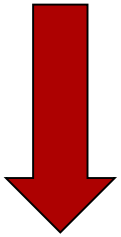
Data Recombination for Neural Semantic Parsing

Robin Jia and Percy Liang
Stanford University

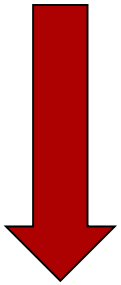
Semantic Parsing

What states border Texas ?



Semantic Parser

And(State, NextTo(StateId(Texas)))



Executor

[New Mexico, Oklahoma, Arkansas, Louisiana]

Zelle and Mooney (1996), Zettlemoyer and Collins (2005, 2007), Liang et al. (2011), Artzi and Zettlemoyer (2013), Berant et al. (2013).

Outline

- Neural Semantic Parser
- Data Recombination
- Discussion

Outline

- Neural Semantic Parser
- Data Recombination
- Discussion

Semantic Parsing

- Traditional semantic parsers are engineering-heavy
- Can we get good performance with a domain-general model?

Semantic Parsing

- Geoquery
  - Input: *what is the population of iowa ?*
  - Output: `_answer ( A , ( _population ( B , A ) , _const ( B , _stateid ( iowa ) ) ) )`
- ATIS
  - Input: *list all flights from chicago to milwaukee*
  - Output: `( _lambda $0 e ( _and ( _flight $0 ) ( _from $0 chicago:_ci ) ( _to $0 milwaukee:_ci ) ) )`
- Overnight
  - Input: *what restaurants have takeout*
  - Output: `( call list ( call filter ( call getProperty ( call singleton en.restaurant ) ( string ! type ) ) ( string takeout ) ) )`

Zelle and Mooney (1996), Dahl et al. (1994), Wang et al. (2015).

Sequence-to-sequence Models

- Machine Translation
  - Input: the blue house
  - Output: la maison bleue
- Syntactic Parsing
  - Input: the dog barked
  - Output: (S (NP DT NN)$_{NP}$ (VP VBD)$_{VP}$)$_{S}$
- And many more…

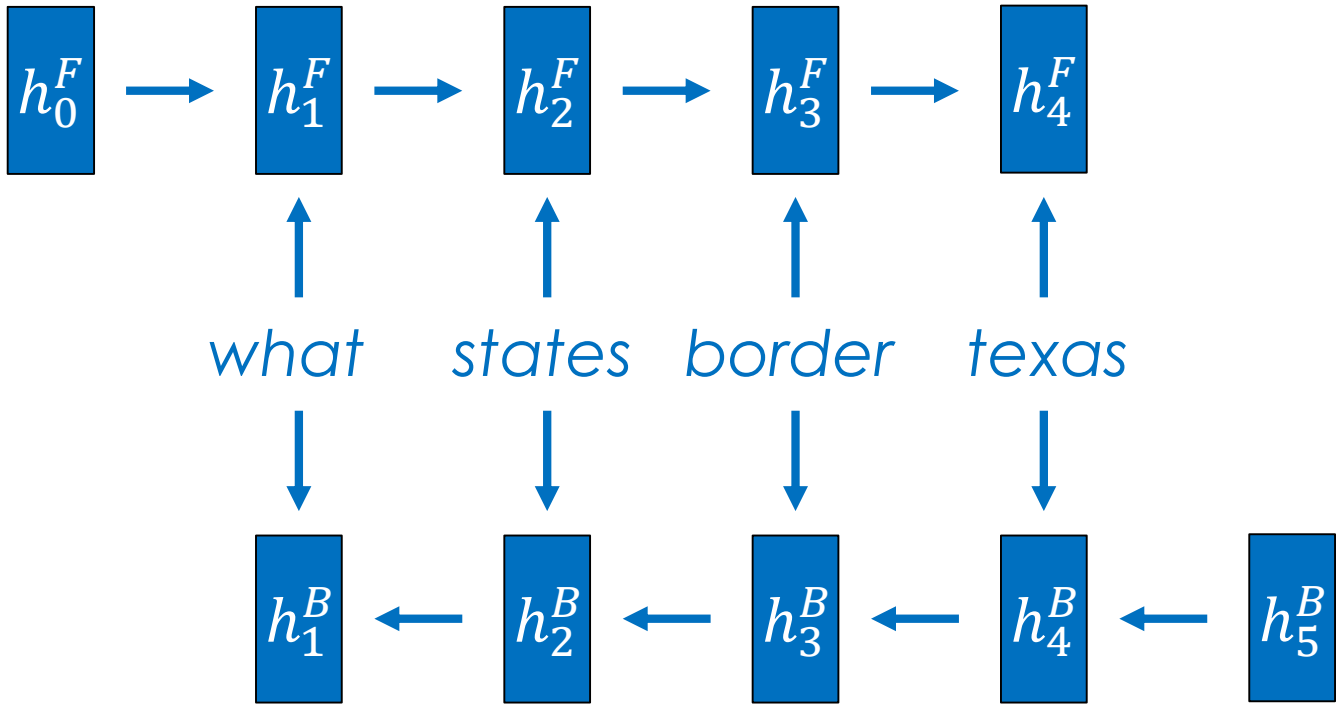
Sutskever et al. (2014), Bahdanau et al. (2014), Vinyals et al. (2015).

Semantic Parsing

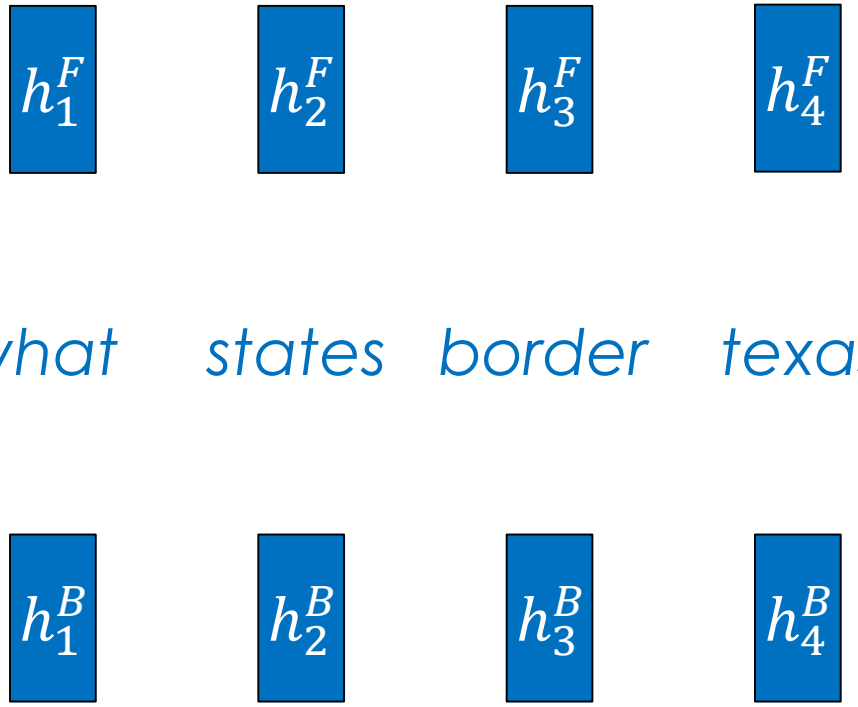
- Treat semantic parsing as a sequence-to-sequence task
- Use domain-general attention-based neural model

Sutskever et al. (2014), Bahdanau et al. (2014), Luong et al. (2015).

Neural Semantic Parser

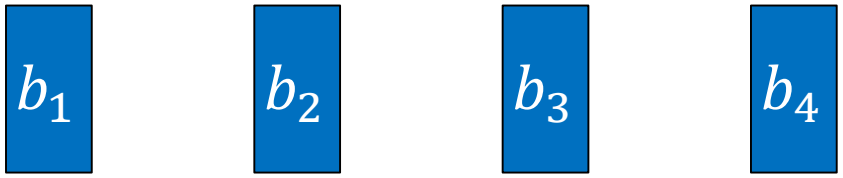


Neural Semantic Parser

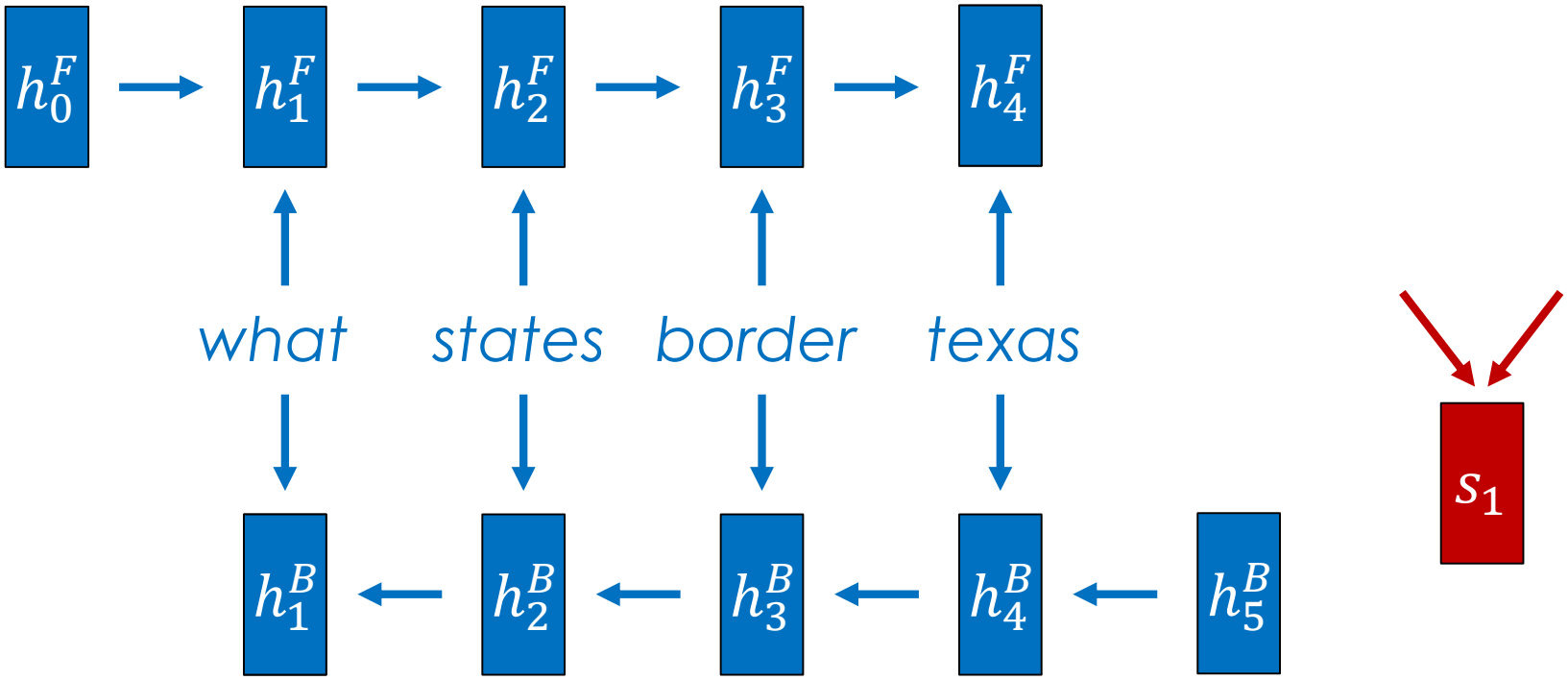


Neural Semantic Parser

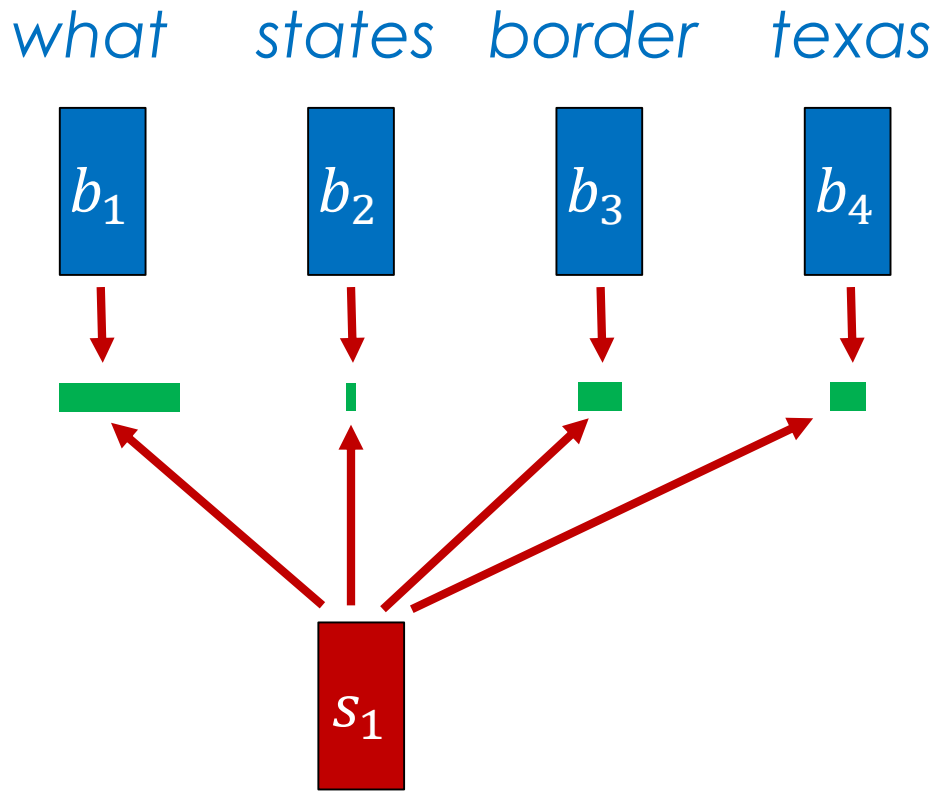
what *states* *border* *texas*



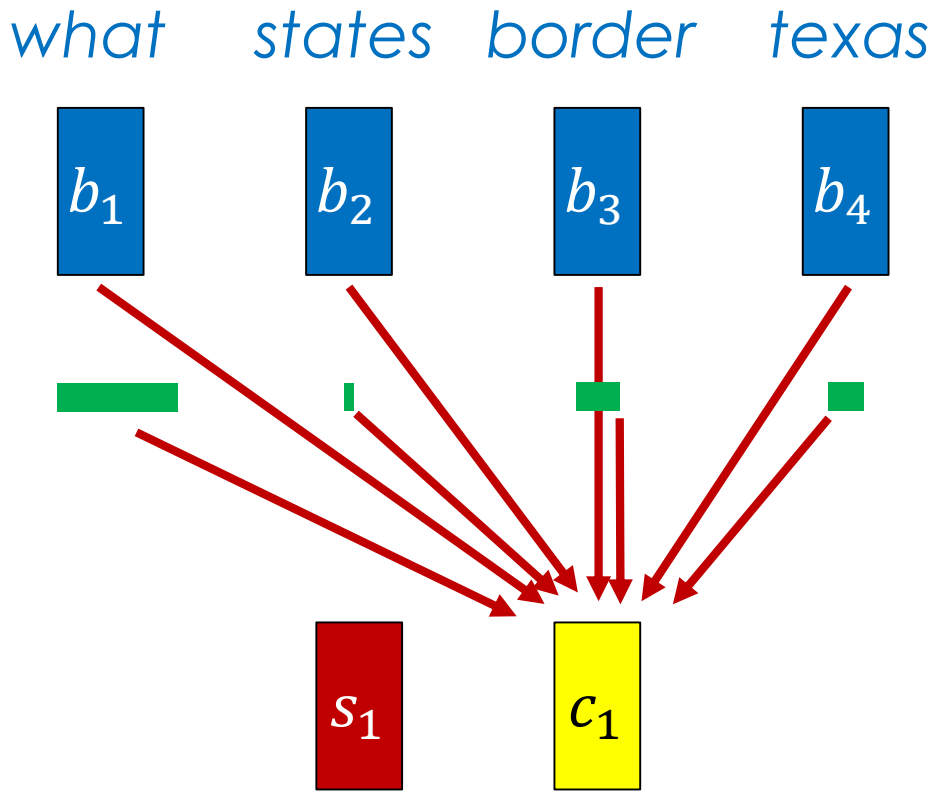
Neural Semantic Parser



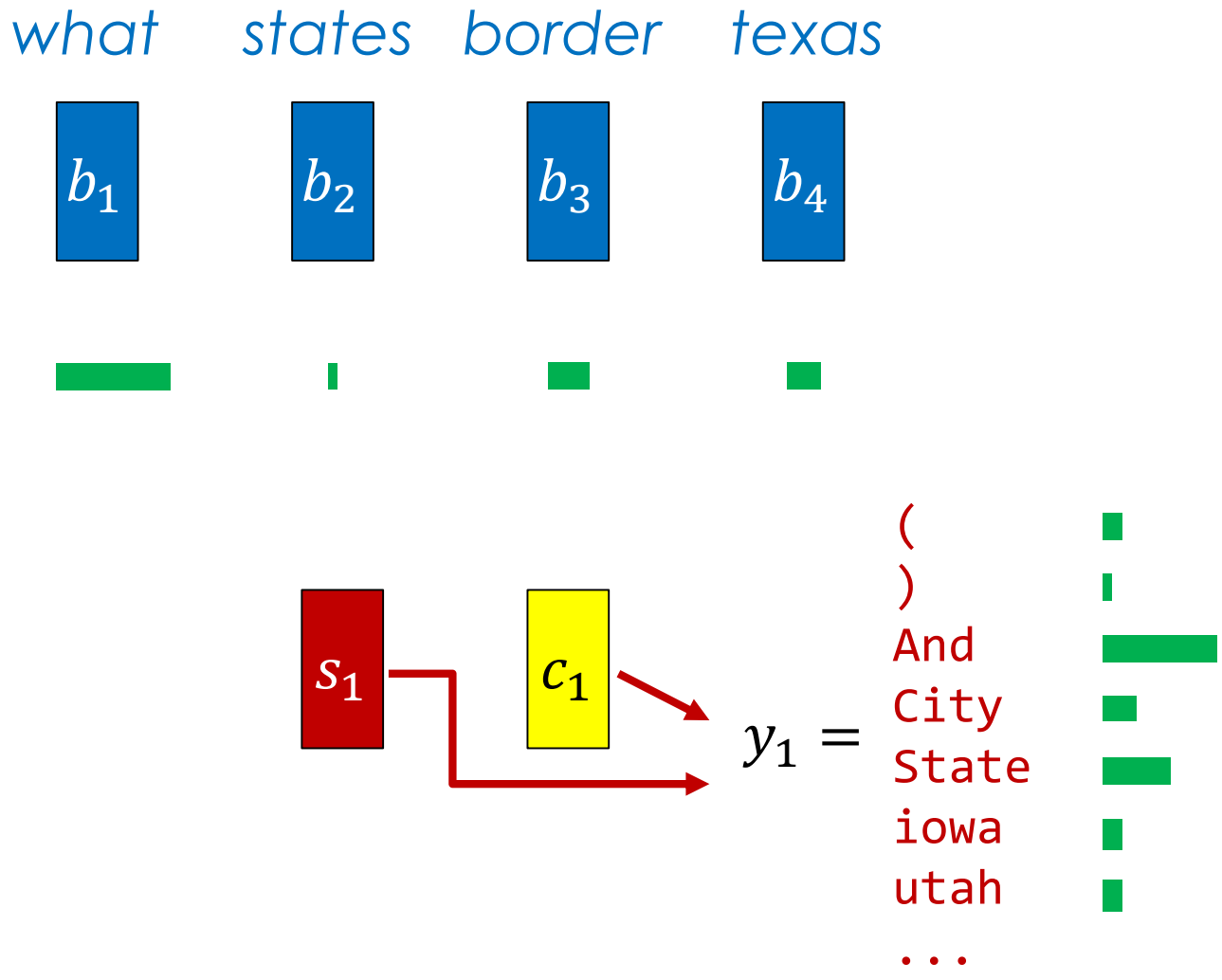
Neural Semantic Parser



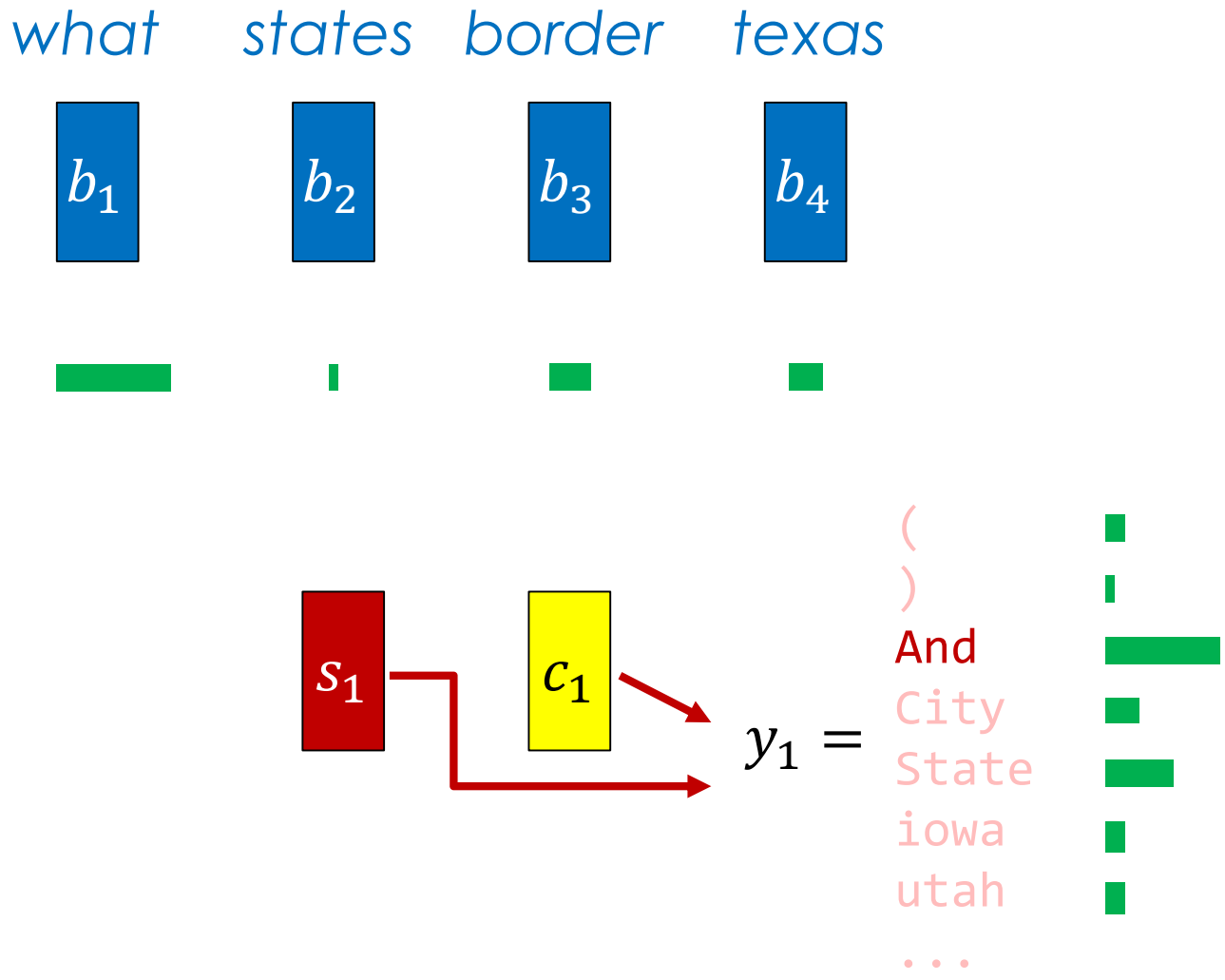
Neural Semantic Parser



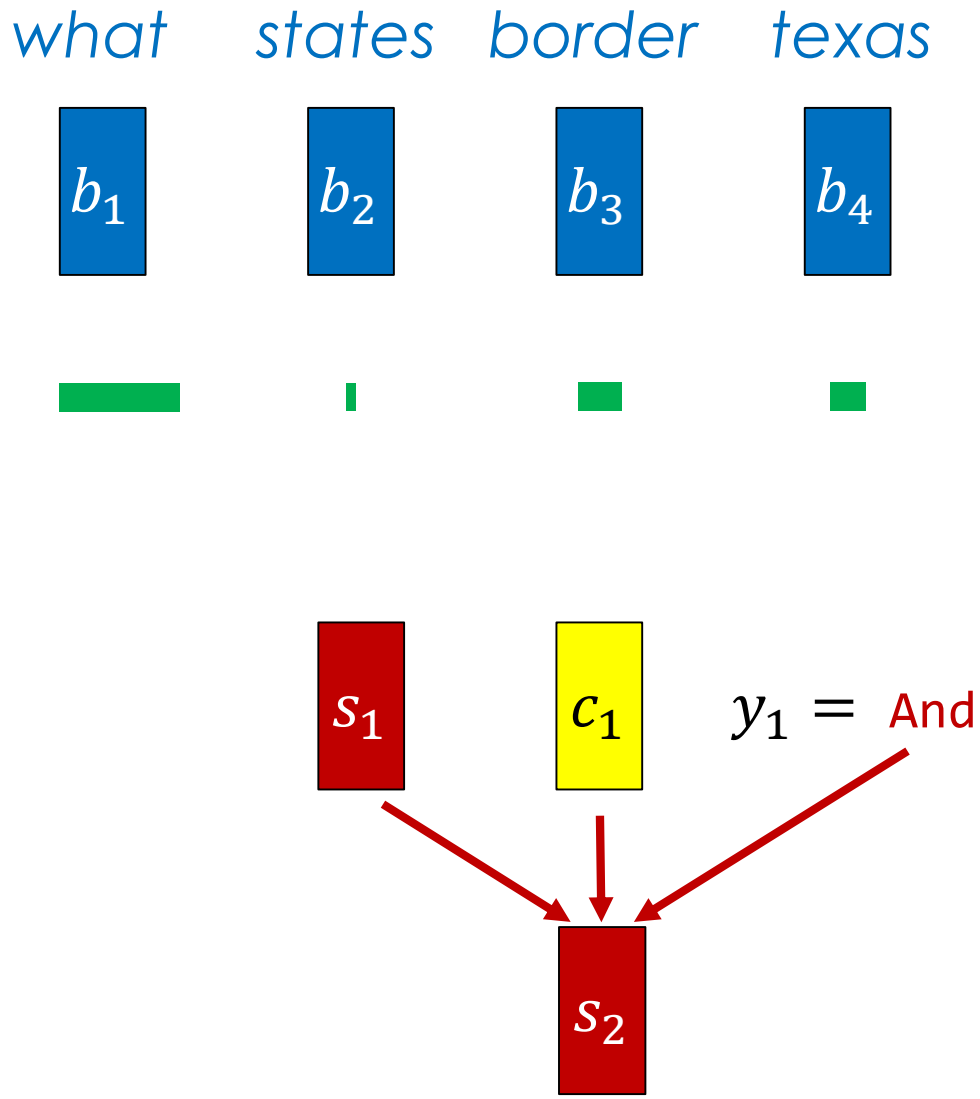
Neural Semantic Parser



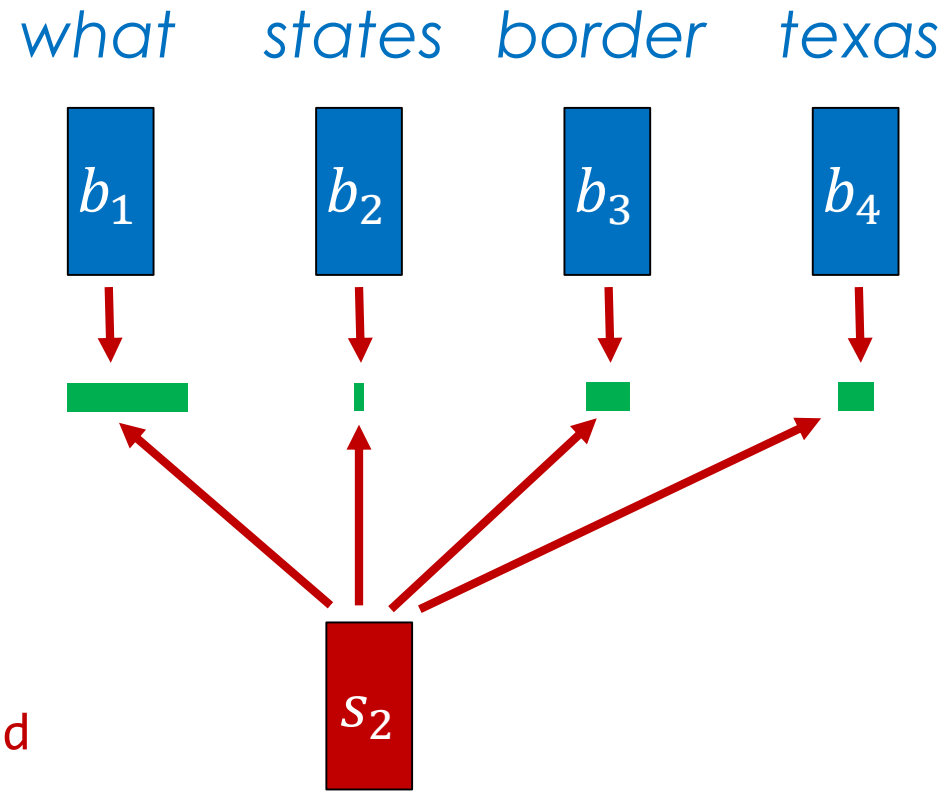
Neural Semantic Parser



Neural Semantic Parser



Neural Semantic Parser



$y =$ And

Rare Entities

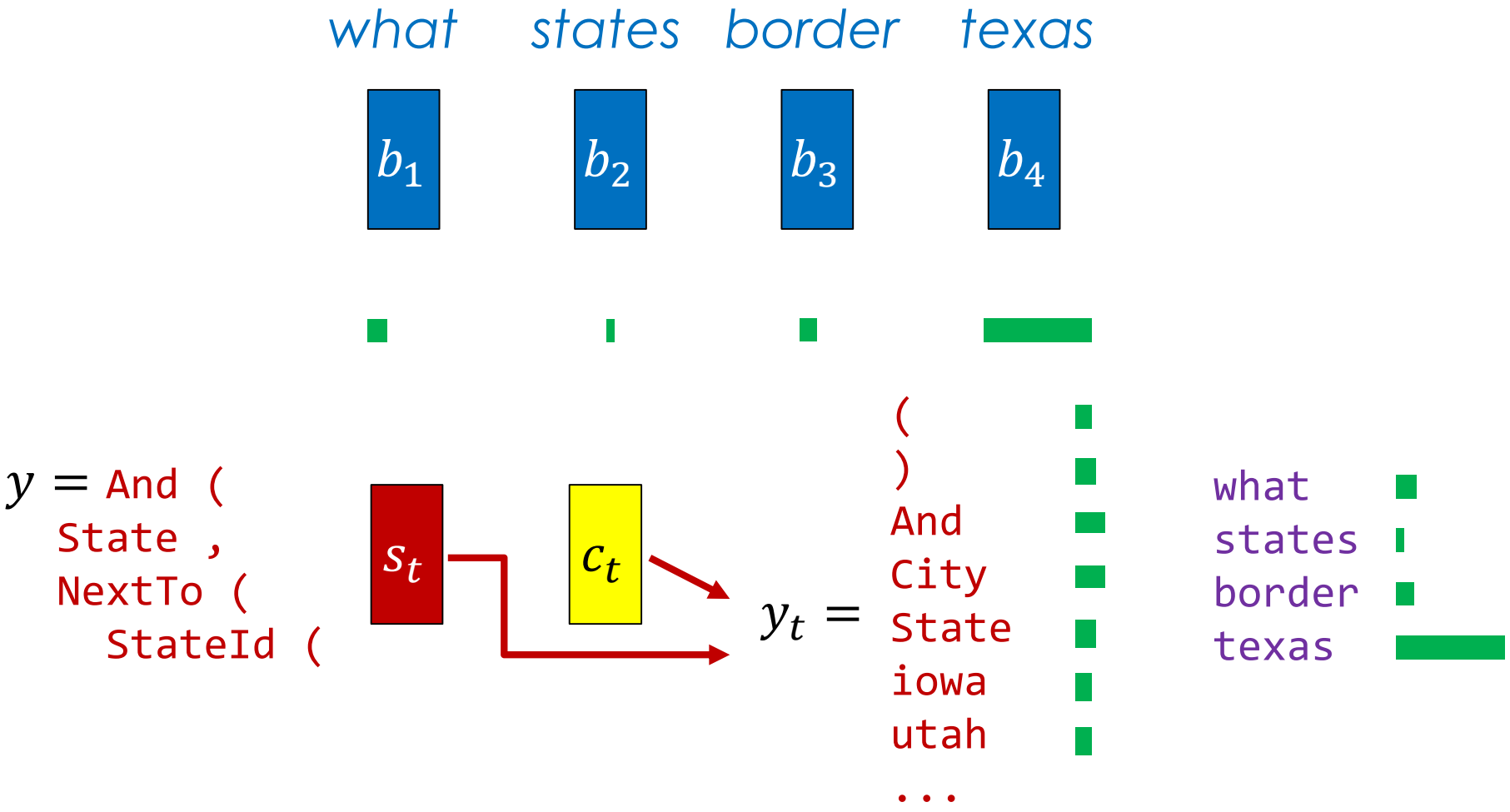
city('alabama','al','birmingham',284413).
city('alabama','al','mobile',200452).
city('alabama','al','montgomery',177857).
city('alabama','al','huntsville',142513).
city('alabama','al','tuscaloosa',75143).
city('alaska','ak','anchorage',174431).
city('arizona','az','phoenix',789704).
city('arizona','az','tucson',330537).
city('arizona','az','mesa',152453).
city('arizona','az','tempe',106919).
city('arizona','az','glendale',96988).
city('arizona','az','scottsdale',88622).
city('arkansas','ar','little rock',158915).
city('arkansas','ar','fort smith',71384).
city('arkansas','ar','north little rock',64388).
city('california','ca','los angeles',2966850).
city('california','ca','san diego',875538).
city('california','ca','san francisco',678974).
city('california','ca','san jose',629442).
city('california','ca','long beach',361334).
city('california','ca','oakland',339337).
city('california','ca','sacramento',275741).
city('california','ca','anaheim',219311).
city('california','ca','fresno',218202).
city('california','ca','santa ana',203713).
city('california','ca','riverside',170876).
city('california','ca','huntington beach',170505).
city('california','ca','stockton',149779).
city('california','ca','glendale',139060).
city('california','ca','fremont',131945).
…

Rare Entities

- Rare entities pose a problem!
  - If we see an entity name we didn't see during training, how can we even generate the right logical form?
- Solution: **Attention-based copying**

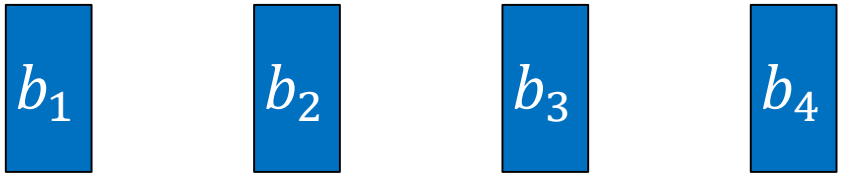
See also Gu et al. (2016), Gulcehre et al. (2016)

Neural Semantic Parser

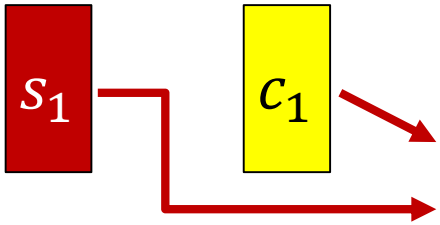


Neural Semantic Parser

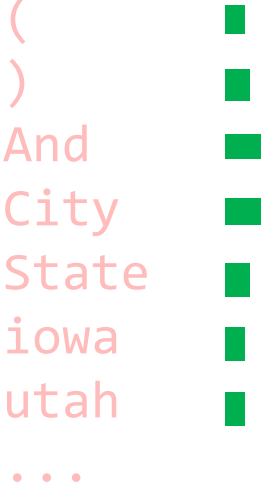
what *states* *border* *texas*



$y =$ And (
State ,
NextTo (
StateId (



$y_1 =$



Implementation Details

- Training: maximize loglikelihood of correct logical form with SGD
- Test time
  - Decode with beam search
  - Add missing parentheses
  - Prune logical forms that result in execution error

Results (Take 1)

System	Geoquery	ATIS	Overnight
Zettlemoyer and Collins (2007)		**84.6**	
Kwiatkowski et al. (2010)	88.9		
Liang et al. (2011)*	**91.1**		
Kwiatkowski et al. (2013)	**89.0**		
Zhao and Huang (2015)	88.9	84.2	
Wang et al. (2015)			58.8
Our RNN Model	85.0	76.3	**75.8**

How can we do better?

* Not directly comparable; used a seed lexicon for predicates.

Outline

- Neural Semantic Parser
- Data Recombination
- Discussion

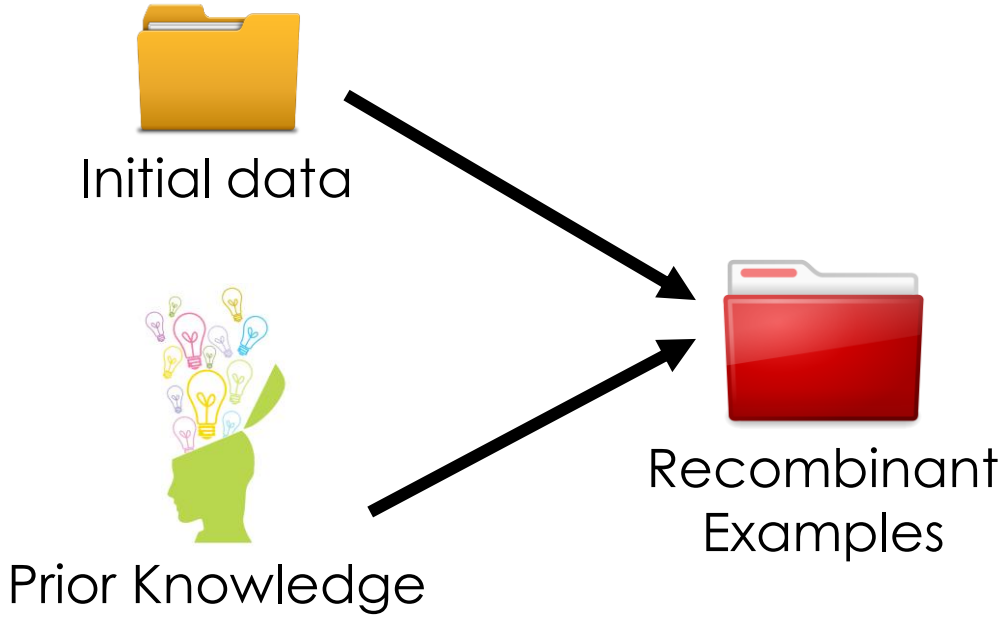
Structural Regularities

- Given
  - *what are the major cities in iowa ?*
    `And(City, Major, LocatedIn(StateId(iowa)))`
- We know how to parse
  - *what are the major cities in* ***texas*** *?*

    `And(City, Major, LocatedIn(StateId(`**`texas`**`)))`

Structural Regularities

- How do we build a neural model that respects compositional structural regularities?



Recombinant Examples

Dataset:

what are the major cities in iowa ?

```
And(City, Major, LocatedIn(StateId( iowa )))
```

what are states that border texas ?

```
And(State, NextTo(StateId( texas )))
```

Recombinant Examples

what are the major cities in iowa ?

`And(City, Major, LocatedIn(StateId(iowa)))`

Step 1: Apply high-precision alignment rules

Recombinant Examples

what are the major cities in iowa ?

```
And(City, Major, LocatedIn(StateId(iowa)))
```

Step 1: Apply high-precision alignment rules

Step 2: Infer types of aligned fragments

Recombinant Examples

ROOT → (*what are the major cities in* STATE *?*,
`And(City, Major, LocatedIn(`STATE`))`)

STATE → (*iowa*, `StateId(iowa)`)

Step 1: Apply high-precision alignment rules

Step 2: Infer types of aligned fragments

Step 3: Generate grammar rules by abstracting aligned fragments

Recombinant Examples

ROOT → (*what are states that border* STATE *?*,
`And(State, NextTo(`STATE`))`)

STATE → (*texas*, `StateId(texas)`)

Step 1: Apply high-precision alignment rules

Step 2: Infer types of aligned fragments

Step 3: Generate grammar rules by abstracting aligned fragments

(Repeat for other examples)

Recombinant Examples

what are the major cities in STATE *?*

`And(City, Major, LocatedIn(`STATE`))`

Step 1: Apply high-precision alignment rules

Step 2: Infer types of aligned fragments

Step 3: Generate grammar rules by abstracting aligned fragments

Step 4: Combine rules from different examples to form new "recombinant" examples

Recombinant Examples

what are the major cities in texas ?

```
And(City, Major, LocatedIn(StateId(texas)))
```

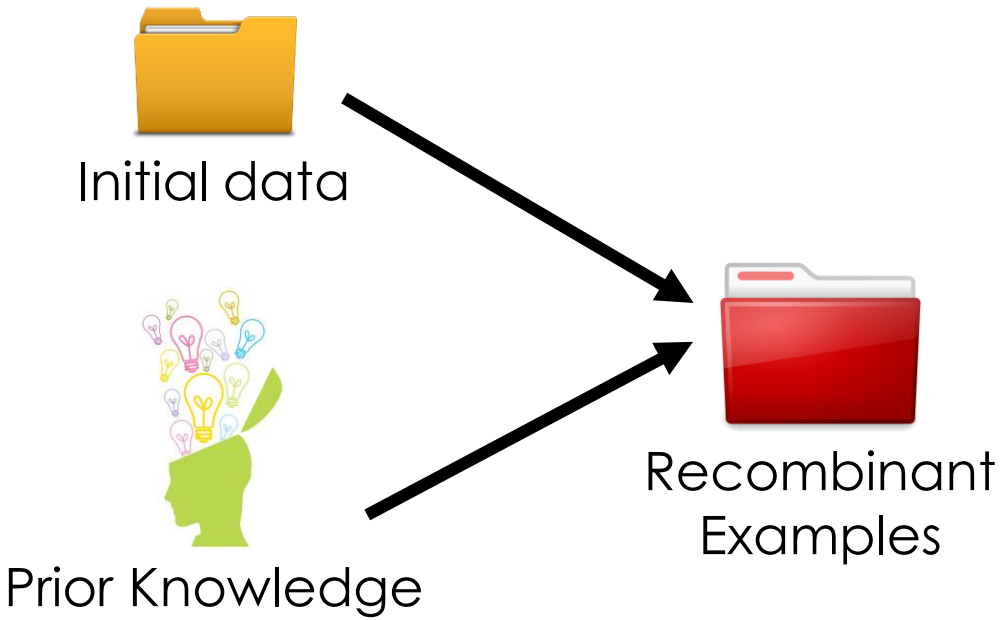
Step 1: Apply high-precision alignment rules

Step 2: Infer types of aligned fragments

Step 3: Generate grammar rules by abstracting aligned fragments

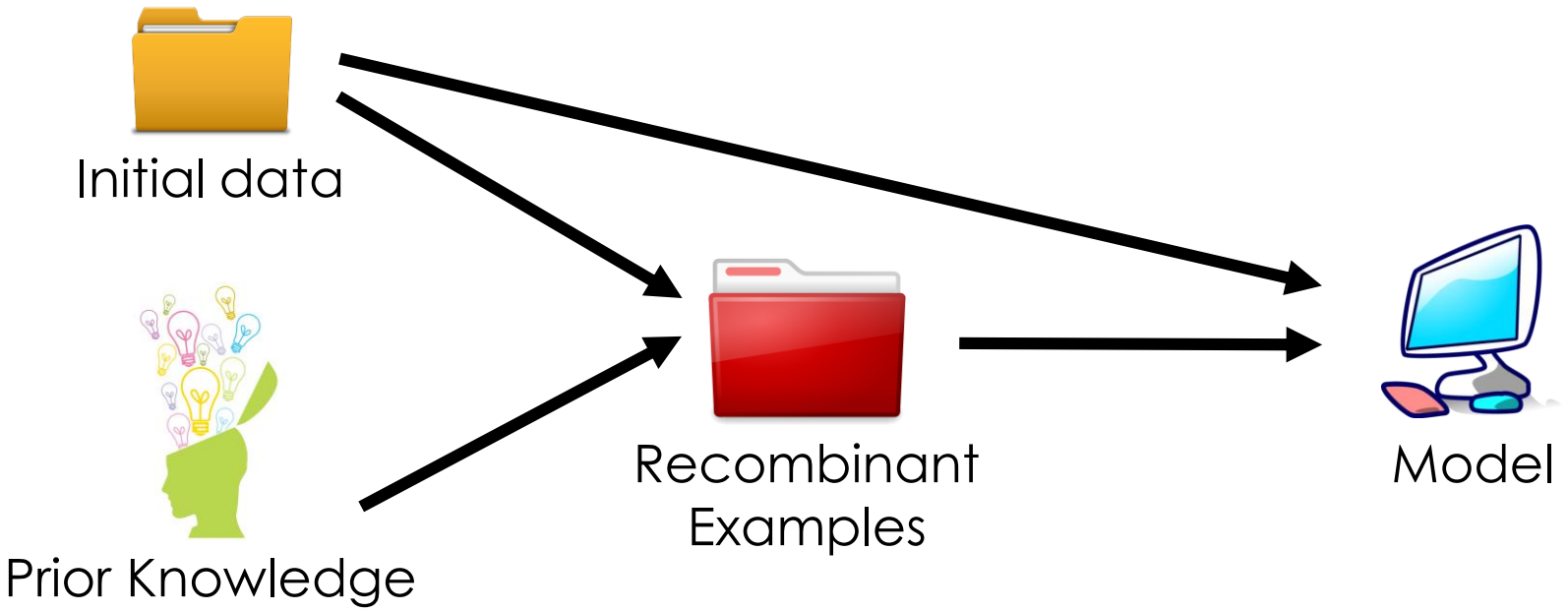
Step 4: Combine rules from different examples to form new "recombinant" examples

Data Recombination



Step 1: **Generate recombinant examples** using training data and prior knowledge about domain

Data Recombination



Step 2: **Train on these examples** to increase model's awareness of task structure

Training the Model

- At each epoch, sample new recombinant examples from grammar
  - Choose production rules uniformly at random
- Train on recombinant examples plus original training examples in 1:1 ratio

Results (Take 2)

System	Geoquery	ATIS	Overnight
Zettlemoyer and Collins (2007)		**84.6**	
Kwiatkowski et al. (2010)	88.9		
Liang et al. (2011)*	**91.1**		
Kwiatkowski et al. (2013)	**89.0**		
Zhao and Huang (2015)	88.9	84.2	
Wang et al. (2015)			58.8
Our RNN Model	85.0	76.3	**75.8**
+ Abstracting Entities	85.4	79.9	75.3

* Not directly comparable; used a seed lexicon for predicates.

Structural Regularities, Part 2

- Given
  - *what are the major cities in iowa ?*
    `And(City, Major, LocatedIn(StateId(iowa)))`
  - *what are states that border texas ?*
    `And(State, NextTo(StateId(texas)))`
- We know how to parse
  - *what are the major cities in*
    states that border texas *?*

```
And(City, Major, LocatedIn(
    And(State, NextTo(StateId(texas)))))
```

Abstracting Whole Phrases

What are states that border texas ?

```
And(State, NextTo(StateId(texas)))
```

Step 1: Apply high-precision alignment rules

Abstracting Whole Phrases

What are states that border texas ?

```
And(State, NextTo(StateId(texas)))
```

Step 1: Apply high-precision alignment rules

Step 2: Infer types of aligned fragments

Abstracting Whole Phrases

What are states that border texas ?

```
And(State, NextTo(StateId(texas)))
```

Step 1: Apply high-precision alignment rules

Step 2: Infer types of aligned fragments

New Grammar Rule

- STATE → (*states that border texas*,
  `And(State, NextTo(StateId(texas)))`)

Results (Take 3)

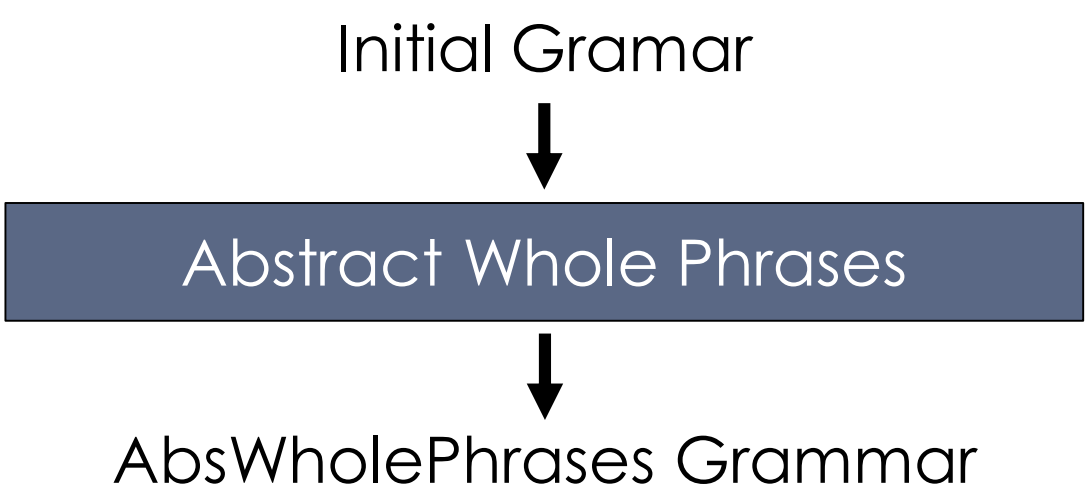
System	Geoquery	ATIS	Overnight
Zettlemoyer and Collins (2007)		**84.6**	
Kwiatkowski et al. (2010)	88.9		
Liang et al. (2011)*	**91.1**		
Kwiatkowski et al. (2013)	**89.0**		
Zhao and Huang (2015)	88.9	84.2	
Wang et al. (2015)			58.8
Our RNN Model	85.0	76.3	75.8
+ AbsEntities	85.4	79.9	75.3
+ AbsWholePhrases	87.5	---	**75.9**

* Not directly comparable; used a seed lexicon for predicates.

Composition of Strategies

- Grammar induction strategies are functions on grammars
  - Process each rule in grammar independently to generate new grammar rules
- Initial grammar
  - ROOT → (*what are the major cities in iowa ?*, `And(City, Major, LocatedIn(StateId(iowa)))`)
  - ROOT → (*what are states that border texas ?*, `And(State, NextTo(StateId(texas)))`)

Composition of Strategies



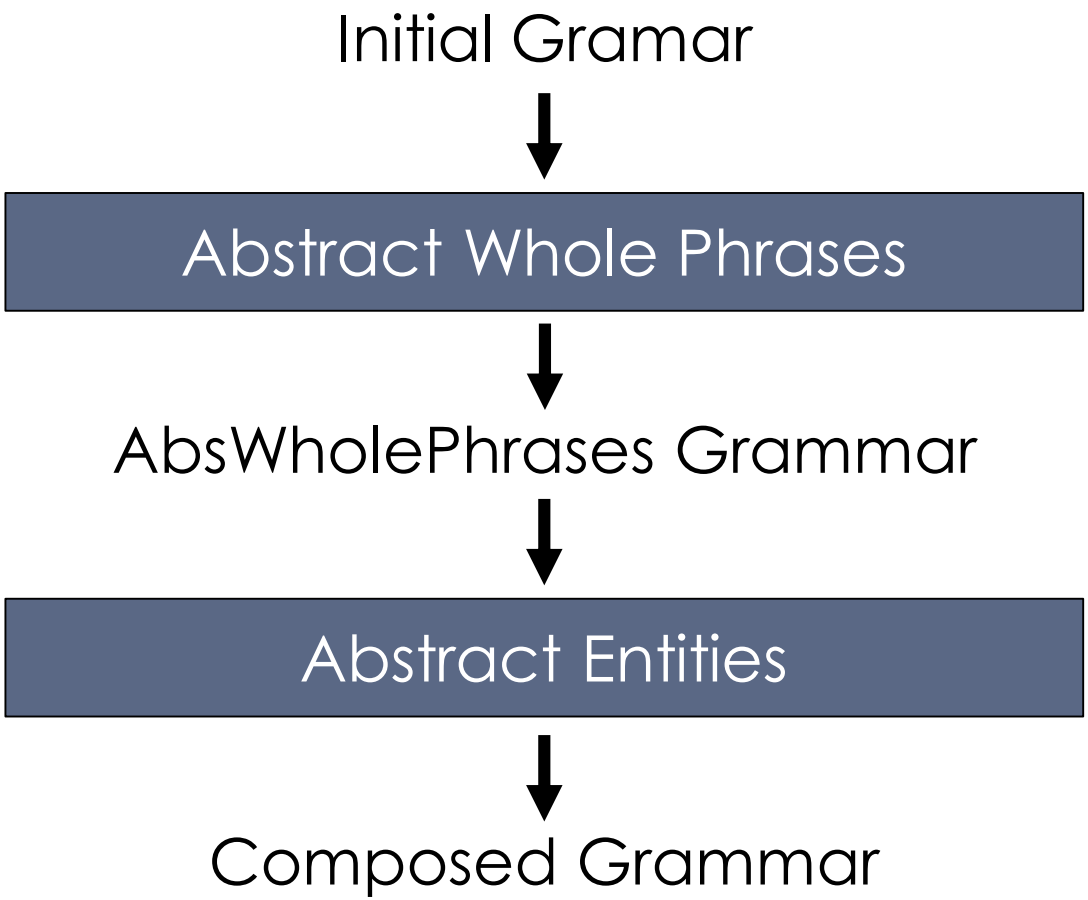
Composition of Strategies

STATE →

(*states that border texas* ,

`And(State, NextTo(StateId(texas)))`)

Composition of Strategies



Results (Take 4)

System	Geoquery	ATIS	Overnight
Zettlemoyer and Collins (2007)		**84.6**	
Kwiatkowski et al. (2010)	88.9		
Liang et al. (2011)*	**91.1**		
Kwiatkowski et al. (2013)	**89.0**		
Zhao and Huang (2015)	88.9	84.2	
Wang et al. (2015)			58.8
Our RNN Model	85.0	76.3	75.8
+ AbsEntities	85.4	79.9	75.3
+ AbsWholePhrases	87.5	---	**75.9**
+ AbsWholePhrases, AbsEntities	88.9	---	75.3

* Not directly comparable; used a seed lexicon for predicates.

What else can we do?

- What happens if we artificially make longer examples?

Concatenation

- Given
  - *what are the major cities in iowa ?*
    `And(City, Major, LocatedIn(StateId(iowa)))`
  - *what are states that border texas ?*
    `And(State, NextTo(StateId(texas)))`
- Create new example
  - *what are the major cities in iowa ? </s> what are states that border texas ?*

    `And(City, Major, LocatedIn(StateId(iowa))) </s> And(State, NextTo(StateId(texas)))`

Concatenation

- Can cast as grammar induction
- Compose with previous grammar induction strategies

Final Results

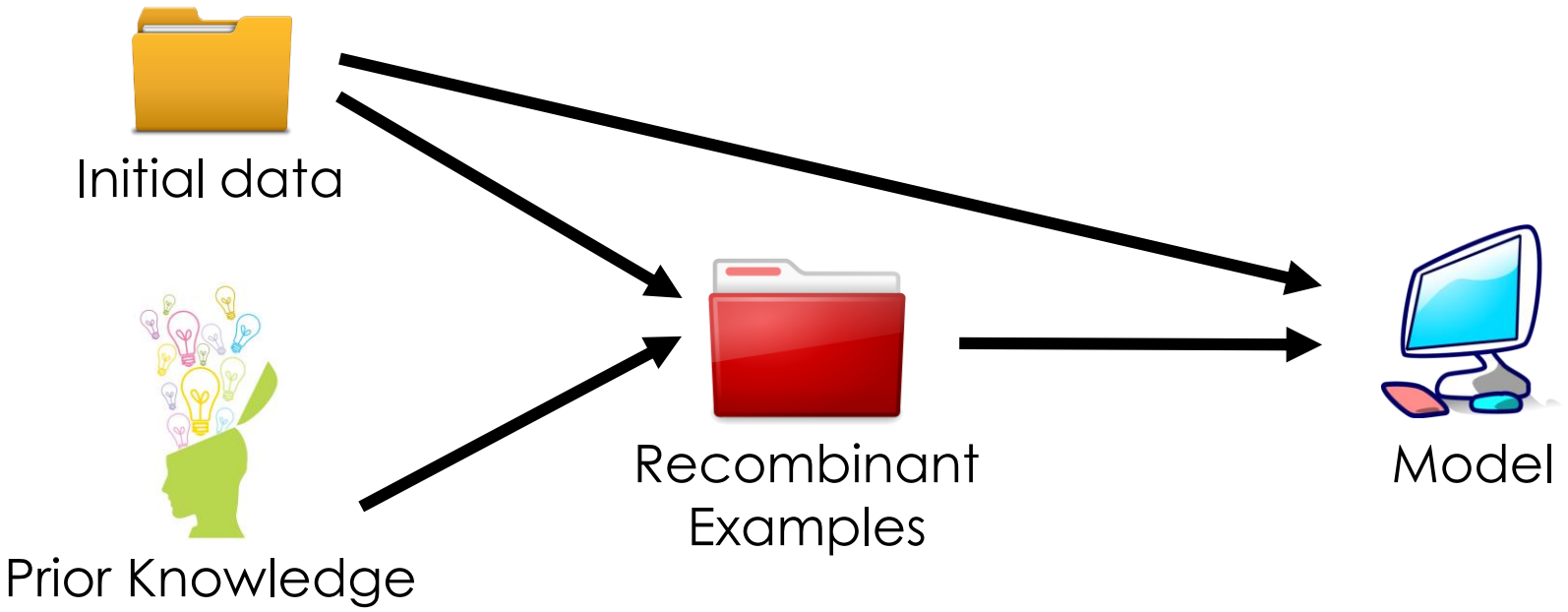
System	Geoquery	ATIS	Overnight
Zettlemoyer and Collins (2007)		**84.6**	
Kwiatkowski et al. (2010)	88.9		
Liang et al. (2011)*	**91.1**		
Kwiatkowski et al. (2013)	89.0		
Zhao and Huang (2015)	88.9	84.2	
Wang et al. (2015)			58.8
Our RNN Model	85.0	76.3	75.8
+ AbsEntities	85.4	79.9	75.3
+ AbsWholePhrases	87.5	---	75.9
+ AbsWholePhrases, AbsEntities	88.9	---	75.3
+ AbsWholePhrases, AbsEntities, Concat	**89.3**	83.3	**77.5**

* Not directly comparable; used a seed lexicon for predicates.

Outline

- Neural Semantic Parser
- Data Recombination
- Discussion

Data Recombination



Step 1: **Generate recombinant examples** using training data and prior knowledge about domain

Data Recombination

- Builds a **generative model** from the data
  - Mixture of the empirical distribution and induced probabilistic grammar
  - Train on examples (x, y) sampled from this model

Data Augmentation

- Apply local transformations to a single example to generate more examples

$x =$ 

$y =$ "puppy"

Krizhevsky et al. (2012).

Data Augmentation

- Apply local transformations to a single example to generate more examples

$x =$ 

$y =$ "puppy"

Horizontal Reflection

Krizhevsky et al. (2012).

Data Augmentation

- Apply local transformations to a single example to generate more examples

$x =$ 

$y =$ "puppy"

Cropping

Krizhevsky et al. (2012).

Data Recombination

- Get a richer generative model by combining information from **multiple examples**

what are the major cities in iowa ?

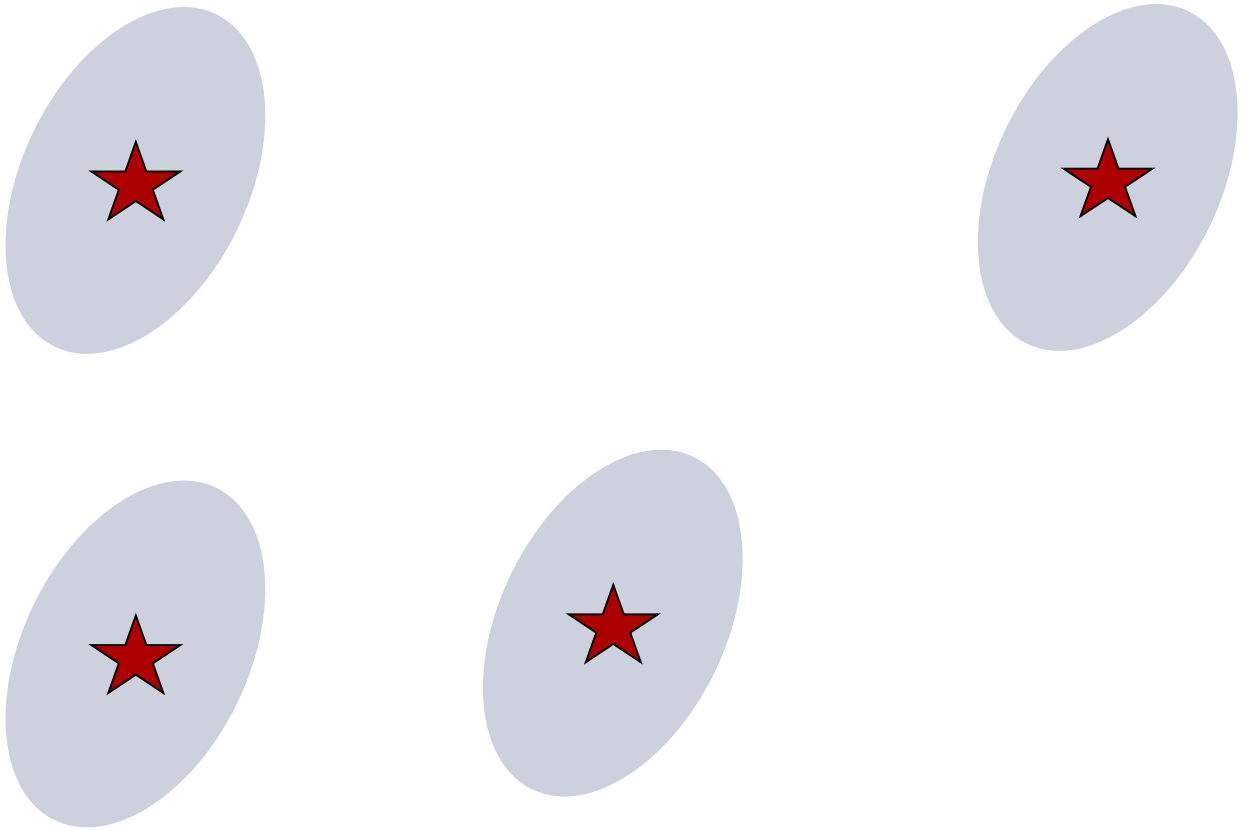
```
And(City, Major, LocatedIn(StateId( iowa )))
```

what are states that border texas ?

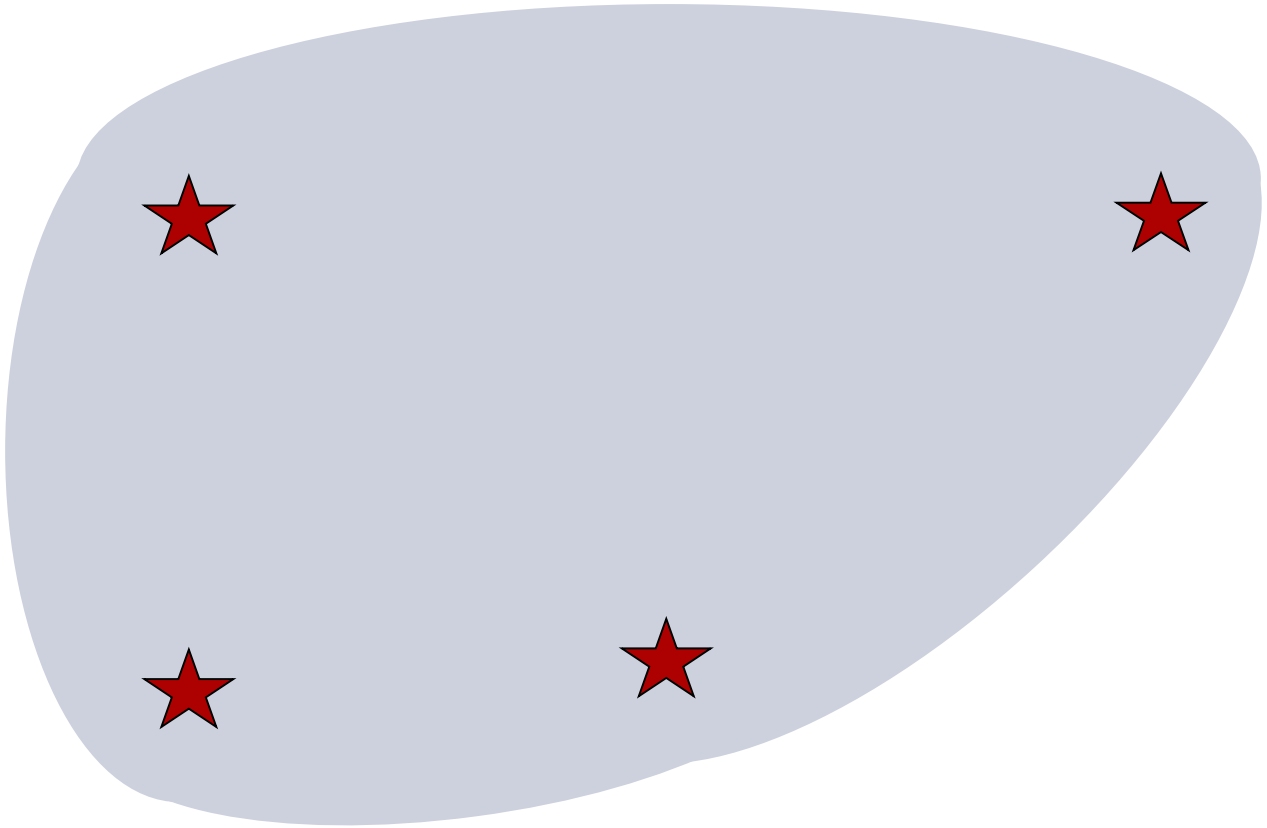
```
And(State, NextTo(StateId( texas )))
```

Empirical Distribution

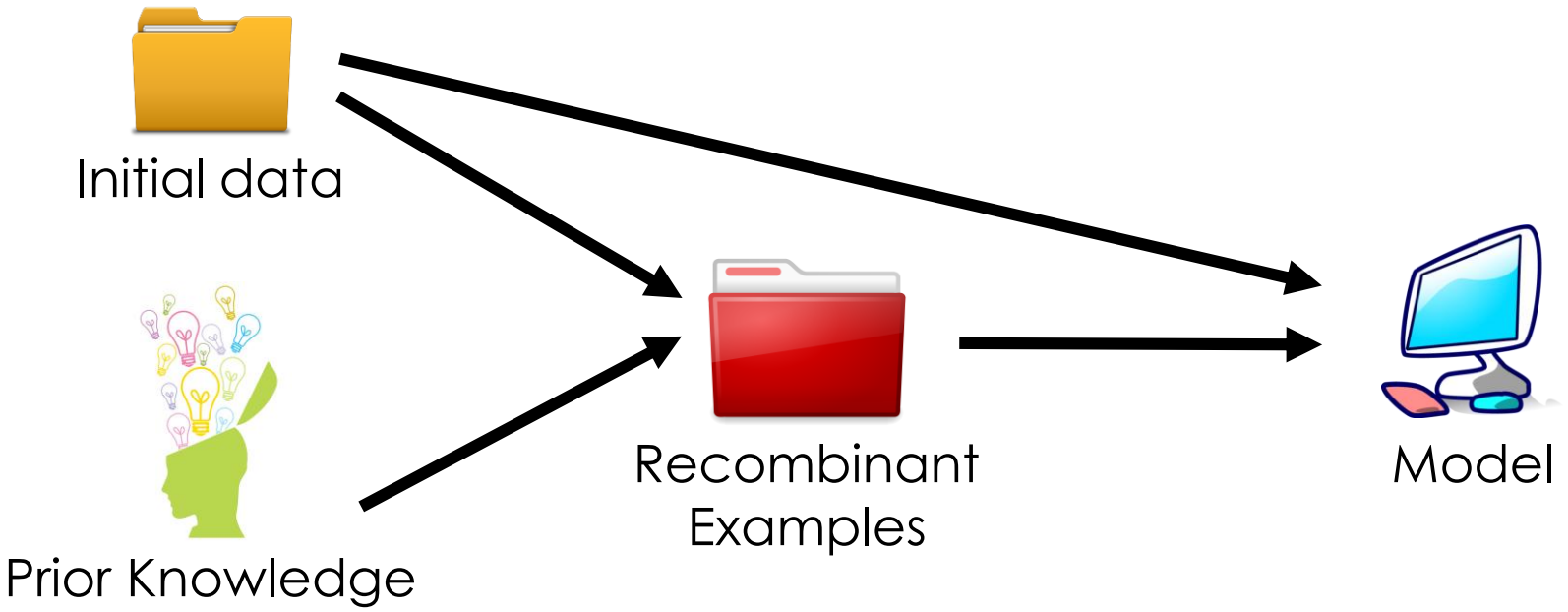
Data Augmentation



Data Recombination



Thank you!



Code, data, and experiments available on CodaLab

https://worksheets.codalab.org/worksheets/0x50757a37779b485f89012e4ba03b6f4f/

Composition of Strategies

STATE →

(*states that border texas* ,

`And(State, NextTo(StateId(texas)))`)

New Grammar Rule

- STATE → (*states that border* STATEID,
  `And(State, NextTo(`STATEID`))`)

An Artificial Experiment

- Make up some simple artificial data
  - *friends of relatives of alice*
    `Friend(Relative(alice))`
  - *colleagues of brothers of bob*
    `Colleague(Brother(bob))`
- Generate recombinant examples to train on
  - **Same length**: only swap entities
    - *friends of relatives of* ***bob***
      `Friend(Relative(bob))`
  - **Longer**: nest whole phrases
    - *friends of relatives of* ***colleagues of brothers of*** *alice*
      `Friend(Relative(Colleague(Brother(alice))))`

An Artificial Experiment

- Test only on "short" examples like those in the original training data
  - *siblings of roommates of eve*
    `Sibling(Roommate(eve))`

An Artificial Experiment