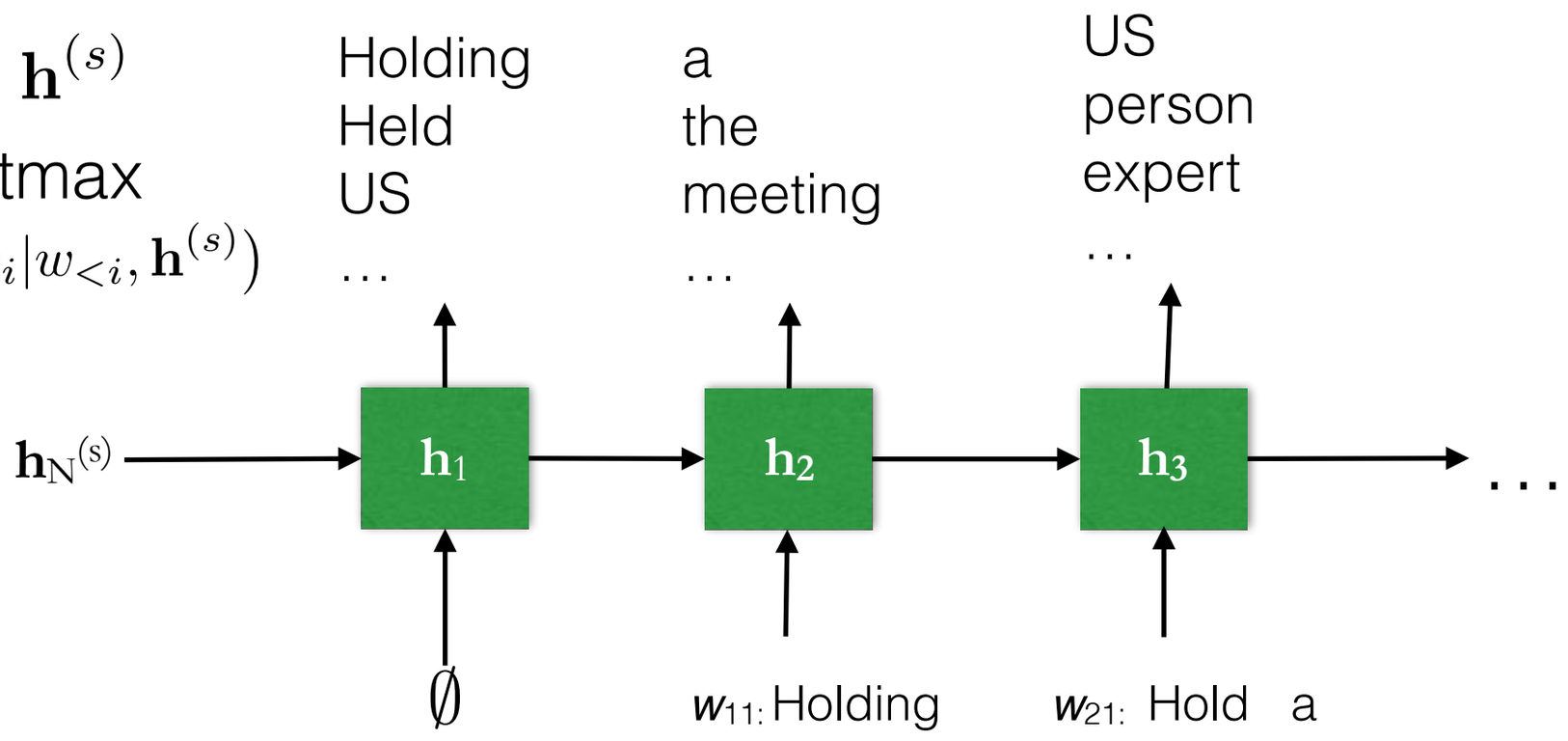
- init $\mathbf{h}^{(s)}$
- softmax
- $p(w_i|w_{<i}, \mathbf{h}^{(s)})$



Decoding

RNN Encoding —> RNN Decoding (Beam search)

- init $\mathbf{h}^{(s)}$
- softmax
- $p(w_i|w_{<i}, \mathbf{h}^{(s)})$



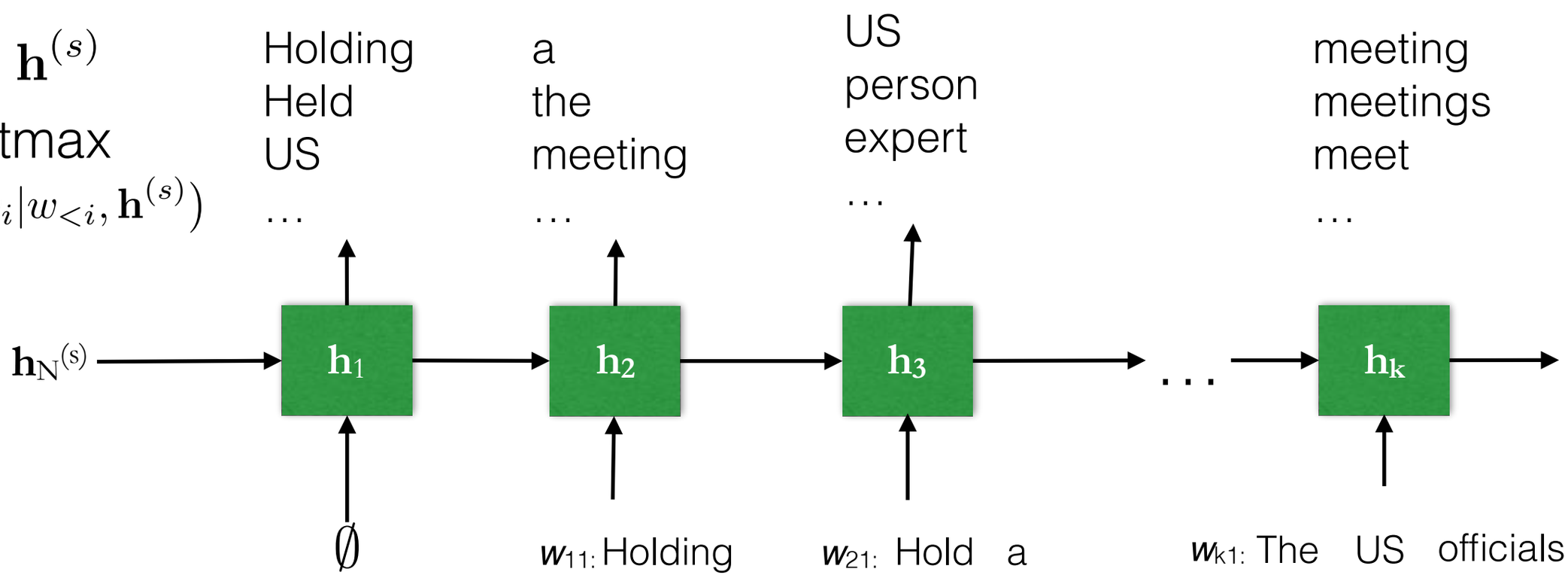
w_{11}: Holding
w_{12}: Helds
w_{13}: Hold
w_{14}: US
…

w_{21}: Hold a
w_{22}: Hold the
w_{23}: Held a
w_{24}: Held the
…

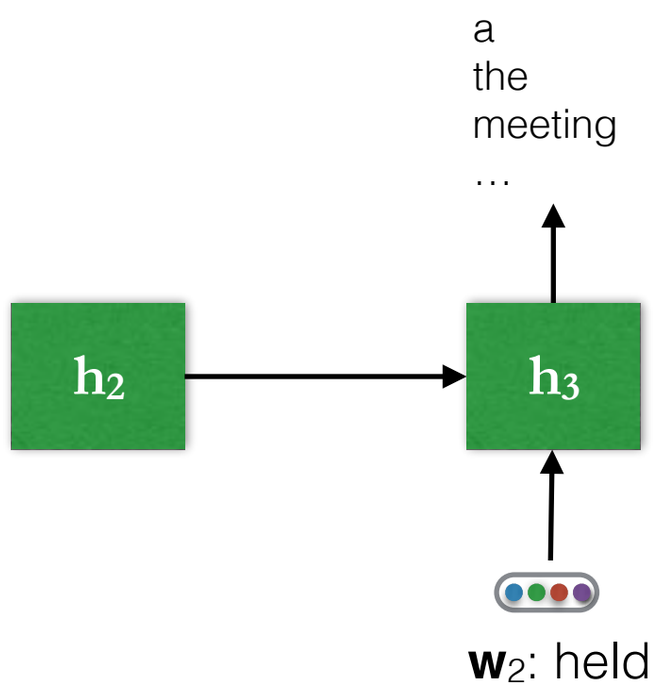
Decoding

RNN Encoding —> RNN Decoding (Beam search)

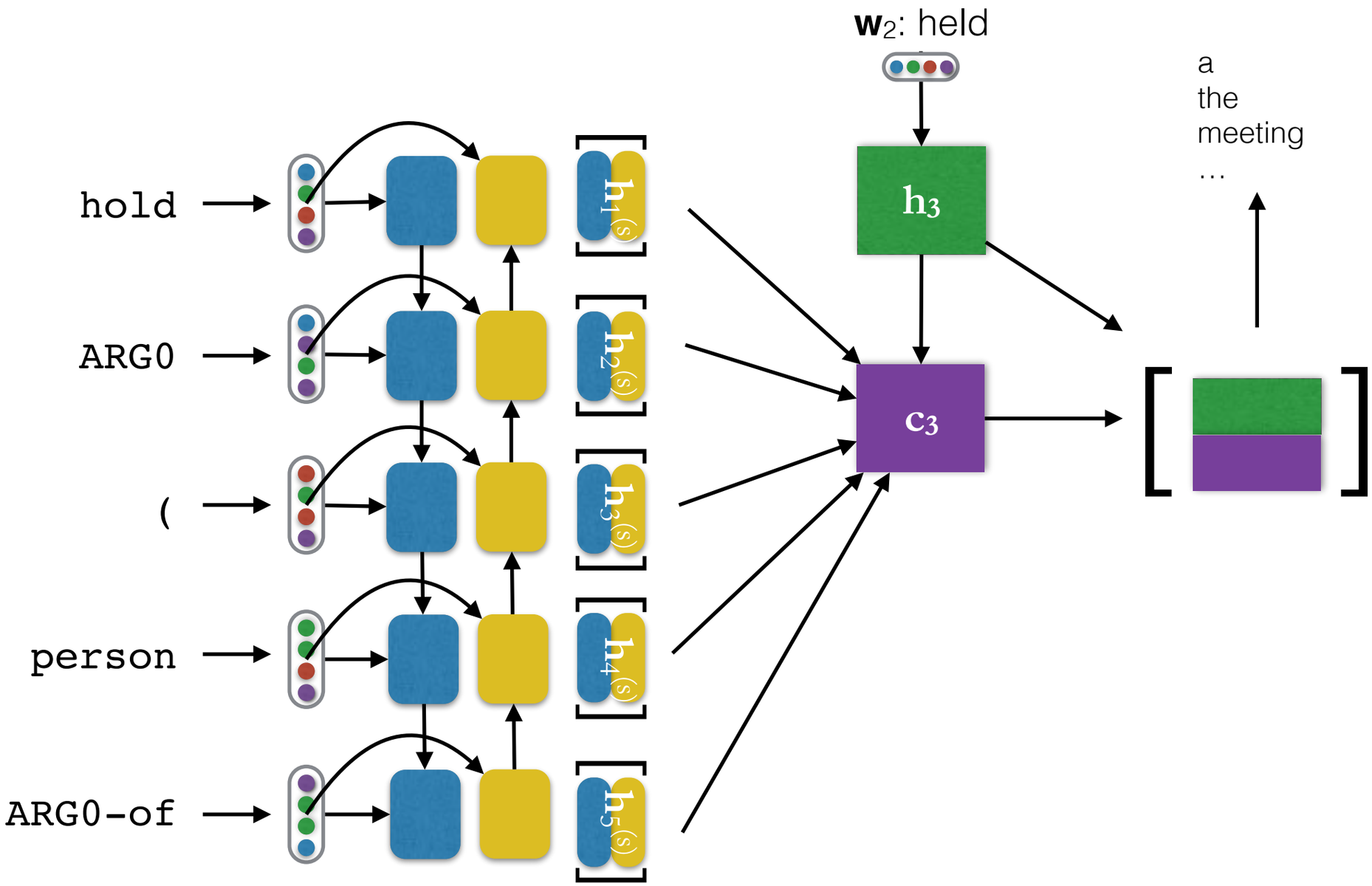
- init $\mathbf{h}^{(s)}$
- softmax
- $p(w_i|w_{<i}, \mathbf{h}^{(s)})$



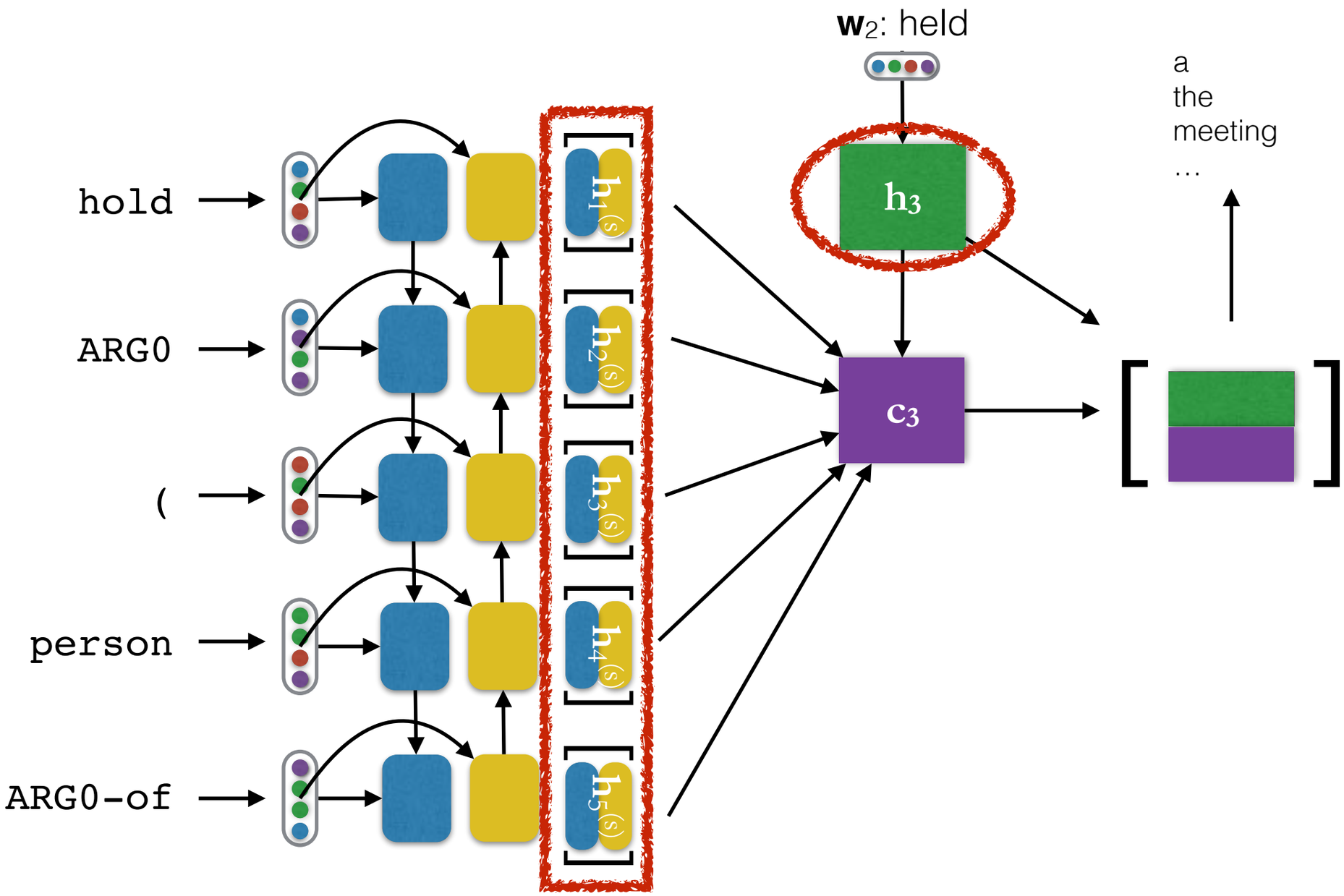
Attention



Attention


w2: held
hold
ARG0
(
person
ARG0-of
h1(s)
h2(s)
h3(s)
h4(s)
h5(s)
h3
c3
a
the
meeting
…

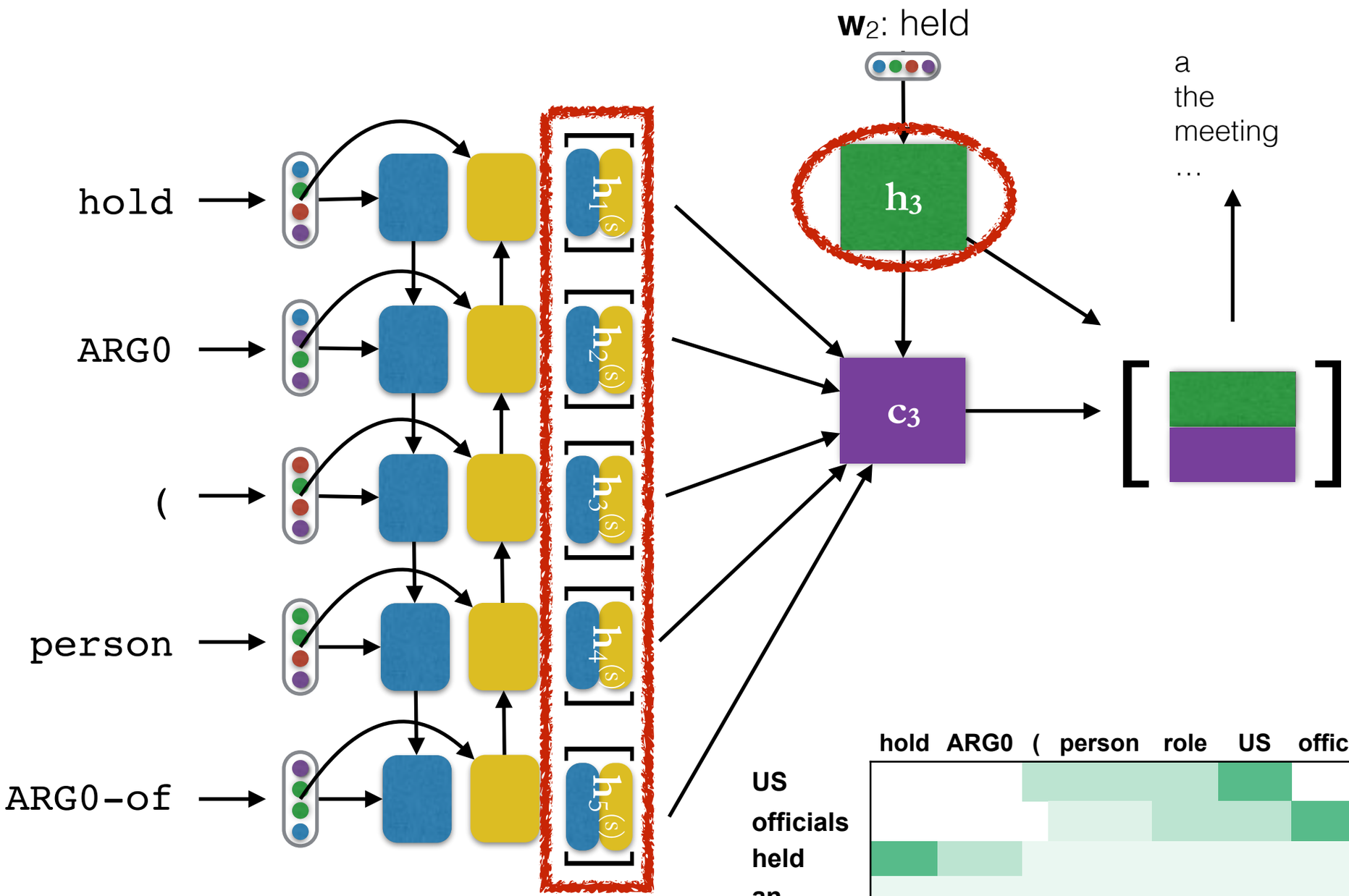
Attention



$$\mathbf{a}_i = \text{soft}\max\left(\mathbf{f}_i\left(\mathbf{h}^{(\mathbf{s})}, \mathbf{h}_i\right)\right)$$

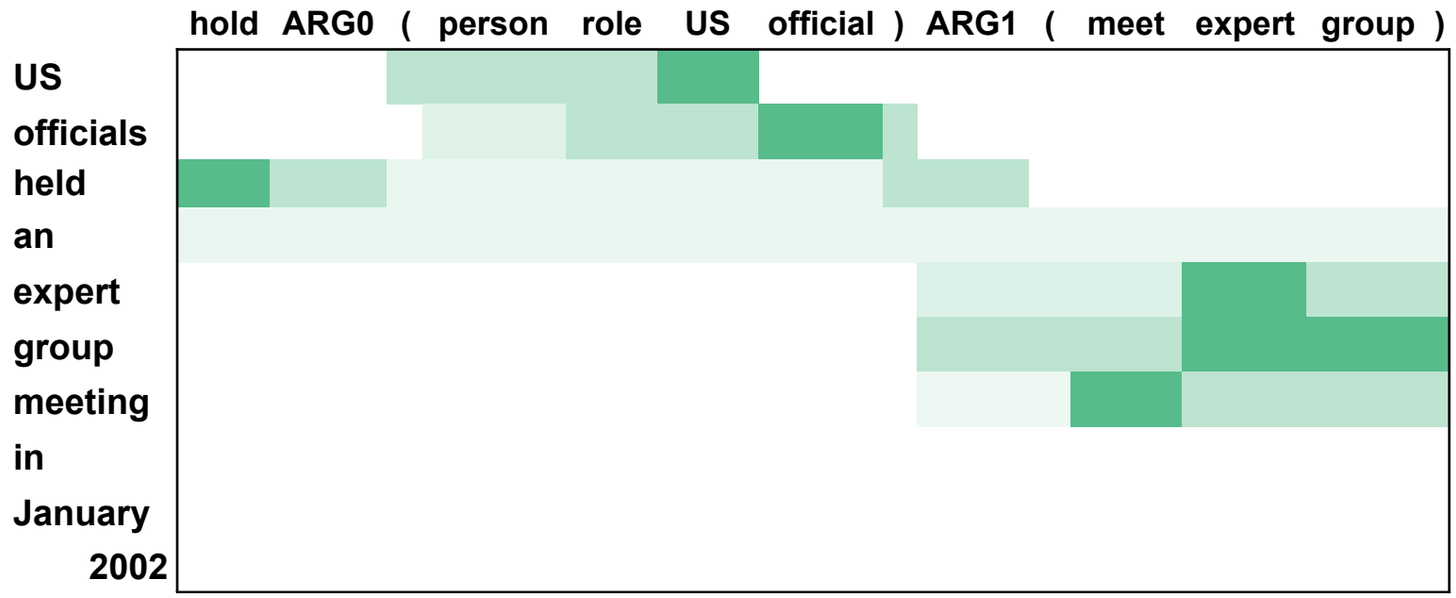
$$\mathbf{c}_i = \sum_i \mathbf{a}_{ij}\mathbf{h}_j^{(s)}$$

Attention

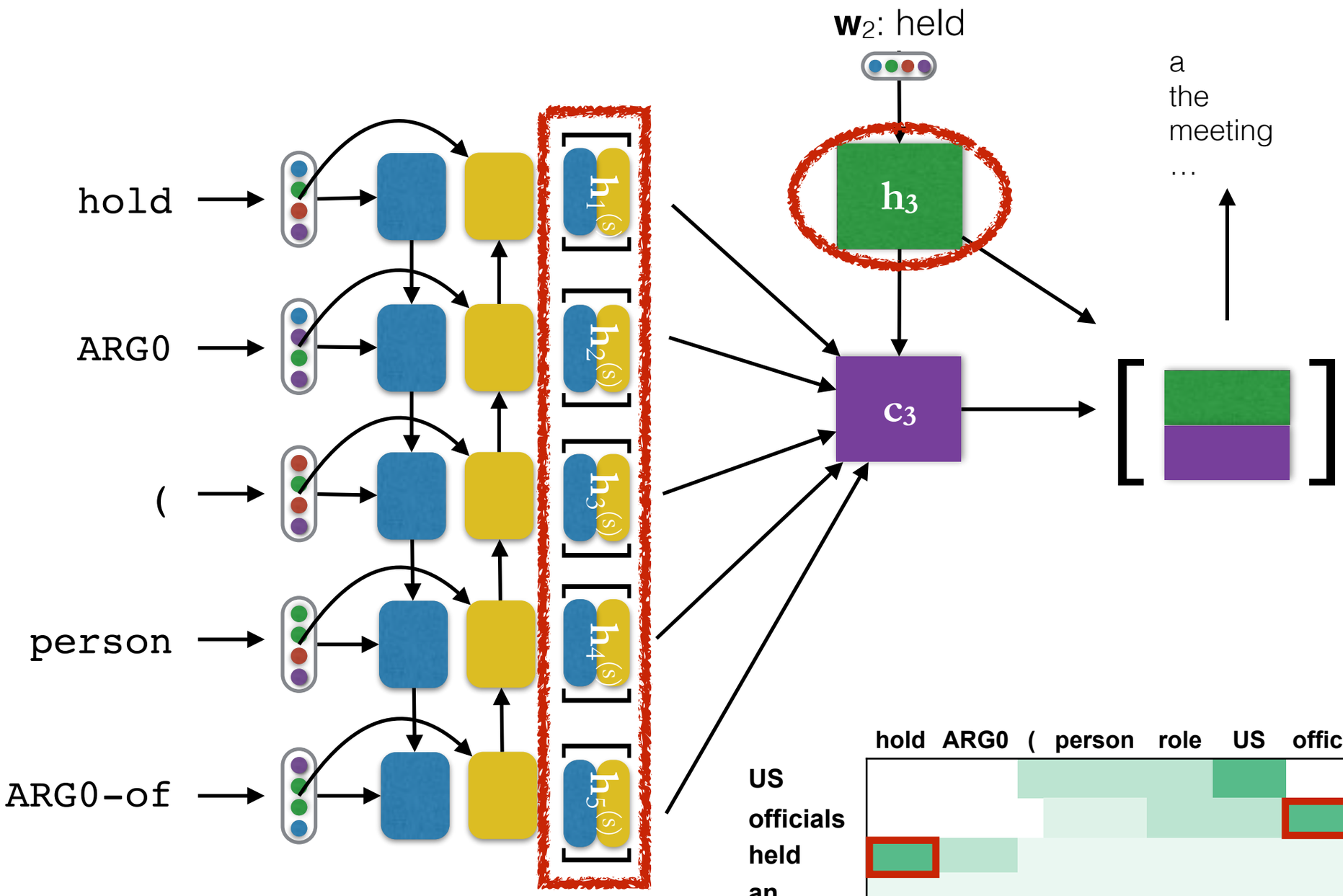


$$\mathbf{a}_i = \text{soft}\max\left(\mathbf{f}_i\left(\mathbf{h}^{(\mathbf{s})}, \mathbf{h}_i\right)\right)$$

$$\mathbf{c}_i = \sum_i \mathbf{a}_{ij}\mathbf{h}_j^{(s)}$$

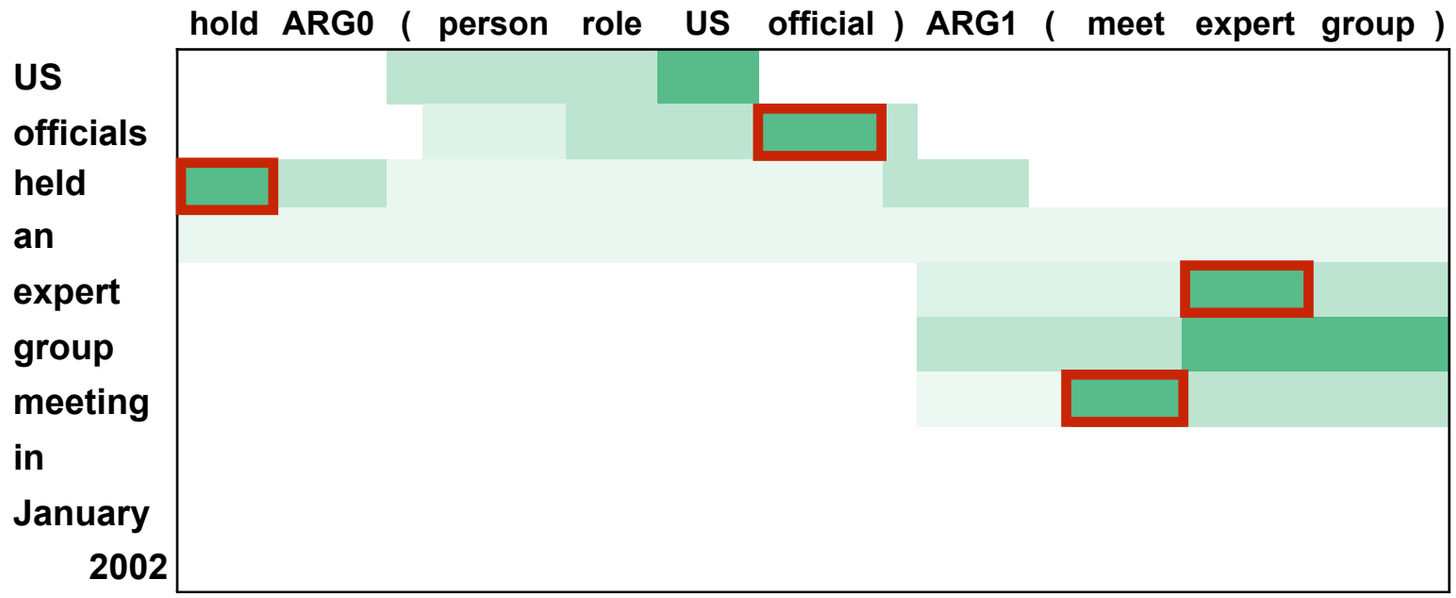


Attention

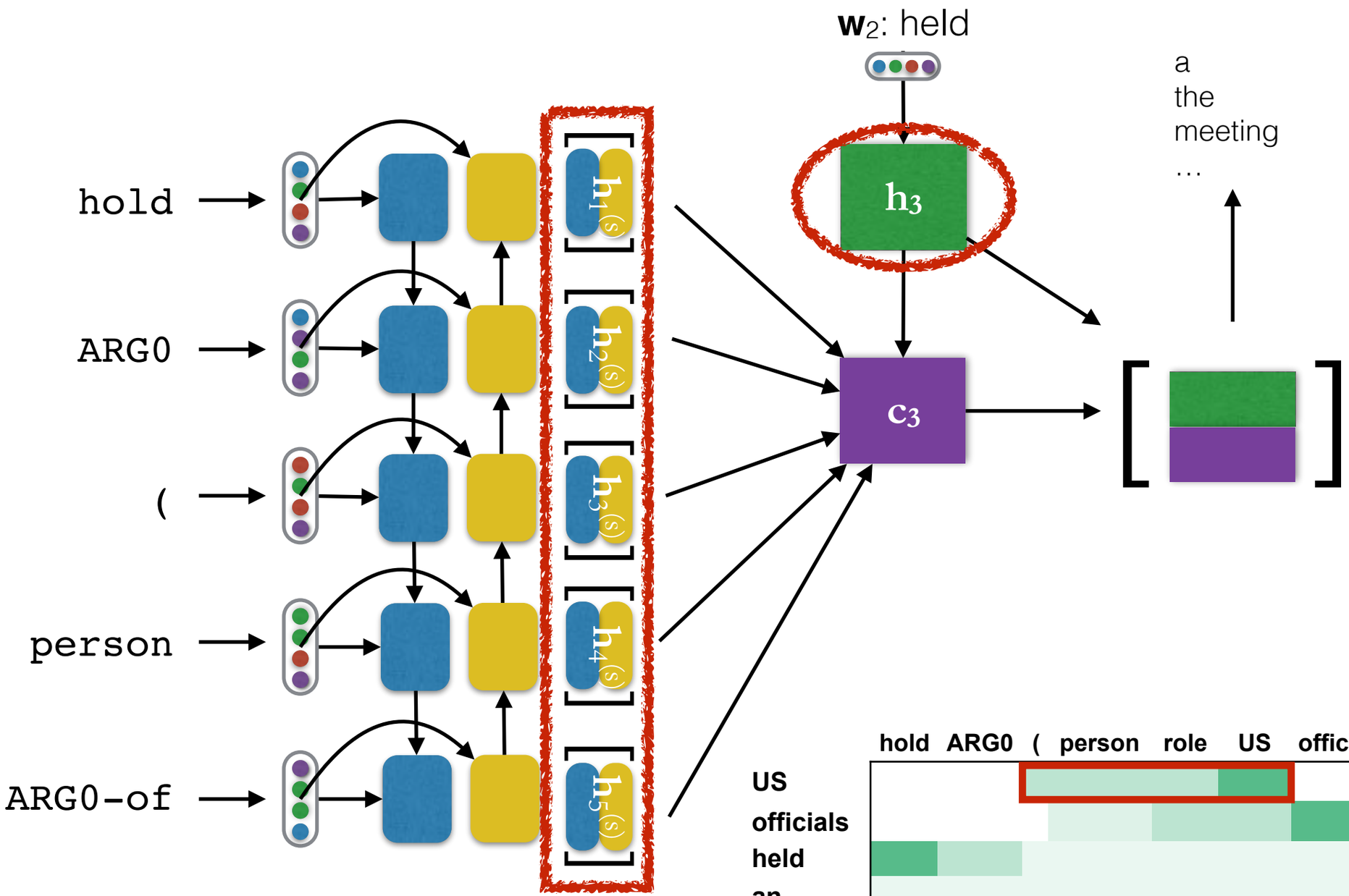


$$\mathbf{a}_i = \text{soft}\max\left(\mathbf{f}_i\left(\mathbf{h}^{(\mathbf{s})}, \mathbf{h}_i\right)\right)$$

$$\mathbf{c}_i = \sum_i \mathbf{a}_{ij}\mathbf{h}_j^{(s)}$$

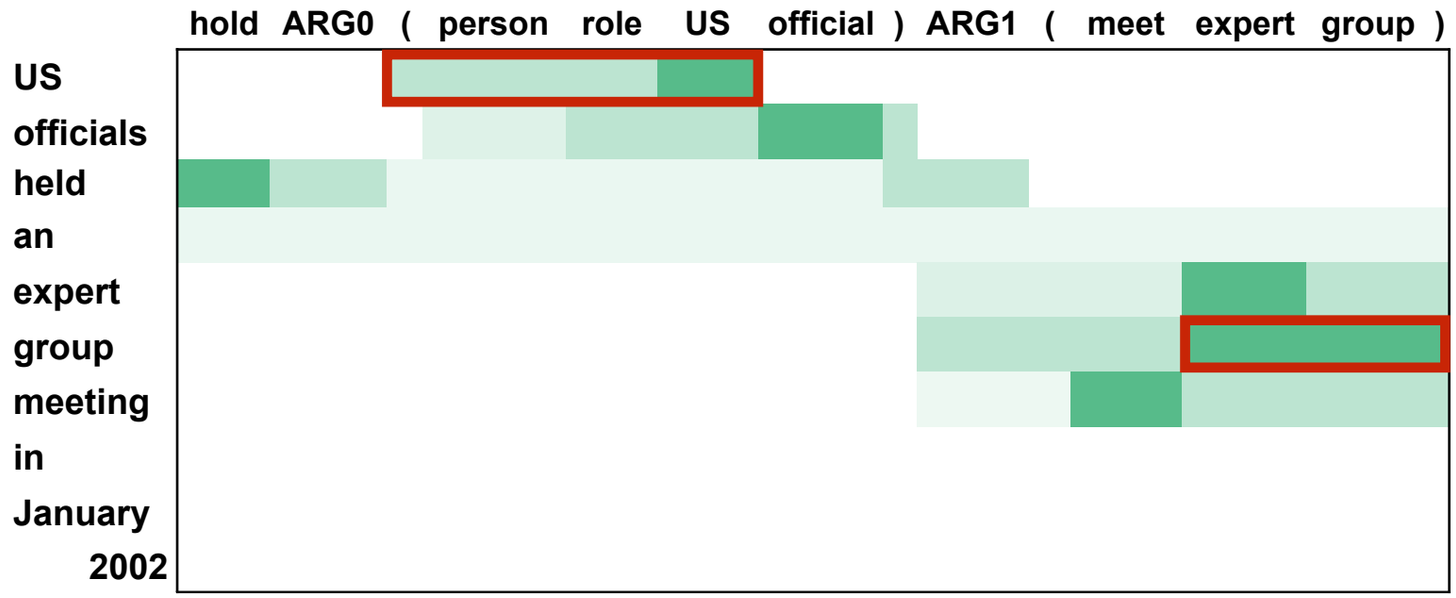


Attention

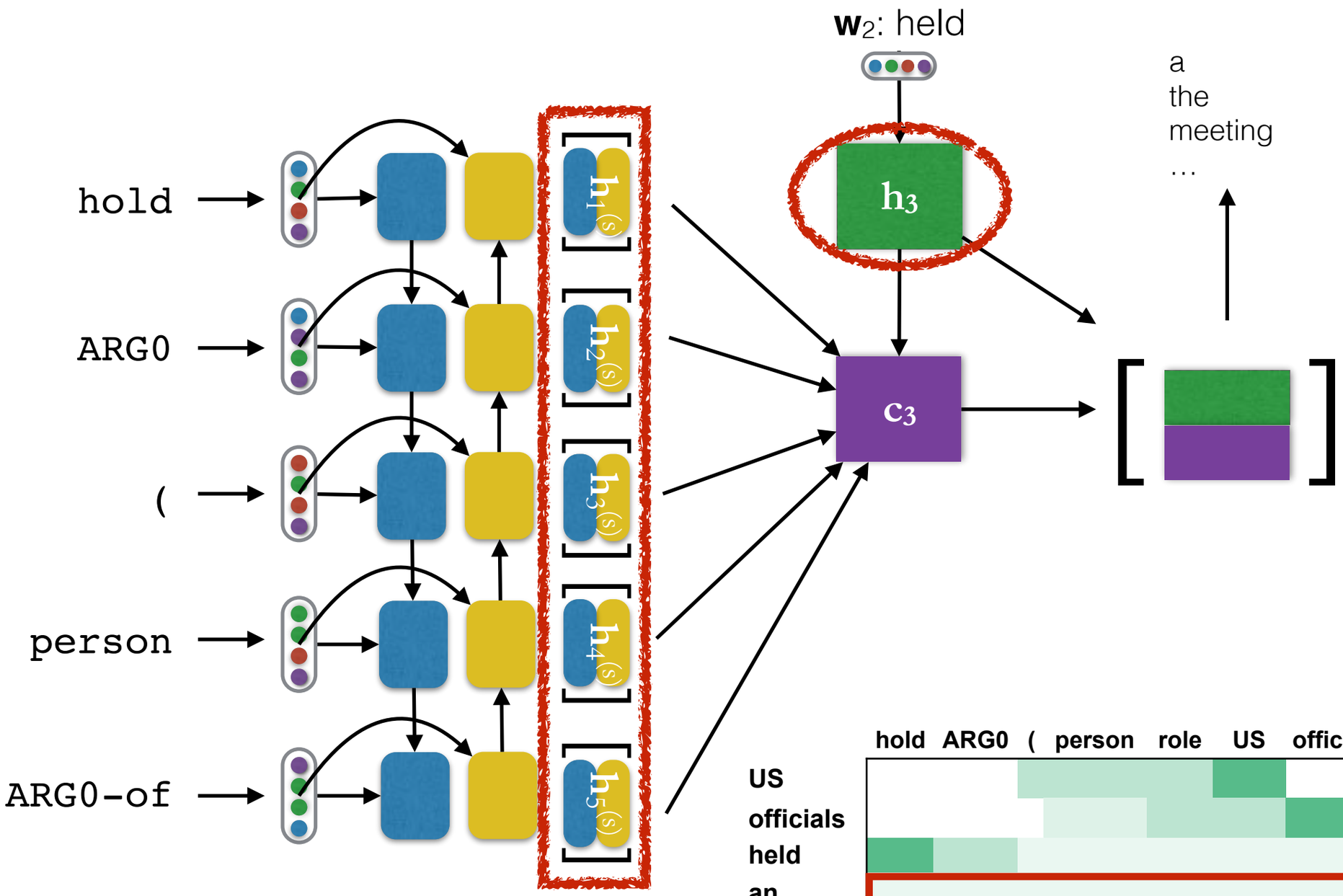


$$\mathbf{a}_i = \text{soft}\max\left(\mathbf{f}_i\left(\mathbf{h}^{(\mathbf{s})}, \mathbf{h}_i\right)\right)$$

$$\mathbf{c}_i = \sum_i \mathbf{a}_{ij}\mathbf{h}_j^{(s)}$$



Attention



$$\mathbf{a}_i = \text{soft}\max\left(\mathbf{f}_i\left(\mathbf{h}^{(\mathbf{s})}, \mathbf{h}_i\right)\right)$$

$$\mathbf{c}_i = \sum_i \mathbf{a}_{ij}\mathbf{h}_j^{(s)}$$