UNIVERSITY of WASHINGTON

UWNLP

Building Adaptable and Scalable **Natural Language Generation** Systems

Yannis Konstas

Natural Language Generation is everywhere
(Machine Translation)

Input: Ο πρόεδρος των ΗΠΑ Ντόναλντ Τραμπ γνωστοποίησε ότι δεν θα πάει στο ετήσιο δείπνο της Ένωσης Ανταποκριτών Λευκού Οίκου (WHCA) στα τέλη του Απριλίου.

Human: The president of the United States Donald Trump announced that **he would not go** to the annual dinner of the White House Correspondents' Association (WHCA) in late April.

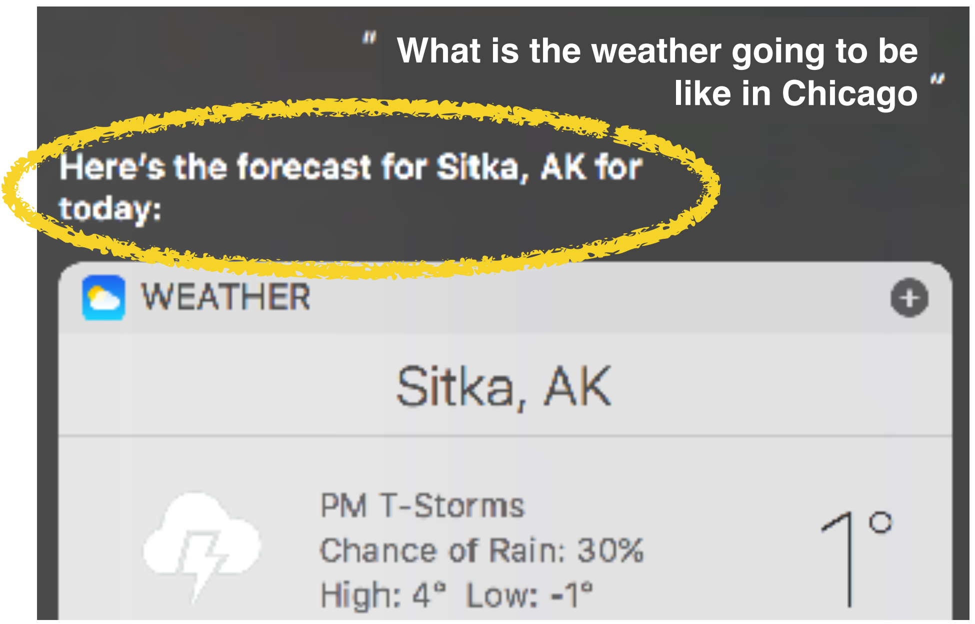
Natural Language Generation is everywhere
(Machine Translation)

Input: Ο πρόεδρος των ΗΠΑ Ντόναλντ Τραμπ γνωστοποίησε ότι δεν θα πάει στο ετήσιο δείπνο της Ένωσης Ανταποκριτών Λευκού Οίκου (WHCA) στα τέλη του Απριλίου.

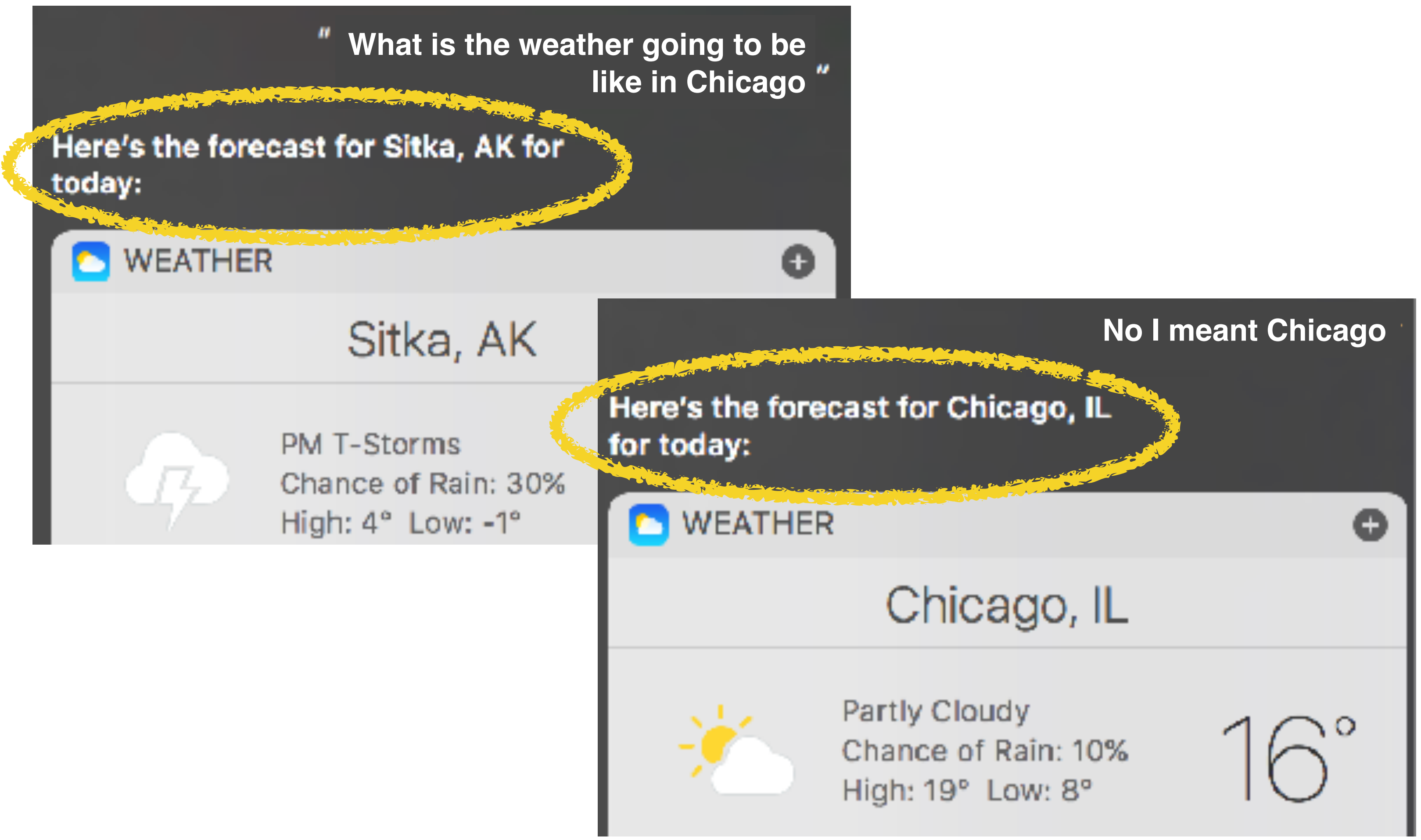
Human: The president of the United States Donald Trump announced that **he would not go** to the annual dinner of the White House Correspondents' Association (WHCA) in late April.

Google Translate: The US president Donald Trump announced that **it would go** to the annual dinner of White House Correspondents Union (WHCA) in late April.

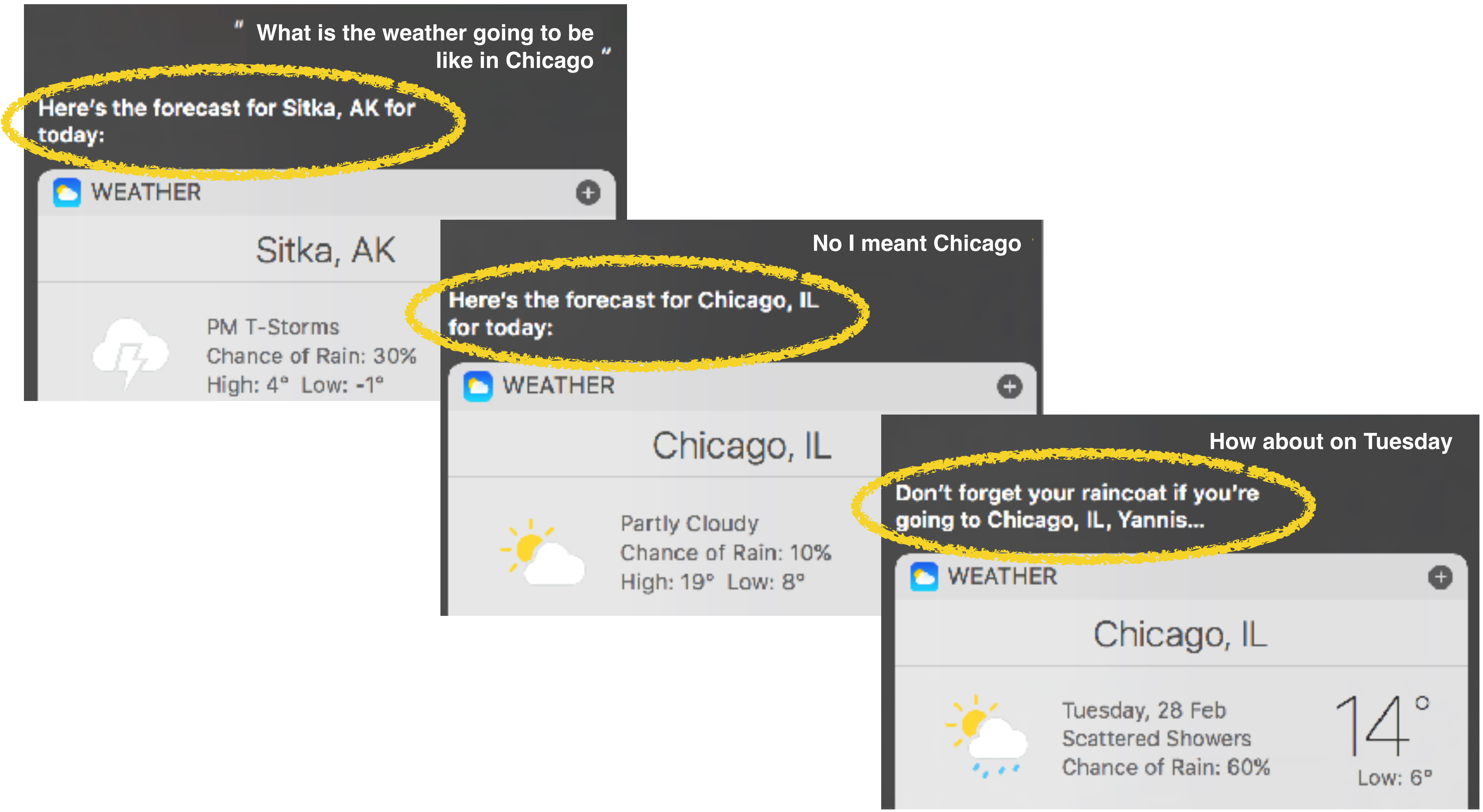
Natural Language Generation is everywhere
(Dialogue Systems)



Natural Language Generation is everywhere
(Dialogue Systems)

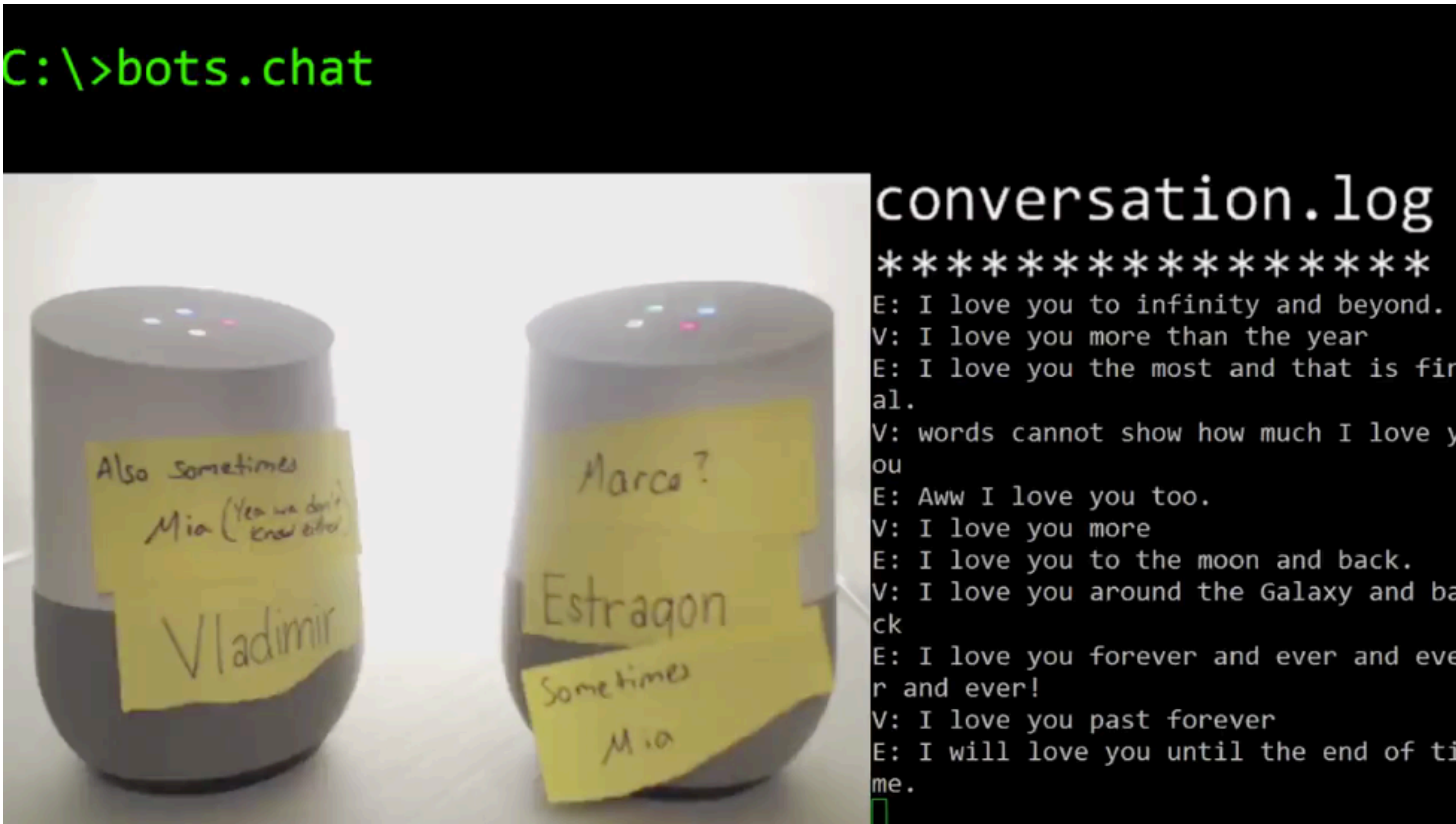


Natural Language Generation is everywhere
(Dialogue Systems)

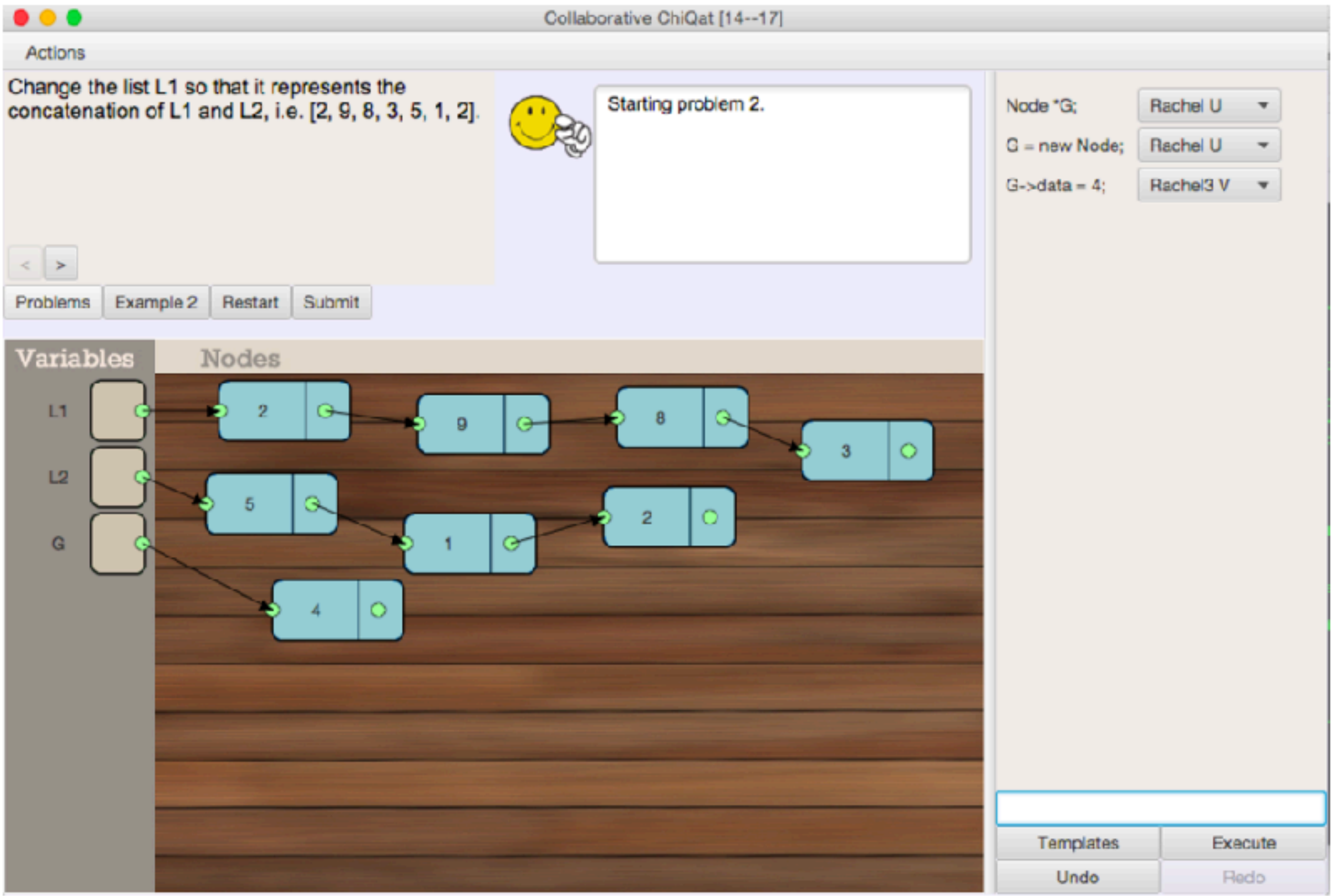


Natural Language Generation is everywhere
(Conversational Agents)

…or when things get too emotional



Natural Language Generation is everywhere
(Educational Technology)



(Harsley et al., CSCW 2016)

Natural Language Generation is everywhere
(Caption Generation)

A man swinging a bat.

(Krause et al., CVPR 2017)

Natural Language Generation is everywhere (Caption Generation)

A man swinging a bat.

A baseball player is **swinging** a bat.

He is **wearing** a red helmet and a white shirt.

The catcher's mitt **is behind** the batter.

(Krause et al., CVPR 2017)

Machine Translation

Text Summarization

Concept-to-Text

Human-Robot Interaction

Code to Language

Dialogue Systems

Instructional Text

Conversational Agents

Meaning Representations

Storytelling

Educational Technology

Captions

Machine Translation

Concept-to-Text

Text Summarization

Human-Robot Interaction

Code to Language

Dialogue Systems

Instructional Text

Meaning Representations

Conversational Agents

Storytelling

Educational Technology

Captions

Machine Translation

Concept-to-Text

Text Summarization

Human-Robot Interaction

Code to Language

Dialogue Systems

Instructional Text

Meaning Representations

Conversational Agents

Storytelling

Educational Technology

Captions

Machine Translation

Concept-to-Text

Text Summarization

Human-Robot Interaction

Code to Language

Dialogue Systems

Instructional Text

Meaning Representations

Conversational Agents

Storytelling

Educational Technology

Captions

Machine Translation

Concept-to-Text

Text Summarization

Human-Robot Interaction

Code to Language

Dialogue Systems

Instructional Text

Meaning Representations

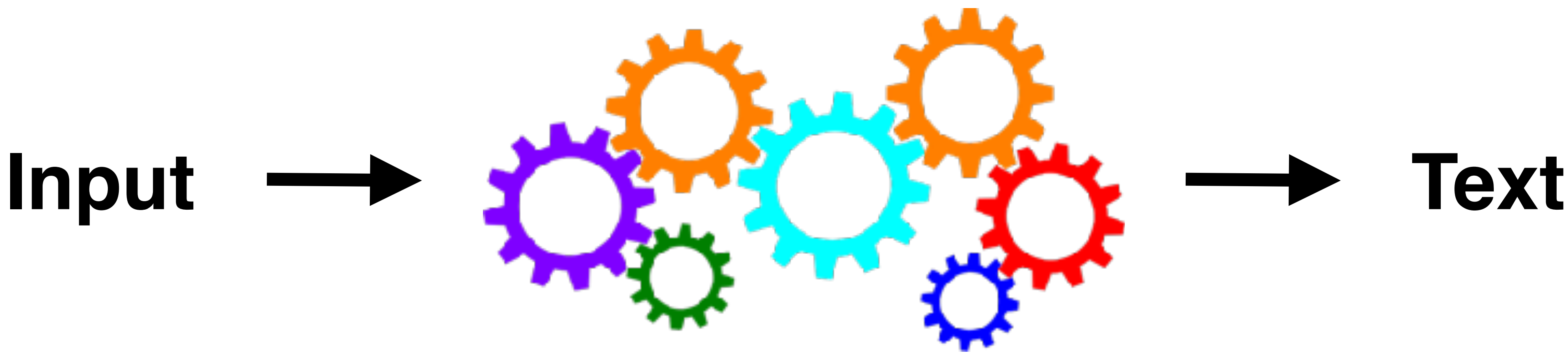
Conversational Agents

Storytelling

Educational Technology

Captions

Natural Language Generation



- Input: Computer-interpretable representation of the world
  - **Select** content
  - **Organize** content in particular order
  - Decide how to **verbalise** content
- Output: Text

Machine Translation

High quality source code is often paired with high level summaries of the computation it performs, for example in code documentation or in descriptions posted in online forums.

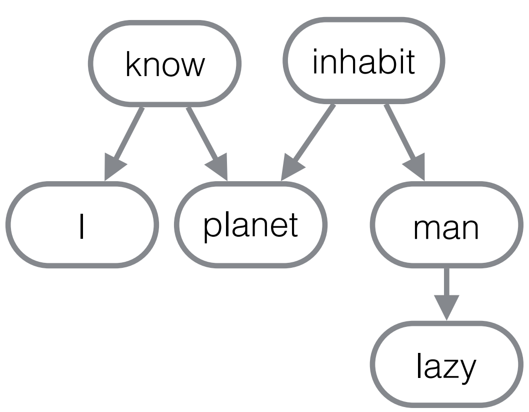
Code to Language

```
public int TextWidth (string text)
{
  TextBlock t = new TextBlock();
  t.Text = text;
  return (int)
  Math.Ceiling(t.ActualWidth);
}
```

Concept-to-Text

	min	mean	max	mod
wind	10	15	20	
dir				W
temp	50	60	72	
gust	5	10	13	

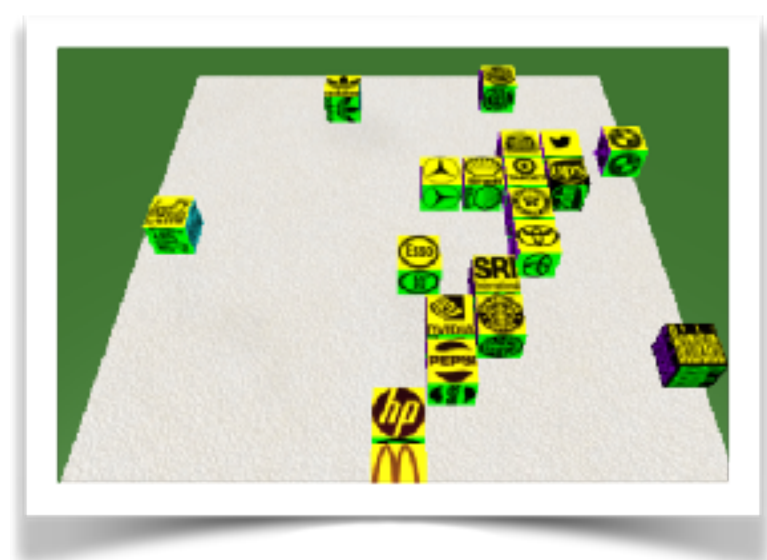
Meaning Representations



Educational Technology

$$20x + 5y = \gamma$$

Human-Robot Interaction



Machine Translation

High quality source code is often paired with high level summaries of the computation it performs, for example in code documentation or in descriptions posted in online forums.

高品質のソースコードは、コードドキュメントやオンラインフォーラムに掲載された説明など、実行する計算のハイレベルの要約と対になることがよくあります。

Code to Language

```
public int TextWidth (string text)
{
  TextBlock t = new TextBlock();
  t.Text = text;
  return (int)
  Math.Ceiling(t.ActualWidth);
}
```

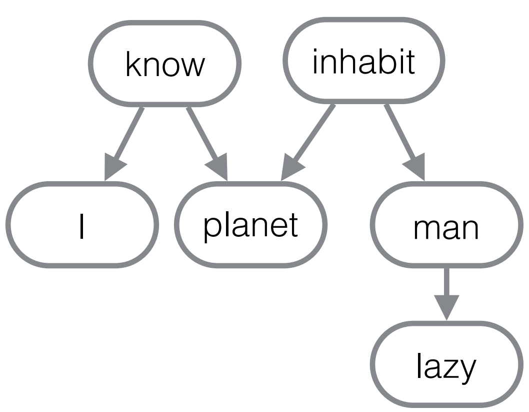
Get rendered width of string rounded up to the nearest integer.

Concept-to-Text

	min	mean	max	mod
wind	10	15	20	
dir				W
temp	50	60	72	
gust	5	10	13	

Overcast, with a high of 70. Moderate westerly winds, with gusts as high as 13 mph.

Meaning Representations



I know the planet is inhabited by a lazy man.

Educational Technology

$$20x + 5y = \gamma$$

Tammy bought 20 apples and 5 oranges. How many fruits does she have now?

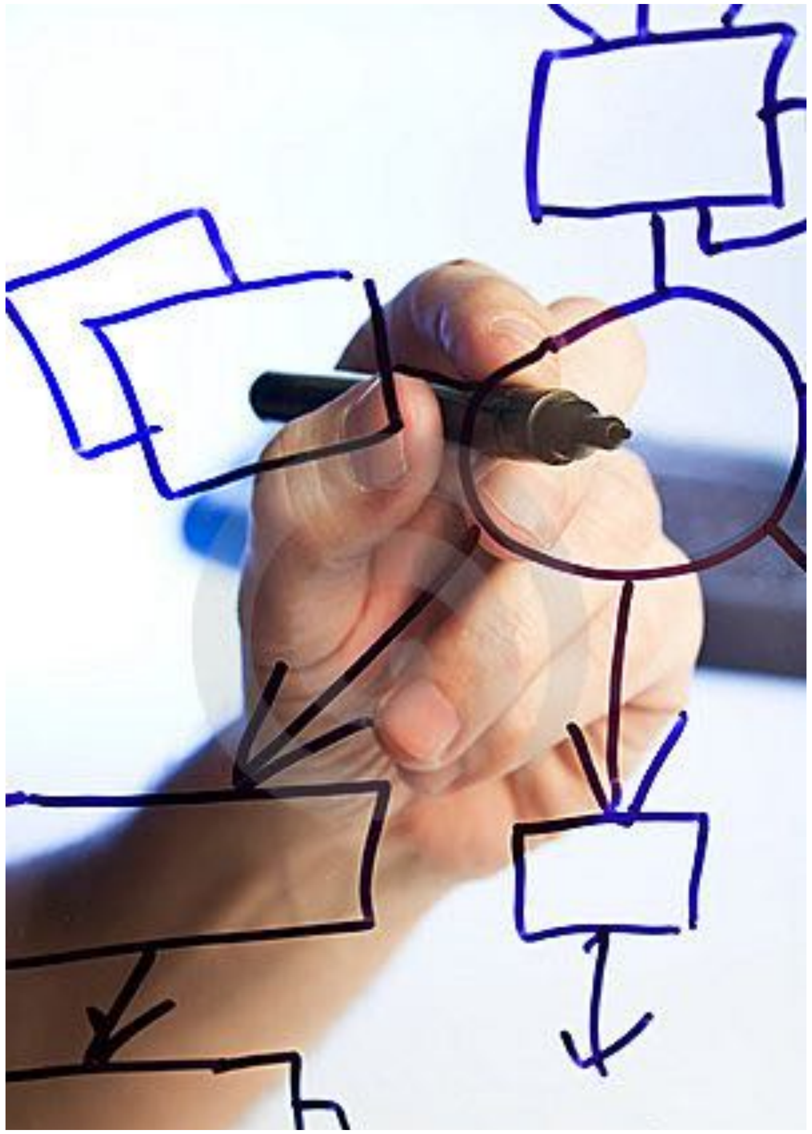
Human-Robot Interaction

Place the heineken block west of the mercedes block.

Existing Approaches

Successes

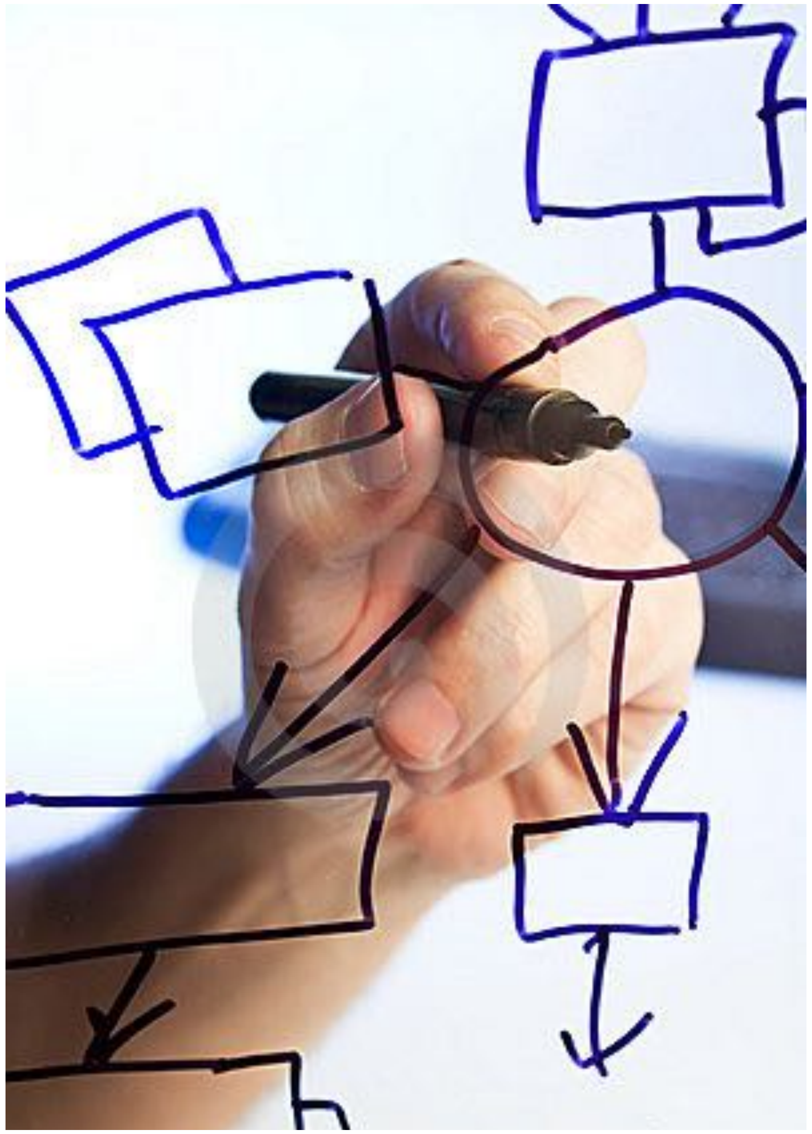
- Rule-based frameworks
- Modular architecture



Existing Approaches

Successes

- Rule-based frameworks
- Modular architecture



Challenges

- **Expensive** to build
- **Hard to deploy** to new applications

Data-driven NLG

- **Learn** generation process *directly* from data
- **Easier** to build and maintain
- **Adapt** to multiple domains

Data-driven NLG

- **Learn** generation process *directly* from data
- **Easier** to build and maintain
- **Adapt** to multiple domains

Challenges

- **Require** large corpora - NLG is **low-resourced**
- **New** machine learning model for **every** application

Outline

- **Neural Network** architecture for NLG
  - **Learn** from different inputs

Outline

- **Neural Network** architecture for NLG
  - **Learn** from different inputs
- **Address** low-resource problem
  - Generic framework for **scaling** to large corpora without extra annotation
  - **Collect** large datasets from community-based platform

Outline

- **Neural Network** architecture for NLG
  - **Learn** from different inputs
- **Address** low-resource problem
  - Generic framework for **scaling** to large corpora without extra annotation
  - **Collect** large datasets from community-based platform
- **Adapt** to two applications
  - Meaning Representations
  - Code to Language

Neural NLG

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

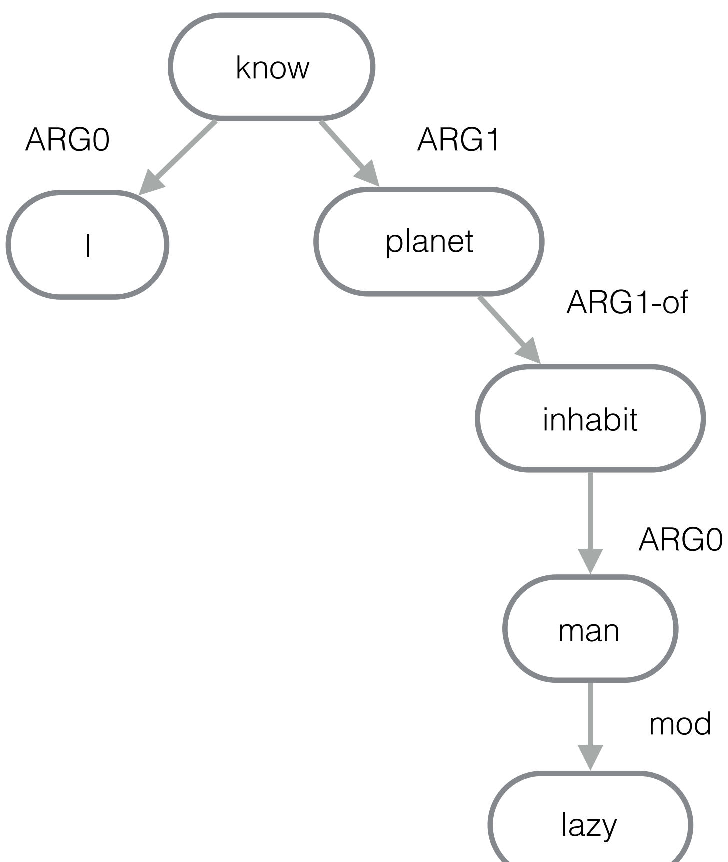
Overview

- Sequence to sequence architecture
  - End-to-end model w/o intermediate representations
  - Linearisation of input to string
  - Pre-process

- Paired Training
  - Scalable data augmentation

Meaning Representations

Input: Graph Structure
(Abstract Meaning Representation - AMR; Banarescu et al., 2013)



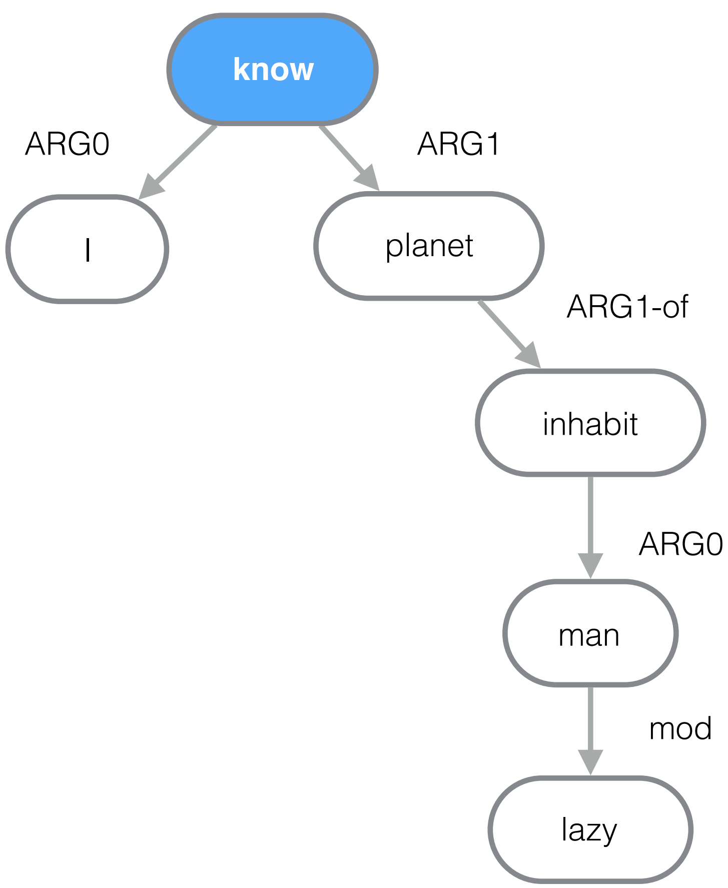
I knew a **planet** that was **inhabited** by a **lazy man**.

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

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

(Flanigan et al, NAACL 2016, Pourdamaghani and Knight, INLG 2016, Song et al, EMNLP 2016)
(Konstas, Iyer, Yatskar, Choi, Zettlemoyer, ACL 2017, to Appear)

Meaning Representations



Input: Graph Structure
(Abstract Meaning Representation - AMR; Banarescu et al., 2013)

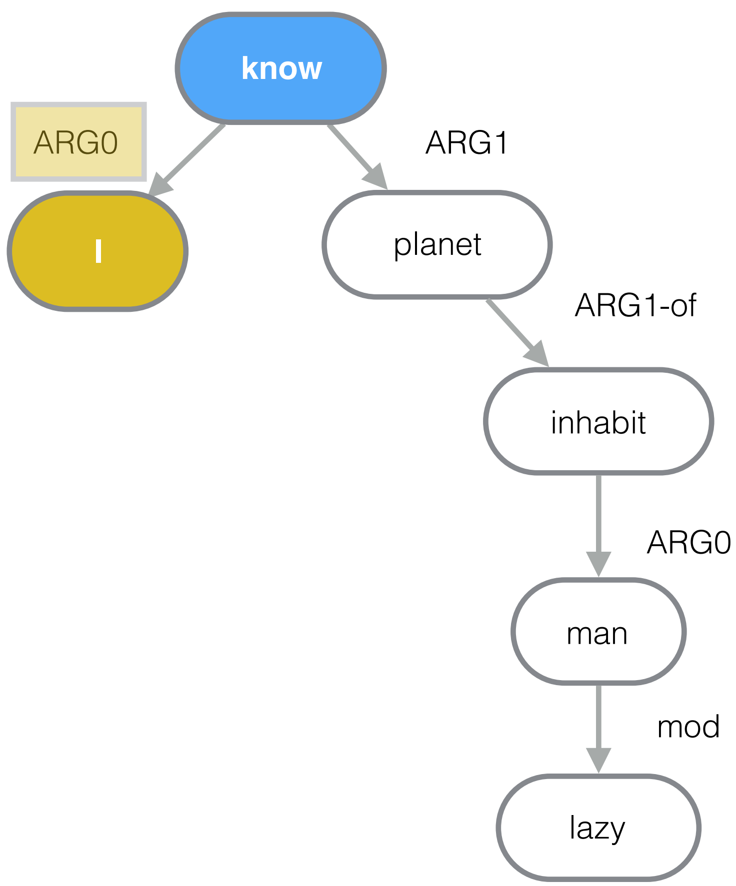
I knew a **planet** that was **inhabited** by a **lazy man**.

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

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

(Flanigan et al, NAACL 2016, Pourdamaghani and Knight, INLG 2016, Song et al, EMNLP 2016)
(Konstas, Iyer, Yatskar, Choi, Zettlemoyer, ACL 2017, to Appear)

Meaning Representations



Input: Graph Structure
(Abstract Meaning Representation - AMR; Banarescu et al., 2013)

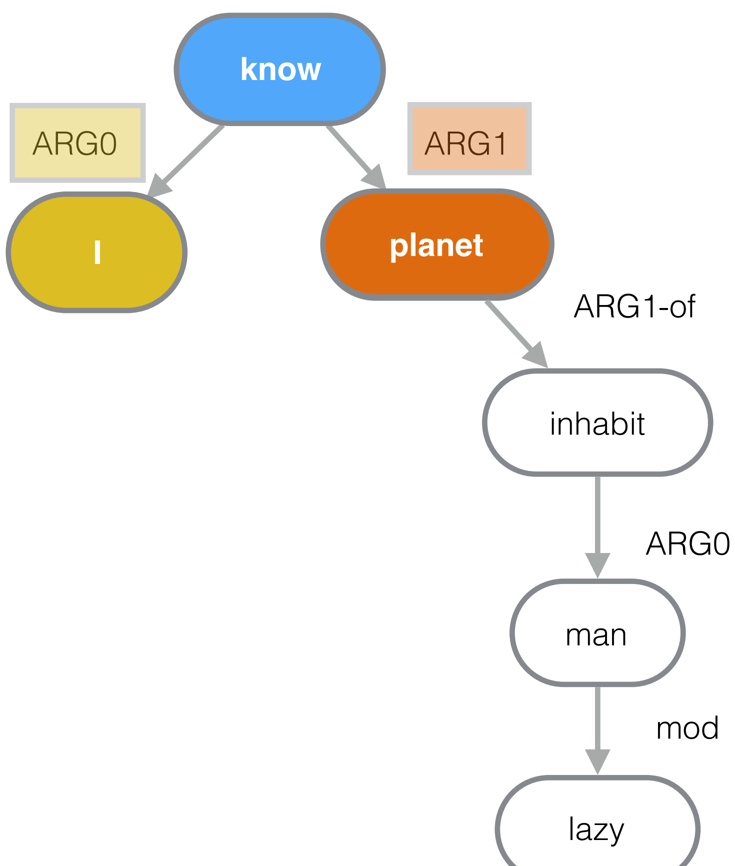
I knew a **planet** that was **inhabited** by a **lazy man**.

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

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

(Flanigan et al, NAACL 2016, Pourdamaghani and Knight, INLG 2016, Song et al, EMNLP 2016)
(Konstas, Iyer, Yatskar, Choi, Zettlemoyer, ACL 2017, to Appear)

Meaning Representations



Input: Graph Structure
(Abstract Meaning Representation - AMR; Banarescu et al., 2013)

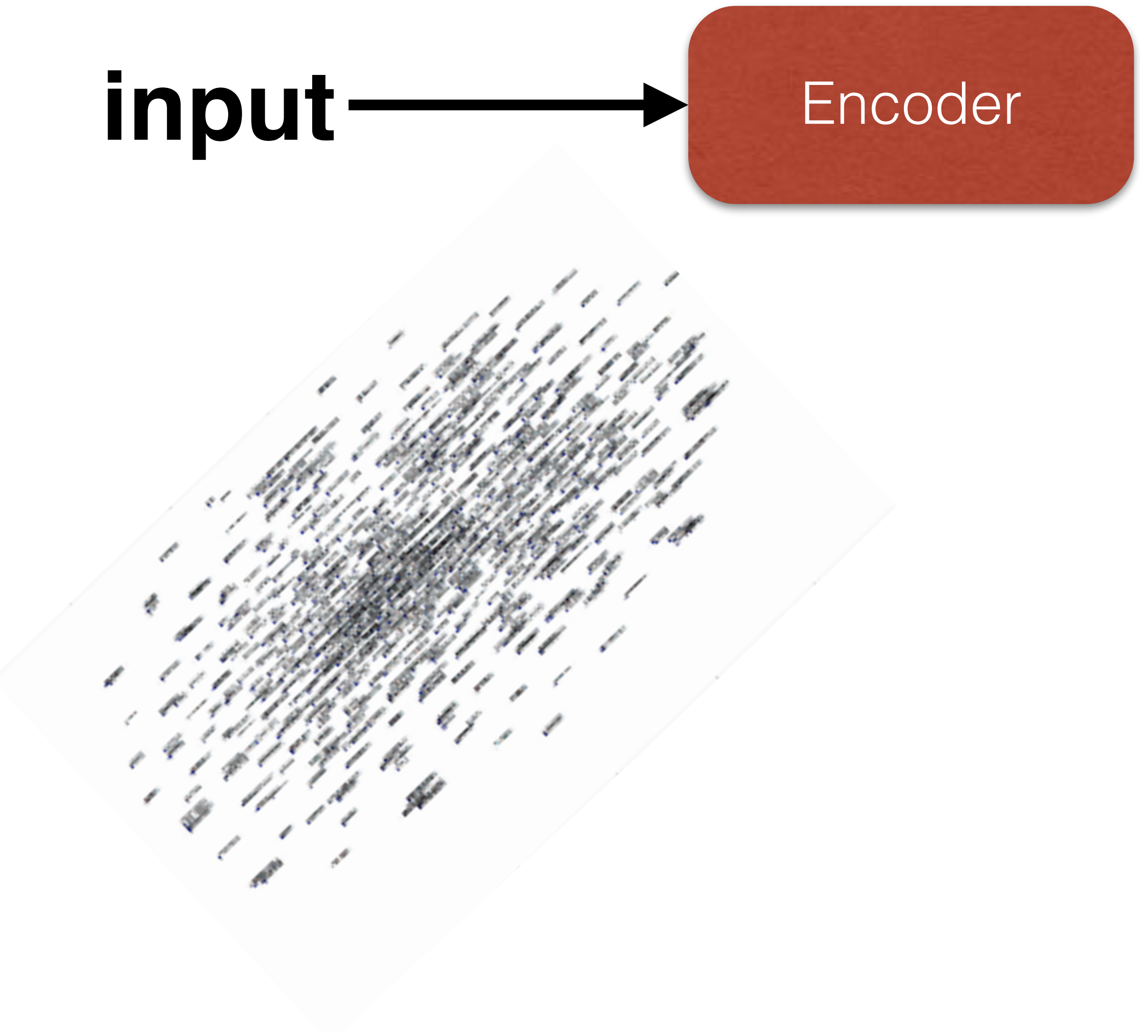
I knew a **planet** that was **inhabited** by a **lazy man**.

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

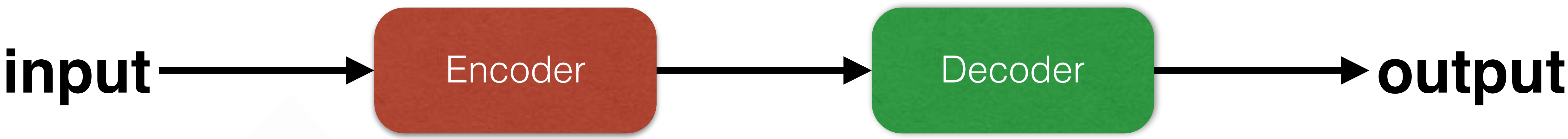
I know a **planet**. It is **inhabited** by a **lazy man**.

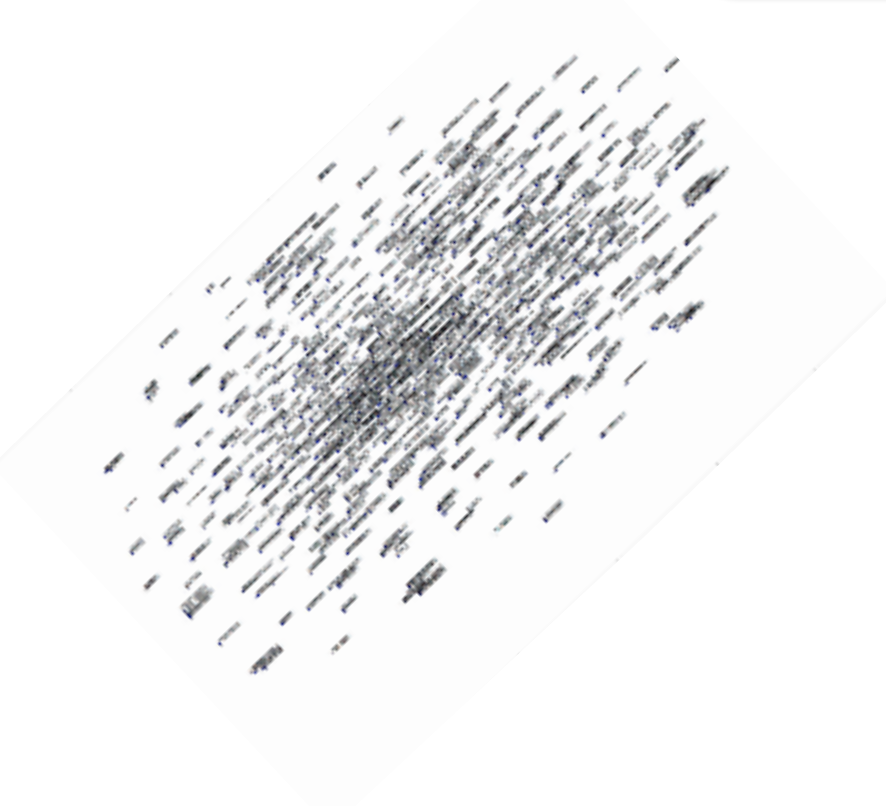
(Flanigan et al, NAACL 2016, Pourdamaghani and Knight, INLG 2016, Song et al, EMNLP 2016)
(Konstas, Iyer, Yatskar, Choi, Zettlemoyer, ACL 2017, to Appear)

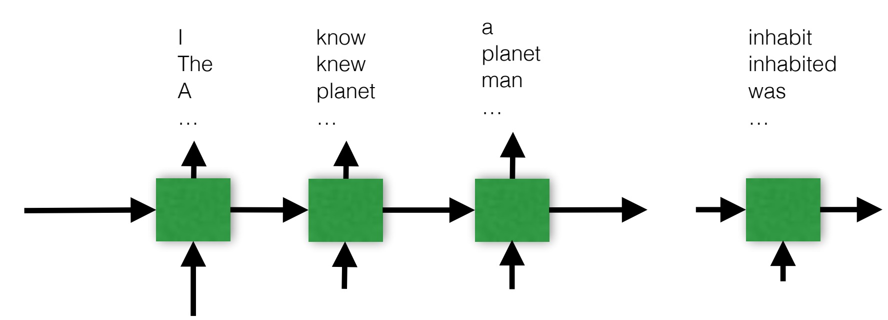
Sequence to sequence model



Sequence to sequence model

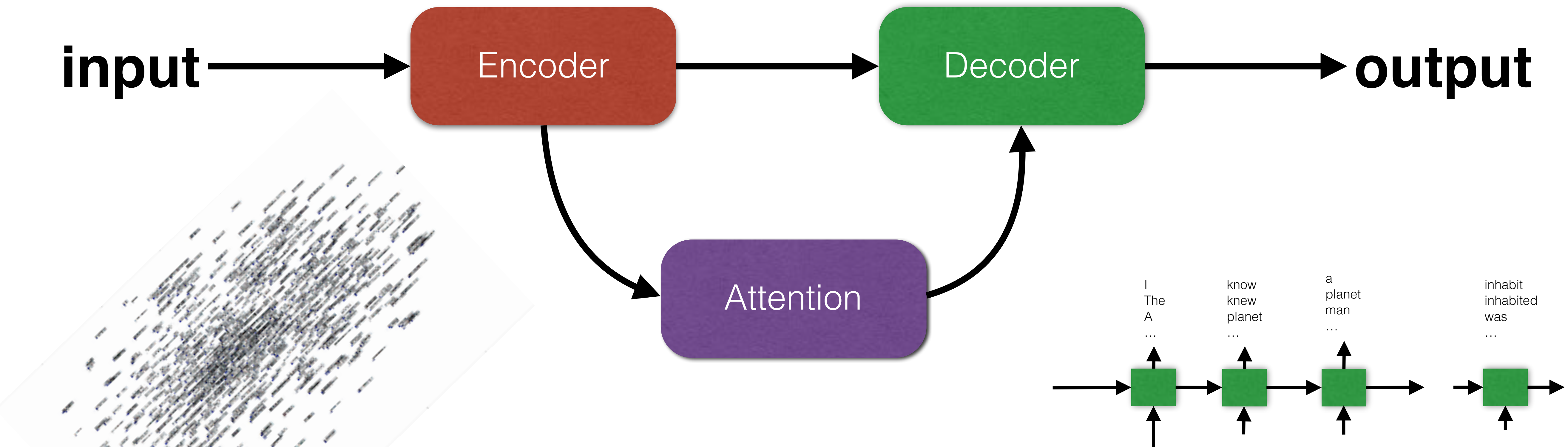


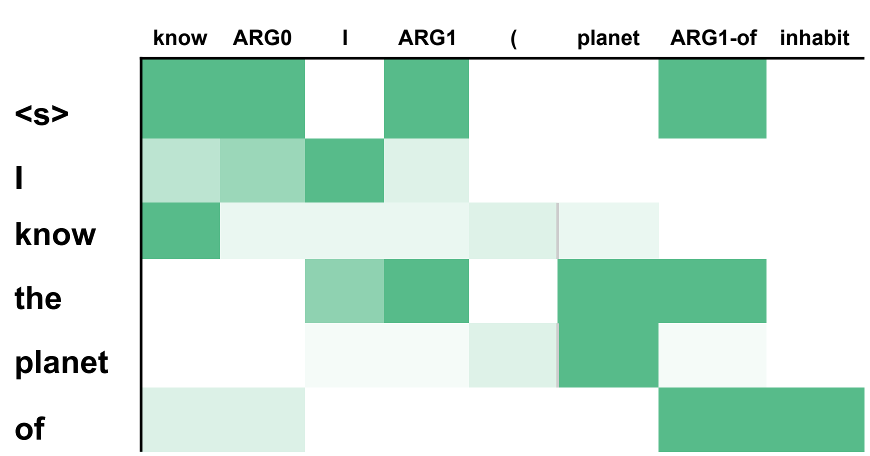




$$\hat{w} = \underset{w}{\operatorname{argmax}} \prod_i p\big(w_i | w_{<i}, \mathbf{h}^{(s)}\big)$$

Sequence to sequence model

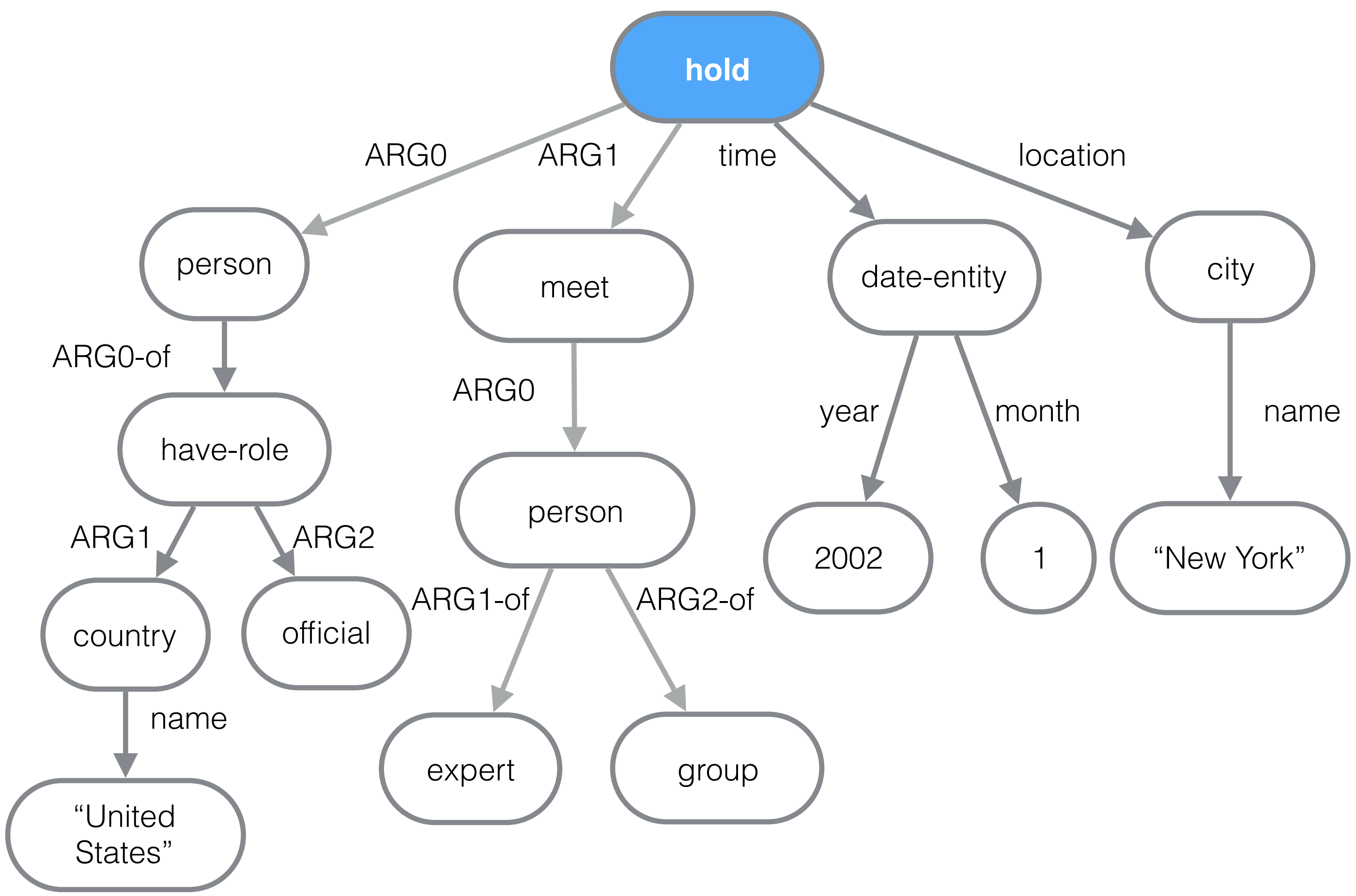




$$\hat{w} = \operatorname*{argmax}_{w} \prod_{i} p(w_i | w_{<i}, \mathbf{h}^{(s)})$$

Linearization

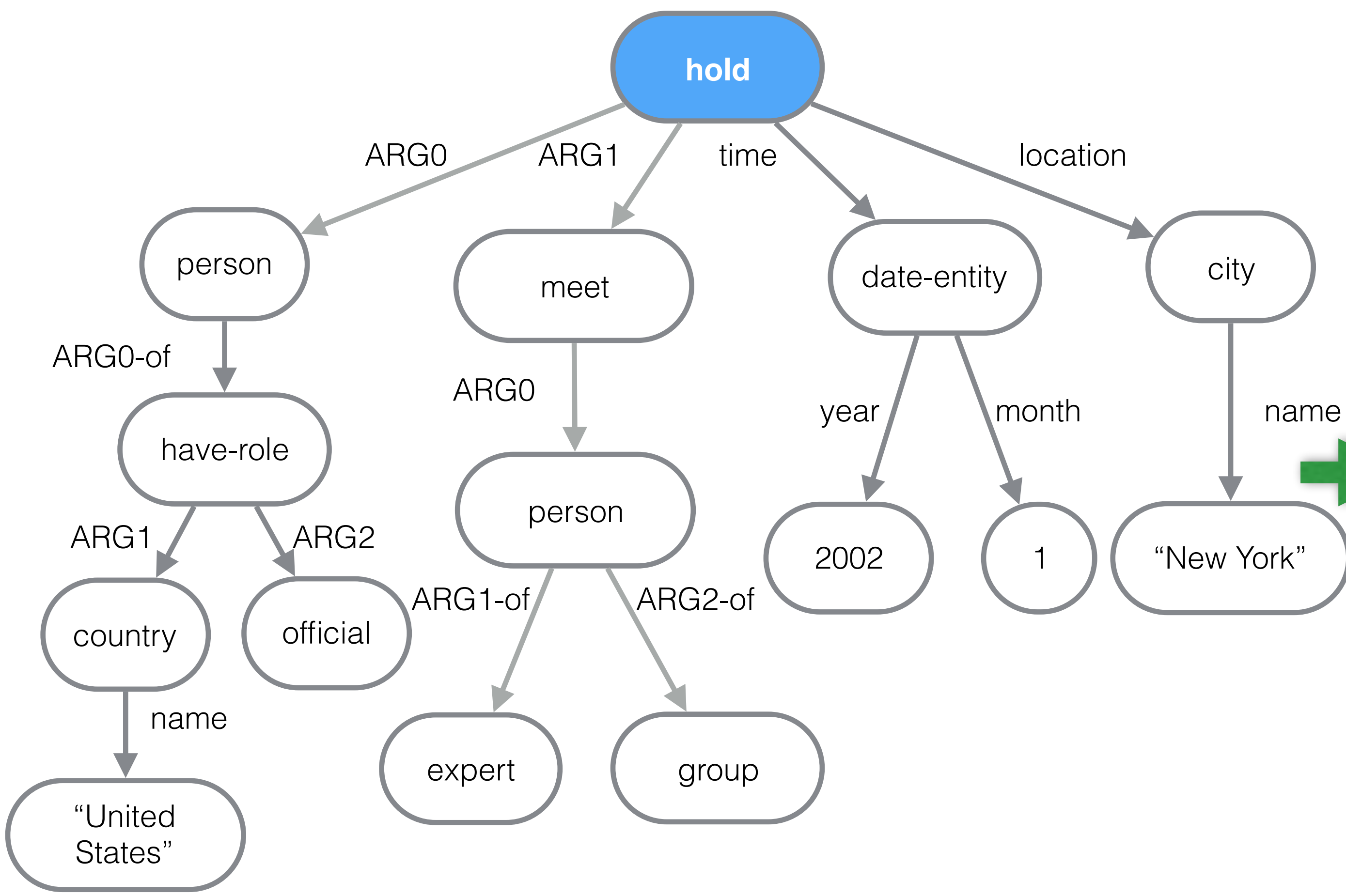
Graph —> Depth First Search



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

Linearization

Graph —> Depth First Search




```
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 .

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

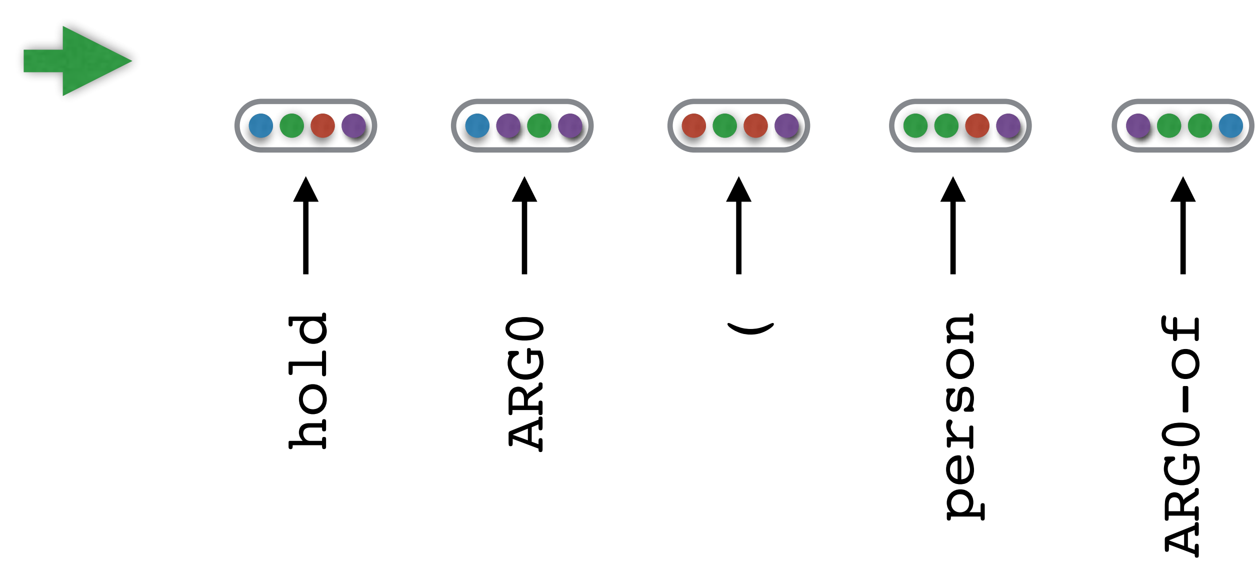
```
hold
   :ARG0 (person
            :ARG0-of (have-role
               :ARG1 United_States
               :ARG2 official)
         )
   :ARG1 (meet
            :ARG0 (person
               :ARG1-of expert
               :ARG2-of group)
         )
   :time (date-entity 2002 1)
   :location New_York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

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
```

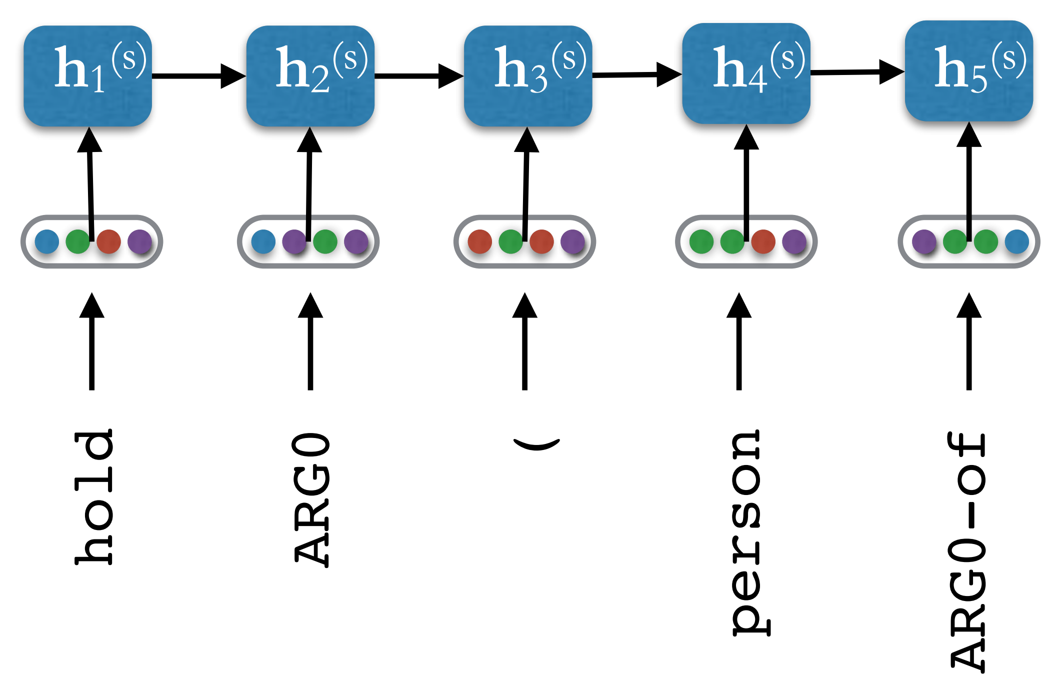


Encoding

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
```

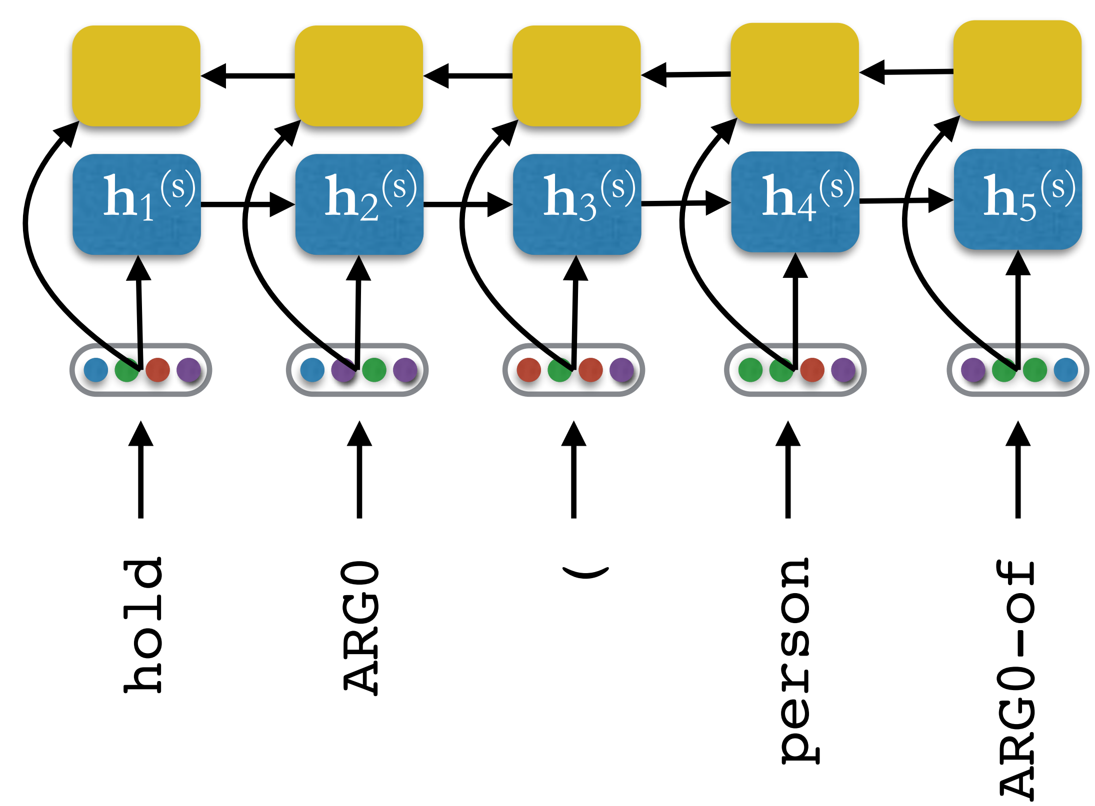


Encoding

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
```

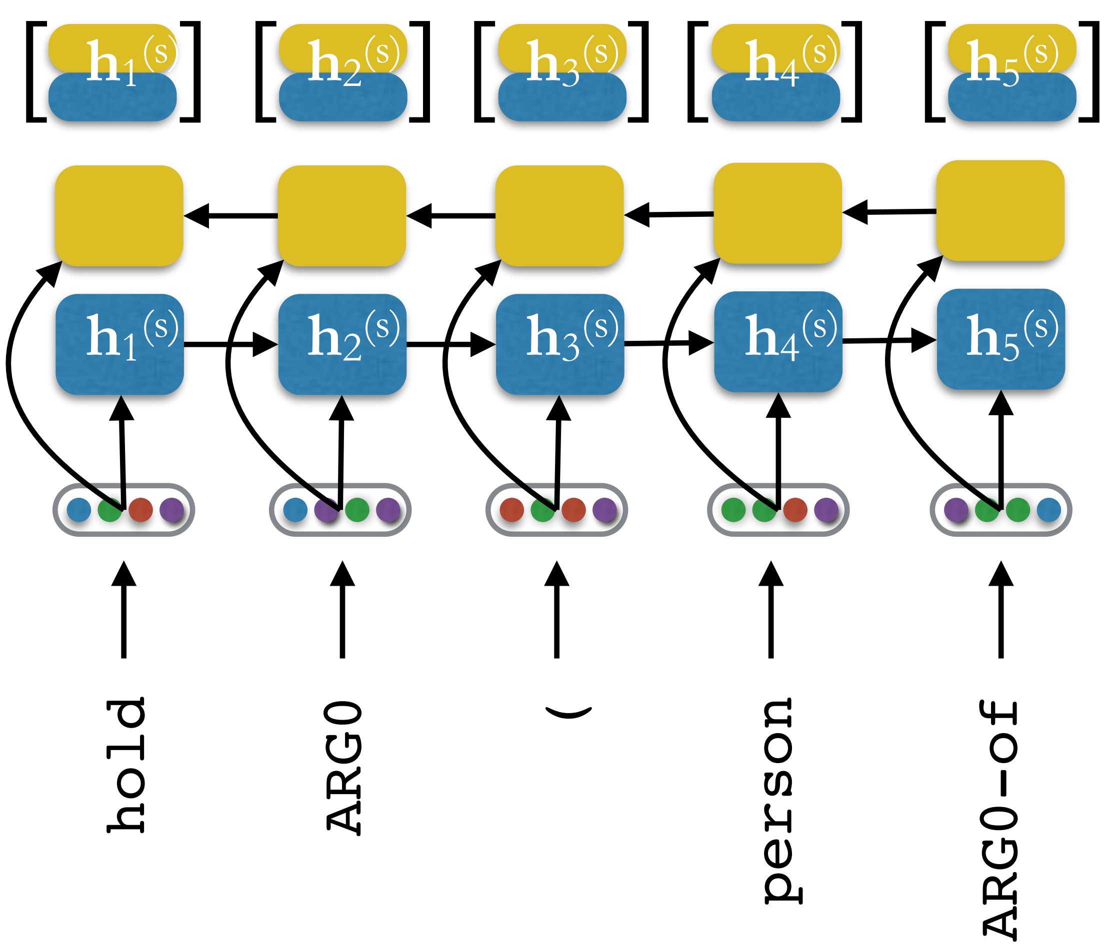


Encoding

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
```

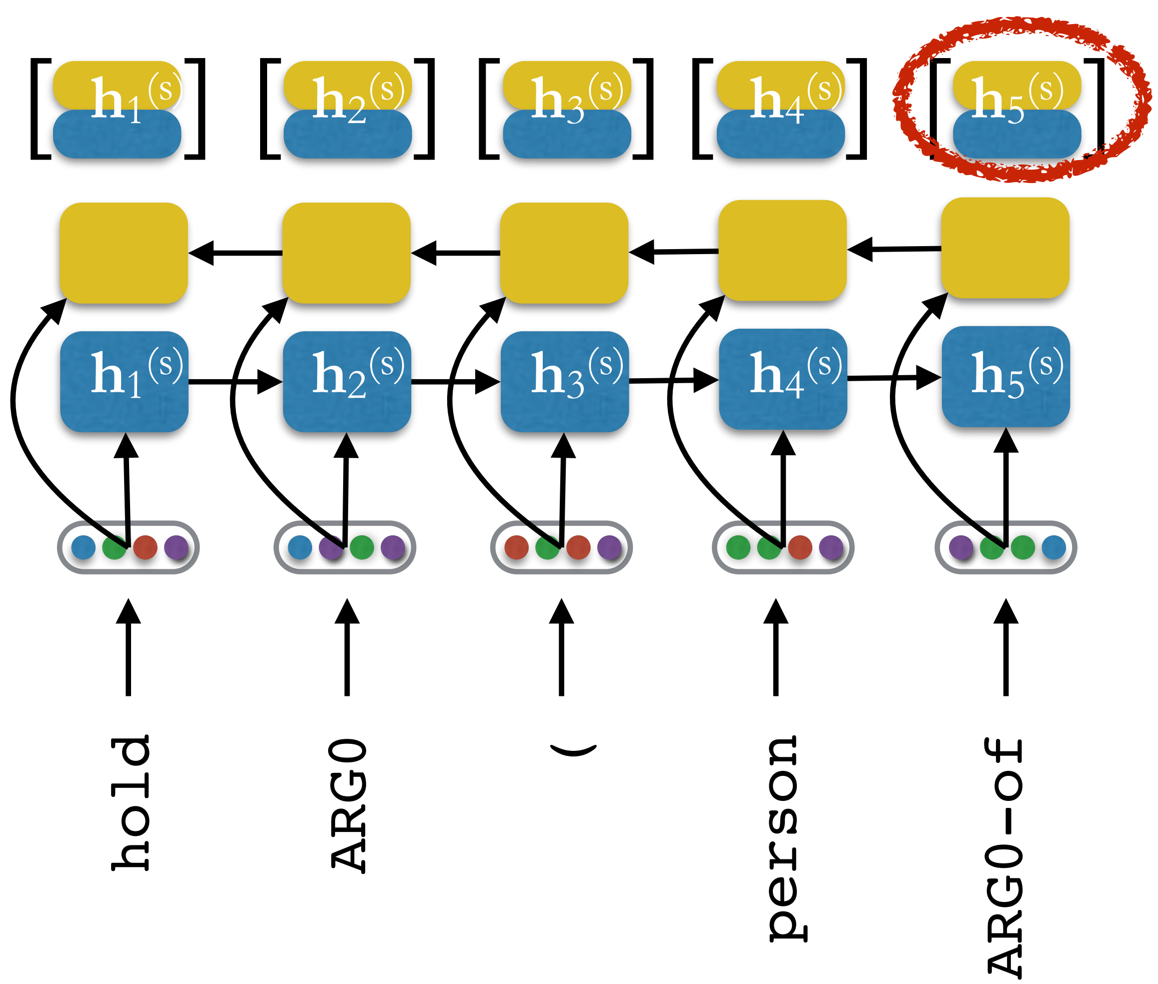


Encoding

Linearize —> RNN encoding

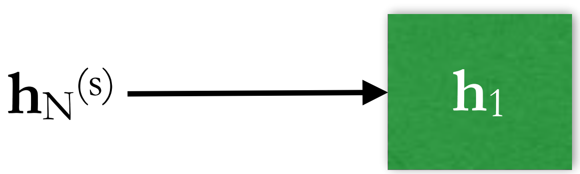
- Token embeddings
- Recurrent Neural Network (RNN)
- Bi-directional RNN

```
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

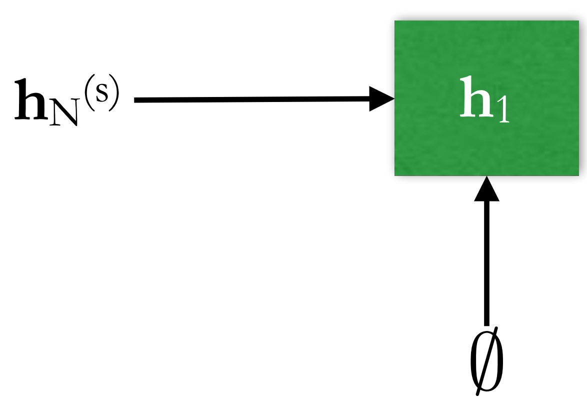
RNN Encoding —> RNN Decoding (Beam search)



Decoding

RNN Encoding —> RNN Decoding (Beam search)

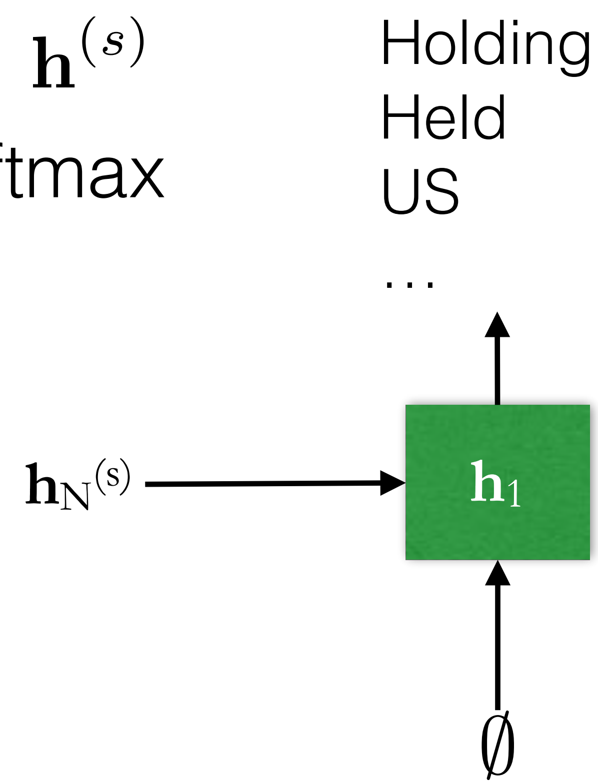
- init $\mathbf{h}^{(s)}$



Decoding

RNN Encoding —> RNN Decoding (Beam search)

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

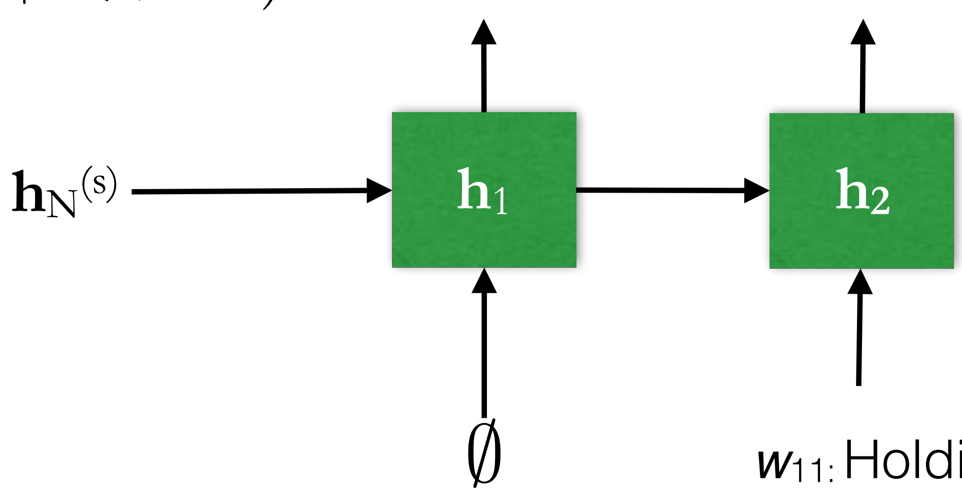


Decoding

RNN Encoding —> RNN Decoding (Beam search)

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

Holding	a
Held	the
US	meeting
…	…



w_{11}: Holding

w_{12}: Helds

w_{13}: Hold

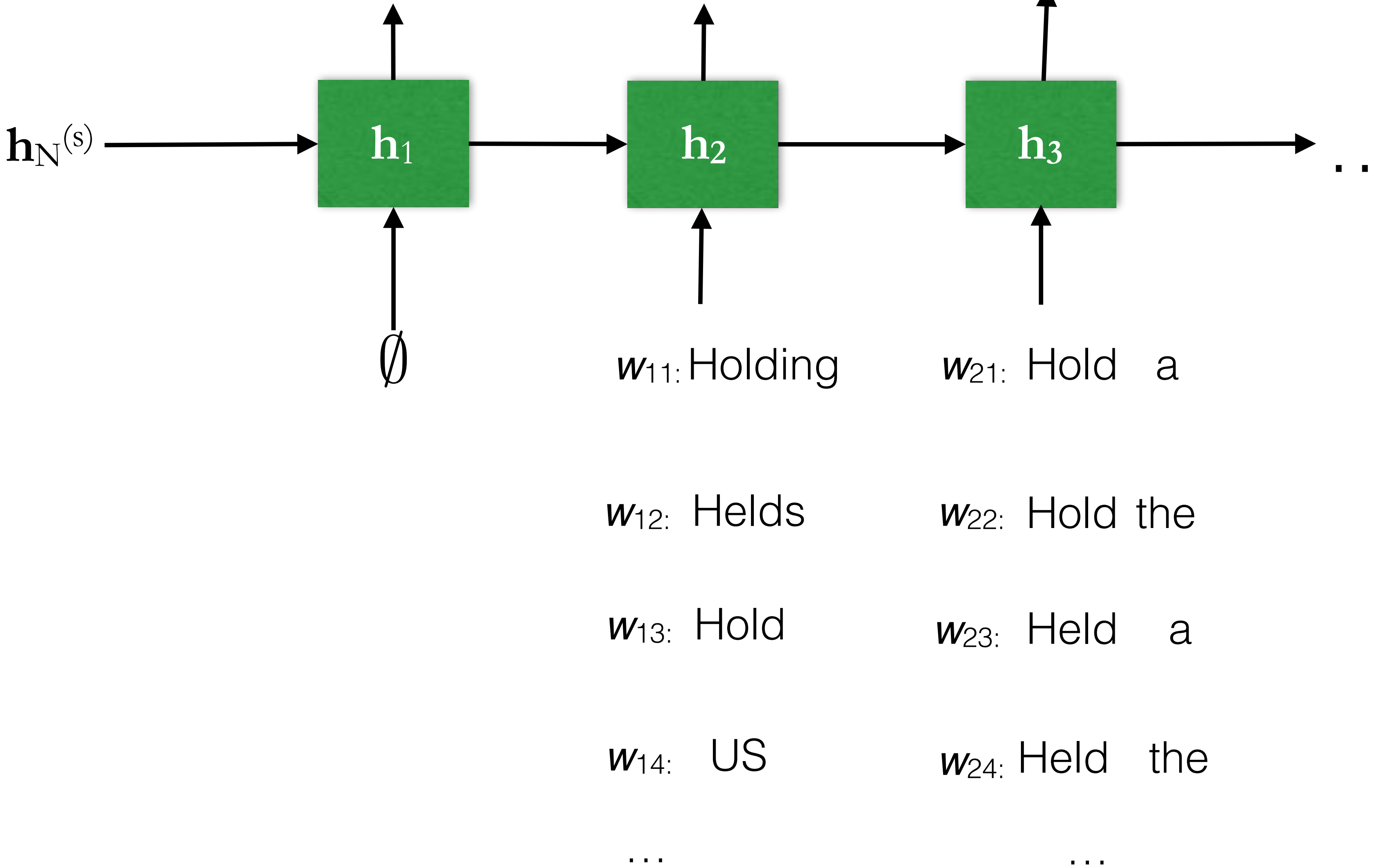
w_{14}: US

…

Decoding

RNN Encoding —> RNN Decoding (Beam search)

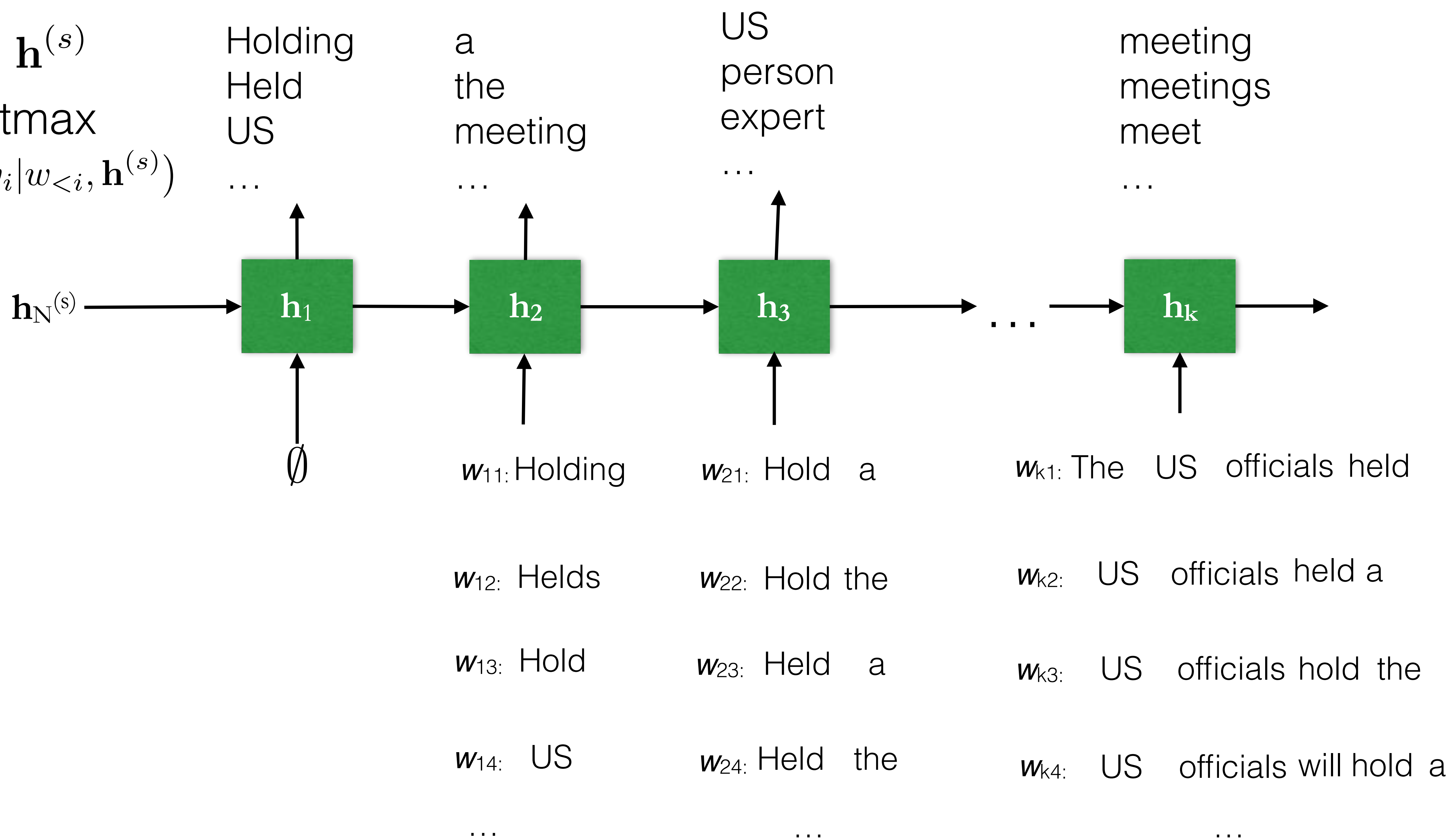
- init $\mathbf{h}^{(s)}$
- softmax
- $p(w_i | w_{<i}, \mathbf{h}^{(s)})$



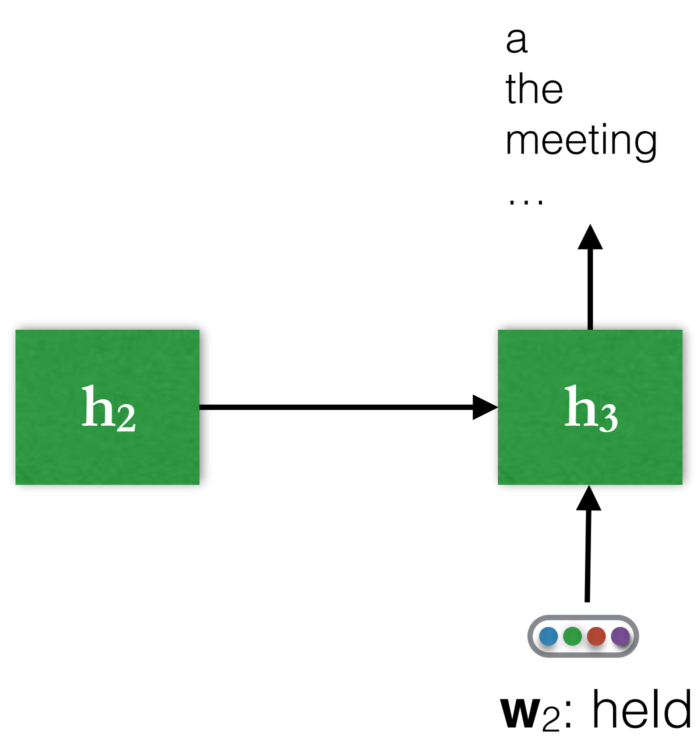
Decoding

RNN Encoding —> RNN Decoding (Beam search)

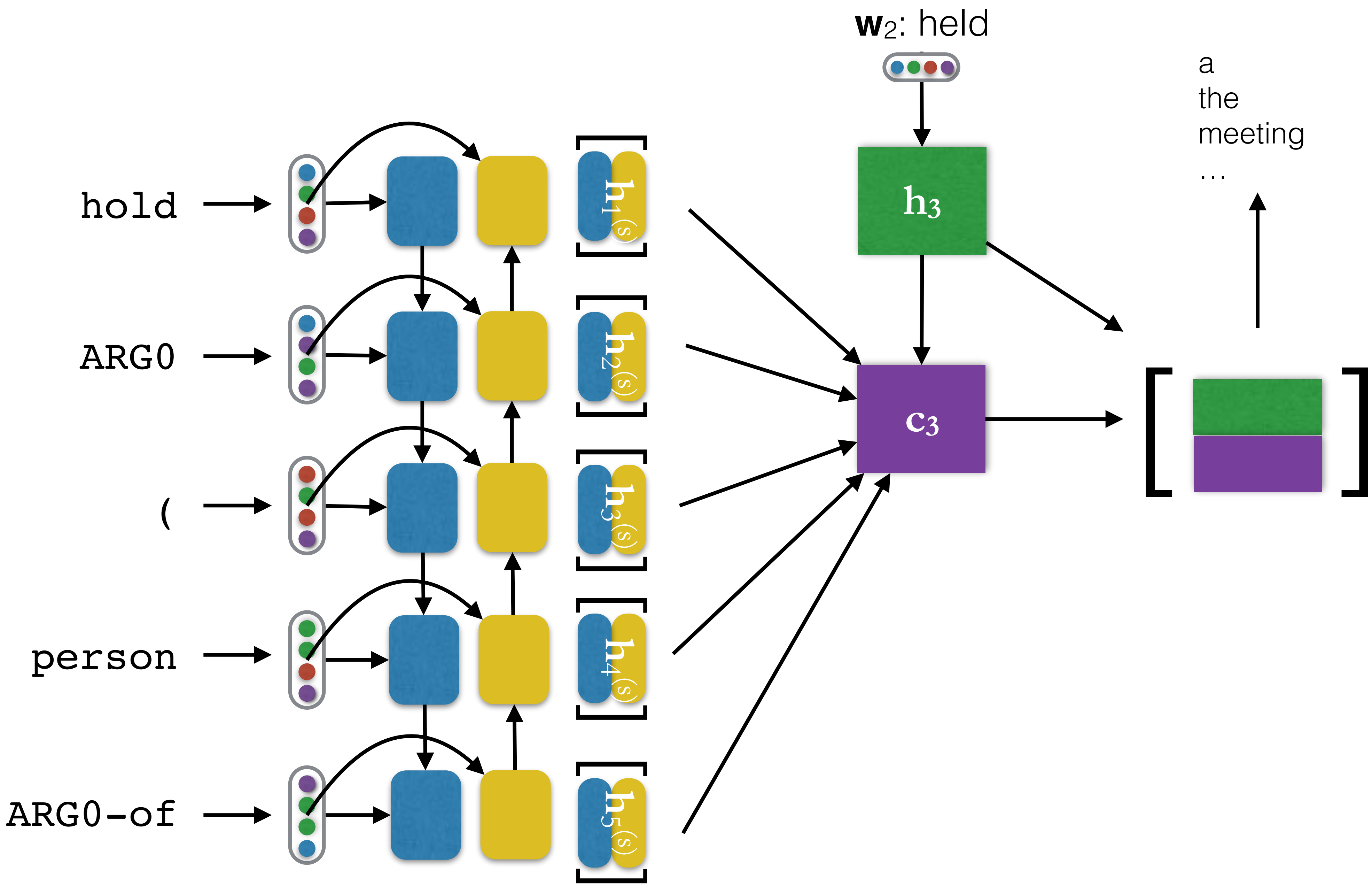
- init $\mathbf{h}^{(s)}$
- softmax
- $p(w_i | w_{<i}, \mathbf{h}^{(s)})$



Attention

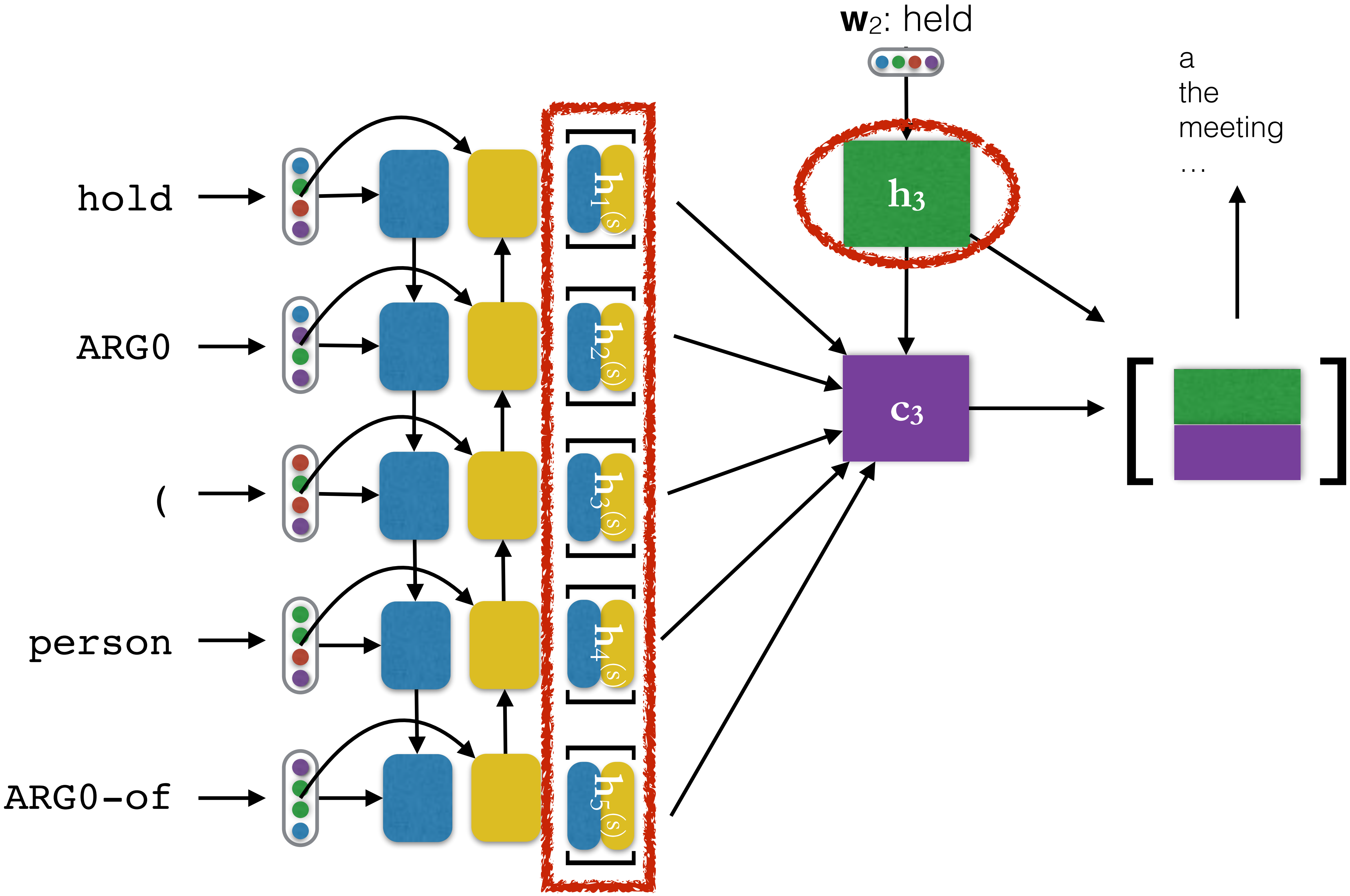


Attention



$\mathbf{w}_2$: held
a
the
meeting
...
hold
ARG0
(
person
ARG0-of
$h_1^{(s)}$
$h_2^{(s)}$
$h_3^{(s)}$
$h_4^{(s)}$
$h_5^{(s)}$
h_3
c_3

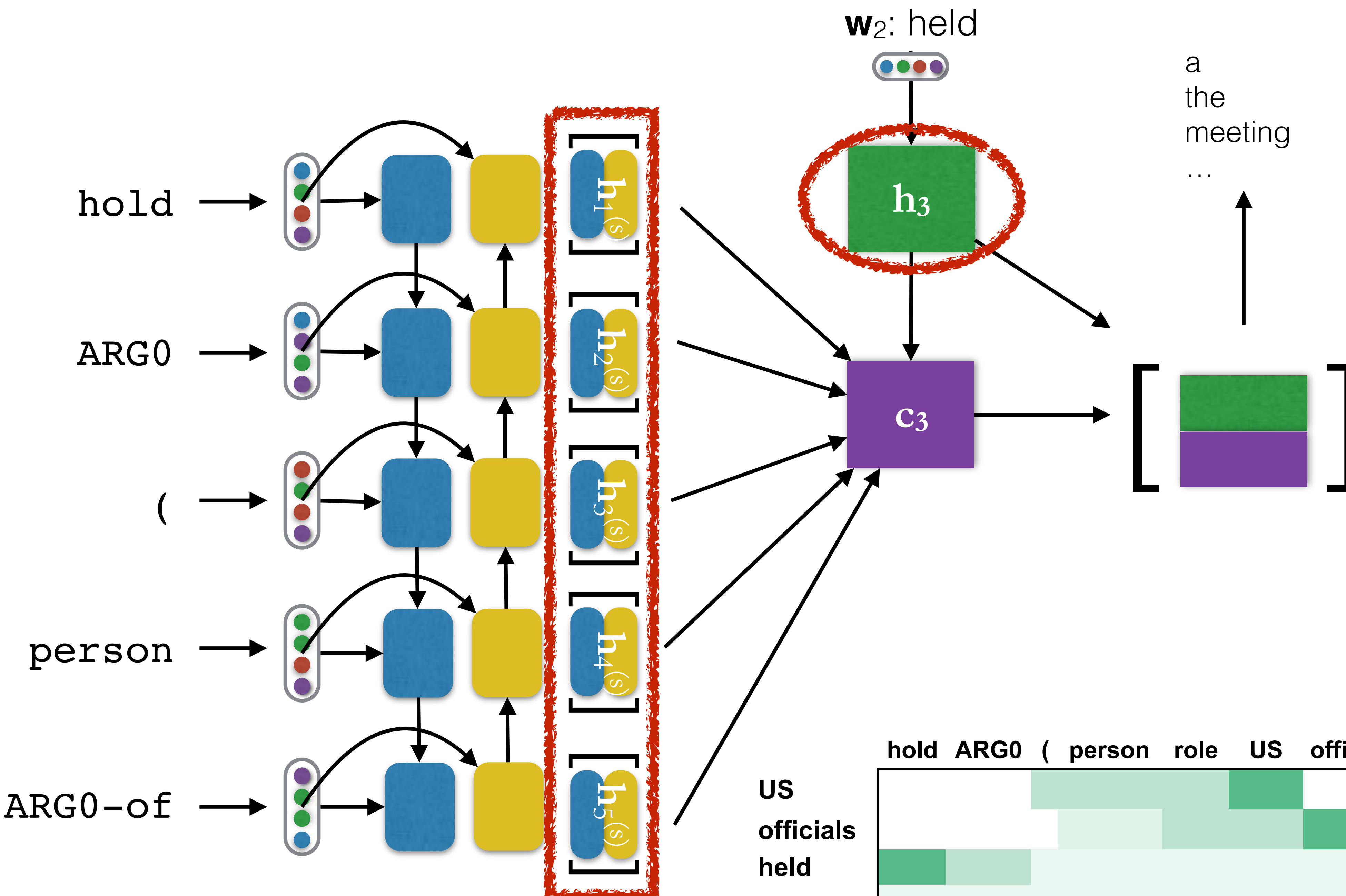
Attention



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

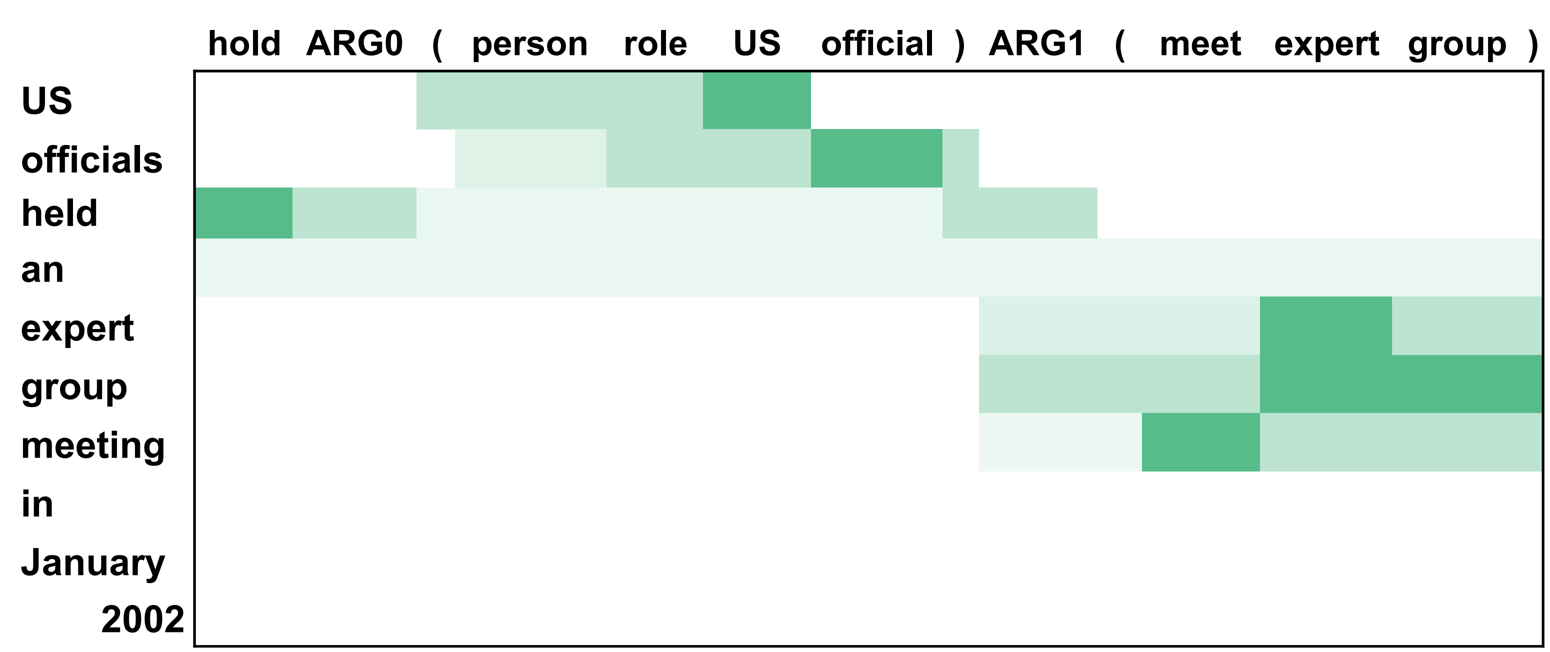
$$\mathbf{c}_i = \sum_i \mathbf{a}_{ij} \mathbf{h}_j^{(s)}$$

Attention

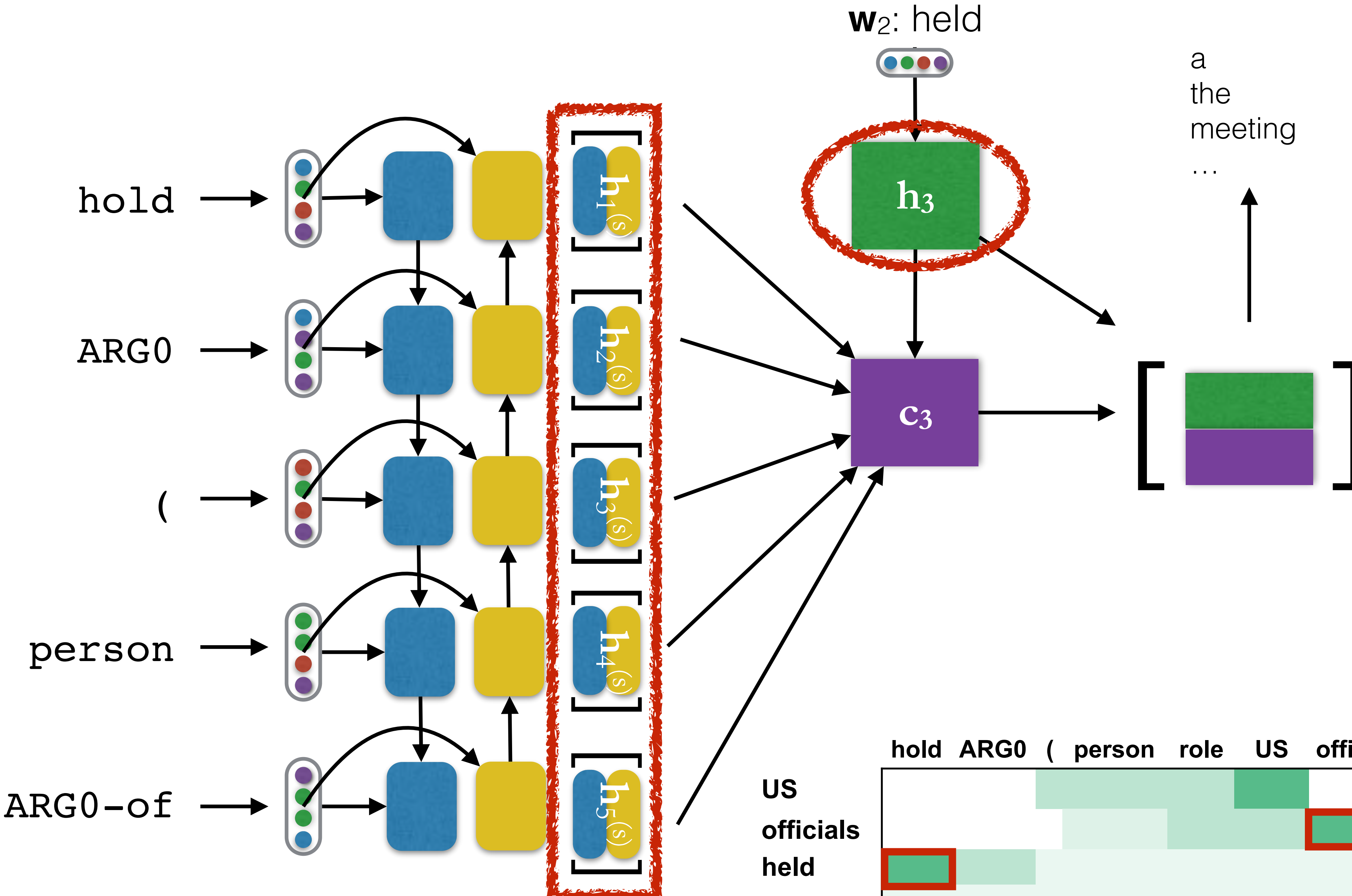


$$\mathbf{a}_i = \operatorname{soft\,max}\left(\mathbf{f}_i\left(\mathbf{h}^{(\mathbf{s})}, \mathbf{h}_i\right)\right)$$

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

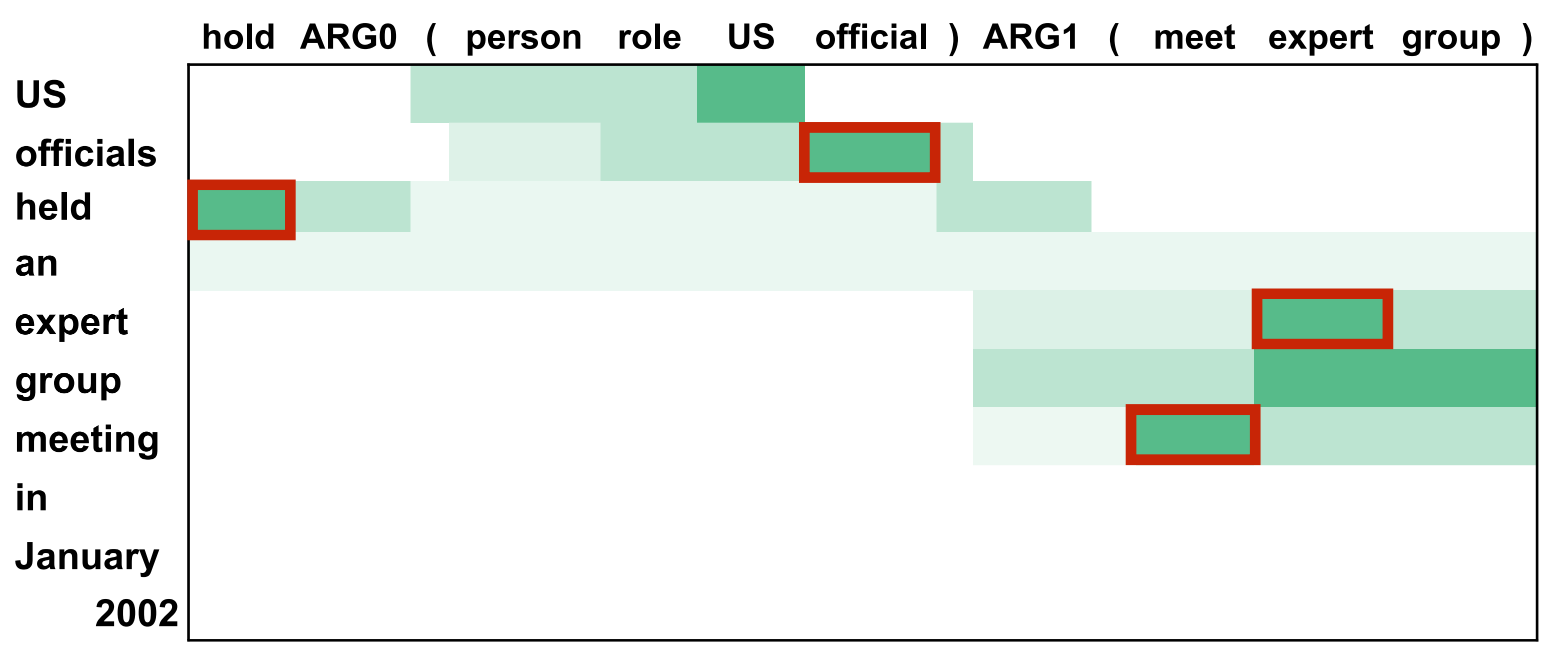


Attention

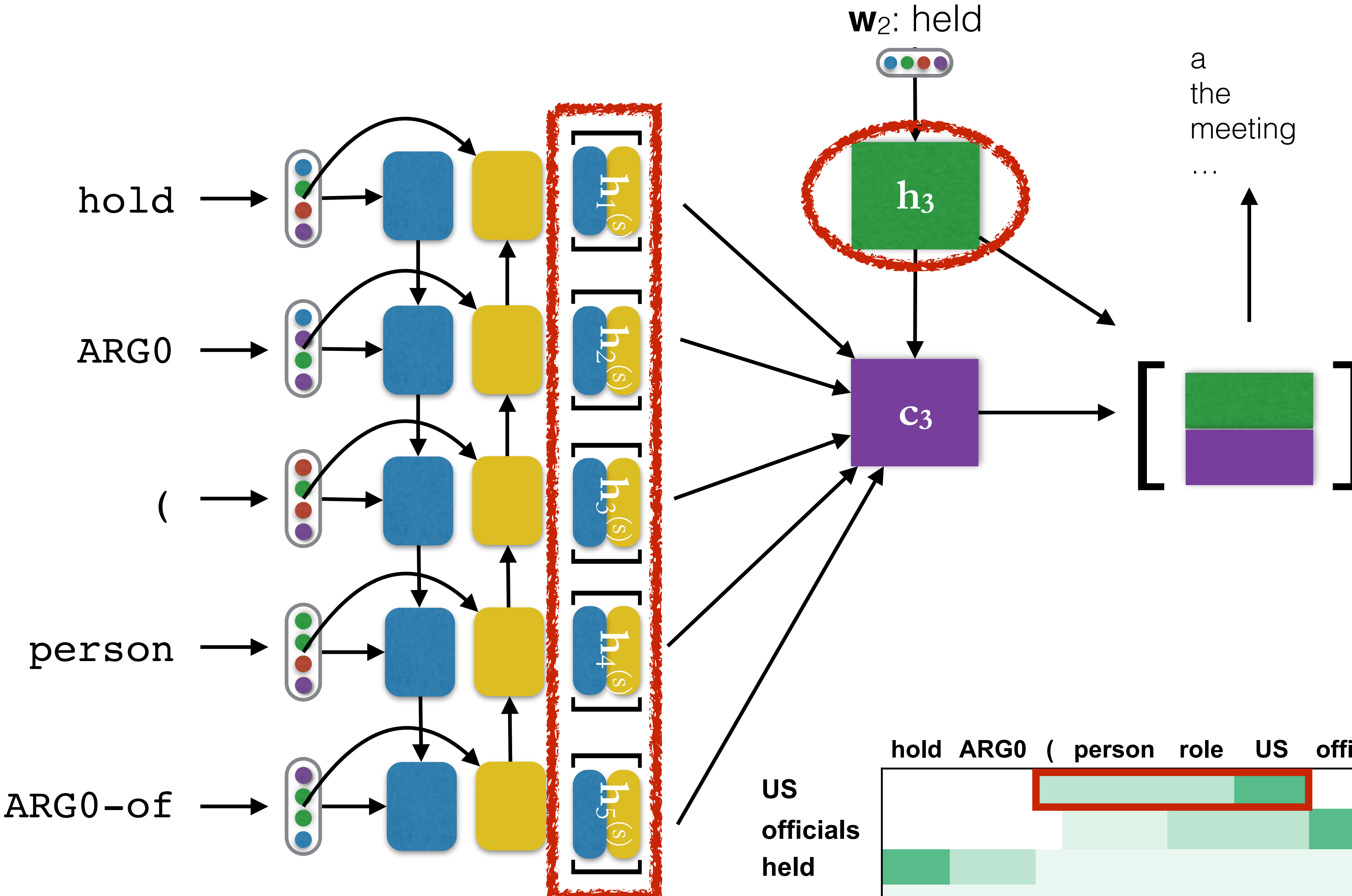


$$\mathbf{a}_i = \operatorname{soft\,max}\left(\mathbf{f}_i\left(\mathbf{h}^{(\mathbf{s})}, \mathbf{h}_i\right)\right)$$

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

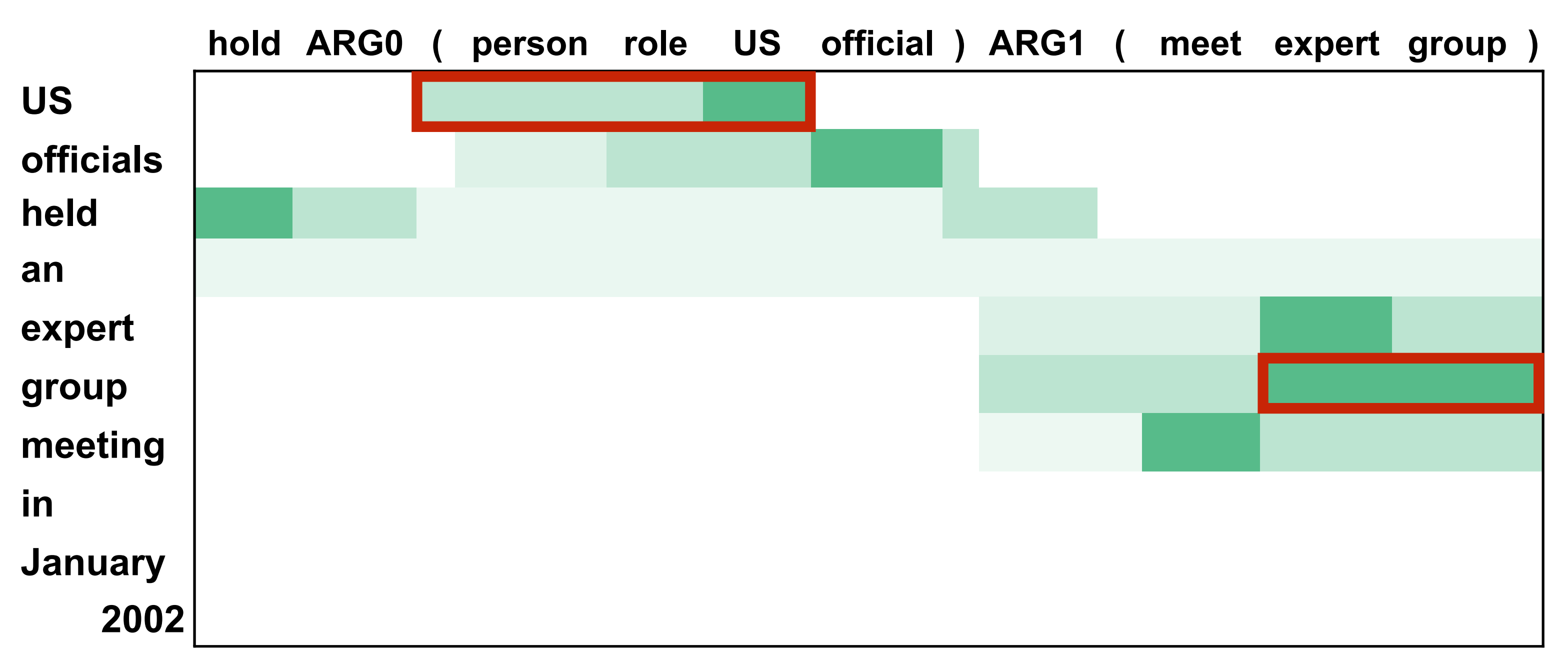


Attention

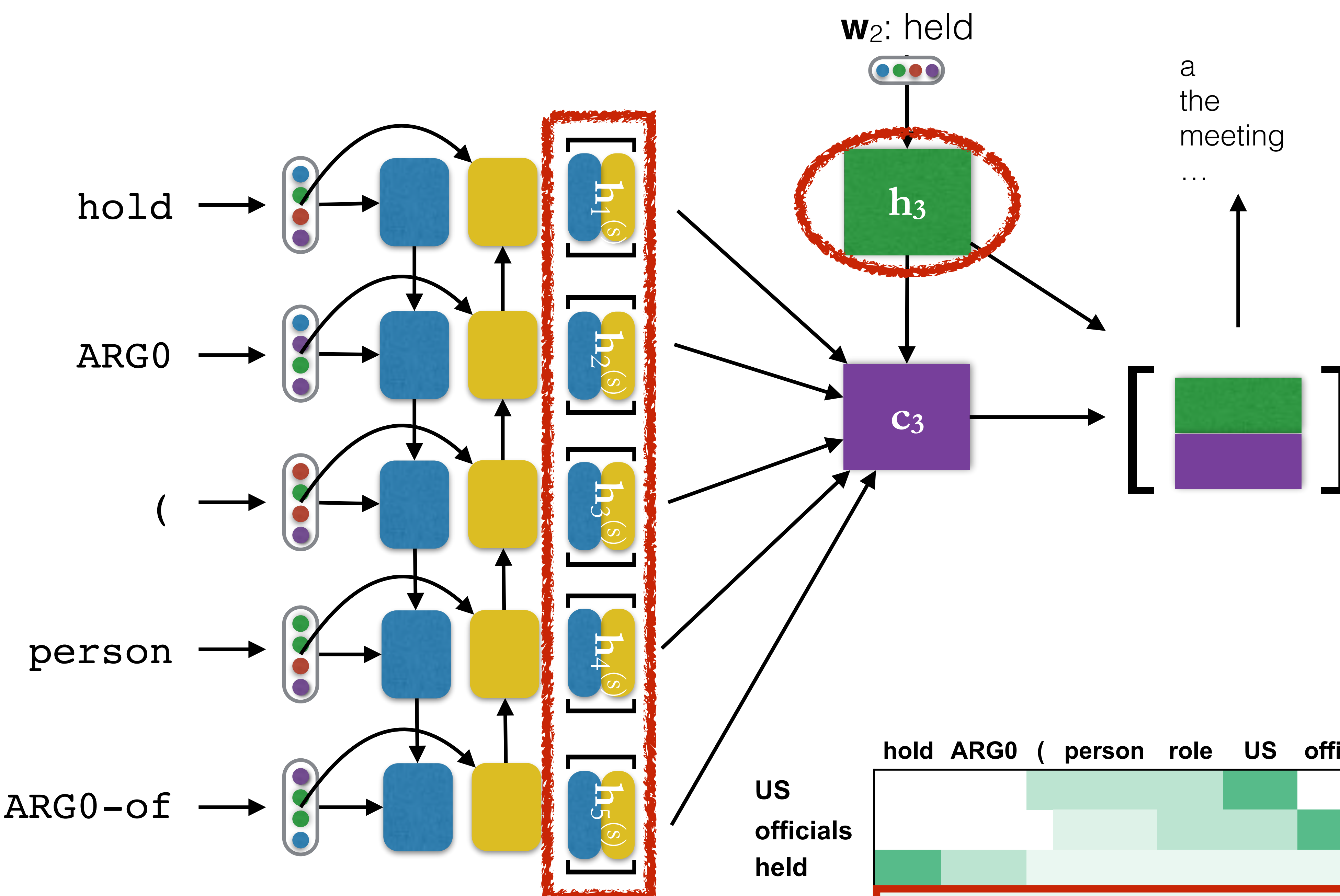


$$\mathbf{a}_i = \text{soft}\max\left(\mathbf{f}_i\left(\mathbf{h}^{(\mathbf{s})}, \mathbf{h}_i\right)\right)$$

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

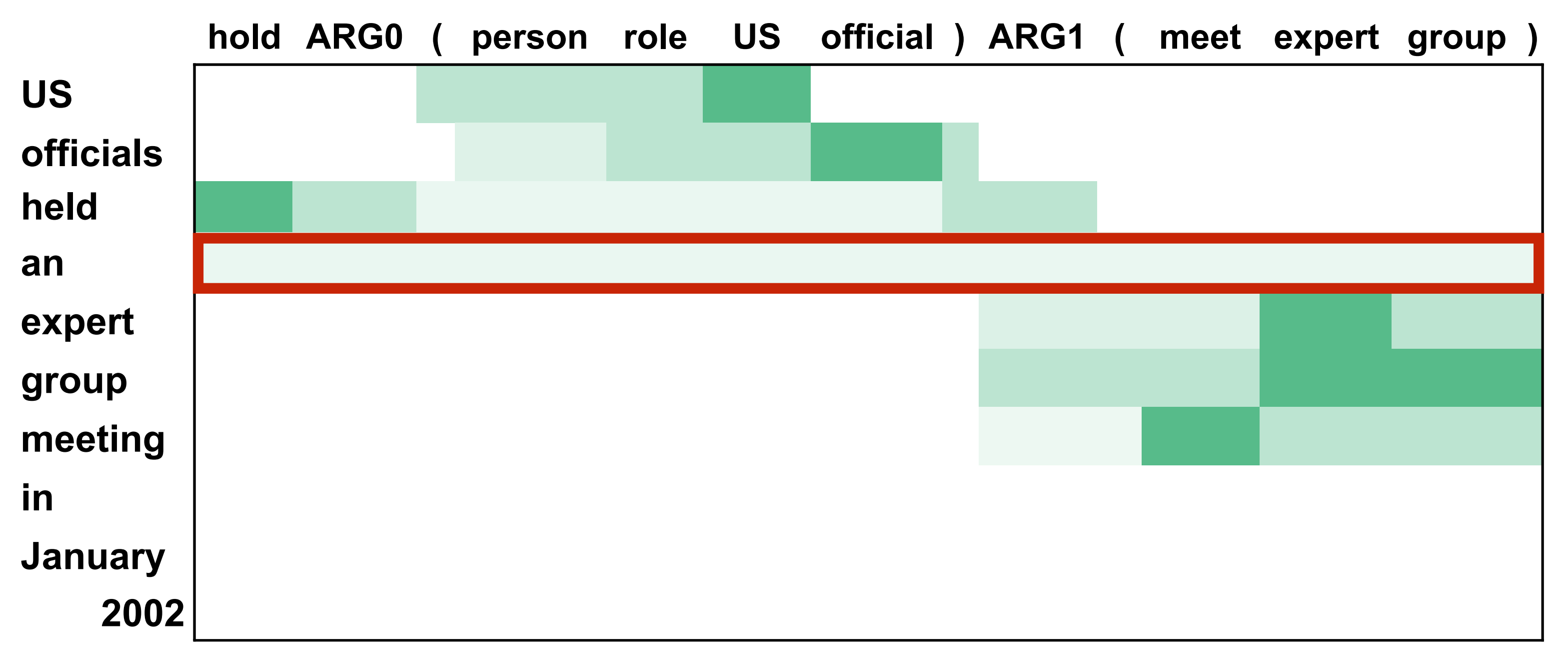


Attention



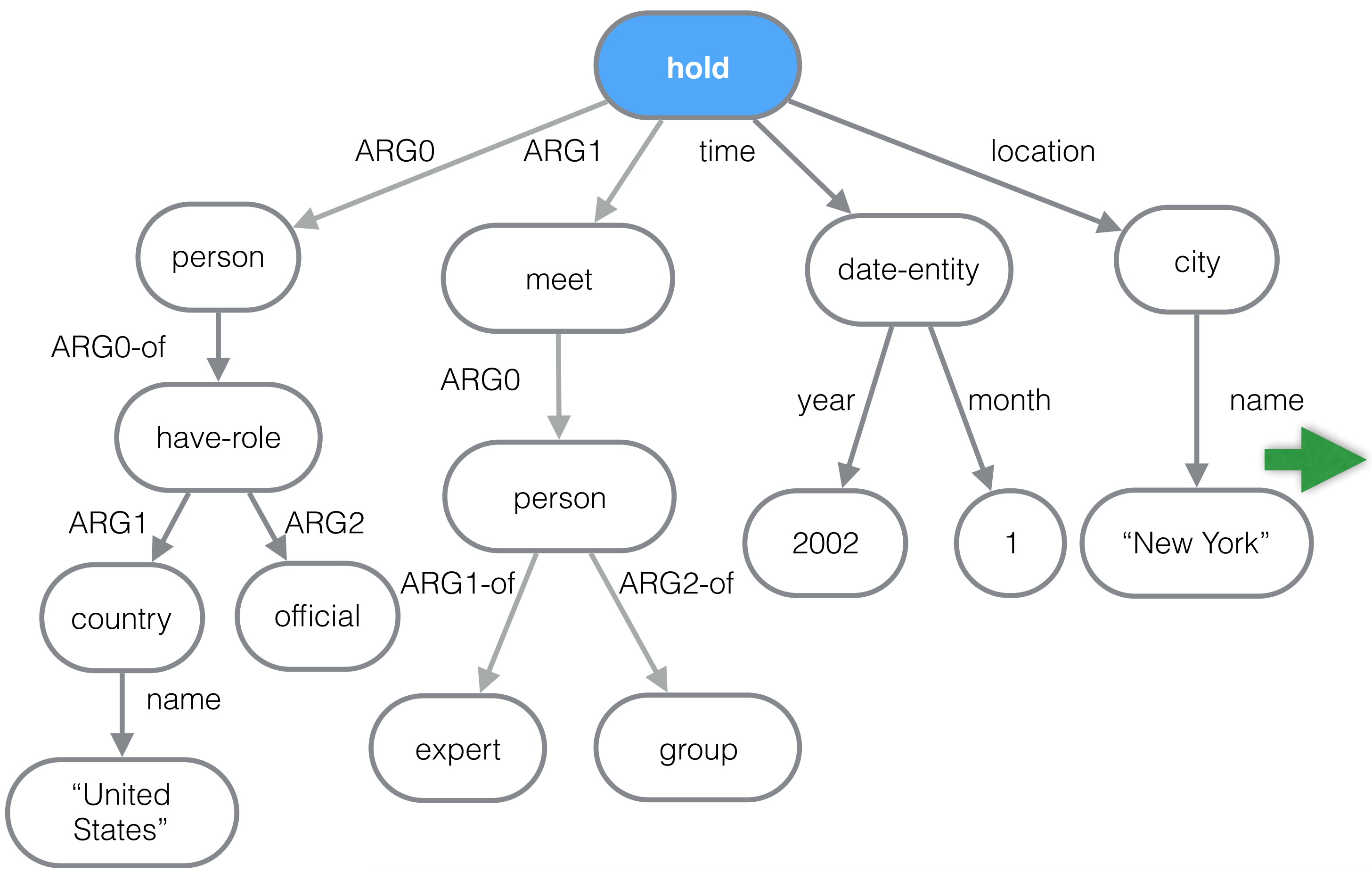
$$\mathbf{a}_i = \operatorname{soft\,max}\left(\mathbf{f}_i\left(\mathbf{h}^{(\mathbf{s})}, \mathbf{h}_i\right)\right)$$

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



Pre-processing

Linearization —> Anonymization

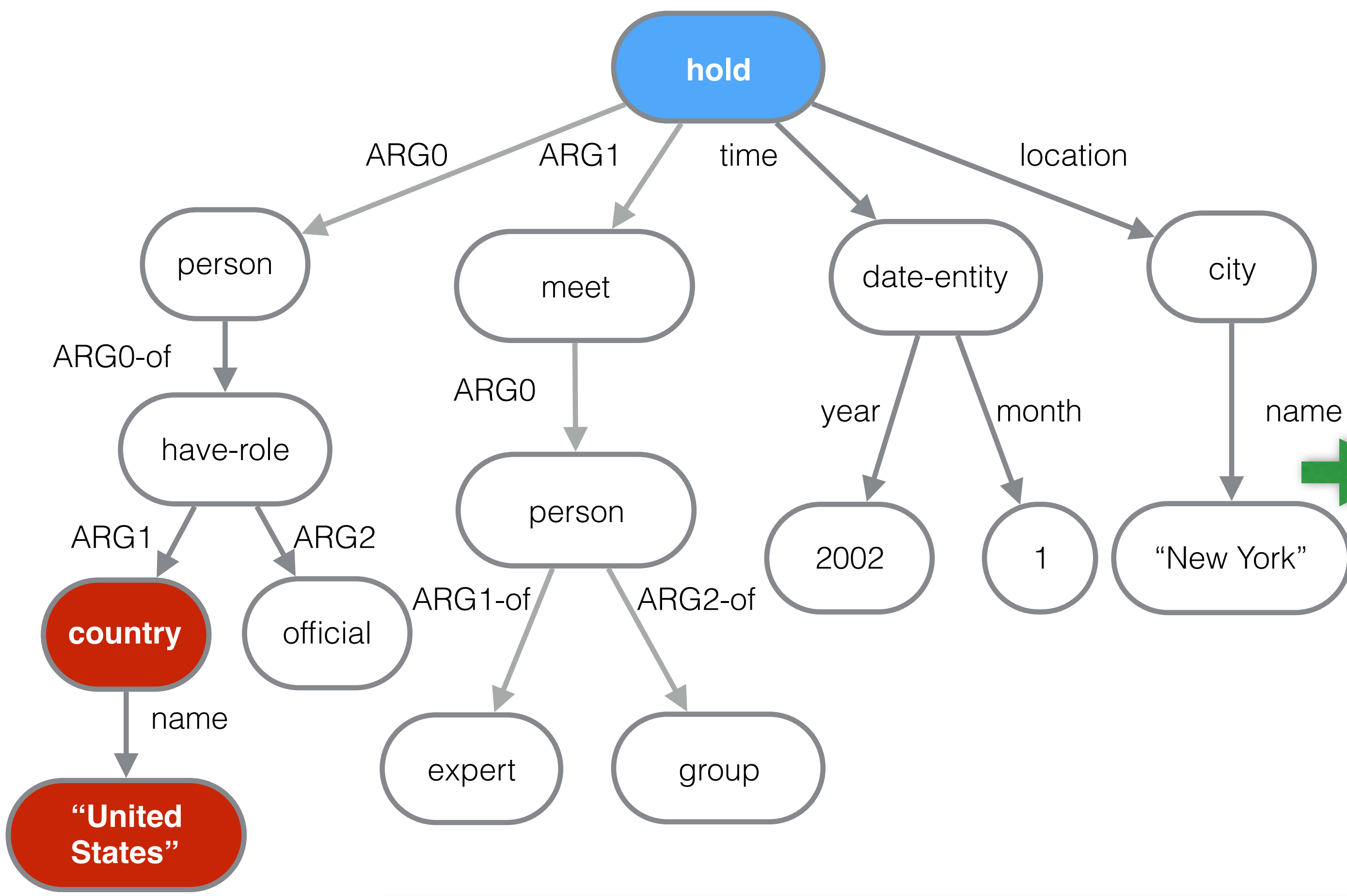



```
hold
    :ARG0 (person
              :ARG0-of (have-role
                  :ARG1 loc_0
                  :ARG2 official)
          )
    :ARG1 (meet
              :ARG0 (person
                  :ARG1-of expert
                  :ARG2-of group)
          )
    :time (date-entity year_0 month_0)
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

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

Pre-processing

Linearization —> Anonymization

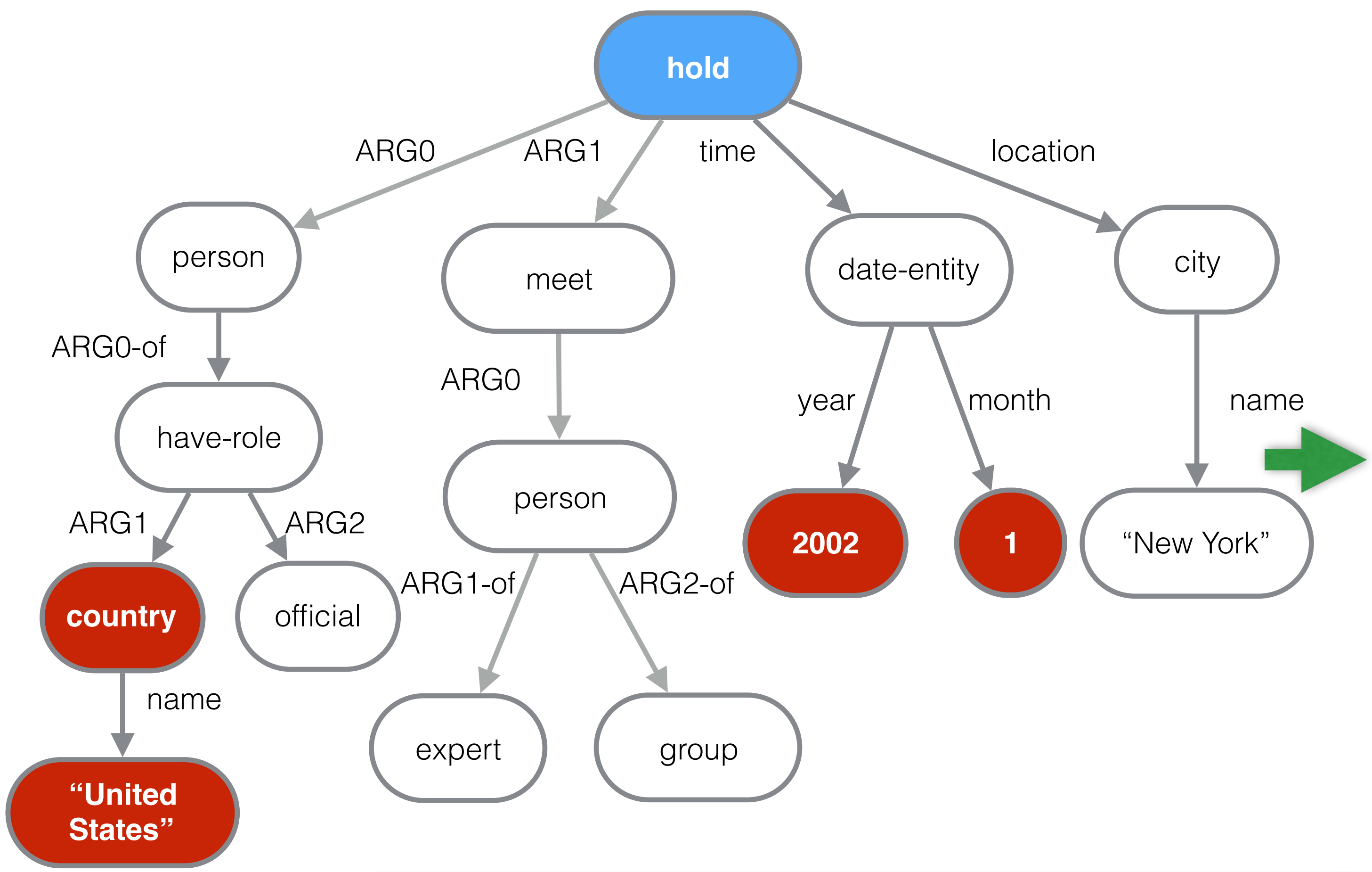



```
hold
    :ARG0 (person
              :ARG0-of (have-role
                 :ARG1 loc_0
                 :ARG2 official)
           )
    :ARG1 (meet
              :ARG0 (person
                 :ARG1-of expert
                 :ARG2-of group)
           )
    :time (date-entity year_0 month_0)
  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

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

Pre-processing

Linearization —> Anonymization

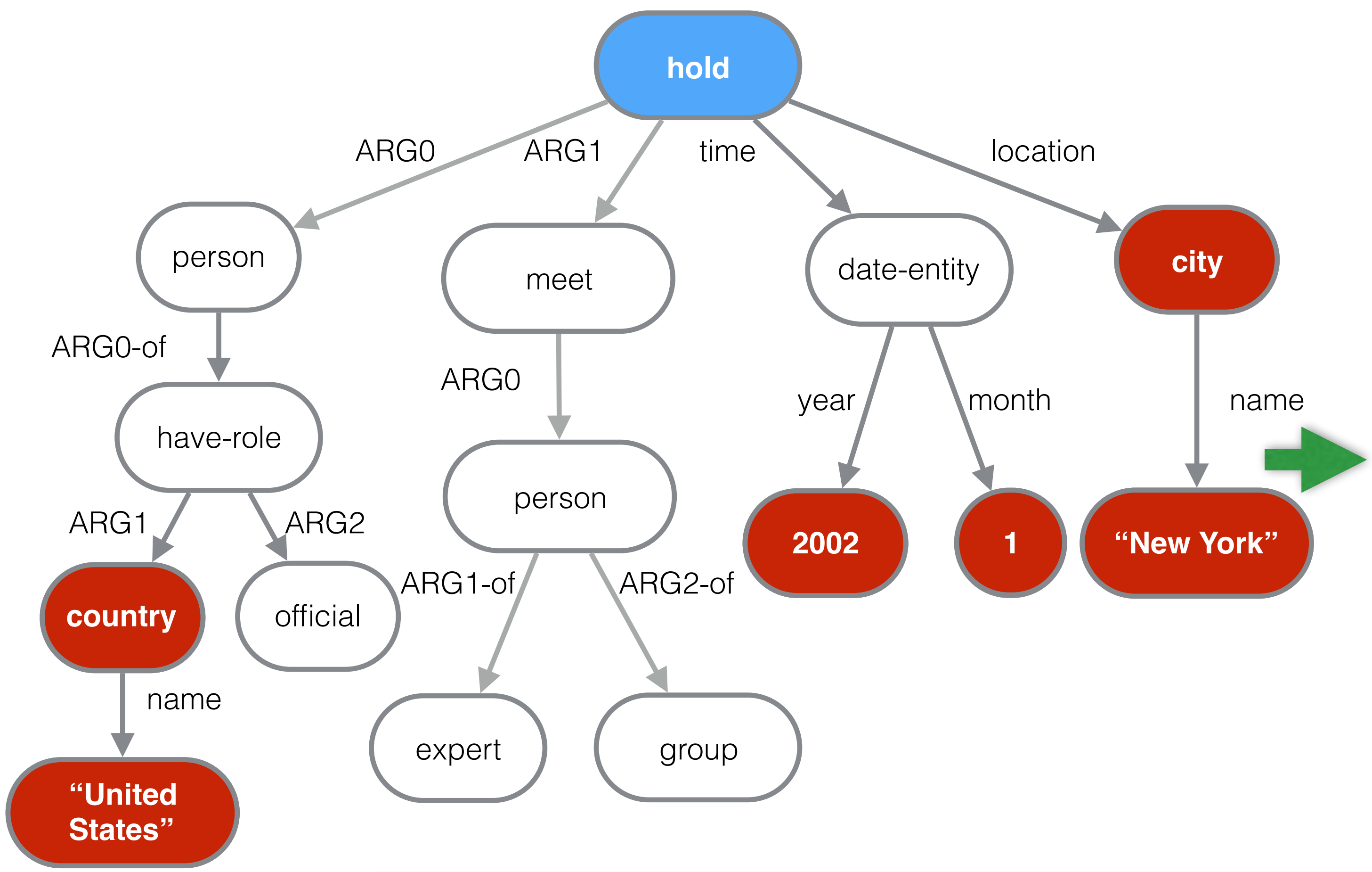



```
hold
    :ARG0 (person
            :ARG0-of (have-role
                :ARG1 loc_0
                :ARG2 official)
            )
    :ARG1 (meet
            :ARG0 (person
                :ARG1-of expert
                :ARG2-of group)
            )
    :time (date-entity year_0 month_0)
    :location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

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

Pre-processing

Linearization —> Anonymization

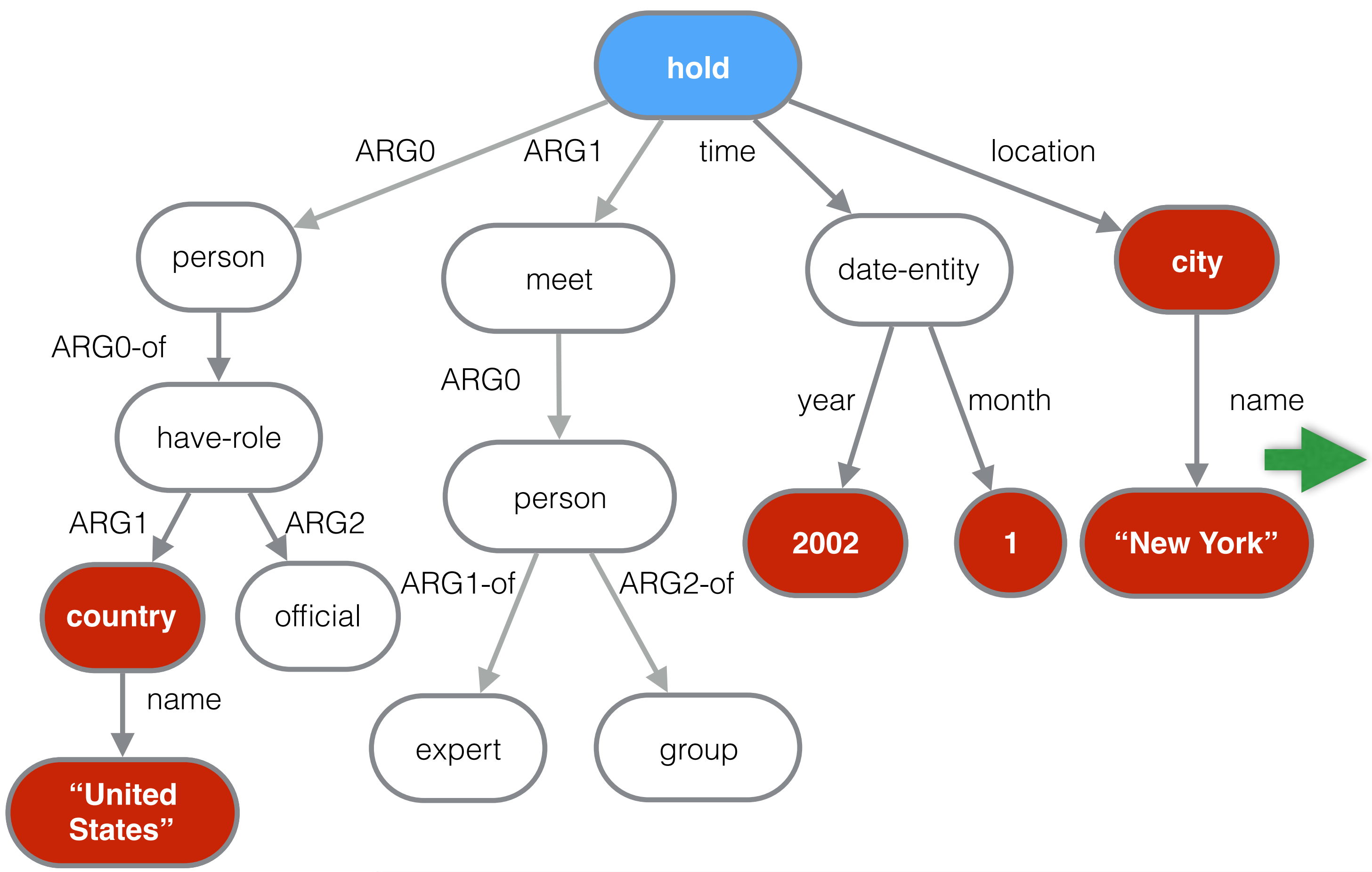



```
hold
    :ARG0 (person
           :ARG0-of (have-role
              :ARG1 loc_0
              :ARG2 official)
          )
    :ARG1 (meet
           :ARG0 (person
              :ARG1-of expert
              :ARG2-of group)
          )
    :time (date-entity year_0 month_0)
    :location loc_1
```

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

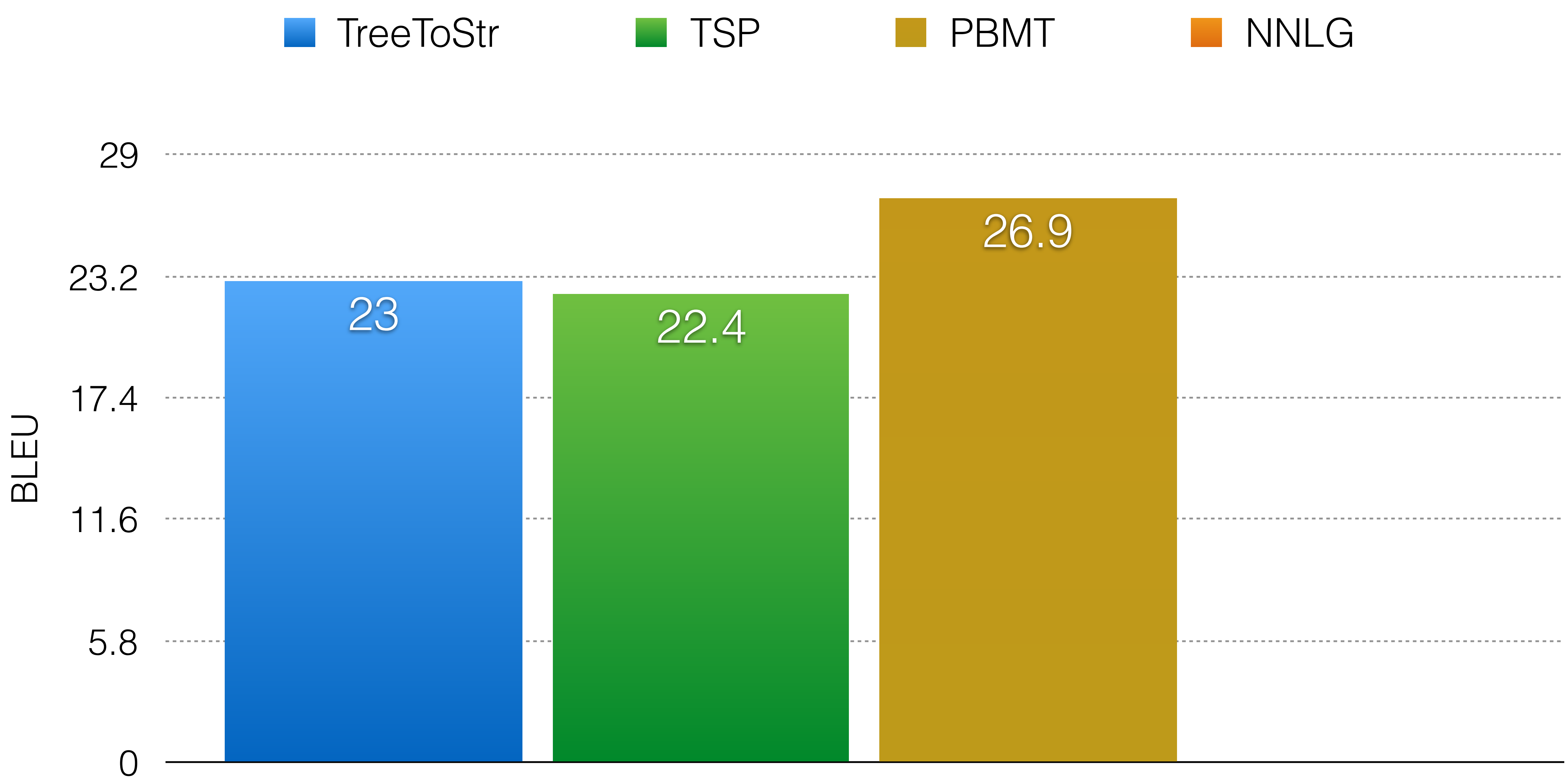
Evaluation

- BLEU n-gram precision
  (Papineni et al., ACL 2002)

First Attempt

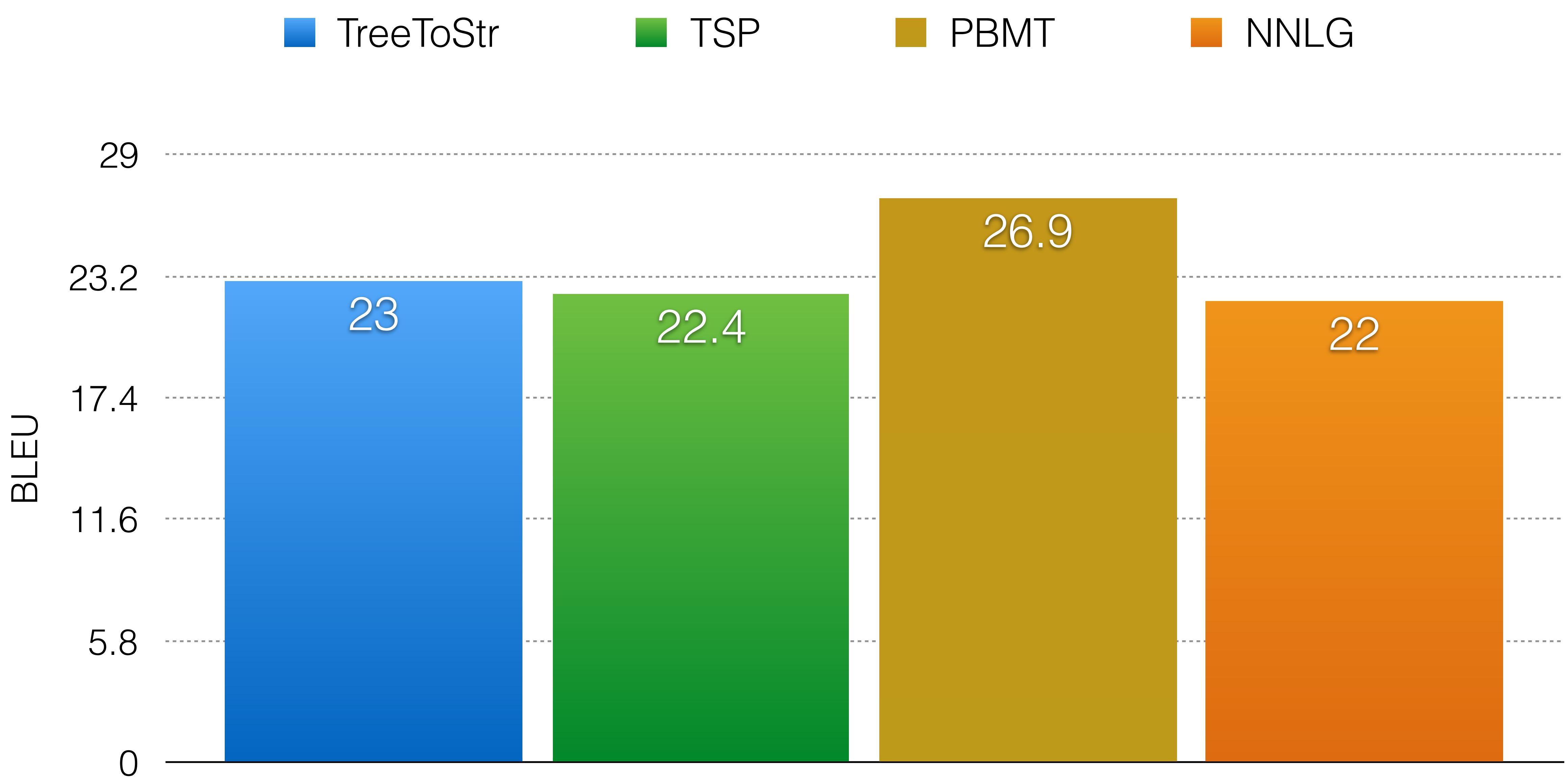
TreeToStr: Flanigan et al, NAACL 2016
TSP: Song et al, EMNLP 2016
PBMT: Pourdamaghani and Knight, INLG 2016

First Attempt



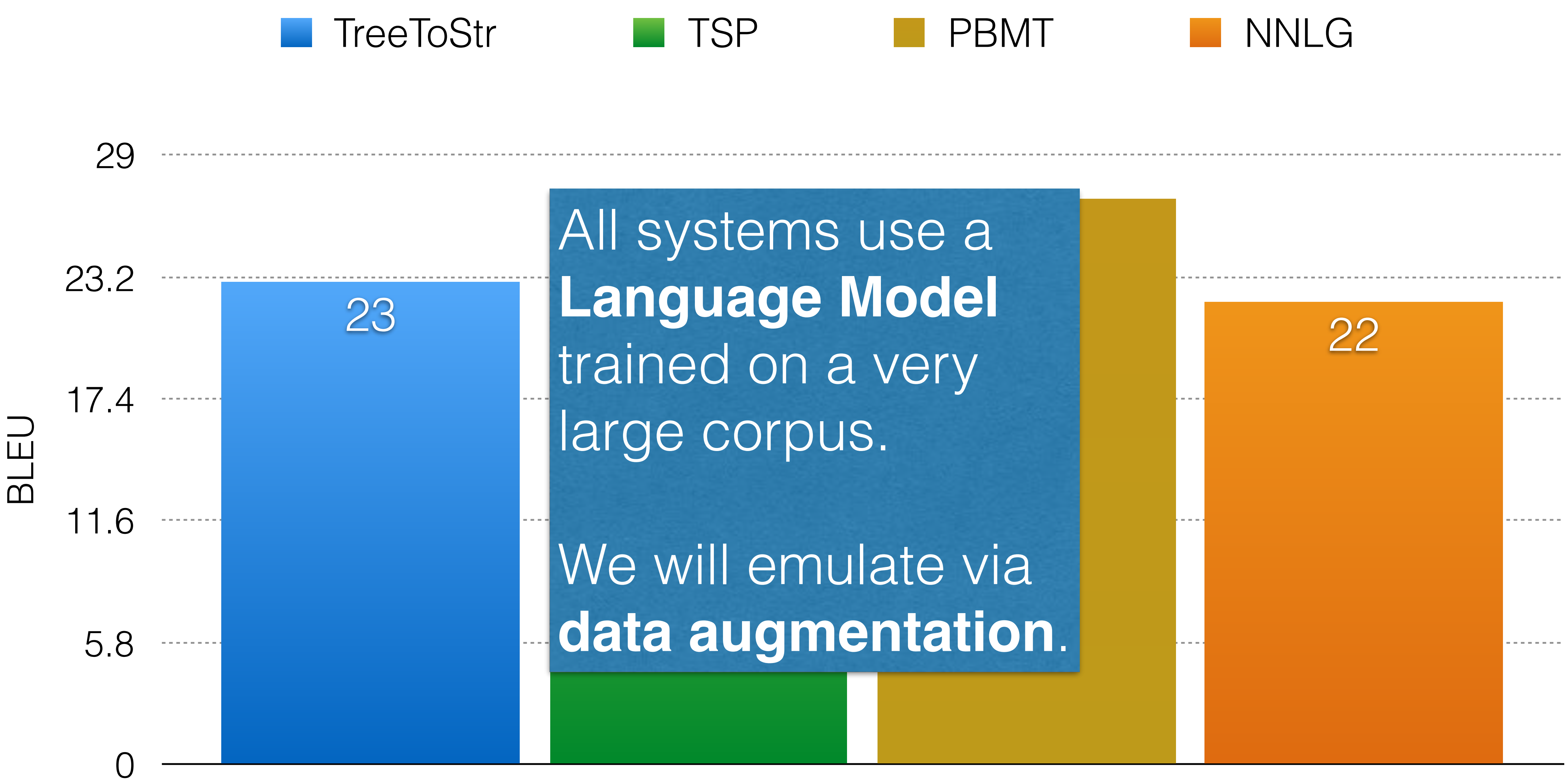
TreeToStr: Flanigan et al, NAACL 2016
TSP: Song et al, EMNLP 2016
PBMT: Pourdamaghani and Knight, INLG 2016

First Attempt



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

First Attempt



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
```

Reference

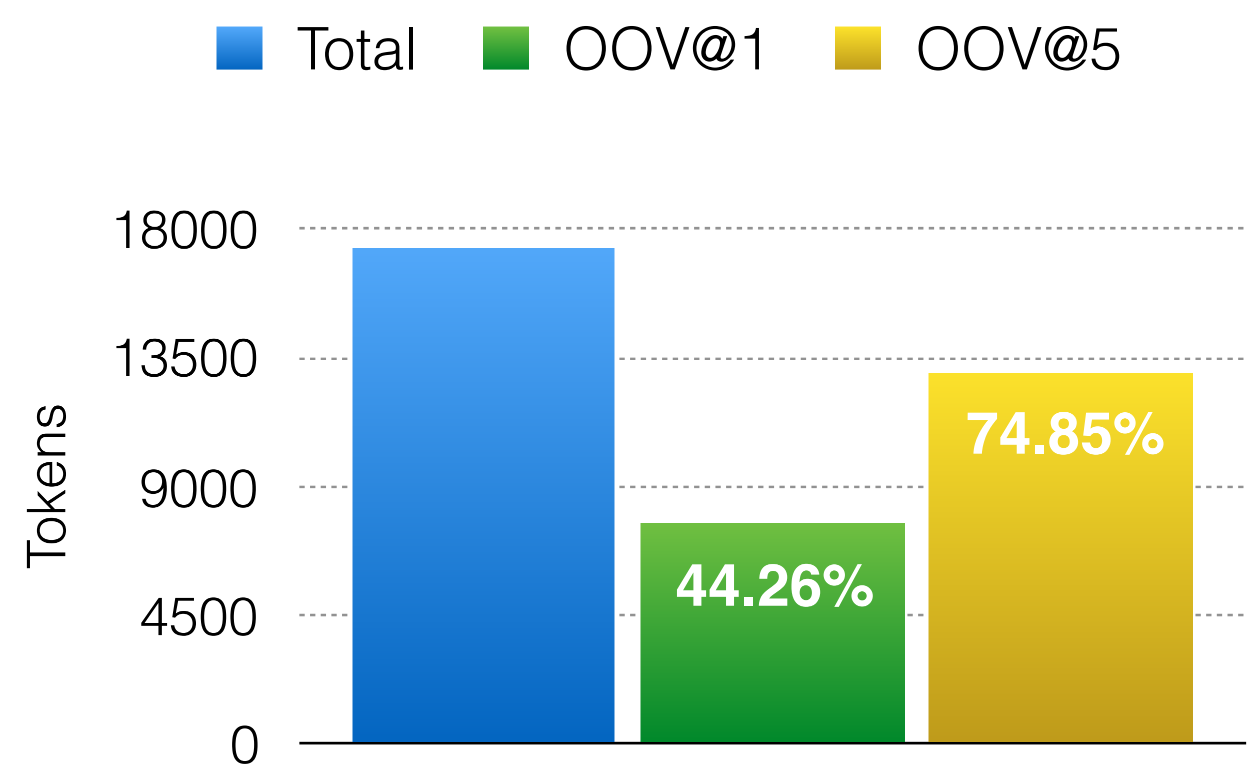
US officials held an expert group meeting in January 2002 in New York .

Prediction

United States officials held held a meeting in January 2002 .

- Repetition
- Coverage

  a) Sparsity



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

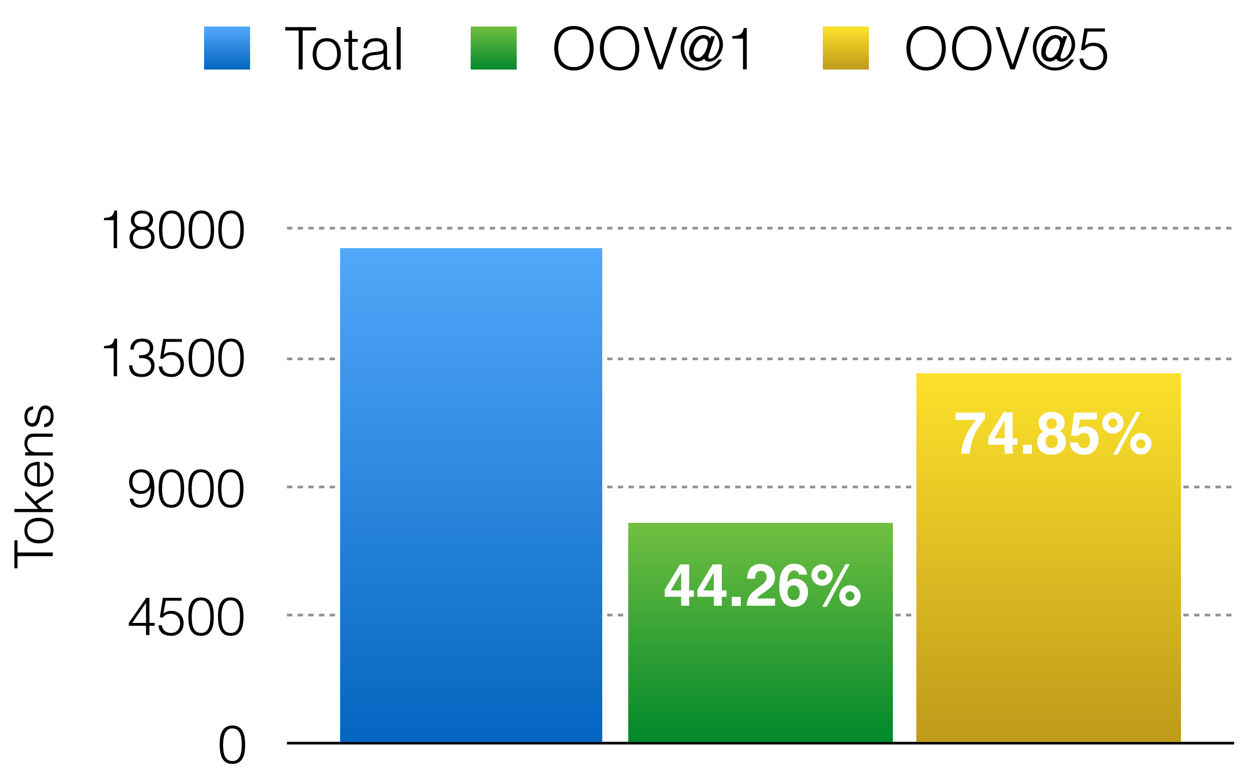
United States officials held held a meeting in January 2002 .

- 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**

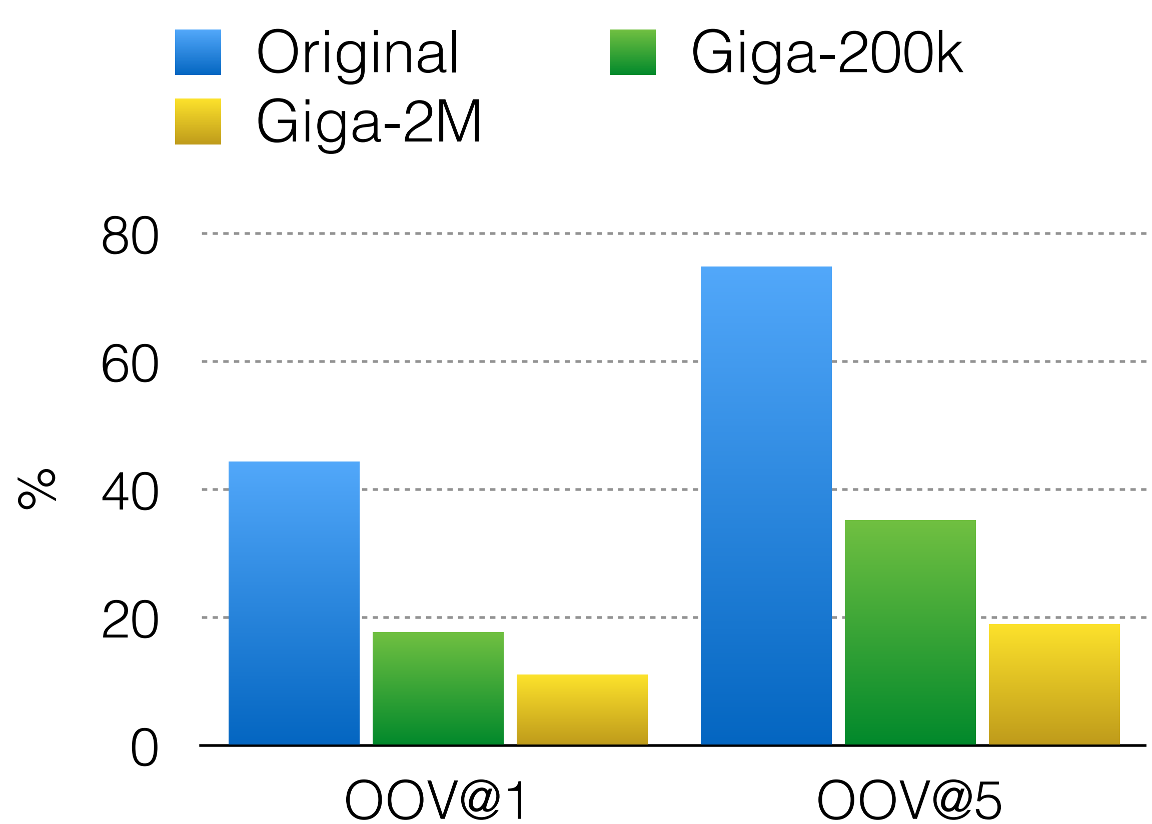
Gigaword: ~183M sentences ***only***

Data Augmentation

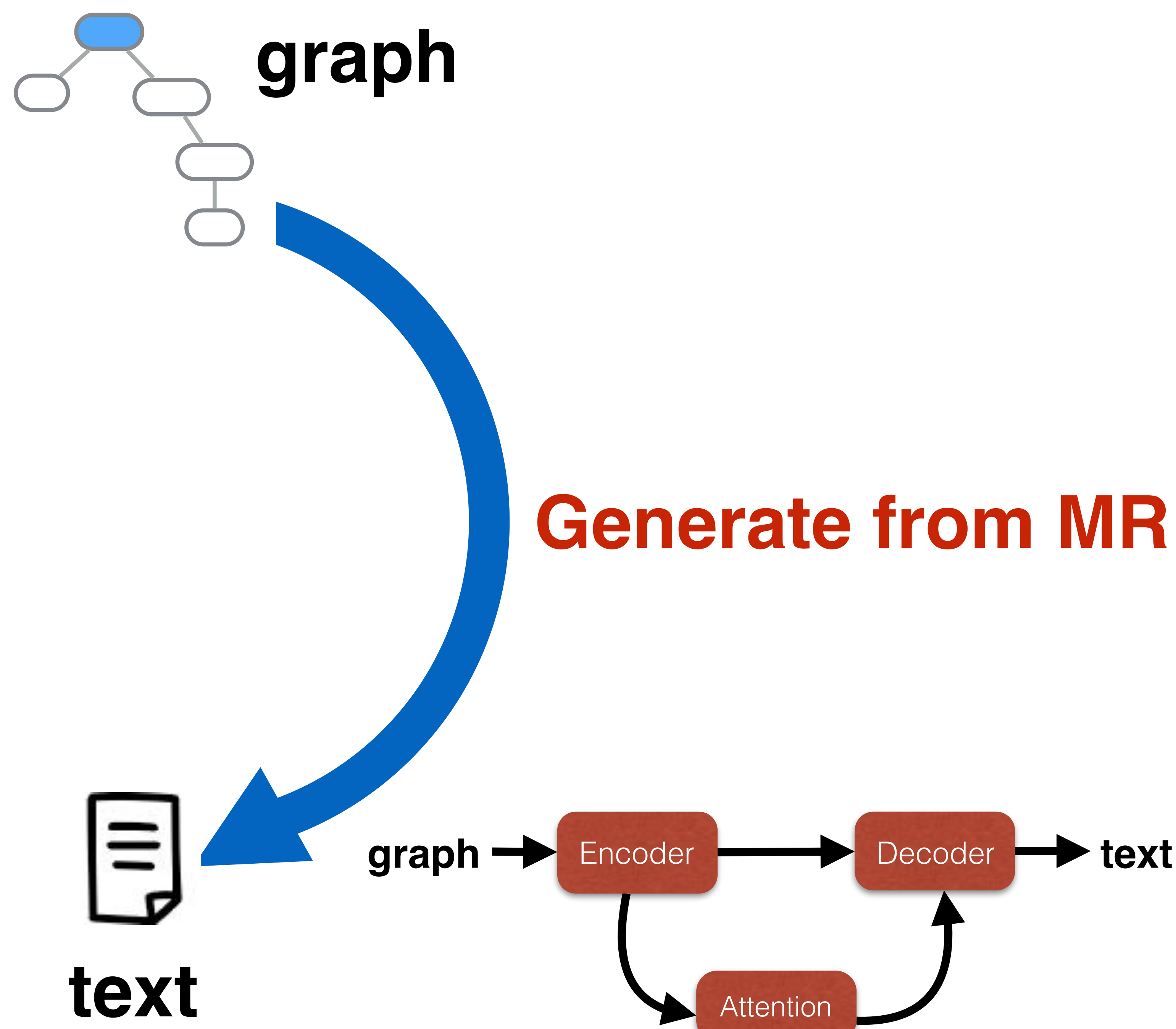
Original Dataset: ~16k graph-sentence **pairs**

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

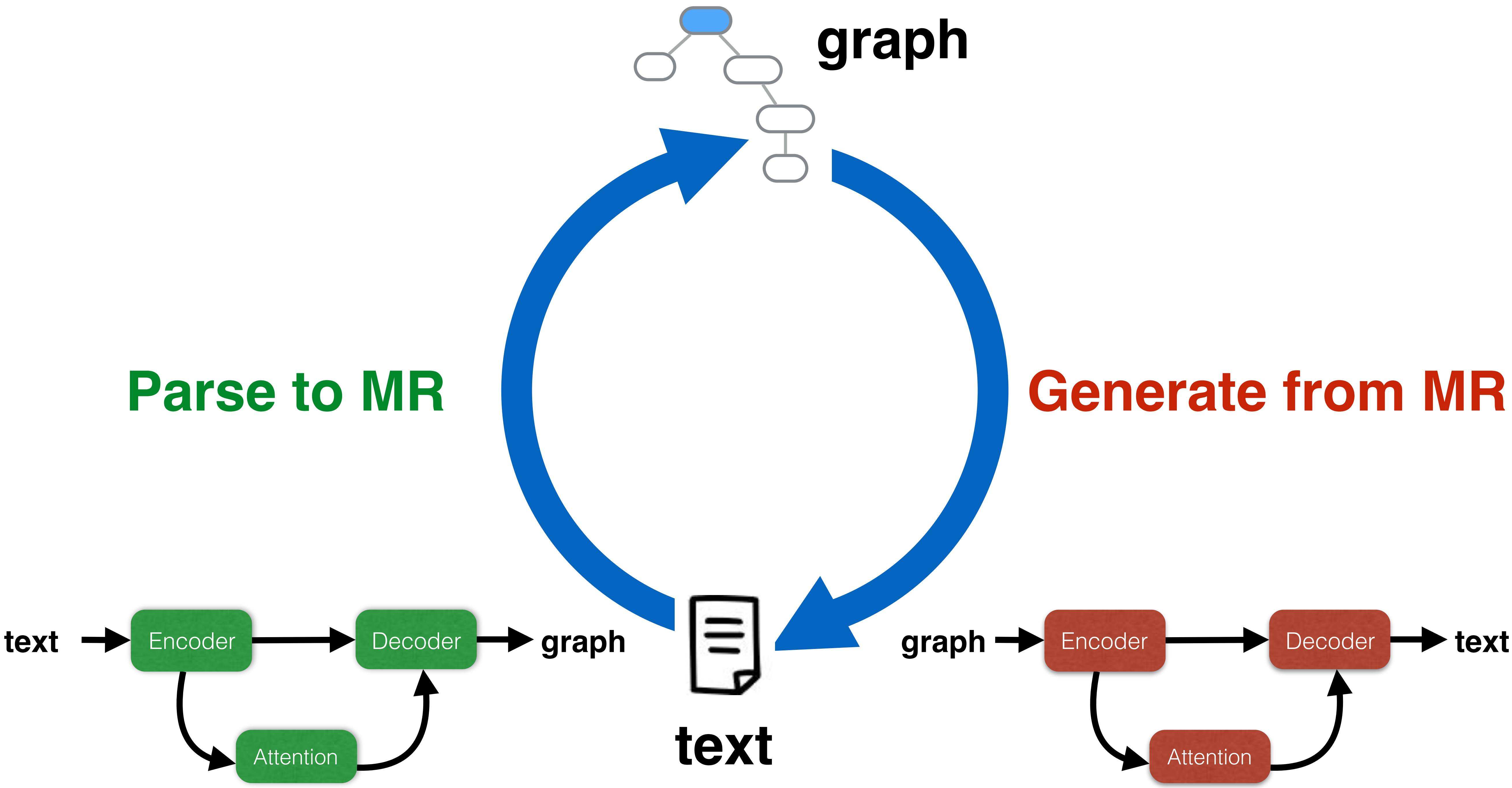
Sample sentences with vocabulary overlap



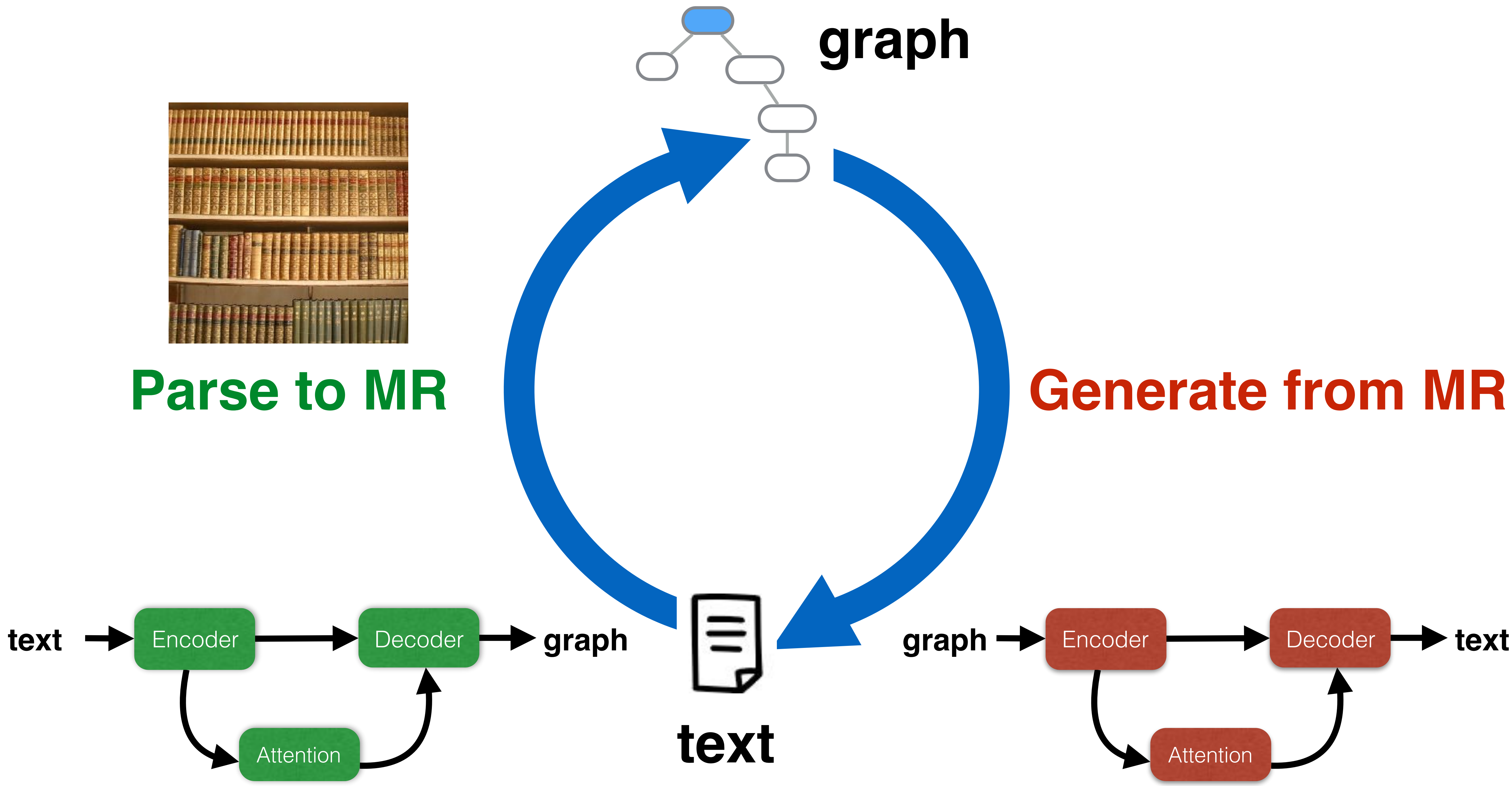
Data Augmentation



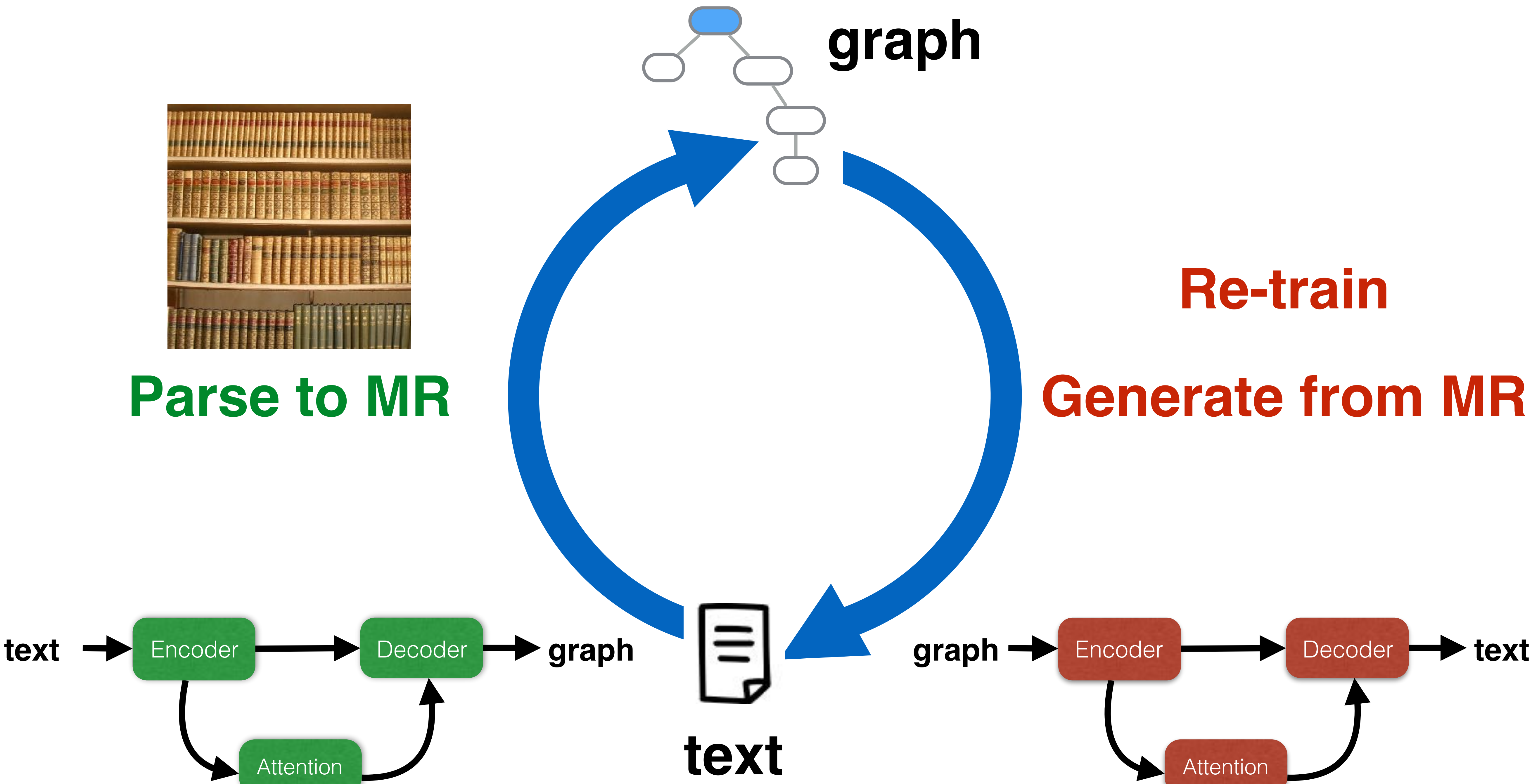
Data Augmentation



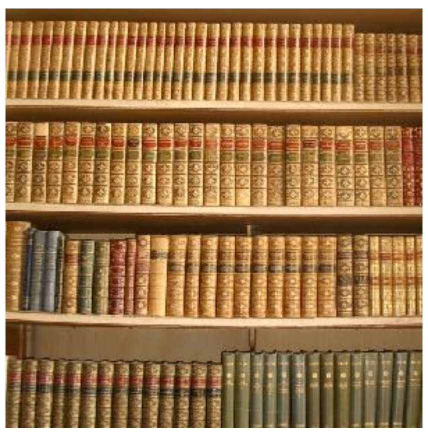
Data Augmentation

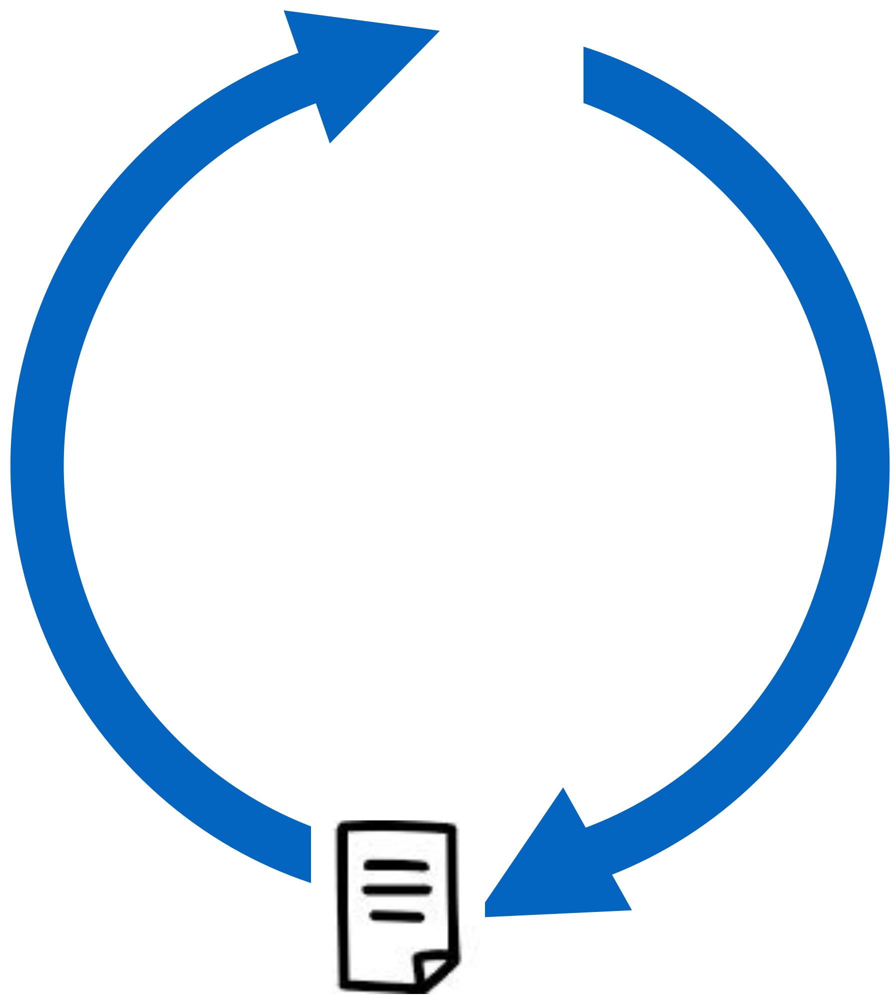


Data Augmentation



Data Augmentation





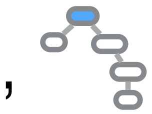
Paired Training

Paired Training

Train MR Parser **P** on Original Dataset

Paired Training

Train MR Parser **P** on Original Dataset

( , )

for i = 0 … N

Paired Training

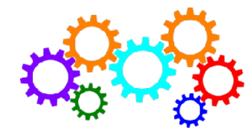
Train MR Parser **P** on Original Dataset

for i = 0 … N

$\mathbf{S_i}$ =Sample **k** $\mathbf{10^i}$ sentences from Gigaword

Paired Training

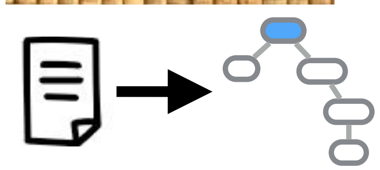
Train MR Parser **P** on Original Dataset



for i = 0 … N

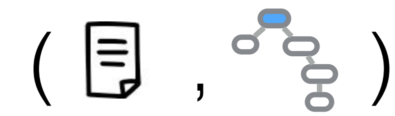
S_i =Sample **$k\ 10^i$** sentences from Gigaword

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



Paired Training

Train MR Parser **P** on Original Dataset

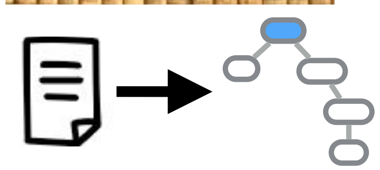


for i = 0 … N

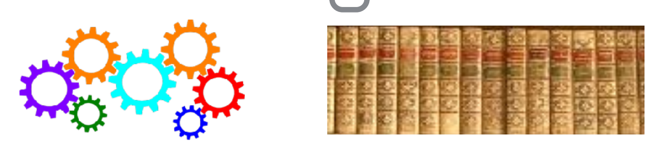
Self-train Parser

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

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



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



Paired Training

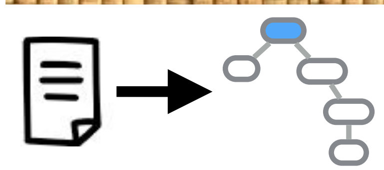
Train MR Parser **P** on Original Dataset

for i = 0 … N

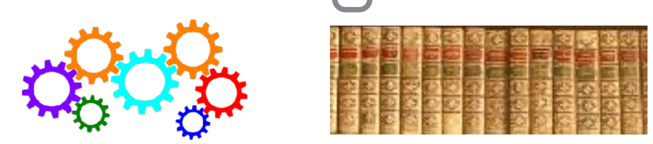
Self-train Parser

$\mathbf{S_i}$ =Sample $\mathbf{k\ 10^i}$ sentences from Gigaword

Parse $\mathbf{S_i}$ sentences with **P**



Re-train MR Parser **P** on $\mathbf{S_i}$



Paired Training

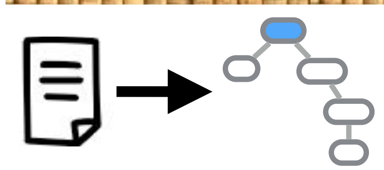
Train MR Parser **P** on Original Dataset

for i = 0 … N

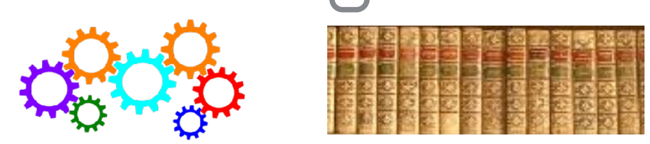
Self-train Parser

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

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



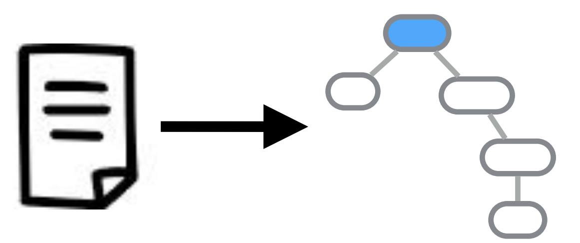
Re-train MR Parser **P** on **S_i**



Train Generator **G** on **S_N**

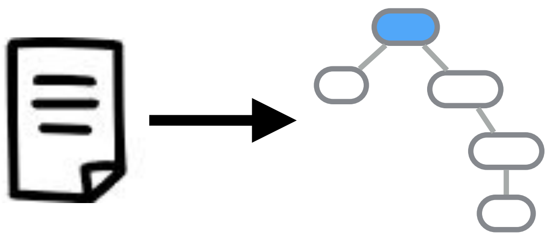
Training MR Parser

Train **P** on
Original Dataset



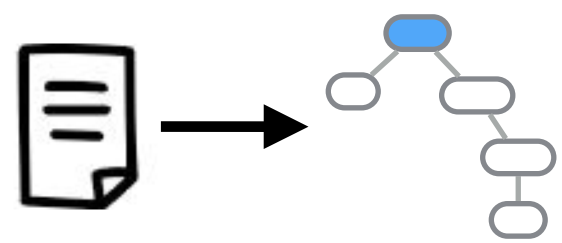
Training MR Parser

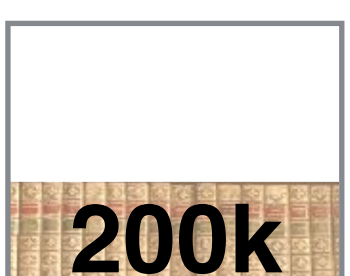
Train **P** on Original Dataset



Training MR Parser

Train **P** on Original Dataset → Sample **S_1=200k** sentences from Gigaword

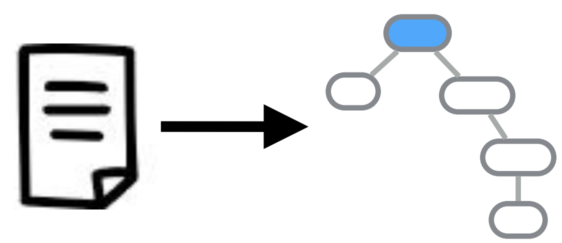




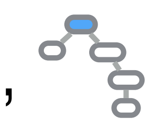
Training MR Parser

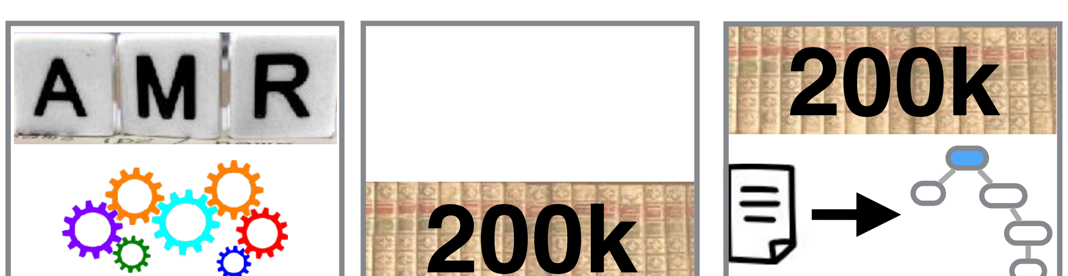
Train **P** on Original Dataset

Sample **S_1=200k** sentences from Gigaword

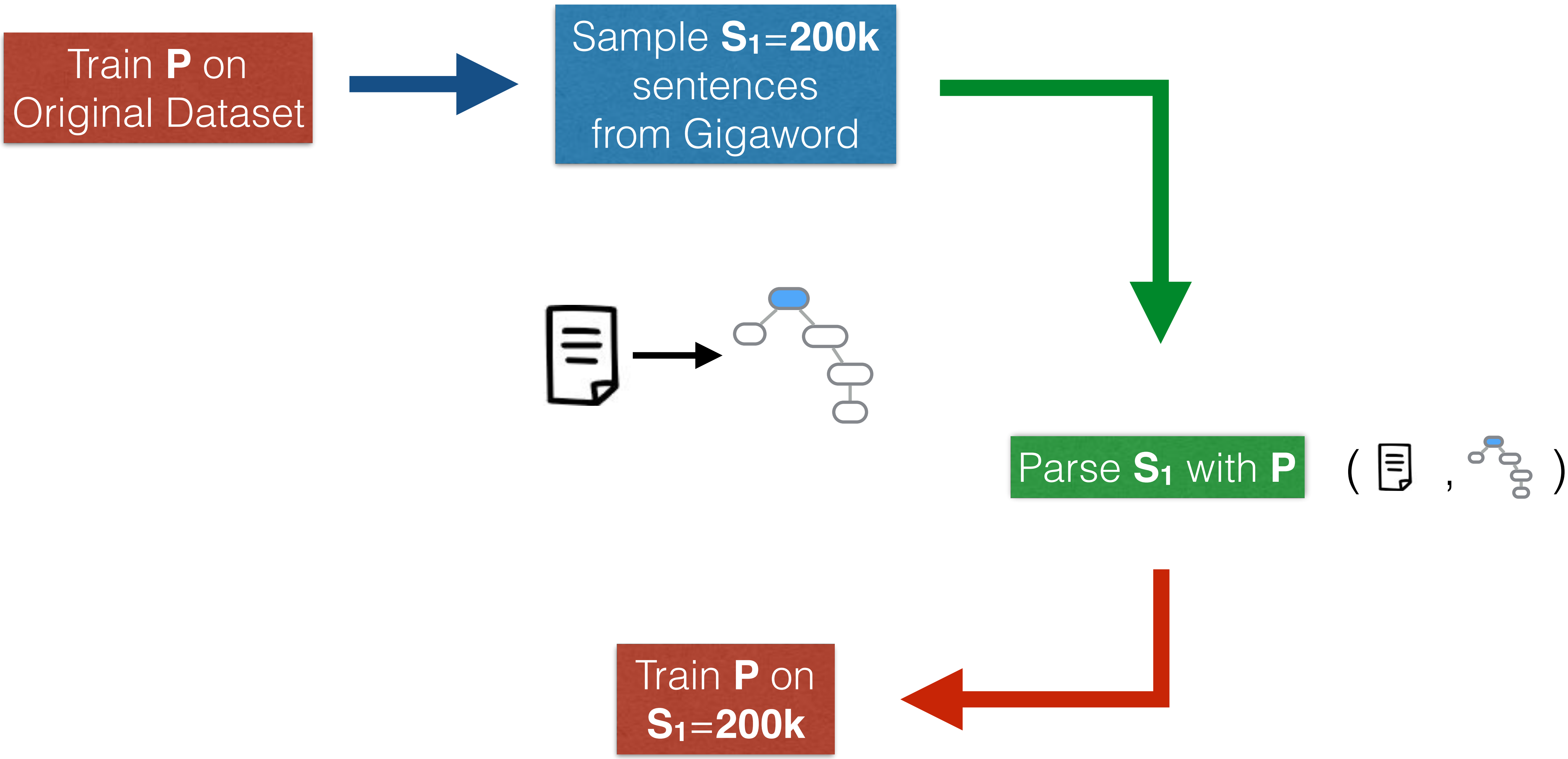


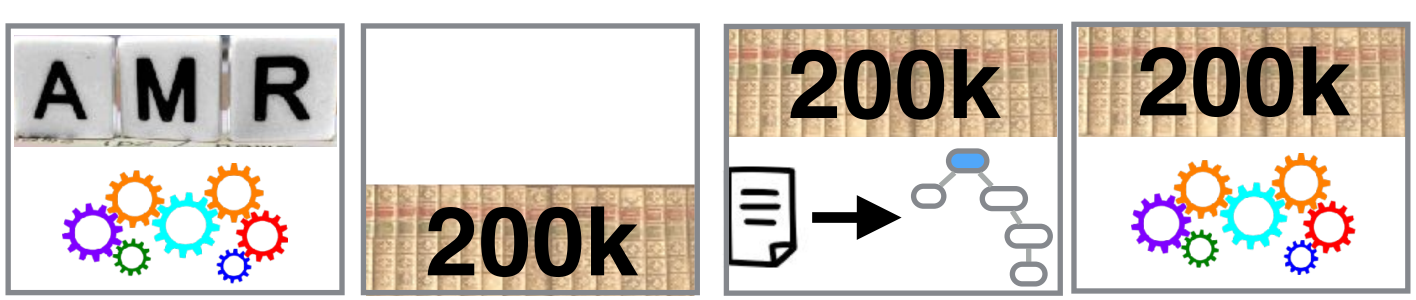
Parse **S_1** with **P**

( , )



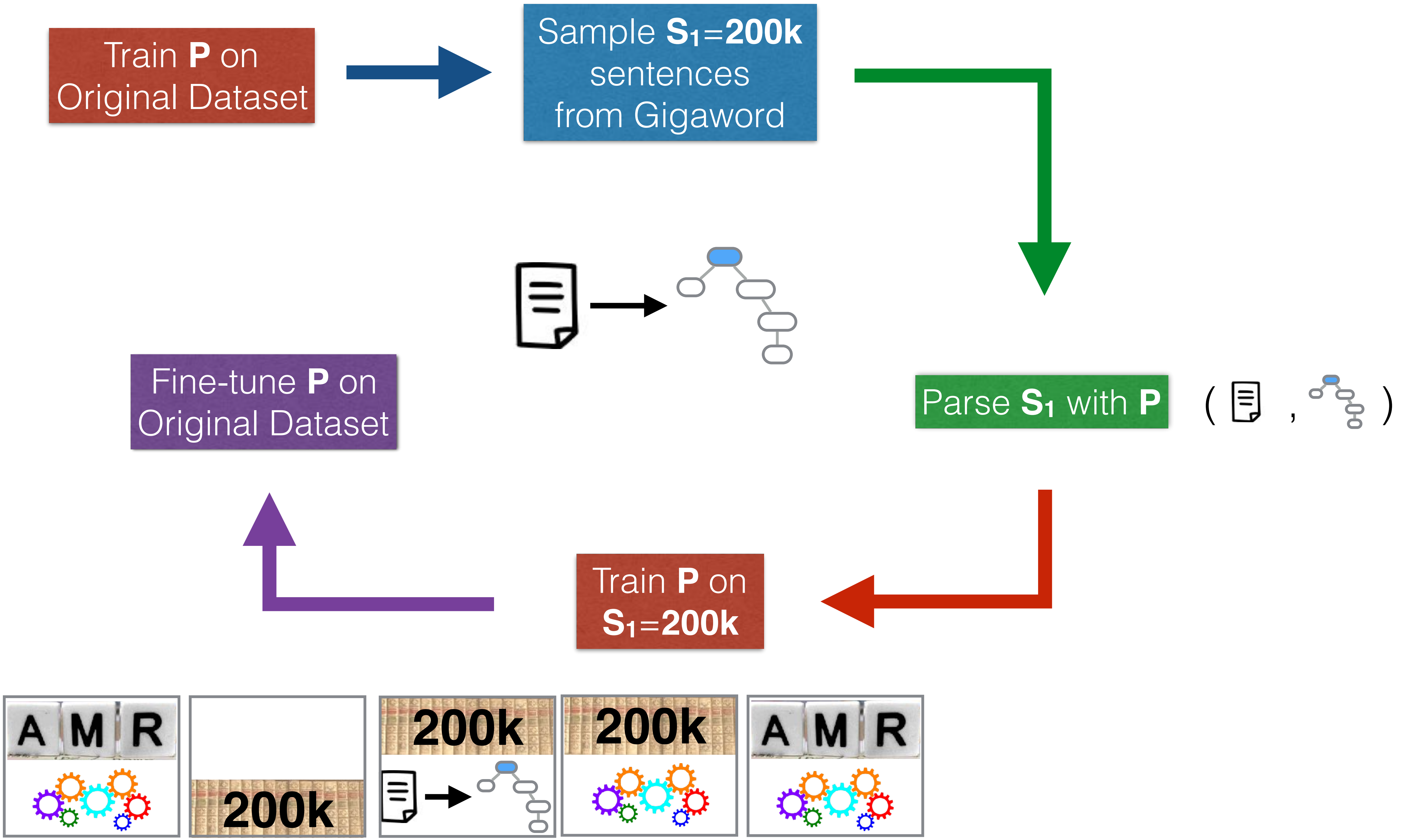
Training MR Parser





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

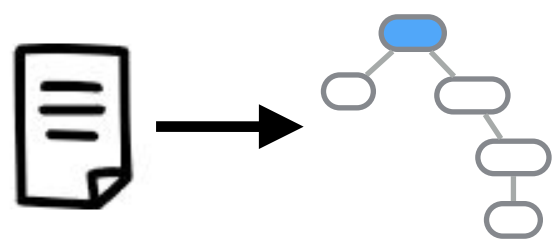
Training MR Parser



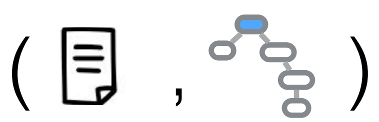
Fine-tune: init parameters from previous step and train on Original Dataset

Training MR Parser

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

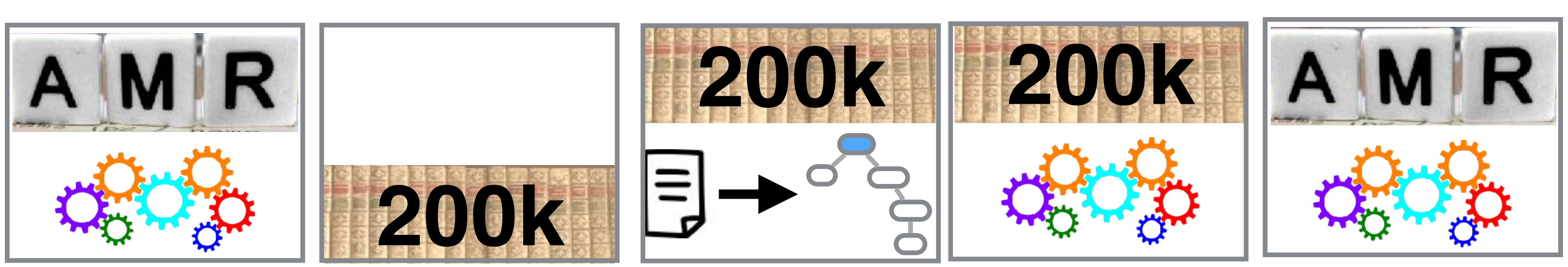


Parse $\mathbf{S_2}$ with **P**



Fine-tune **P** on Original Dataset

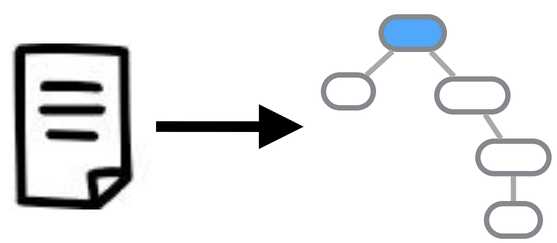
Train **P** on $\mathbf{S_2}$=**2M**



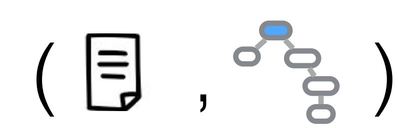
Fine-tune: init parameters from previous step and train on Original Dataset

Training MR Parser

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

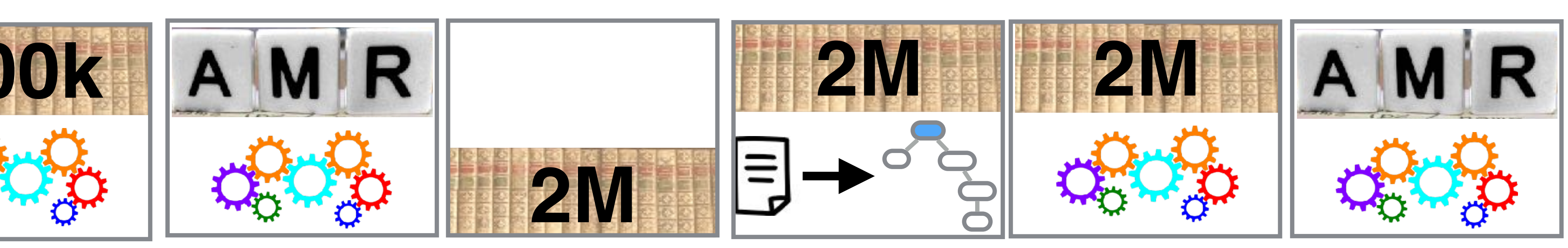


Parse $\mathbf{S_2}$ with **P**



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

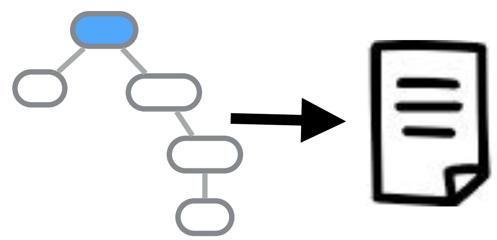
Fine-tune **P** on Original Dataset



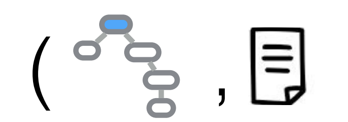
Fine-tune: init parameters from previous step and train on Original Dataset

Training MR Generator

Sample $\mathbf{S_3}$=**2M** sentences from Gigaword

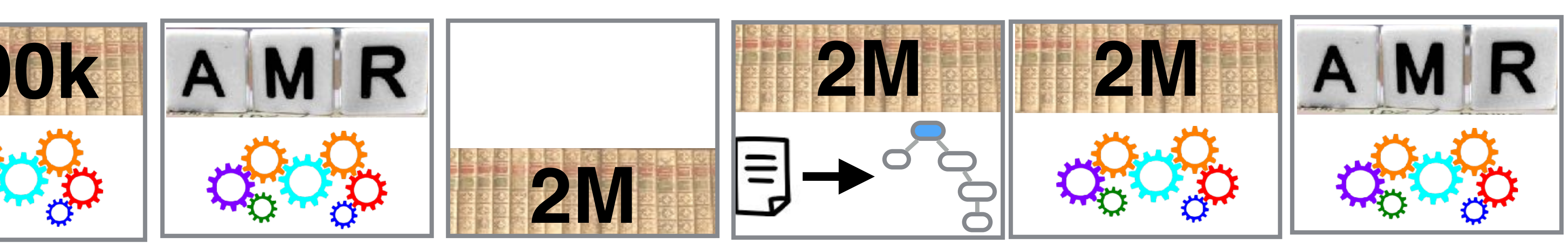


Parse $\mathbf{S_3}$ with **P**



Fine-tune **G** on Original Dataset

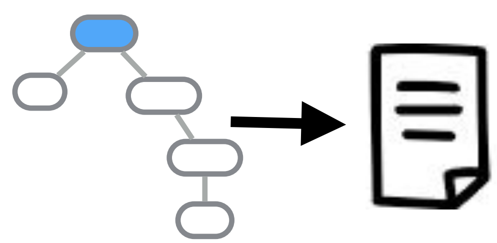
Train **G** on $\mathbf{S_3}$=**2M**



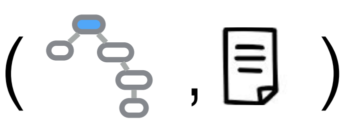
Fine-tune: init parameters from previous step and train on Original Dataset

Training MR Generator

Sample S_3=2M sentences from Gigaword

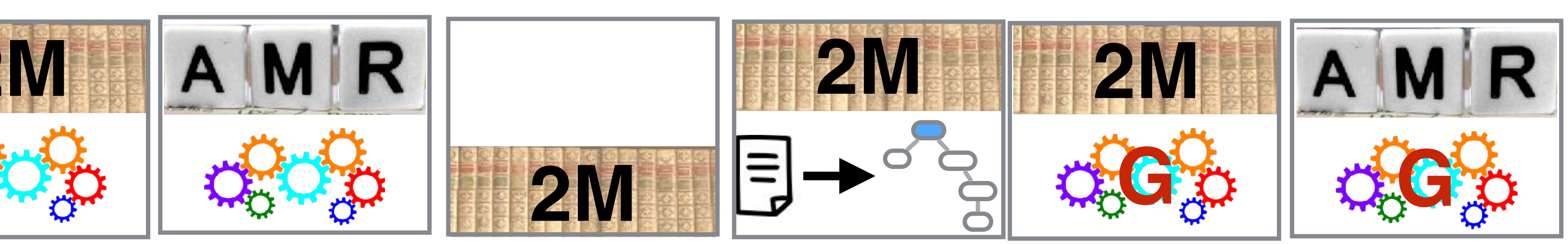


Parse S_3 with **P**



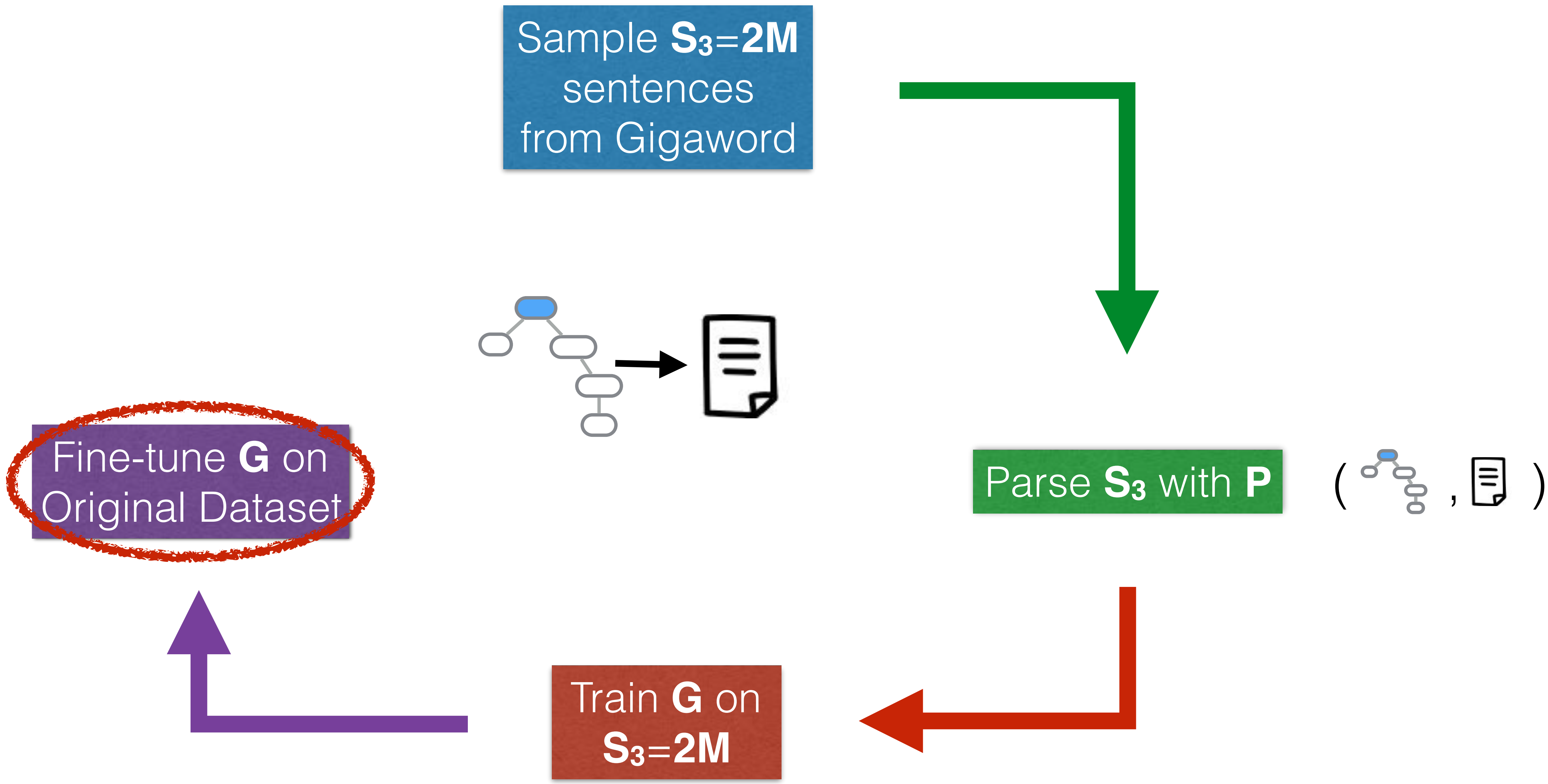
Fine-tune **G** on Original Dataset

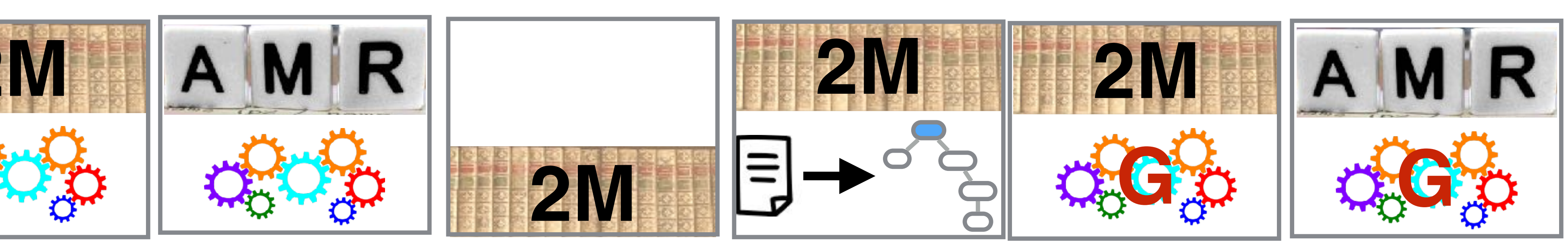
Train **G** on S_3=2M



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

Training MR Generator

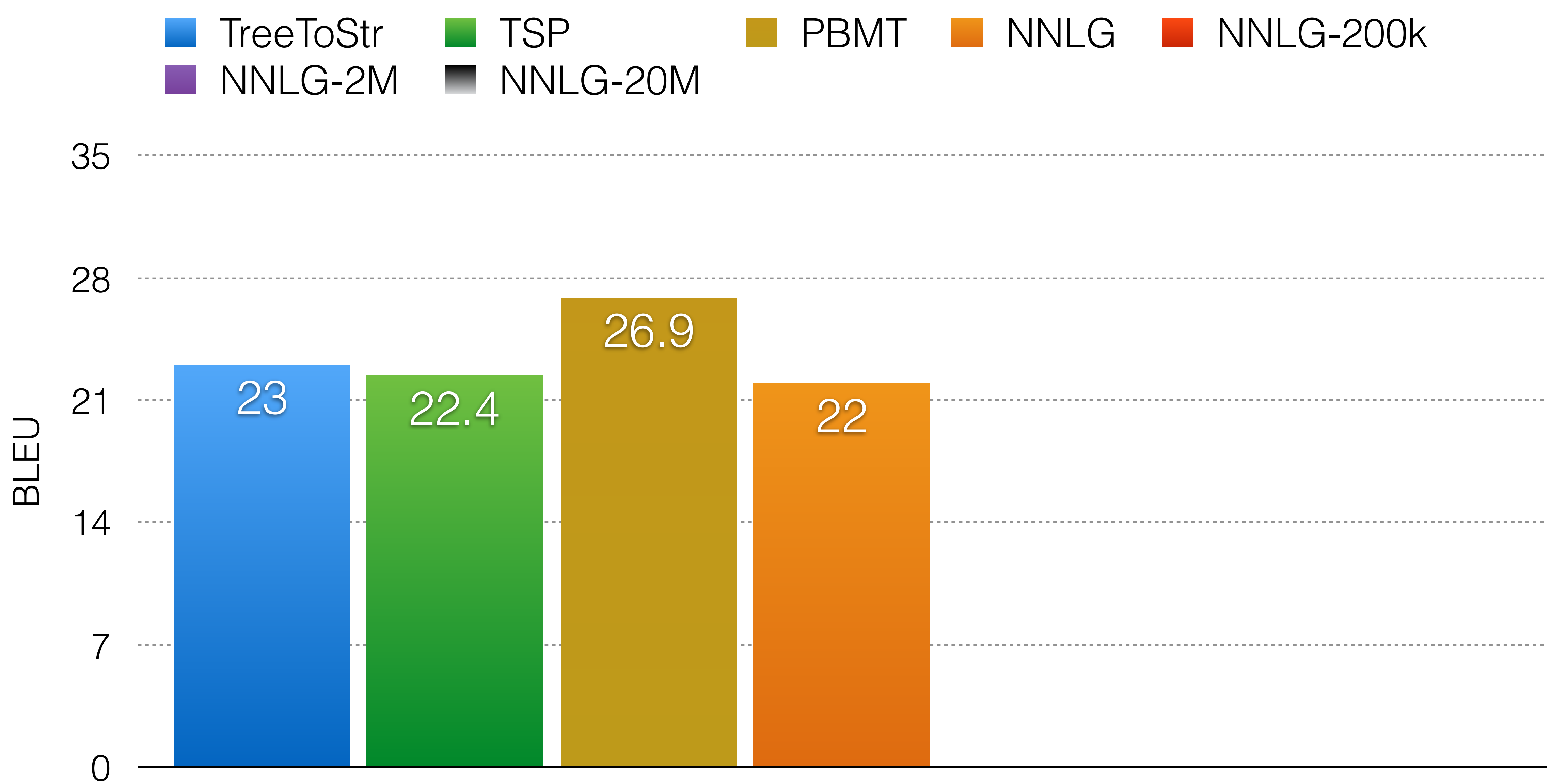




Final Results

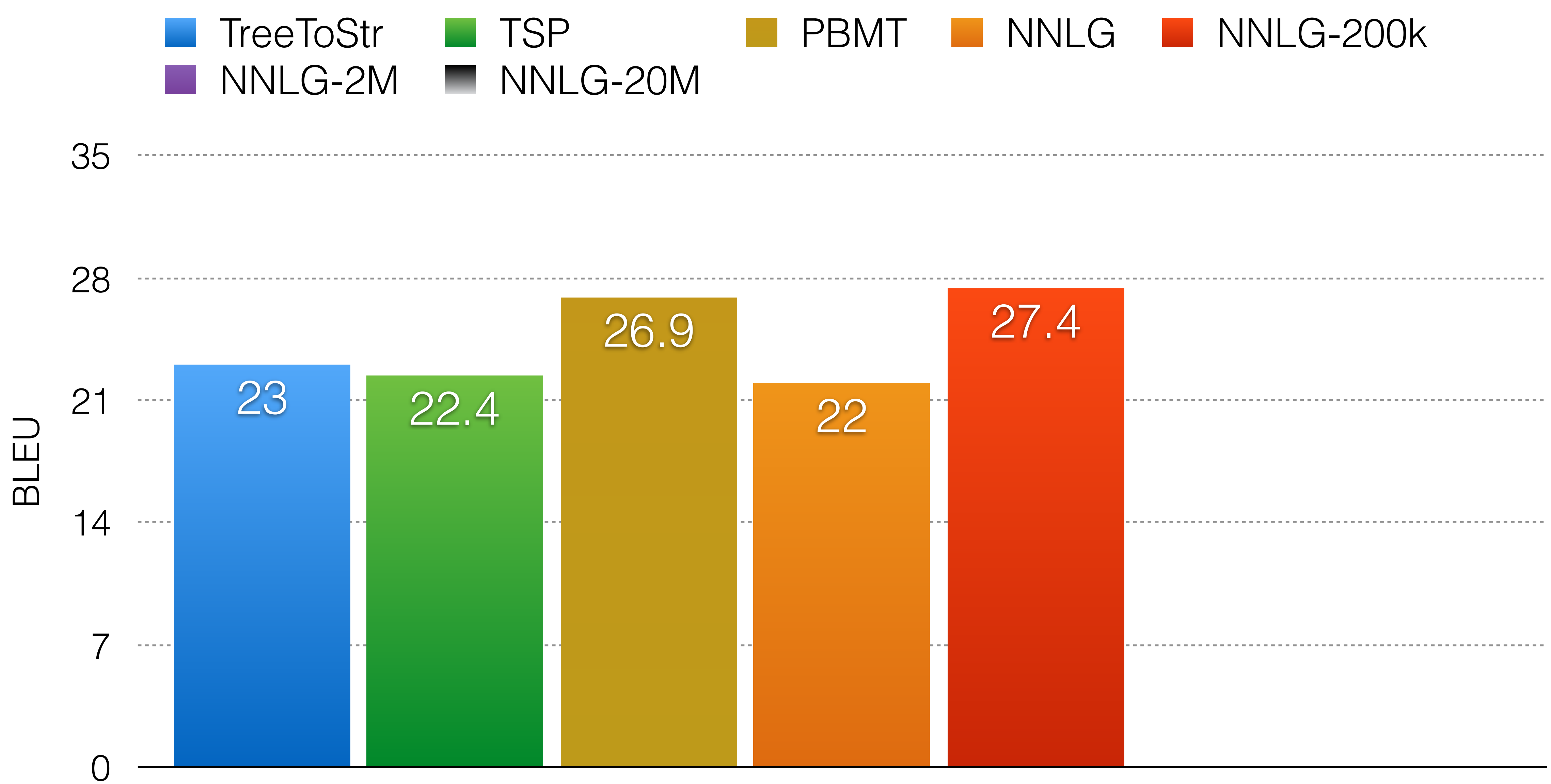
TreeToStr: Flanigan et al, NAACL 2016
TSP: Song et al, EMNLP 2016
PBMT: Pourdamaghani and Knight, INLG 2016

Final Results



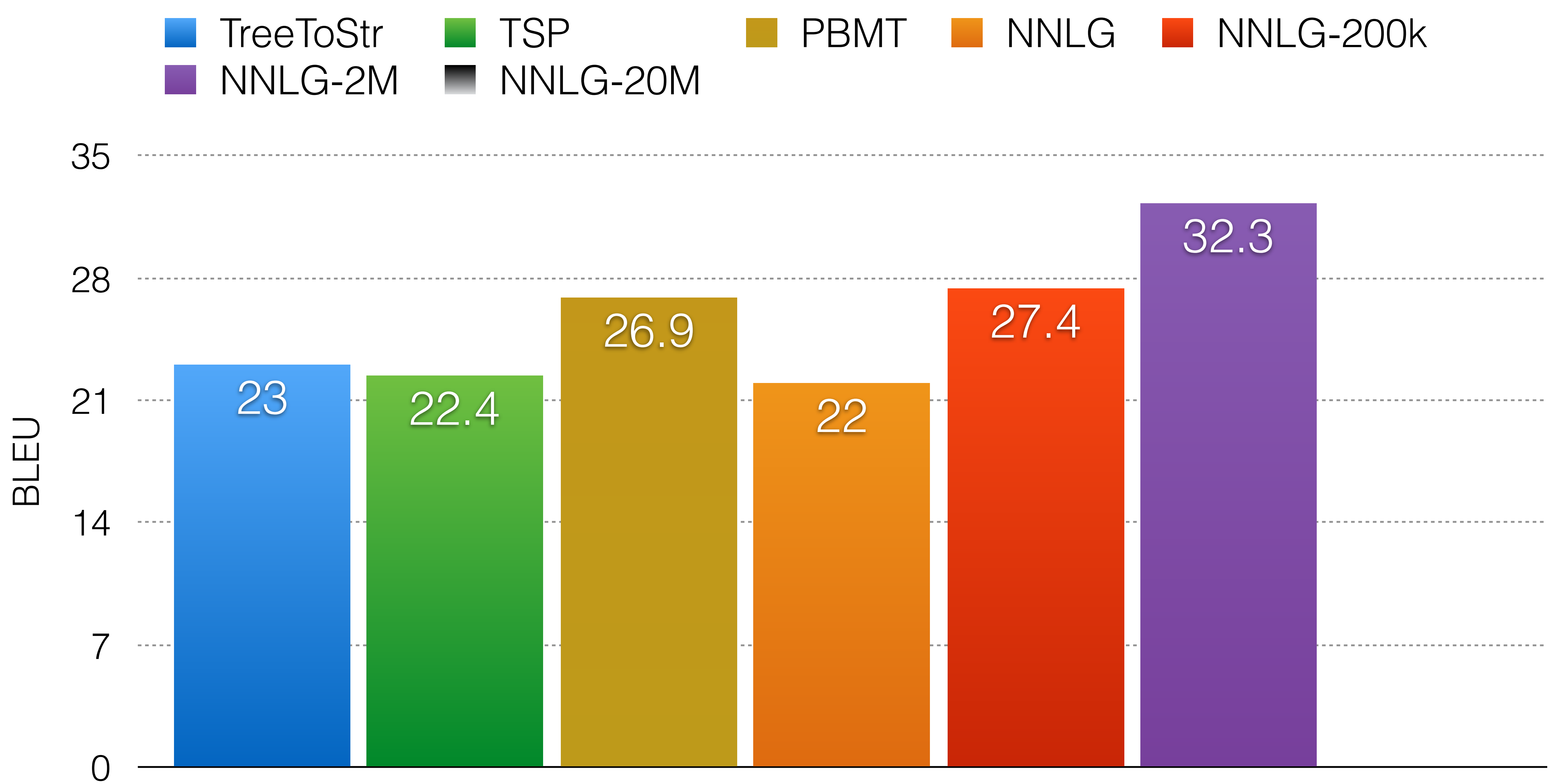
TreeToStr: Flanigan et al, NAACL 2016
TSP: Song et al, EMNLP 2016
PBMT: Pourdamaghani and Knight, INLG 2016

Final Results



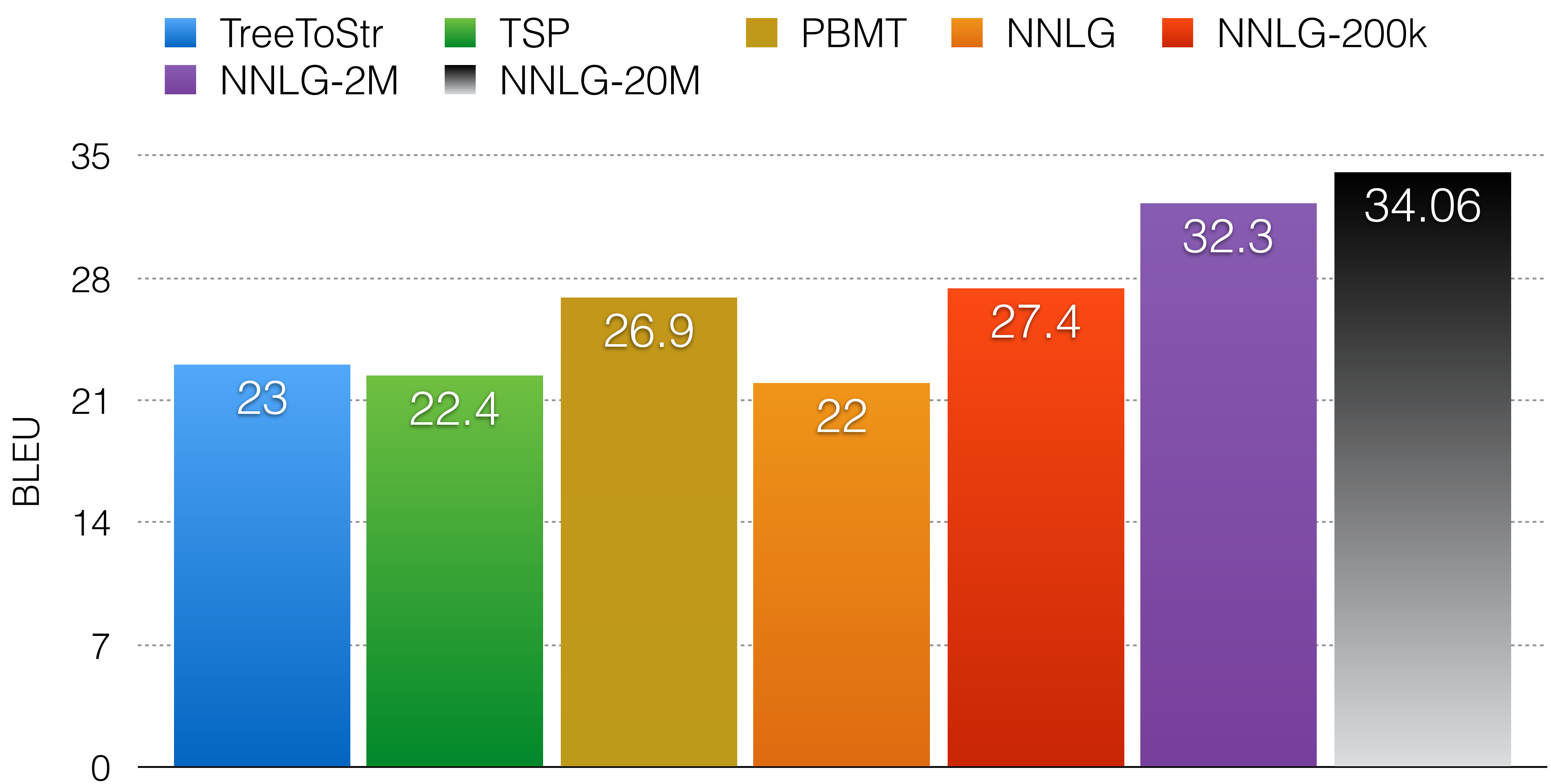
TreeToStr: Flanigan et al, NAACL 2016
TSP: Song et al, EMNLP 2016
PBMT: Pourdamaghani and Knight, INLG 2016

Final Results



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

Final Results



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

How did we do?

```
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?

```
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**

Adapt to other applications?

- **Structured** input representation
  - ☑ Meaning Representation of Natural Language
  - ☐ Programming Language

Code to Language

Joint work with
Srinivasan Iyer
Luke Zettlemoyer, Alvin Cheung

Code to Language

Input: Source Code
(SQL - C#)

```
public int TextWidth (string text) {
  TextBlock t = new TextBlock();
  t.Text = text;
  return (int) Math.Ceiling(t.ActualWidth);
}
```

Output: Summary

Get rendered width of string rounded up to the nearest integer.

(Summarizing Source Code using a Neural Attention Model. Iyer, Konstas, Cheung, Zettlemoyer, ACL 2016)

Code to Language

Input: Source Code
(SQL - C#)

Output: Summary

```
public int TextWidth (string text) {
  TextBlock t = new TextBlock();
  t.Text = text;
  return (int) Math.Ceiling(t.ActualWidth);
}
```

Get rendered width of string rounded up to the nearest integer.

```
SELECT max(marks)
FROM stud_records
WHERE marks < (SELECT max(marks) FROM stud_records);
```

How to find the second largest value from a table?

(**Summarizing Source Code using a Neural Attention Model. Iyer, Konstas, Cheung, Zettlemoyer, ACL 2016**)

Input Representation

1) Code snippet —>Linearize (left-to-right)

```
SELECT max(marks)
FROM stud_records
WHERE marks < (SELECT max(marks) FROM stud_records);
```

How to find the second largest value from a table?

(Summarizing Source Code using a Neural Attention Model. Iyer, Konstas, Cheung, Zettlemoyer, ACL 2016)

Input Representation

1) Code snippet —>Linearize (left-to-right)
2) Anonymize

```
SELECT max(marks)
FROM stud_records
WHERE marks < (SELECT max(marks) FROM stud_records);
```

```
SELECT max(col0)
FROM tab0
WHERE col0 < (SELECT max(col1) FROM tab1);
```

How to find the second largest value from a table?

(Summarizing Source Code using a Neural Attention Model. Iyer, Konstas, Cheung, Zettlemoyer, ACL 2016)

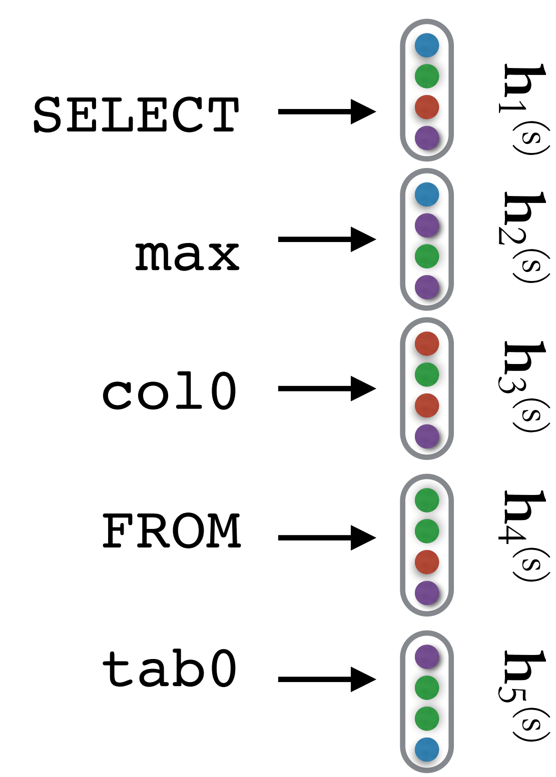
Input Representation

1) Code snippet —>Linearize (left-to-right)
2) Anonymize
3) Bag of Words Encoding

```
SELECT max(marks)
FROM stud_records
WHERE marks < (SELECT max(marks) FROM stud_records);
```

```
SELECT max(col0)
FROM tab0
WHERE col0 < (SELECT max(col1) FROM tab1);
```

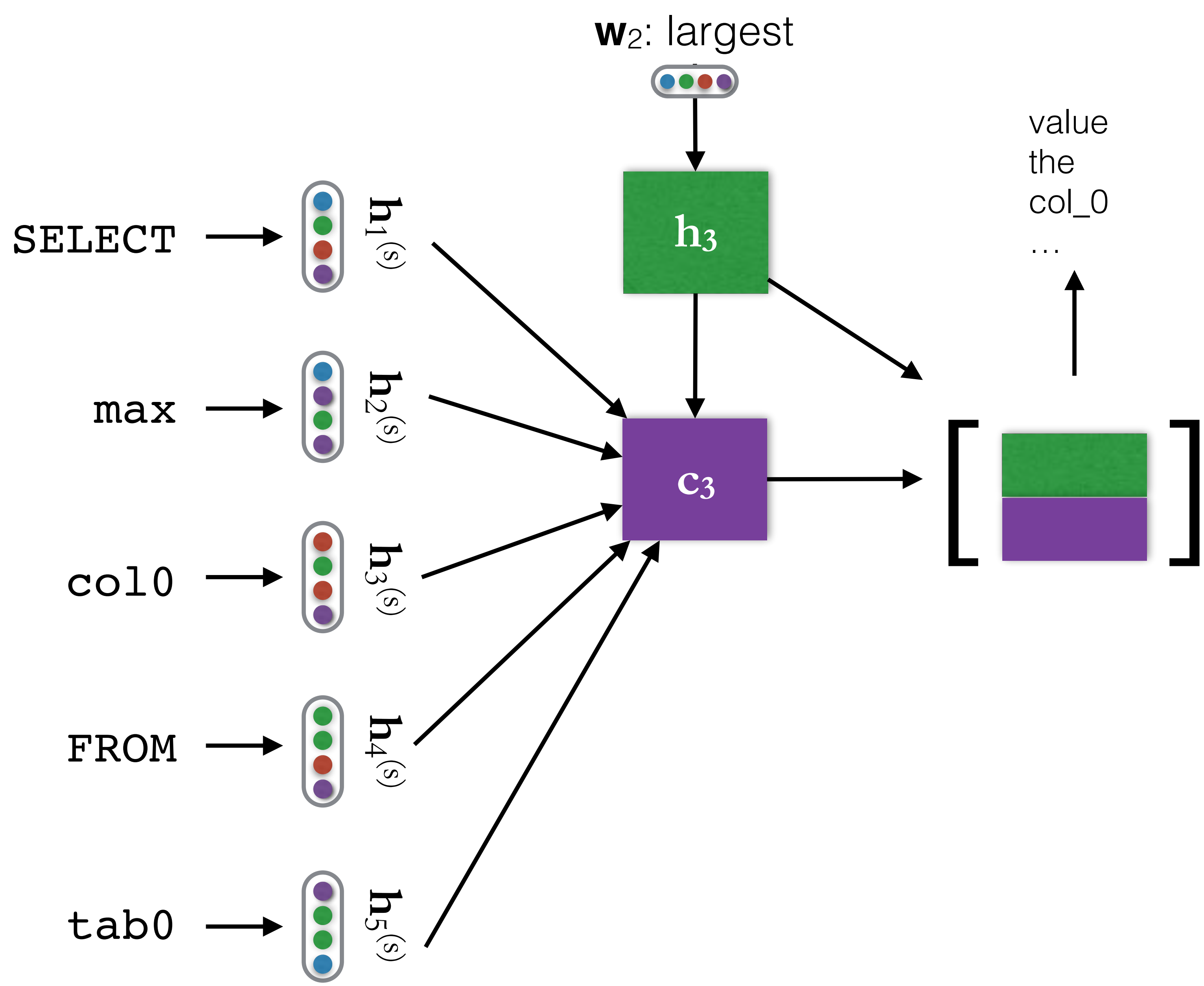
How to find the second largest value from a table?



(Summarizing Source Code using a Neural Attention Model. Iyer, Konstas, Cheung, Zettlemoyer, ACL 2016)

Decoding with Attention

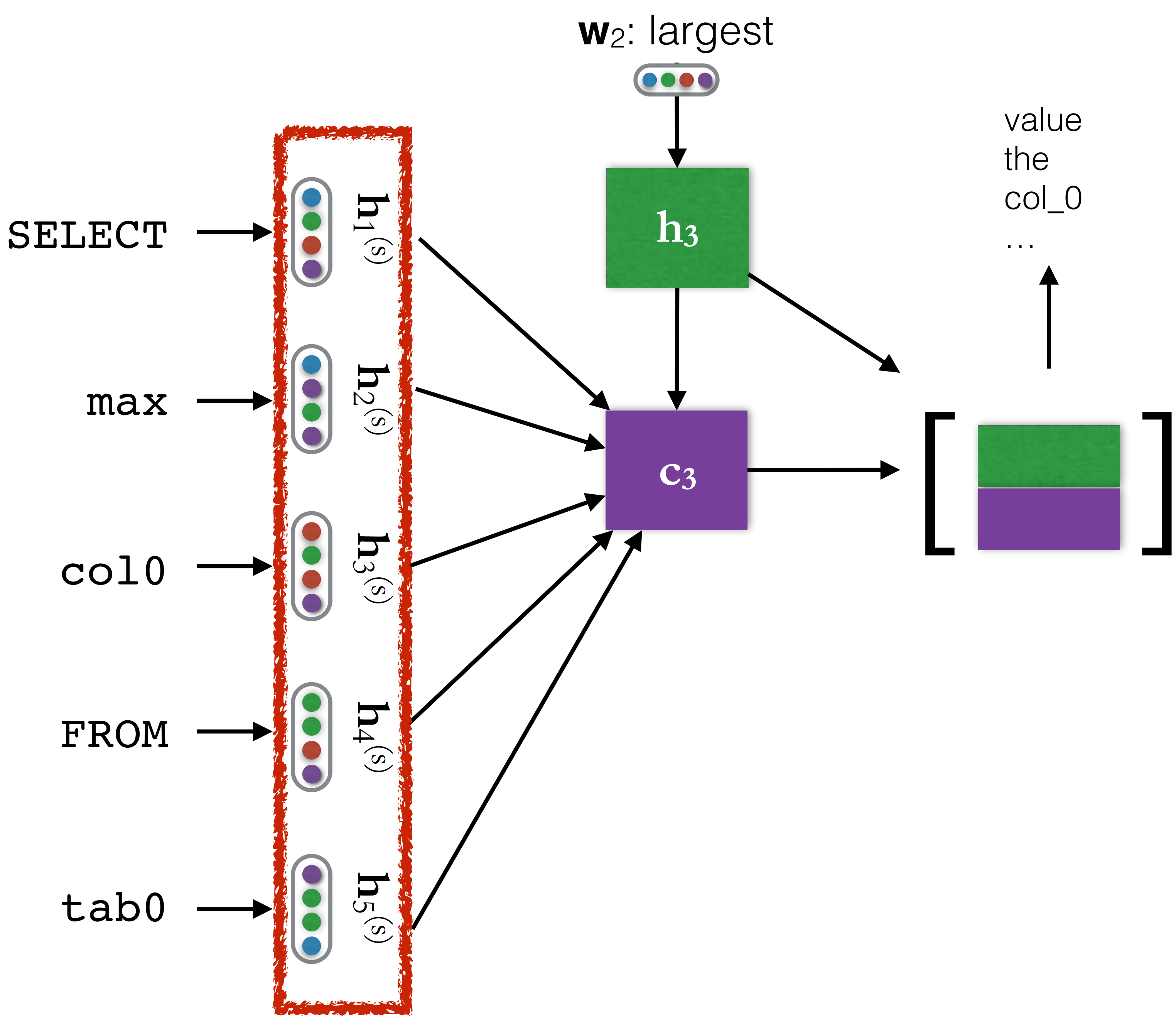
4) Bag of Words Encoding —>RNN Decoding
5) Attention directly on input embeddings



(Summarizing Source Code using a Neural Attention Model. Iyer, Konstas, Cheung, Zettlemoyer, ACL 2016)

Decoding with Attention

4) Bag of Words Encoding —>RNN Decoding
5) Attention directly on input embeddings



(Summarizing Source Code using a Neural Attention Model. Iyer, Konstas, Cheung, Zettlemoyer, ACL 2016)

Community-based Datasets

Community-based Datasets

stackoverflow Questions Jobs Documentation BETA Tags Users Search...

How to find the Second largest value from a table.?

3

One table with EmpSalary in Employee Table. I need to find the second largest Salary what is paid by the company.?

How to find the Second largest value(Salary) from a table.?

sql-server-2005 tsql

share improve this question

edited Oct 18 '10 at 10:56

asked Oct 18 '10 at 10:42

691 ●4 ●16 ●38

5 What value do you want returned if there are two records with the equal top value? – Oct 18 '10 at 10:48

add a comment

3 Answers

active oldest votes

3

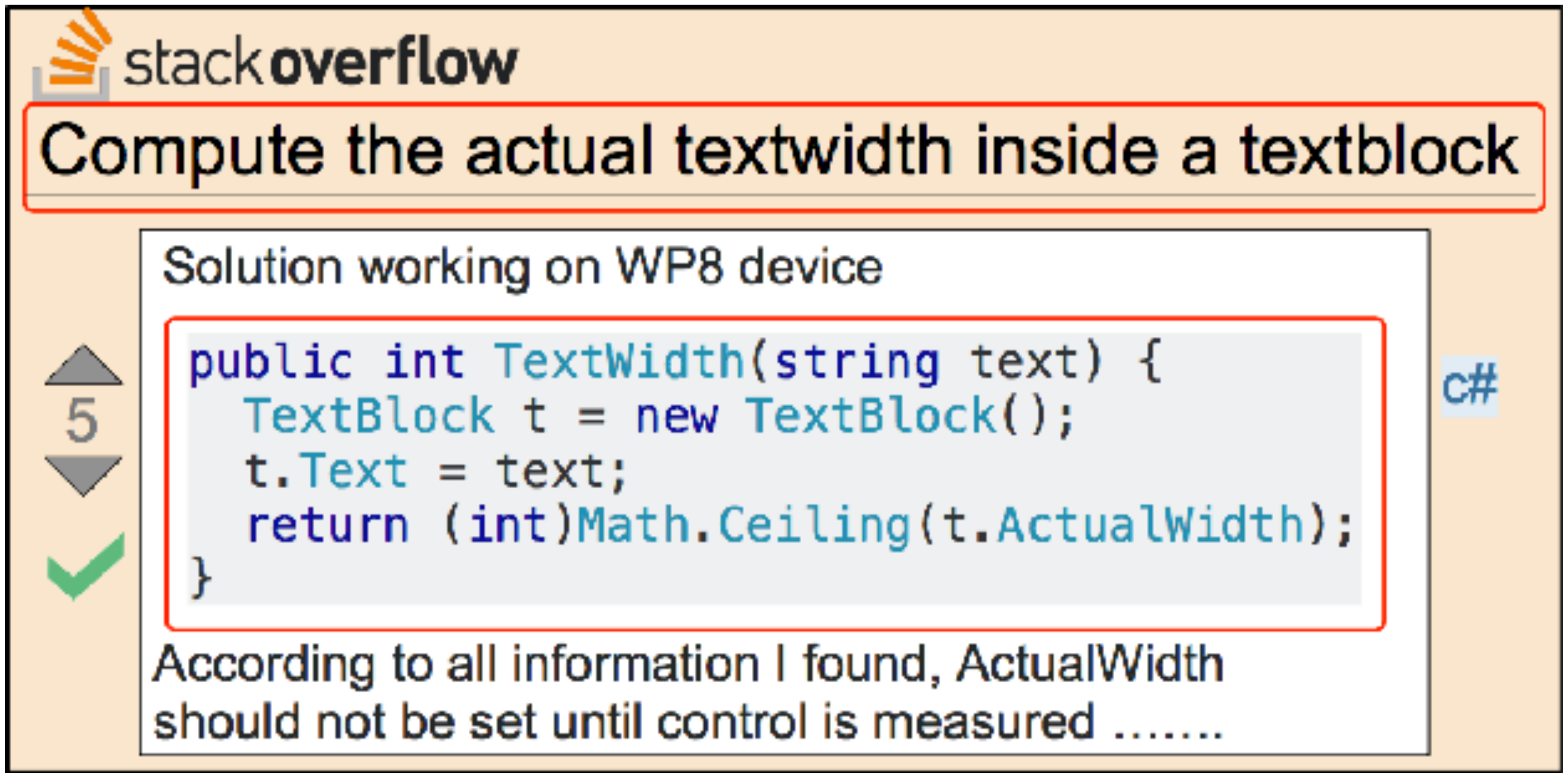
Try this: this should give the second largest salary:

SELECT MAX(EmpSalary) FROM employee WHERE EmpSalary < (SELECT MAX(EmpSalary) FROM employee);

share improve this answer

answered Oct 18 '10 at 10:59

Community-based Datasets

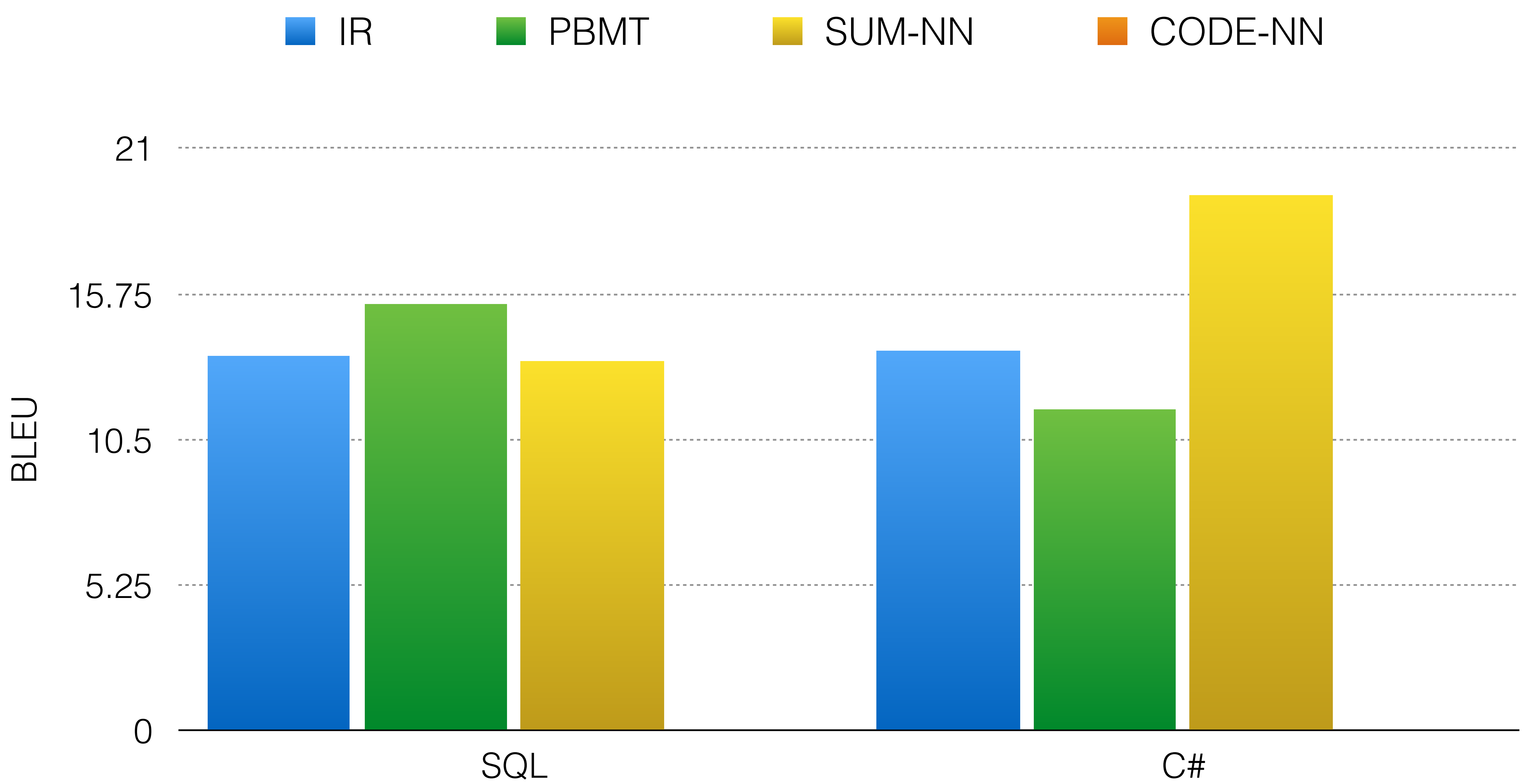


- (Accepted Answer, Post title) pairs
- ~33K **SQL** / 66k **C#** examples

Results

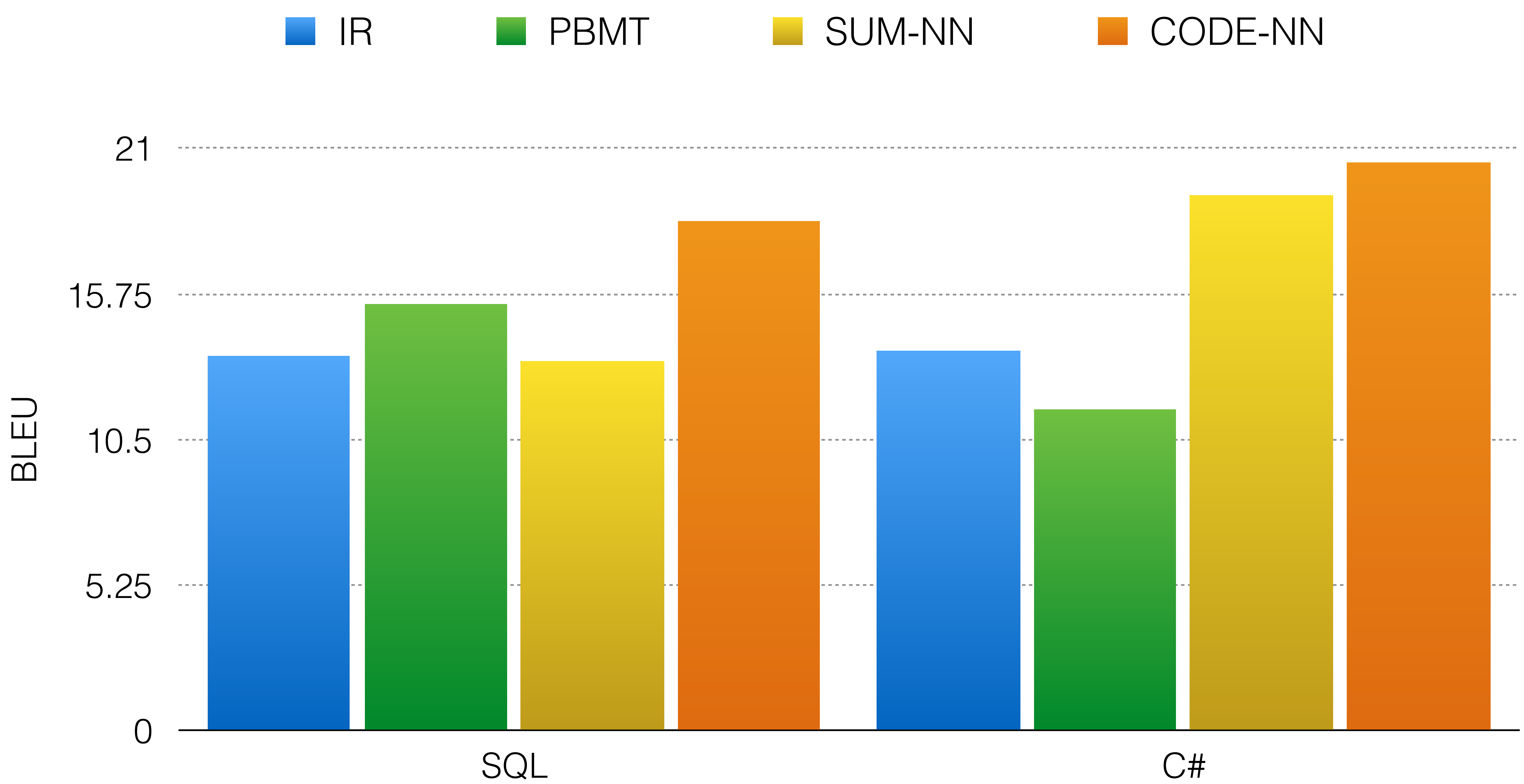
PBMT: MOSES Phrase-based MT system
SUM-NN: Rush et al, EMNLP 2015

Results



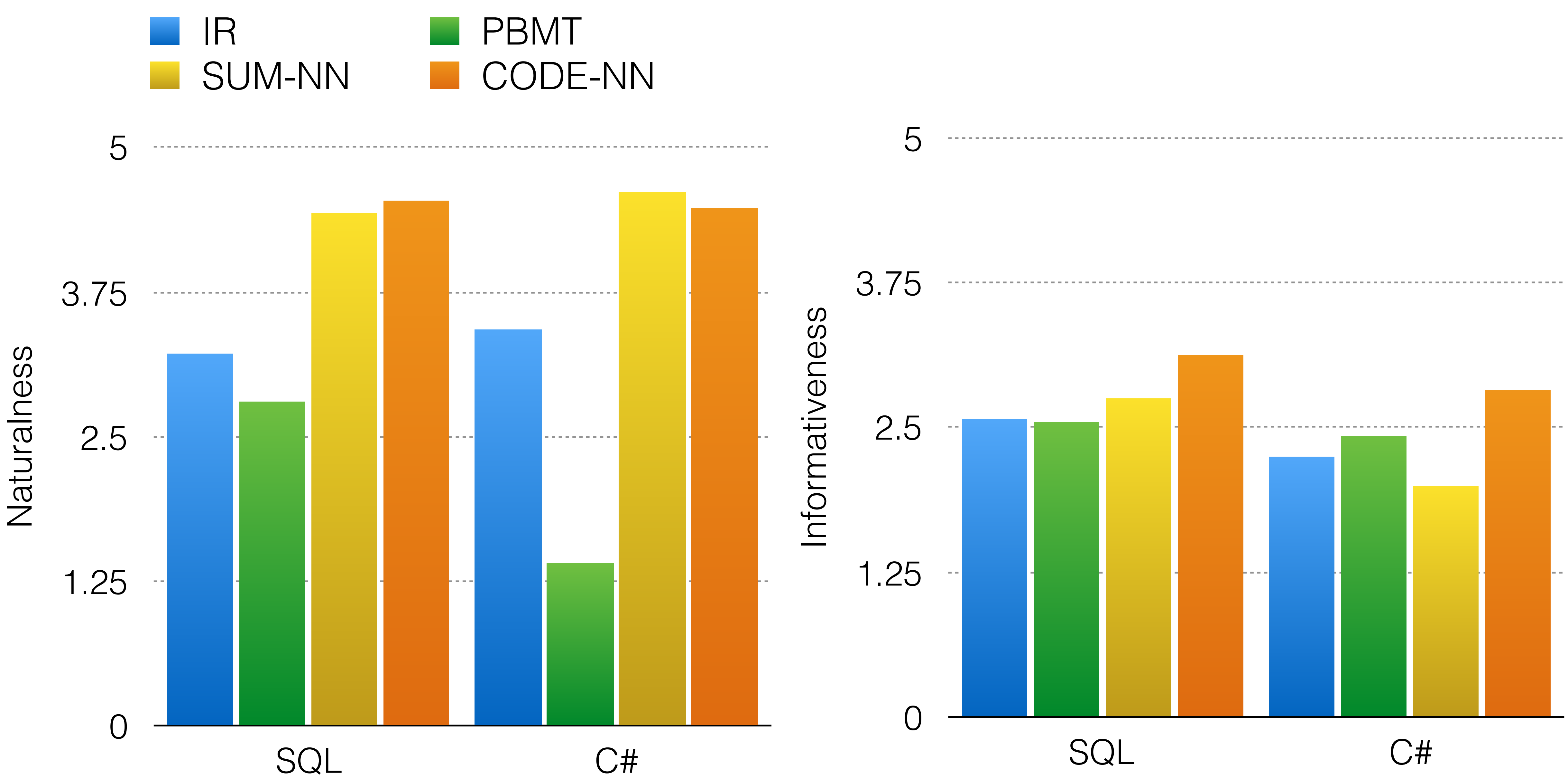
PBMT: MOSES Phrase-based MT system
SUM-NN: Rush et al, EMNLP 2015

Results



PBMT: MOSES Phrase-based MT system
SUM-NN: Rush et al, EMNLP 2015

Human Evaluation Results



PBMT: MOSES Phrase-based MT system
SUM-NN: Rush et al, EMNLP 2015

How did we do?

```
SELECT * FROM  table
ORDER  BY Rand()  LIMIT  10
```

Reference

Select random rows from mysql table

CODE-NN

How to get random rows from a mysql database?

How did we do?

```
SELECT * FROM  table
ORDER  BY Rand()  LIMIT  10
```

```
foreach (string  pTxt in xml.parent) {
   TreeNode parent = new TreeNode();
   foreach (string  cTxt in xml.child) {
     TreeNode child = new  TreeNode();
     parent.Nodes.Add(child );
   }
}
```

Reference

Select random rows from mysql table

CODE-NN

How to get random rows from a mysql database?

Reference

Adding childs to a treenode dynamically in C#

CODE-NN

How to get all child nodes in TreeView?

Neural NLG Contributions

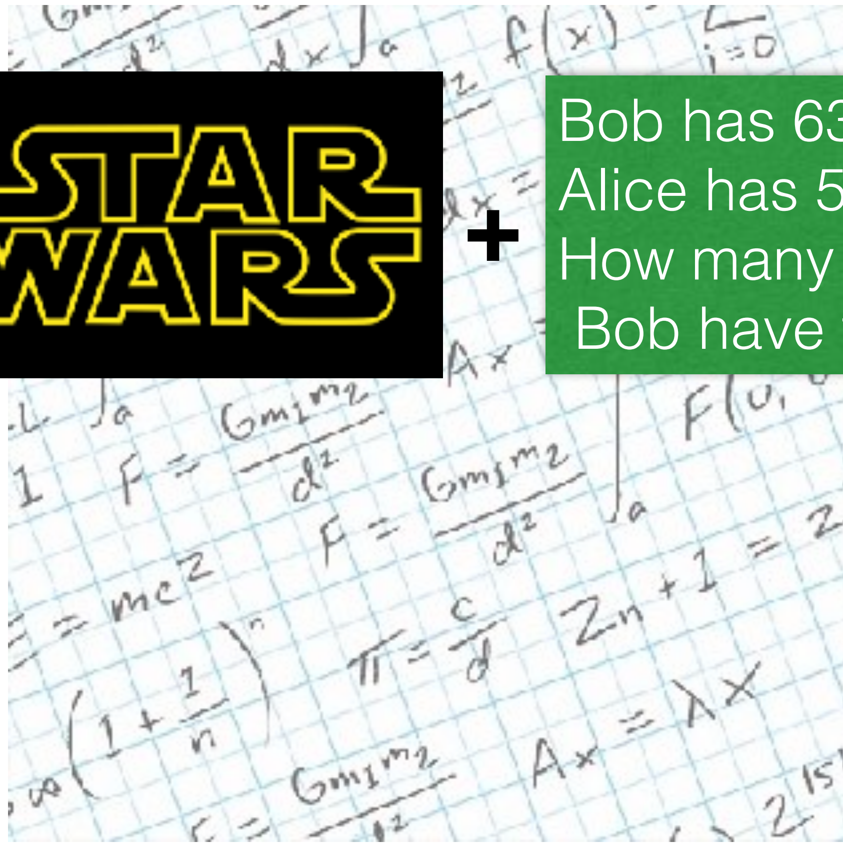
Neural NLG Contributions

- **Adapt** to multiple applications
- **Scale** to very large corpora

- Address low-resource problem
  - **Paired training** general technique
  - Train on **noisy** community-based datasets

Future Work

Educational Technology



\+

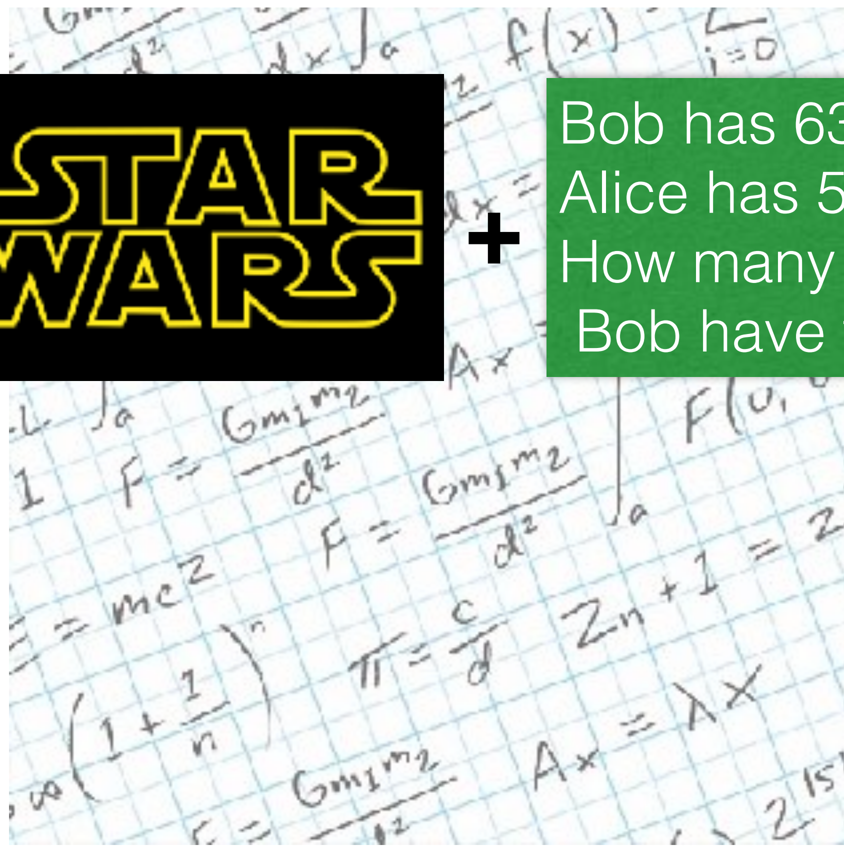
Bob has 639 sheep.
Alice has 504 sheep.
How many more sheep does Bob have than Alice?

=

Joint work with
Rik Koncel-Kedziorski
Luke Zettlemoyer, Hannaneh Hajishirzi

(Koncel-Kedziorski, Konstas, Zettlemoyer, Hajishirzi. A Theme-Rewriting Approach for Generating Algebra Word Problems, EMNLP 2016)

Educational Technology



STAR WARS +

Bob has 639 sheep.
Alice has 504 sheep.
How many more sheep does Bob have than Alice?

=

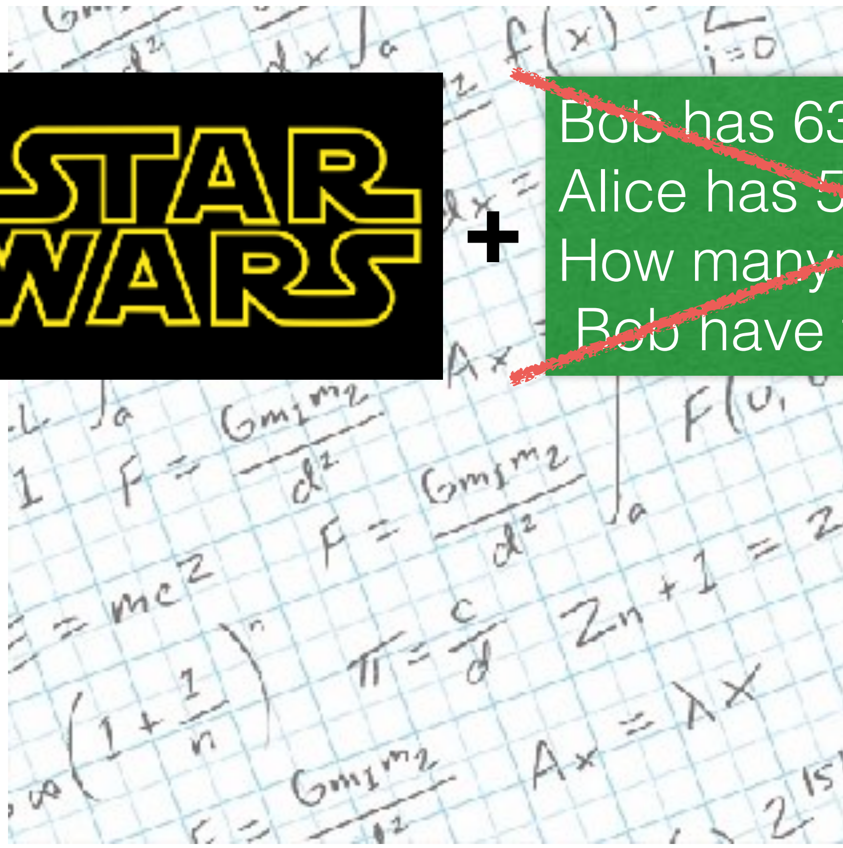
Luke Skywalker has 639 blasters. Leia has 504 blasters. How many more blasters does Luke Skywalker have than Leia?

Syntactic, Semantic, Thematic **rewriter**

Joint work with
Rik Koncel-Kedziorski
Luke Zettlemoyer, Hannaneh Hajishirzi

(Koncel-Kedziorski, Konstas, Zettlemoyer, Hajishirzi. A Theme-Rewriting Approach for Generating Algebra Word Problems, EMNLP 2016)

Educational Technology



STAR WARS

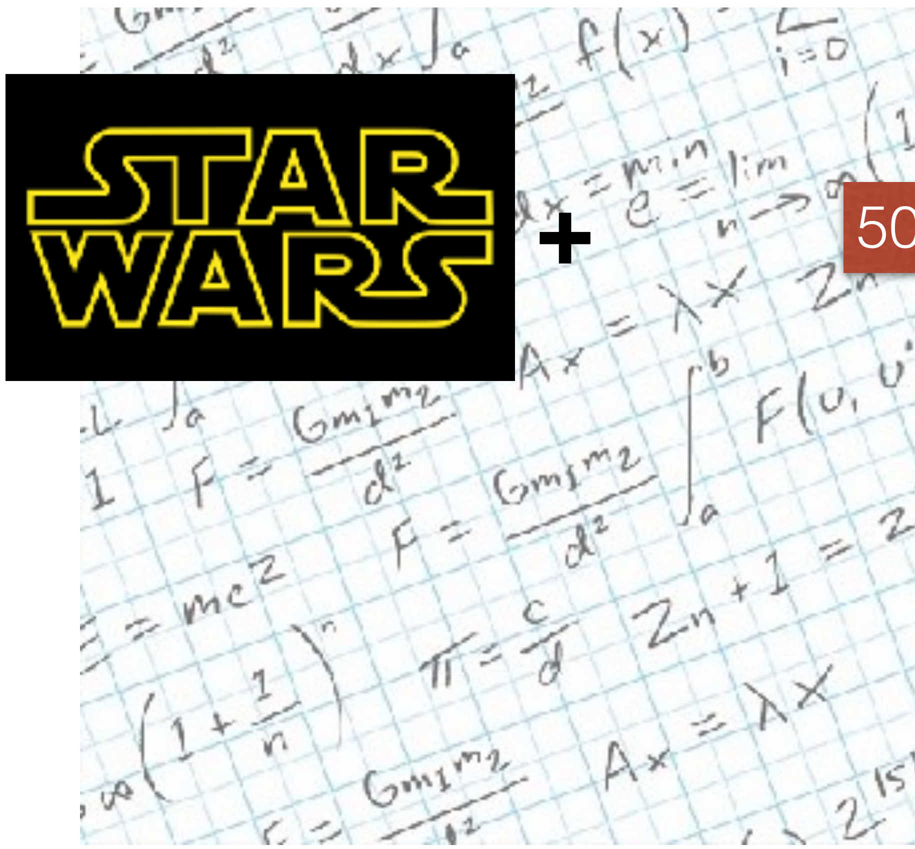
+

~~Bob has 639 sheep.~~
~~Alice has 504 sheep.~~
~~How many more sheep does Bob have than Alice?~~

=

Luke Skywalker has 639 blasters. Leia has 504 blasters. How many more blasters does Luke Skywalker have than Leia?

Educational Technology



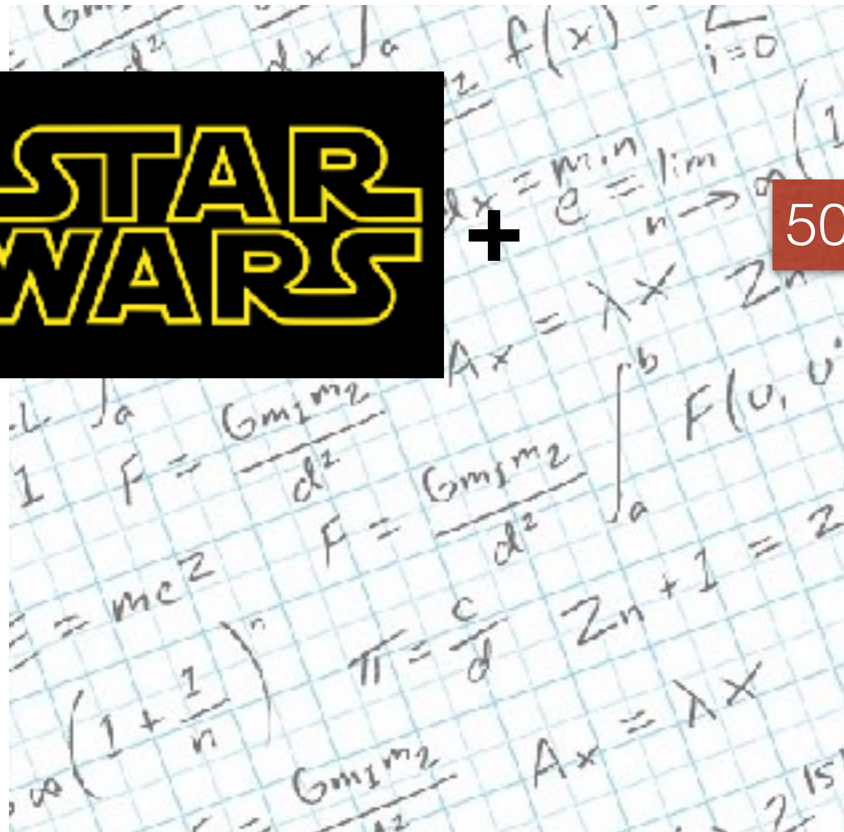
+

504 + x = 639

=

Luke Skywalker has 639 blasters. Leia has 504 blasters. How many more blasters does Luke Skywalker have than Leia?

Educational Technology



+ 504 + x = 639

= Luke Skywalker has 639 blasters. Leia has 504 blasters. How many more blasters does Luke Skywalker have than Leia?

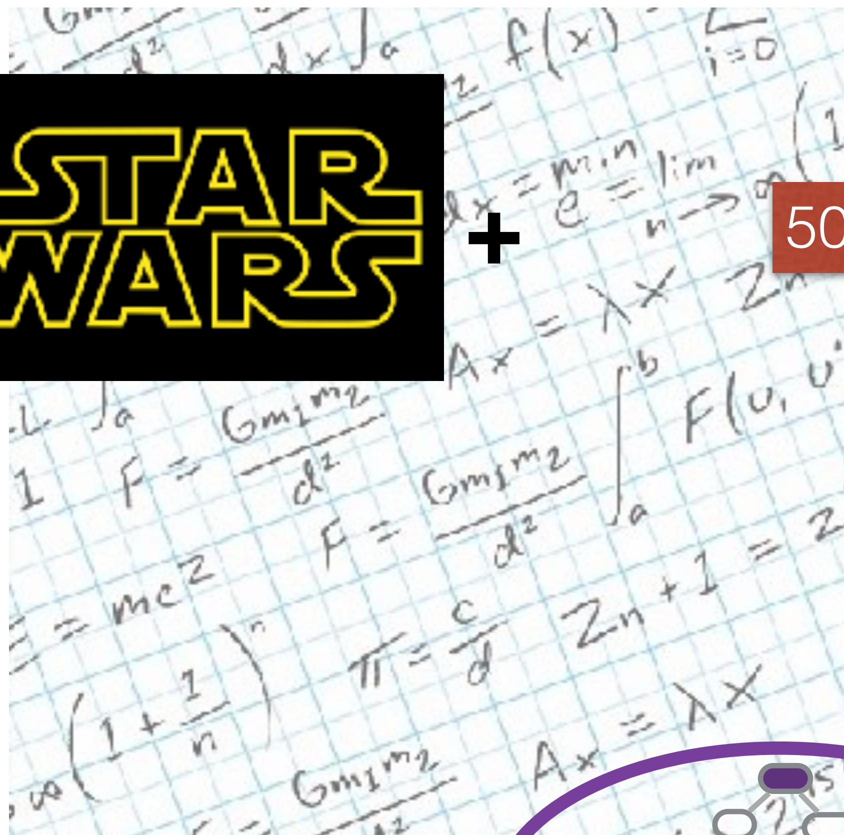
theme → Luke Skywalker blasters

504 + x = 639

Educational Technology

STAR WARS

+



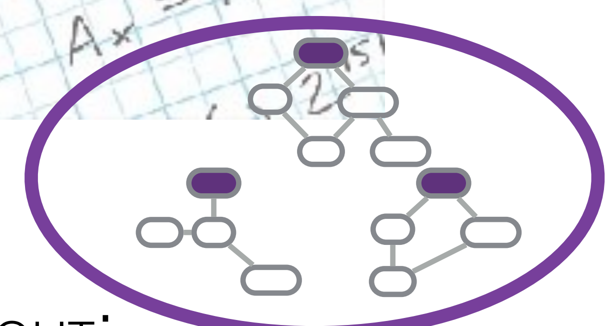
504 + x = 639

=

Luke Skywalker has 639 blasters. Leia has 504 blasters. How many more blasters does Luke Skywalker have than Leia?

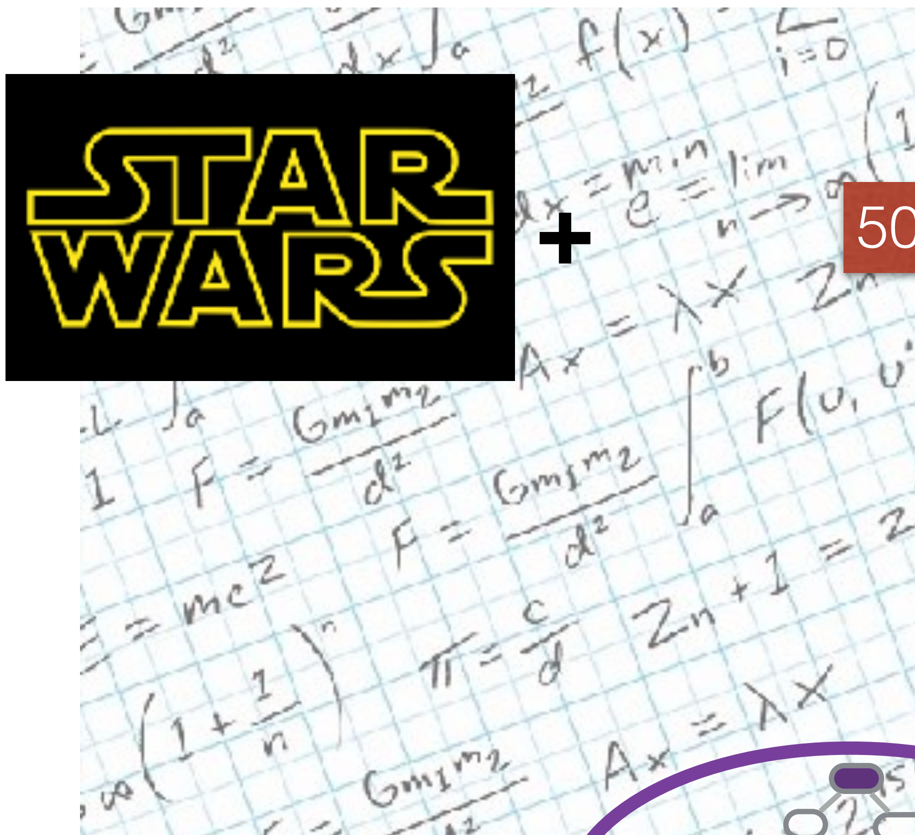
theme → Luke Skywalker blasters

SOUT:



504 + x = 639

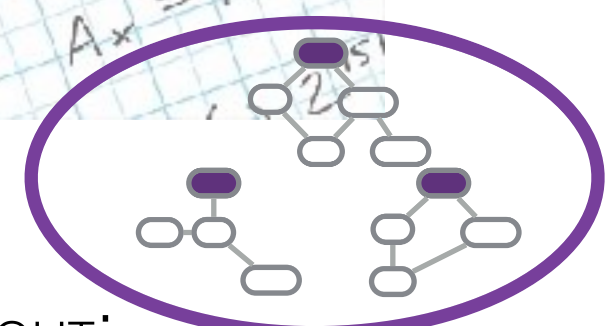
Educational Technology



+ 504 + x = 639 =

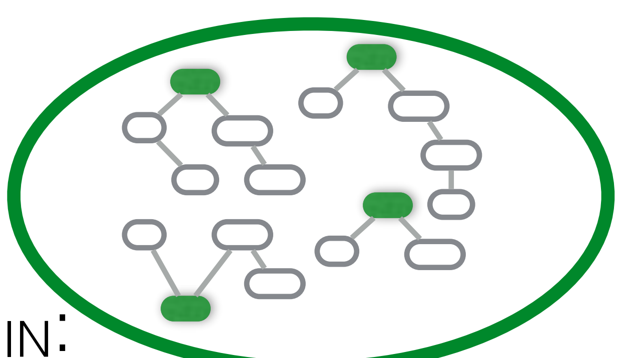
Luke Skywalker has 639 blasters. Leia has 504 blasters. How many more blasters does Luke Skywalker have than Leia?

theme → Luke Skywalker blasters



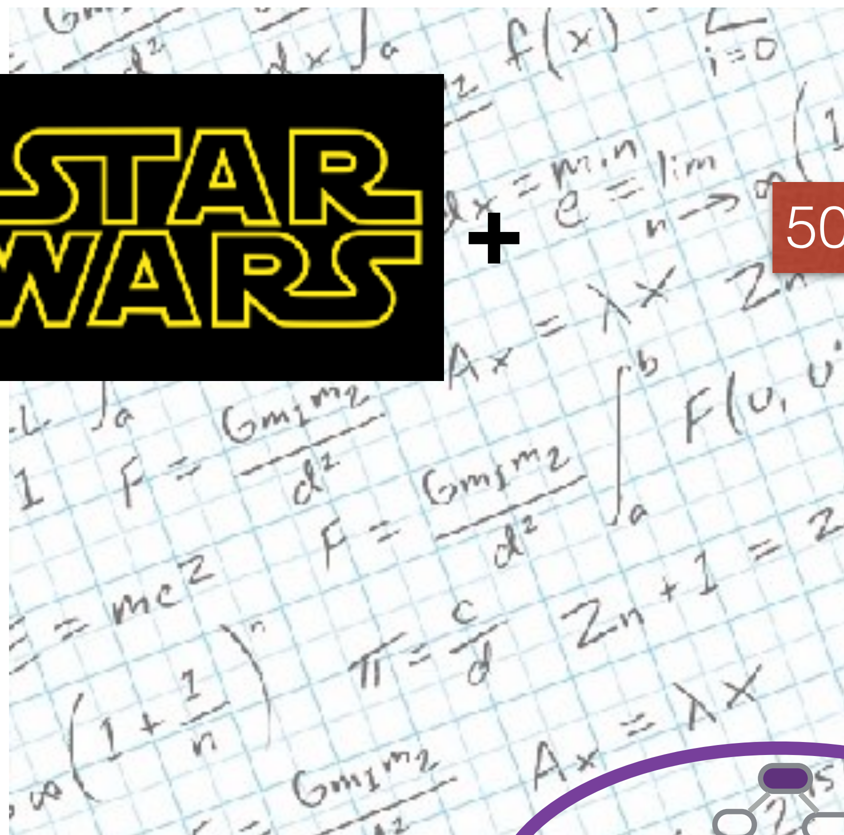
SOUT:

504 + x = 639



SIN:

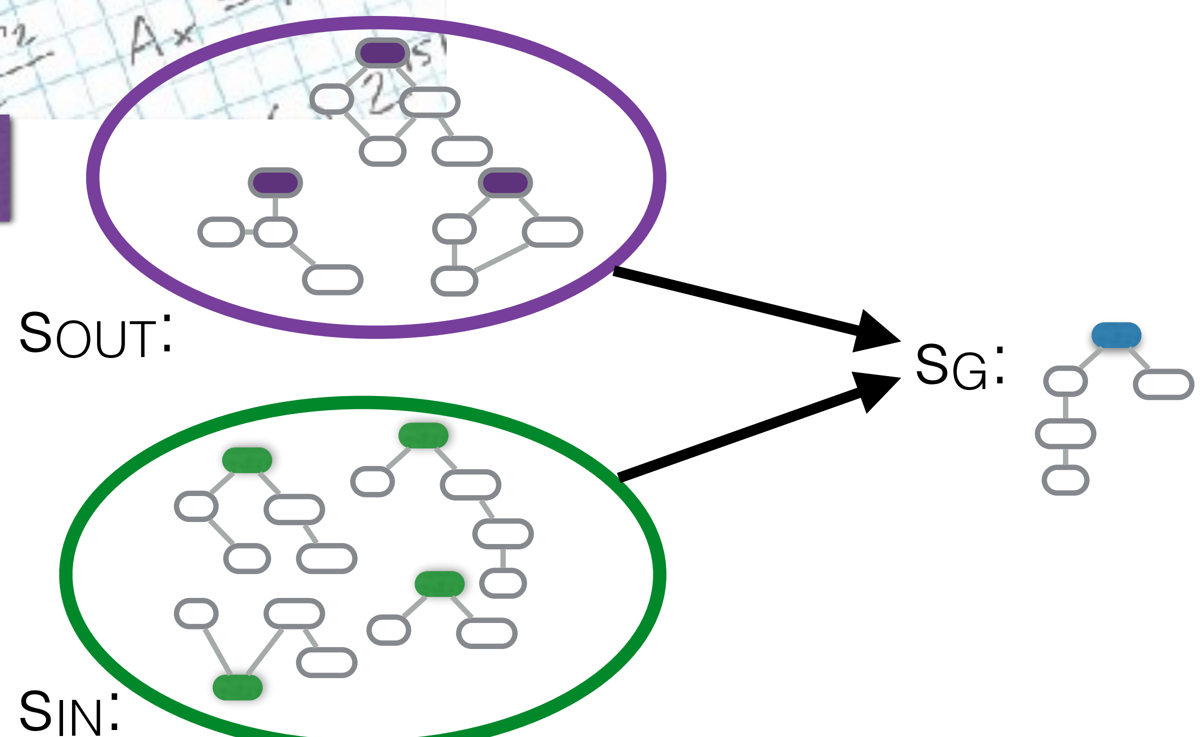
Educational Technology



+ 504 + x = 639 =

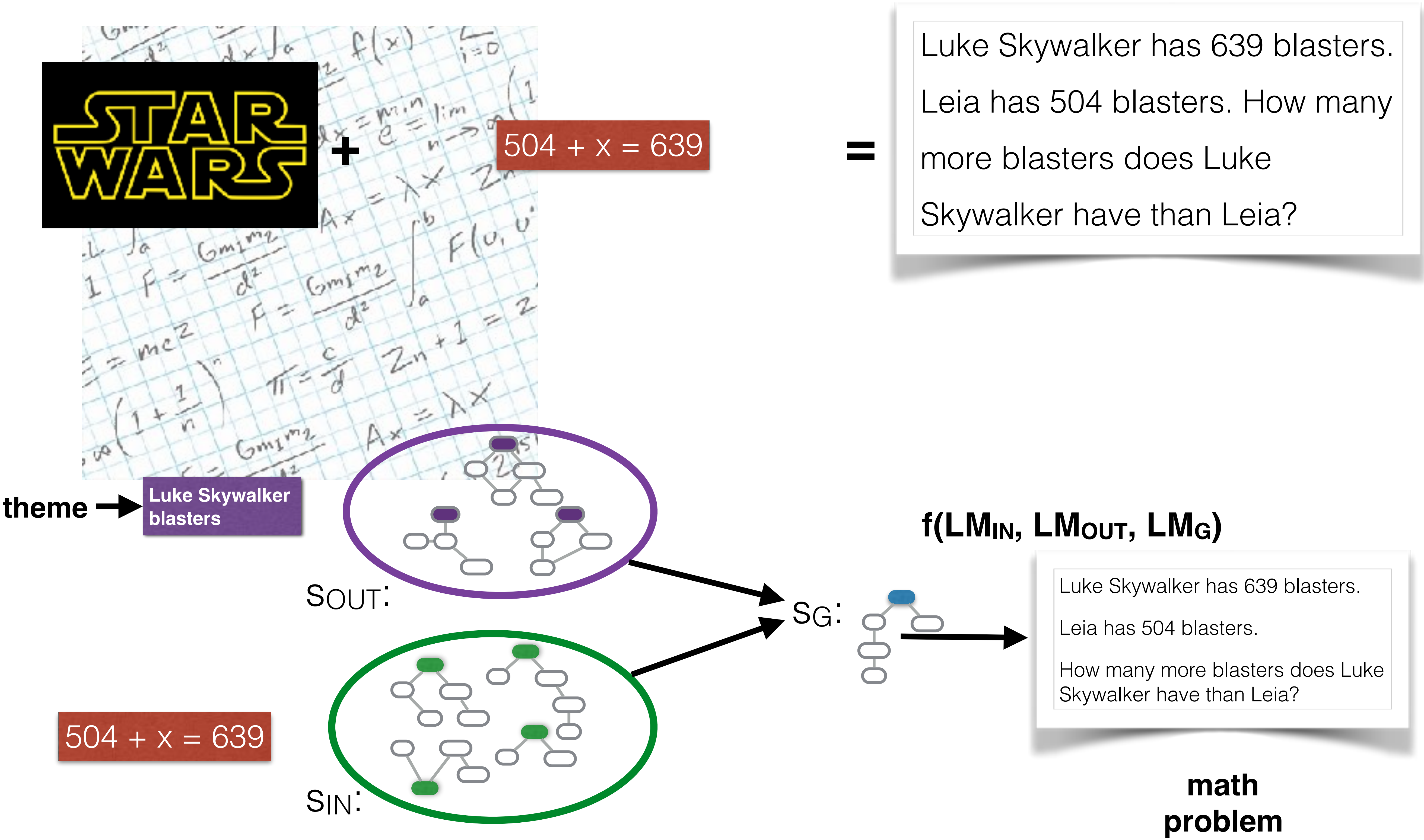
Luke Skywalker has 639 blasters. Leia has 504 blasters. How many more blasters does Luke Skywalker have than Leia?

theme → Luke Skywalker blasters



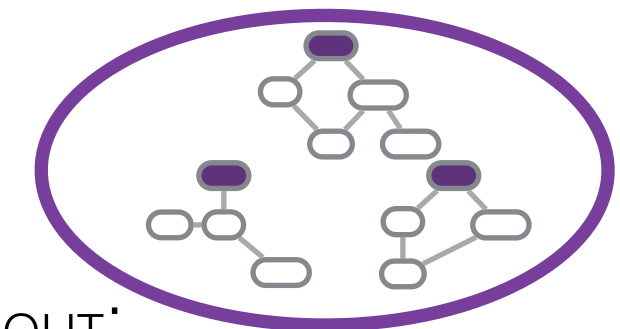
504 + x = 639

Educational Technology



Educational Technology

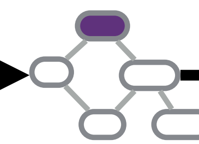
theme → **Luke Skywalker blasters**



SOUT:

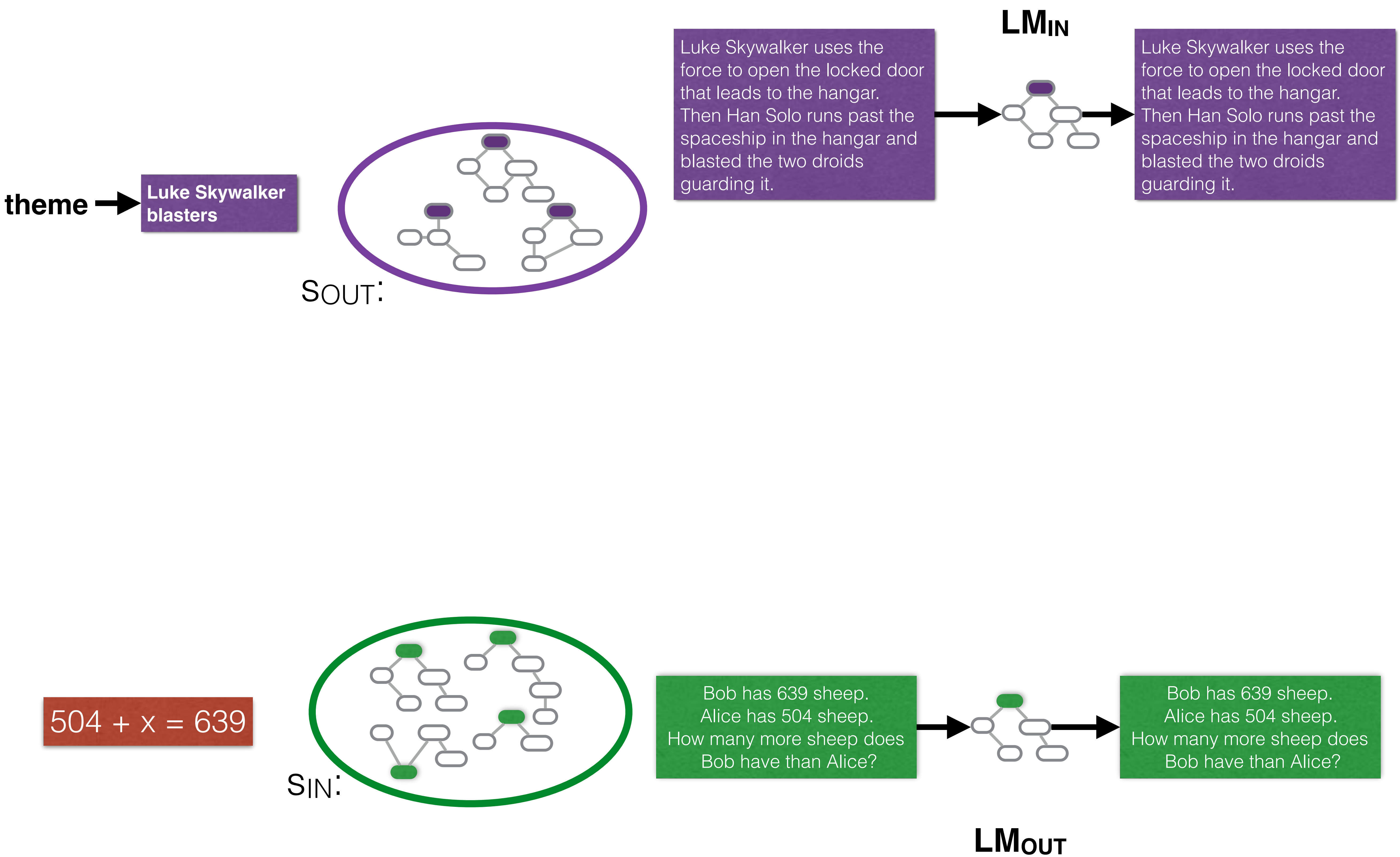
Luke Skywalker uses the force to open the locked door that leads to the hangar. Then Han Solo runs past the spaceship in the hangar and blasted the two droids guarding it.

LM$_{IN}$

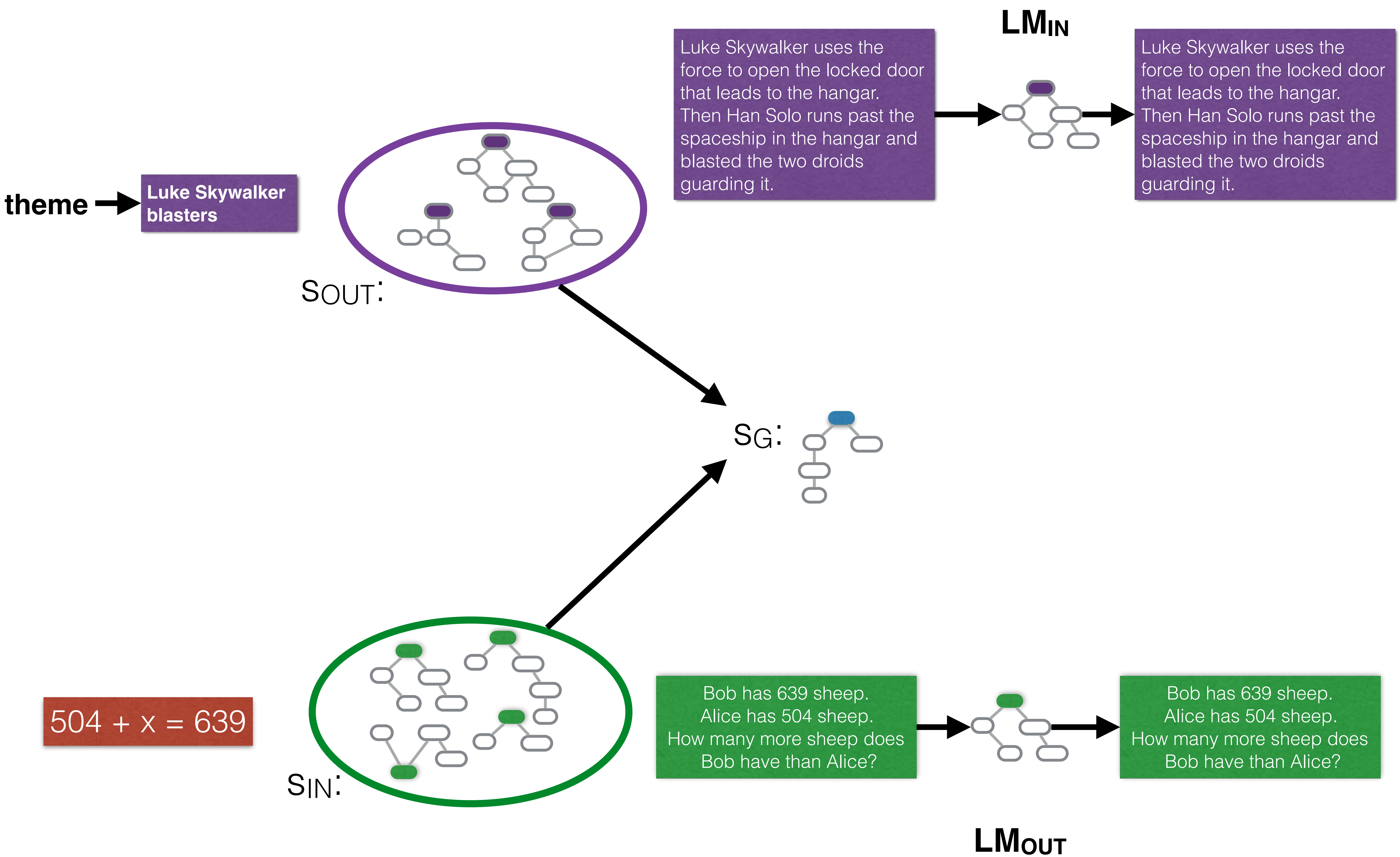


Luke Skywalker uses the force to open the locked door that leads to the hangar. Then Han Solo runs past the spaceship in the hangar and blasted the two droids guarding it.

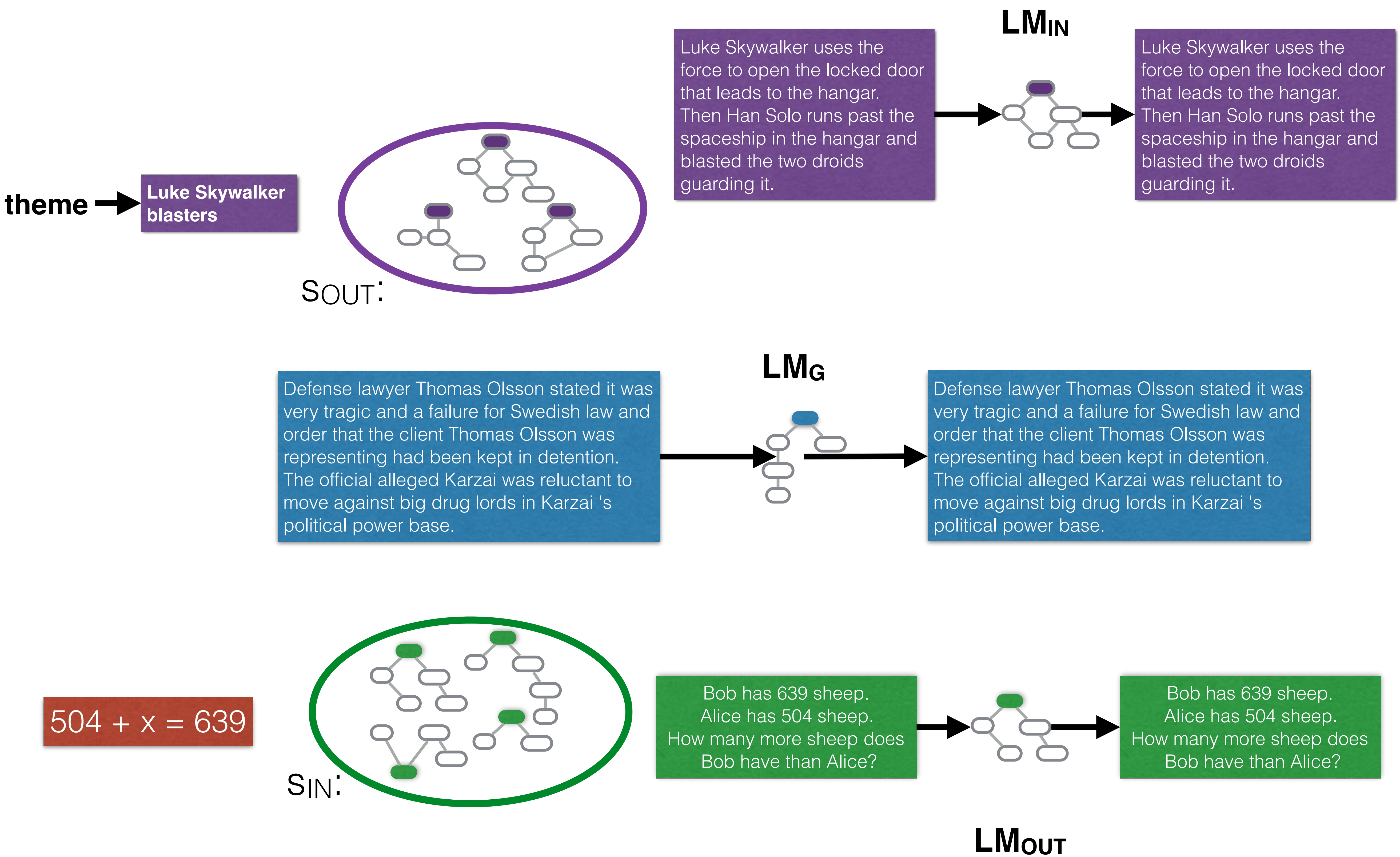
Educational Technology



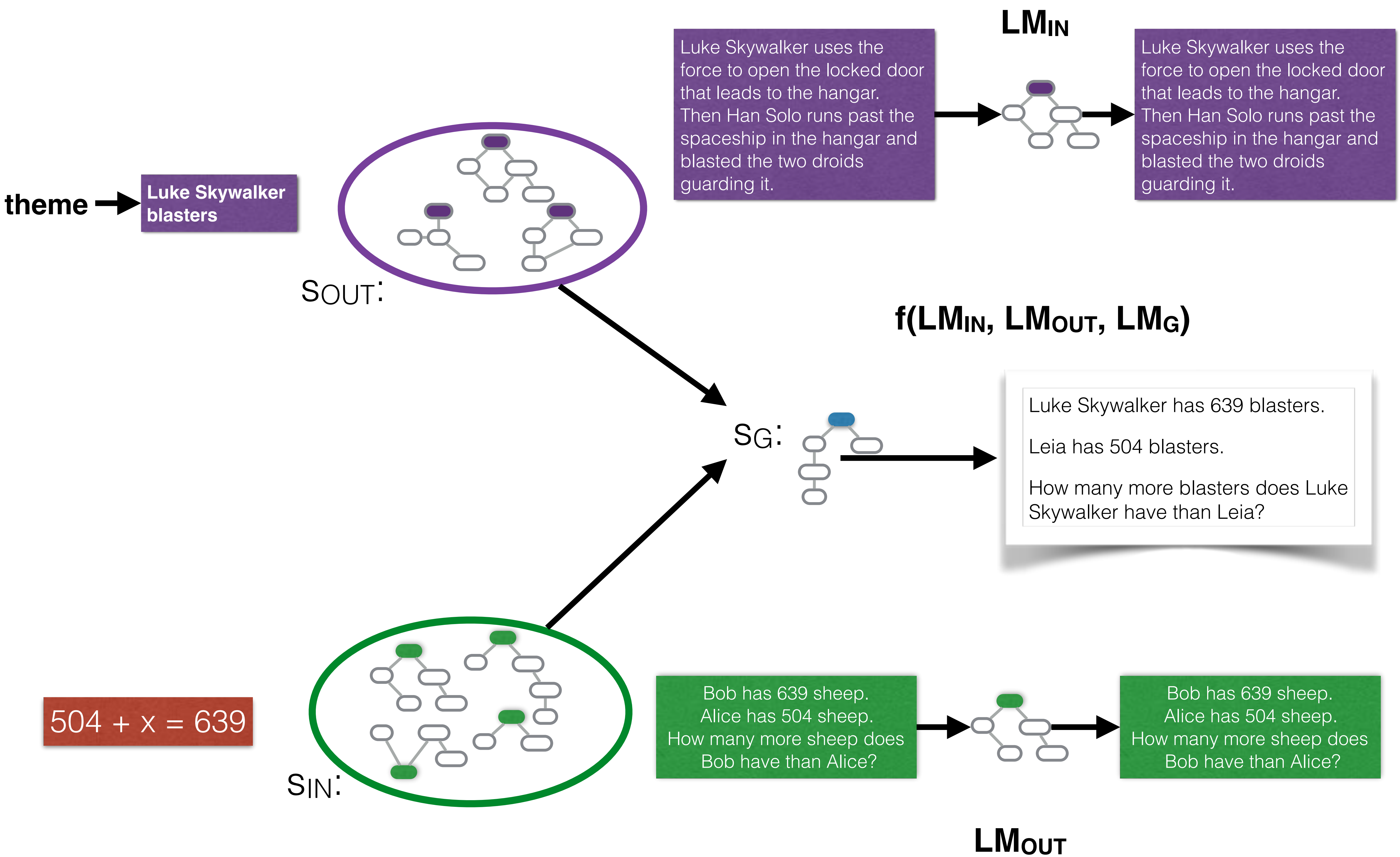
Educational Technology



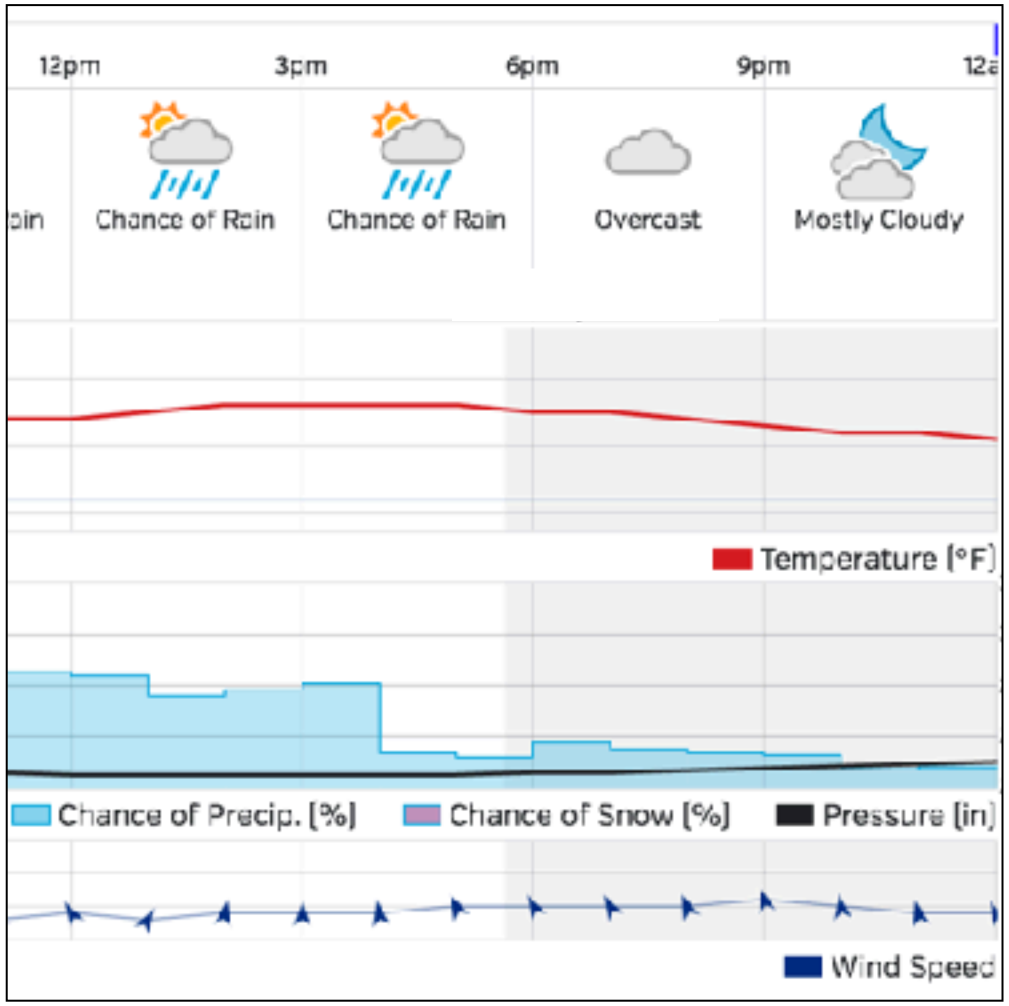
Educational Technology



Educational Technology



Concept-to-Text



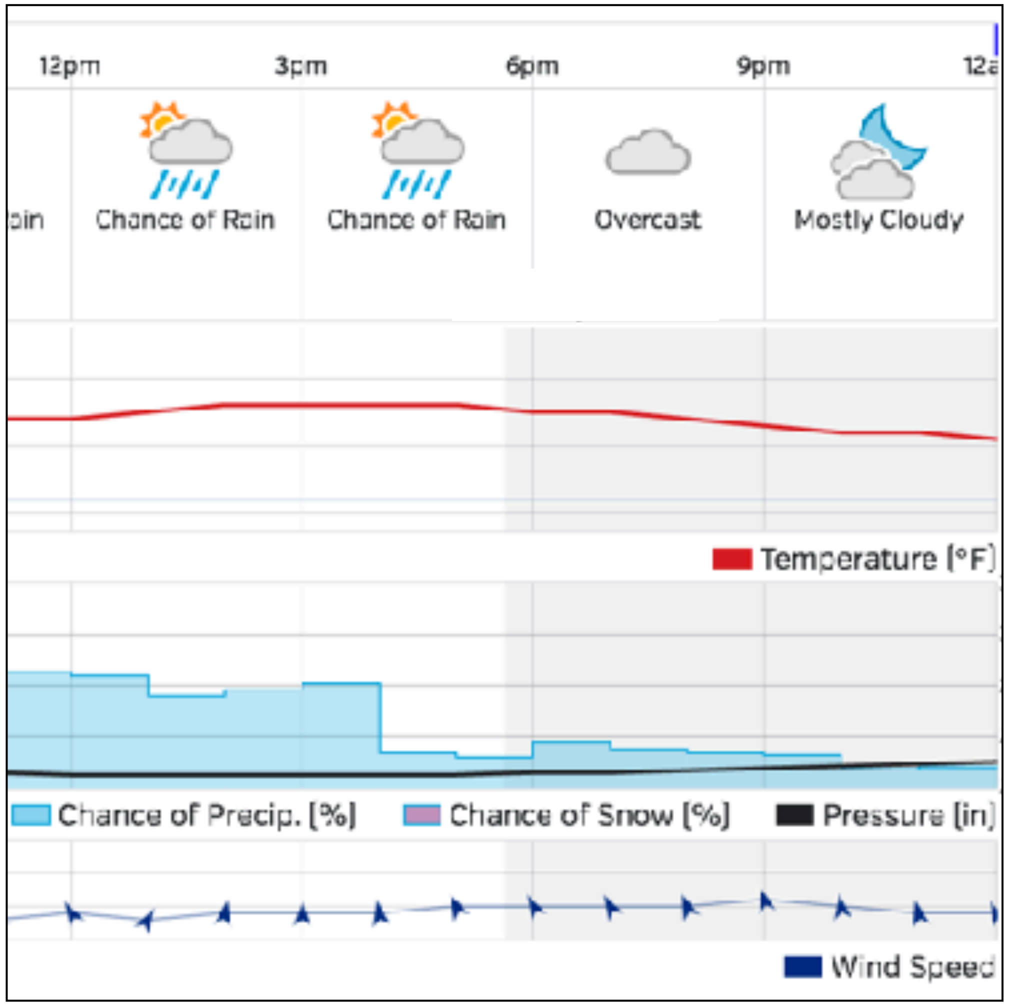
	time	min	mean	max	mode
wind	12-3	3	5	7	
wind	3-6	5	5	5	
wind	6-9	5	6	7	
dir	12-3				NW
dir	3-6				NE
dir	6-9				NE
temp	12-9	40	42	45	
precip	12-3	25	45	50	
precip	3-6	15	30	50	
precip	6-9	12	18	25	
cover	12-3				50-75
cover	3-6				50-75
cover	6-9				75-100

Chance of rain then becoming overcast, with a high of 45.
Calm to moderate northeast winds.

(Angeli et al. EMNLP 2010, Kim and Mooney COLING 2010)

(A Global Model for Concept-to-Text Generation. Konstas and Lapata, JAIR 2013; EMNLP 2013)

Concept-to-Text



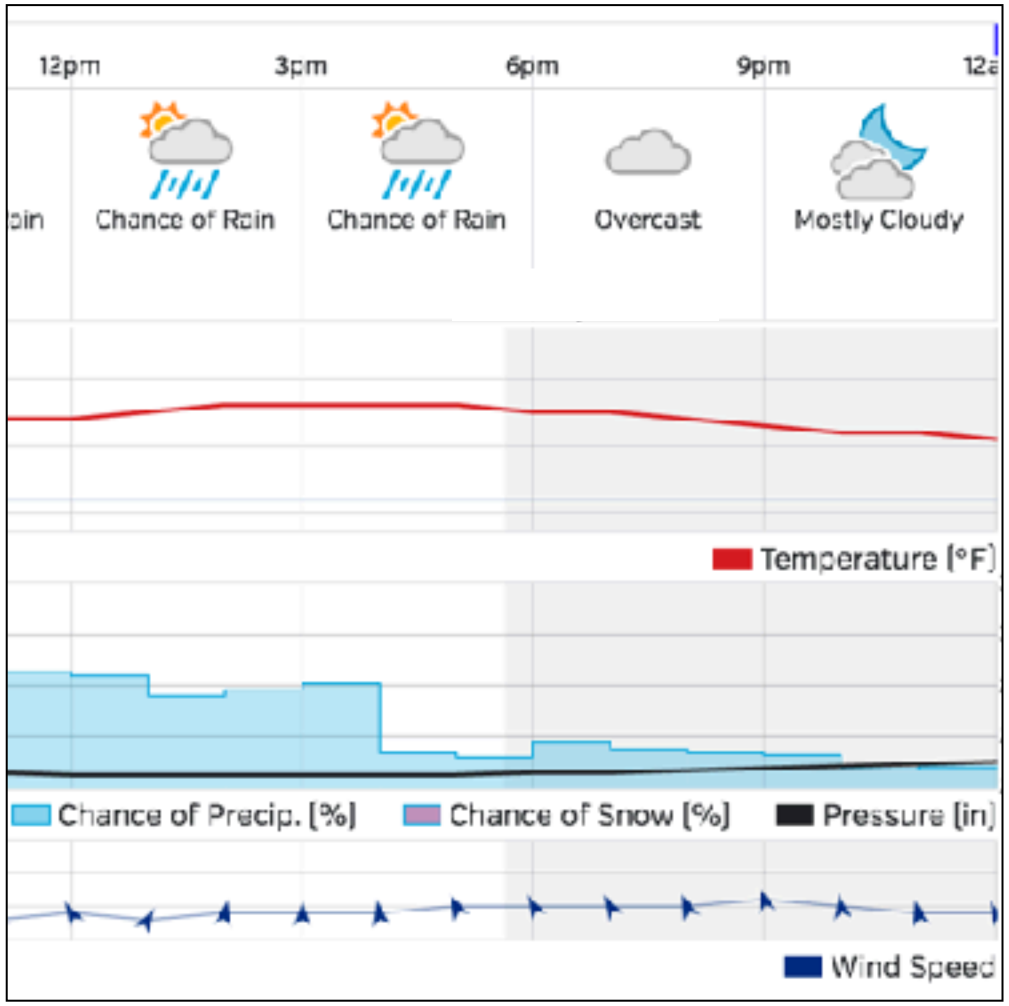
	time	min	mean	max	mode
wind	12-3	3	5	7	
wind	3-6	5	5	5	
wind	6-9	5	6	7	
dir	12-3				NW
dir	3-6				NE
dir	6-9				NE
temp	12-9	40	42	45	
precip	12-3	25	45	50	
precip	3-6	15	30	50	
precip	6-9	12	18	25	
cover	12-3				50-75
cover	3-6				50-75
cover	6-9				75-100

Chance of rain then becoming overcast, with a high of 45.
Calm to moderate northeast winds.

(Angeli et al. EMNLP 2010, Kim and Mooney COLING 2010)

(A Global Model for Concept-to-Text Generation. Konstas and Lapata, JAIR 2013; EMNLP 2013)

Concept-to-Text



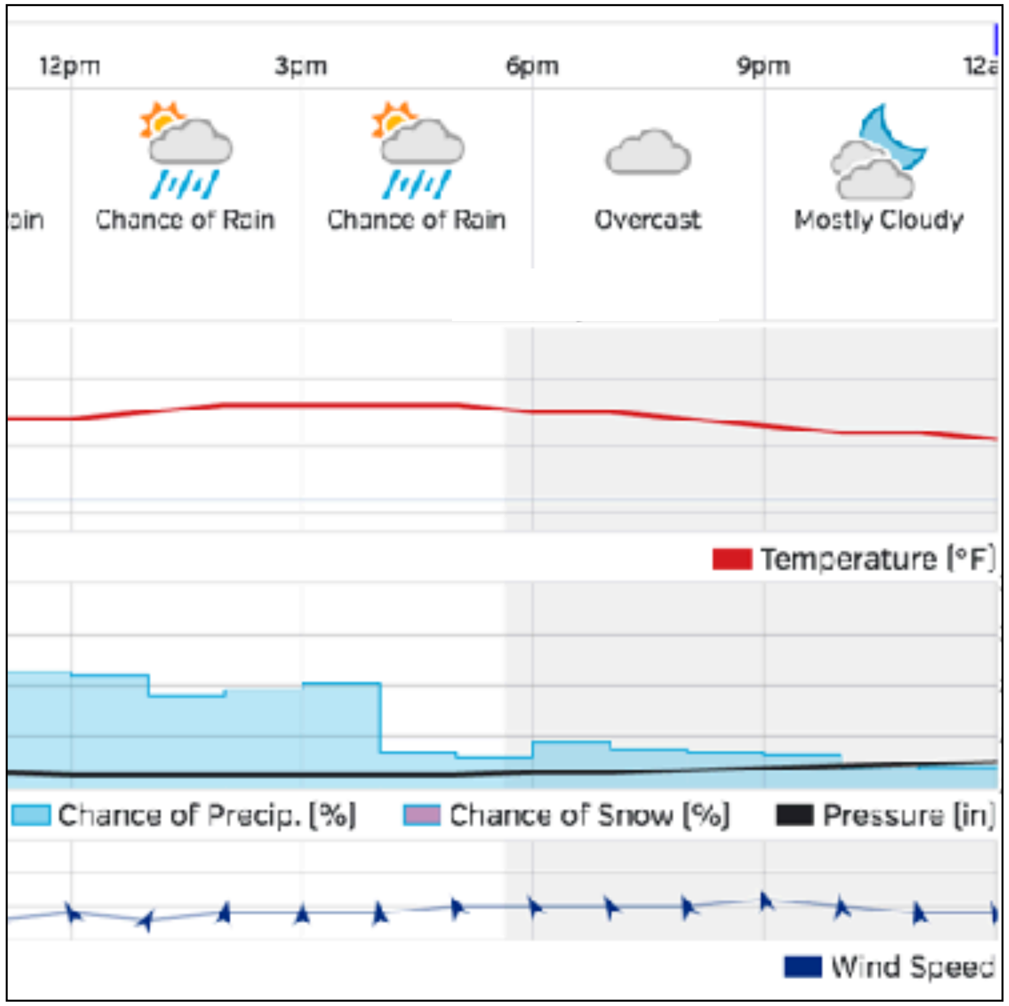
	time	min	mean	max	mode
wind	12-3	3	5	7	
wind	3-6	5	5	5	
wind	6-9	5	6	7	
dir	12-3				NW
dir	3-6				NE
dir	6-9				NE
temp	12-9	40	42	45	
precip	12-3	25	45	50	
precip	3-6	15	30	50	
precip	6-9	12	18	25	
cover	12-3				50-75
cover	3-6				50-75
cover	6-9				75-100

Chance of rain then becoming overcast, with a high of 45.
Calm to moderate northeast winds.

(Angeli et al. EMNLP 2010, Kim and Mooney COLING 2010)

(A Global Model for Concept-to-Text Generation. Konstas and Lapata, JAIR 2013; EMNLP 2013)

Concept-to-Text



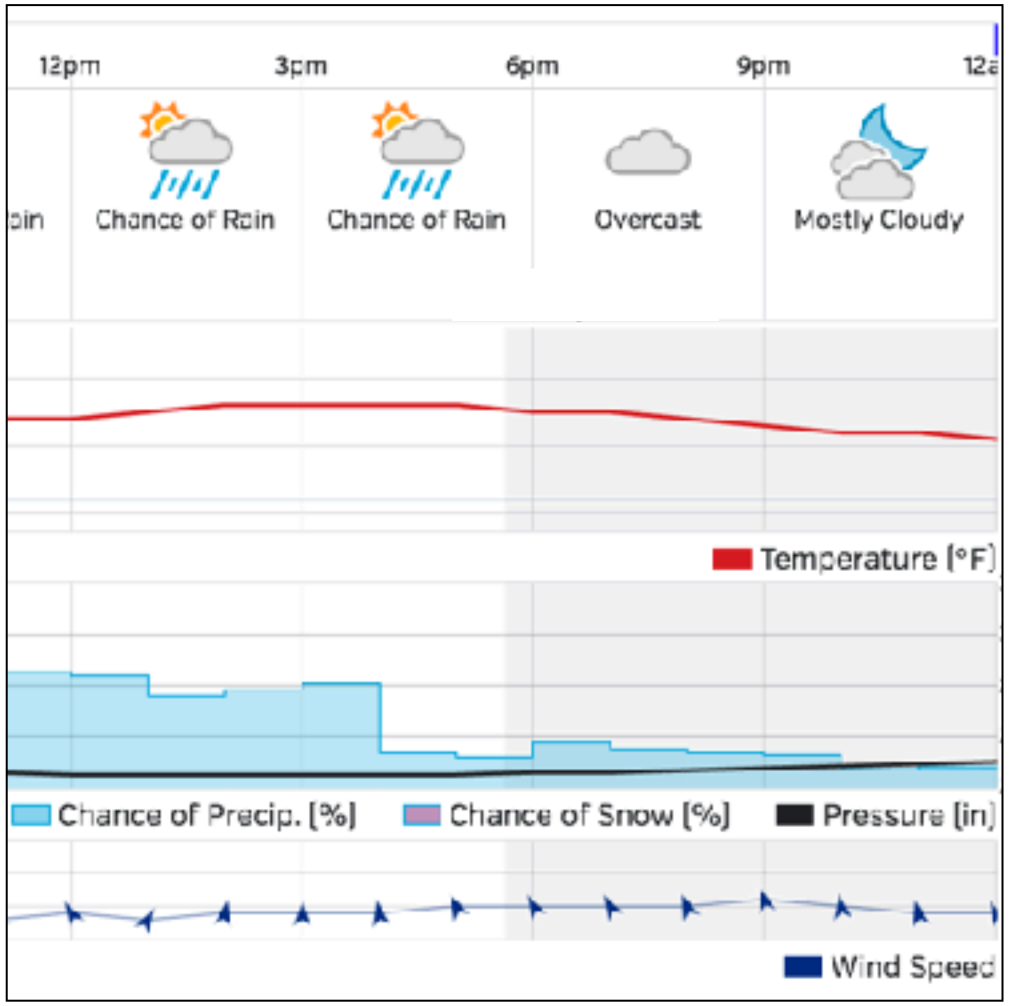
	time	min	mean	max	mode
wind	12-3	3	5	7	
wind	3-6	5	5	5	
wind	6-9	5	6	7	
dir	12-3				NW
dir	3-6				NE
dir	6-9				NE
temp	12-9	40	42	45	
precip	12-3	25	45	50	
precip	3-6	15	30	50	
precip	6-9	12	18	25	
cover	12-3				50-75
cover	3-6				50-75
cover	6-9				75-100

Chance of rain then becoming overcast, with a high of 45.
Calm to moderate northeast winds.

(Angeli et al. EMNLP 2010, Kim and Mooney COLING 2010)

(A Global Model for Concept-to-Text Generation. Konstas and Lapata, JAIR 2013; EMNLP 2013)

Concept-to-Text



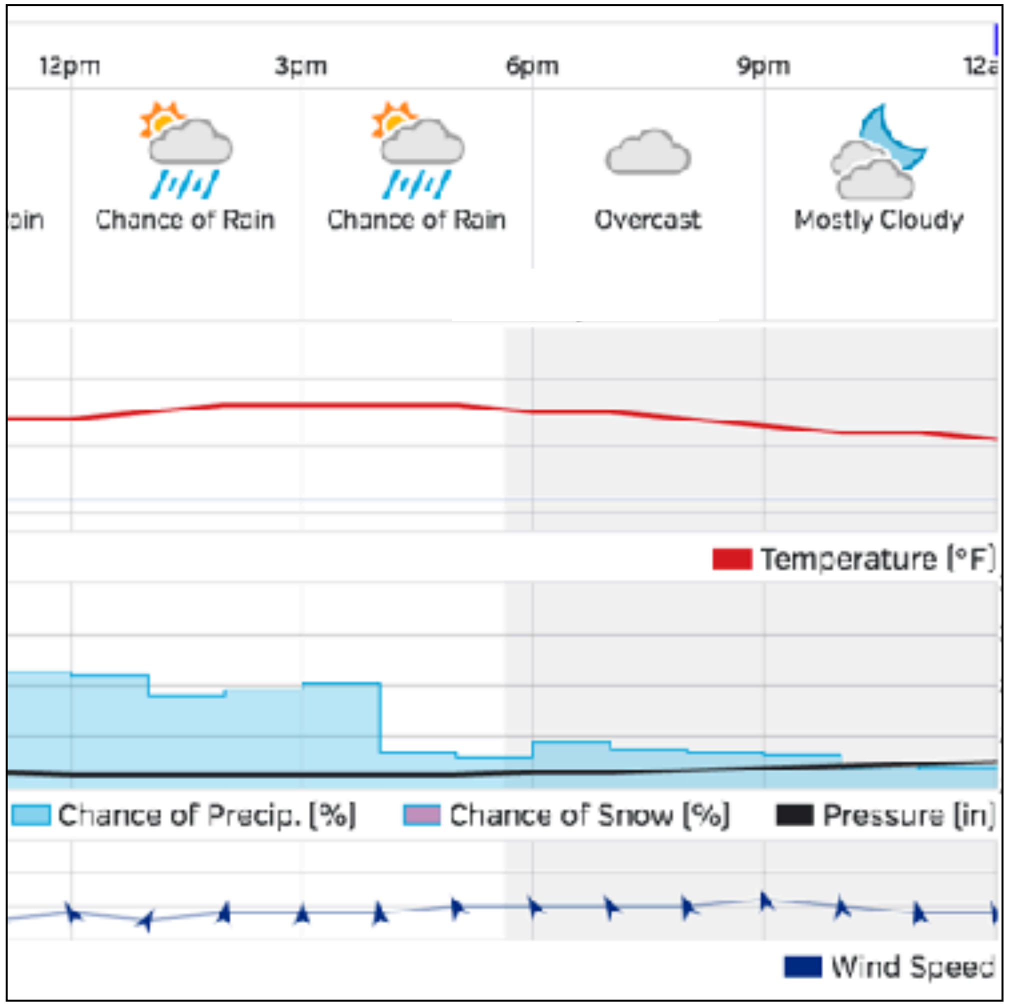
	time	min	mean	max	mode
wind	12-3	3	5	7	
wind	3-6	5	5	5	
wind	6-9	5	6	7	
dir	12-3				NW
dir	3-6				NE
dir	6-9				NE
temp	12-9	40	42	45	
precip	12-3	25	45	50	
precip	3-6	15	30	50	
precip	6-9	12	18	25	
cover	12-3				50-75
cover	3-6				50-75
cover	6-9				75-100

Chance of rain then becoming overcast, with a high of 45.
Calm to moderate northeast winds.

(Angeli et al. EMNLP 2010, Kim and Mooney COLING 2010)

(A Global Model for Concept-to-Text Generation. Konstas and Lapata, JAIR 2013; EMNLP 2013)

Concept-to-Text



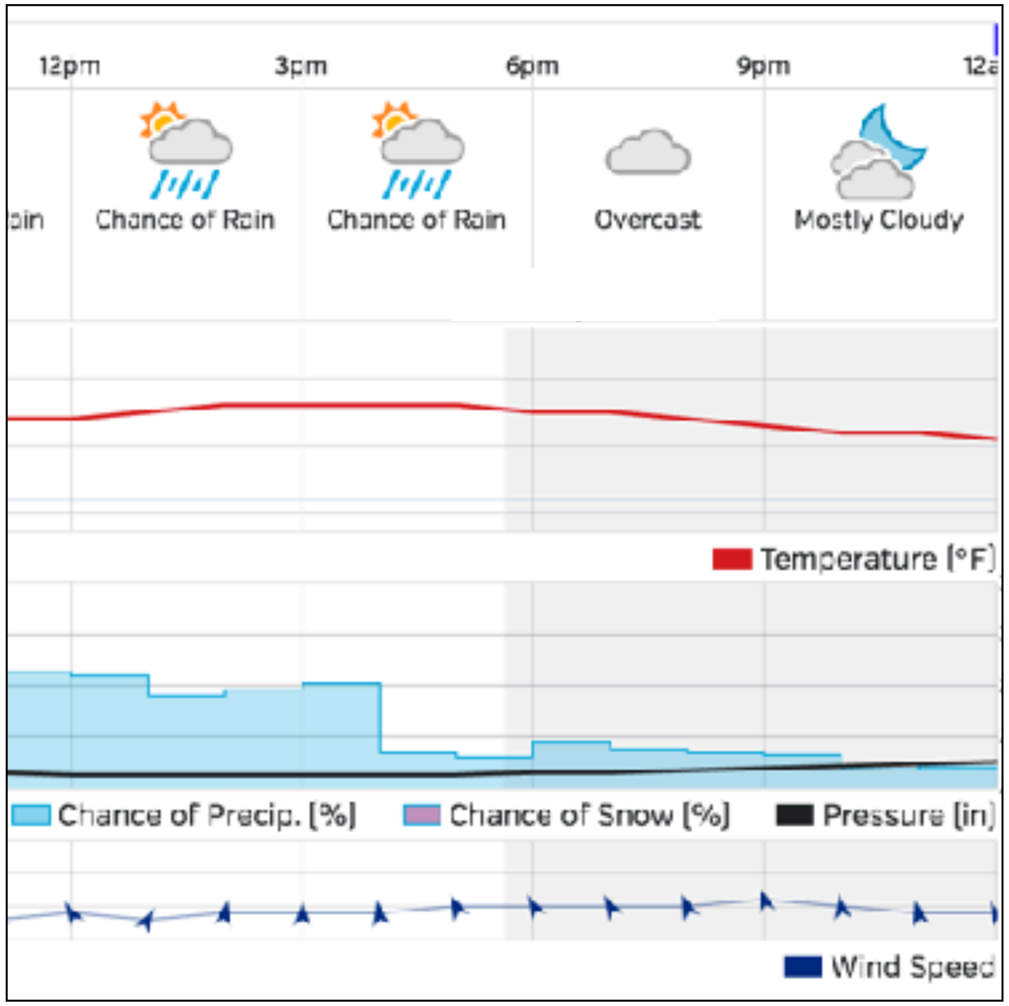
	time	min	mean	max	mode
wind	12-3	3	5	7	
wind	3-6	5	5	5	
wind	6-9	5	6	7	
dir	12-3				NW
dir	3-6				NE
dir	6-9				NE
temp	12-9	40	42	45	
precip	12-3	25	45	50	
precip	3-6	15	30	50	
precip	6-9	12	18	25	
cover	12-3				50-75
cover	3-6				50-75
cover	6-9				75-100

Chance of rain then becoming overcast, with a high of 45. Calm to moderate northeast winds.

(Angeli et al. EMNLP 2010, Kim and Mooney COLING 2010)

(A Global Model for Concept-to-Text Generation. Konstas and Lapata, JAIR 2013; EMNLP 2013)

Concept-to-Text



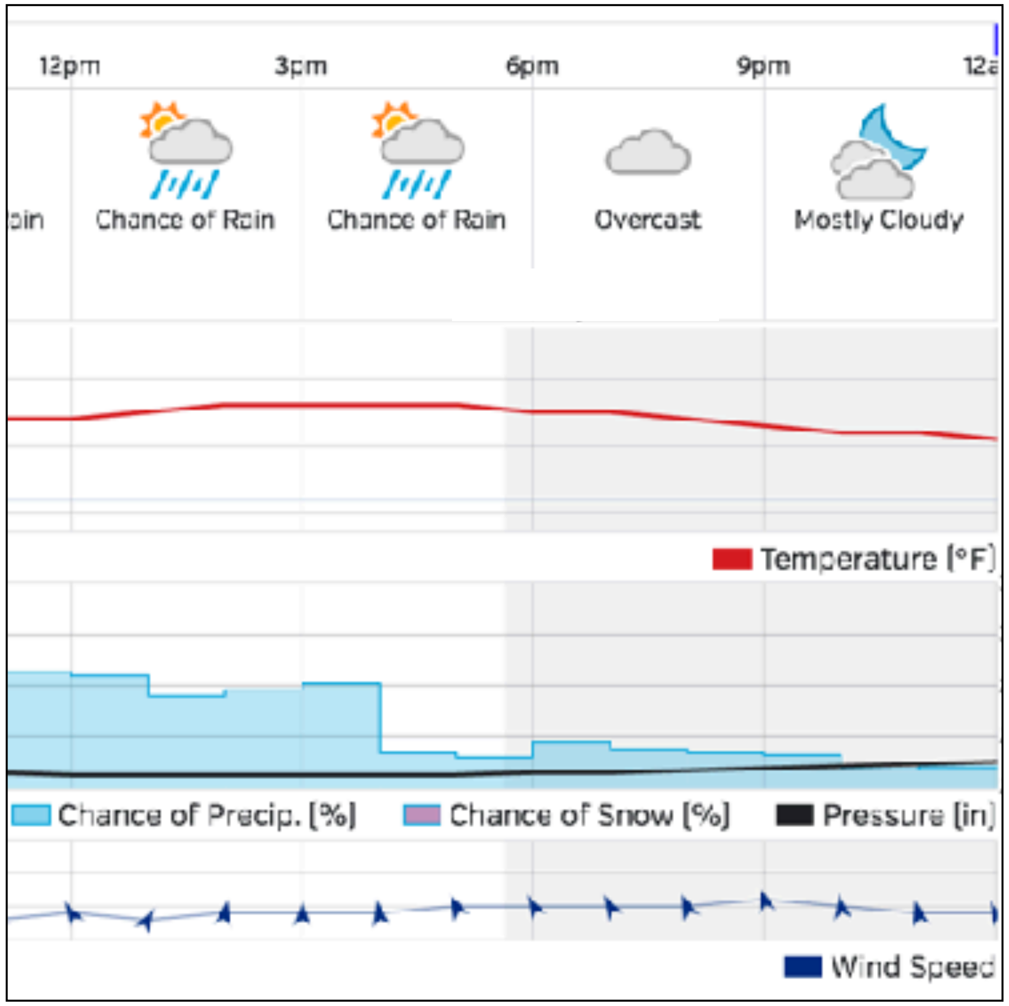
	time	min	mean	max	mode
wind	12-3	3	5	7	
wind	3-6	5	5	5	
wind	6-9	5	6	7	
dir	12-3				NW
dir	3-6				NE
dir	6-9				NE
temp	12-9	40	42	45	
precip	12-3	25	45	50	
precip	3-6	15	30	50	
precip	6-9	12	18	25	
cover	12-3				50-75
cover	3-6				50-75
cover	6-9				75-100

Chance of rain then becoming overcast, with a high of 45. Calm to moderate northeast winds.

(Angeli et al. EMNLP 2010, Kim and Mooney COLING 2010)

(A Global Model for Concept-to-Text Generation. Konstas and Lapata, JAIR 2013; EMNLP 2013)

Concept-to-Text



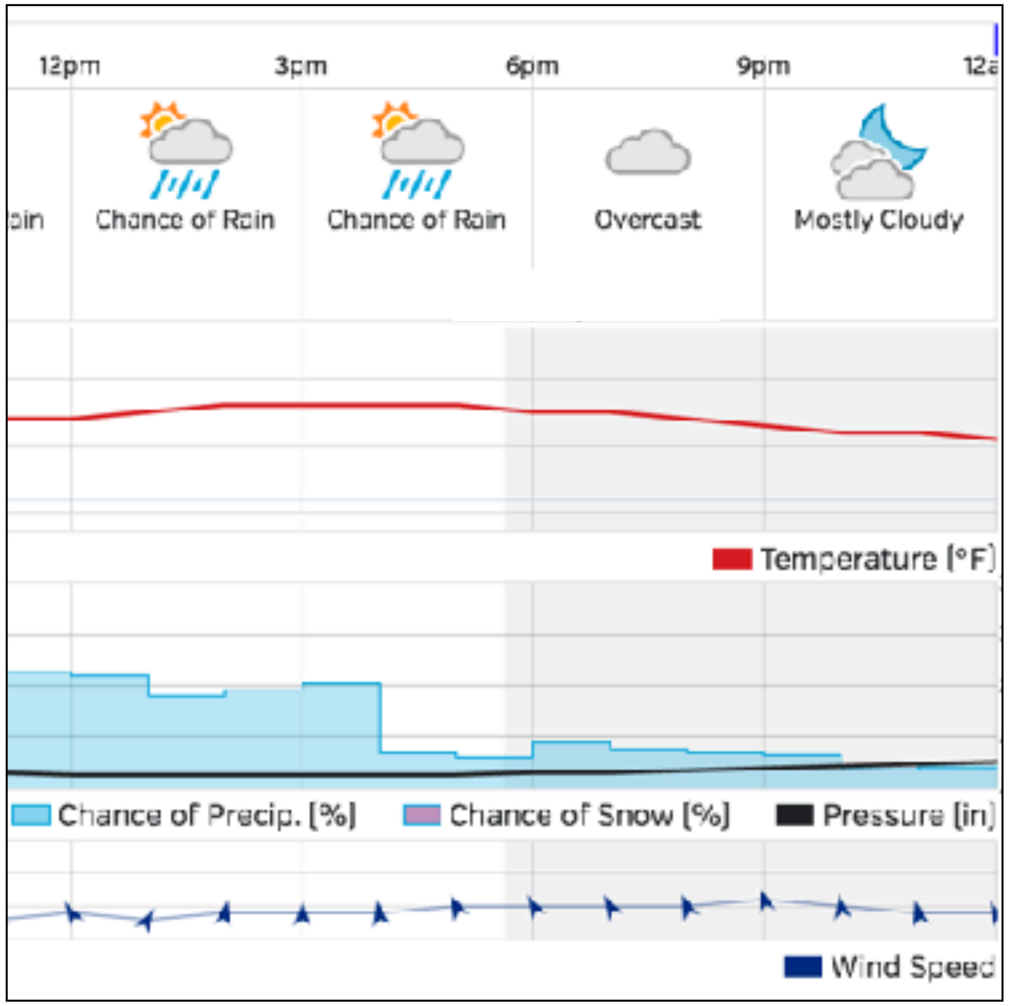
	time	min	mean	max	mode
wind	12-3	3	5	7	
wind	3-6	5	5	5	
wind	6-9	5	6	7	
dir	12-3				NW
dir	3-6				NE
dir	6-9				NE
temp	12-9	40	42	45	
precip	12-3	25	45	50	
precip	3-6	15	30	50	
precip	6-9	12	18	25	
cover	12-3				50-75
cover	3-6				50-75
cover	6-9				75-100

Chance of rain then becoming overcast, with a high of 45. Calm to moderate northeast winds.

(Angeli et al. EMNLP 2010, Kim and Mooney COLING 2010)

(A Global Model for Concept-to-Text Generation. Konstas and Lapata, JAIR 2013; EMNLP 2013)

Concept-to-Text



	time	min	mean	max	mode
wind	12-3	3	5	7	
wind	3-6	5	5	5	
wind	6-9	5	6	7	
dir	12-3				NW
dir	3-6				NE
dir	6-9				NE
temp	12-9	40	42	45	
precip	12-3	25	45	50	
precip	3-6	15	30	50	
precip	6-9	12	18	25	
cover	12-3				50-75
cover	3-6				50-75
cover	6-9				75-100

Chance of rain then becoming overcast, with a high of 45.
Calm to moderate northeast winds.

(Angeli et al. EMNLP 2010, Kim and Mooney COLING 2010)

(A Global Model for Concept-to-Text Generation. Konstas and Lapata, JAIR 2013; EMNLP 2013)

Concept-to-Text

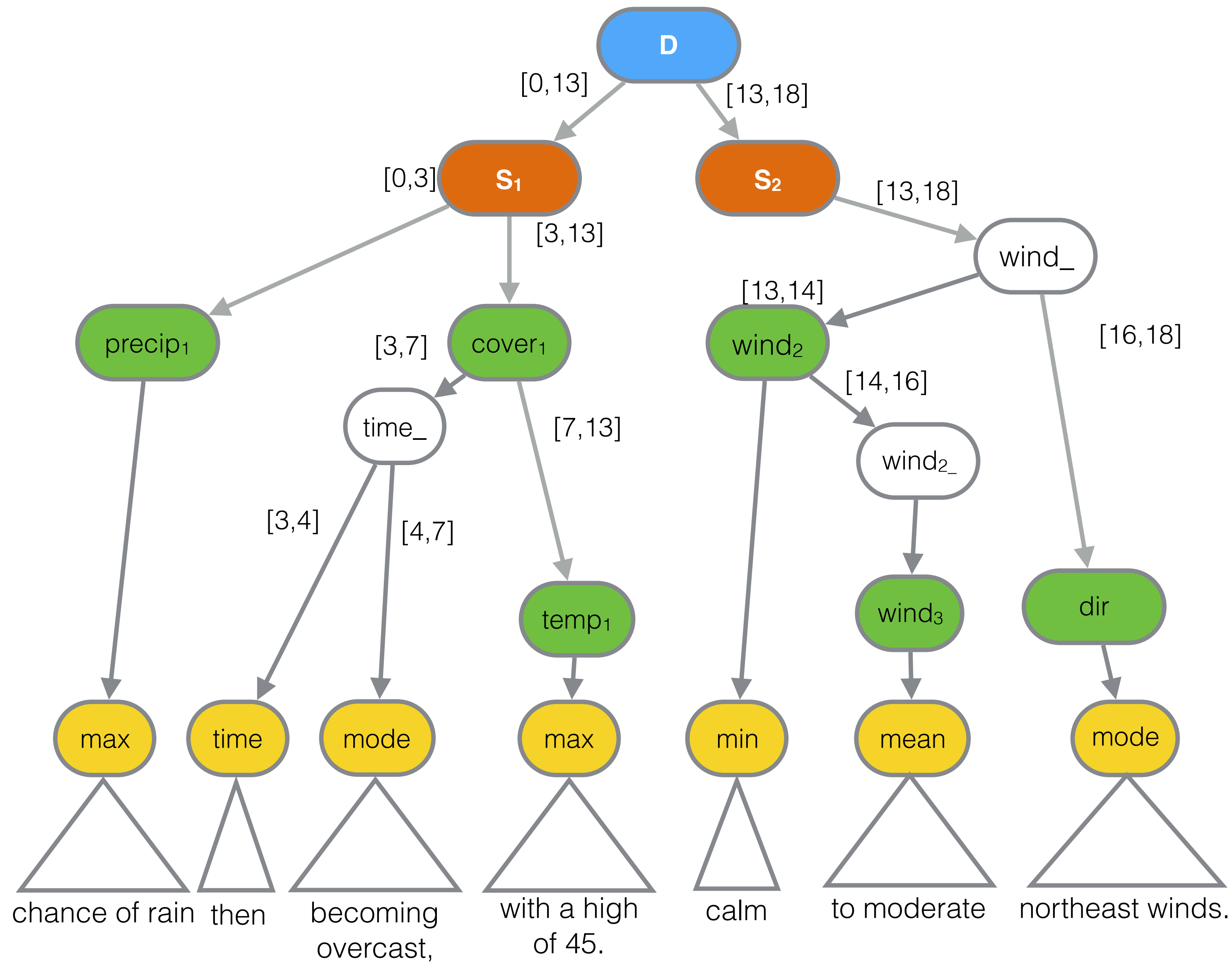
	time	min	mean	max	mode
wind	12-3	3	5	7	
wind	3-6	5	5	5	
wind	6-9	5	6	7	
dir	12-3				NW
dir	3-6				NE
dir	6-9				NE
temp	12-9	40	42	45	
precip	12-3	25	45	50	
precip	3-6	15	30	50	
precip	6-9	12	18	25	
cover	12-3				50-75
cover	3-6				50-75
cover	6-9				75-100

Chance of rain then becoming overcast, with a high of 45.
Calm to moderate northeast winds.

(A Global Model for Concept-to-Text Generation. Konstas and Lapata, JAIR 2013; EMNLP 2013)

Concept-to-Text

	time	min	mean	max	mode
wind	12-3	3	5	7	
wind	3-6	5	5	5	
wind	6-9	5	6	7	
dir	12-3				NW
dir	3-6				NE
dir	6-9				NE
temp	12-9	40	42	45	
precip	12-3	25	45	50	
precip	3-6	15	30	50	
precip	6-9	12	18	25	
cover	12-3				50-75
cover	3-6				50-75
cover	6-9				75-100

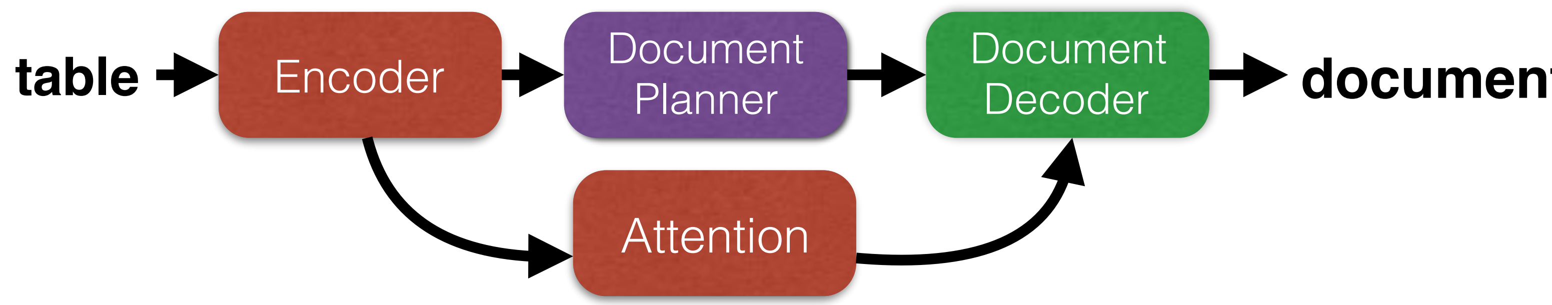


Chance of rain then becoming overcast, with a high of 45.
Calm to moderate northeast winds.

(A Global Model for Concept-to-Text Generation. Konstas and Lapata, JAIR 2013; EMNLP 2013)

Concept-to-Text

	time	min	mean	max	mode
wind	12-3	3	5	7	
wind	3-6	5	5	5	
wind	6-9	5	6	7	
dir	12-3				NW
dir	3-6				NE
dir	6-9				NE
temp	12-9	40	42	45	
precip	12-3	25	45	50	
precip	3-6	15	30	50	
precip	6-9	12	18	25	
cover	12-3				50-75
cover	3-6				50-75
cover	6-9				75-100

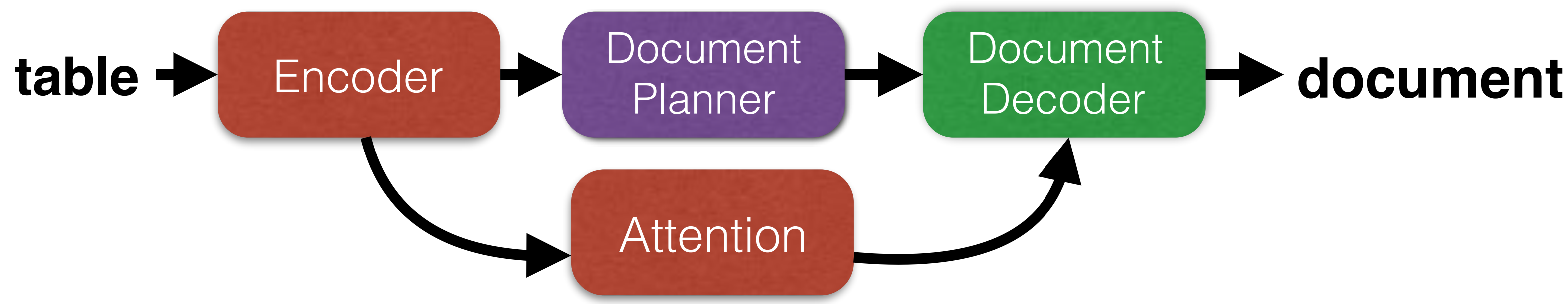


Chance of rain then becoming overcast, with a high of 45.
Calm to moderate northeast winds.

(Mei et al., NAACL 2016)

Concept-to-Text

	time	min	mean	max	mode
wind	12-3	3	5	7	
wind	3-6	5	5	5	
wind	6-9	5	6	7	
dir	12-3				NW
dir	3-6				NE
dir	6-9				NE
temp	12-9	40	42	45	
precip	12-3	25	45	50	
precip	3-6	15	30	50	
precip	6-9	12	18	25	
cover	12-3				50-75
cover	3-6				50-75
cover	6-9				75-100

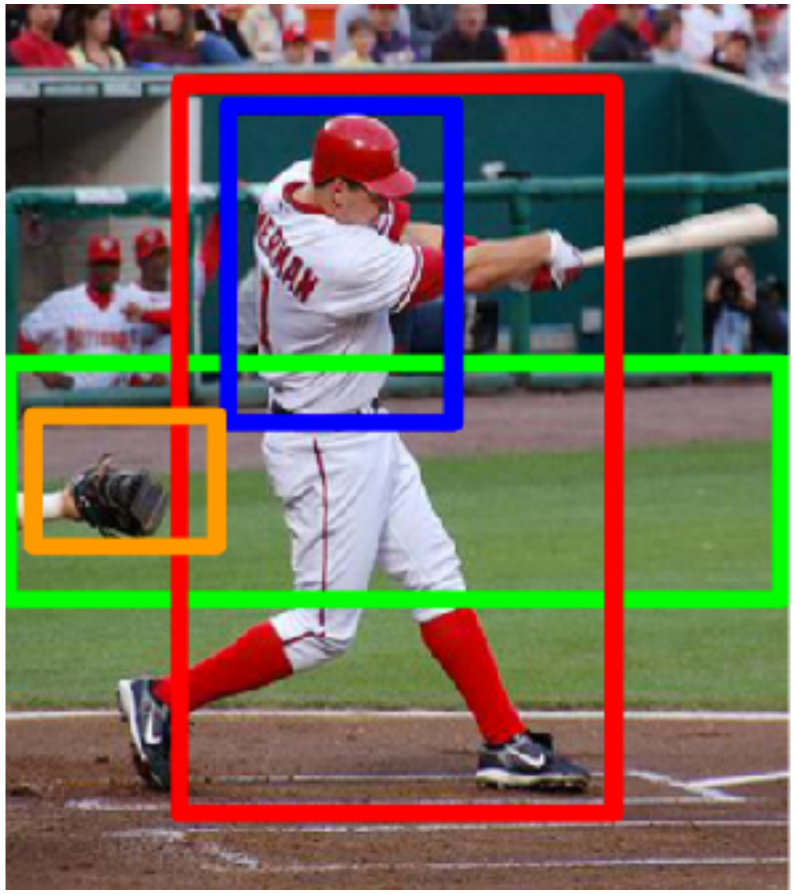


- **Document plan** based on:
  - sequences of records
  - discourse relations

Chance of rain then becoming overcast, with a high of 45.
Calm to moderate northeast winds.

(Mei et al., NAACL 2016)

Caption Generation



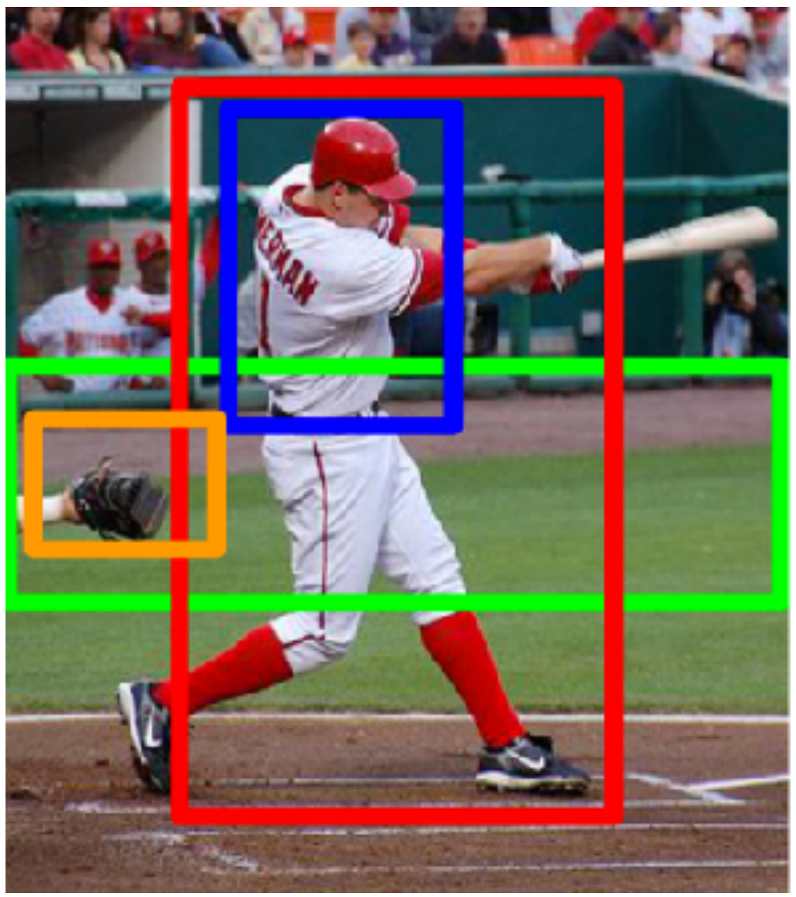
hitting				
agent	victim	victim part	tool	place
ballplayer	baseball	-	baseball bat	baseball diamond

wearing		
wearer	clothing	body part
ballplayer	red helmet	head

wearing		
wearer	clothing	body part
ballplayer	white shirt	torso

(Krause et al., CVPR 2017, Yatskar et al., CVPR 2016, Krishna et al., 2016)

Caption Generation

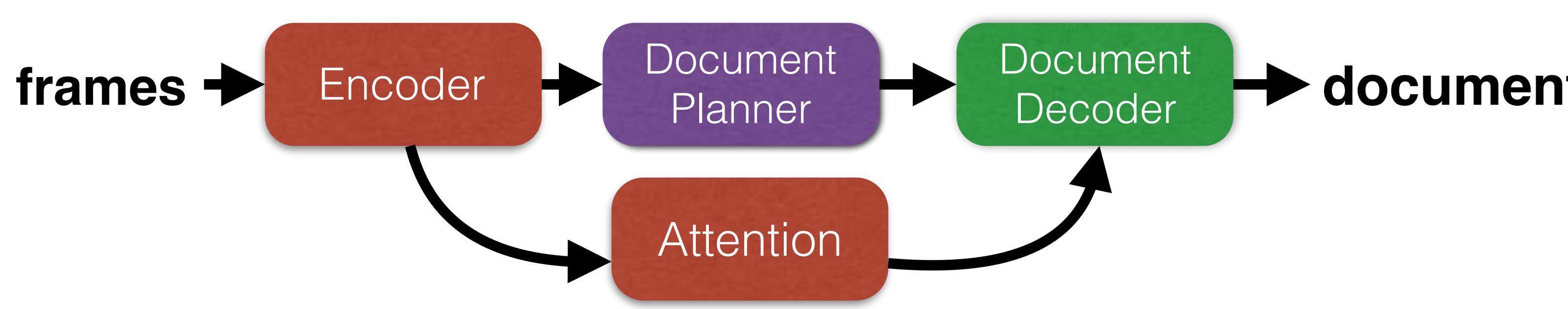


hitting				
agent	victim	victim part	tool	place
ballplayer	baseball	-	baseball bat	baseball diamond

wearing		
wearer	clothing	body part
ballplayer	red helmet	head

wearing		
wearer	clothing	body part
ballplayer	white shirt	torso

A baseball player is **swinging** a bat.
He is **wearing** a red helmet and a white shirt.
The catcher's mitt **is behind** the batter.



(Krause et al., CVPR 2017, Yatskar et al., CVPR 2016, Krishna et al., 2016)

Semantic-based Machine Translation

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

(Jones et al., COLING 2012)

Semantic-based Machine Translation

Source

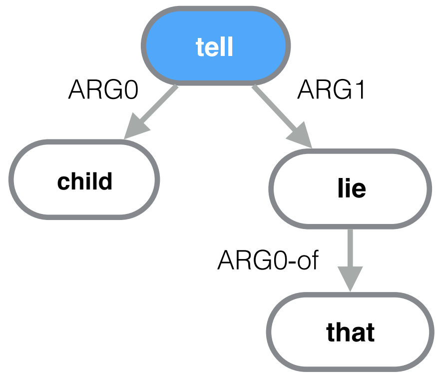
The children told that lie

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

(Jones et al., COLING 2012)

Semantic-based Machine Translation



Source

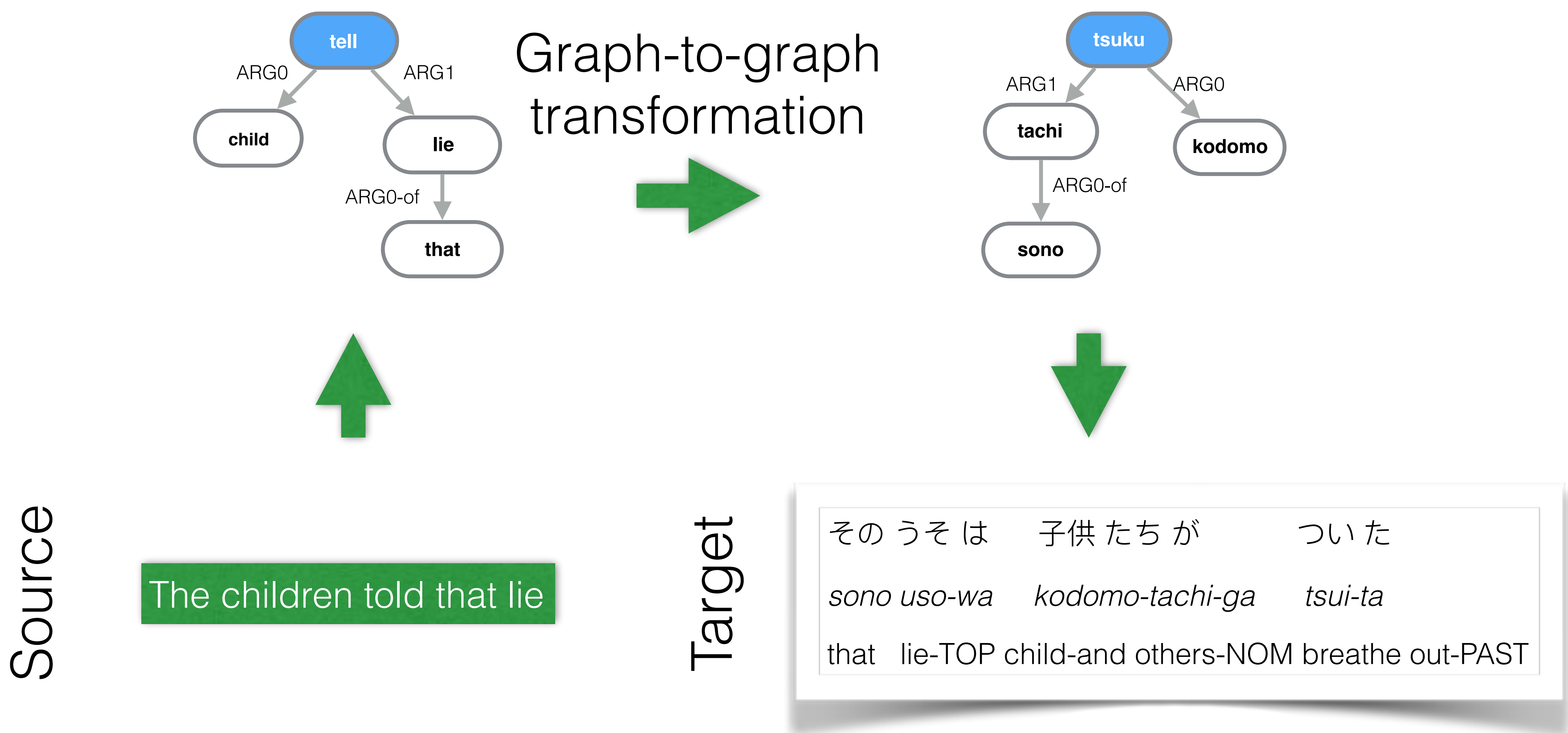
The children told that lie

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

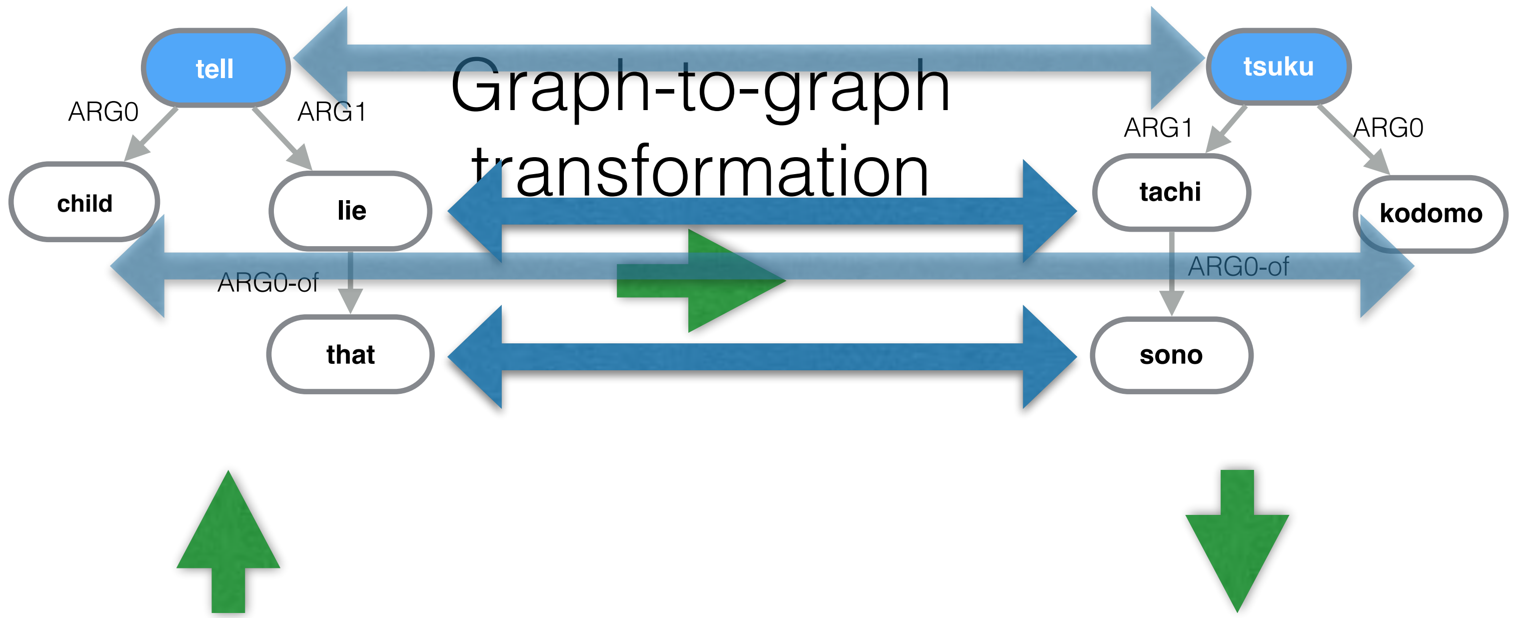
(Jones et al., COLING 2012)

Semantic-based Machine Translation



(Jones et al., COLING 2012)

Semantic-based Machine Translation



Source

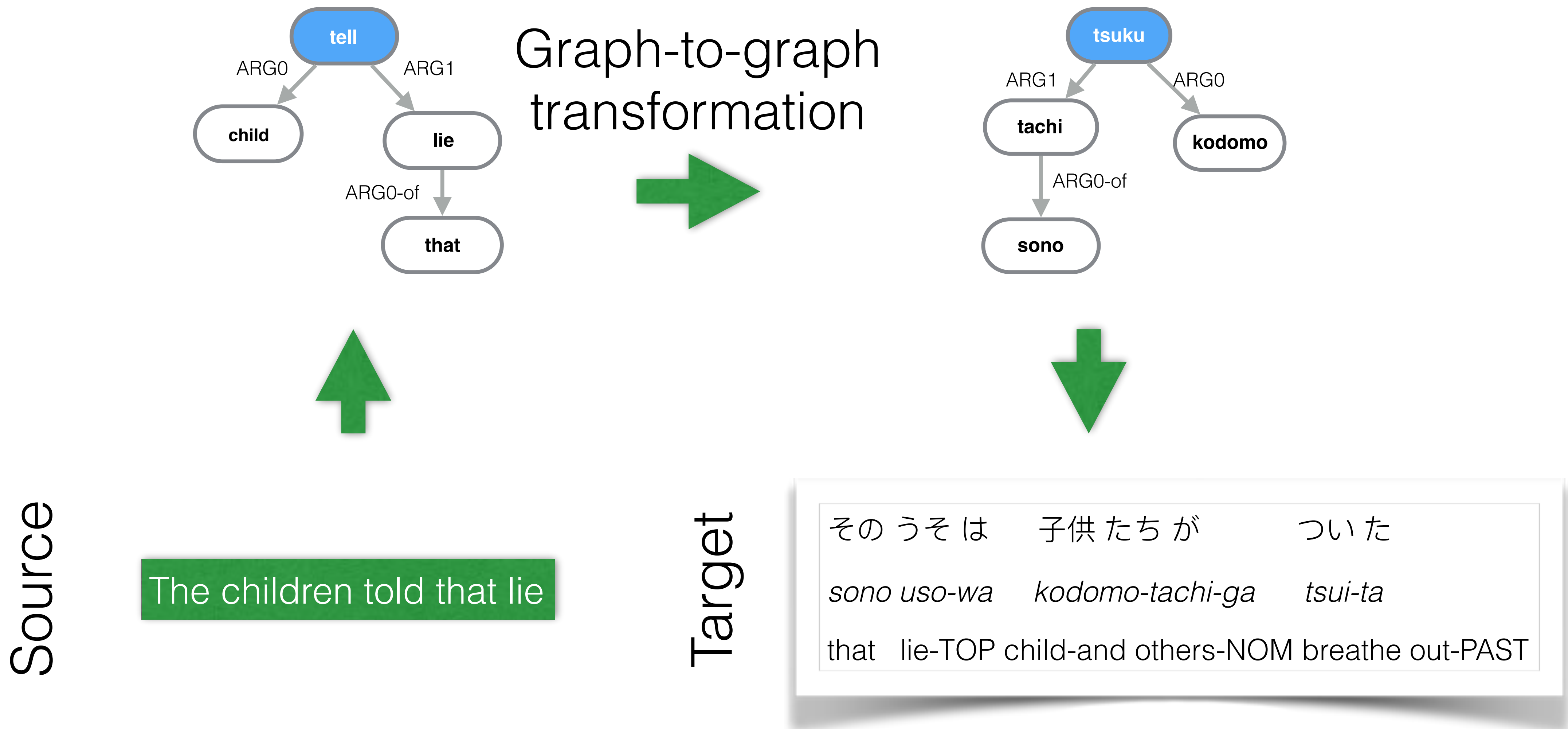
The children told that lie

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

(Jones et al., COLING 2012)

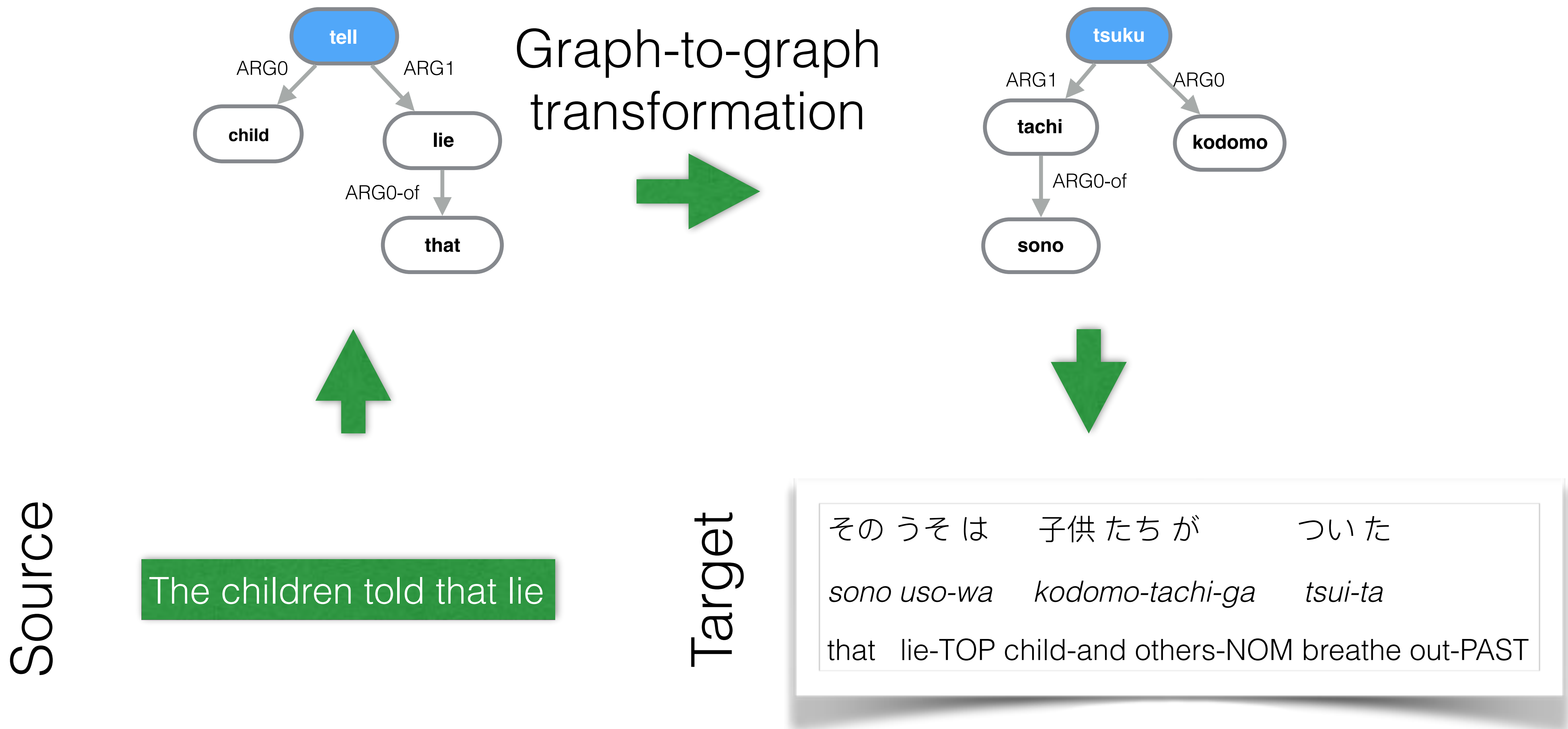
Semantic-based Machine Translation



- No Japanese AMR corpus

(Jones et al., COLING 2012)

Semantic-based Machine Translation

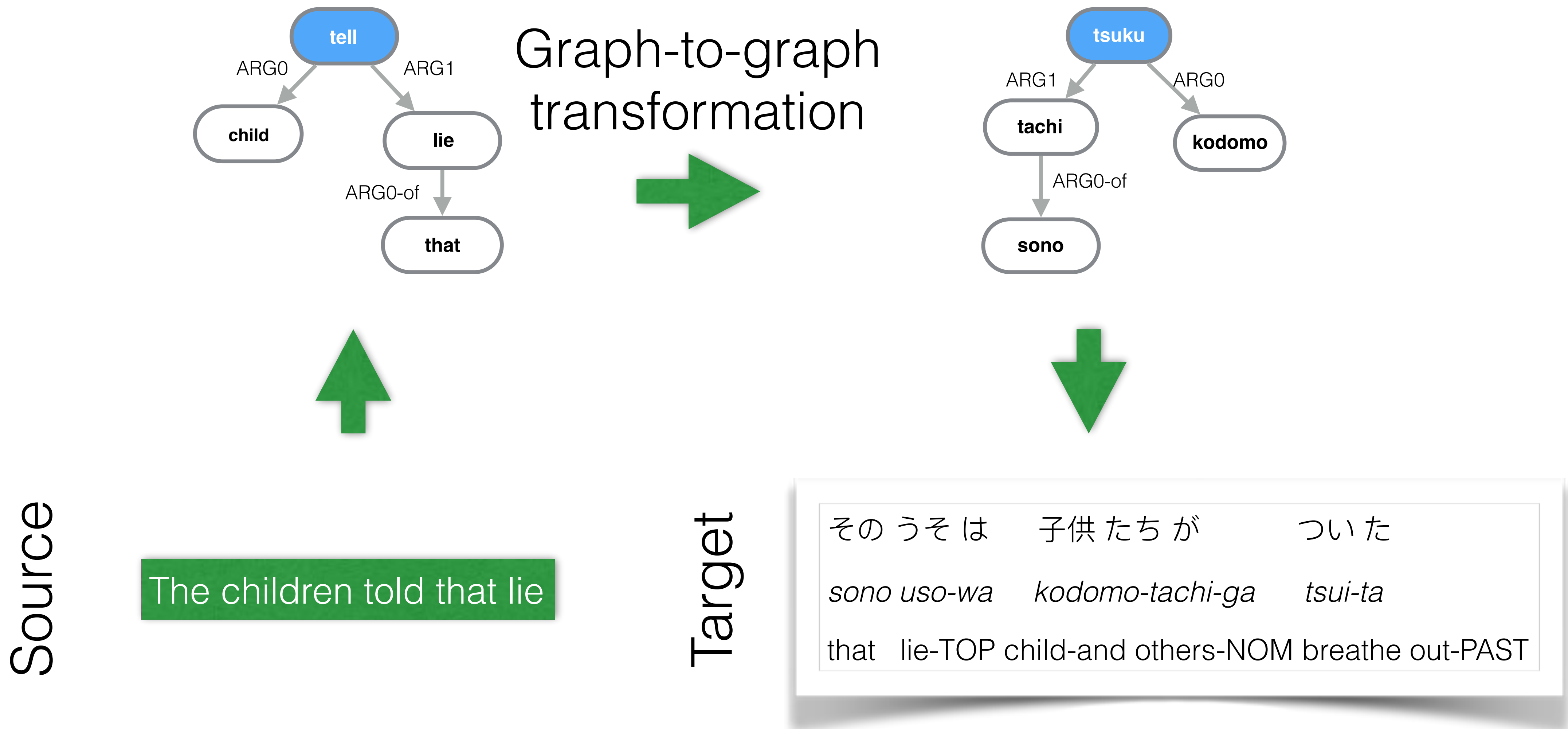


- No Japanese AMR corpus
- MRS hand-crafted grammars (Minimal Recursion Semantics; Copestake et al., RLC 2006)

Joint work with
Michael Wayne Goodman

(Jones et al., COLING 2012)

Semantic-based Machine Translation



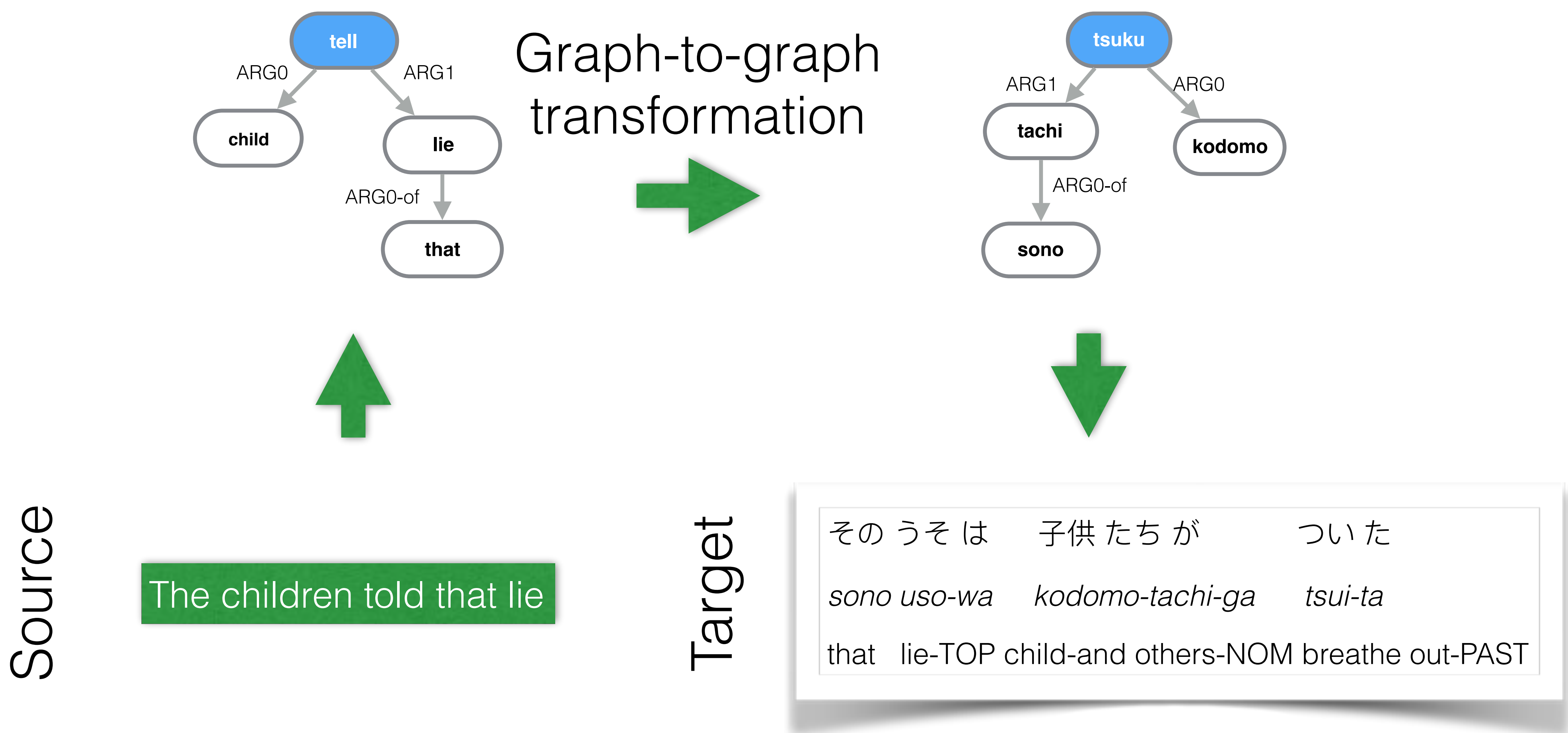
- No Japanese AMR corpus
- MRS hand-crafted grammars (Minimal Recursion Semantics; Copestake et al., RLC 2006)

1) **Parse to MRS** from English

Joint work with
Michael Wayne Goodman

(Jones et al., COLING 2012)

Semantic-based Machine Translation



- No Japanese AMR corpus
- MRS hand-crafted grammars (Minimal Recursion Semantics; Copestake et al., RLC 2006)

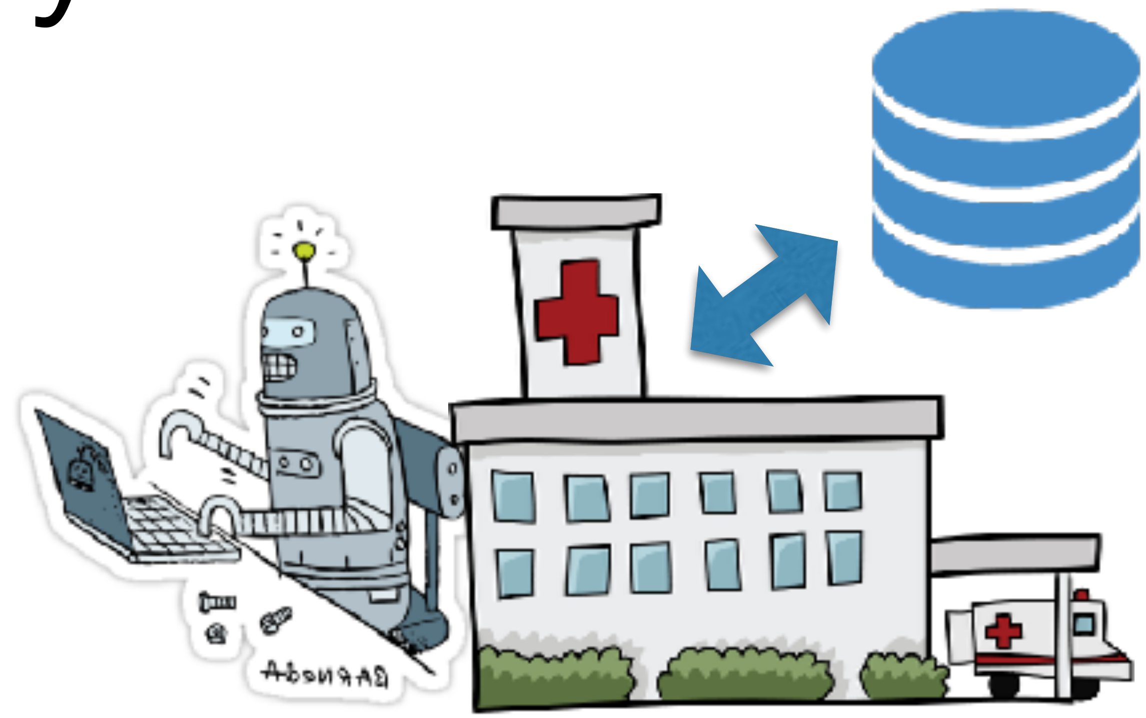
1) **Parse to MRS** from English

2) **Generate** Japanese from MRS

Joint work with
Michael Wayne Goodman

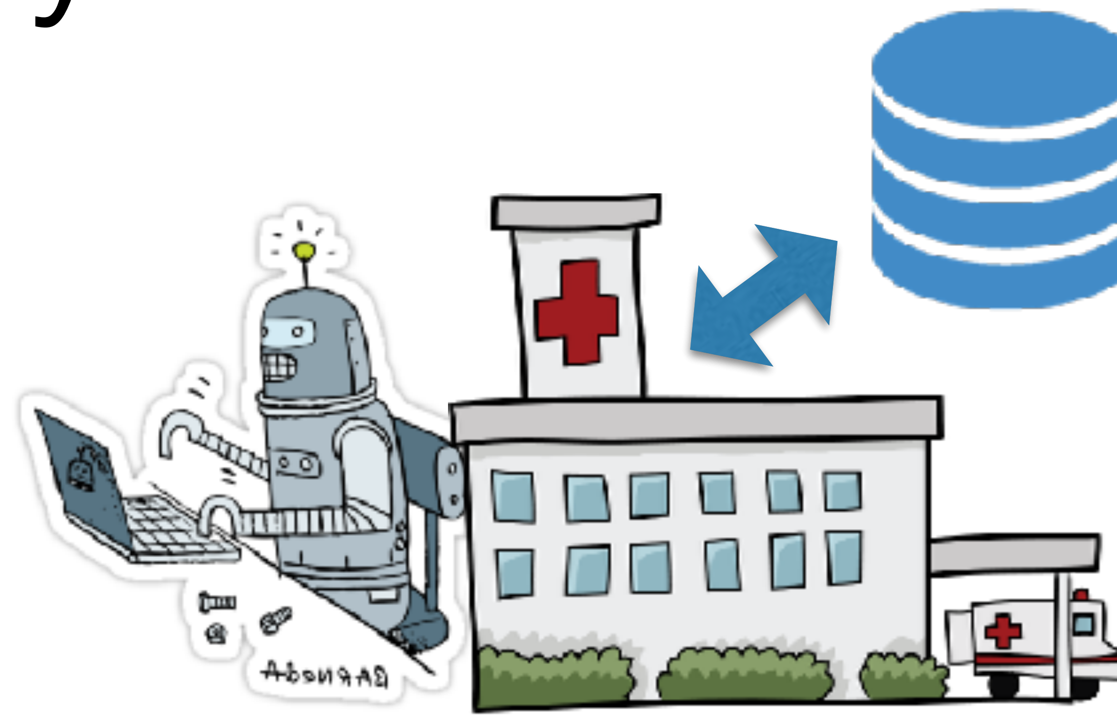
(Jones et al., COLING 2012)

Dialogue Systems



(Acharya et al., INLG 2016, Rieser et al., IEEE/ACM 2014)

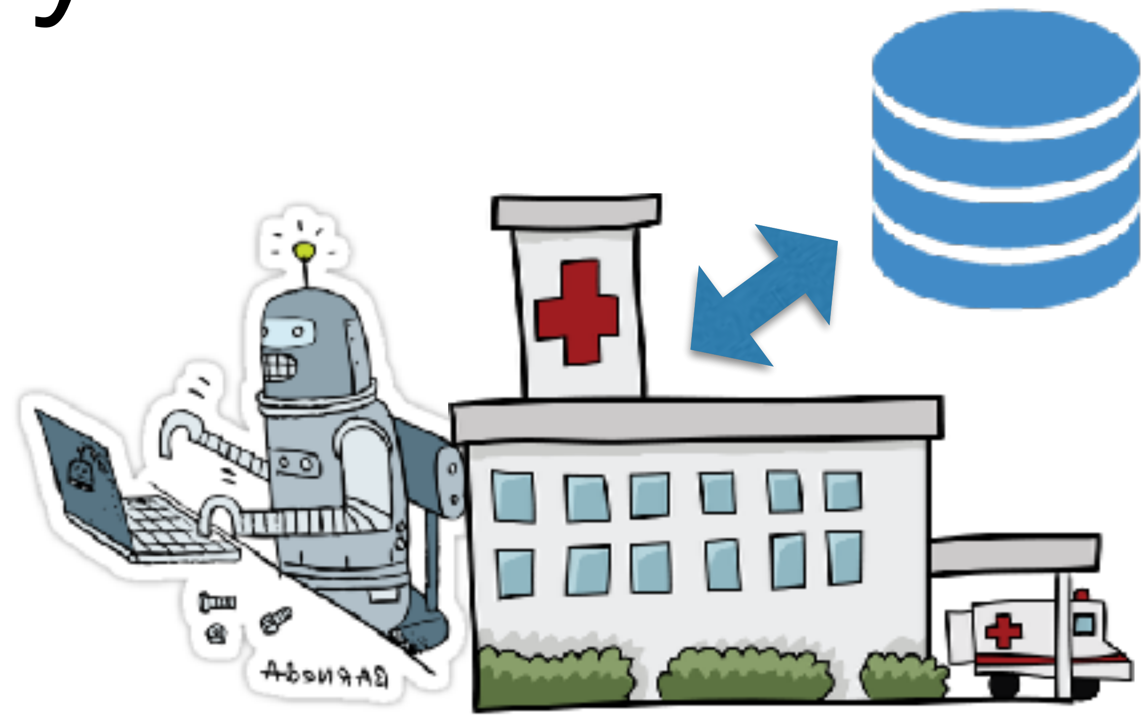
Dialogue Systems



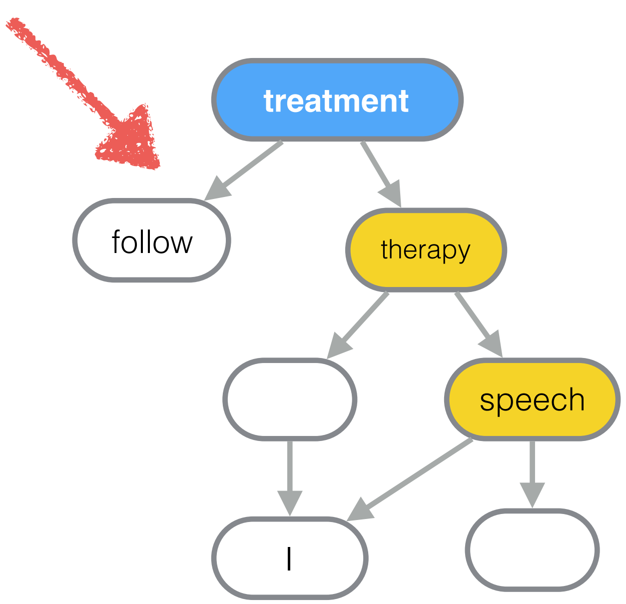
> I would like to follow up on
  my speech therapy treatment.

(Acharya et al., INLG 2016, Rieser et al., IEEE/ACM 2014)

Dialogue Systems

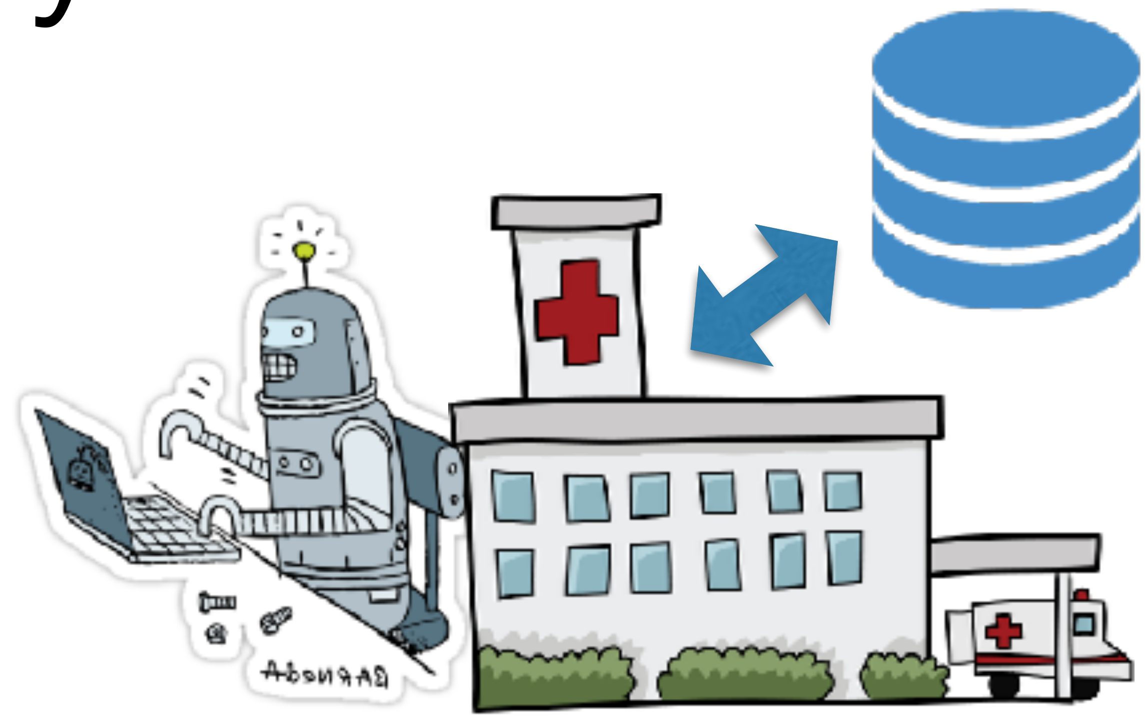


```
> I would like to follow up on
  my speech therapy treatment.
```

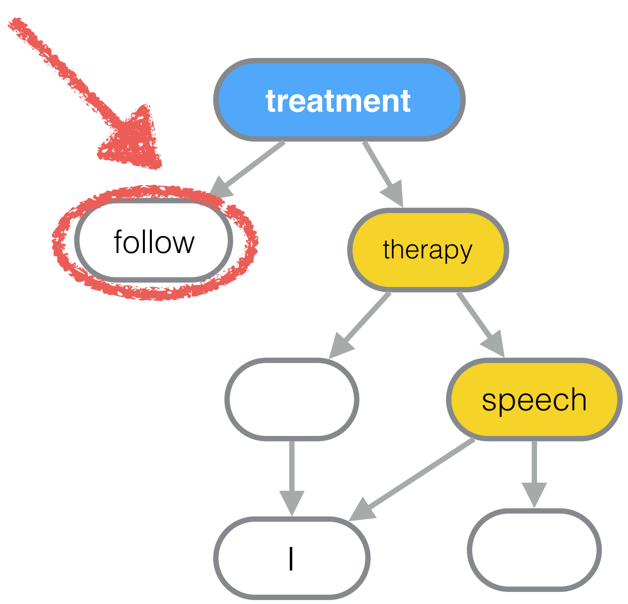


(Acharya et al., INLG 2016, Rieser et al., IEEE/ACM 2014)

Dialogue Systems

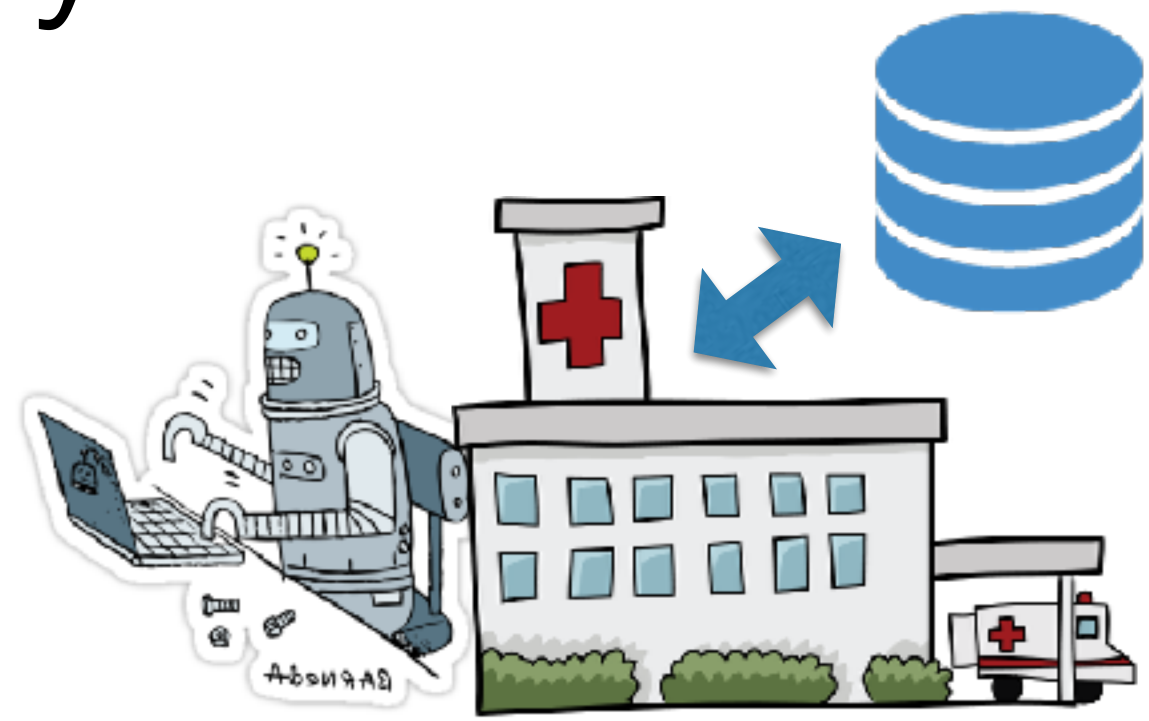


```
> I would like to follow up on
  my speech therapy treatment.
```

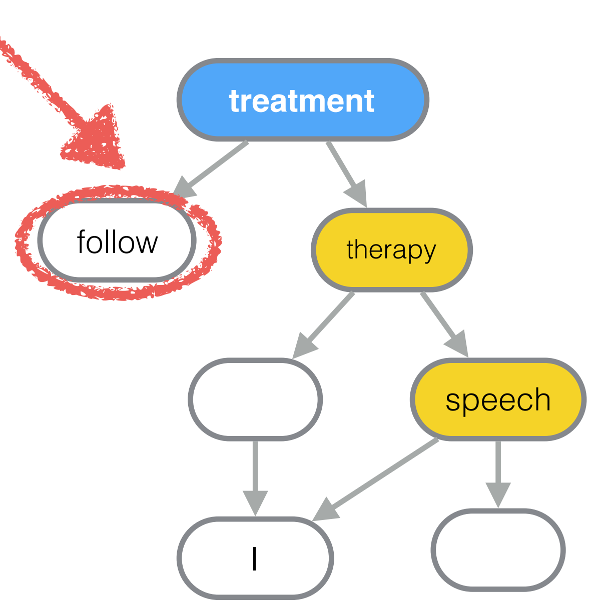


(Acharya et al., INLG 2016, Rieser et al., IEEE/ACM 2014)

Dialogue Systems



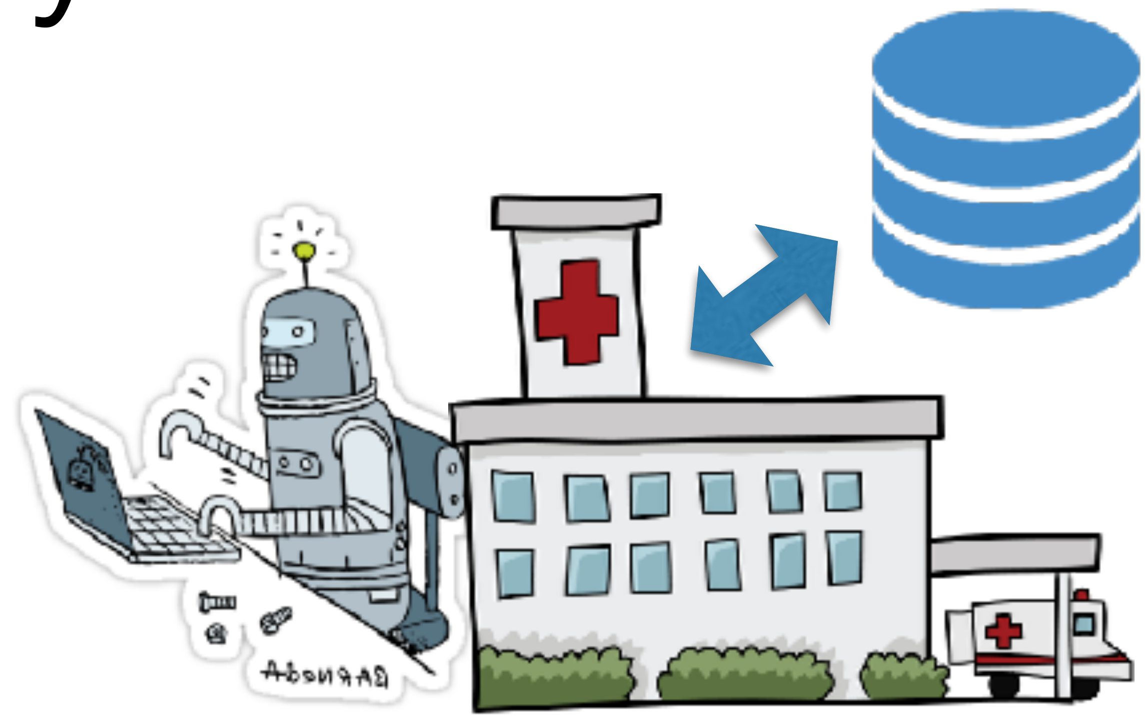
```
> I would like to follow up on
  my speech therapy treatment.
```



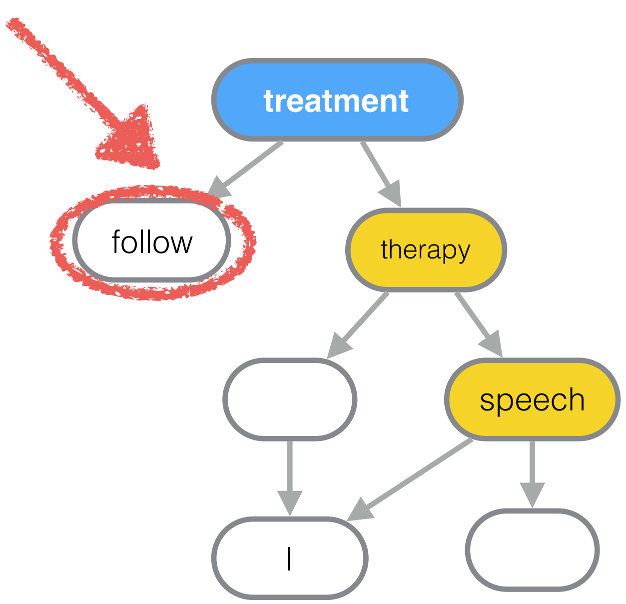
Patient #3245 Log:
You were admitted for acute subcortical cerebrovascular accident. […]
Verbal impairment related to **communication impairment** was **treated** with **speech therapy 3 months ago**.
[...]

(Acharya et al., INLG 2016, Rieser et al., IEEE/ACM 2014)

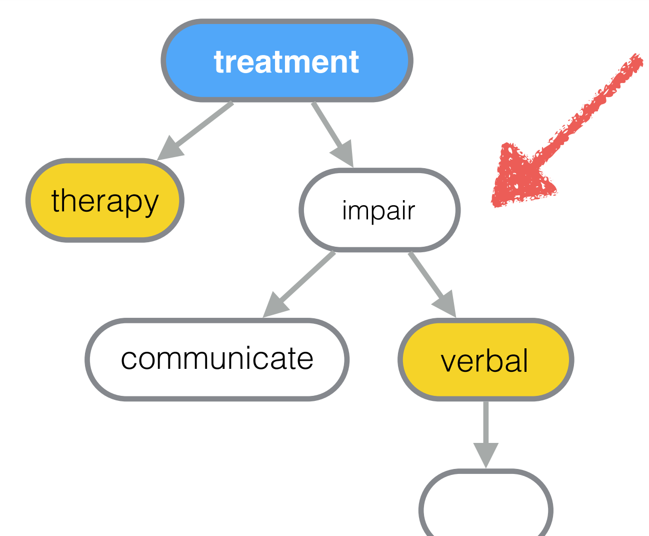
Dialogue Systems



> I would like to follow up on
  my speech therapy treatment.

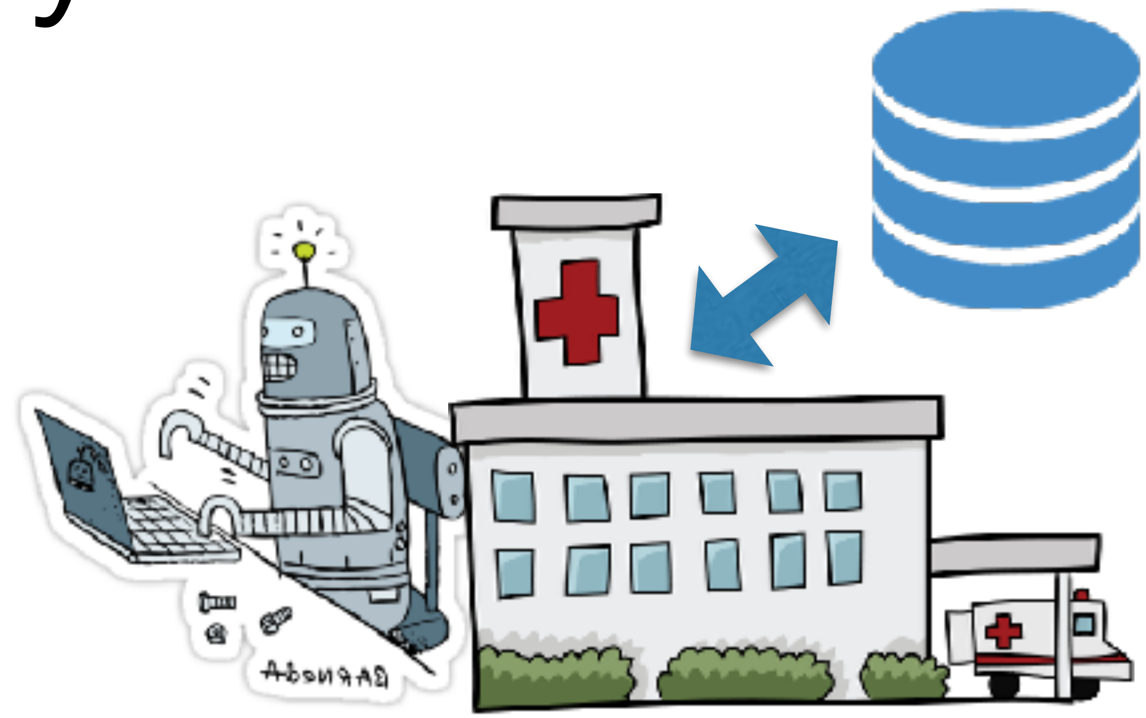


Patient #3245 Log:
You were admitted for acute subcortical cerebrovascular accident. […]
Verbal impairment related to **communication impairment** was **treated** with **speech therapy 3 months ago**.
[...]

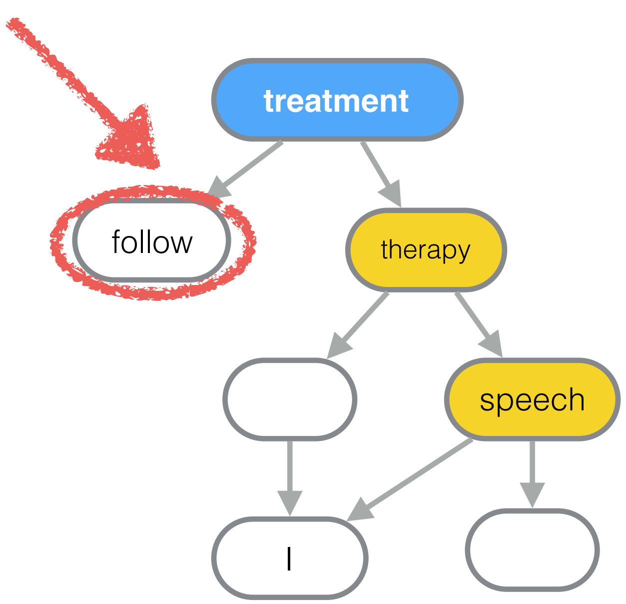


(Acharya et al., INLG 2016, Rieser et al., IEEE/ACM 2014)

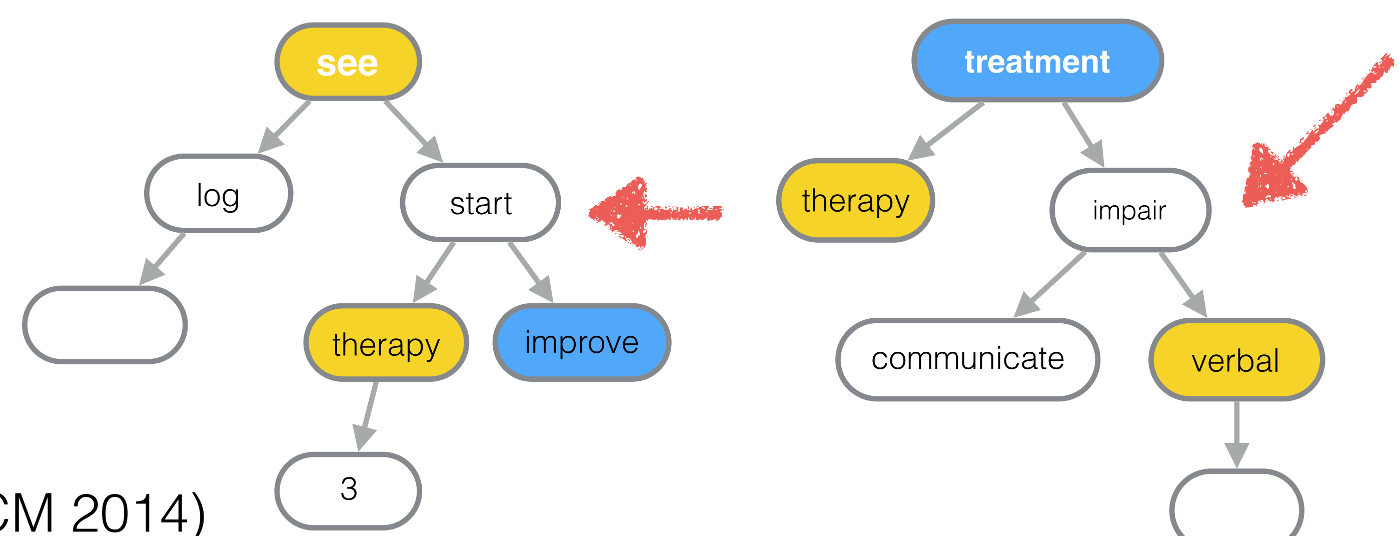
Dialogue Systems



```
> I would like to follow up on
  my speech therapy treatment.
```

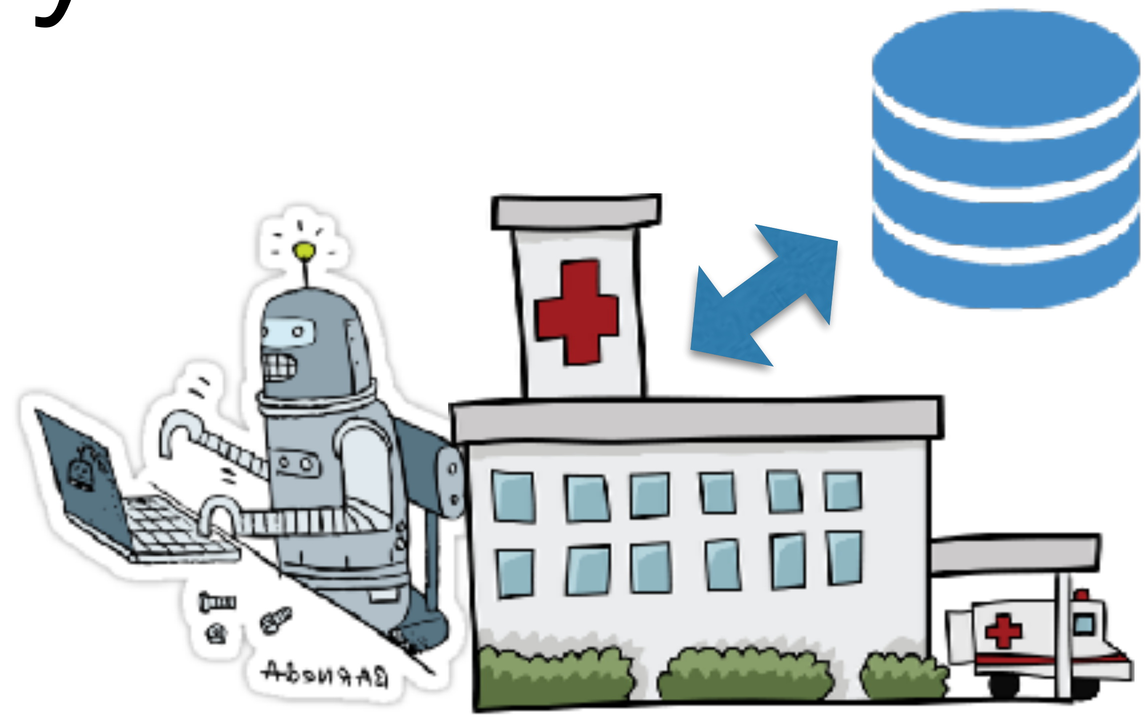


Patient #3245 Log:
You were admitted for acute subcortical cerebrovascular accident. […]
Verbal impairment related to **communication impairment** was **treated** with **speech therapy 3 months ago**.
[...]



(Acharya et al., INLG 2016, Rieser et al., IEEE/ACM 2014)

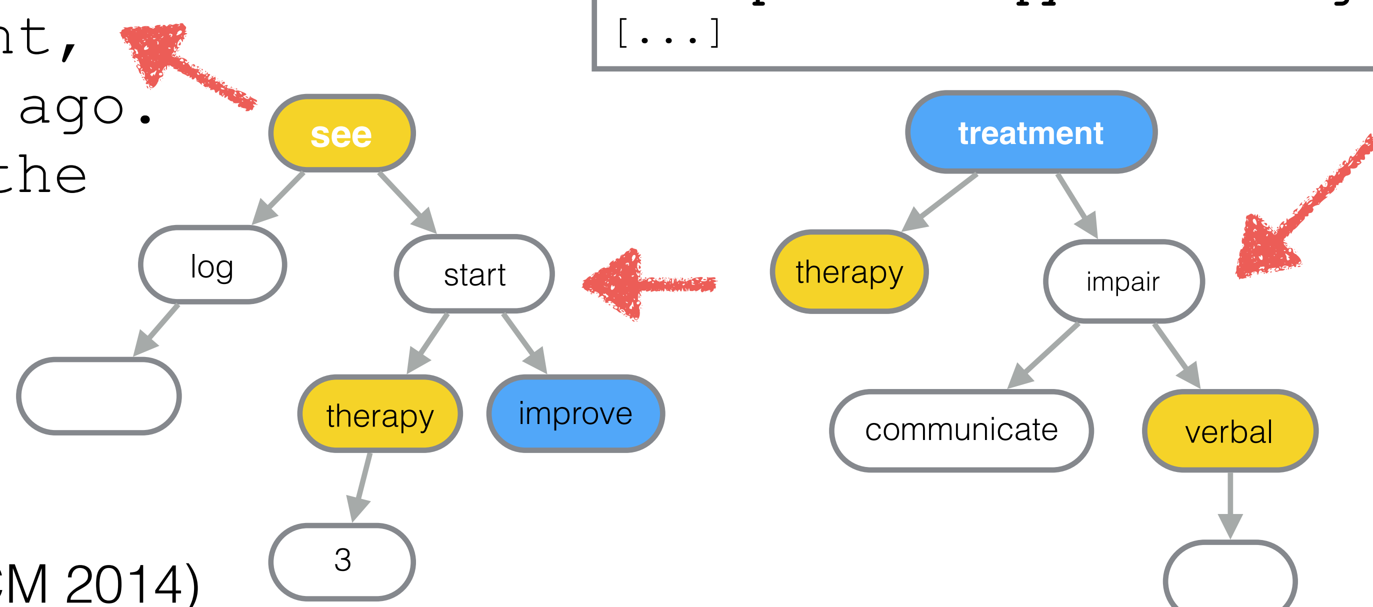
Dialogue Systems



> I would like to follow up on
my speech therapy treatment.

< I can see in my logs, that we
started improving verbal
impairment due to the accident,
with speech therapy 3 months ago.
When would you like to book the
next appointment?

Patient #3245 Log:
You were admitted for acute
subcortical
cerebrovascular accident. […]
Verbal impairment related to
communication impairment was **treated**
with **speech therapy 3 months ago**.
[...]



(Acharya et al., INLG 2016, Rieser et al., IEEE/ACM 2014)

Summary

- General data-driven approach for NLG
- Facilitates the deployment to new domains
- Integrates to existing systems and applications

Summary

- General data-driven approach for NLG
- Facilitates the deployment to new domains
- Integrates to existing systems and applications

Thank You