Combinatory Categorial Grammar

Abstract Meaning Representation

Broad-coverage CCG Semantic Parsing with AMR

Yoav Artzi
Cornell University

Kenton Lee Luke Zettlemoyer
University of Washington

Semantic Parsing

Show me all papers about semantic parsing

Grammar

$$\lambda x.\text{paper}(x) \wedge \text{topic}(x, \text{SEMANTIC\_PARSING})$$

Less Supervision	**More Domains**	**Situated Parsing**
Answers	Databases	Spatial Observations
Demonstrations	Large Knowledge-bases	Linguistic Context
Conversations	Instructions	Database Content
	Web Tables	
	Time	

Semantic Parsing

Show me all papers about semantic parsing

Grammar

$$\lambda x.\text{paper}(x) \wedge \text{topic}(x, \text{SEMANTIC\_PARSING})$$

Less
Su

More Domains

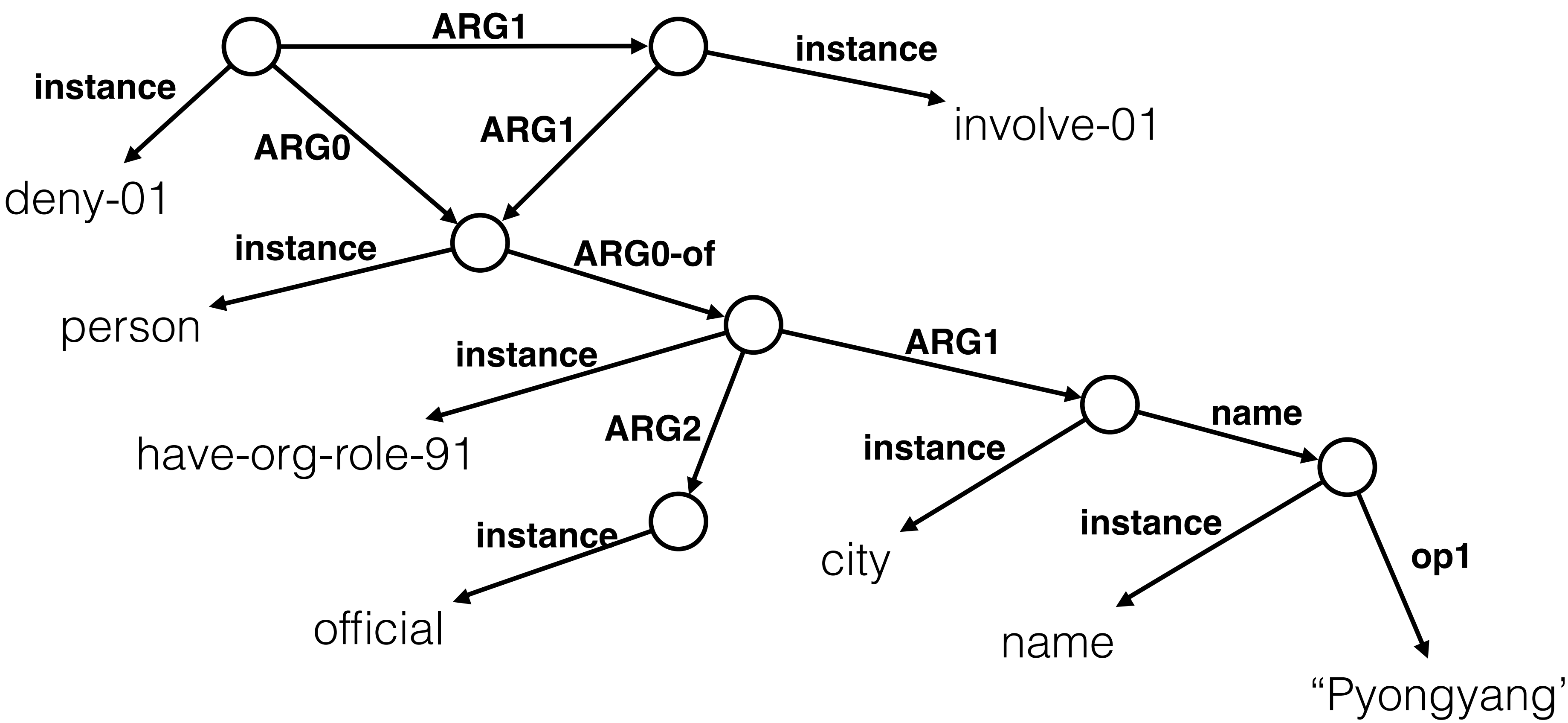
Situated Parsing

Non-compositional Semantics

Broad-coverage Grammar Induction

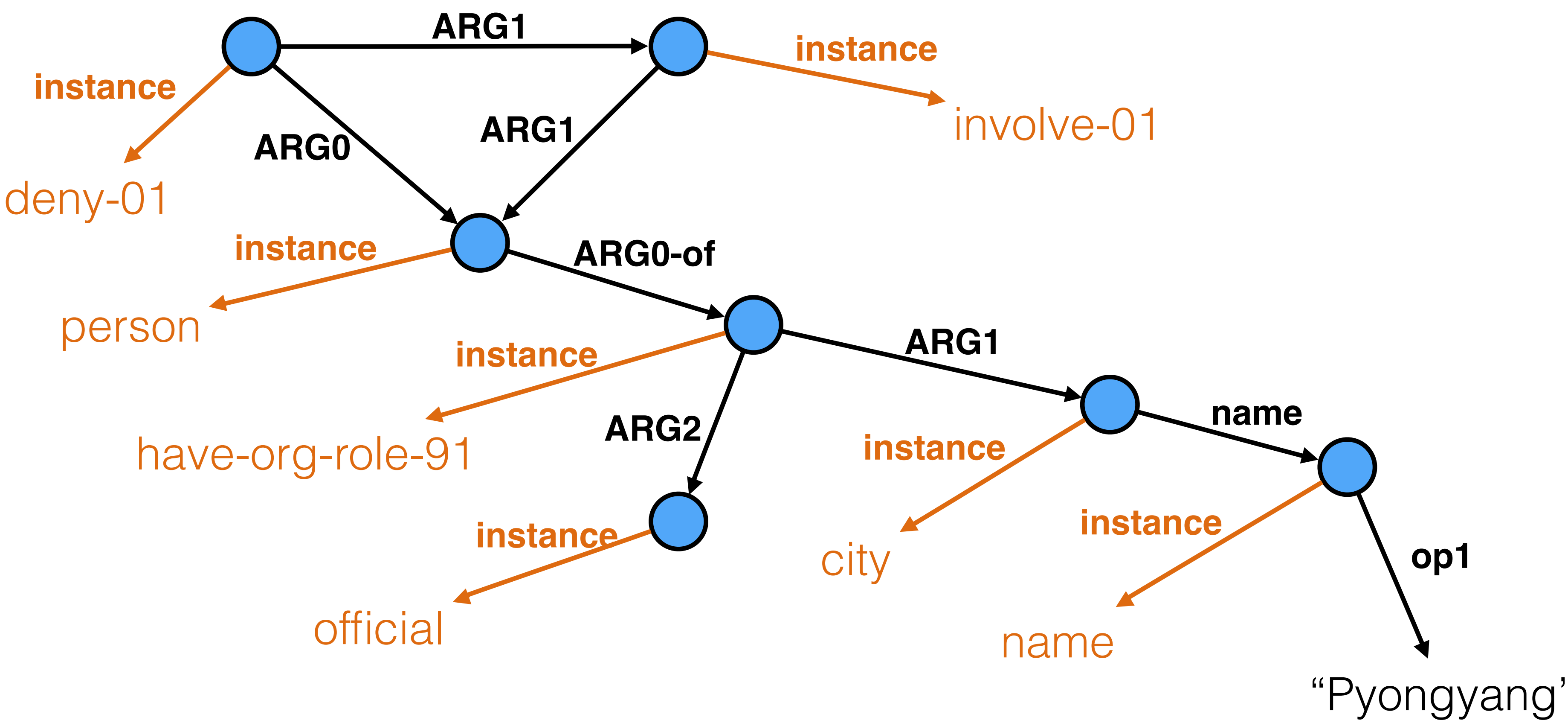
Abstract Meaning Representation

Pyongyang officials denied their involvement

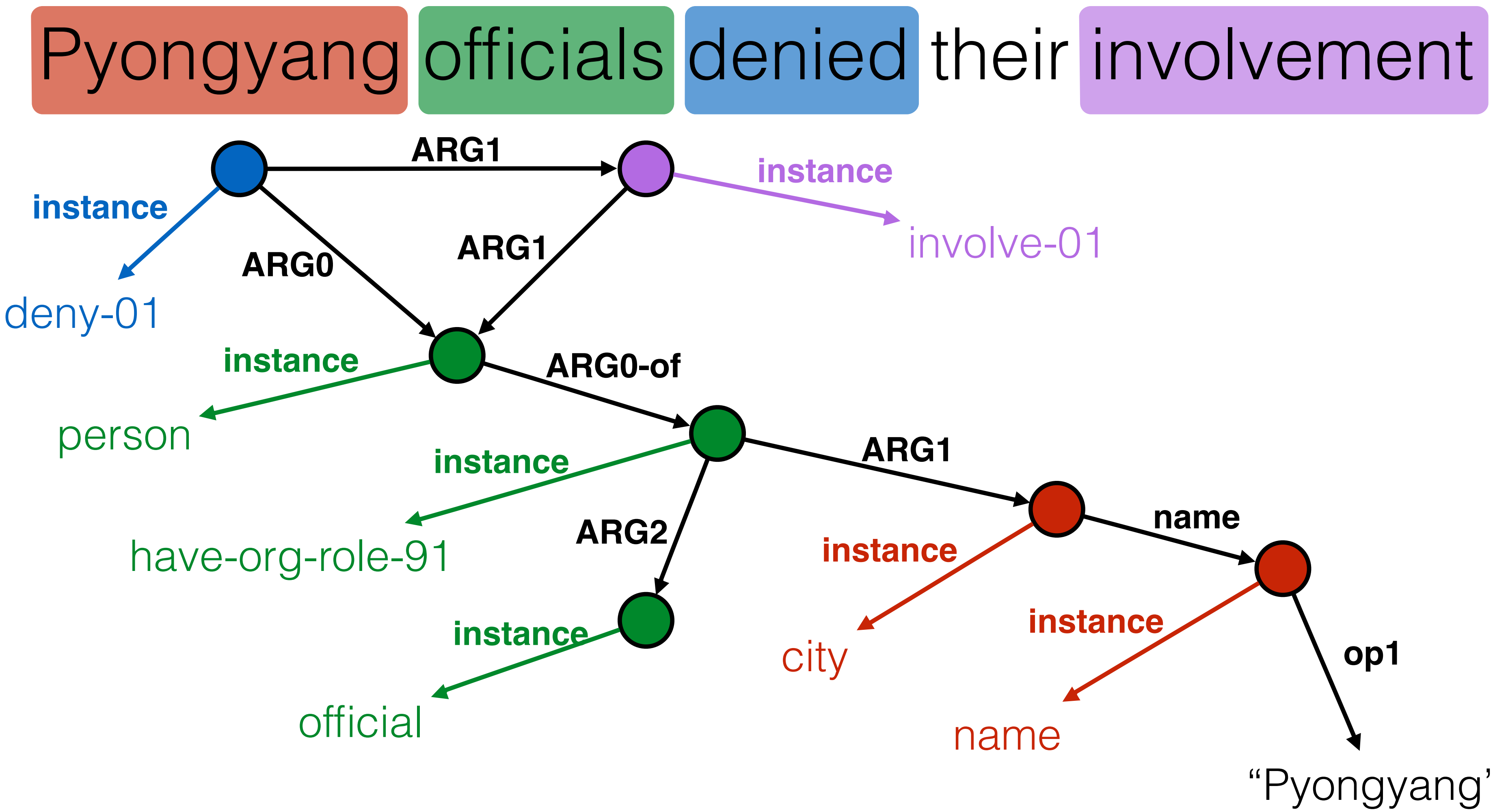


AMR: Instances

Pyongyang officials denied their involvement

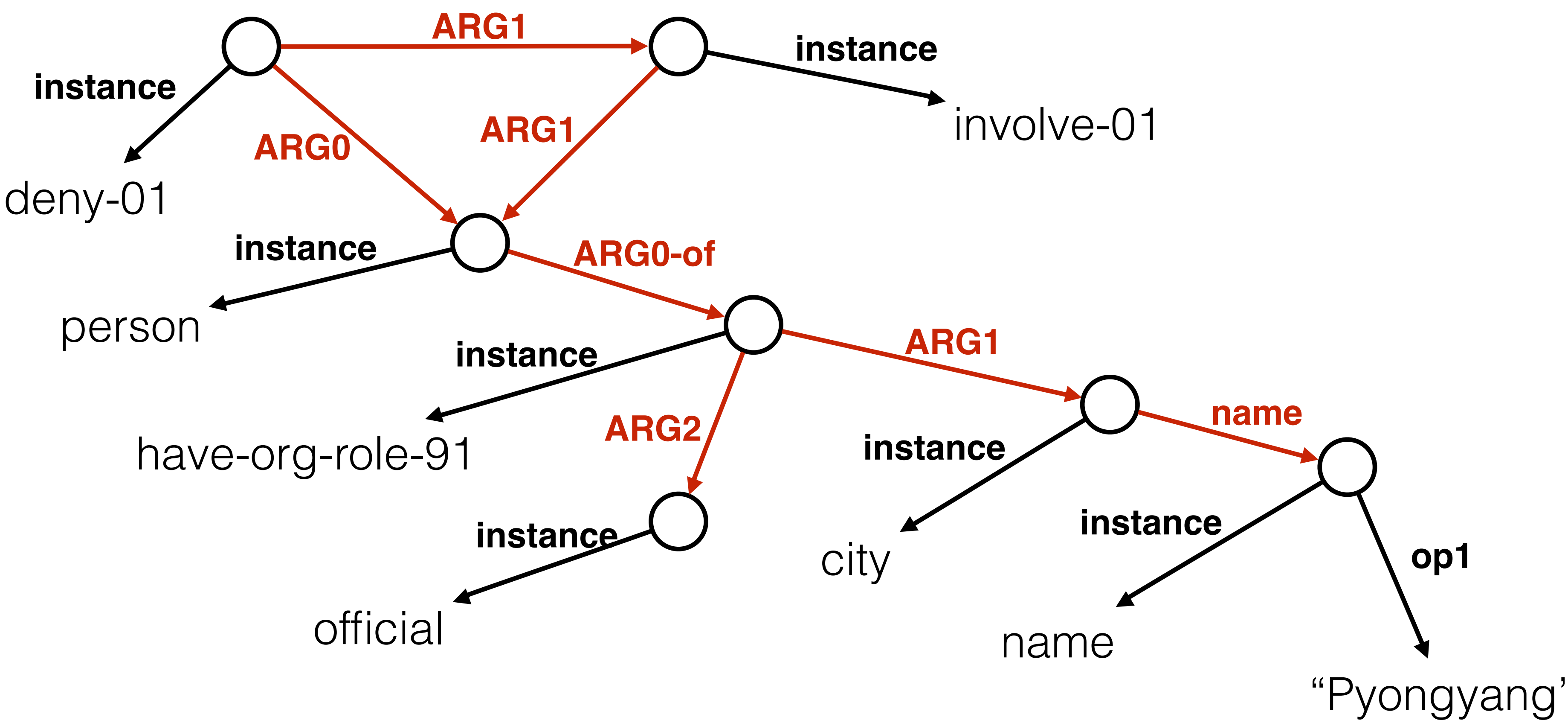


AMR: Instances



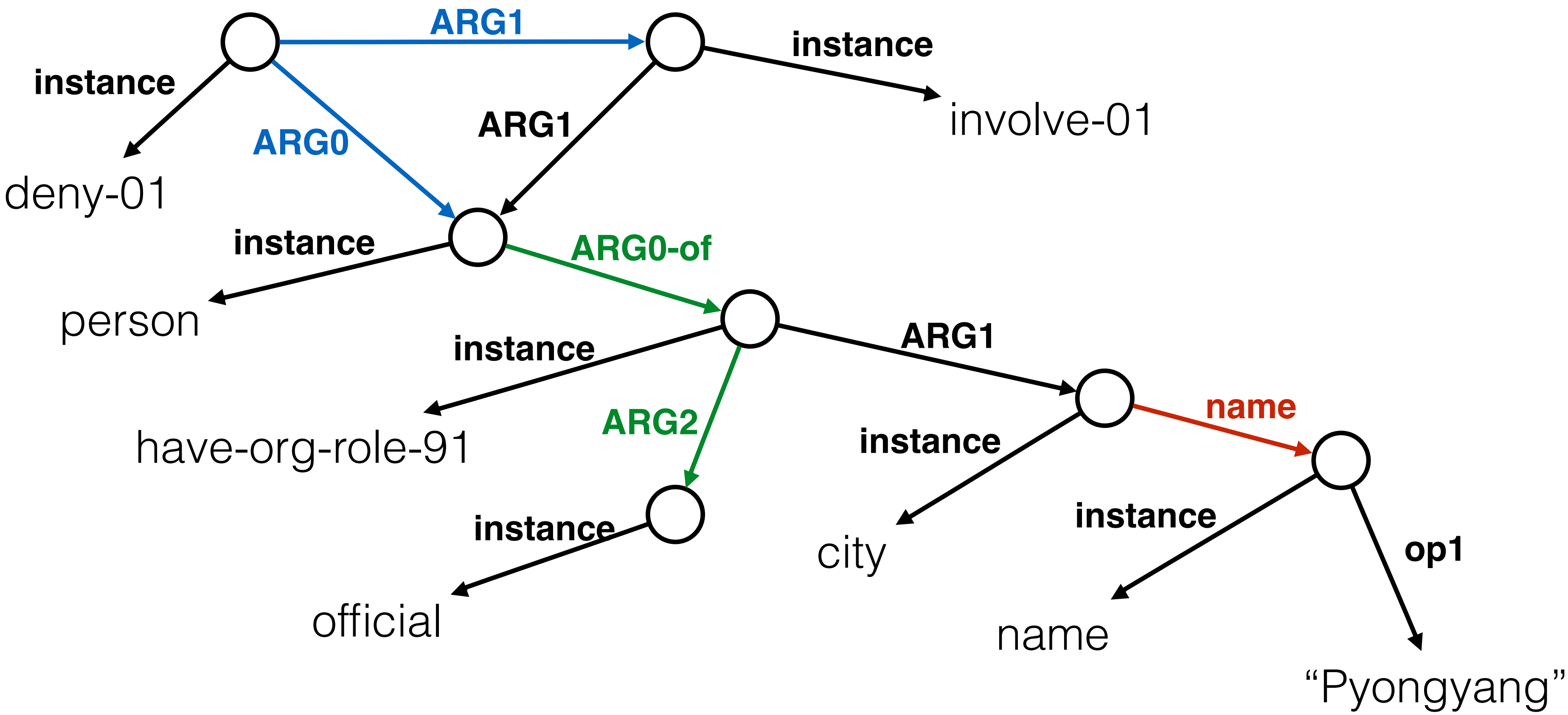
AMR: Relations

Pyongyang officials denied their involvement



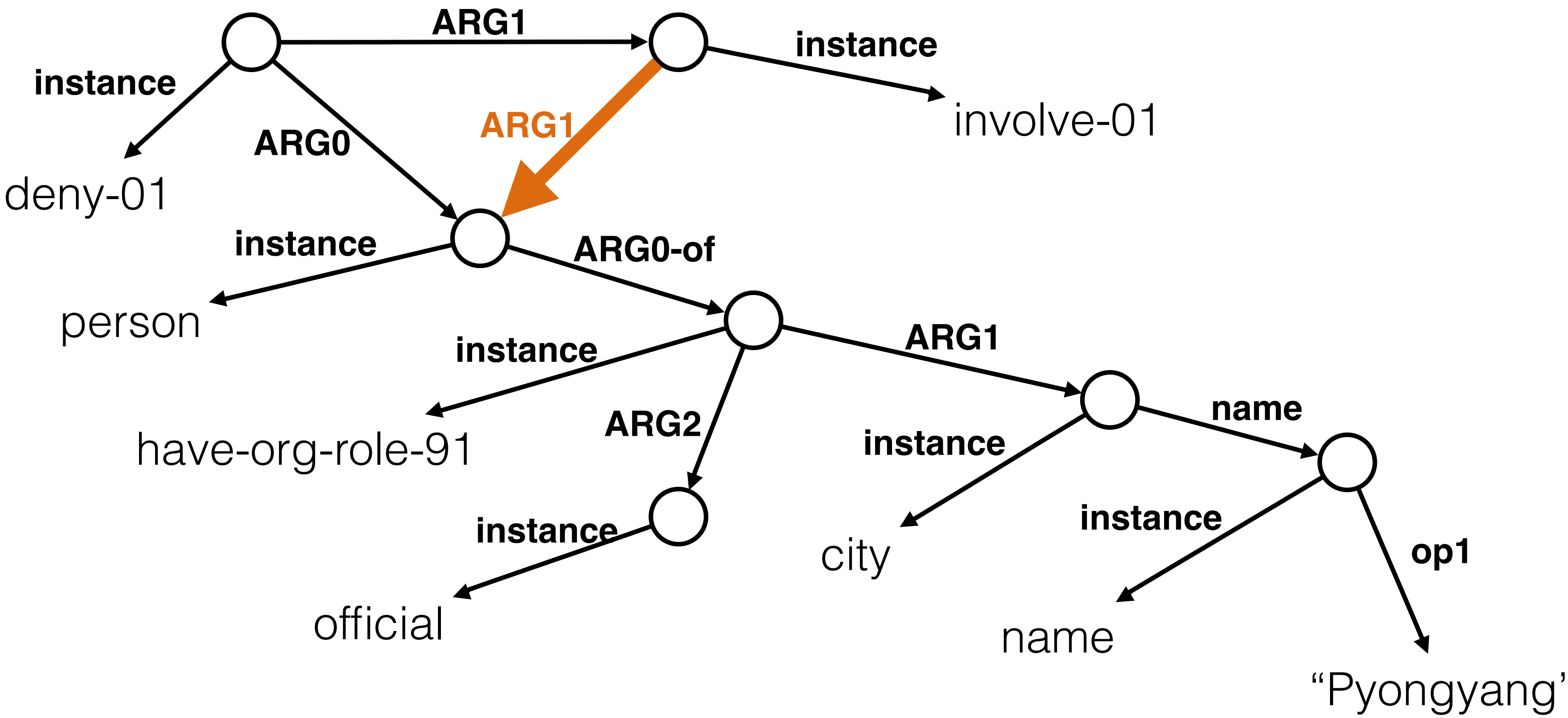
AMR: Relations

Pyongyang officials denied their involvement



AMR: Relations

Pyongyang officials denied their involvement



AMR and Combinatory Categorial Grammar

Great opportunity study CCG semantic parsing at scale

Challenges:

- Distant non-compositional dependencies
- Longer sentences
- Higher syntactic variability

Parsing Approach

- Use CCG to recover compositional parse structure
- Second stage to resolve non-compositional phenomena, such as co-reference resolution

Combinatory Categorial Grammar

Category

$$\underbrace{S\backslash NP/NP}_{\text{Syntax}} : \underbrace{\lambda x.\lambda y.\lambda d.\text{deny-01}(d) \wedge \text{ARG0}(d, y) \wedge \text{ARG1}(d, x)}_{\text{Semantics}}$$

Lexicon

Assign category to words

denied $\rightarrow$ $S\backslash NP/NP : \lambda x.\lambda y.\lambda d.\text{deny-01}(d) \wedge \text{ARG0}(d, y) \wedge \text{ARG1}(d, x)$

Combinators

Unary and binary operators to combine categories

CCG

Entries from Lexicon

$$
\begin{array}{ccc}
\text{CCG} & \text{is} & \text{fun} \\
\hline
NP & S\backslash NP/ADJ & ADJ \\
\text{CCG} & \lambda f.\lambda x.f(x) & \lambda x.\text{fun}(x) \\
 & \multicolumn{2}{c}{\xrightarrow{\hspace{12em}} >} \\
 & \multicolumn{2}{c}{S\backslash NP} \\
 & \multicolumn{2}{c}{\lambda x.\text{fun}(x)} \\
\multicolumn{3}{c}{\xrightarrow{\hspace{20em}} <} \\
\multicolumn{3}{c}{S} \\
\multicolumn{3}{c}{\text{fun}(\text{CCG})}
\end{array}
$$

Parse Steps

Logical Form

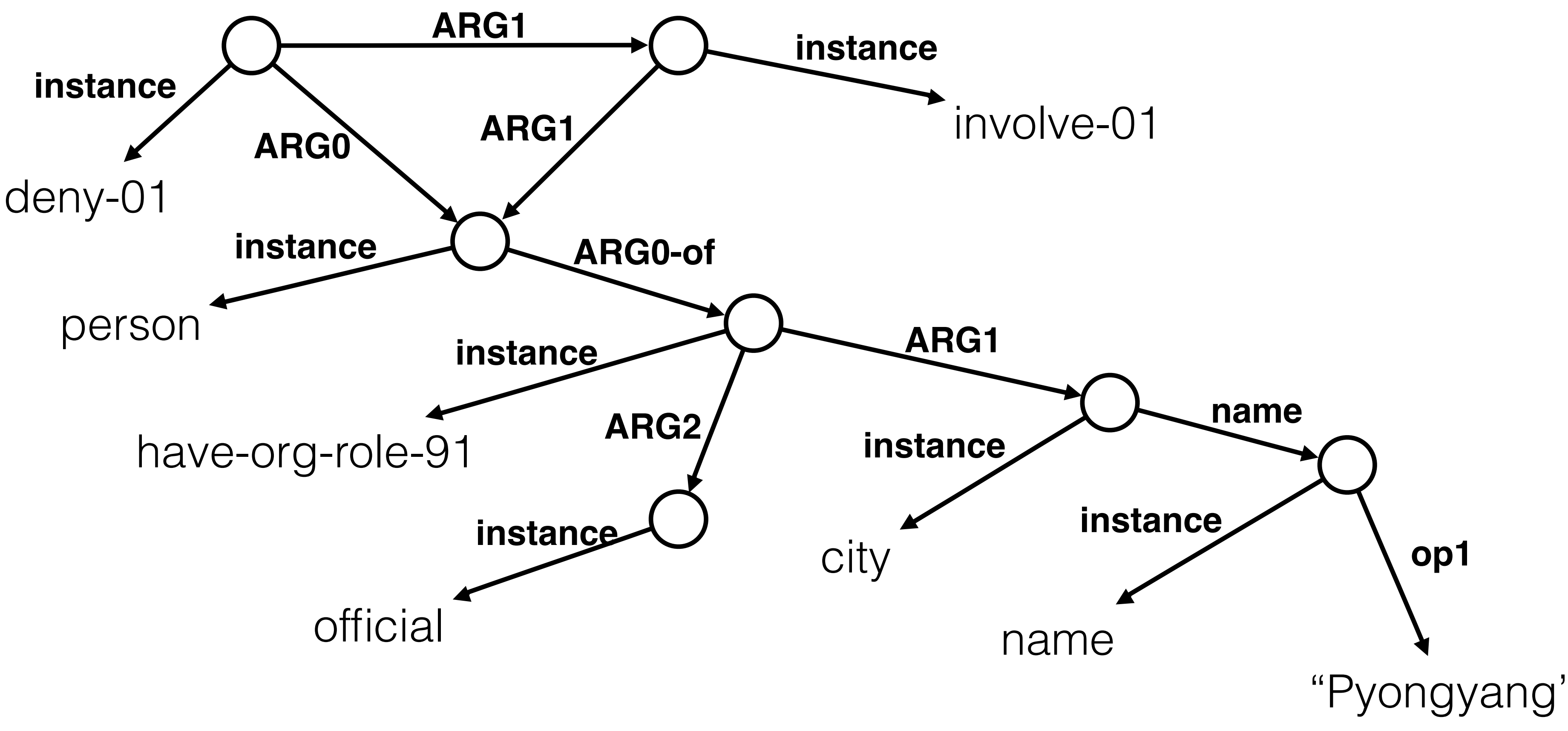
Combinators

Lexicon

Learned

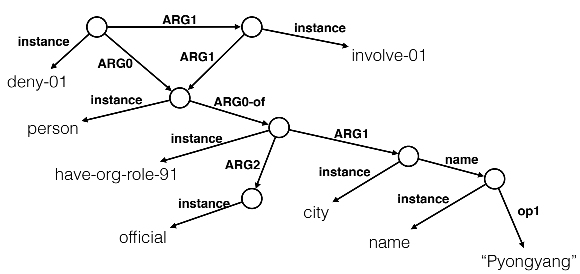
AMR to Lambda Calculus

Pyongyang officials denied their involvement



AMR to Lambda Calculus

Pyongyang officials denied their involvement



$$
\begin{aligned}
&\mathcal{A}_1(\lambda d.\text{deny-01}(d)\wedge\\
&\quad \text{ARG0}(d, \mathcal{A}_2(\lambda p.\text{person}(p)\wedge\\
&\qquad \text{ARG0-of}(p, \mathcal{A}_3(\lambda h.\text{have-org-role-91}(h)\wedge\\
&\qquad\quad \text{ARG1}(h, \mathcal{A}_4(\lambda c.\text{city}(c)\wedge\\
&\qquad\qquad \text{name}(c, \mathcal{A}_5(\lambda n.\text{name}(n) \wedge \text{op1}(n, \text{PYONGYANG})))))\wedge\\
&\qquad\quad \text{ARG2}(h, \mathcal{A}_6(\lambda o.\text{official}(o))))))\wedge\\
&\quad \text{ARG1}(d, \mathcal{A}_7(\lambda i.\text{involve-01}(i) \wedge \text{ARG1}(i, \mathcal{R}(2))))))
\end{aligned}
$$

AMR ⟷ **Deterministic Conversion** ⟷ **Lambda Calculus**

AMR Lambda Calculus: Instances

Pyongyang officials denied their involvement

$$
\begin{aligned}
&\mathcal{A}_1(\lambda d.\text{deny-01}(d)\wedge \\
&\quad \text{ARG0}(d, \mathcal{A}_2(\lambda p.\text{person}(p)\wedge \\
&\quad\quad \text{ARG0-of}(p, \mathcal{A}_3(\lambda h.\text{have-org-role-91}(h)\wedge \\
&\quad\quad\quad \text{ARG1}(h, \mathcal{A}_4(\lambda c.\text{city}(c)\wedge \\
&\quad\quad\quad\quad \text{name}(c, \mathcal{A}_5(\lambda n.\text{name}(n) \wedge \text{op1}(n, \text{PYONGYANG})))))\wedge \\
&\quad\quad\quad \text{ARG2}(h, \mathcal{A}_6(\lambda o.\text{official}(o)))))))\wedge \\
&\quad \text{ARG1}(d, \mathcal{A}_7(\lambda i.\text{involve-01}(i) \wedge \text{ARG1}(i, \mathcal{R}(2))))))
\end{aligned}
$$

AMR Lambda Calculus: Relations

Pyongyang officials denied their involvement

$$
\begin{aligned}
&\mathcal{A}_1(\lambda d.\text{deny-01}(d)\wedge\\
&\quad \text{ARG0}(d, \mathcal{A}_2(\lambda p.\text{person}(p)\wedge\\
&\qquad \text{ARG0-of}(p, \mathcal{A}_3(\lambda h.\text{have-org-role-91}(h)\wedge\\
&\qquad\quad \text{ARG1}(h, \mathcal{A}_4(\lambda c.\text{city}(c)\wedge\\
&\qquad\qquad \text{name}(c, \mathcal{A}_5(\lambda n.\text{name}(n) \wedge \text{op1}(n, \text{PYONGYANG})))))\wedge\\
&\qquad\quad \text{ARG2}(h, \mathcal{A}_6(\lambda o.\text{official}(o))))))\wedge\\
&\quad \text{ARG1}(d, \mathcal{A}_7(\lambda i.\text{involve-01}(i) \wedge \text{ARG1}(i, \mathcal{R}(2))))))
\end{aligned}
$$

AMR Lambda Calculus: Instances

Pyongyang officials denied their involvement

$$
\begin{aligned}
&\boxed{\mathcal{A}_1}(\lambda d.\text{deny-01}(d) \wedge \\
&\quad \text{ARG0}(d, \boxed{\mathcal{A}_2}(\lambda p.\text{person}(p) \wedge \\
&\quad\quad \text{ARG0-of}(p, \boxed{\mathcal{A}_3}(\lambda h.\text{have-org-role-91}(h) \wedge \\
&\quad\quad\quad \text{ARG1}(h, \boxed{\mathcal{A}_4}(\lambda c.\text{city}(c) \wedge \\
&\quad\quad\quad\quad \text{name}(c, \boxed{\mathcal{A}_5}(\lambda n.\text{name}(n) \wedge \text{op1}(n, \text{PYONGYANG}))))) \wedge \\
&\quad\quad\quad \text{ARG2}(h, \boxed{\mathcal{A}_6}(\lambda o.\text{official}(o)))))) \wedge \\
&\quad \text{ARG1}(d, \boxed{\mathcal{A}_7}(\lambda i.\text{involve-01}(i) \wedge \text{ARG1}(i, \mathcal{R}(2))))))
\end{aligned}
$$

$$\mathcal{A}_2$$

Skolem ID

Instance Quantifier

AMR Lambda Calculus: References

Reference Predicate

Pyongyang officials denied their involvement

$$\boxed{\mathcal{R}(2)}$$

$$
\begin{aligned}
&\boxed{\mathcal{A}_1}(\lambda d.\text{deny-01}(d) \wedge \\
&\quad \text{ARG0}(d, \boxed{\mathcal{A}_2}(\lambda p.\text{person}(p) \wedge \\
&\quad\quad \text{ARG0-of}(p, \boxed{\mathcal{A}_3}(\lambda h.\text{have-org-role-91}(h) \wedge \\
&\quad\quad\quad \text{ARG1}(h, \boxed{\mathcal{A}_4}(\lambda c.\text{city}(c) \wedge \\
&\quad\quad\quad\quad \text{name}(c, \boxed{\mathcal{A}_5}(\lambda n.\text{name}(n) \wedge \text{op1}(n, \text{PYONGYANG}))))) \wedge \\
&\quad\quad\quad \text{ARG2}(h, \boxed{\mathcal{A}_6}(\lambda o.\text{official}(o)))))) \wedge \\
&\quad \text{ARG1}(d, \boxed{\mathcal{A}_7}(\lambda i.\text{involve-01}(i) \wedge \text{ARG1}(i, \boxed{\mathcal{R}(2)})))))
\end{aligned}
$$

$$\boxed{\mathcal{A}_2}$$

Skolem ID

Instance Quantifier

Model

Pyongyang officials denied their involvement

↓ **CCG?**

$$
\begin{aligned}
&\mathcal{A}_1(\lambda d.\text{deny-01}(d)\wedge\\
&\quad \text{ARG0}(d, \mathcal{A}_2(\lambda p.\text{person}(p)\wedge\\
&\qquad \text{ARG0-of}(p, \mathcal{A}_3(\lambda h.\text{have-org-role-91}(h)\wedge\\
&\qquad\quad \text{ARG1}(h, \mathcal{A}_4(\lambda c.\text{city}(c)\wedge\\
&\qquad\qquad \text{name}(c, \mathcal{A}_5(\lambda n.\text{name}(n) \wedge \text{op1}(n, \text{PYONGYANG})))))\wedge\\
&\qquad\quad \text{ARG2}(h, \mathcal{A}_6(\lambda o.\text{official}(o))))))\wedge\\
&\quad \text{ARG1}(d, \mathcal{A}_7(\lambda i.\text{involve-01}(i) \wedge \text{ARG1}(i, \mathcal{R}(2))))))
\end{aligned}
$$

Model

Pyongyang officials denied their involvement

CCG Parse

$$
\begin{array}{l}
\mathcal{A}_1(\lambda d.\text{deny-01}(d)\wedge \\
\quad \text{ARG0}(d, \mathcal{A}_2(\lambda p.\text{person}(p)\wedge \\
\quad\quad \text{REL-of}(p, \mathcal{A}_3(\lambda h.\text{have-org-role-91}(h)\wedge \\
\quad\quad\quad \text{ARG1}(h, \mathcal{A}_4(\lambda c.\text{city}(c)\wedge \\
\quad\quad\quad\quad \text{name}(c, \mathcal{A}_5(\lambda n.\text{name}(n) \wedge \text{op1}(n, \text{PYONGYANG})))))\wedge \\
\quad\quad\quad \text{REL}(h, \mathcal{A}_6(\lambda o.\text{official}(o))))))\wedge \\
\quad \text{ARG1}(d, \mathcal{A}_7(\lambda i.\text{involve-01}(i) \wedge \text{ARG1}(i, \mathcal{R}(\text{ID})))))
\end{array}
$$

Underspecified Logical Form

Constant Mapping

$$
\begin{array}{l}
\mathcal{A}_1(\lambda d.\text{deny-01}(d)\wedge \\
\quad \text{ARG0}(d, \mathcal{A}_2(\lambda p.\text{person}(p)\wedge \\
\quad\quad \text{ARG0-of}(p, \mathcal{A}_3(\lambda h.\text{have-org-role-91}(h)\wedge \\
\quad\quad\quad \text{ARG1}(h, \mathcal{A}_4(\lambda c.\text{city}(c)\wedge \\
\quad\quad\quad\quad \text{name}(c, \mathcal{A}_5(\lambda n.\text{name}(n) \wedge \text{op1}(n, \text{PYONGYANG})))))\wedge \\
\quad\quad\quad \text{ARG2}(h, \mathcal{A}_6(\lambda o.\text{official}(o))))))\wedge \\
\quad \text{ARG1}(d, \mathcal{A}_7(\lambda i.\text{involve-01}(i) \wedge \text{ARG1}(i, \mathcal{R}(2)))))
\end{array}
$$

Model

Pyongyang officials denied their involvement

$$
\begin{aligned}
&\mathcal{A}_1(\lambda d.\text{deny-01}(d)\wedge\\
&\quad \text{ARG0}(d, \mathcal{A}_2(\lambda p.\text{person}(p)\wedge\\
&\quad\quad \boxed{\text{REL-of}}(p, \mathcal{A}_3(\lambda h.\text{have-org-role-91}(h)\wedge\\
&\quad\quad\quad \text{ARG1}(h, \mathcal{A}_4(\lambda c.\text{city}(c)\wedge\\
&\quad\quad\quad\quad \text{name}(c, \mathcal{A}_5(\lambda n.\text{name}(n) \wedge \text{op1}(n, \text{PYONGYANG})))))\wedge\\
&\quad\quad\quad \boxed{\text{REL}}(h, \mathcal{A}_6(\lambda o.\text{official}(o))))))\wedge\\
&\quad \text{ARG1}(d, \mathcal{A}_7(\lambda i.\text{involve-01}(i) \wedge \text{ARG1}(i, \boxed{\mathcal{R}(\text{ID})}))))))
\end{aligned}
$$

$\boxed{\text{REL-of}}$ **Passive relation placeholder**

Underspecified Logical Form

$\boxed{\mathcal{R}(\text{ID})}$ **Reference placeholder**

$\boxed{\text{REL}}$ **Active relation placeholder**

$$
\begin{aligned}
&\mathcal{A}_1(\lambda d.\text{deny-01}(d)\wedge\\
&\quad \text{ARG0}(d, \mathcal{A}_2(\lambda p.\text{person}(p)\wedge\\
&\quad\quad \text{ARG0-of}(p, \mathcal{A}_3(\lambda h.\text{have-org-role-91}(h)\wedge\\
&\quad\quad\quad \text{ARG1}(h, \mathcal{A}_4(\lambda c.\text{city}(c)\wedge\\
&\quad\quad\quad\quad \text{name}(c, \mathcal{A}_5(\lambda n.\text{name}(n) \wedge \text{op1}(n, \text{PYONGYANG})))))\wedge\\
&\quad\quad\quad \text{ARG2}(h, \mathcal{A}_6(\lambda o.\text{official}(o))))))\wedge\\
&\quad \text{ARG1}(d, \mathcal{A}_7(\lambda i.\text{involve-01}(i) \wedge \text{ARG1}(i, \mathcal{R}(2)))))))
\end{aligned}
$$

Model

Pyongyang officials denied their involvement

$$
\begin{aligned}
&\mathcal{A}_1(\lambda d.\text{deny-01}(d)\wedge\\
&\quad \text{ARG0}(d, \mathcal{A}_2(\lambda p.\text{person}(p)\wedge\\
&\quad\quad \boxed{\text{REL-of}}(p, \mathcal{A}_3(\lambda h.\text{have-org-role-91}(h)\wedge\\
&\quad\quad\quad \text{ARG1}(h, \mathcal{A}_4(\lambda c.\text{city}(c)\wedge\\
&\quad\quad\quad\quad \text{name}(c, \mathcal{A}_5(\lambda n.\text{name}(n) \wedge \text{op1}(n, \text{PYONGYANG})))))\wedge\\
&\quad\quad\quad \boxed{\text{REL}}(h, \mathcal{A}_6(\lambda o.\text{official}(o))))))\wedge\\
&\quad \text{ARG1}(d, \mathcal{A}_7(\lambda i.\text{involve-01}(i) \wedge \text{ARG1}(i, \mathcal{R}(\text{ID})))))
\end{aligned}
$$

REL-of → ARG0-of ⋮ cause-of

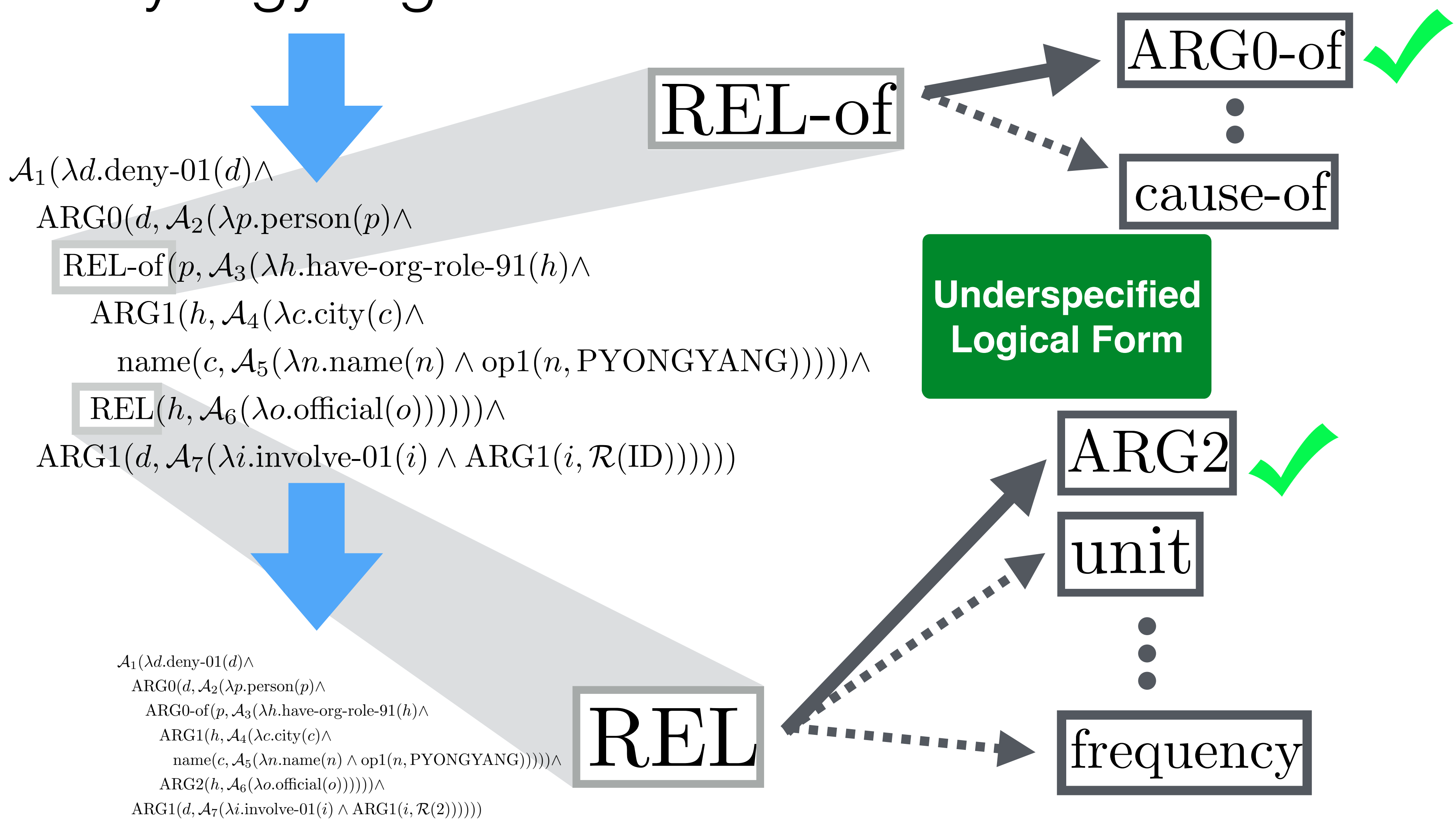
Underspecified Logical Form

REL → ARG2, unit ⋮ frequency

$$
\begin{aligned}
&\mathcal{A}_1(\lambda d.\text{deny-01}(d)\wedge\\
&\quad \text{ARG0}(d, \mathcal{A}_2(\lambda p.\text{person}(p)\wedge\\
&\quad\quad \text{ARG0-of}(p, \mathcal{A}_3(\lambda h.\text{have-org-role-91}(h)\wedge\\
&\quad\quad\quad \text{ARG1}(h, \mathcal{A}_4(\lambda c.\text{city}(c)\wedge\\
&\quad\quad\quad\quad \text{name}(c, \mathcal{A}_5(\lambda n.\text{name}(n) \wedge \text{op1}(n, \text{PYONGYANG})))))\wedge\\
&\quad\quad\quad \text{ARG2}(h, \mathcal{A}_6(\lambda o.\text{official}(o))))))\wedge\\
&\quad \text{ARG1}(d, \mathcal{A}_7(\lambda i.\text{involve-01}(i) \wedge \text{ARG1}(i, \mathcal{R}(2)))))
\end{aligned}
$$

Model

Pyongyang officials denied their involvement



Model

Pyongyang officials denied their involvement

$$
\begin{aligned}
&\boxed{\mathcal{A}_1}(\lambda d.\text{deny-01}(d)\wedge\\
&\quad \text{ARG0}(d, \boxed{\mathcal{A}_2}(\lambda p.\text{person}(p)\wedge\\
&\qquad \text{REL-of}(p, \mathcal{A}_3(\lambda h.\text{have-org-role-91}(h)\wedge\\
&\qquad\quad \text{ARG1}(h, \mathcal{A}_4(\lambda c.\text{city}(c)\wedge\\
&\qquad\qquad \text{name}(c, \mathcal{A}_5(\lambda n.\text{name}(n) \wedge \text{op1}(n, \text{PYONGYANG})))))\wedge\\
&\qquad\quad \text{REL}(h, \mathcal{A}_6(\lambda o.\text{official}(o))))))\wedge\\
&\quad \text{ARG1}(d, \boxed{\mathcal{A}_7}(\lambda i.\text{involve-01}(i) \wedge \text{ARG1}(i, \boxed{\mathcal{R}(\text{ID})})))))
\end{aligned}
$$

Underspecified Logical Form

$$\boxed{\mathcal{R}(\text{ID})} \rightarrow \boxed{\mathcal{R}(1)},\ \boxed{\mathcal{R}(2)},\ \vdots,\ \boxed{\mathcal{R}(7)}$$

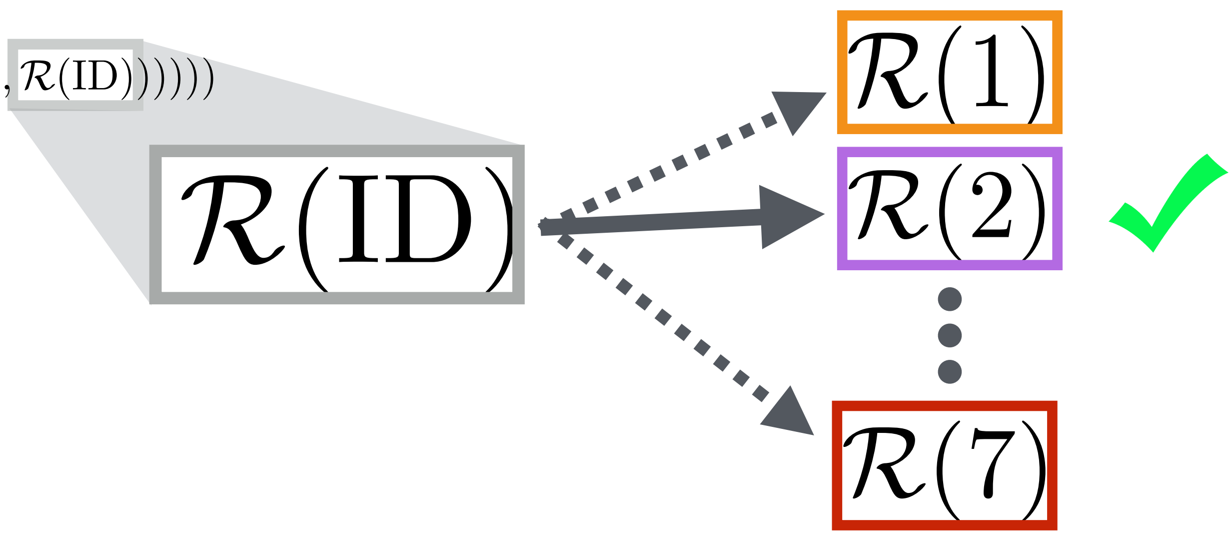
$$
\begin{aligned}
&\mathcal{A}_1(\lambda d.\text{deny-01}(d)\wedge\\
&\quad \text{ARG0}(d, \mathcal{A}_2(\lambda p.\text{person}(p)\wedge\\
&\qquad \text{ARG0-of}(p, \mathcal{A}_3(\lambda h.\text{have-org-role-91}(h)\wedge\\
&\qquad\quad \text{ARG1}(h, \mathcal{A}_4(\lambda c.\text{city}(c)\wedge\\
&\qquad\qquad \text{name}(c, \mathcal{A}_5(\lambda n.\text{name}(n) \wedge \text{op1}(n, \text{PYONGYANG})))))\wedge\\
&\qquad\quad \text{ARG2}(h, \mathcal{A}_6(\lambda o.\text{official}(o))))))\wedge\\
&\quad \text{ARG1}(d, \mathcal{A}_7(\lambda i.\text{involve-01}(i) \wedge \text{ARG1}(i, \mathcal{R}(2))))))
\end{aligned}
$$

Model

Pyongyang officials denied their involvement

$$
\begin{aligned}
&\boxed{\mathcal{A}_1}(\lambda d.\text{deny-01}(d)\wedge\\
&\quad \text{ARG0}(d, \boxed{\mathcal{A}_2}(\lambda p.\text{person}(p)\wedge\\
&\qquad \text{REL-of}(p, \mathcal{A}_3(\lambda h.\text{have-org-role-91}(h)\wedge\\
&\qquad\quad \text{ARG1}(h, \mathcal{A}_4(\lambda c.\text{city}(c)\wedge\\
&\qquad\qquad \text{name}(c, \mathcal{A}_5(\lambda n.\text{name}(n) \wedge \text{op1}(n, \text{PYONGYANG})))))\wedge\\
&\qquad\quad \text{REL}(h, \mathcal{A}_6(\lambda o.\text{official}(o))))))\wedge\\
&\quad \text{ARG1}(d, \boxed{\mathcal{A}_7}(\lambda i.\text{involve-01}(i) \wedge \text{ARG1}(i, \boxed{\mathcal{R}(\text{ID})})))))
\end{aligned}
$$

Underspecified Logical Form



$$
\begin{aligned}
&\mathcal{A}_1(\lambda d.\text{deny-01}(d)\wedge\\
&\quad \text{ARG0}(d, \mathcal{A}_2(\lambda p.\text{person}(p)\wedge\\
&\qquad \text{ARG0-of}(p, \mathcal{A}_3(\lambda h.\text{have-org-role-91}(h)\wedge\\
&\qquad\quad \text{ARG1}(h, \mathcal{A}_4(\lambda c.\text{city}(c)\wedge\\
&\qquad\qquad \text{name}(c, \mathcal{A}_5(\lambda n.\text{name}(n) \wedge \text{op1}(n, \text{PYONGYANG})))))\wedge\\
&\qquad\quad \text{ARG2}(h, \mathcal{A}_6(\lambda o.\text{official}(o))))))\wedge\\
&\quad \text{ARG1}(d, \mathcal{A}_7(\lambda i.\text{involve-01}(i) \wedge \text{ARG1}(i, \mathcal{R}(2))))))
\end{aligned}
$$

Model Advantages

- Reason about non-compositional distant references, including:
    - Co-reference
    - Control structures (often compositional, but not distinguished)
- Defer certain compositional decisions from the difficult CCG parsing problem

Derivation

CCG Parse

Pyongyang	*officials*	*denied*	*their*	*involvement*
$NP_{[sg]}$	$N_{[pl]}\backslash(N_{[pl]}/N_{[pl]})$	$S\backslash NP/NP$	$NP_{[pl]}$	$N_{[nb]}$
$\mathcal{A}_1(\lambda c.\text{city}(c)\wedge \text{name}(c, \mathcal{A}_2(\lambda n.\text{name}(n)\wedge \text{op}(n, \text{PYONGYANG}))))$	$\lambda f.\lambda p.\text{person}(p)\wedge \text{REL-of}(p, \mathcal{A}_3(f(\lambda h.\text{have-org-role-91}(h)\wedge \text{REL}(h, \mathcal{A}_4(\lambda o.\text{official}(o))))))$	$\lambda x.\lambda y.\lambda d.\text{deny-01}(d)\wedge \text{ARG0}(d, y)\wedge \text{ARG1}(d, x)$	$\mathcal{R}(\text{ID})$	$\lambda i.\text{involve-01}(i)$

$<$ $>$ $>$ $<$

A

$$\mathcal{A}_1(\lambda d.\text{deny-01}(d) \wedge \text{ARG0}(d, \mathcal{A}_2(\lambda p.\text{person}(p) \wedge \text{REL-of}(p, \mathcal{A}_3(\lambda h.\text{have-org-role-91}(h)\wedge \text{ARG1}(h, \mathcal{A}_4(\lambda c.\text{city}(c) \wedge \text{name}(c, \mathcal{A}_5(\lambda n.\text{name}(n) \wedge \text{op}(n, \text{PYONGYANG})))))\wedge \text{REL}(h, \mathcal{A}_6(\lambda o.\text{official}(o)))))) \wedge \text{ARG1}(d, \mathcal{A}_7(\lambda i.\text{involve-01}(i) \wedge \text{ARG1}(i, \mathcal{R}(\text{ID})))))$$

$$\mathcal{A}_1(\lambda d.\text{deny-01}(d) \wedge \text{ARG0}(d, \mathcal{A}_2(\lambda p.\text{person}(p) \wedge \text{ARG0-of}(p, \mathcal{A}_3(\lambda h.\text{have-org-role-91}(h)\wedge \text{ARG1}(h, \mathcal{A}_4(\lambda c.\text{city}(c) \wedge \text{name}(c, \mathcal{A}_5(\lambda n.\text{name}(n) \wedge \text{op}(n, \text{PYONGYANG})))))\wedge \text{ARG2}(h, \mathcal{A}_6(\lambda o.\text{official}(o)))))) \wedge \text{ARG1}(d, \mathcal{A}_7(\lambda i.\text{involve-01}(i) \wedge \text{ARG1}(i, \mathcal{R}(2)))))$$

Constant Mapping

Log-linear Model

$\omega \in \mathbb{R}^m$	Λ	$\phi : \mathcal{X} \times \mathcal{D} \rightarrow \mathbb{R}^m$
Weights	CCG Lexicon	Feature Function

- Given a sentence $x \in \mathcal{X}$:
  - The probability of a logical form z is:

$$p(z \mid x; \theta, \Lambda) = \sum_{d \in \mathcal{D}(z)} p(d \mid x; \theta, \Lambda)$$

  - The probability of a derivation $d \in \mathcal{D}$ is:

$$p(d \mid x; \theta, \Lambda) = \frac{e^{\theta \cdot \phi(x,d)}}{\sum_{d' \in \mathcal{D}} e^{\theta \cdot \phi(x,d')}}$$

Inference

Joint scoring

CCG Parse

CKY parsing

Pyongyang

$NP_{[sg]}$

$\mathcal{A}_1(\lambda c.\text{city}(c) \wedge$
$\text{name}(c, \mathcal{A}_2(\lambda n.\text{name}(n) \wedge$
$\text{op}(n, \text{PYONGYANG}))))$

officials

$N_{[pl]} \backslash (N_{[pl]}/N_{[pl]})$

$\lambda f.\lambda p.\text{person}(p) \wedge$
$\text{REL-of}(p, \mathcal{A}_3(f(\lambda h.\text{have-org-role-91}(h) \wedge$
$\text{REL}(h, \mathcal{A}_4(\lambda o.\text{official}(o))))))$

<

A

$\mathcal{A}_1(\lambda d.\text{deny-01}(d) \wedge \text{ARG0}(d, \mathcal{A}_2(\lambda p.\text{person}(p) \wedge$ RE

$\text{ARG1}(h, \mathcal{A}_4(\lambda c.\text{city}(c) \wedge \text{name}(c, \mathcal{A}_5(\lambda n.\text{name}($

$\text{REL}(h, \mathcal{A}_6(\lambda o.\text{official}(o)))))) \wedge \text{ARG1}(d, \mathcal{A}_7(\lambda i.\text{in}$

Factor graph

$\mathcal{A}_1(\lambda d.\text{deny-01}(d) \wedge \text{ARG0}(d, \mathcal{A}_2(\lambda p.\text{person}(p) \wedge$ AR

$\text{ARG1}(h, \mathcal{A}_4(\lambda c.\text{city}(c) \wedge \text{name}(c, \mathcal{A}_5(\lambda n.\text{name}(n$

$\text{ARG2}(h, \mathcal{A}_6(\lambda o.\text{official}(o)))))) \wedge \text{ARG1}(d, \mathcal{A}_7(\lambda i.\text{i}$

Consta

Constant Mapping with a Factor Graph

Build a factor graph for each underspecified logical form

$$
\begin{aligned}
&\mathcal{A}_1(\lambda d.\text{deny-01}(d) \wedge \\
&\qquad \text{ARG0}(d, \mathcal{A}_2(\lambda p.\text{person}(p) \wedge \\
&\qquad\qquad \text{REL-of}(p, \mathcal{A}_3(\lambda h.\text{have-org-role-91}(h) \wedge \\
&\qquad\qquad\qquad \text{ARG1}(h, \mathcal{A}_4(\lambda c.\text{city}(c) \wedge \\
&\qquad\qquad\qquad\qquad \text{name}(c, \mathcal{A}_5(\lambda n.\text{name}(n) \wedge \text{op}(n, \text{PYONGYANG}))))) \wedge \\
&\qquad\qquad\qquad \text{REL}(h, \mathcal{A}_6(\lambda o.\text{official}(o)))))) \wedge \\
&\qquad \text{ARG1}(d, \mathcal{A}_7(\lambda i.\text{involve-01}(i) \wedge \\
&\qquad\qquad \text{ARG1}(i, \mathcal{R}(\text{ID}))))))
\end{aligned}
$$

Factor Graph

Each constant is a random variable

$$
\begin{aligned}
&\mathcal{A}_1(\lambda d.\text{deny-01}(d) \wedge \\
&\quad \text{ARG0}(d, \mathcal{A}_2(\lambda p.\text{person}(p) \wedge \\
&\quad\quad \text{REL-of}(p, \mathcal{A}_3(\lambda h.\text{have-org-role-91}(h) \wedge \\
&\quad\quad\quad \text{ARG1}(h, \mathcal{A}_4(\lambda c.\text{city}(c) \wedge \\
&\quad\quad\quad\quad \text{name}(c, \mathcal{A}_5(\lambda n.\text{name}(n) \wedge \text{op}(n, \text{PYONGYANG}))))) \wedge \\
&\quad\quad\quad \text{REL}(h, \mathcal{A}_6(\lambda o.\text{official}(o))))))) \wedge \\
&\quad \text{ARG1}(d, \mathcal{A}_7(\lambda i.\text{involve-01}(i) \wedge \\
&\quad\quad \text{ARG1}(i, \mathcal{R}(\text{ID}))))))
\end{aligned}
$$

Factor Graph

Potential mapping of placeholders defines assignments

unit-of, prep-with-of, frequency-of,
prep-against-of, compared-to-of,
employed-by-of, ARG0-of, . . .

$$
\begin{aligned}
&\mathcal{A}_1(\lambda d.\text{deny-01}(d) \wedge \\
&\quad \text{ARG0}(d, \mathcal{A}_2(\lambda p.\text{person}(p) \wedge \\
&\qquad \text{REL-of}(p, \mathcal{A}_3(\lambda h.\text{have-org-role-91}(h) \wedge \\
&\qquad\quad \text{ARG1}(h, \mathcal{A}_4(\lambda c.\text{city}(c) \wedge \\
&\qquad\qquad \text{name}(c, \mathcal{A}_5(\lambda n.\text{name}(n) \wedge \text{op}(n, \text{PYONGYANG}))))) \wedge \\
&\qquad\quad \text{REL}(h, \mathcal{A}_6(\lambda o.\text{official}(o)))))) \wedge \\
&\quad \text{ARG1}(d, \mathcal{A}_7(\lambda i.\text{involve-01}(i) \wedge \\
&\qquad \text{ARG1}(i, \mathcal{R}(\text{ID}))))))
\end{aligned}
$$

unit, prep-with, frequency,
prep-against, compared-to,
employed-by, ARG2, . . .

$1, 2, 3, 4, 5, 6, 7$

Factor Graph

Features define factors to resolve underspecified constants

$$
\begin{aligned}
&\mathcal{A}_1(\lambda d.\text{deny-01}(d)\wedge\\
&\quad \text{ARG0}(d, \mathcal{A}_2(\lambda p.\text{person}(p)\wedge\\
&\qquad \text{REL-of}(p, \mathcal{A}_3(\lambda h.\text{have-org-role-91}(h)\wedge\\
&\qquad\quad \text{ARG1}(h, \mathcal{A}_4(\lambda c.\text{city}(c)\wedge\\
&\qquad\qquad \text{name}(c, \mathcal{A}_5(\lambda n.\text{name}(n) \wedge \text{op}(n, \text{PYONGYANG})))))\wedge\\
&\qquad\quad \text{REL}(h, \mathcal{A}_6(\lambda o.\text{official}(o))))))\wedge\\
&\quad \text{ARG1}(d, \mathcal{A}_7(\lambda i.\text{involve-01}(i)\wedge\\
&\qquad \text{ARG1}(i, \mathcal{R}(\text{ID})))))))
\end{aligned}
$$

A

Selectional preference features to specify REL to one of 67 active relations

B

Features for resolving ID to 3

Approach

- Model:
  - Two-stage model for compositional semantics and non-compositional distant references
- Learning:
  - Lexicon induction
  - Parameter estimation

Learning Algorithm Sketch

For T iterations:

Lexicon Induction

- For each training sample:
  - Two-pass generation of new lexical entries
- Update the model lexicon

Parameter Estimation

- For each mini-batch of size M
  - Compute gradient with early updates
  - Apply update with AdaGrad

Learning Algorithm Sketch

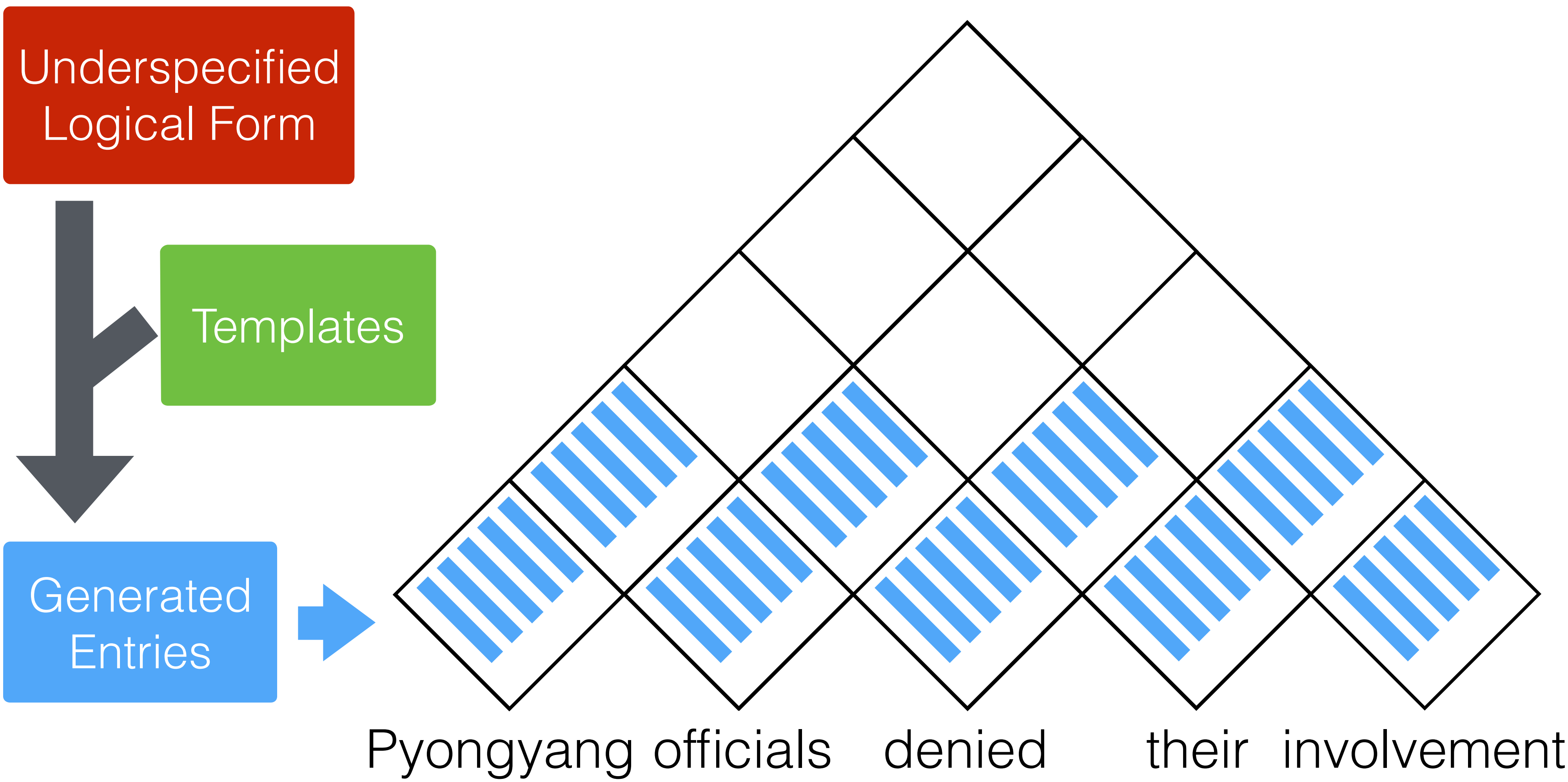
For T iterations:

- For each training sample:
  - Two-pass generation of new lexical entries
- Update the model lexicon
- For each mini-batch of size M
  - Compute gradient with early updates
  - Apply update with AdaGrad

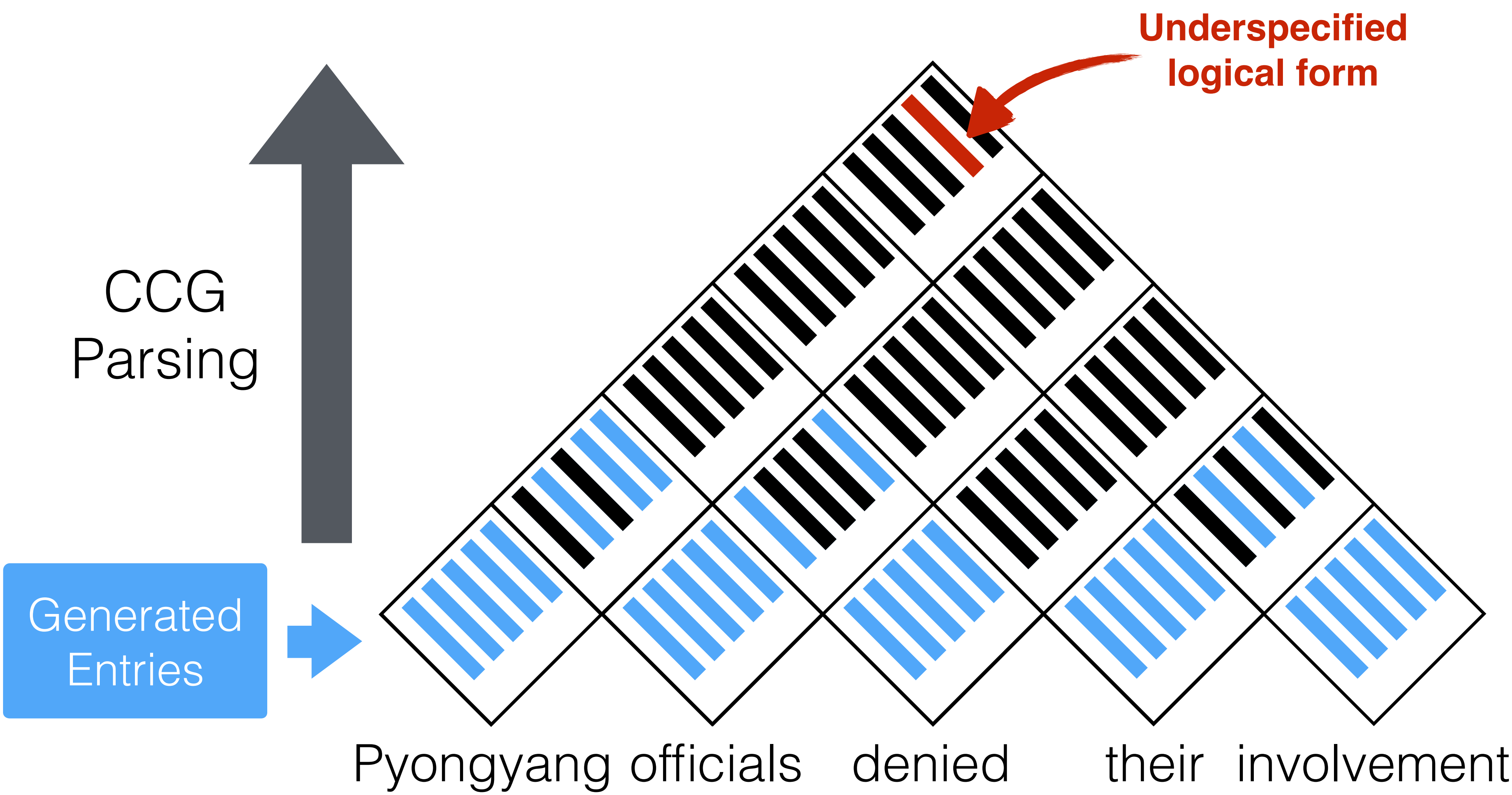
Two-pass Lexical Generation

- Bottom-up: over-generate new entries and parse
- Top-down: recursive splitting to complete partial derivations

Bottom-up Pass

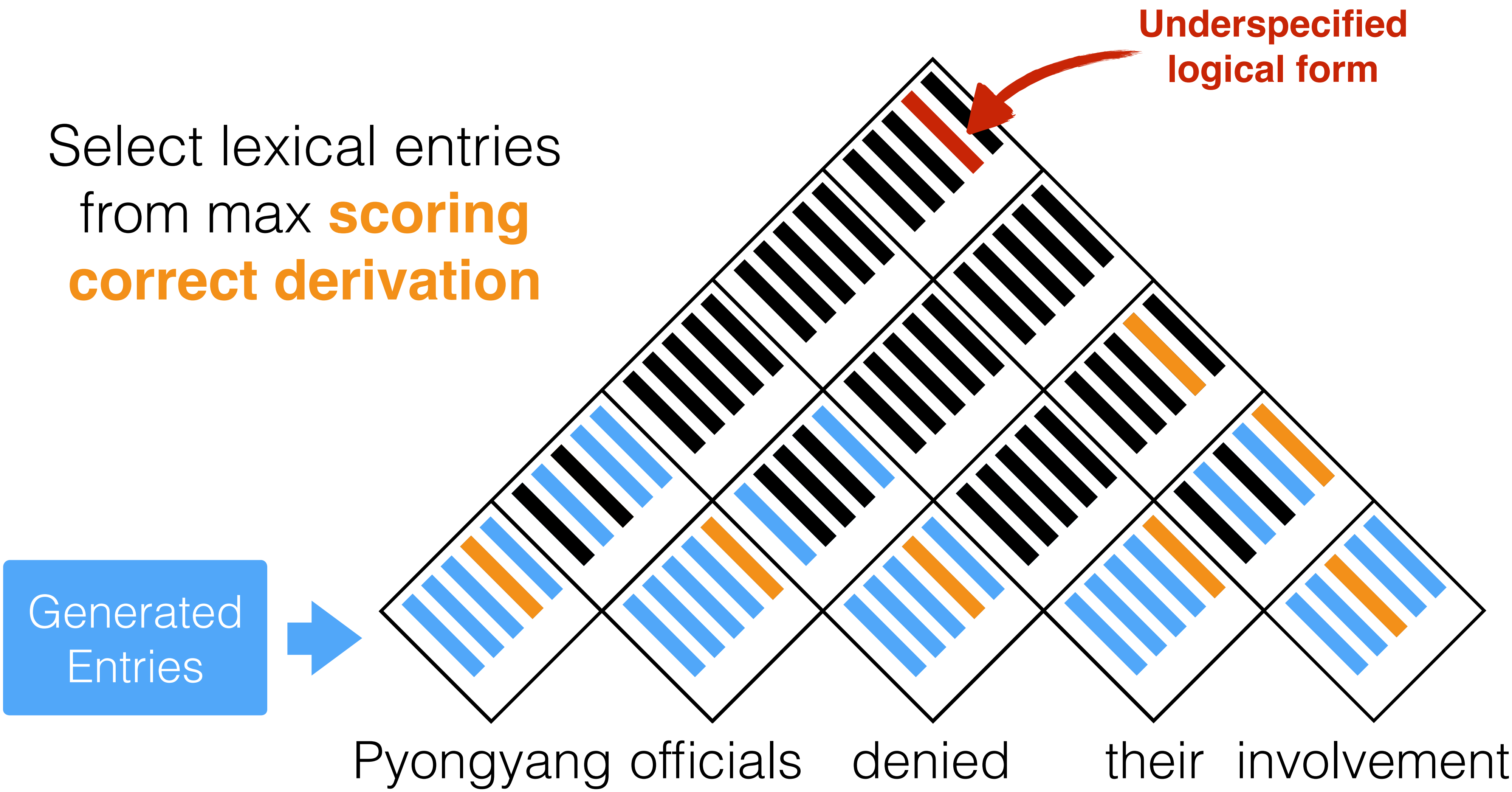


Bottom-up Pass



Bottom-up Pass

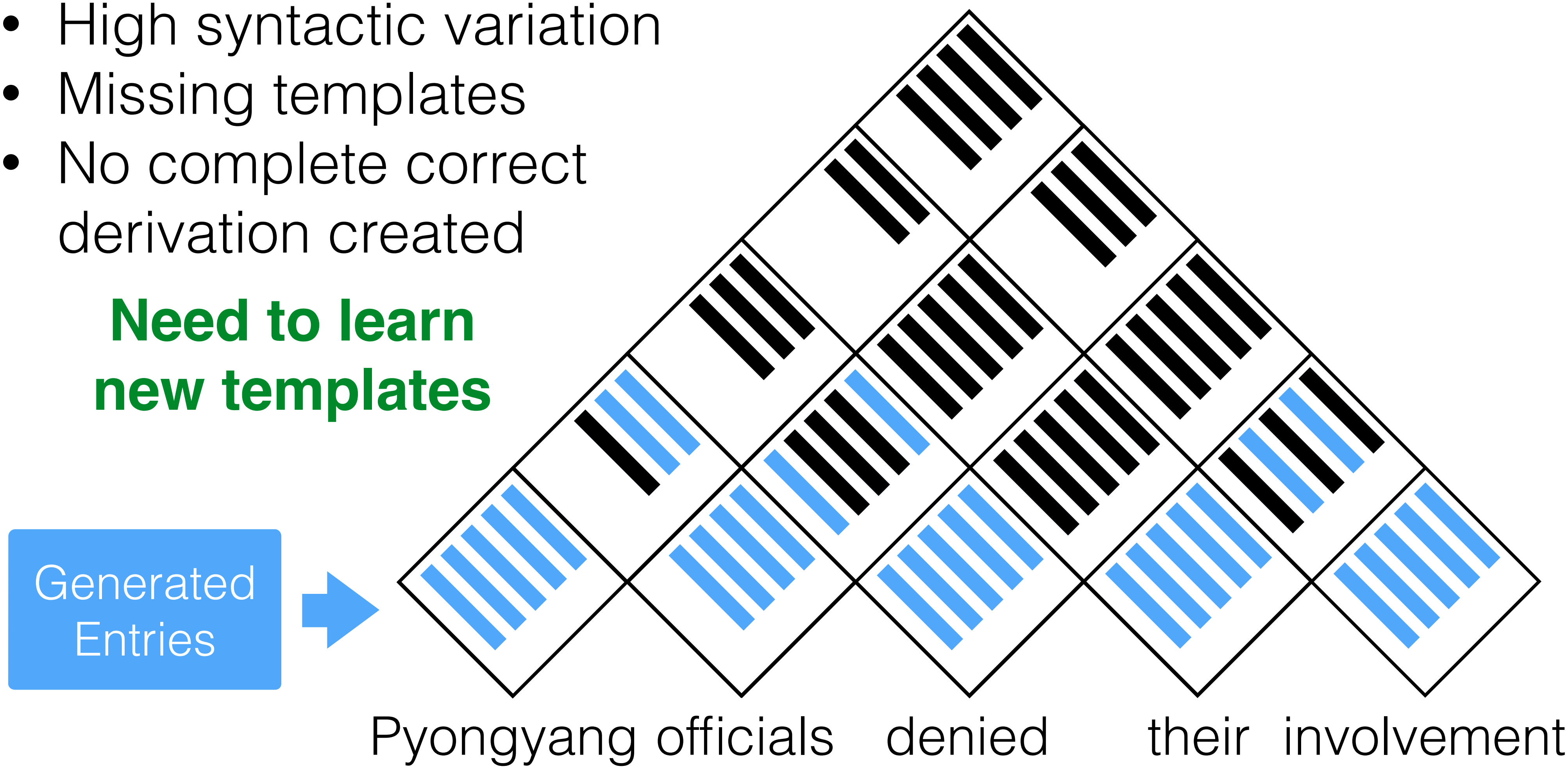
Select lexical entries from max **scoring correct derivation**



Common Failure

- High syntactic variation
- Missing templates
- No complete correct derivation created

Need to learn new templates



Splitting CCG Categories

- Introduced by Kwiatkowski et al. 2010
- Approximately reverses CCG parsing operations
- Explore new syntactic structures, learn new templates

Splitting CCG Categories

Given a CCG category $C : h$:

1. Split logical form h to f and g s.t.:

$$f(g) = h \text{ or } \lambda x. f(g(x)) = h$$

$$NP_{[nb]} : \lambda i.\text{involve-01}(i) \wedge \text{ARG1}(i, \mathcal{R}(\text{ID}))$$

$\lambda f.\lambda i.f(i) \wedge \text{ARG1}(i, \mathcal{R}(\text{ID}))$

$\lambda i.\text{involve-01}(i)$

$\mathcal{R}(\text{ID})$

$\lambda x.\lambda i.\text{involve-01}(i) \wedge \text{ARG1}(i, x)$

Splitting CCG Categories

Given a CCG category $C : h$:

1. Split logical form h to f and g s.t.:

$$f(g) = h \text{ or } \lambda x.f(g(x)) = h$$

2. Infer syntax from logical form type

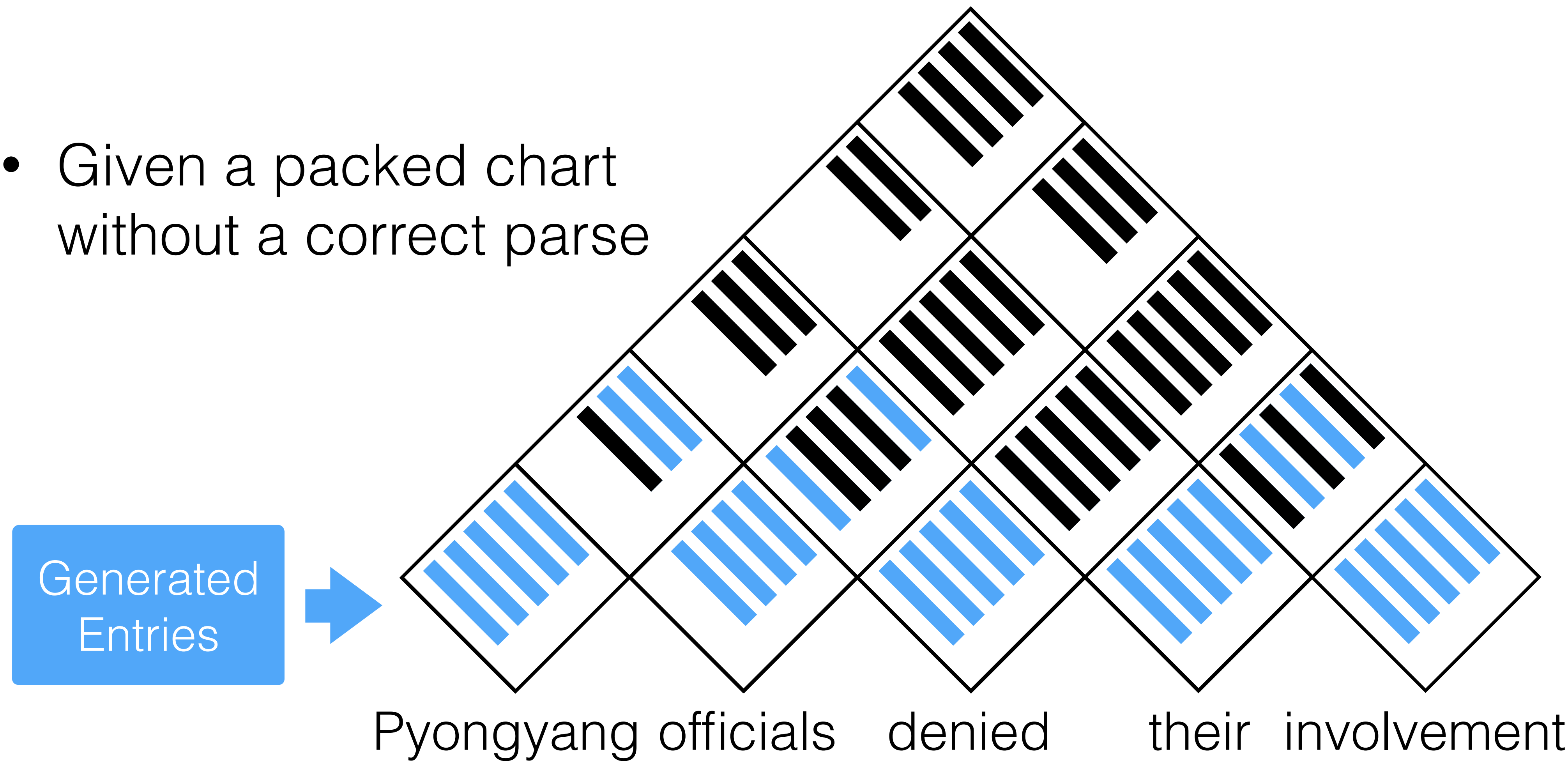
$$NP_{[nb]} : \lambda i.\text{involve-01}(i) \wedge \text{ARG1}(i, \mathcal{R}(\text{ID}))$$

$$NP_{[x]}/N_{[x]} : \lambda f.\lambda i.f(i) \wedge \text{ARG1}(i, \mathcal{R}(\text{ID}))$$
$$N_{[nb]} : \lambda i.\text{involve-01}(i)$$

$$NP_{[pl]} : \mathcal{R}(\text{ID})$$
$$NP_{[nb]}\backslash NP : \lambda x.\lambda i.\text{involve-01}(i) \wedge \text{ARG1}(i, x)$$

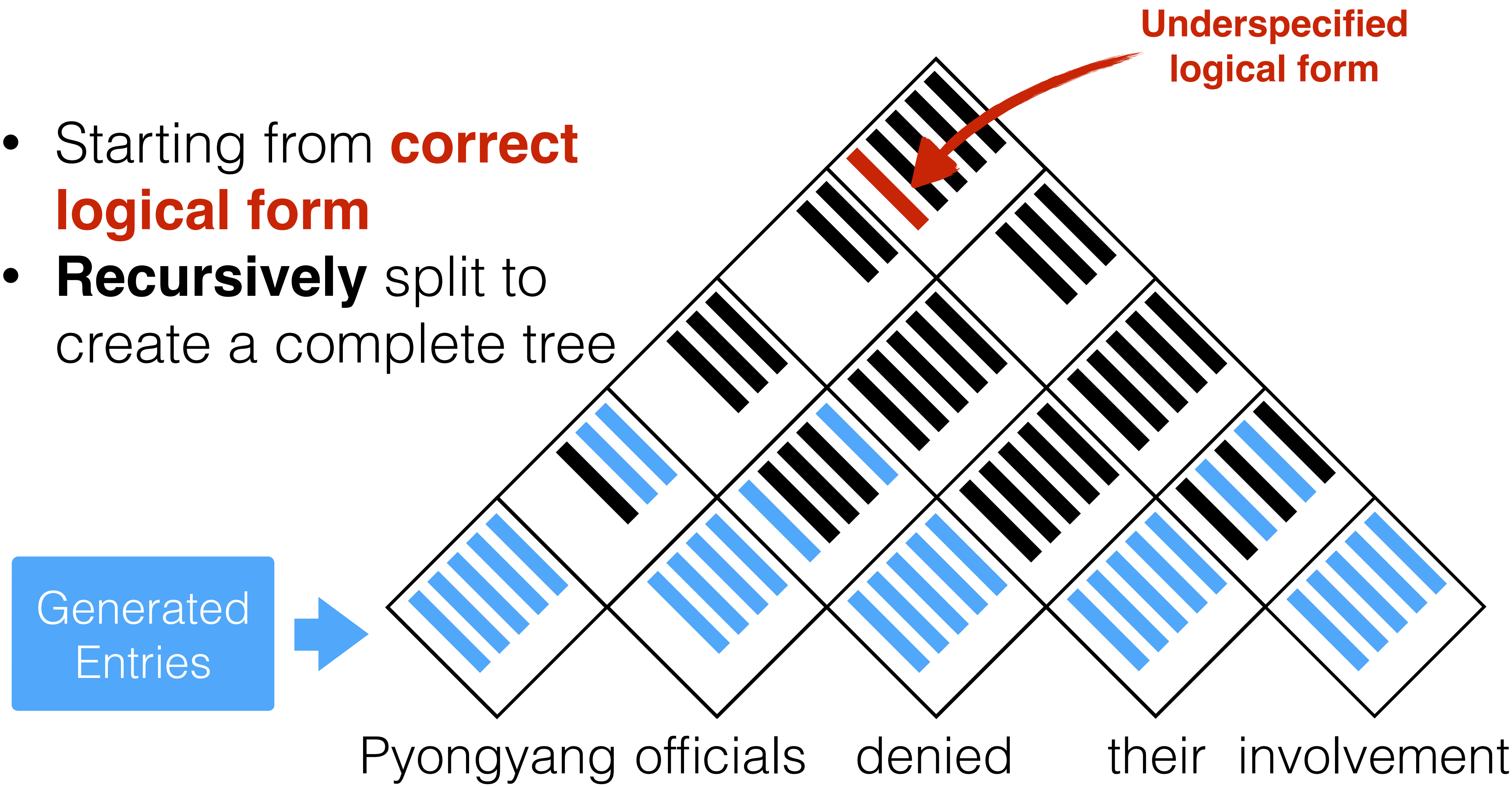
Top-down Pass

- Given a packed chart without a correct parse



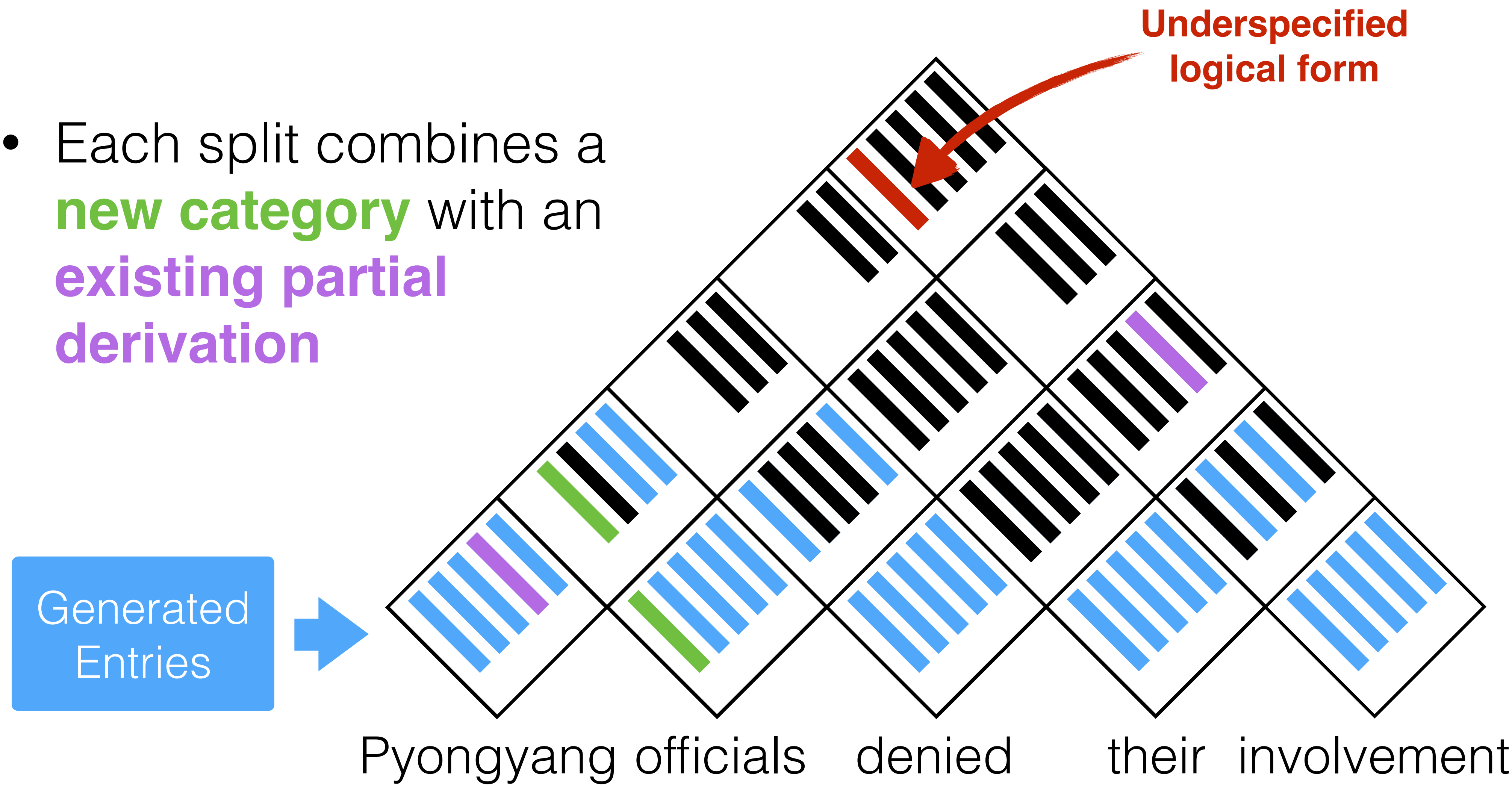
Top-down Pass

- Starting from **correct logical form**
- **Recursively** split to create a complete tree



Top-down Pass

- Each split combines a **new category** with an **existing partial derivation**



Splitting for CCG Induction

- Kwiatkowski et al. 2010:
  - No restriction on result categories
  - Applied up to depth one
- Our approach:
  - Combined with bottom-up template approach
  - Must connect to an existing partial derivation
  - Applied recursively

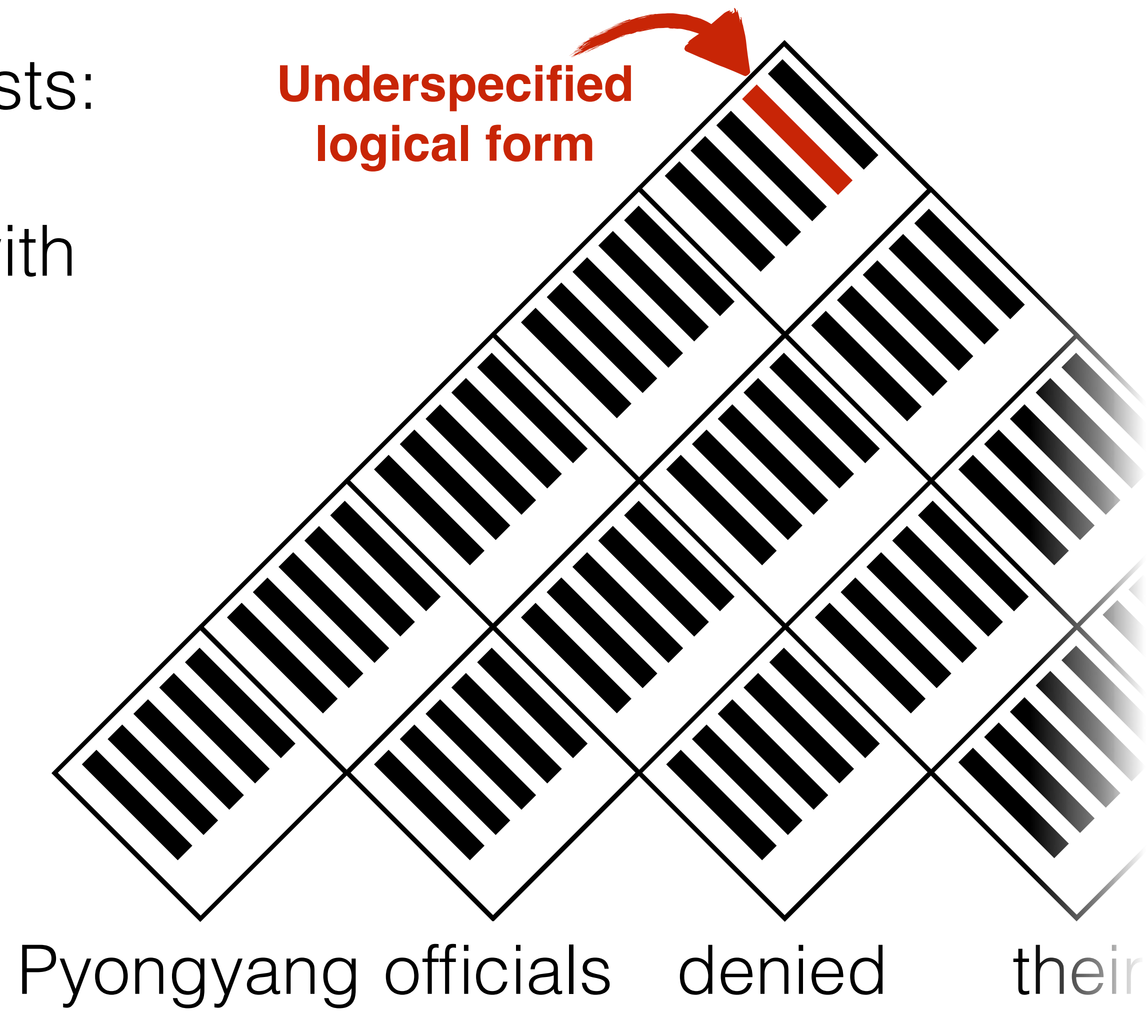
Learning Algorithm Sketch

For T iterations:

- For each training sample:
  - Two-pass generation of new lexical entries
- Update the model lexicon
- For each mini-batch of size M
  - Compute gradient with early updates
  - Apply update with AdaGrad

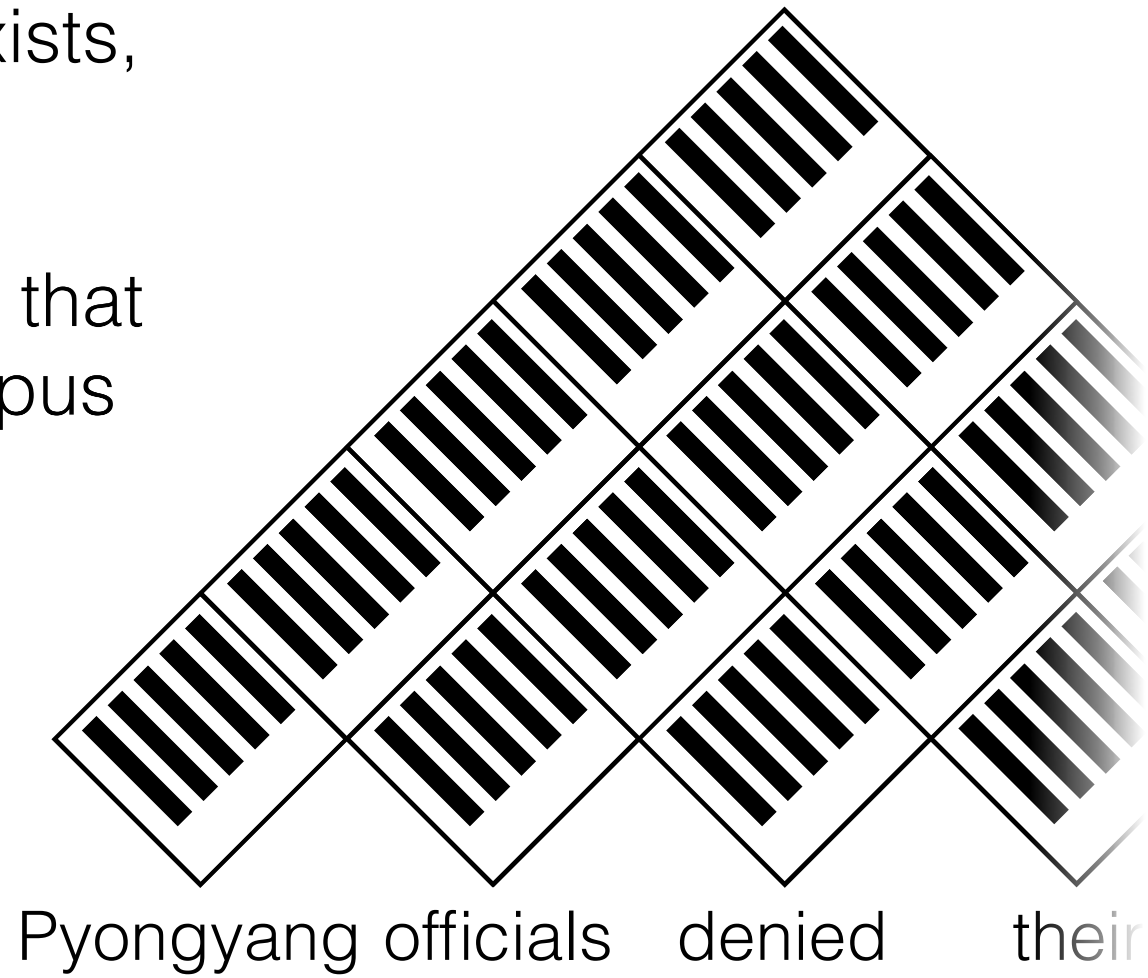
Gradient Computation

- If a correct derivation exists:
  - Compute gradient with inside-outside
  - Re-normalize with constant mapping features



Common Failure

- No correct derivation exists, ~40% of training data
- Previous work assumed that all (or at least most) corpus can be parsed
- Instead: early updates

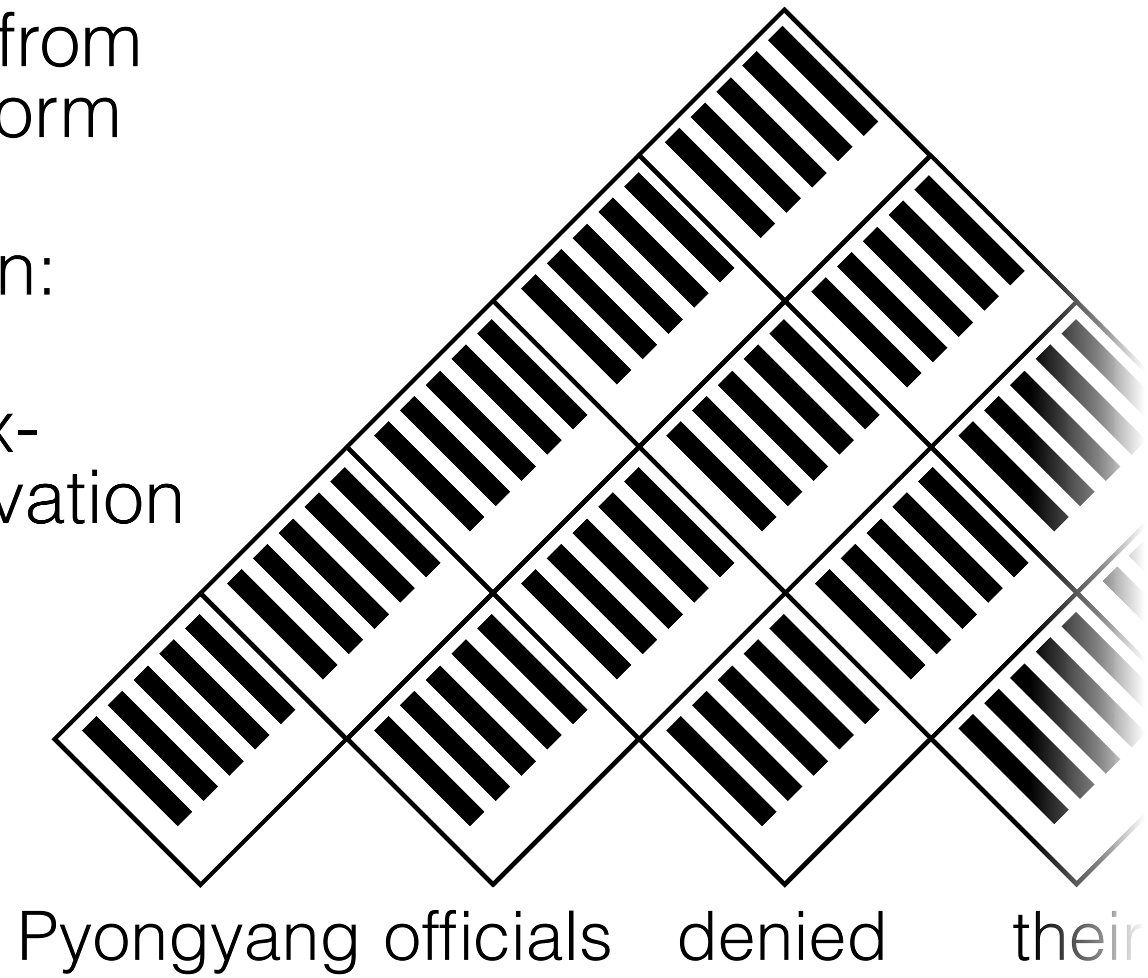


Early Updates

- Collins and Roark (2004):
  - Given fully labeled parse trees
  - Update with partial derivations
- Challenge: derivation is latent

Early Update with Latent Structures

- Extract sub-expression from underspecified logical form
- For each sub-expression:
    - Identify largest max-scoring partial derivation
    - Compute gradient



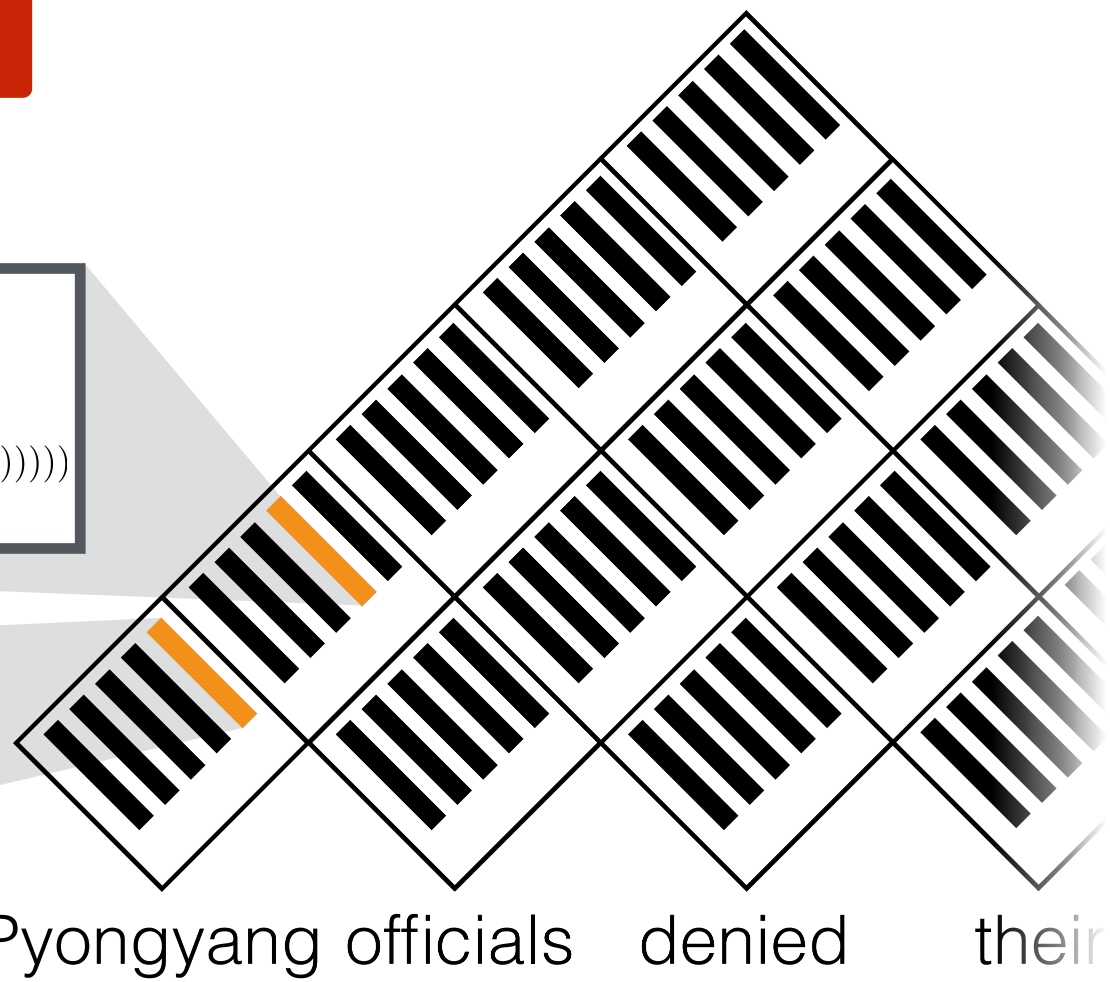
Early Update with Latent Structures

Underspecified Logical Form

$$
\begin{aligned}
&\mathcal{A}_2(\lambda p.\text{person}(p)\wedge\\
&\quad \text{REL-of}(p, \mathcal{A}_3(\lambda h.\text{have-org-role-91}(h)\wedge\\
&\qquad \text{ARG1}(h, \mathcal{A}_4(\lambda c.\text{city}(c)\wedge\\
&\qquad\quad \text{name}(c, \mathcal{A}_5(\lambda n.\text{name}(n) \wedge \text{op}(n, \text{PYONGYANG})))))\\
&\qquad \text{REL}(h, \mathcal{A}_6(\lambda o.\text{official}(o))))))
\end{aligned}
$$

$$
\begin{aligned}
&\mathcal{A}_4(\lambda c.\text{city}(c)\wedge\\
&\quad \text{name}(c, \mathcal{A}_5(\lambda n.\text{name}(n) \wedge \text{op}(n, \text{PYONGYANG}))))
\end{aligned}
$$

$\vdots$



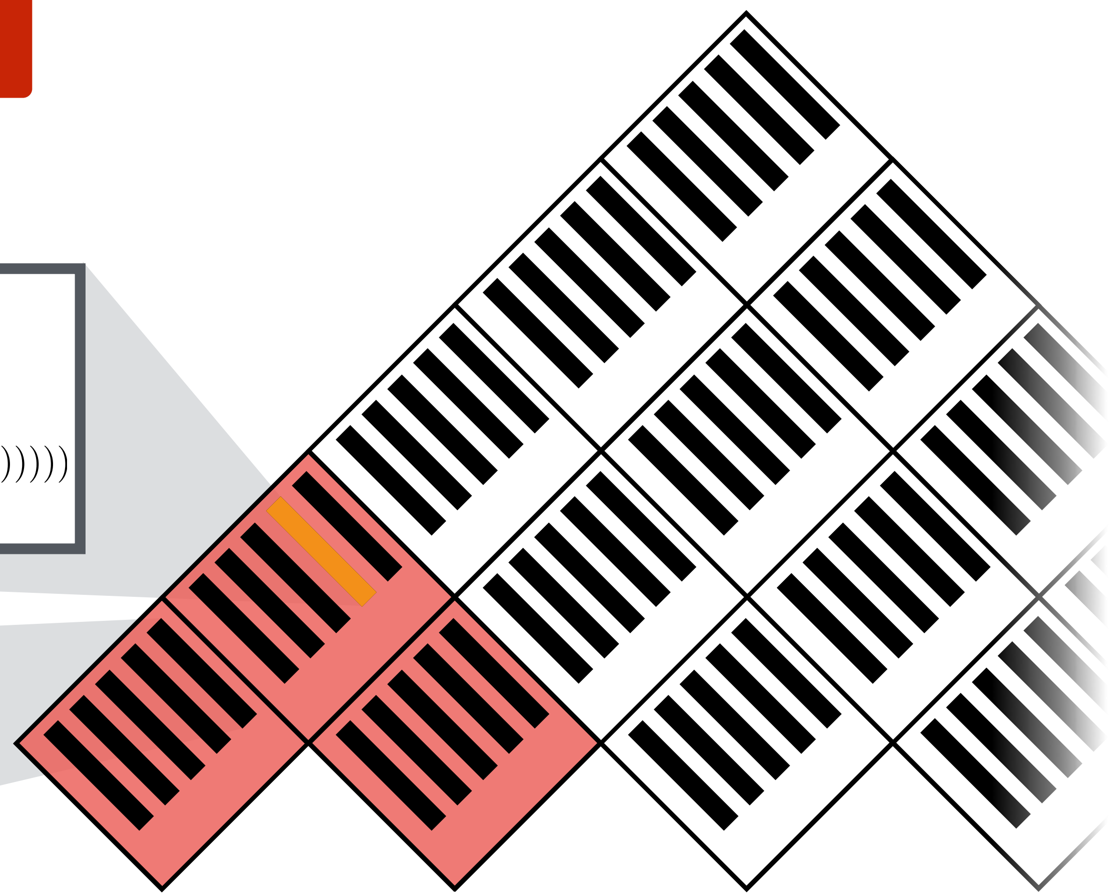
Early Update with Latent Structures

Underspecified Logical Form

$$
\begin{aligned}
&\mathcal{A}_2(\lambda p.\text{person}(p) \wedge \\
&\quad \text{REL-of}(p, \mathcal{A}_3(\lambda h.\text{have-org-role-91}(h) \wedge \\
&\quad\quad \text{ARG1}(h, \mathcal{A}_4(\lambda c.\text{city}(c) \wedge \\
&\quad\quad\quad \text{name}(c, \mathcal{A}_5(\lambda n.\text{name}(n) \wedge \text{op}(n, \text{PYONGYANG}))))) \\
&\quad\quad \text{REL}(h, \mathcal{A}_6(\lambda o.\text{official}(o))))))
\end{aligned}
$$

$$
\begin{aligned}
&\mathcal{A}_4(\lambda c.\text{city}(c) \wedge \\
&\quad \text{name}(c, \mathcal{A}_5(\lambda n.\text{name}(n) \wedge \text{op}(n, \text{PYONGYANG}))))
\end{aligned}
$$

⋮



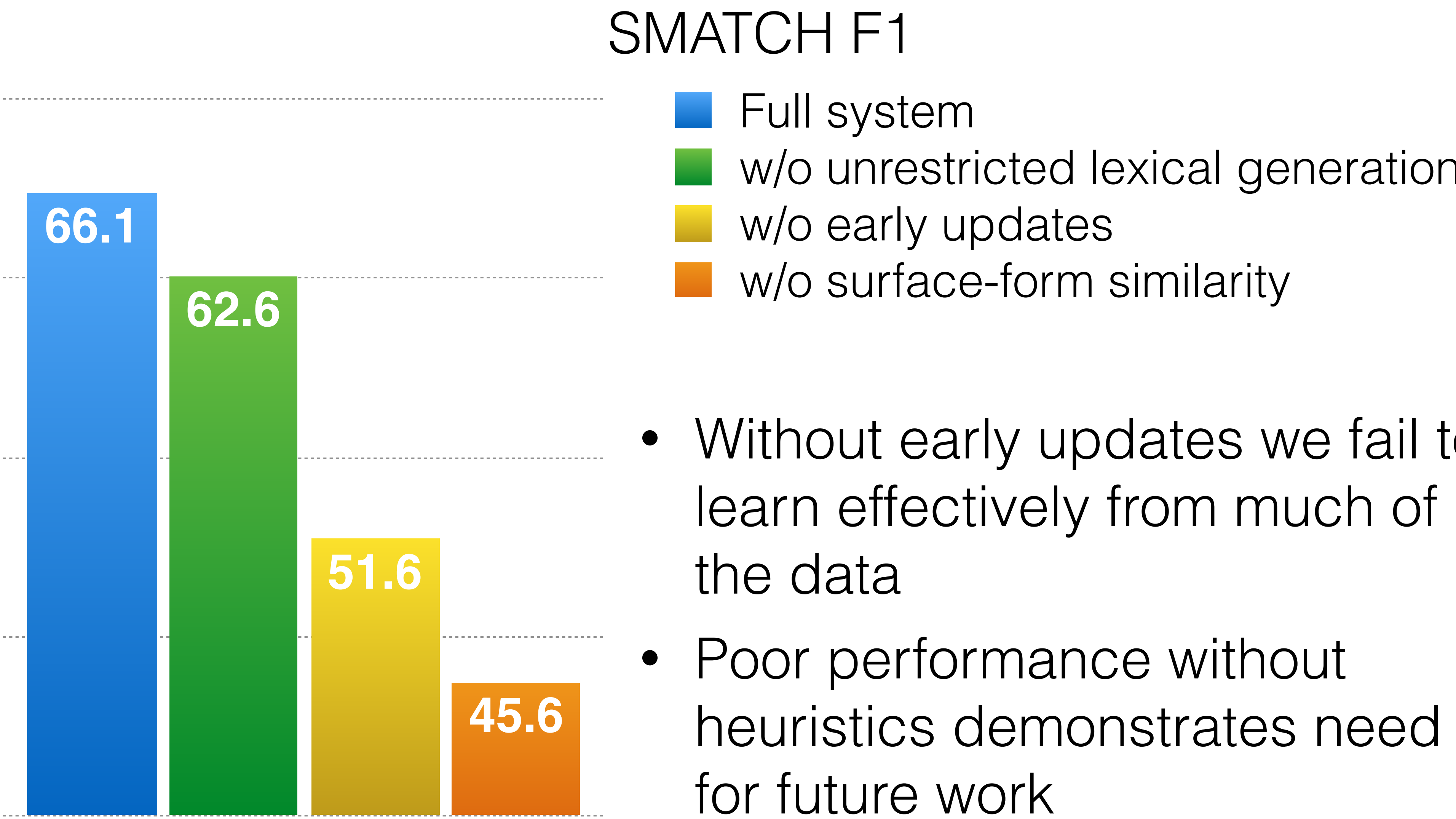
Related Work

CCG Semantic Parsing	[Zettlemoyer and Collins 2005, 2007; Kwiatkowski et al. 2010, 2011; Artzi and Zettlemoyer 2013]
Skolem Terms for CCG	[Steedman 2011]
AMR Evaluation	[Cai and Knight 2013]
Graph-based Parsing for AMR	[Flanigan et al. 2014]
Dependency Structure Transformation for AMR	[Wang et al. 2015a, 2015b]
Syntax-based MT for AMR	[Pust et al. 2015]
Rule-based Parsing for AMR	[Vanderwende et al. 2015]
AMR Applications	[Pan et al. 2015; Lin et al. 2015]

Experimental Setup

- AMR Bank release 1.0, proxy report portion
- Evaluation metric: SMATCH [Cai and Knight 2013]
- Features: lexical features, parsing operations, parsing attachment, selectional preferences, control structures
- Seed lexicon and templates:
    - 50 annotated sentences
    - Heuristic alignment from JAMR [Flanigan et al. 2014]

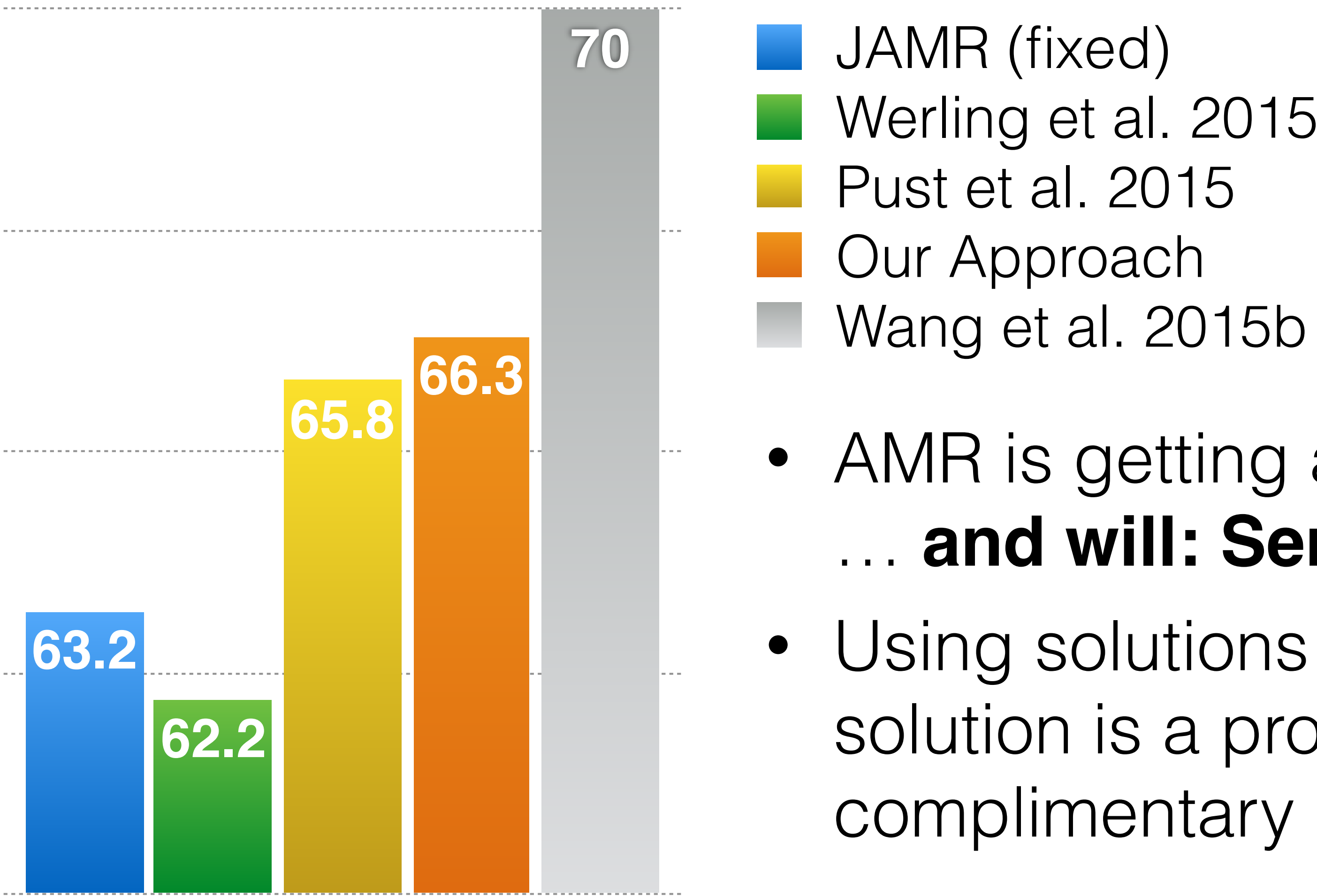
Ablation Results



- Without early updates we fail to learn effectively from much of the data
- Poor performance without heuristics demonstrates need for future work

Results

SMATCH F1



- AMR is getting a lot of attention! … **and will: SemEval 2016**
- Using solutions sub-problem solution is a promising complimentary direction

Contributions

- Joint model for compositional and non-compositional semantics

- Scalable CCG induction for semantic parsing

- First CCG approach to AMR

- **Code and models available in Cornell SPF:** http://yoavartzi.com/spf

Cornell Tech

[fin]