NeuGen

Text Generation from Meaning Representations

Yannis Konstas

Joint work with
Mark Yatskar, Luke Zettlemoyer and Yejin Choi (UW)
~
Yonatan Bisk and Daniel Marcu (ISI)

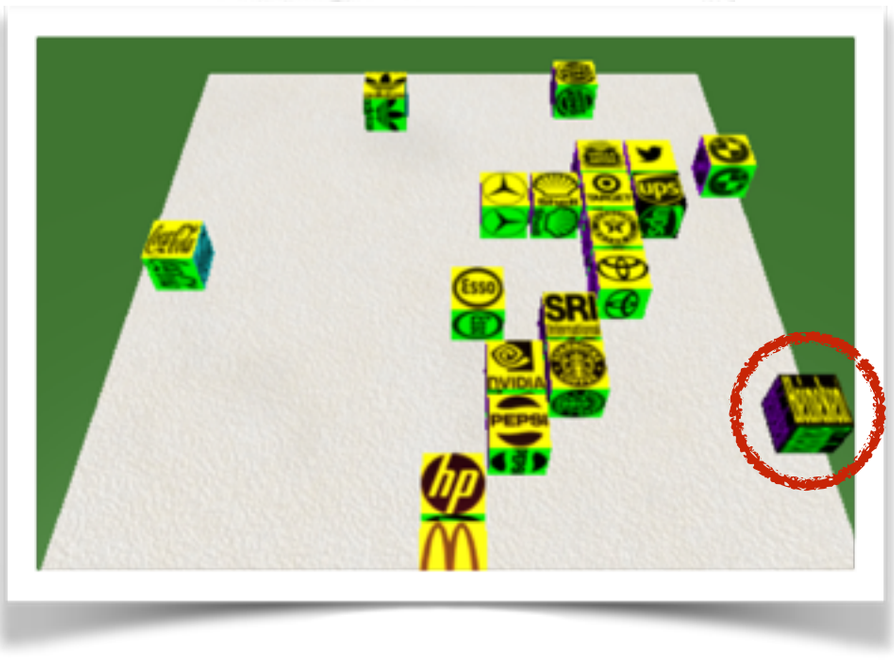
Motivation

Motivation

Machine-generated Representation

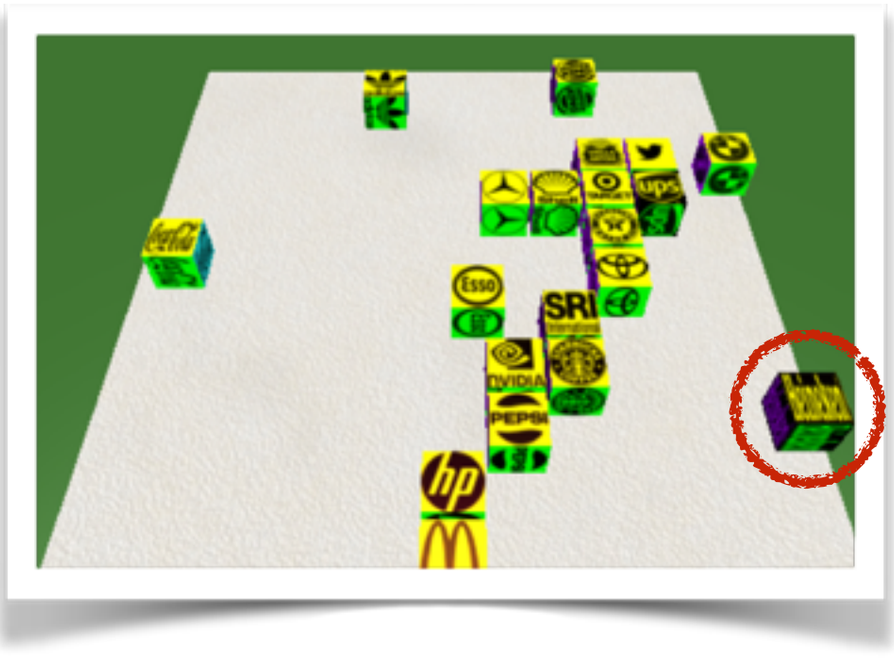
Motivation

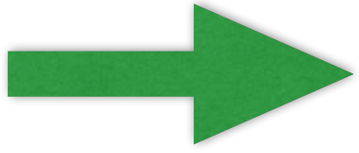
Machine-generated Representation

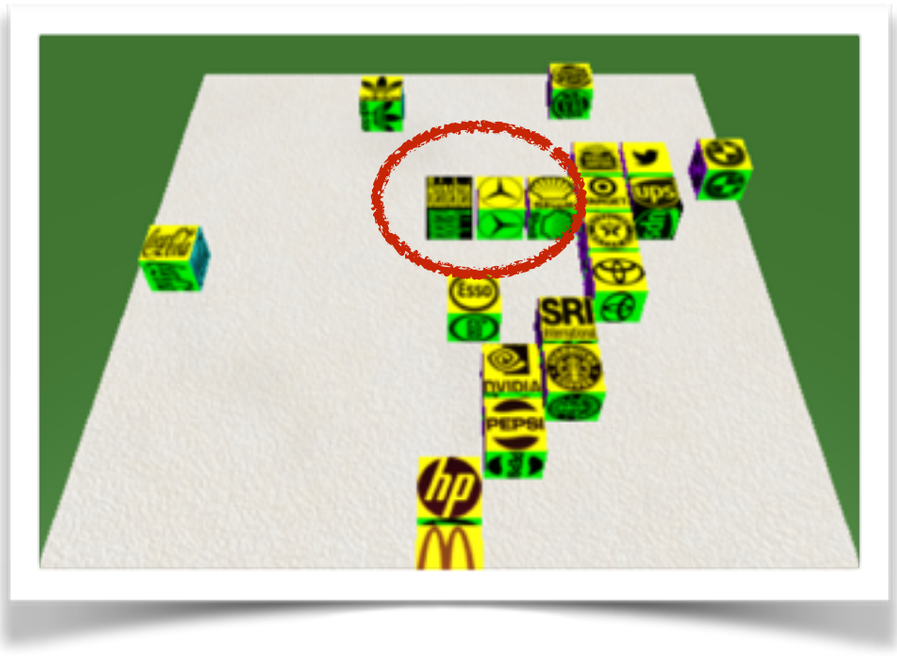


Motivation

Machine-generated Representation

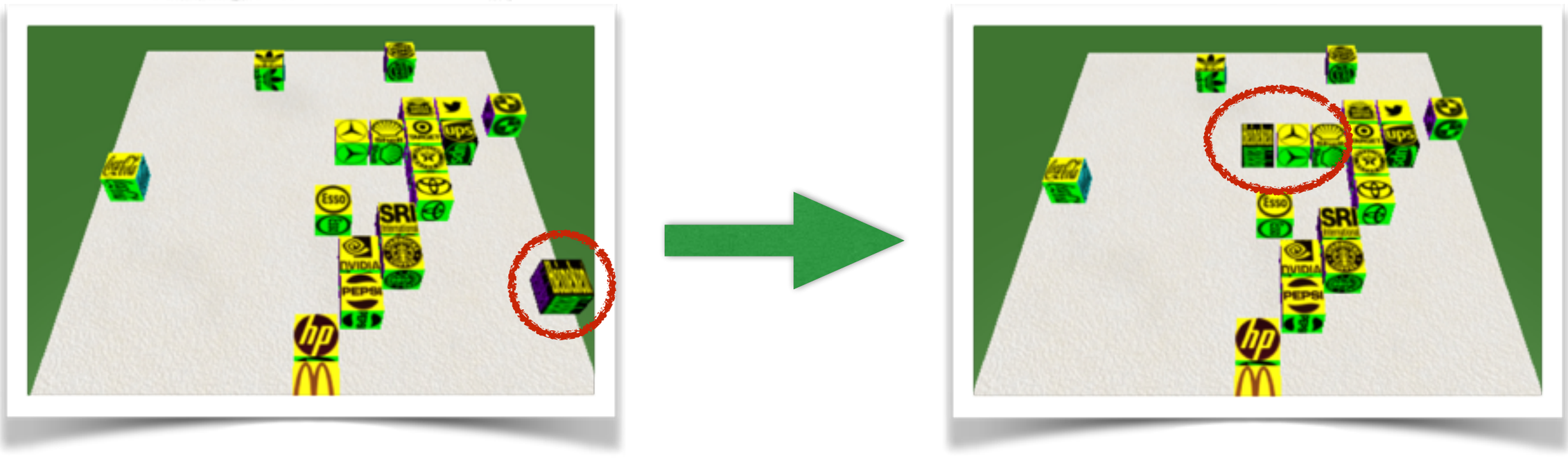






Motivation

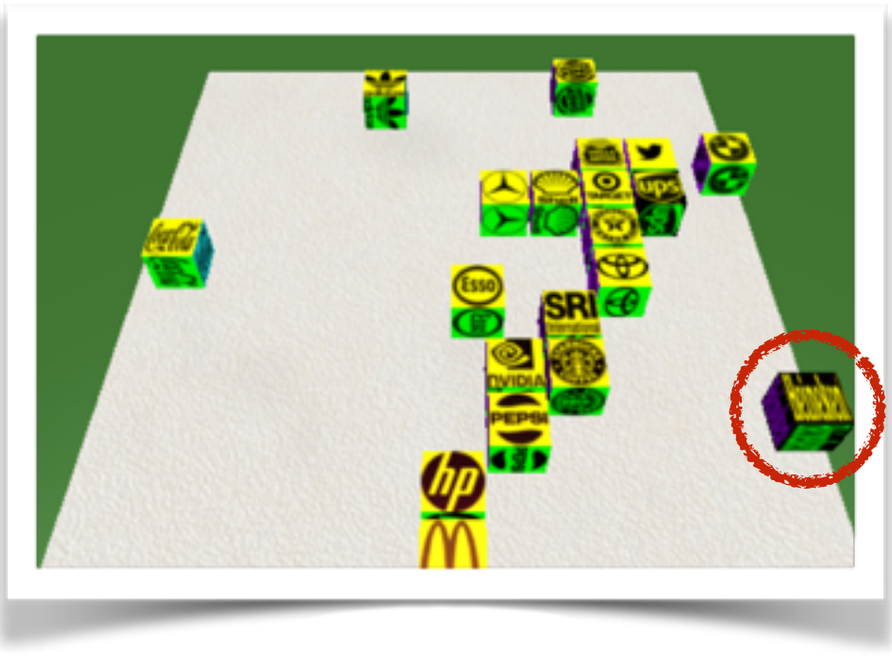
Machine-generated Representation

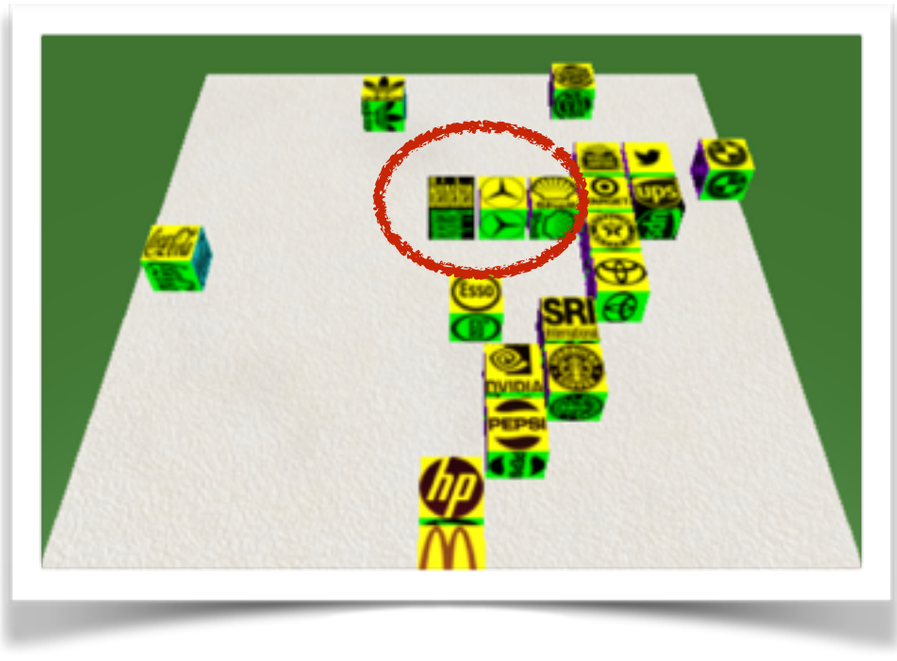


source	block:	hk		
target	block:	ms		
pos	RP:	W	scale:	small

Motivation

Machine-generated Representation



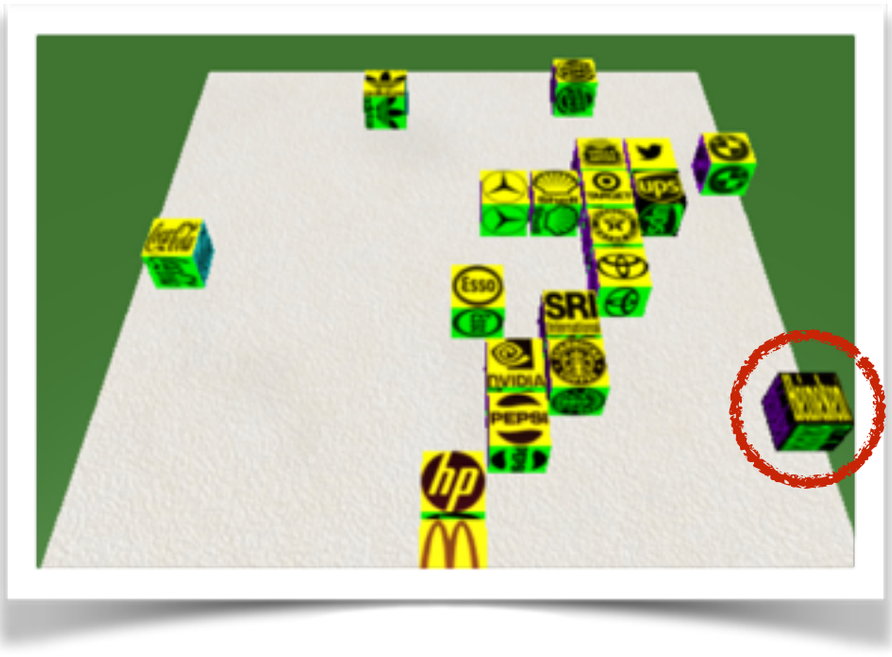


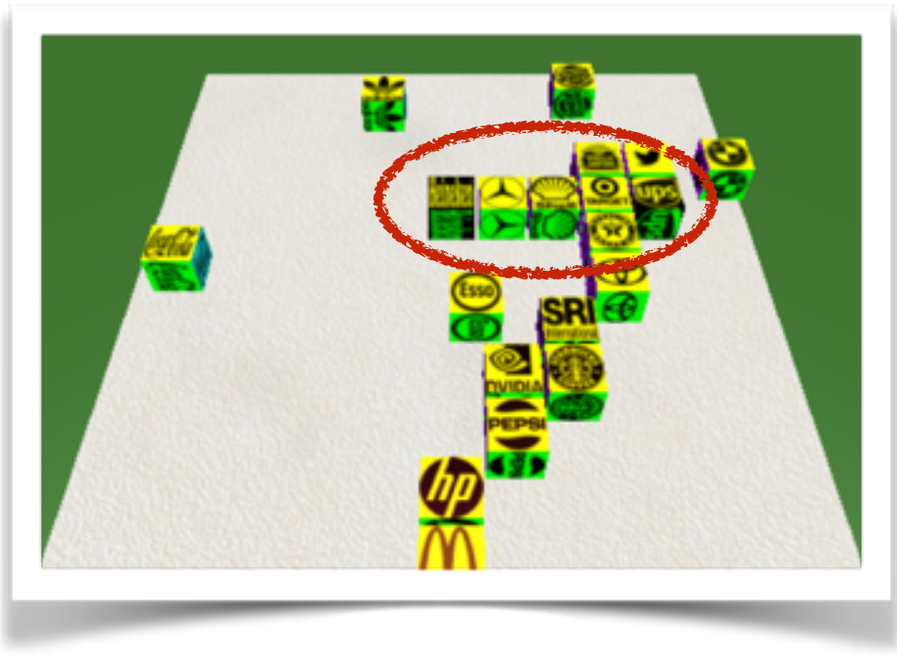
source	block:	hk		
target	block:	ms		
pos	RP:	W	scale:	small

Place the heineken block west of the mercedes block.

Motivation

Machine-generated Representation

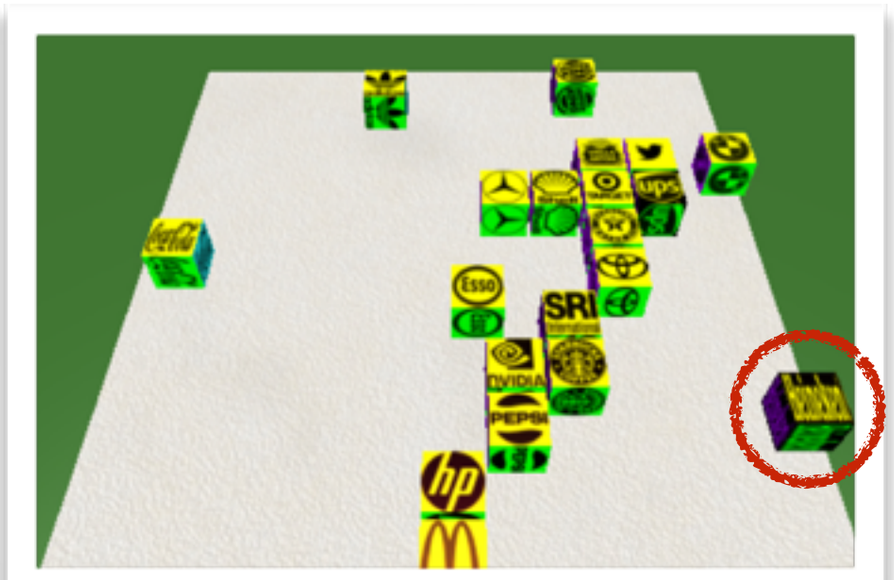


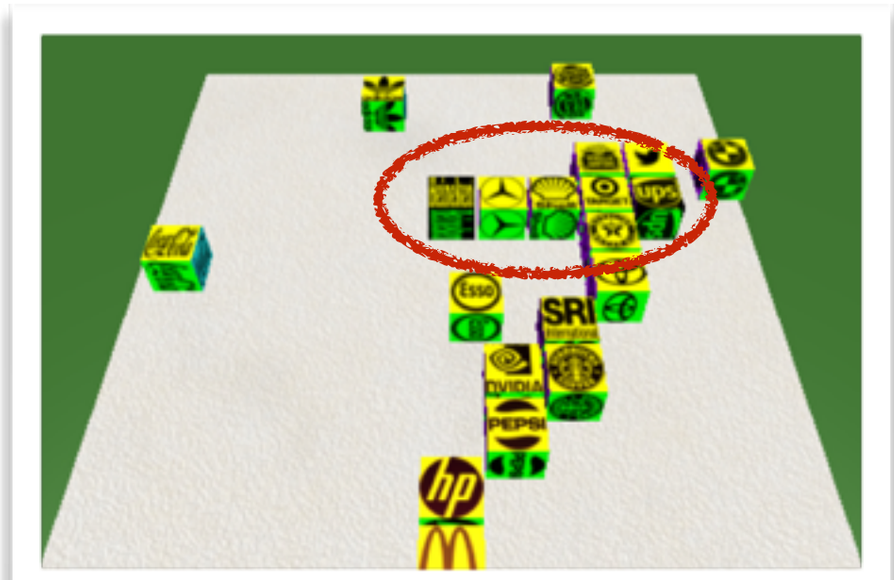


source	block:	hk		
target	block:	**ups**		
pos	RP:	W	scale:	**big**

Motivation

Machine-generated Representation





source	block:	hk		
target	block:	**ups**		
pos	RP:	W	scale:	**big**

Place the heineken block in the first open space to the left of ups block.

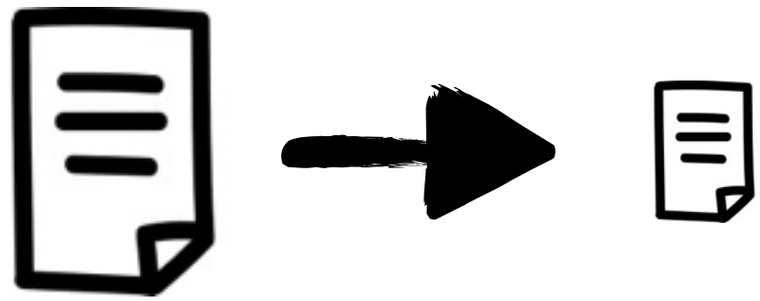
Motivation

Motivation

Machine-generated Representation

Motivation

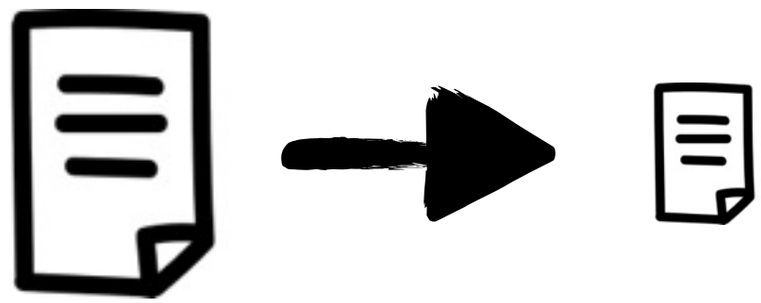
Machine-generated Representation



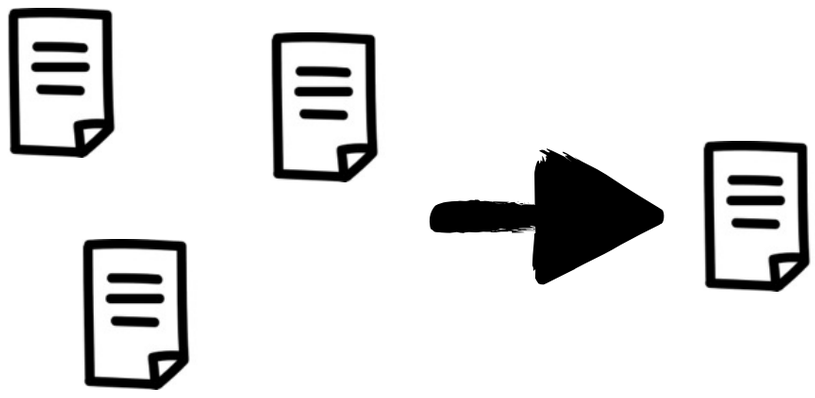
Sentence Compression

Motivation

Machine-generated Representation



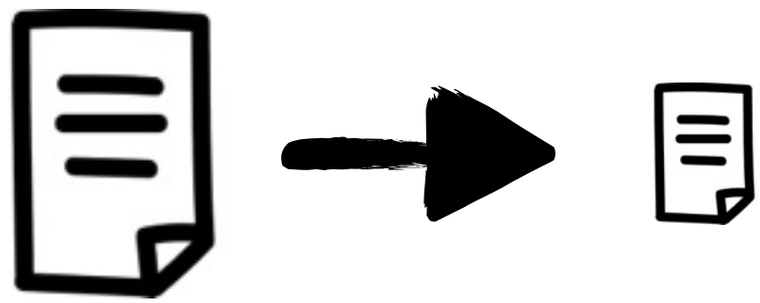
Sentence Compression



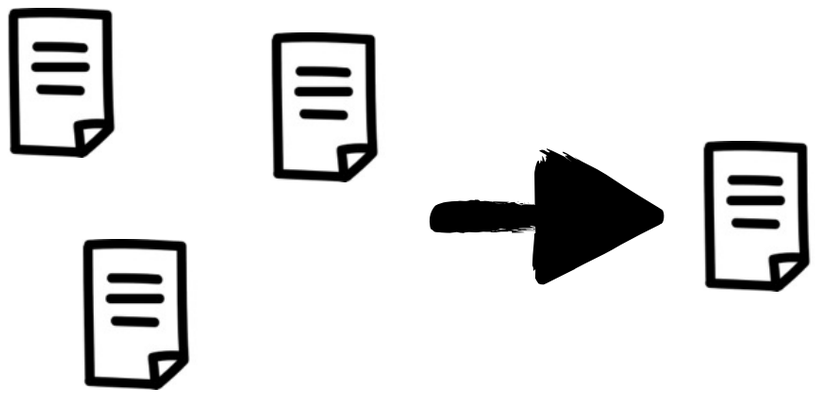
Sentence Fusion

Motivation

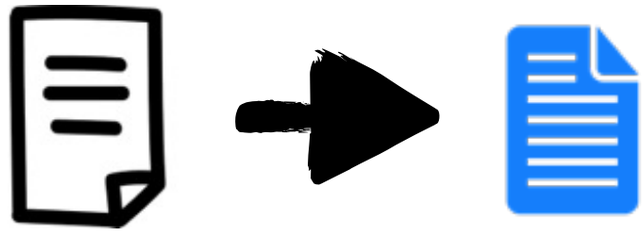
Machine-generated Representation



Sentence Compression

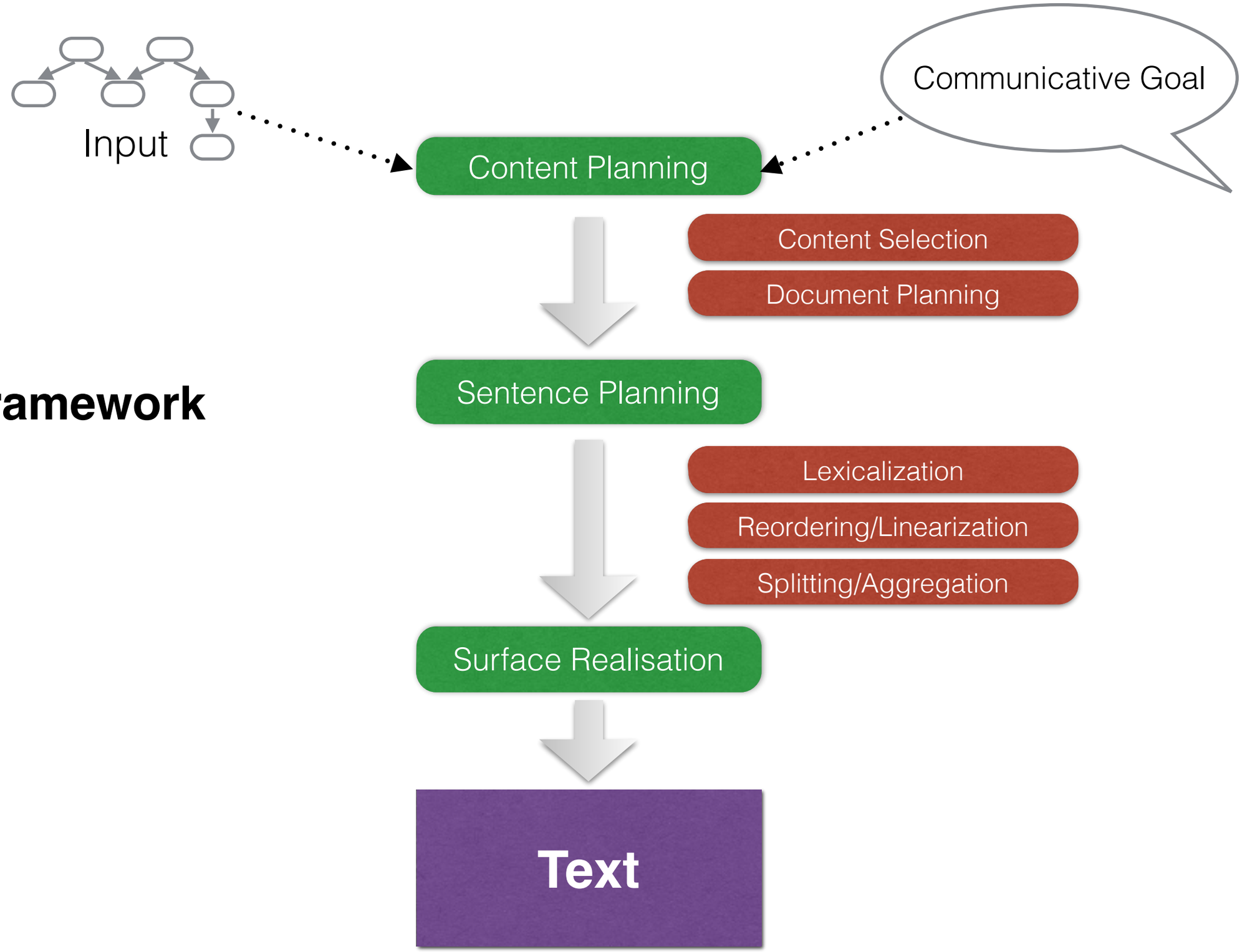


Sentence Fusion



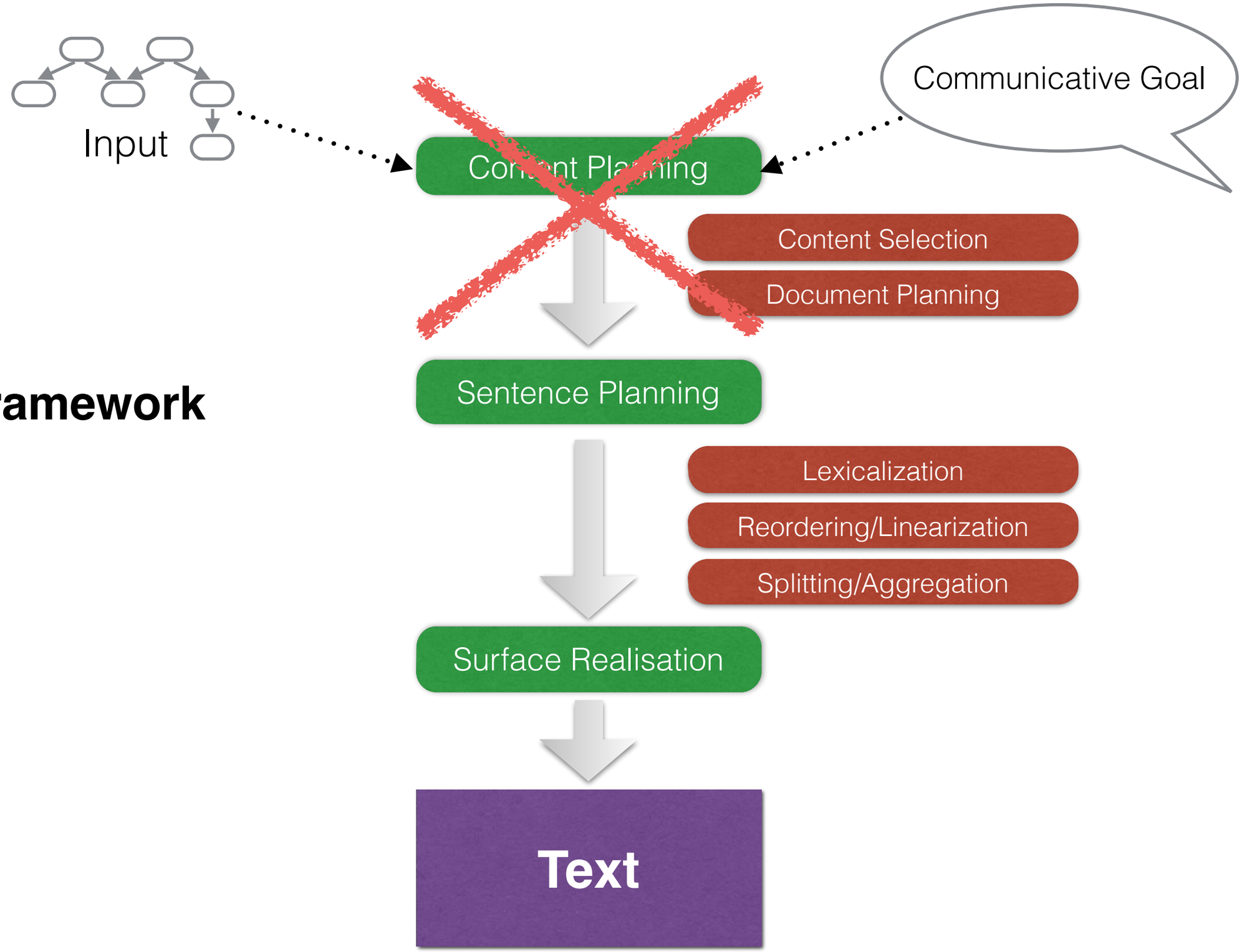
Paraphrasing

NLG Pipeline



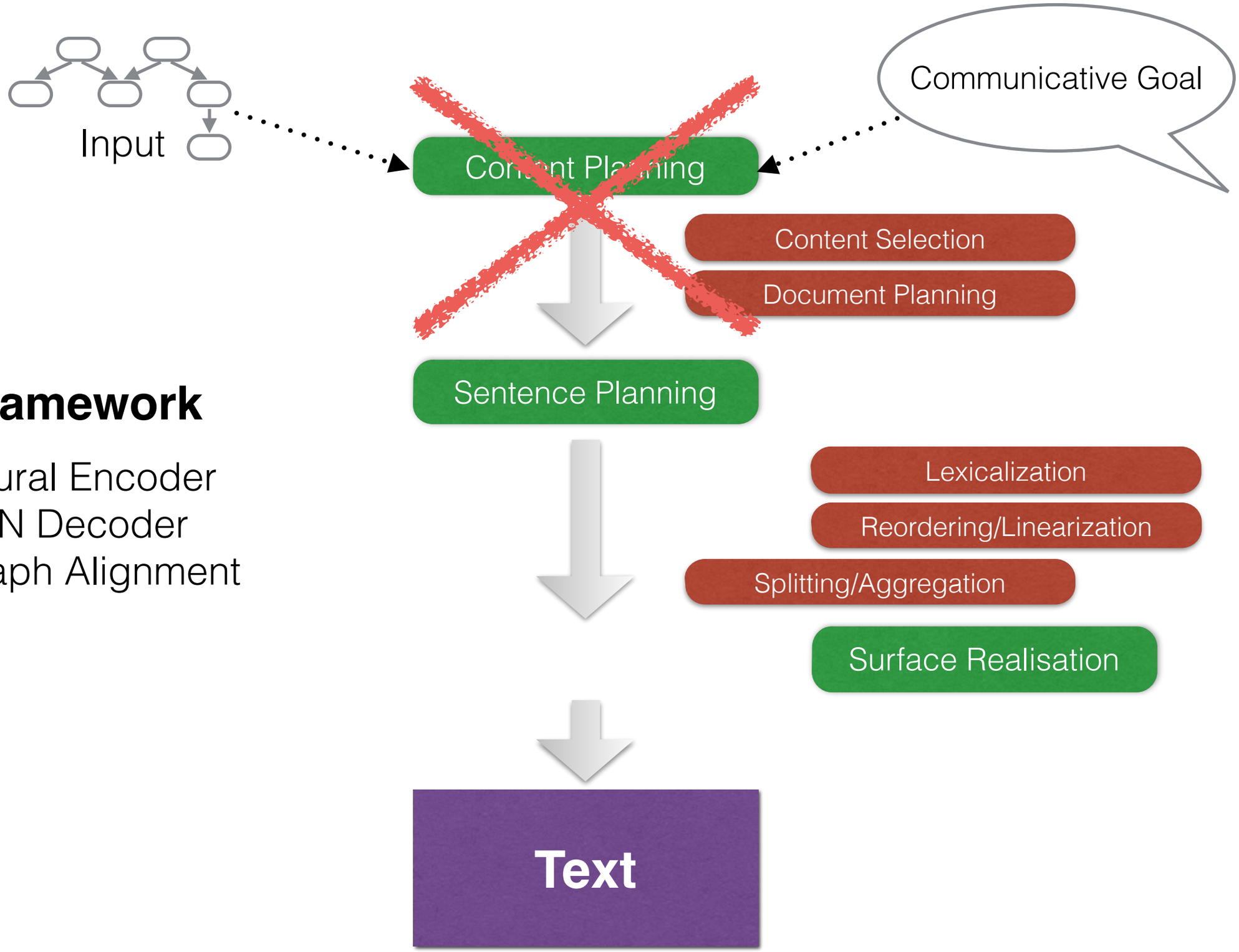
Framework

NLG Pipeline



Framework

NLG Pipeline



Framework

- Neural Encoder
- RNN Decoder
- Graph Alignment

Joint Generator

Training

Temperature

Time	Min	Mean	Max
06:00-21:00	9	15	21

Cloud Sky Cover

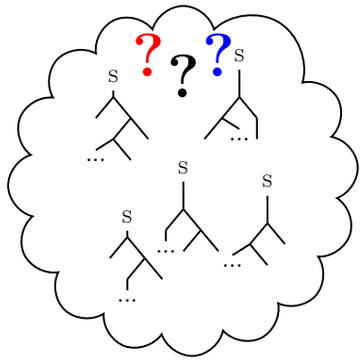
Time	Percent (%)
06:00-09:00	25-50
09:00-12:00	50-75

Wind Speed

Time	Min	Mean	Max
06:00-21:00	15	20	30

Wind Direction

Time	Mode
06:00-21:00	S

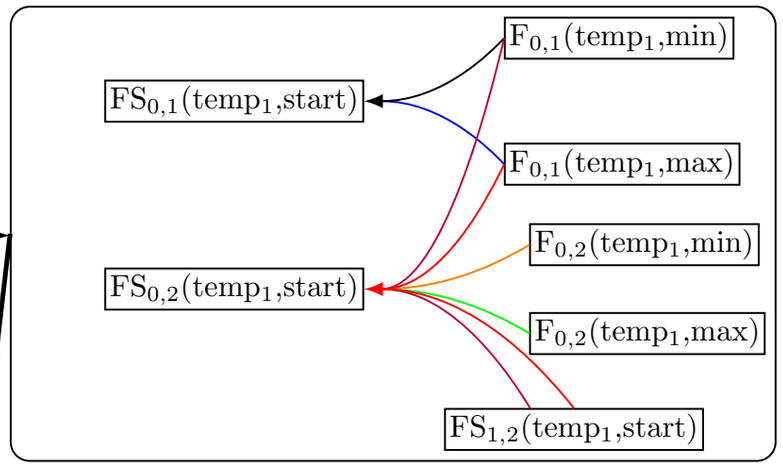


Cloudy, with temperatures between 10 and 20 degrees. South wind around 20 mph.

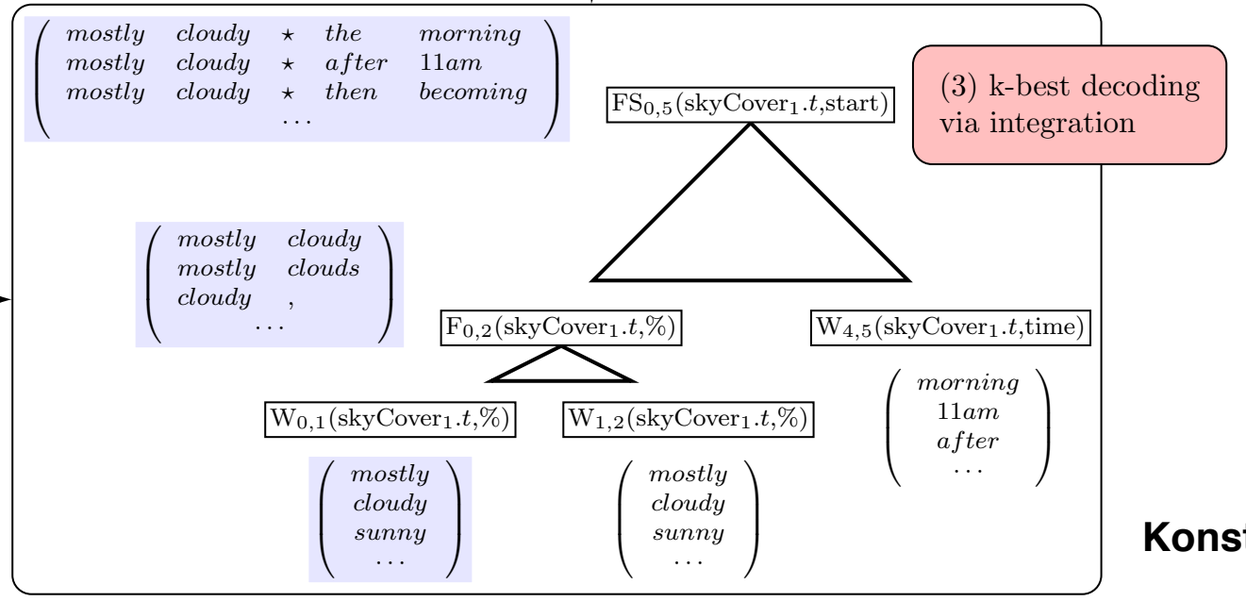
(1) PCFG Grammar

$S \rightarrow R(start)$
$R(r_i.t) \rightarrow FS(r_j, start)R(r_j.t)$
$R(r_i.t) \rightarrow FS(r_j, start)$
$FS(r, r.f_i) \rightarrow F(r, r.f_j)FS(r, r.f_j)$
$FS(r, r.f_i) \rightarrow F(r, r.f_j)$
$F(r, r.f) \rightarrow W(r, r.f)F(r, r.f)$
$F(r, r.f) \rightarrow W(r, r.f)$
$W(r, r.f) \rightarrow \alpha$
$W(r, r.f) \rightarrow g(f.v)$

(2) Hypergraph Representation



Testing



(3) k-best decoding via integration

Konstas and Lapata, 2012, 2013

Joint Generator

Training

Temperature

Time	Min	Mean	Max
06:00-21:00	9	15	21

Cloud Sky Cover

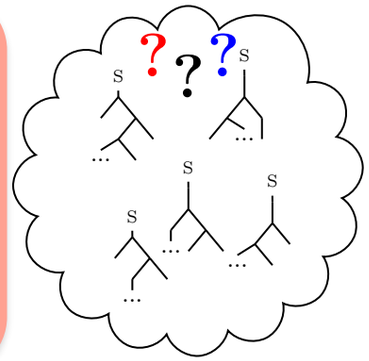
Time	Percent (%)
06:00-09:00	25-50
09:00-12:00	50-75

Wind Speed

Time	Min	Mean	Max
06:00-21:00	15	20	30

Wind Direction

Time	Mode
06:00-21:00	S

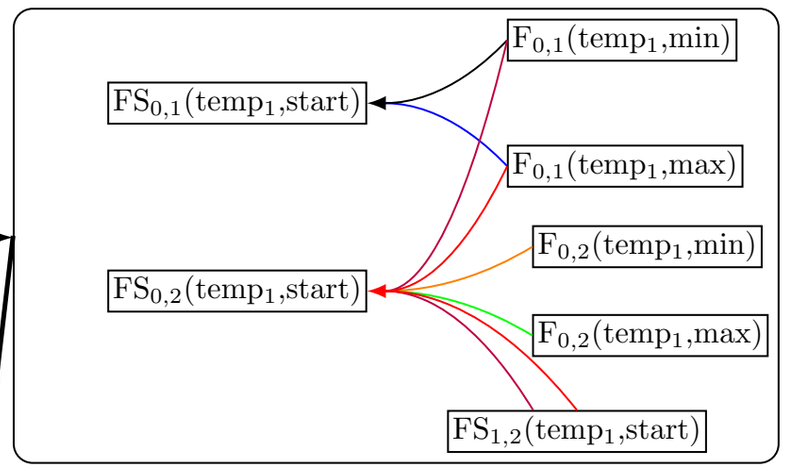


Cloudy, with temperatures between 10 and 20 degrees. South wind around 20 mph.

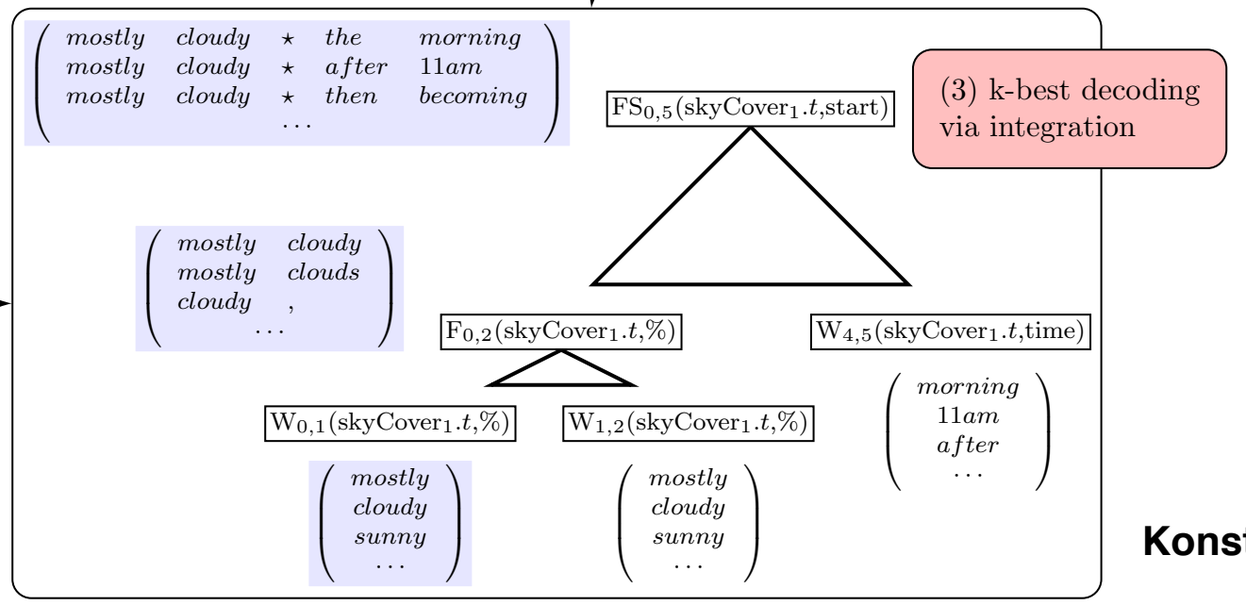
(1) PCFG Grammar

$S \rightarrow R(start)$
$R(r_i.t) \rightarrow FS(r_j, start)R(r_j.t)$
$R(r_i.t) \rightarrow FS(r_j, start)$
$FS(r, r.f_i) \rightarrow F(r, r.f_j)FS(r, r.f_j)$
$FS(r, r.f_i) \rightarrow F(r, r.f_j)$
$F(r, r.f) \rightarrow W(r, r.f)F(r, r.f)$
$F(r, r.f) \rightarrow W(r, r.f)$
$W(r, r.f) \rightarrow \alpha$
$W(r, r.f) \rightarrow g(f.v)$

(2) Hypergraph Representation



Testing



(3) k-best decoding via integration

Konstas and Lapata, 2012, 2013

Joint Generator

Training

Temperature

Time	Min	Mean	Max
06:00-21:00	9	15	21

Cloud Sky Cover

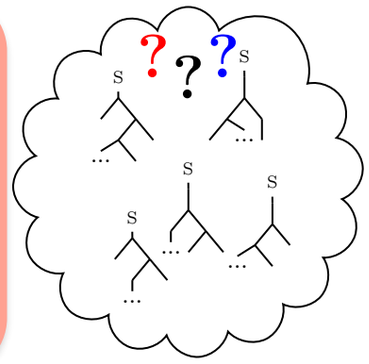
Time	Percent (%)
06:00-09:00	25-50
09:00-12:00	50-75

Wind Speed

Time	Min	Mean	Max
06:00-21:00	15	20	30

Wind Direction

Time	Mode
06:00-21:00	S

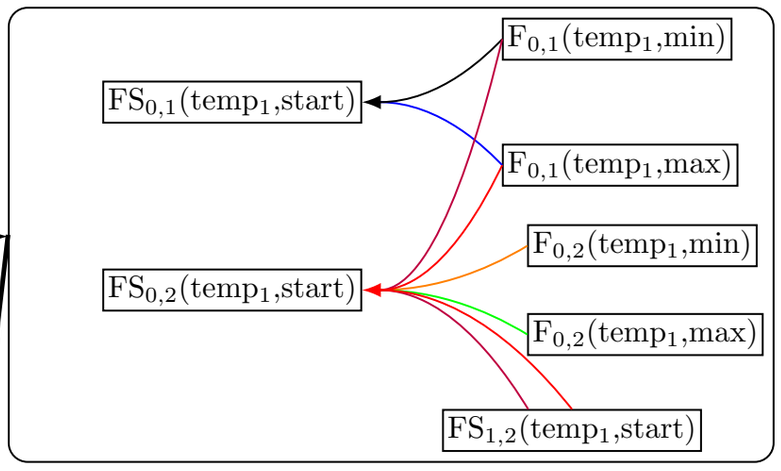


Cloudy, with temperatures between 10 and 20 degrees. South wind around 20 mph.

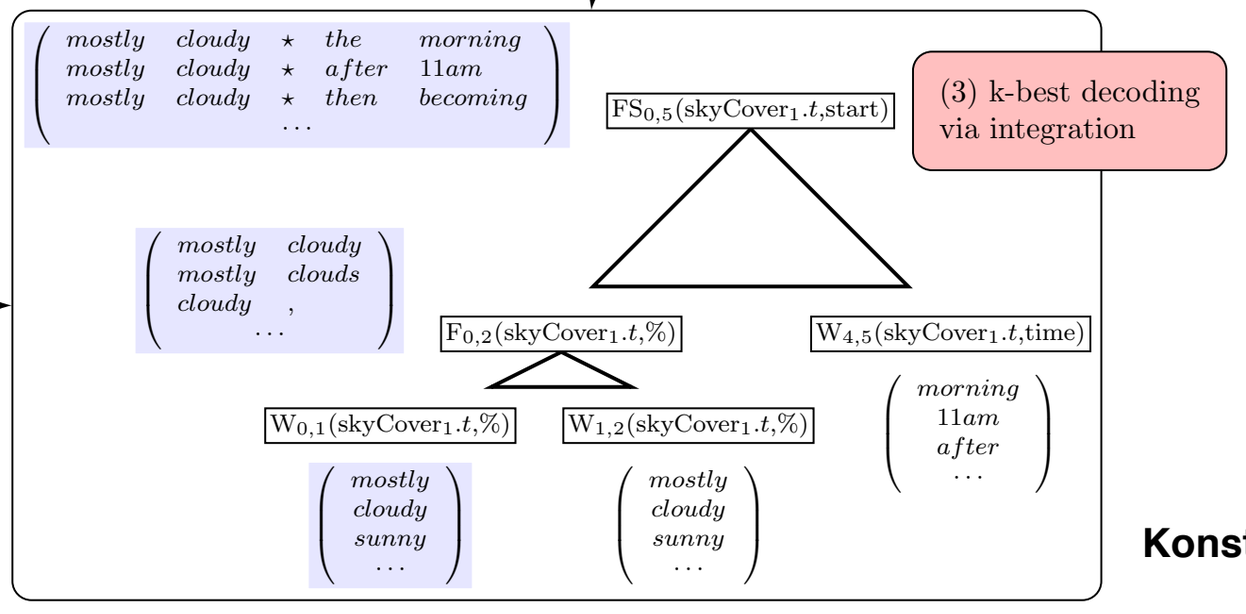
(1) PCFG Grammar

$S \rightarrow R(start)$
$R(r_i.t) \rightarrow FS(r_j, start)R(r_j.t)$
$R(r_i.t) \rightarrow FS(r_j, start)$
$FS(r, r.f_i) \rightarrow F(r, r.f_j)FS(r, r.f_j)$
$FS(r, r.f_i) \rightarrow F(r, r.f_j)$
$F(r, r.f) \rightarrow W(r, r.f)F(r, r.f)$
$F(r, r.f) \rightarrow W(r, r.f)$
$W(r, r.f) \rightarrow \alpha$
$W(r, r.f) \rightarrow g(f.v)$

(2) Hypergraph Representation



Testing



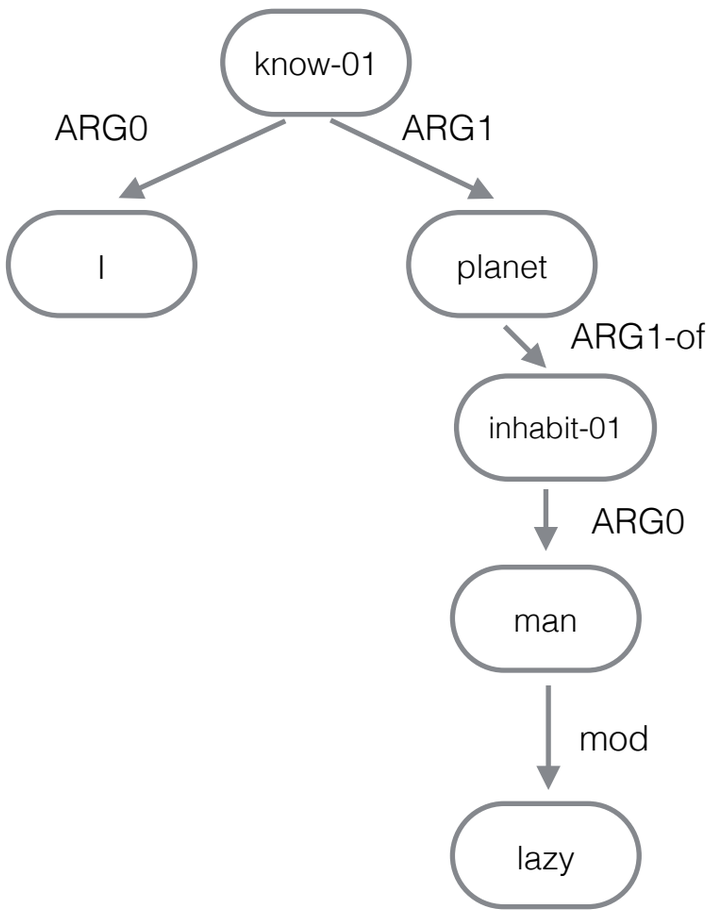
(3) k-best decoding via integration

Konstas and Lapata, 2012, 2013

More Complex Structure

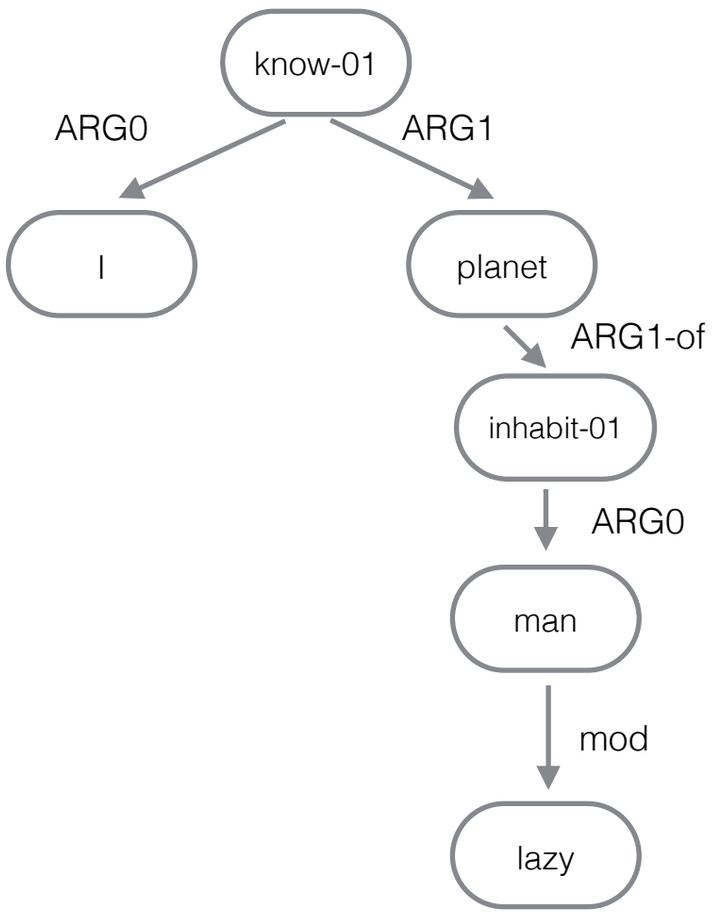
More Complex Structure

Input (Graph or Tree)



More Complex Structure

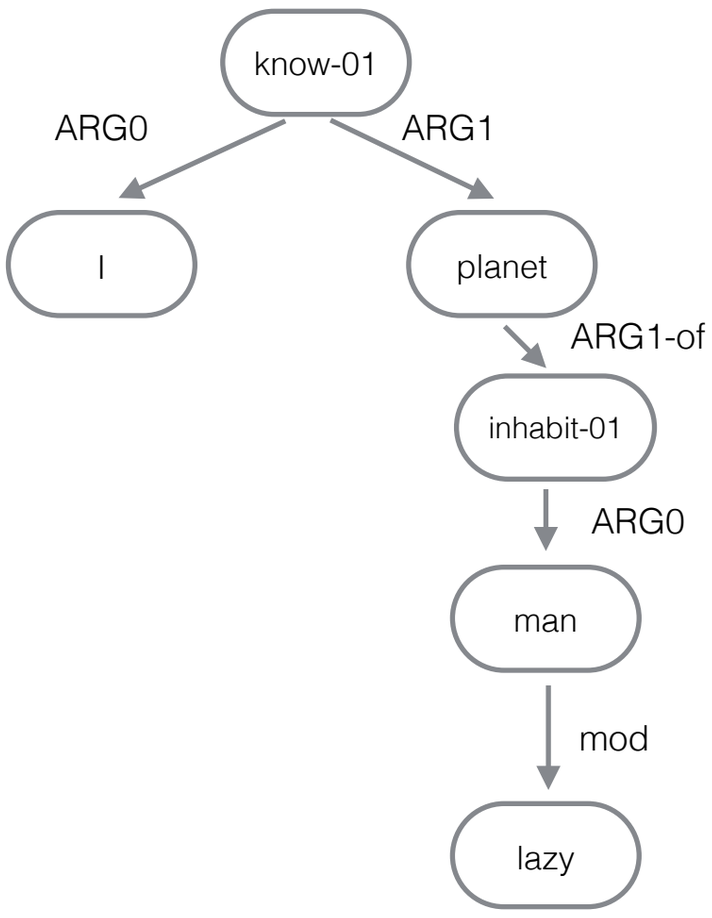
Input (Graph or Tree)



AMR	DAG

More Complex Structure

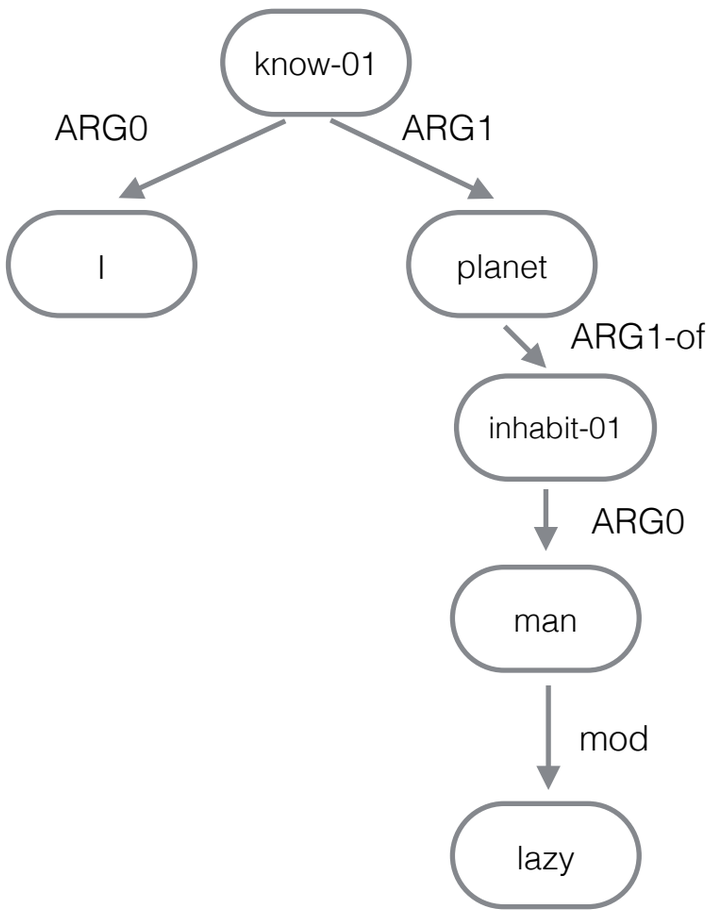
Input (Graph or Tree)



AMR	DAG
λ-calculus Expression	Tree

More Complex Structure

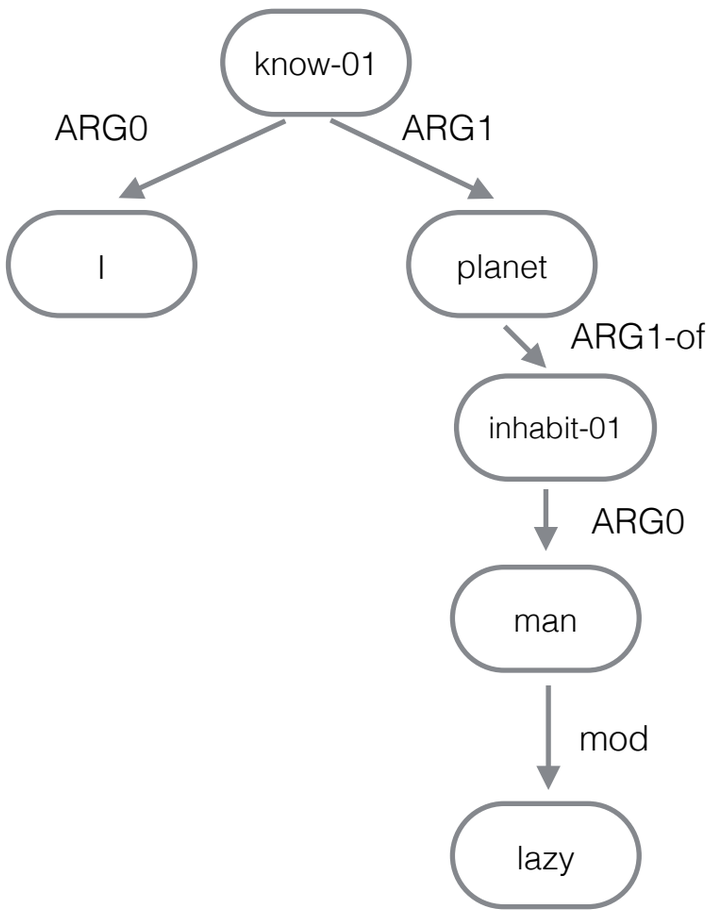
Input (Graph or Tree)



AMR	DAG
λ-calculus Expression	Tree
ECIs	Tree

More Complex Structure

Input (Graph or Tree)



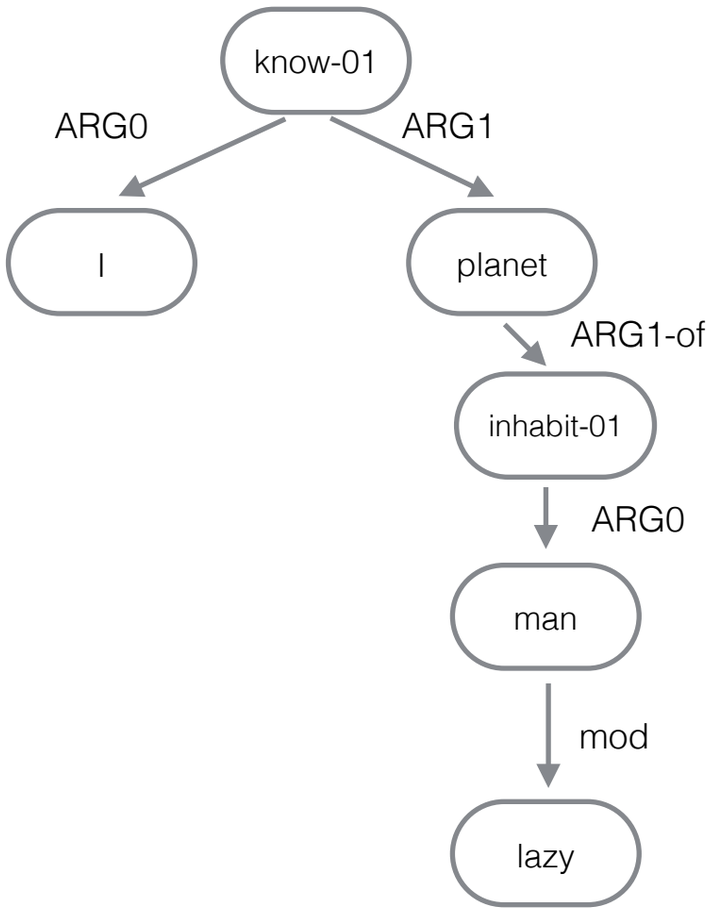
Output (Text)

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

AMR	DAG
λ-calculus Expression	Tree
ECIs	Tree

More Complex Structure

Input (Graph or Tree)



AMR	DAG
λ-calculus Expression	Tree
ECIs	Tree

Output (Text)

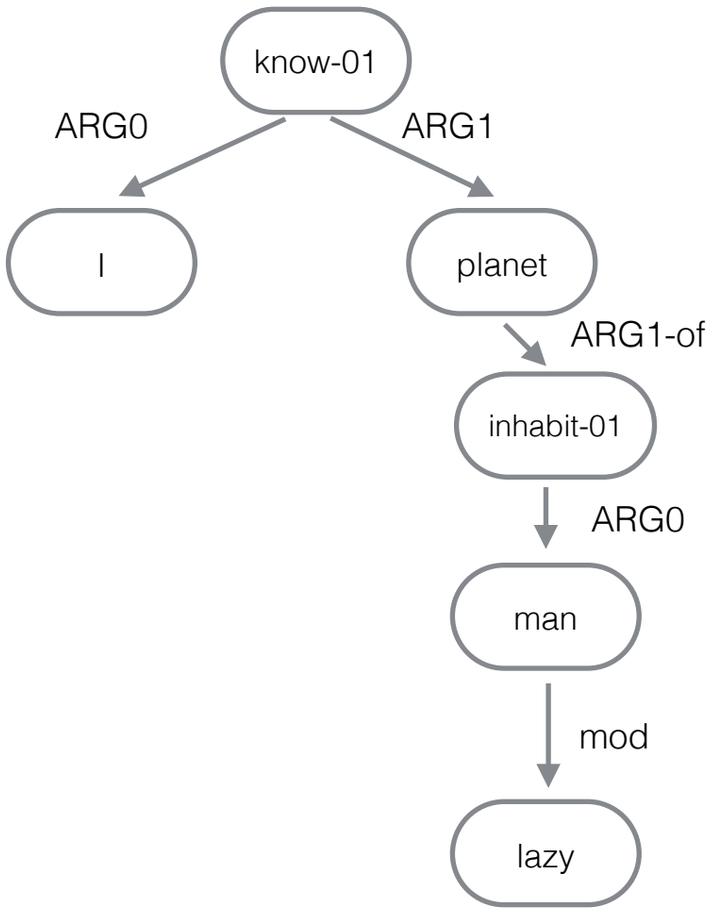
I know a **planet** that is **inhabited** by a **lazy man**.

I knew a **planet** that was **inhabited** by a **lazy man**. **(lpp_1943.249)**

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

More Complex Structure

Input (Graph or Tree)



AMR	DAG
λ-calculus Expression	Tree
ECIs	Tree

Output (Text)

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

I knew a **planet** that was **inhabited** by a **lazy man**. **(lpp_1943.249)**

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

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

NEUGEN Framework

NeuGen Framework

NeuGen Framework

Encoder

- Bag of Words: Concepts and roles
- Attention mechanism

NeuGen Framework

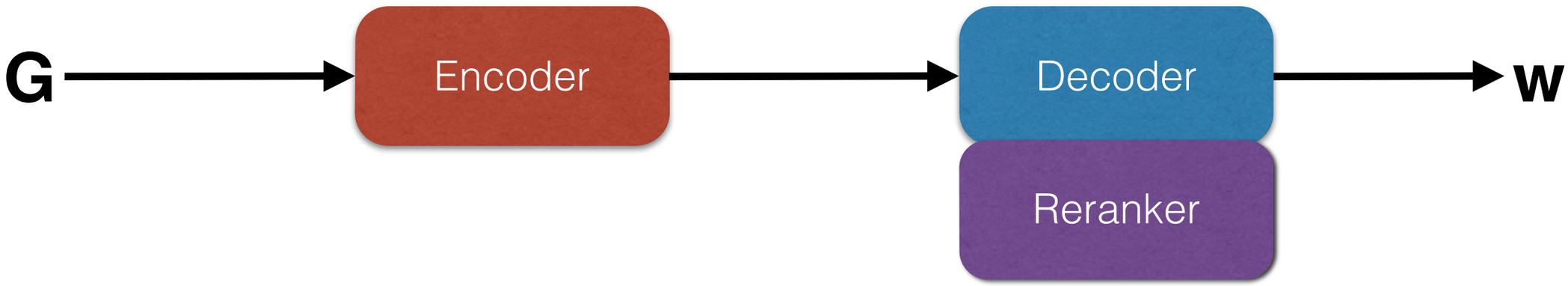
Encoder

- Bag of Words: Concepts and roles
- Attention mechanism

Decoder

- Left-to-right word LSTM
- Greedy beam search

NeuGen Framework



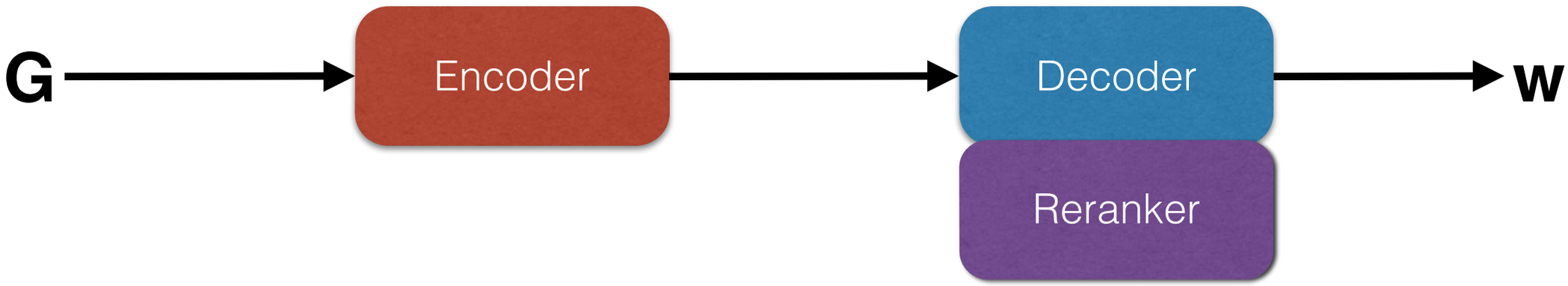
Encoder

- Bag of Words: Concepts and roles
- Attention mechanism

Decoder

- Left-to-right word LSTM
- Greedy beam search

NeuGen Framework



Encoder

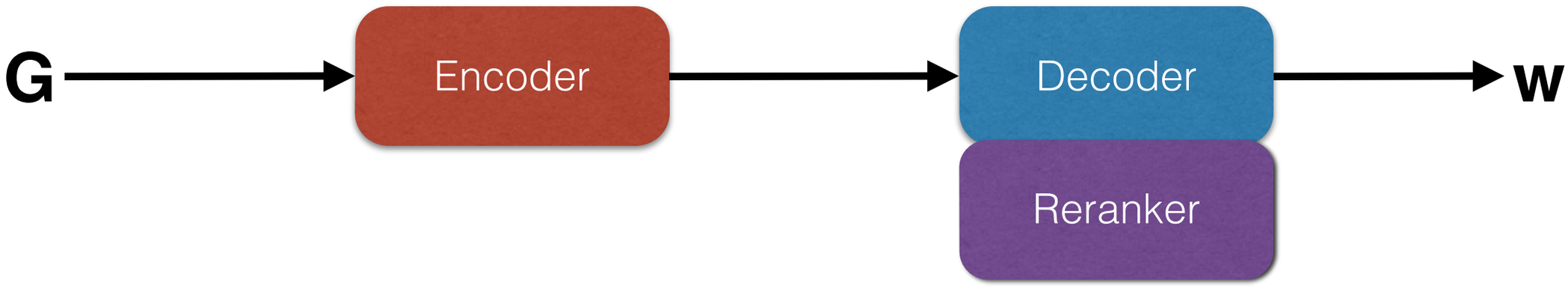
- Bag of Words: Concepts and roles
- Attention mechanism

Decoder

- Left-to-right word LSTM
- Greedy beam search

Reranker

NeuGen Framework



Encoder

- Bag of Words: Concepts and roles
- Attention mechanism

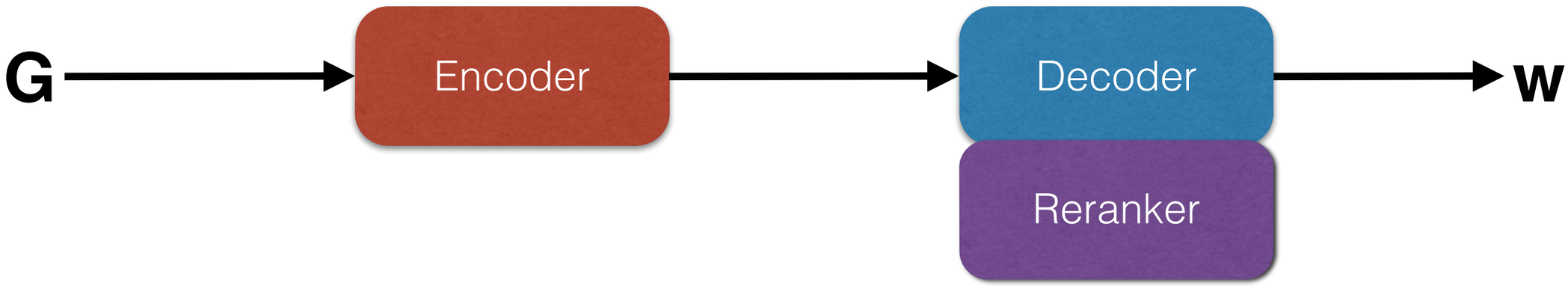
Decoder

- Left-to-right word LSTM
- Greedy beam search

Reranker

- CCG parser -> semantics

NeuGen Framework



Encoder

- Bag of Words: Concepts and roles
- Attention mechanism

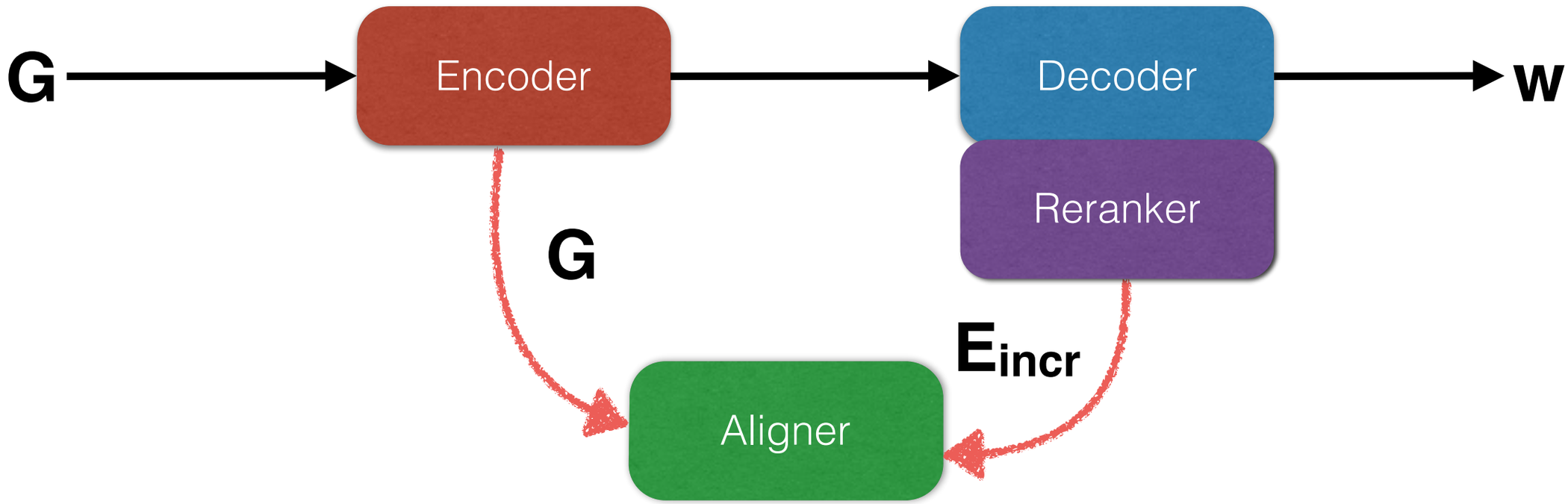
Decoder

- Left-to-right word LSTM
- Greedy beam search

Reranker

- CCG parser -> semantics
- Partial propositions graph $\mathbf{E_{incr}}$

NEUGEN Framework



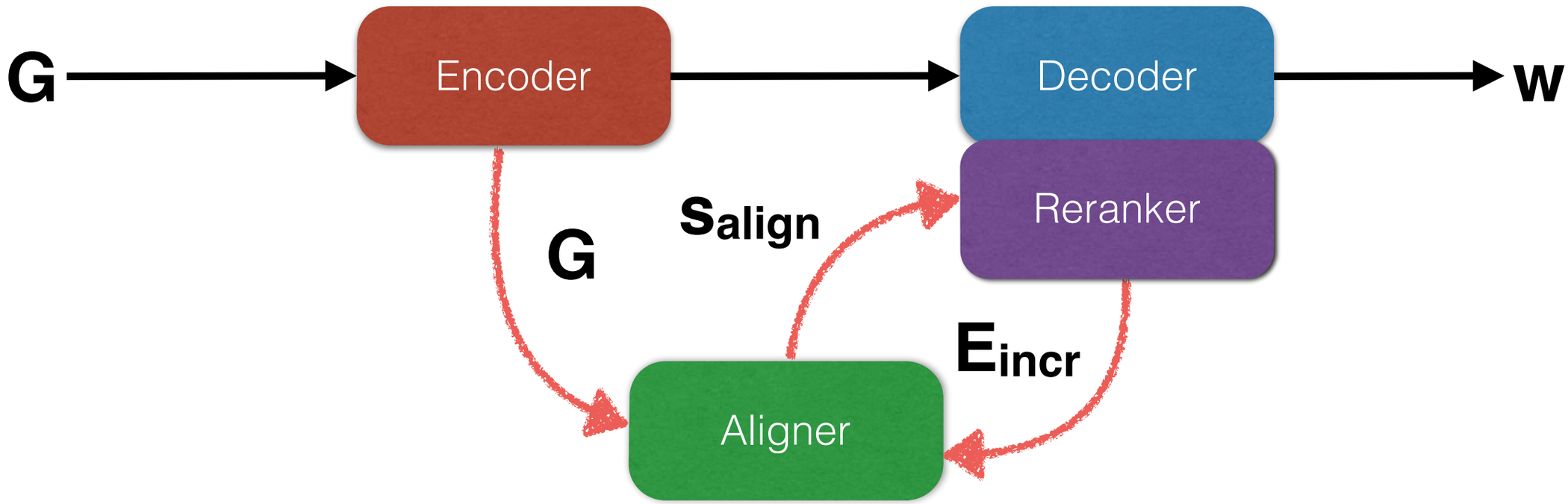
Encoder
- Bag of Words: Concepts and roles
- Attention mechanism

Aligner
- Feed-forward NN with attention

Decoder
- Left-to-right word LSTM
- Greedy beam search

Reranker
- CCG parser -> semantics
- Partial propositions graph $\mathbf{E_{incr}}$

NeuGen Framework



Encoder
- Bag of Words: Concepts and roles
- Attention mechanism

Aligner
- Feed-forward NN with attention
- $\mathbf{S}_{\mathbf{align}} = \mathit{fit}(\mathbf{G}, \mathbf{E}_{\mathbf{incr}})$

Decoder
- Left-to-right word LSTM
- Greedy beam search

Reranker
- CCG parser -> semantics
- Partial propositions graph $\mathbf{E}_{\mathbf{incr}}$

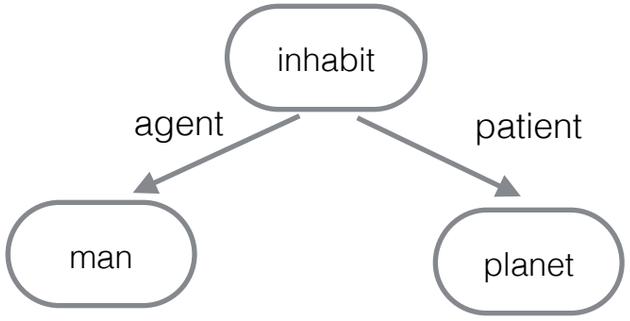
Alignment

Alignment

w: A **man inhabited** a **planet**

Alignment

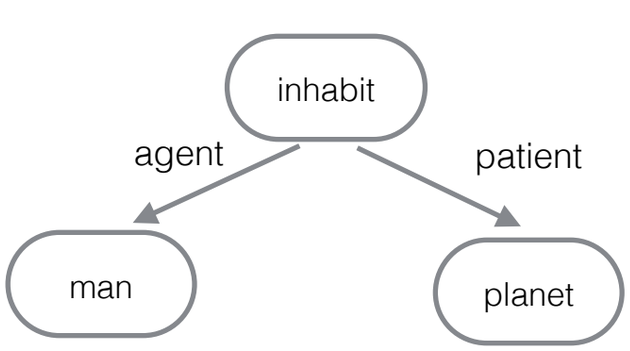
w: A **man inhabited** a **planet**



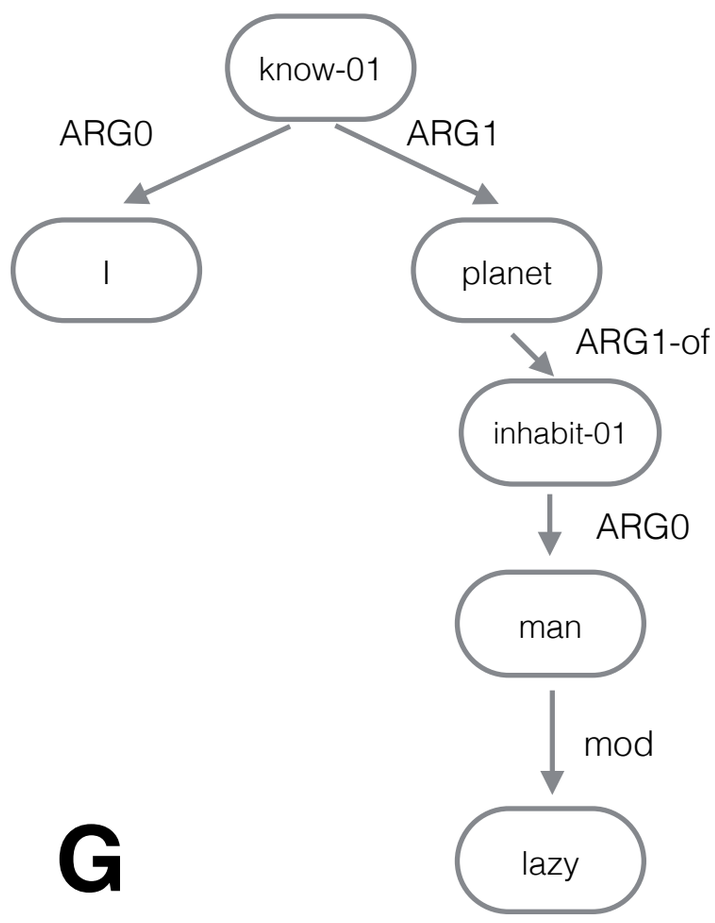
E_{incr}

Alignment

w: A **man inhabited** a **planet**



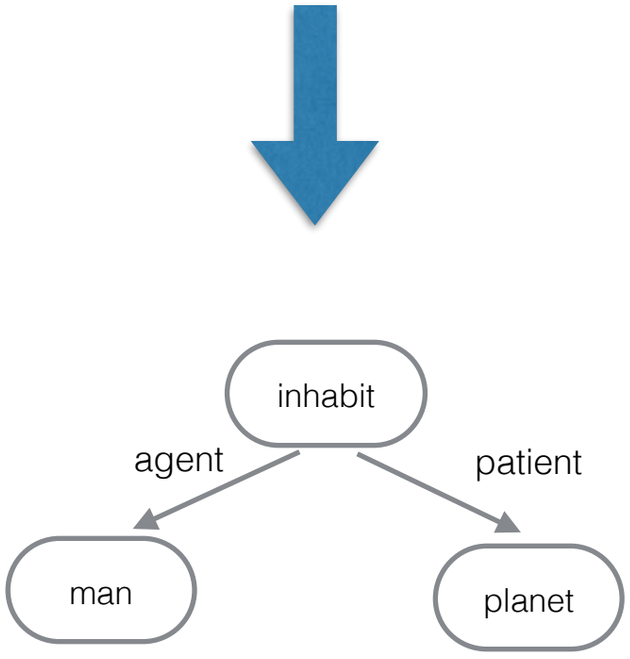
E$_{\textbf{incr}}$



G

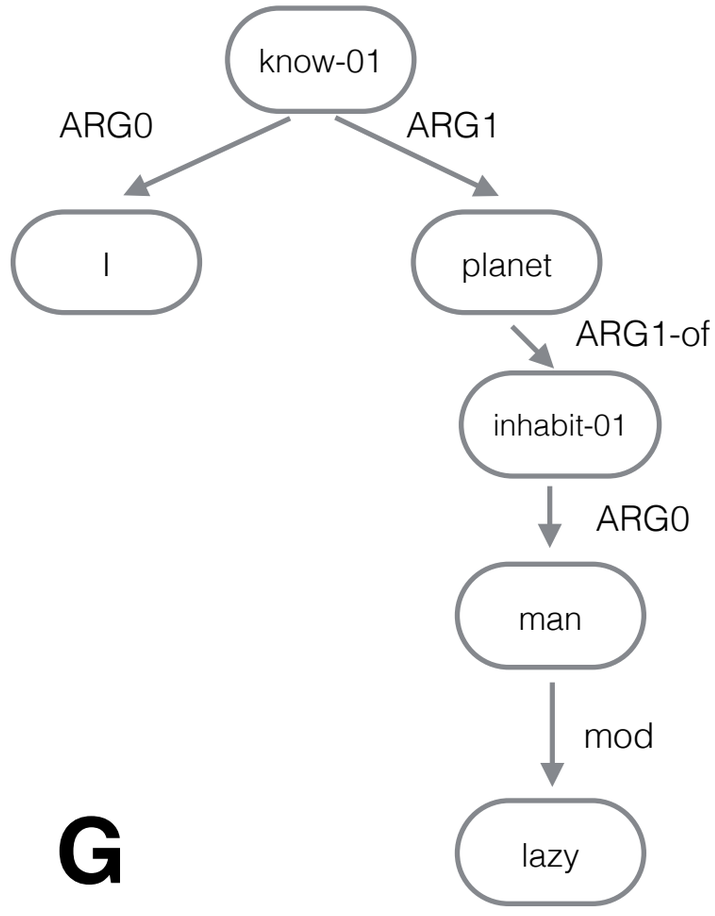
Alignment

w: A **man inhabited** a **planet**



E$_{\textbf{incr}}$

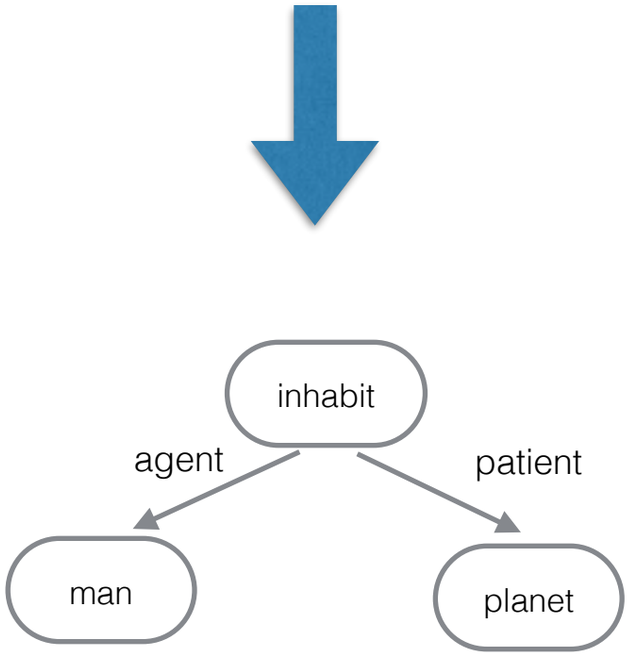
fit(**G**, **E**$_{\textbf{incr}}$)



G

Alignment

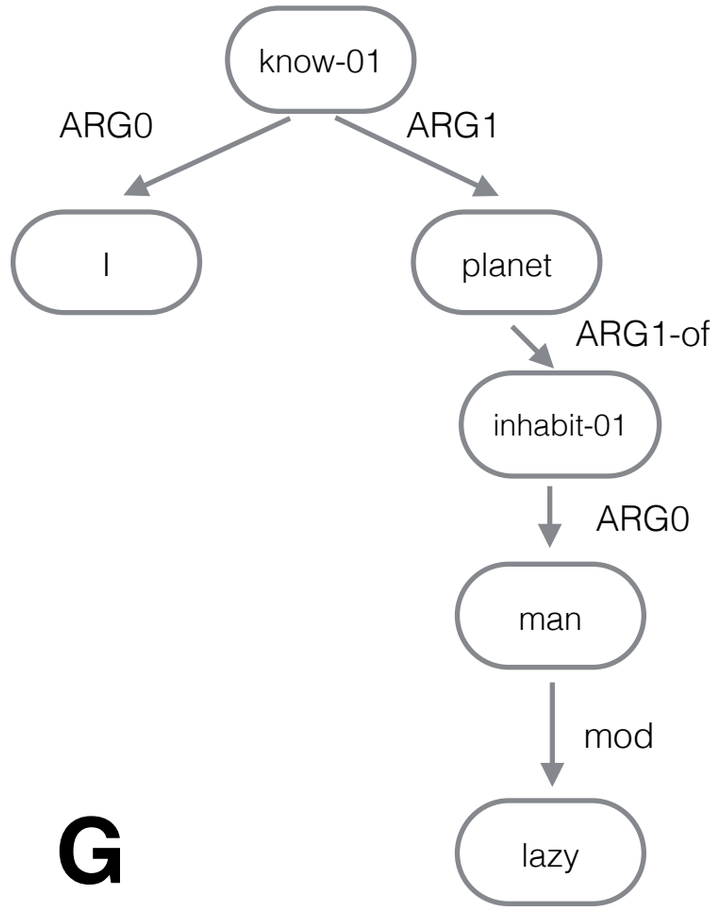
w: A **man inhabited** a **planet**



E$_{\textbf{incr}}$

```
<inhabit, man, agent>
<inhabit, planet, patient>
```

fit(**G**, **E**$_{\textbf{incr}}$)

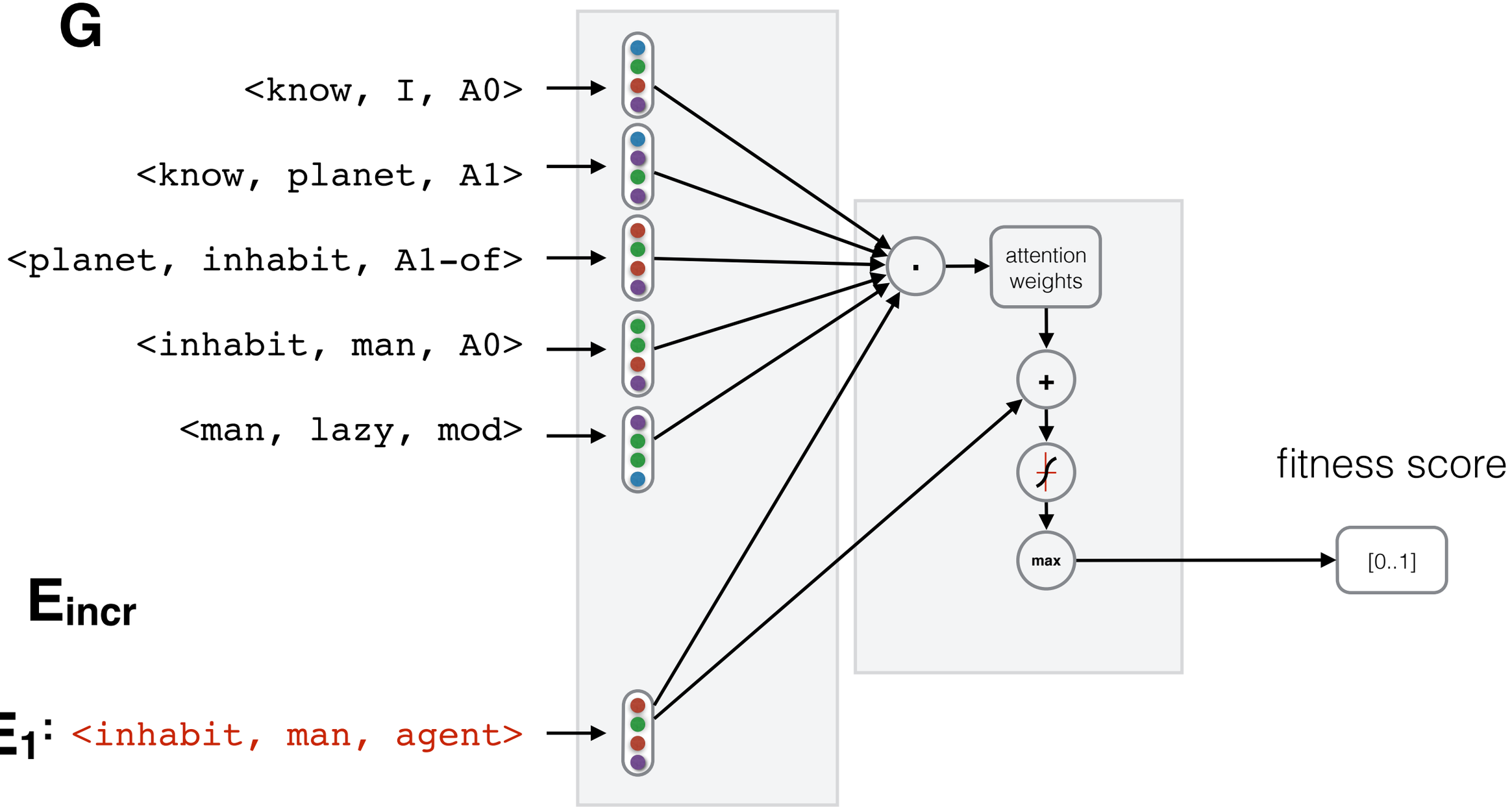


G

```
<know-01, I, A0>
<know-01, planet, A1>
<planet, inhabit-01, A1-of>
<inhabit-01, man, A0>
<man, lazy, mod>
```

Alignment

G



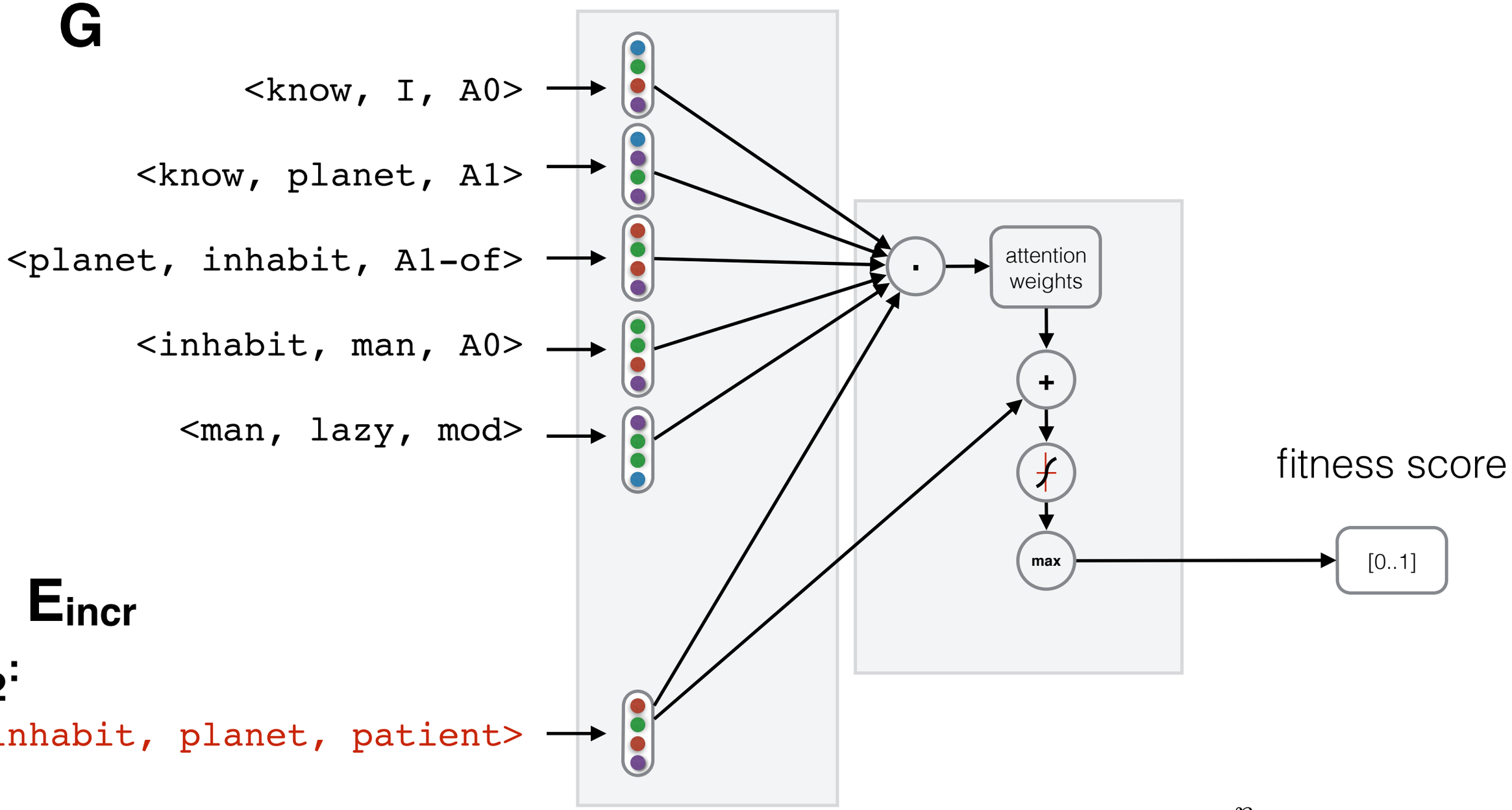
E_{incr}

E_1: <inhabit, man, agent>

<inhabit, planet, patient>

Alignment

G



E_{incr}

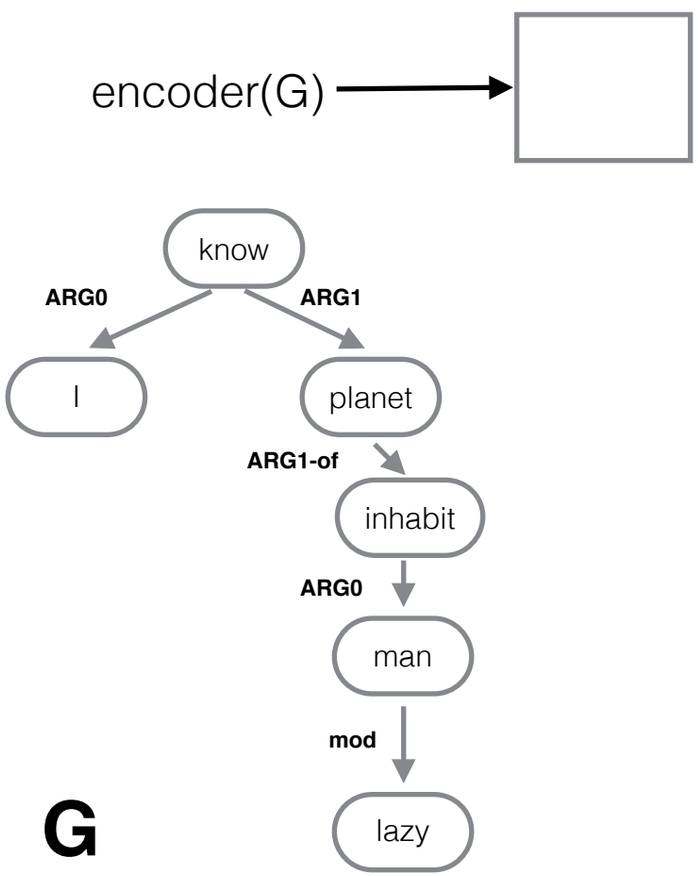
E_2:

<inhabit, planet, patient>

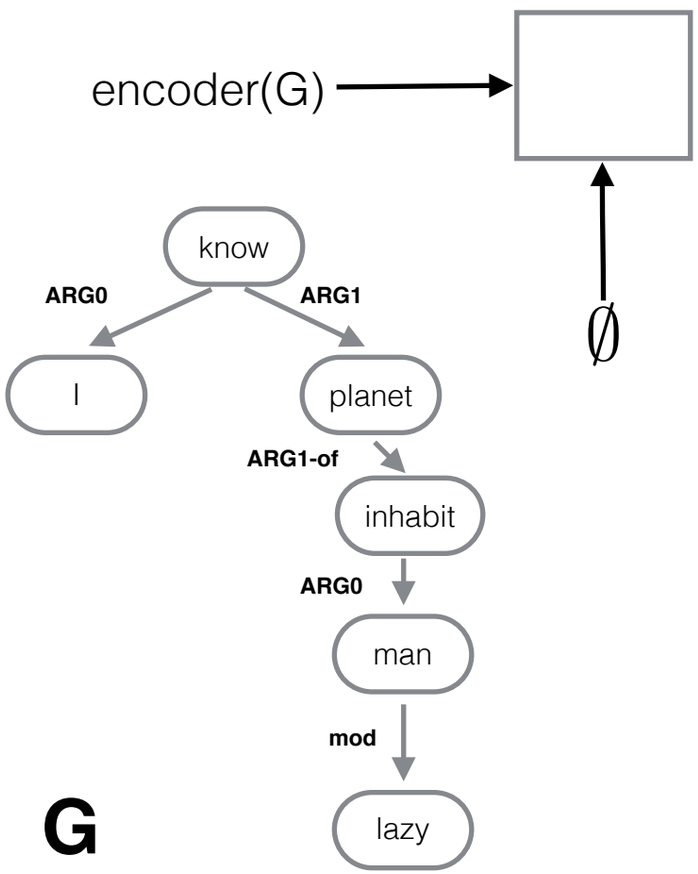
<inhabit, man, agent>

$$fit(G, E_{incr}) = \frac{1}{n} \sum_{i=0}^{n} fit(G, E_i)$$

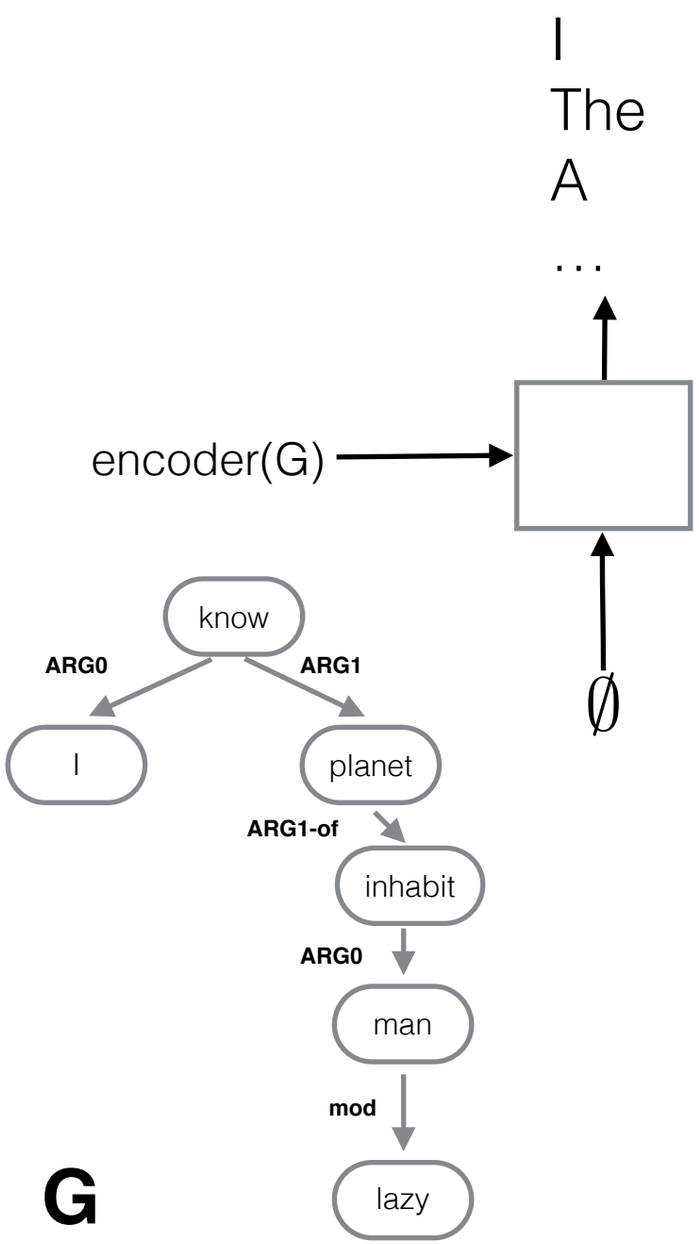
Decoding


encoder(G)
know
ARG0
ARG1
I
planet
ARG1-of
inhabit
ARG0
man
mod
lazy
G

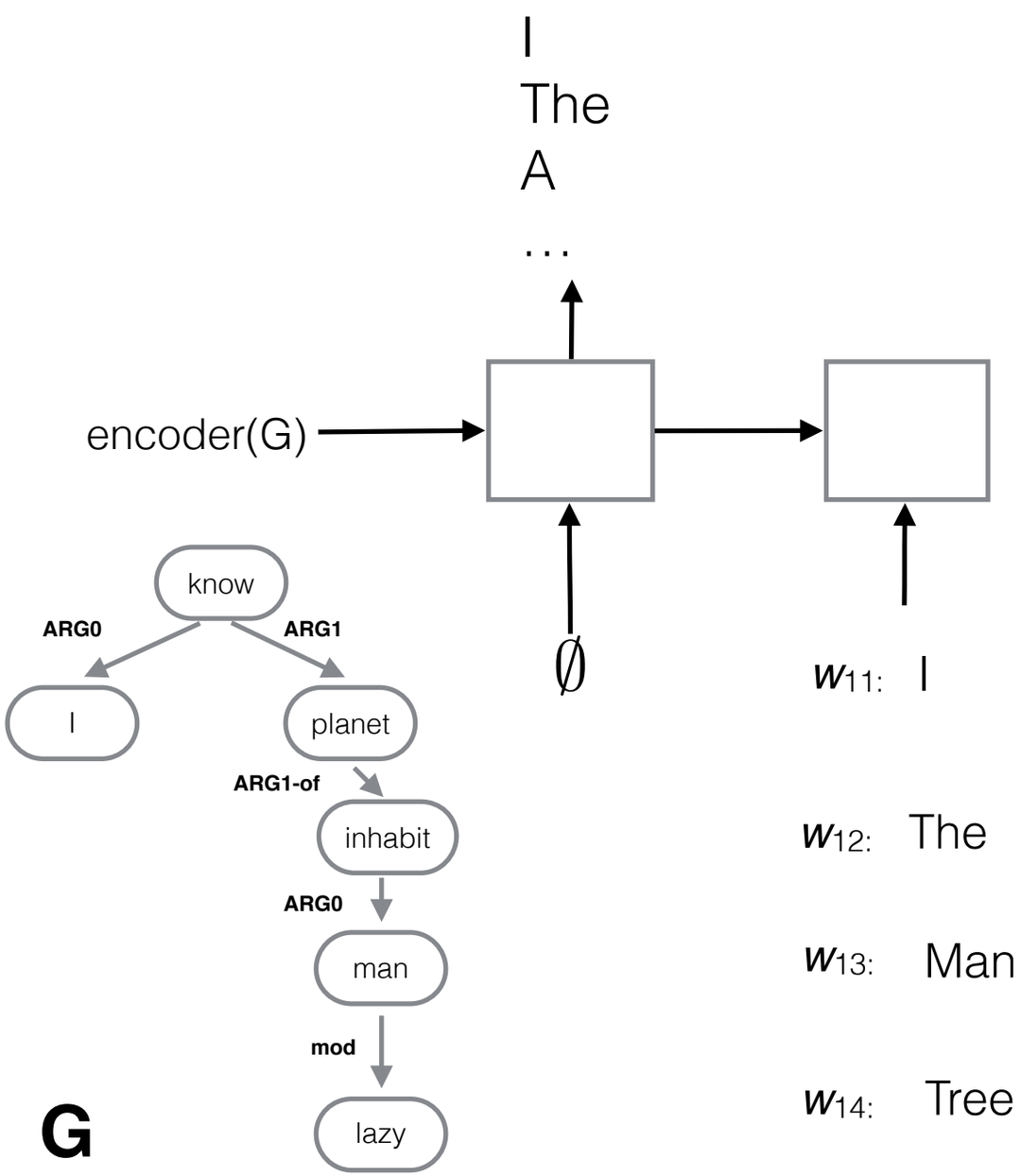
Decoding


encoder(G)
Ø
know
ARG0
ARG1
I
planet
ARG1-of
inhabit
ARG0
man
mod
lazy
G

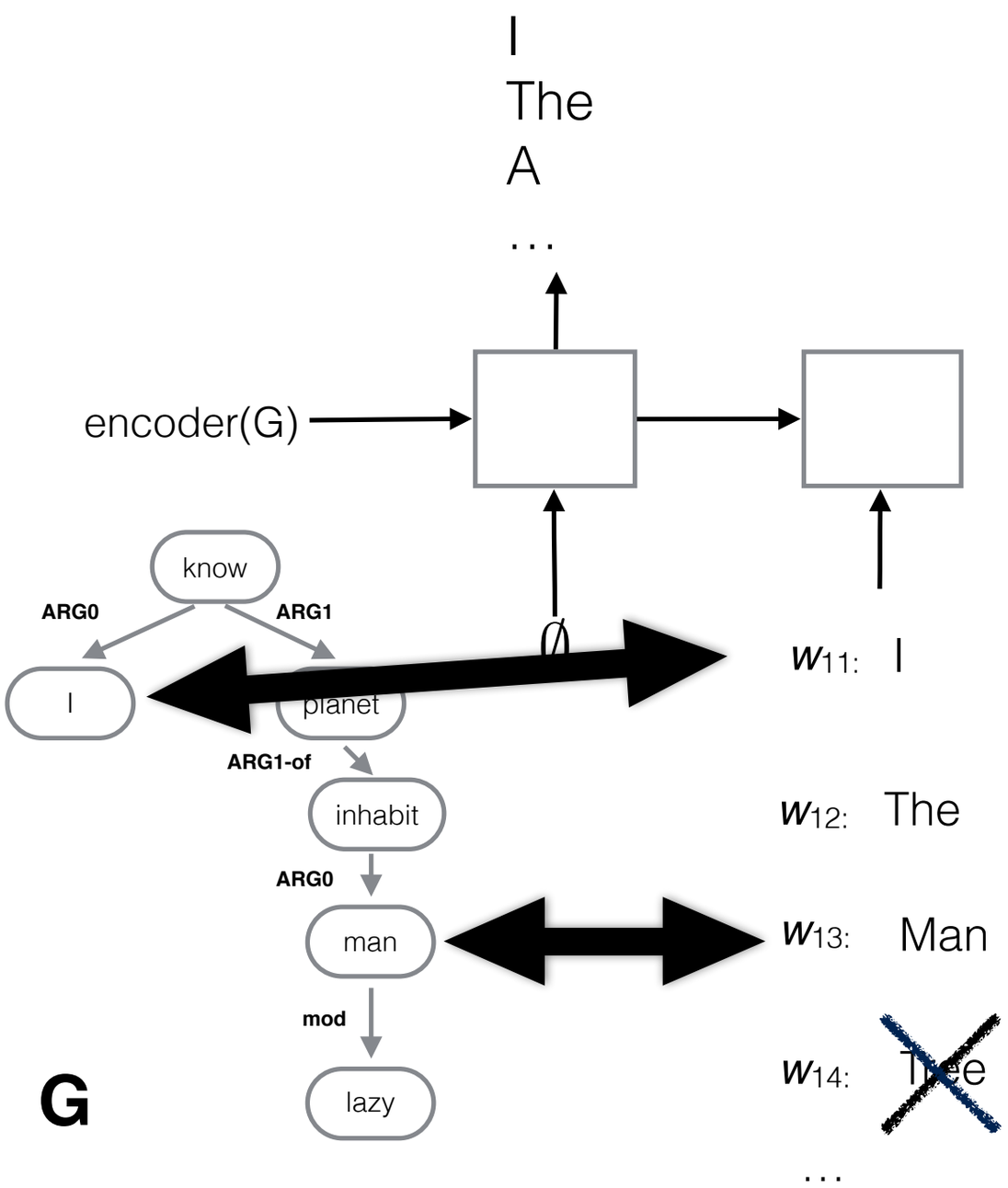
Decoding


I
The
A
…
encoder(G)
Ø
know
ARG0
ARG1
I
planet
ARG1-of
inhabit
ARG0
man
mod
lazy
G

Decoding

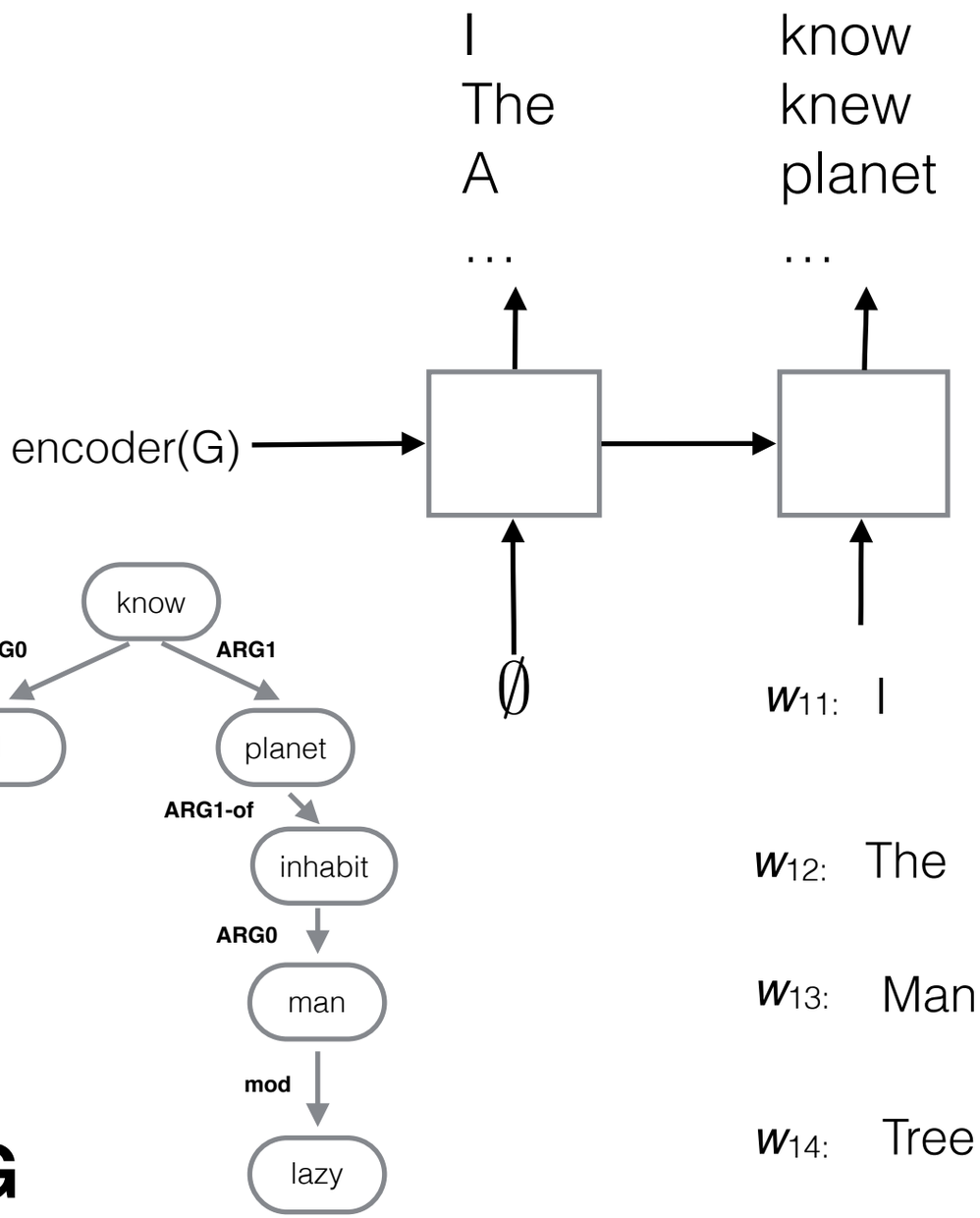

I
The
A
…
encoder(G)
know
ARG0
ARG1
I
planet
ARG1-of
inhabit
ARG0
man
mod
lazy
G
Ø
w_{11}: I
w_{12}: The
w_{13}: Man
w_{14}: Tree
…

Decoding



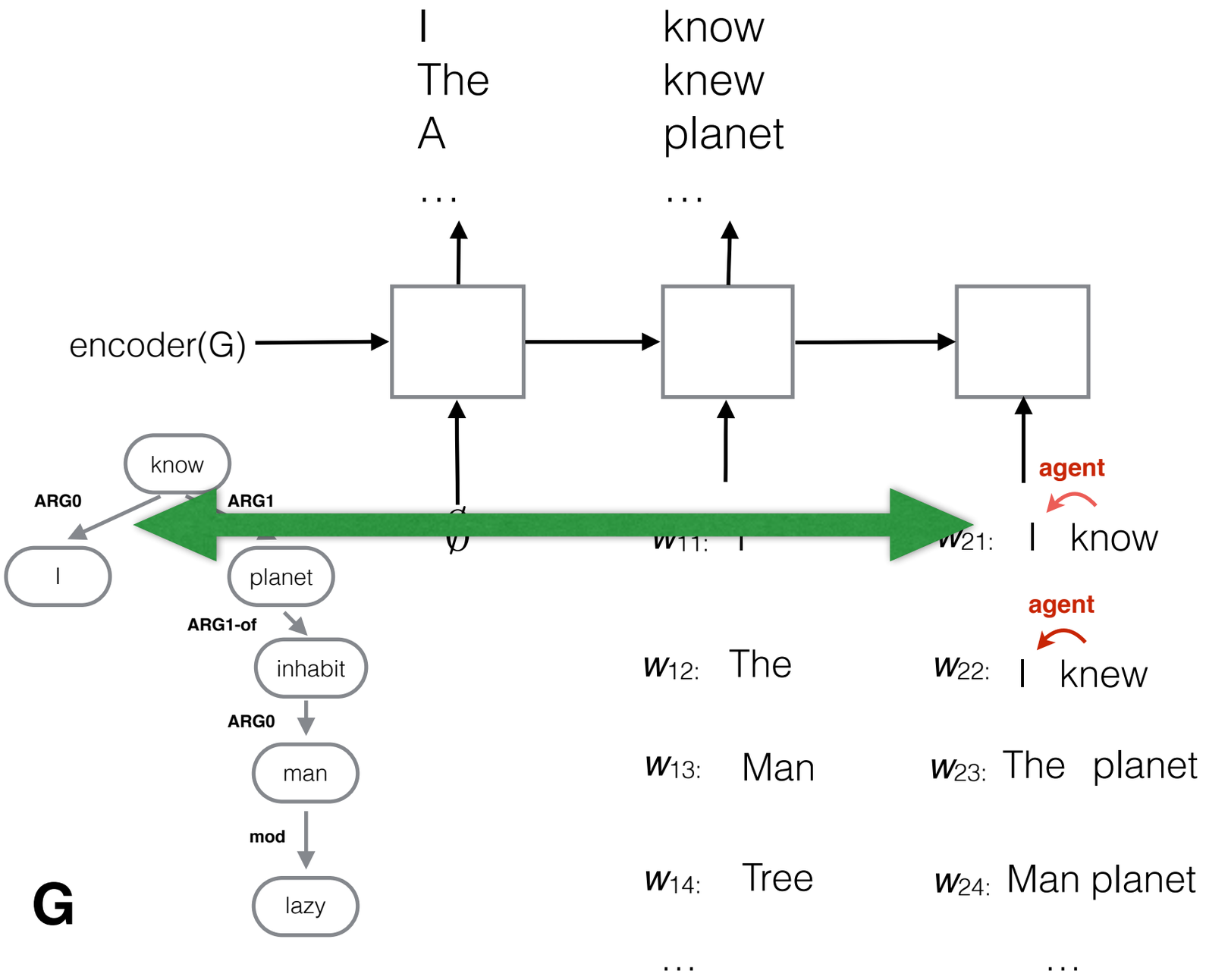
re-score: $\mathbf{S_{LSTM}}$

Decoding



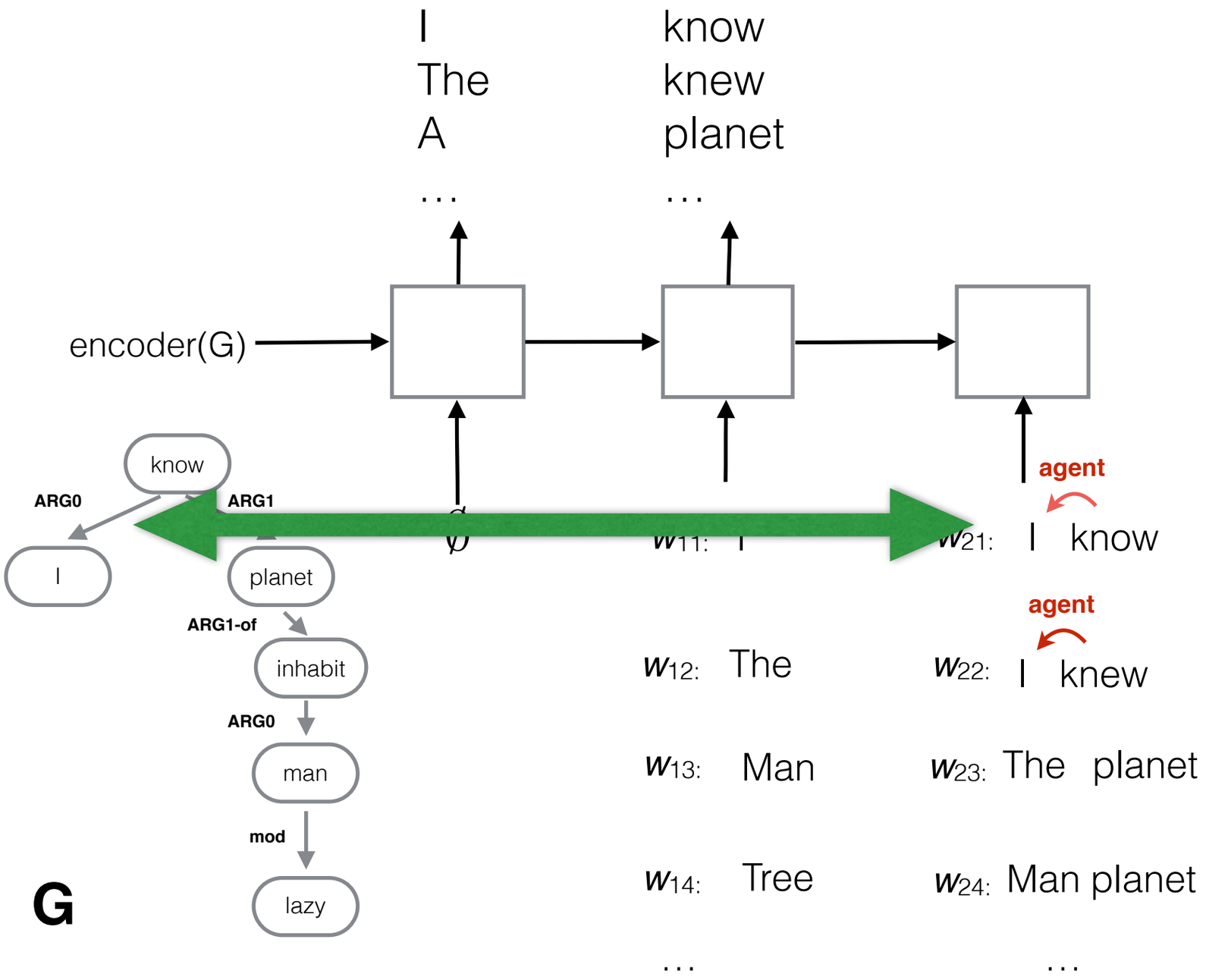
re-score: $\mathbf{S_{LSTM}}$

Decoding



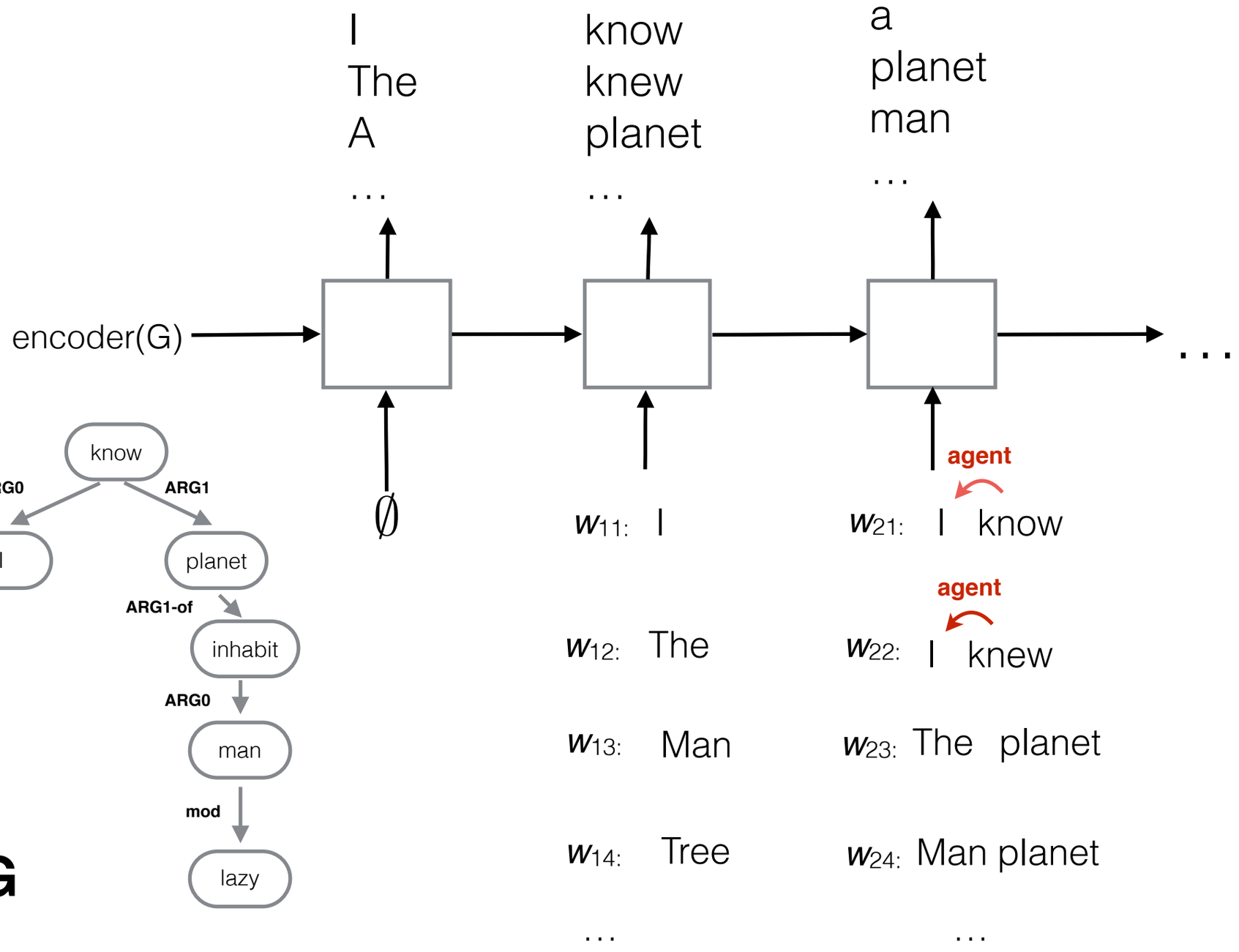
re-score: S_{LSTM} + S_{align}

Decoding



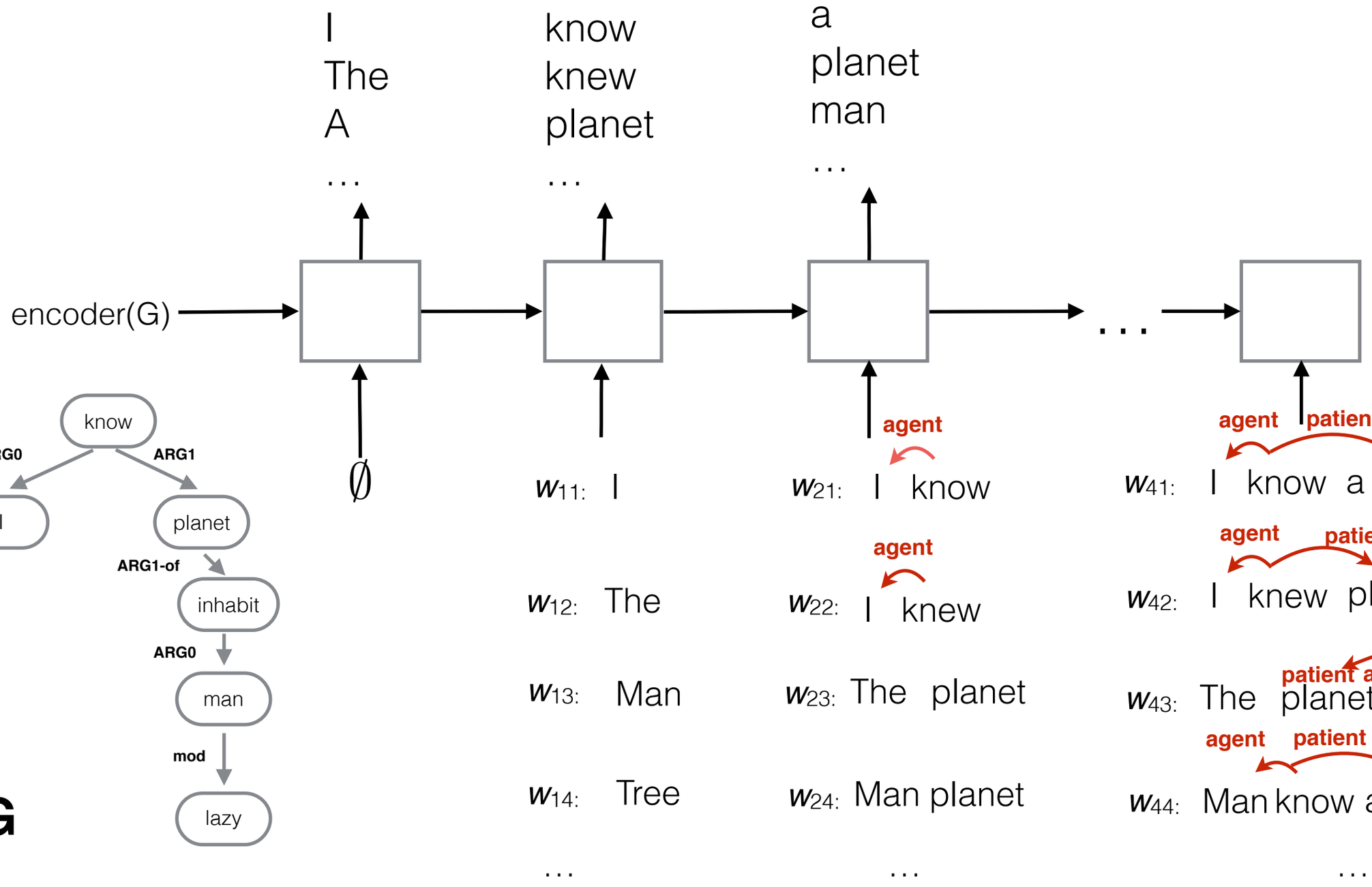
re-score: S_{LSTM} + S_{align}

Decoding



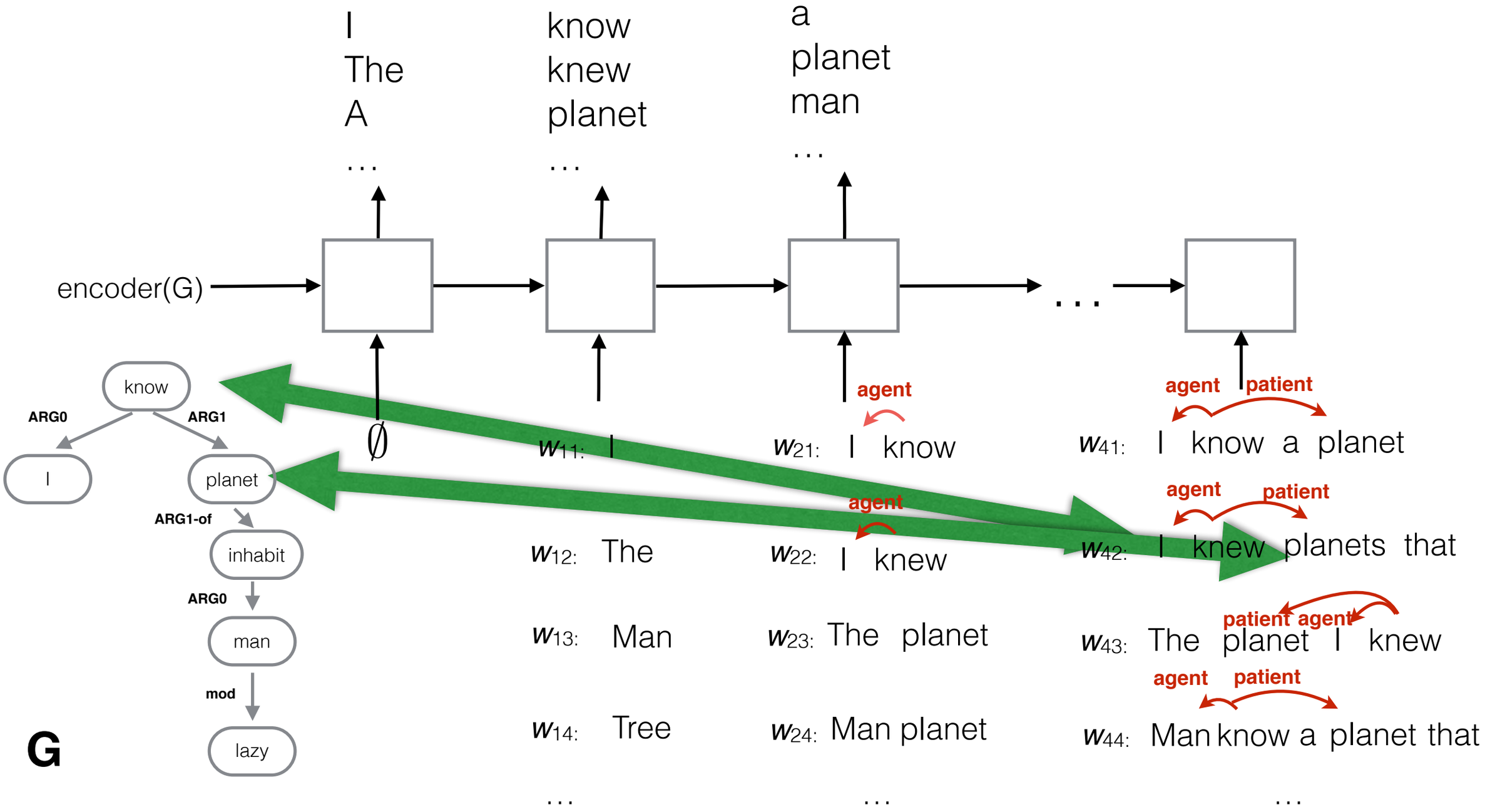
re-score: $\mathbf{S_{LSTM}} + \mathbf{S_{align}}$

Decoding



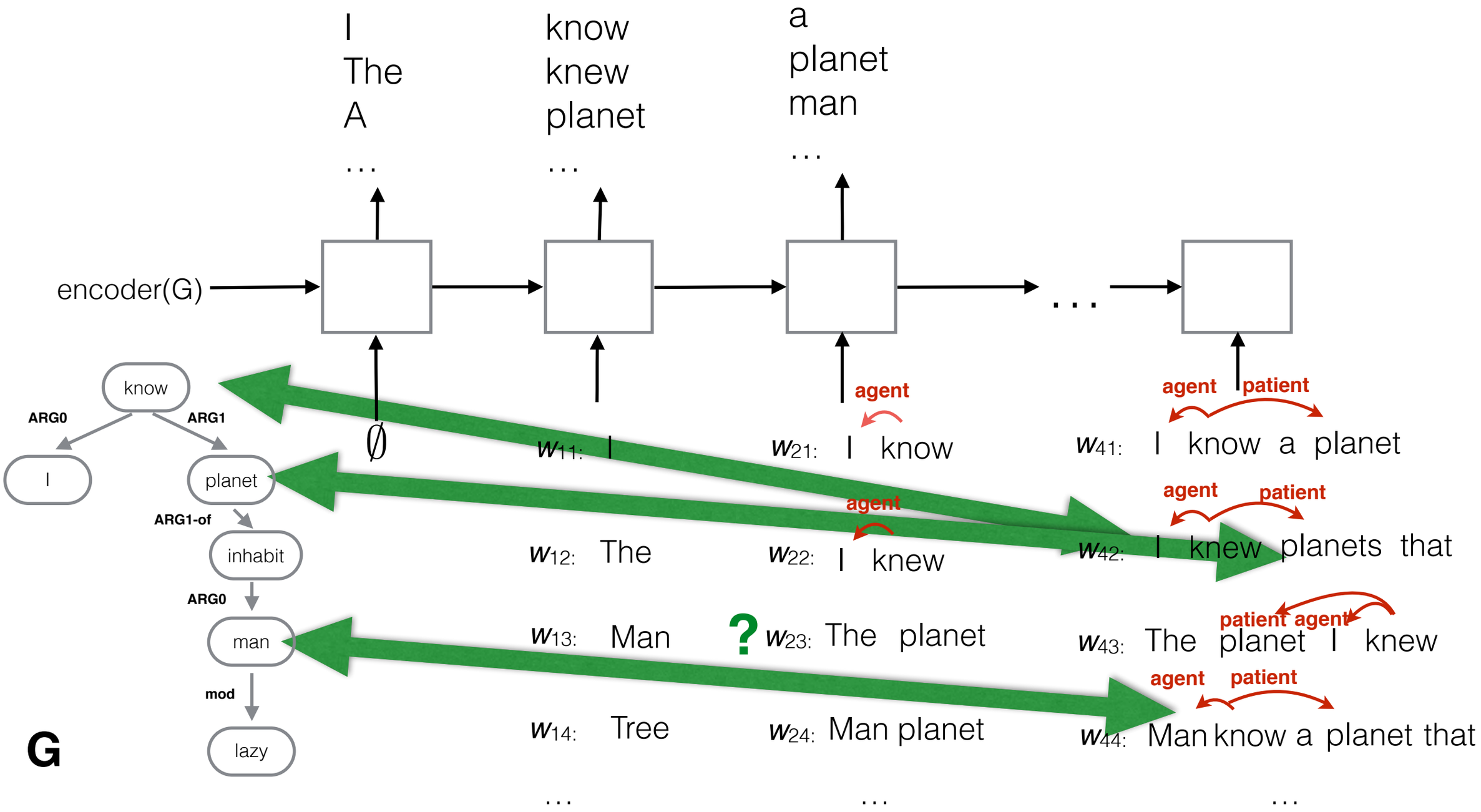
re-score: $\mathbf{S_{LSTM}}$ + $\mathbf{S_{align}}$

Decoding



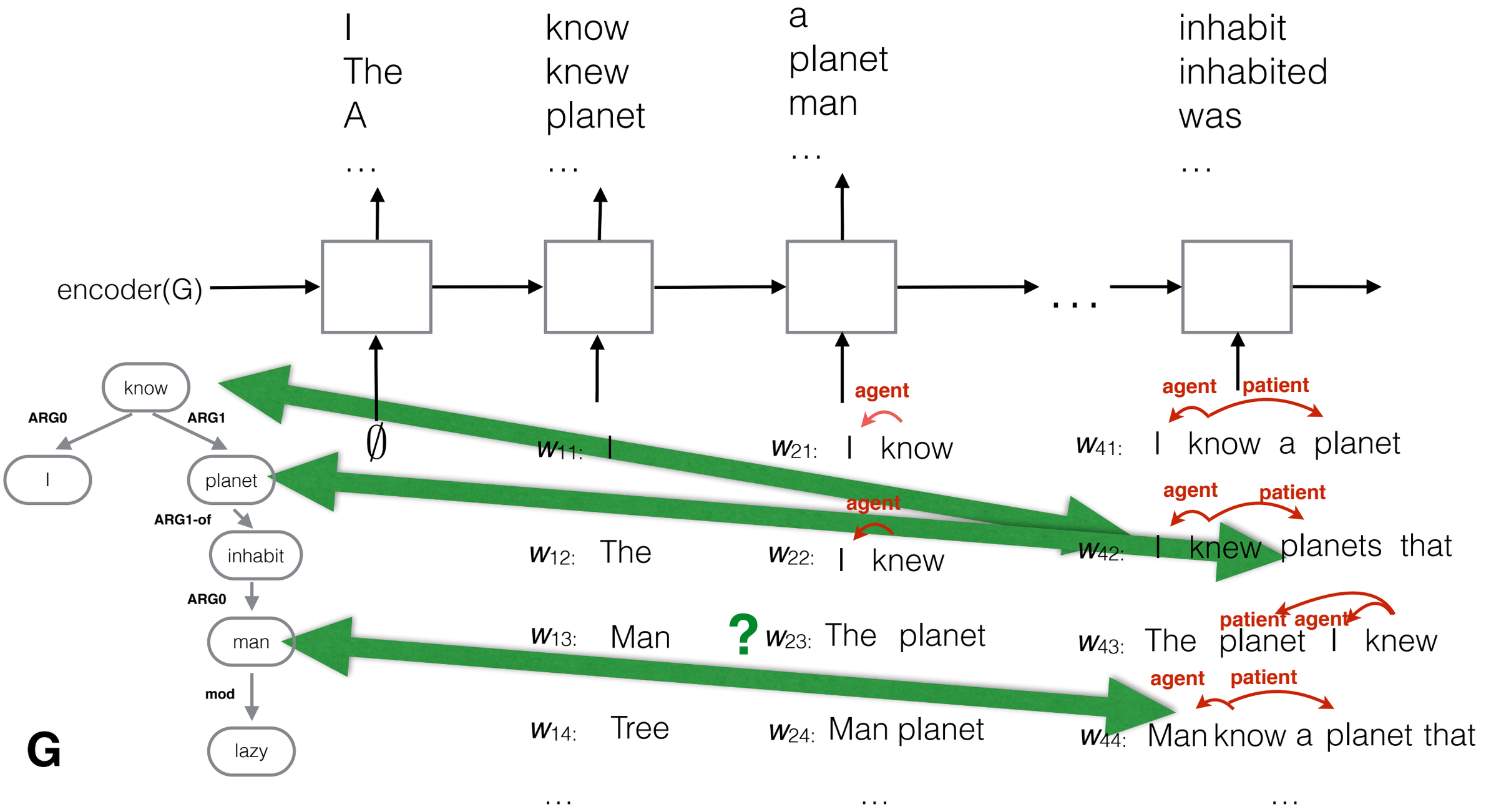
re-score: S_{LSTM} + S_{align}

Decoding



re-score: $\mathbf{S_{LSTM}} + \mathbf{S_{align}}$

Decoding



re-score: $\mathbf{S_{LSTM}} + \mathbf{S_{align}}$

Preliminary Results

Preliminary Results

Experimental Setup

Preliminary Results

Experimental Setup

- Baseline RNN w/o alignment
- NeuGen

Preliminary Results

Experimental Setup

- Baseline RNN w/o alignment
- NeuGen

Dataset

Preliminary Results

Experimental Setup

- Baseline RNN w/o alignment
- NeuGen

Dataset

- AMR LDC2015E86
- Train: 15k, Dev: 1.2k, Test: 1.2k

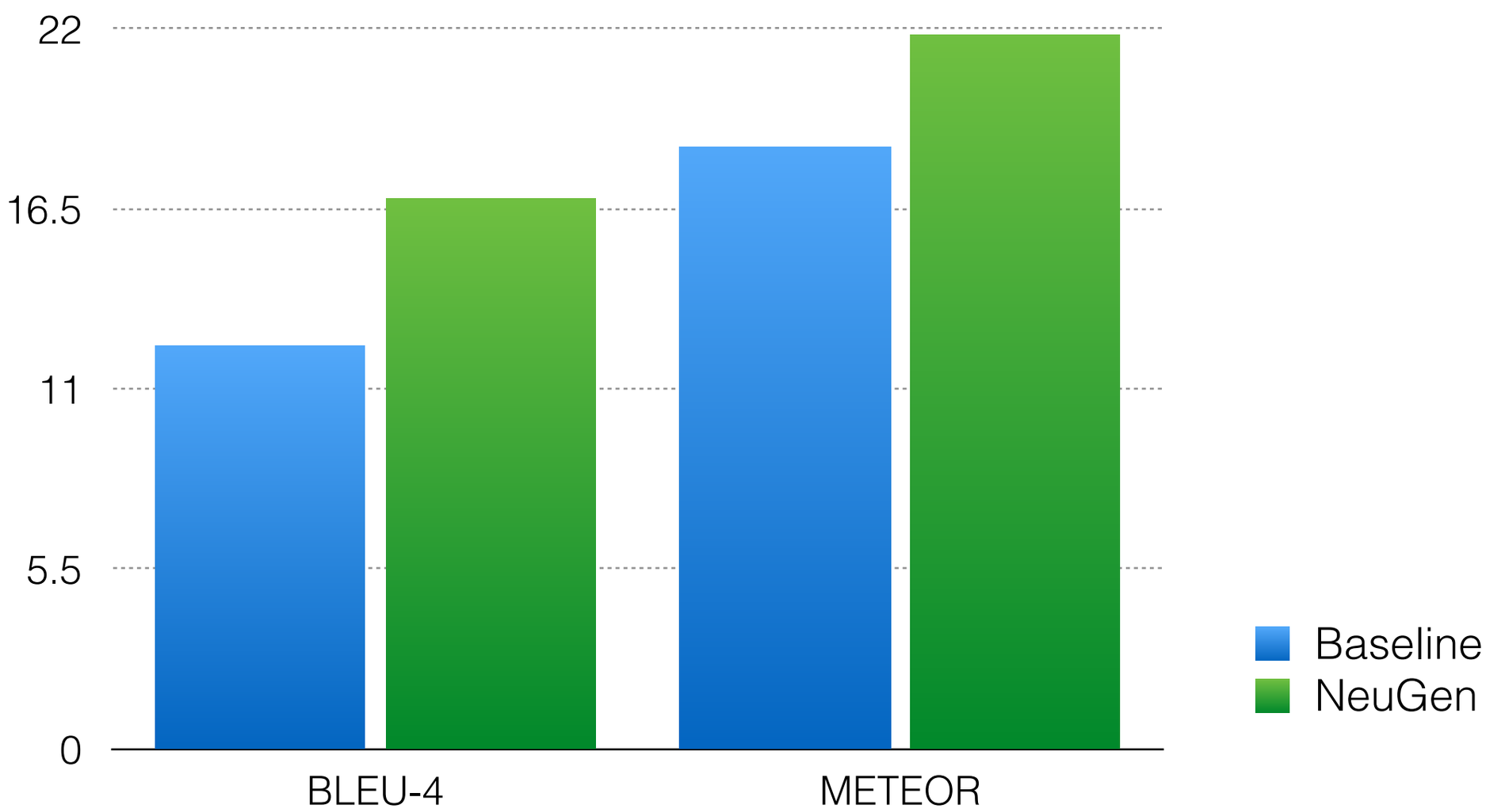
Preliminary Results

Experimental Setup

- Baseline RNN w/o alignment
- NeuGen

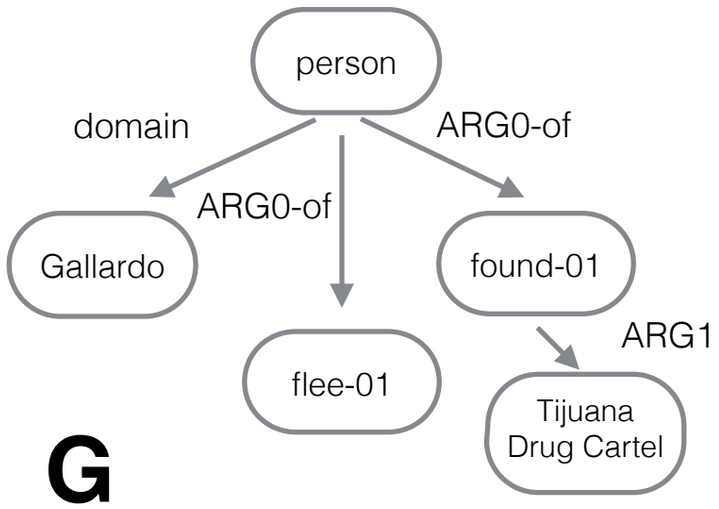
Dataset

- AMR LDC2015E86
- Train: 15k, Dev: 1.2k, Test: 1.2k



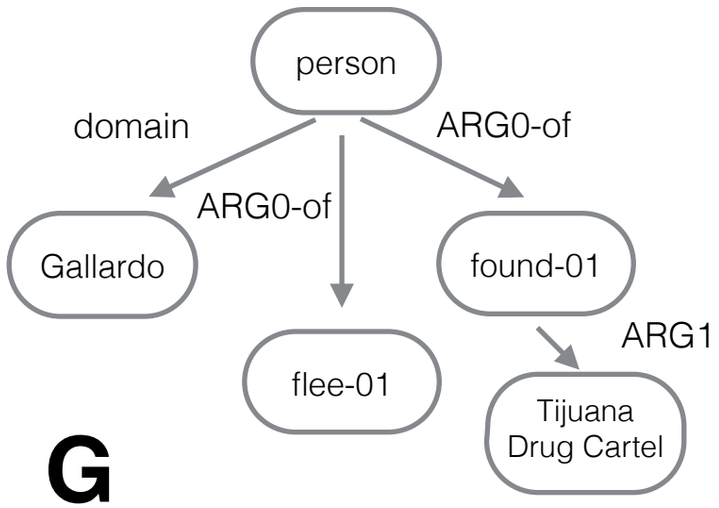
Example Output

Example Output



G

Example Output



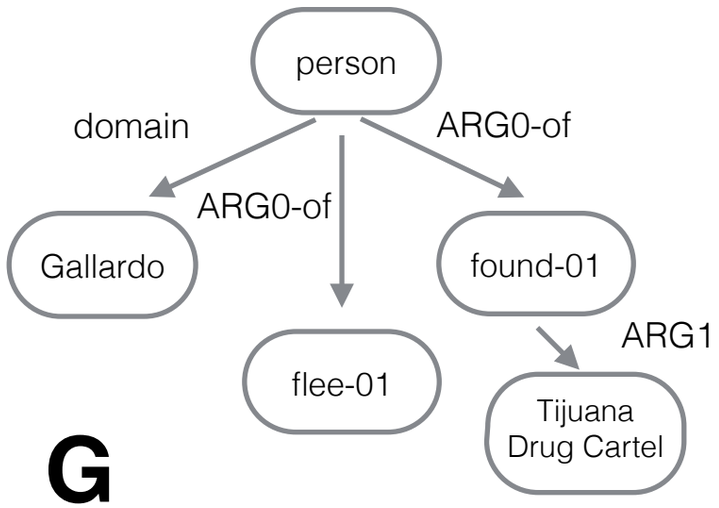
G

Gold: `Gallardo is the fugitive founder of Tijuana Drug Cartel.`

Baseline: `Gallardo is the founder of Tijuana Drug Cartel.`

NeuGen: `Tijunana Drug Cartel is the founder of Gallardo.`

Example Output



G

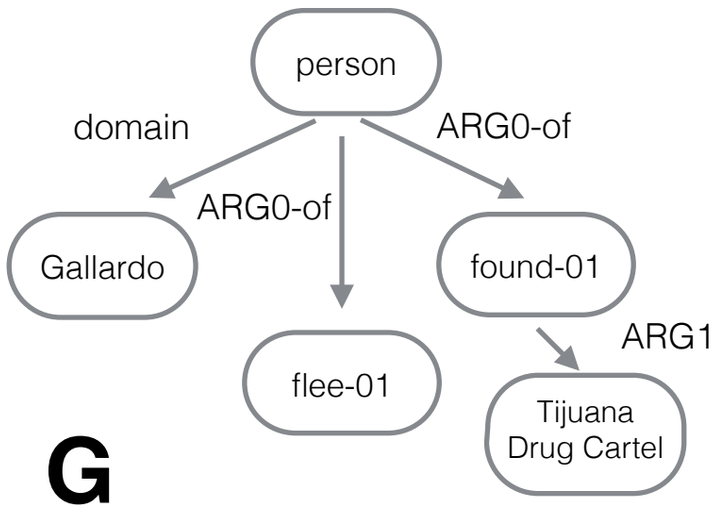
Gold: `Gallardo is the fugitive founder of Tijuana Drug Cartel.`

Baseline: `Gallardo is the founder of Tijuana Drug Cartel.`

NeuGen: `Tijunana Drug Cartel is the founder of Gallardo.`

```
Gallardo

Tijuana Drug Cartel

State

Country

…
```

Example Output

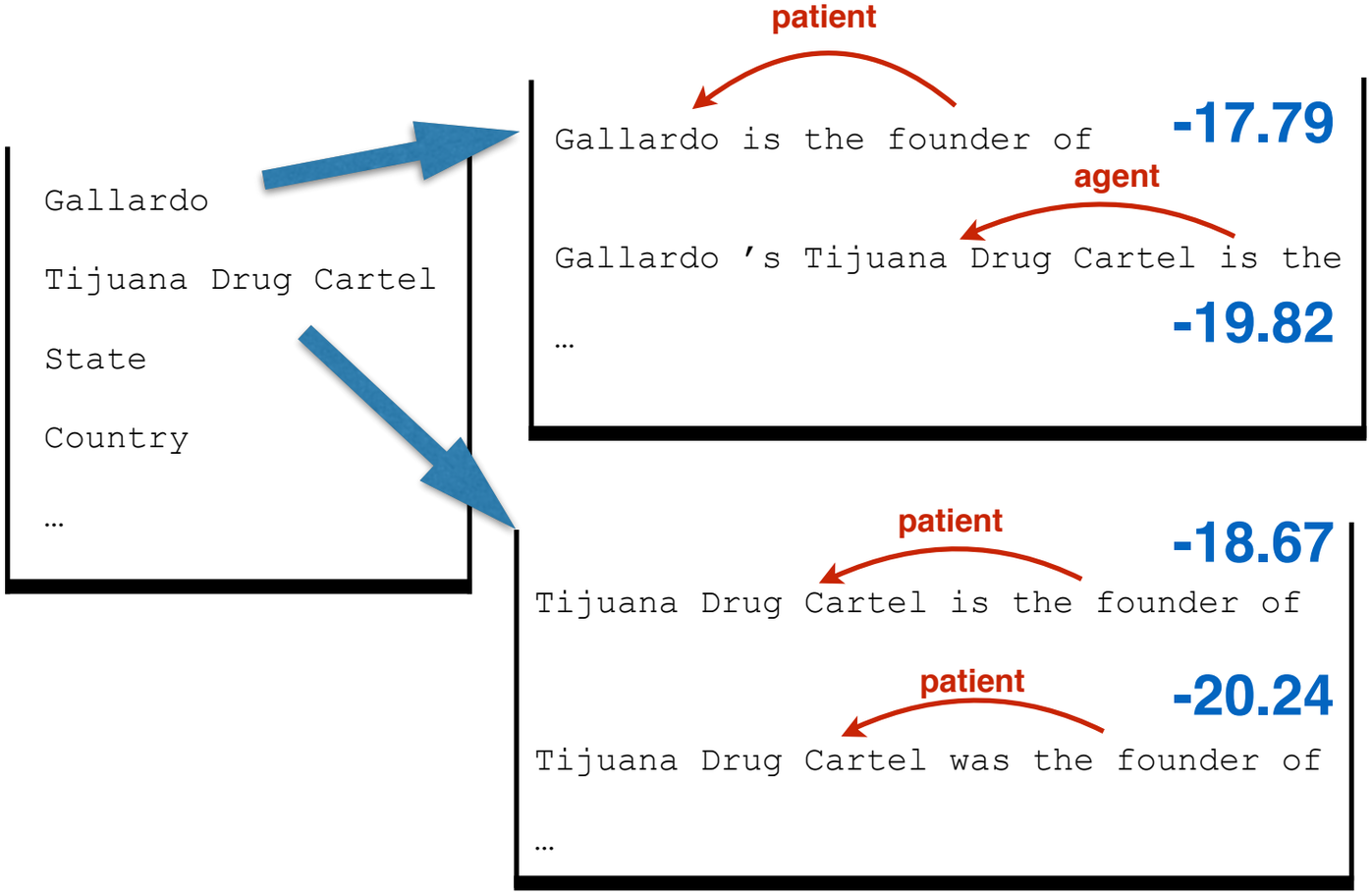


G

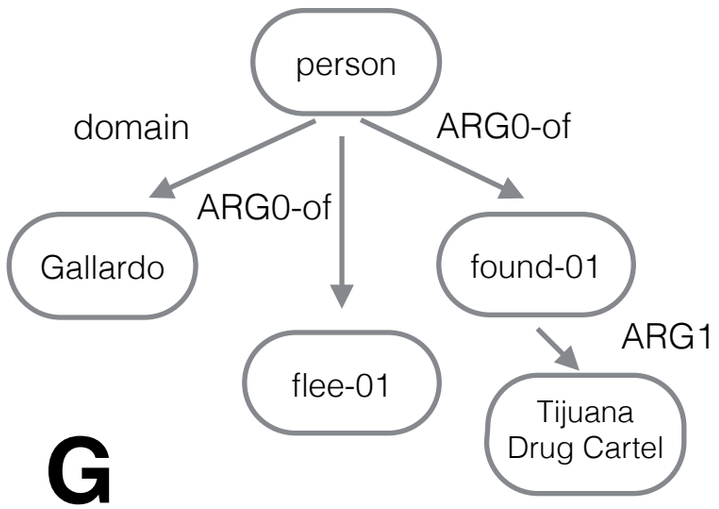
Gold: `Gallardo is the fugitive founder of Tijuana Drug Cartel.`

Baseline: `Gallardo is the founder of Tijuana Drug Cartel.`

NeuGen: `Tijunana Drug Cartel is the founder of Gallardo.`



Example Output

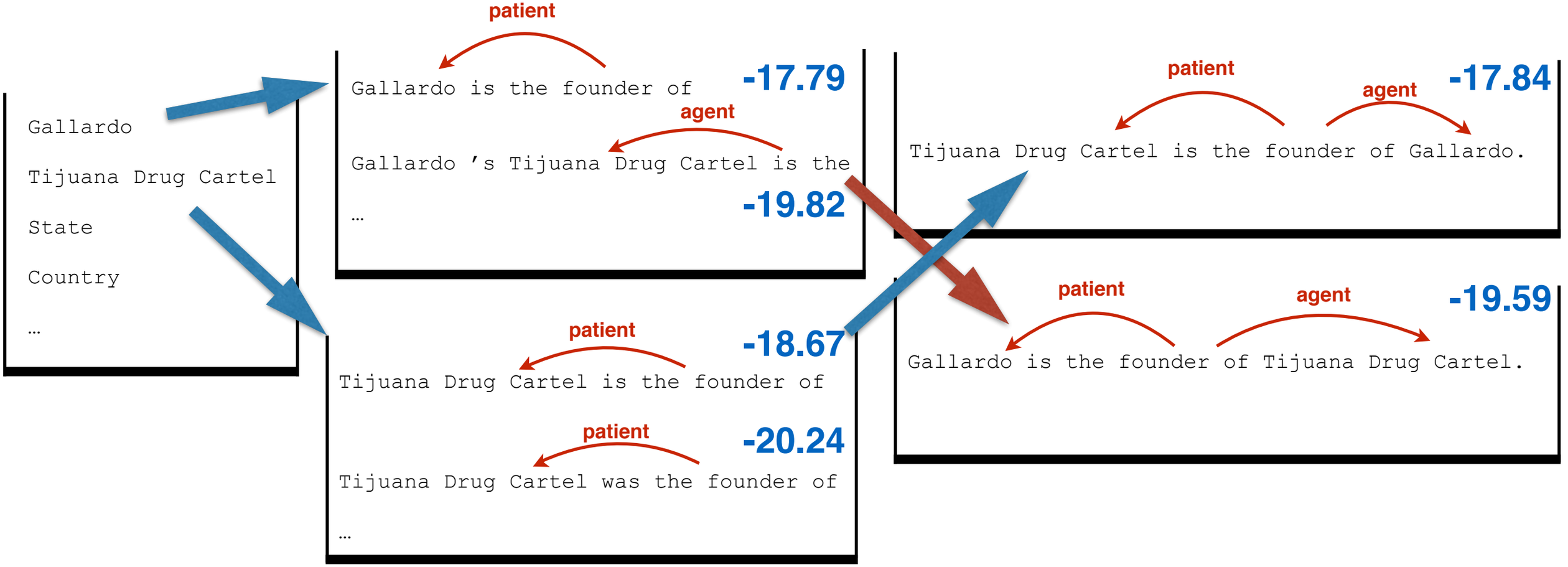


G

Gold: `Gallardo is the fugitive founder of Tijuana Drug Cartel.`

Baseline: `Gallardo is the founder of Tijuana Drug Cartel.`

NeuGen: `Tijunana Drug Cartel is the founder of Gallardo.`



Next Steps

Next Steps

- “Machine-generated” representations
  - λ-calculus (ATIS)
  - Blocks World complex descriptions (from ISI)

Next Steps

- “Machine-generated” representations
  - λ-calculus (ATIS)
  - Blocks World complex descriptions (from ISI)

Next Steps

- “Machine-generated” representations
  - λ-calculus (ATIS)
  - Blocks World complex descriptions (from ISI)
- Downstream Text-to-Text Applications

Thank You