

# Type-driven Incremental Semantic Parsing with Polymorphism

Kai Zhao and Liang Huang

City University of New York

# The Problem

- **fully supervised** semantic parsing

- analogous to MT

Input

What states border Texas?

Output

$\lambda x.state(x) \wedge borders(x, texas)$

- **weakly supervised** semantic parsing

- aka. parsing from Q/A pairs

Input

What states border Texas?

Output

{ LA, AR, OK, NM }

# The Problem

- **fully supervised** semantic parsing
  - analogous to MT

Input

What states border Texas?

Output

$\lambda x.state(x) \wedge borders(x, texas)$

# The Problem

- **fully supervised** semantic parsing
  - analogous to MT

Input

What states border Texas?

Output

$\lambda x.state(x) \wedge borders(x, texas)$

# Challenges

- Unknown Derivation
  - Use Latent Variable
- Unknown Grammar
  - i.e., English phrases  $\leftrightarrow$  predicates
- Learn both derivation & grammar

# Related Work

- Kwiatkowski et al., 2013
  - learn from Q/A pairs
  - based on CCG
  - CKY parsing => underspecified semantics
  - Structural transform to match underspecified semantics to known predicates
    - collapse/expand underspecified predicates

# Our Work

- **incremental** parsing (aka shift-reduce)
- abandon CCG, use **type** to guide parsing
- subtyping **hierarchy** and **polymorphic** functions
- simultaneous grounding

# Decoding

- Type-driven Incremental Parsing
  - **Stack + Queue**
  - Actions:
    - **SHIFT**: pop a word from queue, push its *grounded* semantic expr. onto stack
      - use templates triggered by POS tags/patterns
    - **REDUCE**: function application (*type-driven*)
    - **SKIP**



# A Running Example

**Queue:**

What is the capital of the largest state by area ?  
WP VBZ DT NN IN DT JJS NN IN NN

**Stack**

**Type Constraints**

$\phi$

# A Running Example

Queue:

What is the capital of the largest state by area ?  
WP VBZ DT NN IN DT JJS NN IN NN

**SKIP**

**Stack**

**Type Constraints**

$\phi$

# A Running Example

Queue:

What is the capital of the largest state by area ?  
WP VBZ DT NN IN DT JJS NN IN NN

**SKIP**

**Stack**

**Type Constraints**

$\phi$

# A Running Example

What is the capital of the largest state by area ?  
WP VBZ DT NN IN DT JJS NN IN NN

**SKIP** the

**Queue:** capital ...

**Stack**

**Type Constraints**

$\phi$

# A Running Example

What is the capital of the largest state by area ?  
WP VBZ DT **NN** IN DT JJS NN IN NN

**SHIFT** capital

**Queue:** of ...

**Stack**

**Type Constraints**

$\lambda P : 'b \rightarrow 'c . (\lambda x : 'b . (P x))$

# A Running Example

What is the **capital** of the largest state by area ?

WP VBZ DT **NN** IN DT JJS NN IN NN

**SHIFT** capital

**Queue:** of ...

**Stack**

**Type Constraints**

$\lambda P : 'b \rightarrow 'c . (\lambda x : 'b . (P x))$

**capital** : st  $\rightarrow$  ct

# A Running Example

What is the **capital** of the largest state by area ?

WP VBZ DT **NN** IN DT JJS NN IN NN

**SHIFT** capital

**Queue:** of ...

**Stack**

**Type Constraints**

$\lambda x : 'b . (\mathbf{capital} : st \rightarrow ct \ x)$

$st \rightarrow ct <: 'b \rightarrow 'c$

# A Running Example

What is the **capital** of the largest state by area ?

WP VBZ DT **NN** IN DT JJS NN IN NN

**SHIFT** capital

**Queue:** of ...

**Stack**

**Type Constraints**

$\lambda x : 'b . (\mathbf{capital} : st \rightarrow ct\ x)$

$'b <: st$



# A Running Example

What is the capital of the largest state by area ?  
WP VBZ DT NN IN DT JJS NN IN NN

**SKIP** of  
**Queue:** the ...

**Stack**

**Type Constraints**

$\lambda x : 'b . (\mathbf{capital} : st \rightarrow ct\ x)$

$'b <: st$

# A Running Example

What is the capital of the largest state by area ?  
WP VBZ DT NN IN DT JJS NN IN NN

**SKIP** the

**Queue:** largest ...

**Stack**

**Type Constraints**

$\lambda x : 'b . (\mathbf{capital} : st \rightarrow ct \ x)$

$'b <: st$

# A Running Example

What is the capital of the largest state by area ?  
WP VBZ DT NN IN DT **JJS** NN IN NN

**SHIFT** largest  
**Queue:** state ...

**Stack**

**Type Constraints**

$\lambda x : 'b . (\mathbf{capital} : st \rightarrow ct \ x)$

$'b <: st$

$\lambda P : ('a \rightarrow t) \rightarrow ('a \rightarrow i) \rightarrow 'a .$

$\lambda x : 'a \rightarrow t . \lambda y : 'a \rightarrow i . (P \ x \ y)$

# A Running Example

What is the capital of the **largest** state by area ?  
WP VBZ DT NN IN DT **JJS** NN IN NN

**SHIFT** largest  
**Queue:** state ...

**Stack**

**Type Constraints**

$\lambda x : 'b . (\mathbf{capital} : st \rightarrow ct \ x)$

$'b <: st$

$\lambda P : ('a \rightarrow t) \rightarrow ('a \rightarrow i) \rightarrow 'a .$

$\lambda x : 'a \rightarrow t . \lambda y : 'a \rightarrow i . (P \ x \ y)$

$\mathbf{argmax} : ('b \rightarrow t) \rightarrow ('b \rightarrow i) \rightarrow 'b$

# A Running Example

What is the capital of the **largest** state by area ?

WP VBZ DT NN IN DT **JJS** NN IN NN

**SHIFT** largest  
**Queue:** state ...

**Stack**

**Type Constraints**

$\lambda x : 'b . (\mathbf{capital} : st \rightarrow ct\ x)$

$'b <: st$

$\lambda x : 'a \rightarrow t . \lambda y : 'a \rightarrow i .$

$(\mathbf{argmax} : ('a \rightarrow t) \rightarrow ('a \rightarrow i) \rightarrow 'a$   
 $x\ y)$

$('b \rightarrow t) \rightarrow ('b \rightarrow i) \rightarrow 'b <:$

$('a \rightarrow t) \rightarrow ('a \rightarrow i) \rightarrow 'a$

# A Running Example

What is the capital of the **largest** state by area ?

WP VBZ DT NN IN DT **JJS** NN IN NN

**SHIFT** largest  
**Queue:** state ...

**Stack**

**Type Constraints**

$\lambda x : 'b . (\mathbf{capital} : st \rightarrow ct \ x)$

$'b <: st$

$\lambda x : 'a \rightarrow t . \lambda y : 'a \rightarrow i .$

$(\mathbf{argmax} : ('a \rightarrow t) \rightarrow ('a \rightarrow i) \rightarrow 'a$   
 $\quad x \ y)$

# A Running Example

What is the capital of the largest state by area ?

WP VBZ DT NN IN DT JJS NN IN NN

**SHIFT** state

**Queue:** by ...

**Stack**

**Type Constraints**

$\lambda x : 'b . (\mathbf{capital} : st \rightarrow ct \ x)$

$'b <: st$

$\lambda x : 'a \rightarrow t . \lambda y : 'a \rightarrow i .$

$(\mathbf{argmax} : ('a \rightarrow t) \rightarrow ('a \rightarrow i) \rightarrow 'a$   
 $\quad \quad \quad x \ y)$

$\lambda x : 'c . (\mathbf{state} : st \rightarrow t \ x)$

$'c <: st$

# A Running Example

What is the capital of the largest state by area ?

WP VBZ DT NN IN DT JJS NN IN NN

**SHIFT** state

**Queue:** by ...

**Stack**

**Type Constraints**

$\lambda x : 'b . (\mathbf{capital} : st \rightarrow ct \ x)$

$'b <: st$

$\lambda x : 'a \rightarrow t . \lambda y : 'a \rightarrow i .$

$(\mathbf{argmax} : ('a \rightarrow t) \rightarrow ('a \rightarrow i) \rightarrow 'a$

$x \ y) \quad \lambda x : 'c . (\mathbf{state} : st \rightarrow t \ x)$

$'c <: st$



# A Running Example

What is the capital of the largest state by area ?

WP VBZ DT NN IN DT JJS NN IN NN

**SHIFT** state

**Queue:** by ...

**Stack**

**Type Constraints**

$\lambda x : 'b . (\mathbf{capital} : st \rightarrow ct \ x)$

$'b <: st$

$\lambda y : 'a \rightarrow i . (\mathbf{argmax} : ('a \rightarrow t) \rightarrow ('a \rightarrow i) \rightarrow 'a$

$'c <: st$

$\lambda x : 'c . (\mathbf{state} : st \rightarrow t \ x) \ y)$

$'c \rightarrow t <: 'a \rightarrow t$

# A Running Example

What is the capital of the largest state by area ?

WP VBZ DT NN IN DT JJS NN IN NN

**SHIFT** state

**Queue:** by ...

**Stack**

**Type Constraints**

$\lambda x : 'b . (\mathbf{capital} : st \rightarrow ct \ x)$

$'b <: st$

$\lambda y : 'a \rightarrow i . (\mathbf{argmax} : ('a \rightarrow t) \rightarrow ('a \rightarrow i) \rightarrow 'a$

$'c <: st$

$\lambda x : 'c . (\mathbf{state} : st \rightarrow t \ x) \ y)$

$'a <: 'c$

# A Running Example

What is the capital of the largest state by area ?

WP VBZ DT NN IN DT JJS NN IN NN

**SKIP** by

**Queue:** area ...

**Stack**

**Type Constraints**

$\lambda x : 'b . (\mathbf{capital} : st \rightarrow ct \ x)$

$'b <: st$

$\lambda y : 'a \rightarrow i . (\mathbf{argmax} : ('a \rightarrow t) \rightarrow ('a \rightarrow i) \rightarrow 'a$

$'c <: st$

$\lambda x : 'c . (\mathbf{state} : st \rightarrow t \ x) \ y)$

$'a <: 'c$

# A Running Example

What is the capital of the largest state by area ?

WP VBZ DT NN IN DT JJS NN IN NN

**SHIFT** area

**Queue:**  $\phi$

**Stack**

**Type Constraints**

$\lambda x : 'b . (\mathbf{capital} : \text{st} \rightarrow \text{ct } x)$

$'b <: \text{st}$

$\lambda y : 'a \rightarrow i . (\mathbf{argmax} : ('a \rightarrow \text{t}) \rightarrow ('a \rightarrow i) \rightarrow 'a$

$'c <: \text{st}$

$\lambda x : 'c . (\mathbf{state} : \text{st} \rightarrow \text{t } x) \ y)$

$'a <: 'c$

$\lambda x : 'd . (\mathbf{size} : \text{lo} \rightarrow i \ x)$

$'d <: \text{lo}$

# A Running Example

What is the capital of the largest state by area ?

WP VBZ DT NN IN DT JJS NN IN NN

**REDUCE**

**Queue:**  $\phi$

**Stack**

**Type Constraints**

$\lambda x : 'b . (\mathbf{capital} : \text{st} \rightarrow \text{ct } x)$

$'b <: \text{st}$

$\lambda y : 'a \rightarrow i . (\mathbf{argmax} : ('a \rightarrow \text{t}) \rightarrow ('a \rightarrow i) \rightarrow 'a$

$'c <: \text{st}$

$\lambda x : 'c . (\mathbf{state} : \text{st} \rightarrow \text{t } x) \ y)$

$'a <: 'c$

$\lambda x : 'd . (\mathbf{size} : \text{lo} \rightarrow i \ x)$

$'d <: \text{lo}$

# A Running Example

What is the capital of the largest state by area ?

WP VBZ DT NN IN DT JJS NN IN NN

**REDUCE**

**Queue:**  $\phi$

**Stack**

**Type Constraints**

$\lambda x : 'b . (\mathbf{capital} : st \rightarrow ct\ x)$

$'b <: st$

$(\mathbf{argmax} : ('a \rightarrow t) \rightarrow ('a \rightarrow i) \rightarrow 'a$

$'c <: st$

$(\lambda x : 'c . (\mathbf{state} : st \rightarrow t\ x))$

$'a <: 'c$

$(\lambda y : 'd . (\mathbf{size} : lo \rightarrow i\ y)))$

$'d <: lo$

# A Running Example

What is the capital of the largest state by area ?

WP VBZ DT NN IN DT JJS NN IN NN

**REDUCE**

**Queue:**  $\phi$

Stack	Type Constraints
<b>capital</b> : $st \rightarrow ct$ ( <b>argmax</b> : $('a \rightarrow t) \rightarrow ('a \rightarrow i) \rightarrow 'a$ $(\lambda x : 'c . (\mathbf{state} : st \rightarrow t x))$ $(\lambda y : 'd . (\mathbf{size} : lo \rightarrow i y))))$	$'c <: st$ $'a <: 'c$ $'d <: lo$

# A Running Example

What is the capital of the largest state by area ?  
WP VBZ DT NN IN DT JJS NN IN NN

(**capital** : st → ct (**argmax** : ('a → t) → ('a → i) → 'a 'c <: st  
(λx : 'c . (**state** : st → t x)) 'a <: 'c  
(λy : 'd . (**size** : lo → i y)))) 'd <: lo



# Learning

- Both derivation/Grammar are Unknown
- Spurious Ambiguity
  - Various derivations/groundings lead to the same logical form
- Latent Variable
  - Structured Perceptron =>  
Latent Variable Structured Perceptron

# Learning

What is the capital of the largest state by area ?

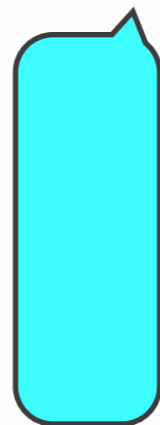



(**capital** : st → ct (**argmax** : ('a → t) → ('a → i) → 'a  
(λx : 'c . (**state** : st → t x))  
(λy : 'd . (**size** : lo → i y))))

# Learning

What is the capital of the largest state by area ?

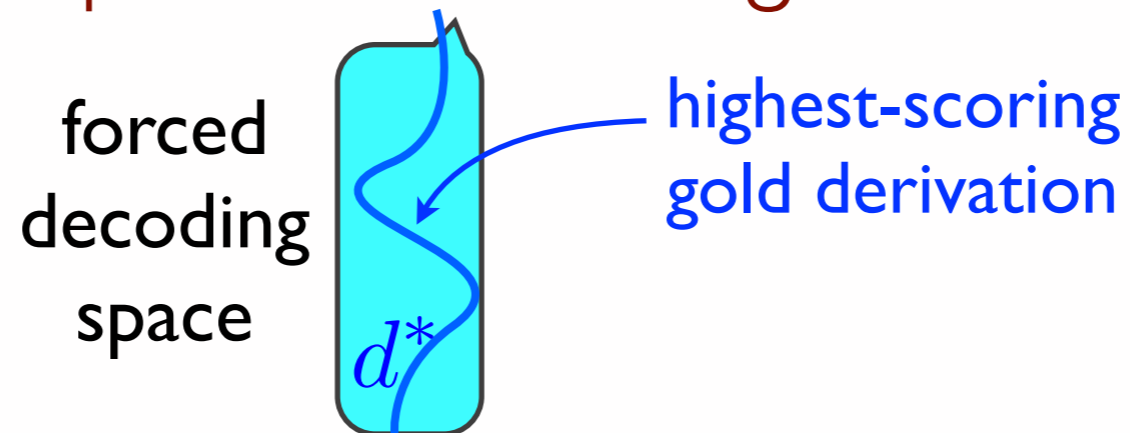
forced  
decoding  
space



 (**capital** : st → ct (**argmax** : ('a → t) → ('a → i) → 'a  
(λx : 'c . (**state** : st → t x))  
(λy : 'd . (**size** : lo → i y))))

# Learning

What is the capital of the largest state by area ?



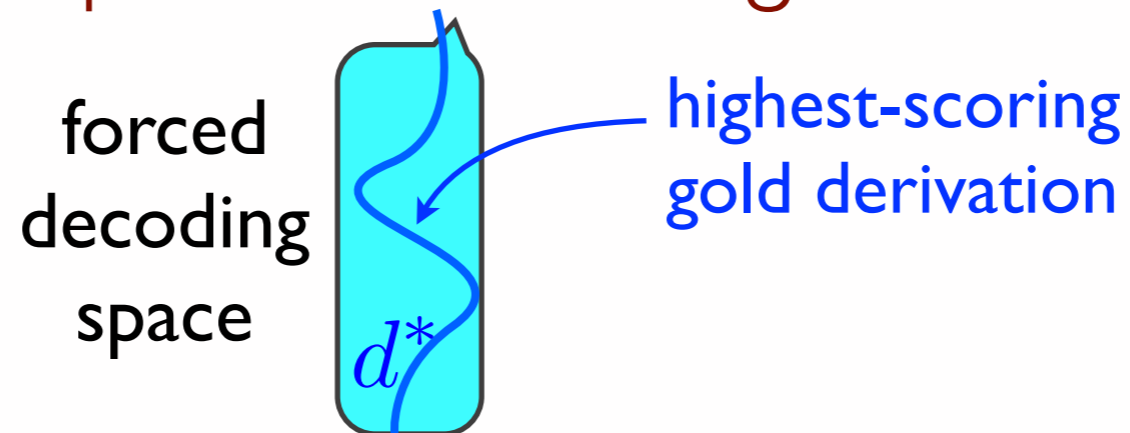
(**capital** : st  $\rightarrow$  ct (**argmax** : ('a  $\rightarrow$  t)  $\rightarrow$  ('a  $\rightarrow$  i)  $\rightarrow$  'a  
( $\lambda x$  : 'c . (**state** : st  $\rightarrow$  t x))  
( $\lambda y$  : 'd . (**size** : lo  $\rightarrow$  i y))))

# Learning



(**size** : st  $\rightarrow$  i (**argmin** : ('a  $\rightarrow$  t)  $\rightarrow$  ('a  $\rightarrow$  i)  $\rightarrow$  'a  
( $\lambda x$  : 'c . (**city** : ct  $\rightarrow$  t  $x$ ))  
( $\lambda y$  : 'd . (**population** : au  $\rightarrow$  i  $y$ ))))))

What is the capital of the largest state by area ?



(**capital** : st  $\rightarrow$  ct (**argmax** : ('a  $\rightarrow$  t)  $\rightarrow$  ('a  $\rightarrow$  i)  $\rightarrow$  'a  
( $\lambda x$  : 'c . (**state** : st  $\rightarrow$  t  $x$ ))  
( $\lambda y$  : 'd . (**size** : lo  $\rightarrow$  i  $y$ ))))))

# Learning



(**size** : st  $\rightarrow$  i (**argmin** : ('a  $\rightarrow$  t)  $\rightarrow$  ('a  $\rightarrow$  i)  $\rightarrow$  'a  
( $\lambda x$  : 'c . (**city** : ct  $\rightarrow$  t x))  
( $\lambda y$  : 'd . (**population** : au  $\rightarrow$  i y))))))

full  
search  
space

What is the capital of the largest state by area ?

forced  
decoding  
space



highest-scoring  
gold derivation

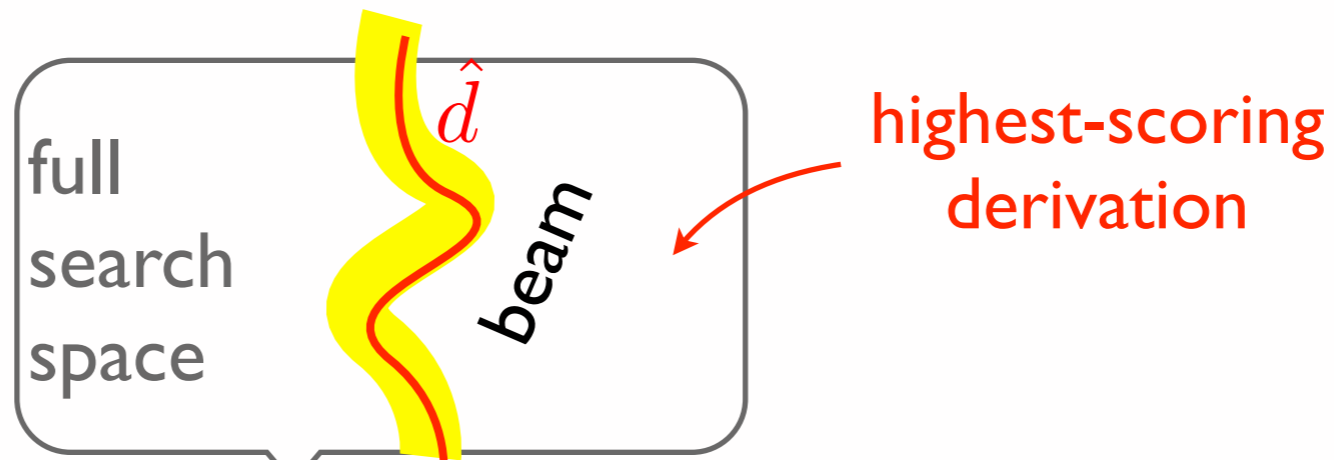


(**capital** : st  $\rightarrow$  ct (**argmax** : ('a  $\rightarrow$  t)  $\rightarrow$  ('a  $\rightarrow$  i)  $\rightarrow$  'a  
( $\lambda x$  : 'c . (**state** : st  $\rightarrow$  t x))  
( $\lambda y$  : 'd . (**size** : lo  $\rightarrow$  i y))))))

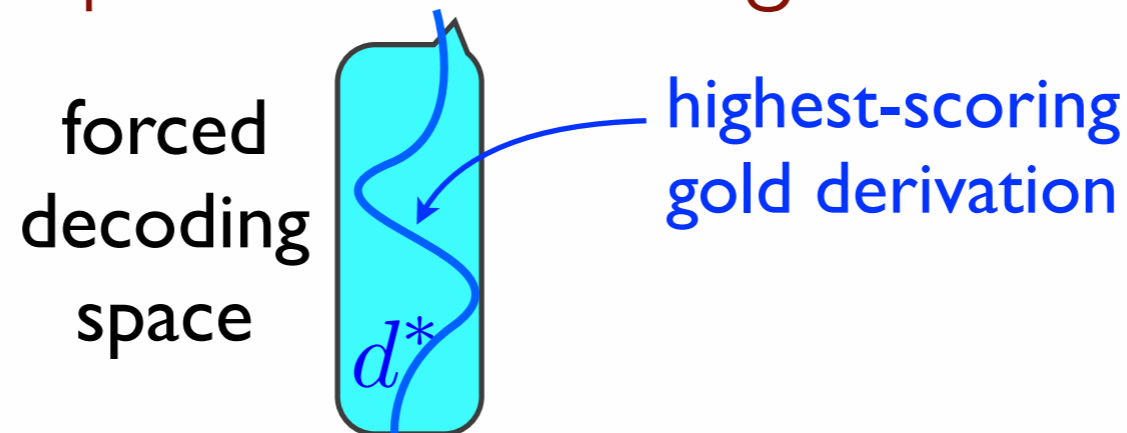
# Learning



(**size** : st  $\rightarrow$  i (**argmin** : ('a  $\rightarrow$  t)  $\rightarrow$  ('a  $\rightarrow$  i)  $\rightarrow$  'a  
( $\lambda x$  : 'c . (**city** : ct  $\rightarrow$  t  $x$ ))  
( $\lambda y$  : 'd . (**population** : au  $\rightarrow$  i  $y$ ))))))



What is the capital of the largest state by area ?

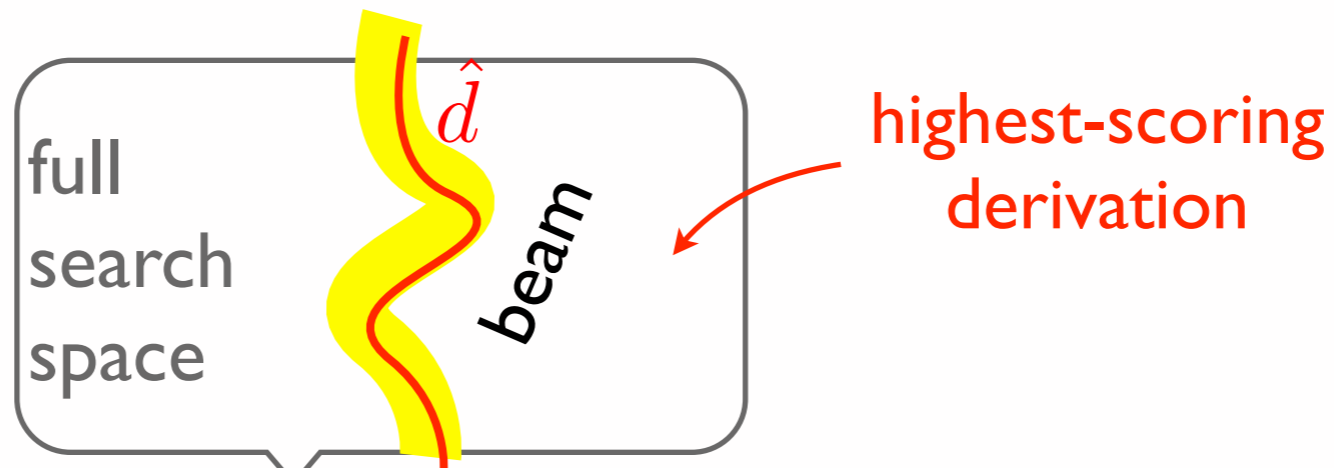


(**capital** : st  $\rightarrow$  ct (**argmax** : ('a  $\rightarrow$  t)  $\rightarrow$  ('a  $\rightarrow$  i)  $\rightarrow$  'a  
( $\lambda x$  : 'c . (**state** : st  $\rightarrow$  t  $x$ ))  
( $\lambda y$  : 'd . (**size** : lo  $\rightarrow$  i  $y$ ))))))

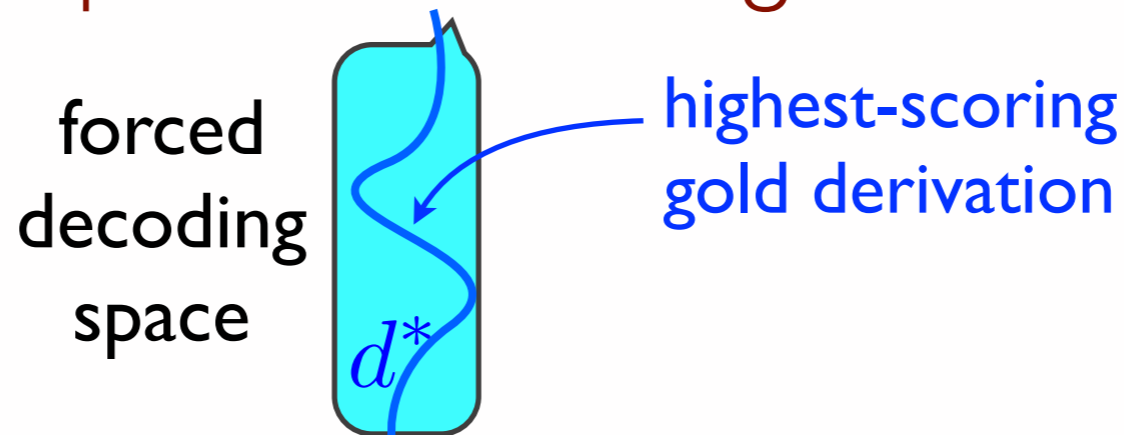
# Learning



(**size** : st  $\rightarrow$  i (**argmin** : ('a  $\rightarrow$  t)  $\rightarrow$  ('a  $\rightarrow$  i)  $\rightarrow$  'a  
( $\lambda x$  : 'c . (**city** : ct  $\rightarrow$  t x))  
( $\lambda y$  : 'd . (**population** : au  $\rightarrow$  i y))))



What is the capital of the largest state by area ?



**REWARD**



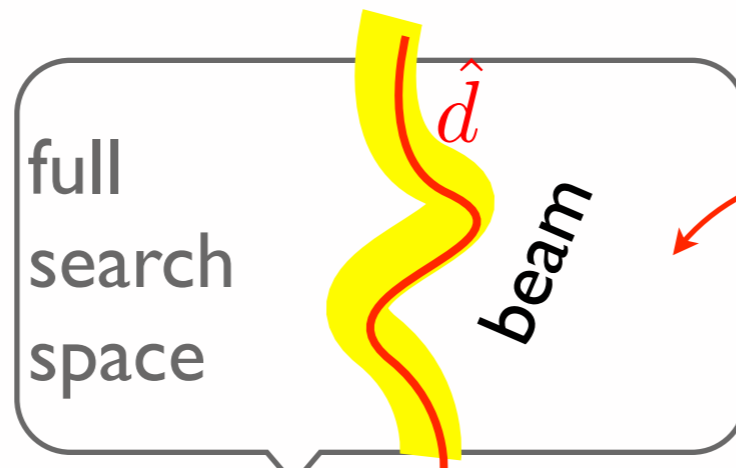
(**capital** : st  $\rightarrow$  ct (**argmax** : ('a  $\rightarrow$  t)  $\rightarrow$  ('a  $\rightarrow$  i)  $\rightarrow$  'a  
( $\lambda x$  : 'c . (**state** : st  $\rightarrow$  t x))  
( $\lambda y$  : 'd . (**size** : lo  $\rightarrow$  i y))))



# Learning



(**size** : st  $\rightarrow$  i (**argmin** : ('a  $\rightarrow$  t)  $\rightarrow$  ('a  $\rightarrow$  i)  $\rightarrow$  'a  
( $\lambda x$  : 'c . (**city** : ct  $\rightarrow$  t x))  
( $\lambda y$  : 'd . (**population** : au  $\rightarrow$  i y))))



highest-scoring  
derivation

**PENALIZE**

What is the capital of the largest state by area ?

forced  
decoding  
space



highest-scoring  
gold derivation

**REWARD**



(**capital** : st  $\rightarrow$  ct (**argmax** : ('a  $\rightarrow$  t)  $\rightarrow$  ('a  $\rightarrow$  i)  $\rightarrow$  'a  
( $\lambda x$  : 'c . (**state** : st  $\rightarrow$  t x))  
( $\lambda y$  : 'd . (**size** : lo  $\rightarrow$  i y))))

# Learning



(**size** : st  $\rightarrow$  i (**argmin** : ('a  $\rightarrow$  t)  $\rightarrow$  ('a  $\rightarrow$  i)  $\rightarrow$  'a  
( $\lambda x$  : 'c . (**city** : ct  $\rightarrow$  t x))  
( $\lambda y$  : 'd . (**population** : au  $\rightarrow$  i y))))

full

$\hat{d}$

highest-scoring  
derivation

**PENALIZE**

$$\mathbf{w} \leftarrow \mathbf{w} + \Phi(x, d^*) - \Phi(x, \hat{d})$$

Wh

reward  
correct

penalize  
wrong

?

space

$d^*$

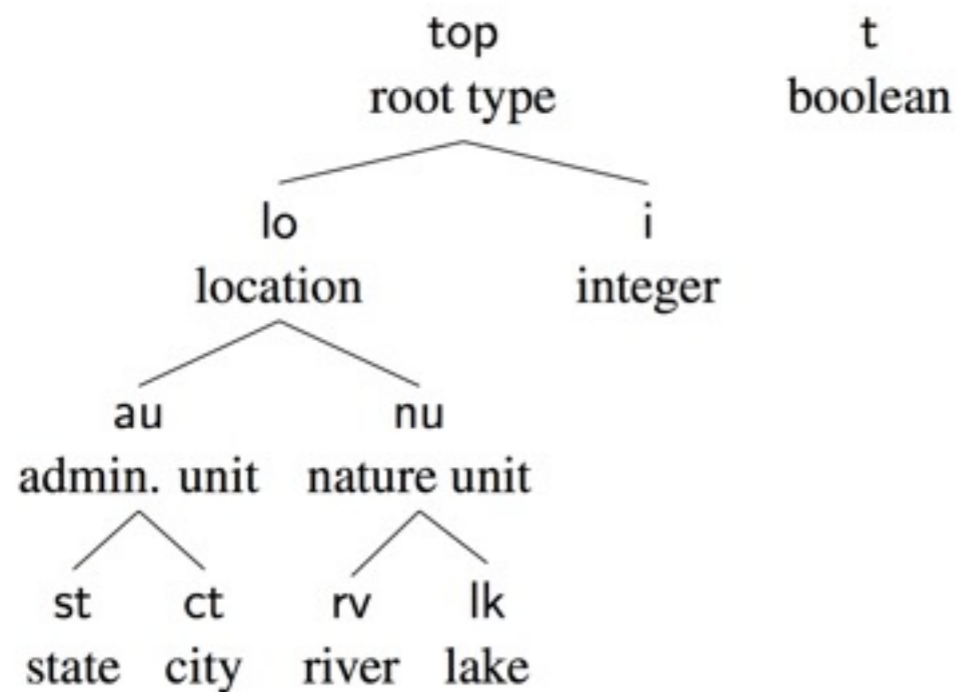
**REWARD**



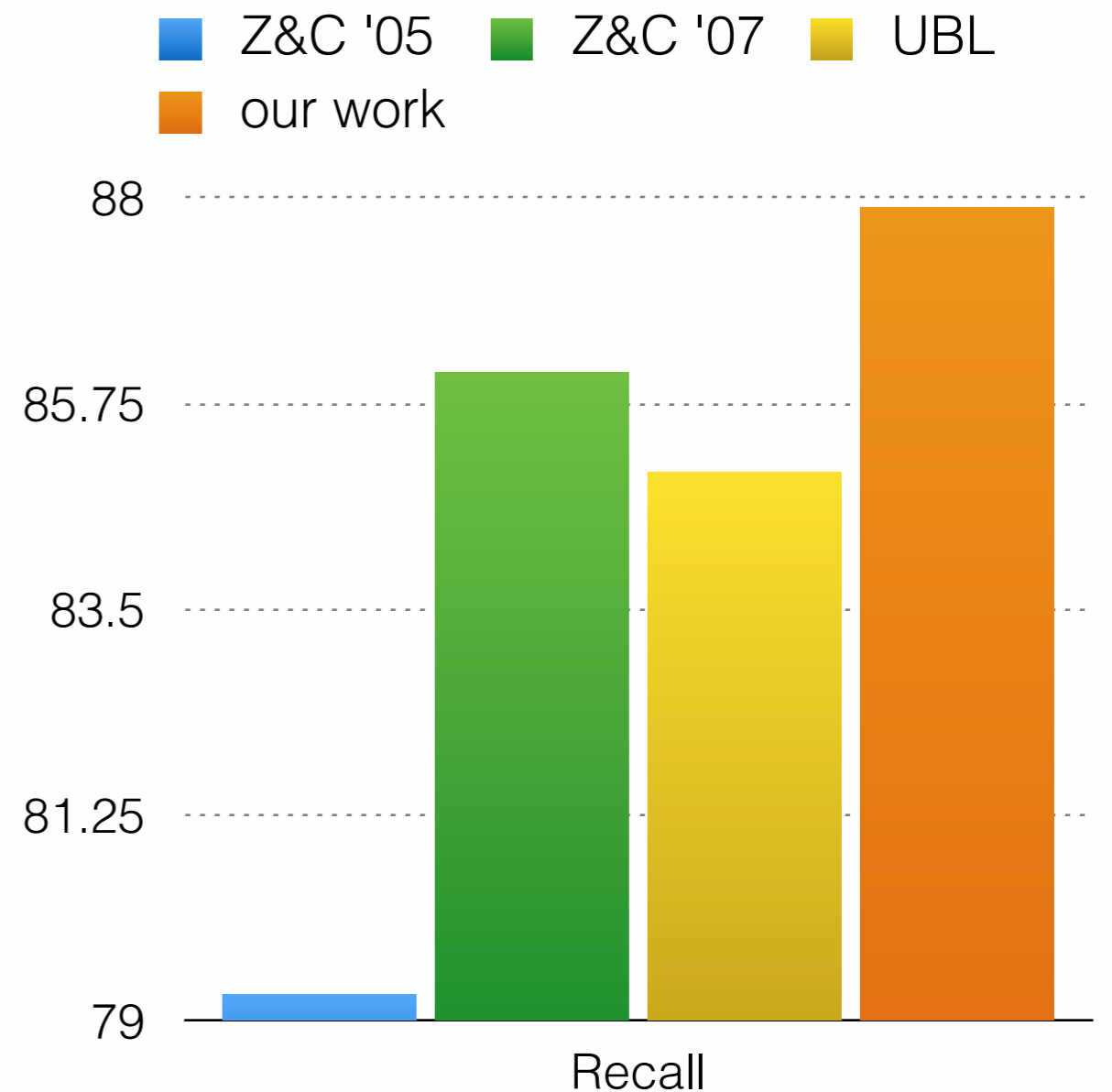
(**capital** : st  $\rightarrow$  ct (**argmax** : ('a  $\rightarrow$  t)  $\rightarrow$  ('a  $\rightarrow$  i)  $\rightarrow$  'a  
( $\lambda x$  : 'c . (**state** : st  $\rightarrow$  t x))  
( $\lambda y$  : 'd . (**size** : lo  $\rightarrow$  i y))))

# Experiments

- High decoding speed; Linear in theory & practice
  - 0.5 sec/sentence
- GeoQuery



Type Hierarchy



# Conclusion

- Polymorphic typing guides the parsing
- Linear time incremental parsing
- Simultaneous grounding
- Learning w/ Latent Variable Structured Perceptron