Learning Distributions over Logical Forms for Referring Expression Generation

Nicholas FitzGerald Yoav Artzi Luke Zettlemoyer

University of Washington Computer Science & Engineering

ISTC PC

Referring Expressions

Photo: pwenzel on Flikr

Referring Expressions

Photo: pwenzel on Flikr

Referring Expressions

Photo: pwenzel on Flikr

Referring Expressions

Photo: pwenzel on Flikr

Referring Expressions

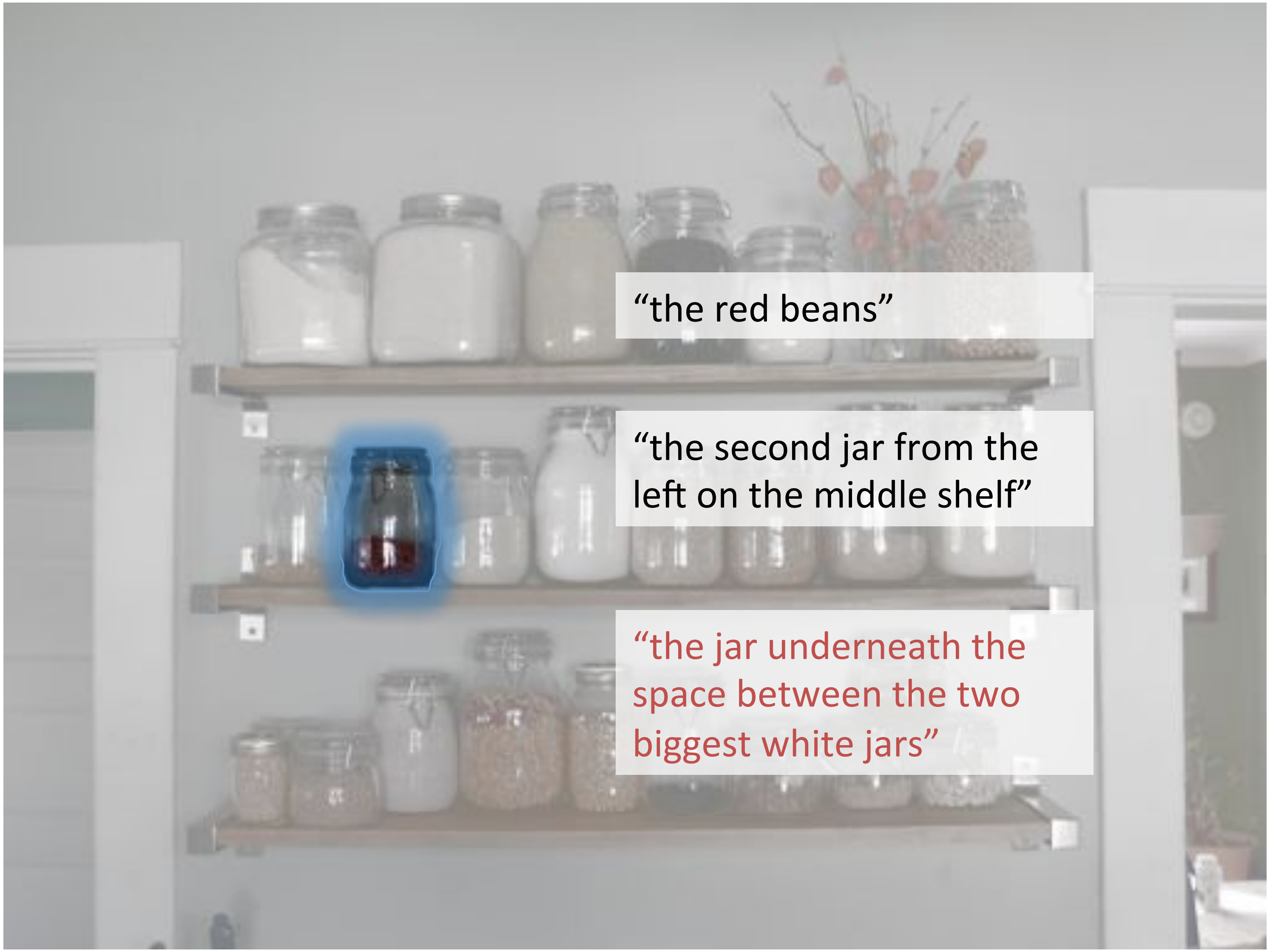


Photo: pwenzel on Flikr

Grounded Language Problems

Physical Scene

Language

``The green and red balls.''

Grounded Language Problems

Physical Scene

Logical Form (LF)

$$\iota x.(green(x) \vee red(x)) \wedge sphere(x)$$

Language

``The green and red balls.''

Grounded Language Problems

Physical Scene

Logical Form (LF)

$$\iota x.(green(x) \vee red(x)) \wedge sphere(x)$$

Language

``The green and red balls.''

Grounded Language Problems

Physical Scene

Logical Form (LF)

$$\iota x.(green(x) \vee red(x)) \wedge sphere(x)$$

Language

``The green and red spheres.''

Grounded Language Problems

Grounded Language Understanding

[Matuszek et. al 2012], [Krishnamurthy et. al 2013]

Physical Scene

Logical Form (LF)

$$\iota x.(green(x) \vee red(x)) \wedge sphere(x)$$

Language

``The green and red spheres.''

Grounded Language Problems

Grounded Language Understanding
[Matuszek et. al 2012], [Krishnamurthy et. al 2013]

Physical Scene

Logical Form (LF)

$$\iota x.(green(x) \vee red(x)) \wedge sphere(x)$$

Language

``The green and red spheres.''

Realization from LF
[White and Rajkumar 2009], [Lu and Ng 2011]

Grounded Language Problems

Grounded Language Understanding
[Matuszek et. al 2012], [Krishnamurthy et. al 2013]

Physical Scene

Logical Form (LF)

$$\iota x.(green(x) \vee red(x)) \wedge sphere(x)$$

Language

``The green and red spheres.''

LF Generation
Focus of this talk:
Generate Distribution

Realization from LF
[White and Rajkumar 2009], [Lu and Ng 2011]

Goal: Generate Distributions over LFs

$\iota x.green(x) \wedge sphere(x) \cup \iota x.red(x) \wedge sphere(x)$ → "The green ball and the red ball."

$\iota x.(green(x) \vee red(x)) \wedge sphere(x)$ → "The green and red spheres." / "The green and red balls."

$\iota x.apple(x) \cup \iota x.pear(x)$ → "The apple and the pear."

$\ldots$

Model how people refer to objects

- Many different expressions for each referent
- Some are more likely to be used in practice
- We need to learn a probability distribution

Goal: Generate Distributions over LFs

P(exp)

0.1 $\iota x.green(x) \wedge sphere(x) \cup \iota x.red(x) \wedge sphere(x)$ → "The green ball and the red ball."

0.3 $\iota x.(green(x) \vee red(x)) \wedge sphere(x)$ → "The green and red spheres." → "The green and red balls."

0.2 $\iota x.apple(x) \cup \iota x.pear(x)$ → "The apple and the pear."

0.4 . . .

Model how people refer to objects

- Many different expressions for each referent
- Some are more likely to be used in practice
- We need to learn a probability distribution

Goal: Generate Distributions over LFs

Input:

Output:

0.3	$\iota x.green(x) \wedge sphere(x)$ $\cup \iota x.red(x) \wedge sphere(x)$
0.2	$\iota x.(green(x) \vee red(x))$ $\wedge sphere(x)$
0.1	$\iota x.apple(x) \cup \iota x.pear(x)$
. . .	. . .

Several advantages

- Natural variation for generation
- Useful prior for understanding systems

Learn from Labeled Examples

{  ,

"The green, red, orange and yellow toys"
"The green, red, yellow, and orange objects"
"All the pieces that are not blue or brown"
"All items that are not brown or blue"
"Everything that is not brown or blue"
..... }

{  ,

"The red and green balls."
"The red and greed spheres."
"The pear and the apple"
"The red ball and the green pear"
"All the balls except the yellow one"
...... }

Lots of variation in practice

- Collected 20 sentences per scene
- Mean of 6 unique logical forms per scene
- Max of 13

Learn from Labeled Examples

{  ,

0.4 $\iota x.(red(x) \vee green(x) \vee orange(x) \vee yellow(x)) \wedge obj(x)$

0.3 $\mathcal{E}x.obj(x) \setminus (\iota x.brown(x) \wedge triangle(x) \cup (\iota x.blue(x) \wedge lego(x))$

. . . }

{  ,

0.1 $\iota x.green(x) \wedge sphere(x) \cup \iota x.red(x) \wedge sphere(x)$

0.3 $\iota x.(green(x) \vee red(x)) \wedge sphere(x)$

0.2 $\iota x.apple(x) \cup \iota x.pear(x)$

. . . }

Lots of variation in practice

- Collected 20 sentences per scene
- Mean of 6 unique logical forms per scene
- Max of 13

Overview

Space of Referring Expressions

Probabilistic Model

Learning

Experiments

Results

Conclusion

Overview

Space of Referring Expressions

- Semantic Modeling
- Enumerating Referring Expressions

Probabilistic Model

Learning

Experiments

Results

Conclusion

Semantic Modeling

- Simply-typed lambda calculus
  - [Steedman 1996], [Carpenter 1997], [Steedman 2011]
- Extended to model set reference
  - Capture distinctions present in data
  - As simple as possible

Semantic Modeling

Two Simple Types:

e : Sets of Objects

t : [True, False]

Semantic Modeling

Attributes

$$< e, t >$$

$$\lambda x.blue(x)$$

Semantic Modeling

Attributes

$$< e, t >$$

$$\lambda x.triangle(x)$$

Semantic Modeling

Attributes

Logical Operators

$$< e, t >$$

$$\lambda x . \neg blue(x)$$

Semantic Modeling

Attributes

Logical Operators

$$< e, t >$$

$$\lambda x.blue(x) \wedge triangle(x)$$

Semantic Modeling

Attributes

Logical Operators

$$< e, t >$$

$$\lambda x.blue(x) \vee triangle(x)$$

Semantic Modeling

Attribute Coordination

Logical Operators

Determiners

$$<< e, t >, e >$$

$$\iota x.triangle(x)$$

Semantic Modeling

Attribute Coordination

Logical Operators

Determiners

$$<< e, t >, e >$$

$$[\iota, \mathcal{E}, \mathcal{A}]$$

$$\mathcal{E} \approx \forall$$

$$\mathcal{A} \approx \exists$$

Semantic Modeling

Attribute Coordination

Logical Operators

Determiners

Set Coordination

$$<< e, e >, e >$$

$$\iota x.triangle(x)$$
$$\cup \iota x.apple(x)$$

Semantic Modeling

Attribute Coordination

Logical Operators

Determiners

Set Coordination

$$<< e, e >, e >$$

$$\iota x.triangle(x)$$
$$\setminus \iota x.blue(x)$$
$$\wedge triangle(x)$$

Semantic Modeling

Attribute Coordination

Logical Operators

Determiners

Set Coordination

Plurality

Cardinality

Semantic Modeling

- Simply-typed lambda calculus
  - [Steedman 1996], [Carpenter 1997], [Steedman 2011]
- Extended to model set reference
  - Capture distinctions present in data
  - As simple as possible
- Two new contributions:
  - Sets as a primitive type (plurals)
  - Coordination

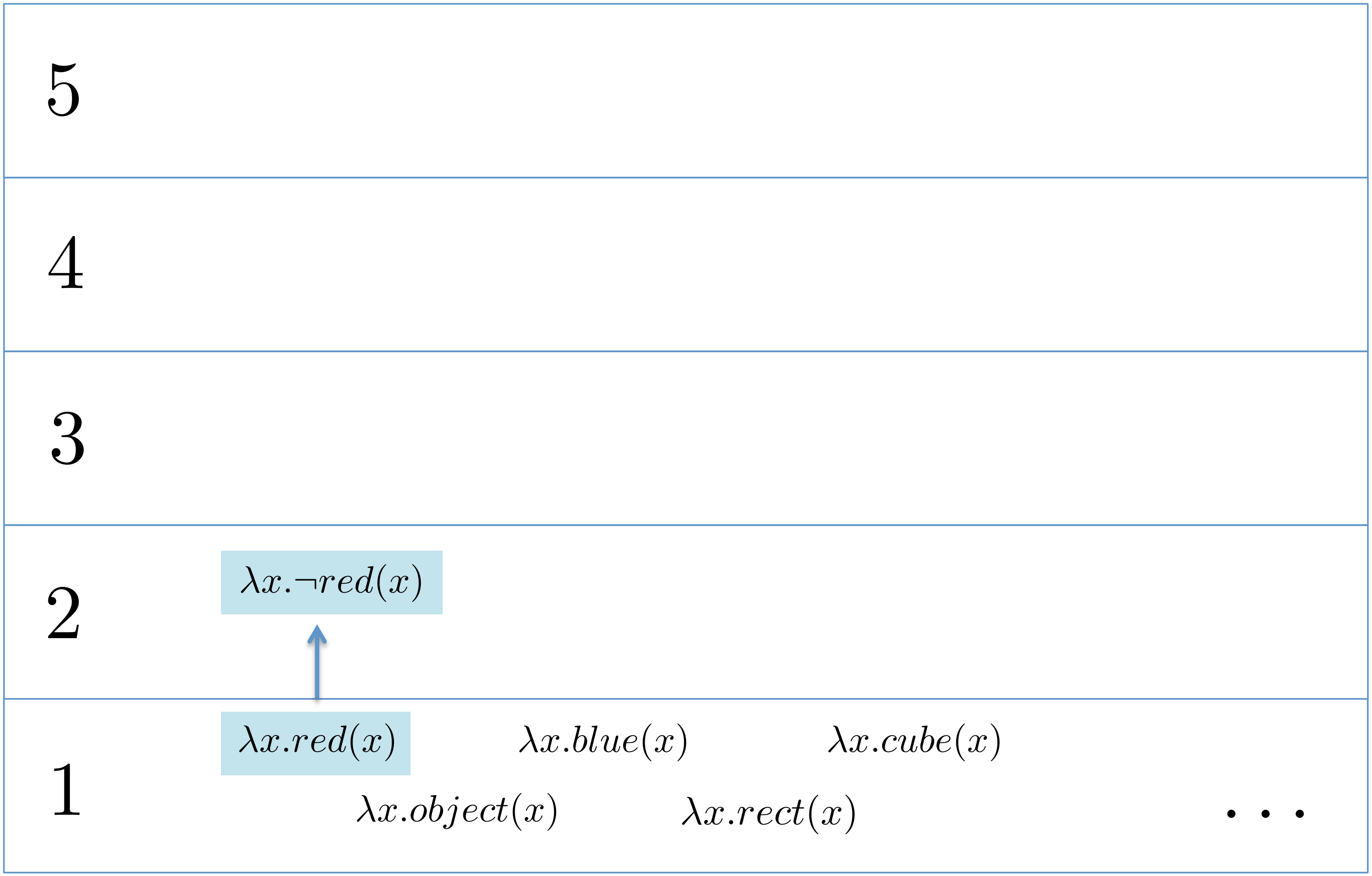
Enumerating Logical Forms

- Enumerate candidate Logical Forms
- Problem:
  - Infinite in general
- Goal:
  - finite set with good empirical coverage
  - strategy for enumerating

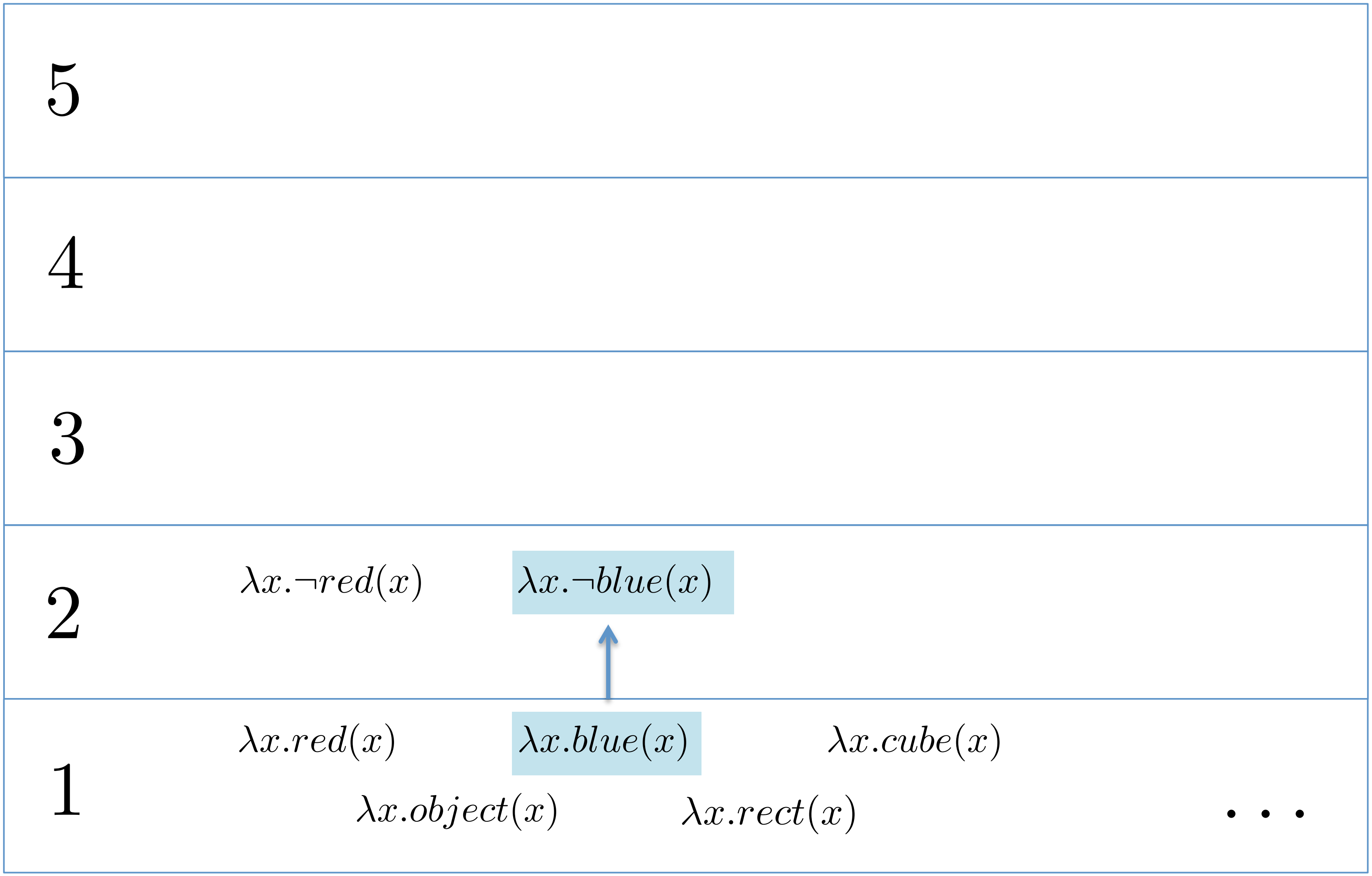
Enumerating Logical Forms

5	
4	
3	
2	
1	$\lambda x.red(x)$ $\lambda x.blue(x)$ $\lambda x.cube(x)$ $\lambda x.object(x)$ $\lambda x.rect(x)$ $\ldots$

Enumerating Logical Forms



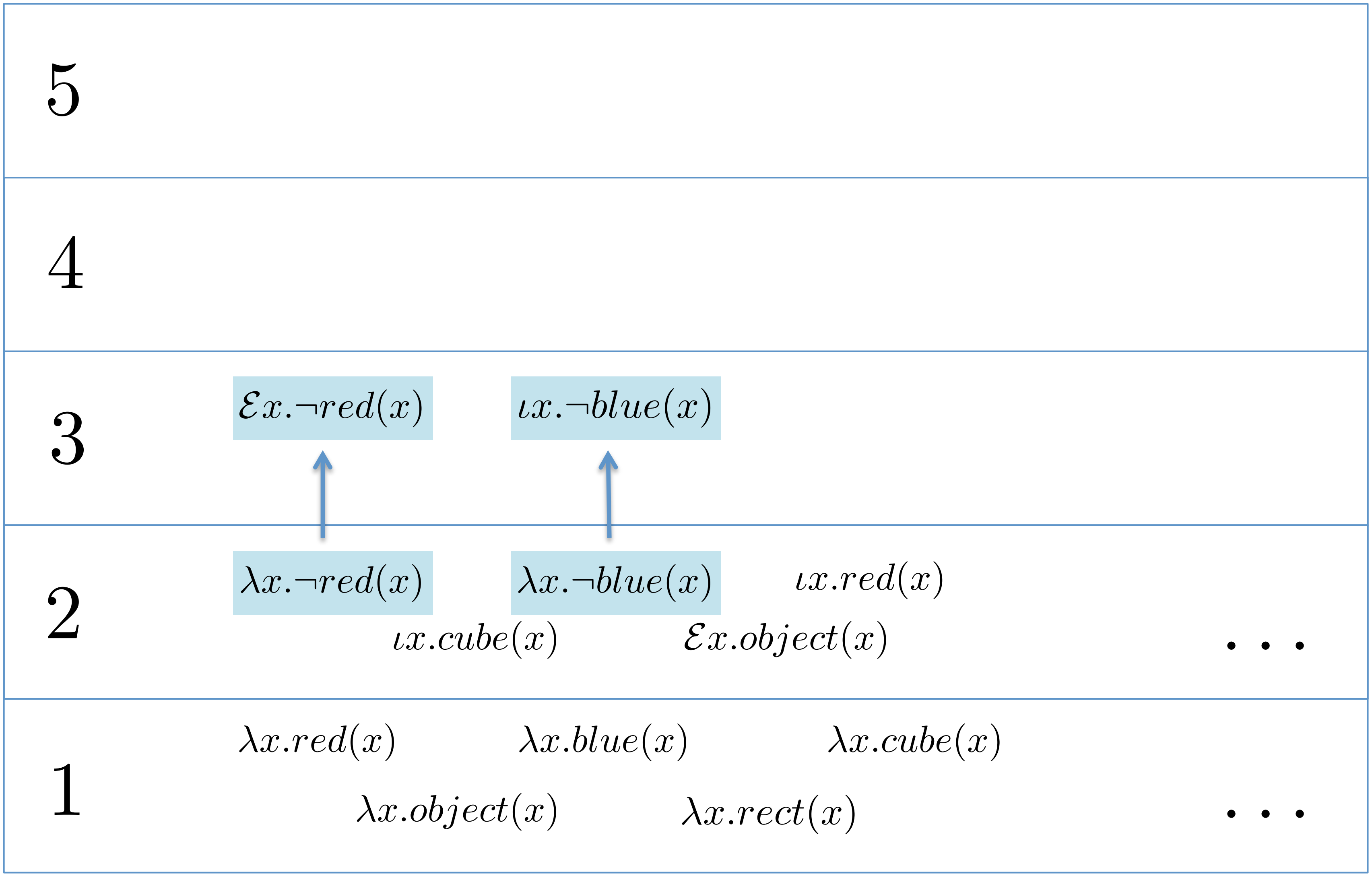
Enumerating Logical Forms



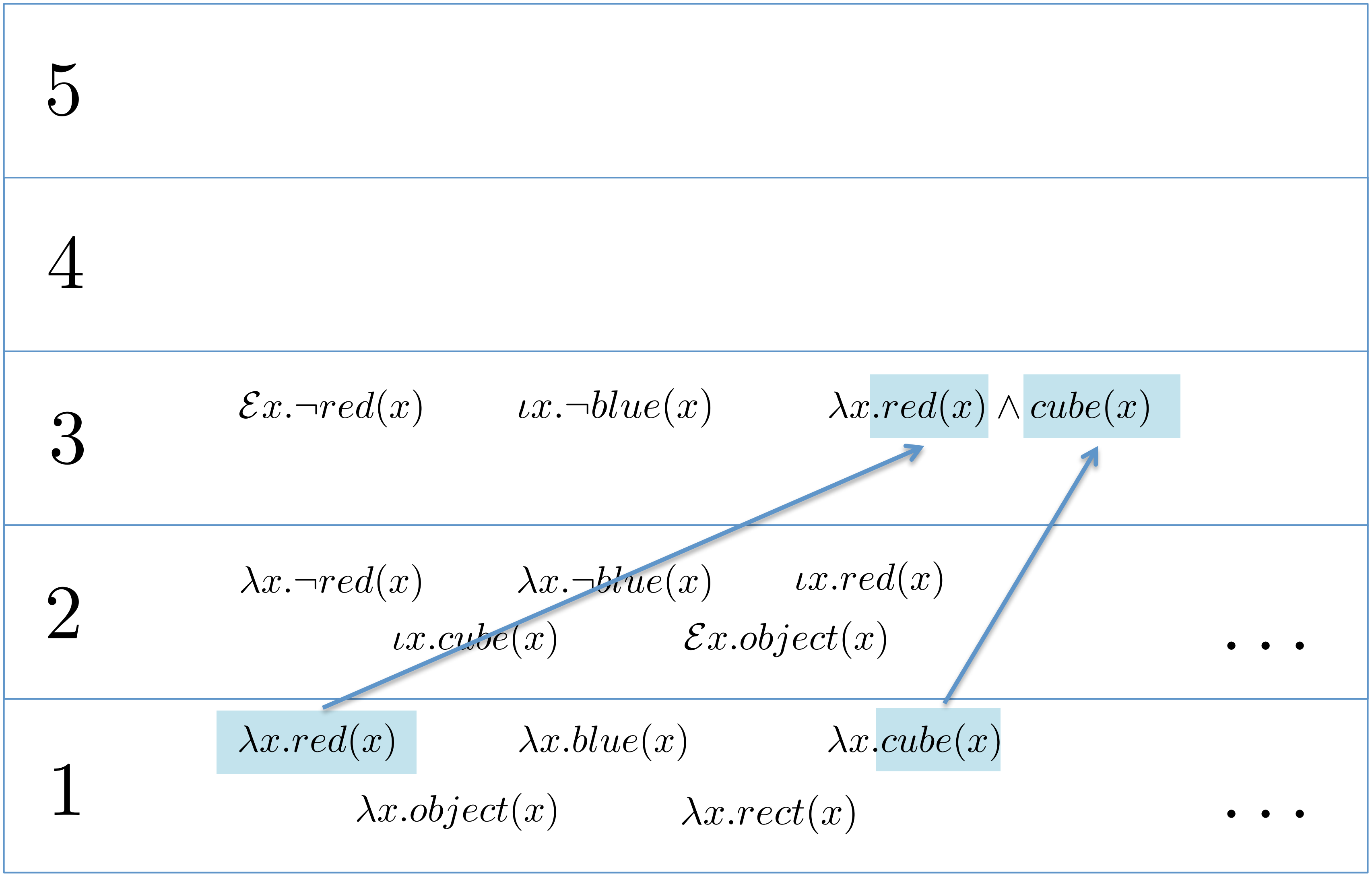
Enumerating Logical Forms

5	
4	
3	
2	$\lambda x.\neg red(x)$ $\lambda x.\neg blue(x)$ $\iota x.red(x)$ $\iota x.cube(x)$ $\mathcal{E} x.object(x)$ $\ldots$
1	$\lambda x.red(x)$ $\lambda x.blue(x)$ $\lambda x.cube(x)$ $\lambda x.object(x)$ $\lambda x.rect(x)$ $\ldots$

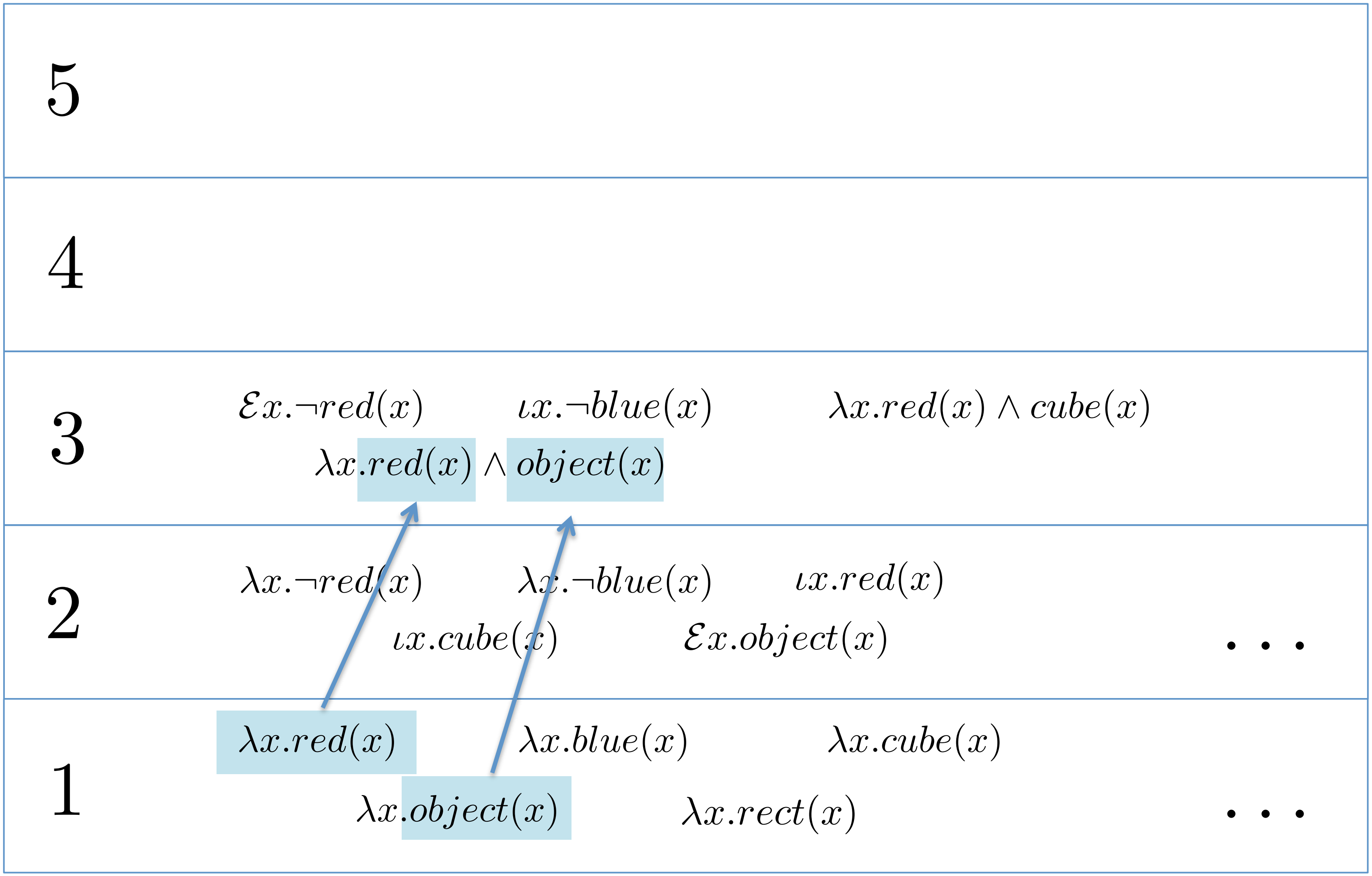
Enumerating Logical Forms



Enumerating Logical Forms



Enumerating Logical Forms



Enumerating Logical Forms

5	
4	
3	$\mathcal{E}x.\neg red(x)$ $\iota x.\neg blue(x)$ $\lambda x.red(x) \wedge cube(x)$ $\lambda x.red(x) \wedge object(x)$ $\lambda x.red(x) \vee cube(x)$ $\ldots$
2	$\lambda x.\neg red(x)$ $\lambda x.\neg blue(x)$ $\iota x.red(x)$ $\iota x.cube(x)$ $\mathcal{E}x.object(x)$ $\ldots$
1	$\lambda x.red(x)$ $\lambda x.blue(x)$ $\lambda x.cube(x)$ $\lambda x.object(x)$ $\lambda x.rect(x)$ $\ldots$

Enumerating Logical Forms

5	
4	$\iota x.red(x) \wedge object(x)$ $\mathcal{A}x.red(x) \vee object(x)$ $\lambda x.\neg cube(x) \wedge object(x)$ $\lambda x.\neg(cube(x) \vee object(x))$ $\ldots$
3	$\mathcal{E}x.\neg red(x)$ $\iota x.\neg blue(x)$ $\lambda x.red(x) \wedge cube(x)$ $\lambda x.red(x) \wedge object(x)$ $\lambda x.red(x) \vee cube(x)$ $\ldots$
2	$\lambda x.\neg red(x)$ $\lambda x.\neg blue(x)$ $\iota x.red(x)$ $\iota x.cube(x)$ $\mathcal{E}x.object(x)$ $\ldots$
1	$\lambda x.red(x)$ $\lambda x.blue(x)$ $\lambda x.cube(x)$ $\lambda x.object(x)$ $\lambda x.rect(x)$ $\ldots$

Enumerating Logical Forms

5	$\iota x.red(x) \cup \iota x.cube(x)$ $\quad$ $\mathcal{E}x.object(x) \setminus \iota x.cube(x)$ $\quad$ $\lambda x.object(x) \wedge equal(x, \mathcal{A}y.cube(y))$ $\quad \ldots$
4	$\iota x.red(x) \wedge object(x)$ $\quad$ $\mathcal{A}x.red(x) \vee object(x)$ $\quad$ $\lambda x.\neg cube(x) \wedge object(x)$ $\quad$ $\lambda x.\neg(cube(x) \vee object(x))$ $\quad \ldots$
3	$\mathcal{E}x.\neg red(x)$ $\quad$ $\iota x.\neg blue(x)$ $\quad$ $\lambda x.red(x) \wedge cube(x)$ $\quad$ $\lambda x.red(x) \wedge object(x)$ $\quad$ $\lambda x.red(x) \vee cube(x)$ $\quad \ldots$
2	$\lambda x.\neg red(x)$ $\quad$ $\lambda x.\neg blue(x)$ $\quad$ $\iota x.red(x)$ $\quad$ $\iota x.cube(x)$ $\quad$ $\mathcal{E}x.object(x)$ $\quad \ldots$
1	$\lambda x.red(x)$ $\quad$ $\lambda x.blue(x)$ $\quad$ $\lambda x.cube(x)$ $\quad$ $\lambda x.object(x)$ $\quad$ $\lambda x.rect(x)$ $\quad \ldots$

M Enumerating Logical Forms

5	$\iota x.red(x) \cup \iota x.cube(x)$ $\mathcal{E}x.object(x) \setminus \iota x.cube(x)$ $\lambda x.object(x) \land equal(x, \mathcal{A}y.cube(y))$ ...
4	$\iota x.red(x) \land object(x)$ $\mathcal{A}x.red(x) \lor object(x)$ $\lambda x.\neg cube(x) \land object(x)$ $\lambda x.\neg(cube(x) \lor object(x))$ ...
3	$\mathcal{E}x.\neg red(x)$ $\iota x.\neg blue(x)$ $\lambda x.red(x) \land cube(x)$ $\lambda x.red(x) \land object(x)$ $\lambda x.red(x) \lor cube(x)$ ...
2	$\lambda x.\neg red(x)$ $\lambda x.\neg blue(x)$ $\iota x.red(x)$ $\iota x.cube(x)$ $\mathcal{E}x.object(x)$ ...
1	$\lambda x.red(x)$ $\lambda x.blue(x)$ $\lambda x.cube(x)$ $\lambda x.object(x)$ $\lambda x.rect(x)$ ...

Overview

Space of Referring Expressions

Probabilistic Model

- Global Model
- Explicit Pruning Model
- Features

Learning

Experiments

Results

Conclusion

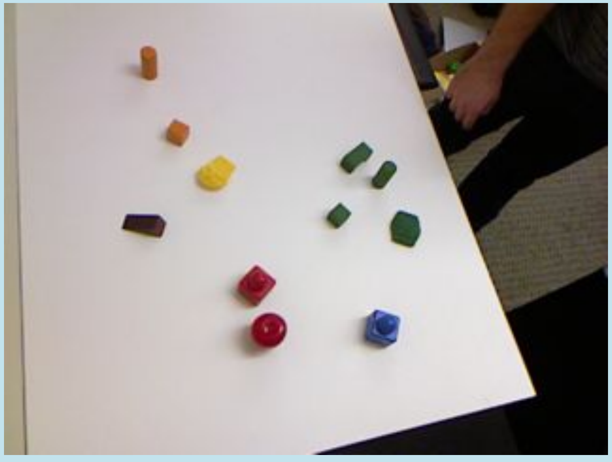
Global Model

$$P(z \mid S, G), z \in \mathcal{Z}$$

Global Model

$$P(z \mid S, G), z \in \mathcal{Z}$$

World State



```
obj1: red, sphere, apple
obj2: brown, triangle
obj3: yellow, fries
…
…
```

Global Model

$$P(z \mid S, G), z \in \mathcal{Z}$$

Target Set

Global Model

- Global Density-Estimation Model
  - Multinomial Log-linear over expressions z that name the set G in state S

$$P_G(z \mid S, G; \theta) = \frac{1}{C} e^{\theta \cdot \phi(z, S, G)}$$

Parameters	$\theta \in \mathcal{R}^n$
Features	$\phi(z, S, G) \in \mathcal{R}^n$
Normalization constant	$C = \sum_{z' \in \mathcal{Z}} e^{\theta \cdot \phi(z', S, G)}$

Global Model

- Global Density-Estimation Model
  - Multinomial Log-linear

$$P_G(z \mid S, G; \theta) = \frac{1}{C} e^{\theta \cdot \phi(z, S, G)}$$

Parameters $\theta \in \mathcal{R}^n$

Features $\phi(z, S, G) \in \mathcal{R}^n$

Normalization constant $C = \sum_{z' \in \mathcal{Z}} e^{\theta \cdot \phi(z', S, G)}$

Too Big!

(Exponential in max number of constants M)

Pruning

M ↑

5	$\iota x.red(x) \cup \iota x.cube(x)$ $\mathcal{E}x.object(x) \setminus \iota x.cube(x)$ $\lambda x.object(x) \wedge equal(x, \mathcal{A}y.cube(y))$ ...
4	$\iota x.red(x) \wedge object(x)$ $\mathcal{A}x.red(x) \vee object(x)$ $\lambda x.\neg cube(x) \wedge object(x)$ $\lambda x.\neg(cube(x) \vee object(x))$ ...
3	$\mathcal{E}x.\neg red(x)$ $\iota x.\neg blue(x)$ $\lambda x.red(x) \wedge cube(x)$ $\lambda x.red(x) \wedge object(x)$ $\lambda x.red(x) \vee cube(x)$ ...
2	$\lambda x.\neg red(x)$ $\lambda x.\neg blue(x)$ $\iota x.red(x)$ $\iota x.cube(x)$ $\mathcal{E}x.object(x)$ ...
1	$\lambda x.red(x)$ $\lambda x.blue(x)$ $\lambda x.cube(x)$ $\lambda x.object(x)$ $\lambda x.rect(x)$ ...

Pruning

M ↑

5	$\iota x.red(x) \cup \iota x.cube(x)$ $\mathcal{E}x.object(x) \setminus \iota x.cube(x)$ $\lambda x.object(x) \wedge equal(x, \mathcal{A}y.cube(y))$	$\ldots$
4	$\iota x.red(x) \wedge object(x)$ $\mathcal{A}x.red(x) \vee object(x)$ $\lambda x.\neg cube(x) \wedge object(x)$ $\lambda x.\neg(cube(x) \vee object(x))$	$\ldots$
3	$\mathcal{E}x.\neg red(x)$ $\iota x.\neg blue(x)$ $\lambda x.red(x) \wedge cube(x)$ $\lambda x.red(x) \wedge object(x)$ $\lambda x.red(x) \vee cube(x)$	$\ldots$
2	$\lambda x.\neg red(x)$ $\lambda x.\neg blue(x)$ $\iota x.red(x)$ $\iota x.cube(x)$ $\mathcal{E}x.object(x)$	$\ldots$
1	$\lambda x.red(x)$ $\lambda x.blue(x)$ $\lambda x.cube(x)$ $\lambda x.object(x)$ $\lambda x.rect(x)$	$\ldots$

Pruning

M ↑

5	$\iota x.red(x) \cup \iota x.cube(x)$ $\mathcal{E}x.object(x) \setminus \iota x.cube(x)$ $\lambda x.object(x) \wedge equal(x, \mathcal{A}y.cube(y))$
4	$\iota x.red(x) \wedge object(x)$ $\mathcal{A}x.red(x) \vee object(x)$ $\lambda x.\neg cube(x) \wedge object(x)$ $\lambda x.\neg(cube(x) \vee object(x))$
3	$\mathcal{E}x.\neg red(x)$ $\iota x.\neg blue(x)$ $\lambda x.red(x) \wedge cube(x)$ $\lambda x.red(x) \wedge object(x)$ $\lambda x.red(x) \vee cube(x)$
2	$\lambda x.\neg red(x)$ $\lambda x.\neg blue(x)$ $\iota x.red(x)$ $\iota x.cube(x)$ $\mathcal{E}x.object(x)$
1	$\lambda x.red(x)$ $\lambda x.blue(x)$ $\lambda x.cube(x)$ $\lambda x.object(x)$ $\lambda x.rect(x)$

Top-k
(Beam Search)

Pruning Model

Good Referring Expression $\neq$ Good Sub-Expression

Pruning Model

Good Referring Expression $\neq$ Good Sub-Expression

$$P_j(a \mid S, G) =$$

Binary probability distribution indicating whether an expression should be pruned at complexity-level j

Pruning Model

- Binary Log-Linear Model for each complexity-level j

$$P_j(a \mid S, G; \pi_j) = \frac{e^{\pi_j \cdot \phi(a,S,G)}}{1 + e^{\pi_j \cdot \phi(a,S,G)}}$$

Parameters π_j

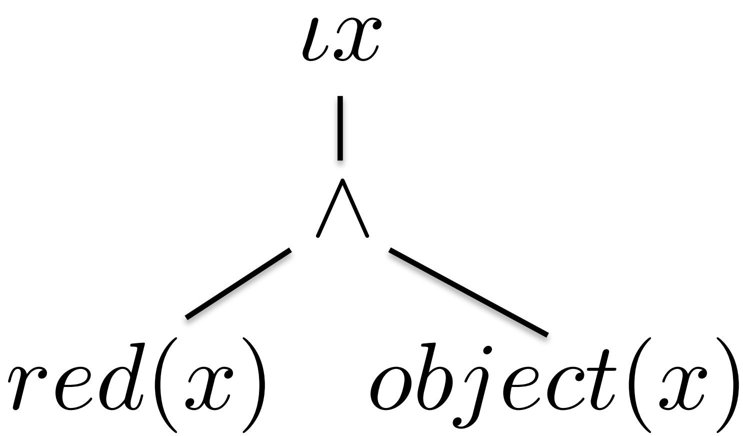
Features

- Structural Features
  - Logical form z only
  - Capture common combinations of predicates
- Situated Features
  - LF z, world-state S and target-set G
  - Capture how sub-expressions of z group sets of objects in the scene
- Complexity Feature

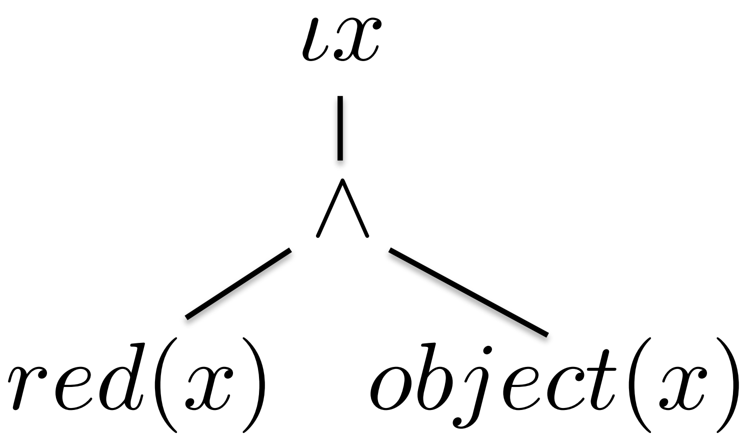
Structural Features

$$\iota x.red(x) \wedge object(x)$$

Structural Features



Structural Features



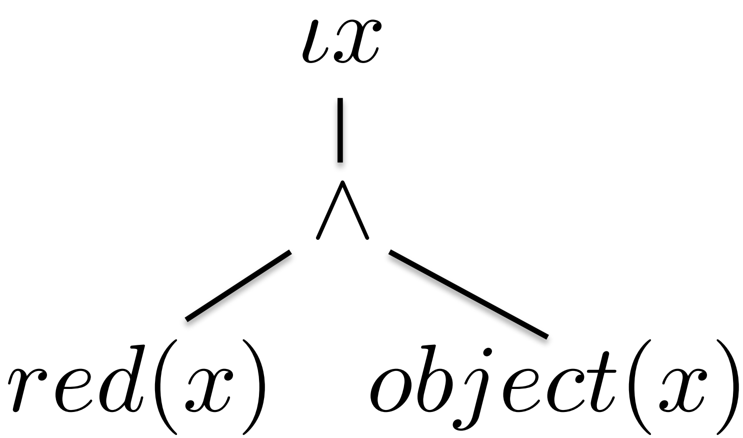
Head Bigram

$$[\iota, \wedge]$$

$$[\wedge, color]$$

$$[\wedge, object]$$

Structural Features

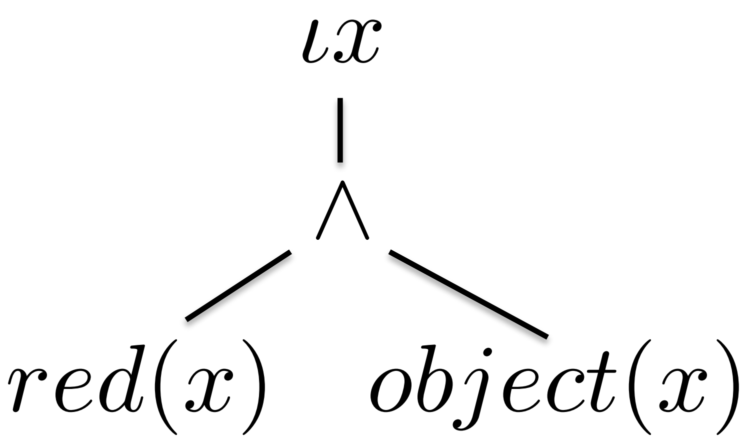


Head Bigram

Coordination Children

$$[\wedge; color, object]$$

Structural Features



Head Bigram

Coordination Children

Coordination Duplicate

Head Predicate

Head Trigram

Situated Features

$G:$ 

<u>Coverage</u>

$$\iota x.red(x) \wedge object(x)$$

Situated Features

$G:$ 

<u>Coverage</u>

SUB

$\iota x.red(x) \wedge object(x)$

Situated Features

$G:$ 

$\iota x.red(x) \wedge object(x)$

<u>Coverage</u>

SUB

SPR

$DISJ$

ALL

$EMPTY$

$OTHER$

Structural Features

$\iota x.red(x) \wedge object(x)$

Head Predicate and Coverage

$[\iota, SUB]$

$[\wedge, SUB]$

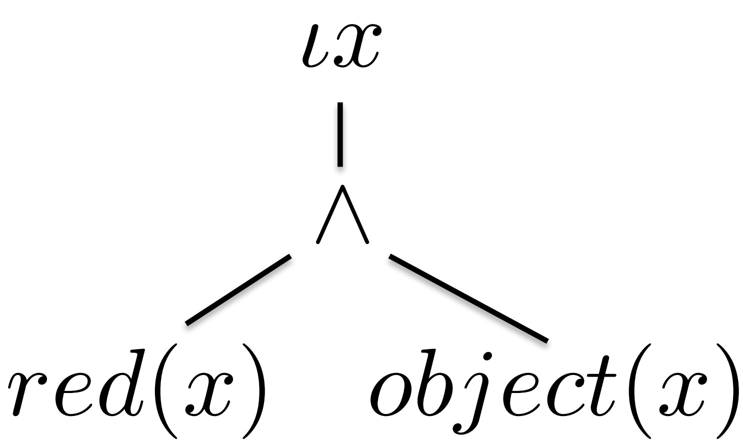
$[color, SUB]$

$[object, ALL]$

Situated Features

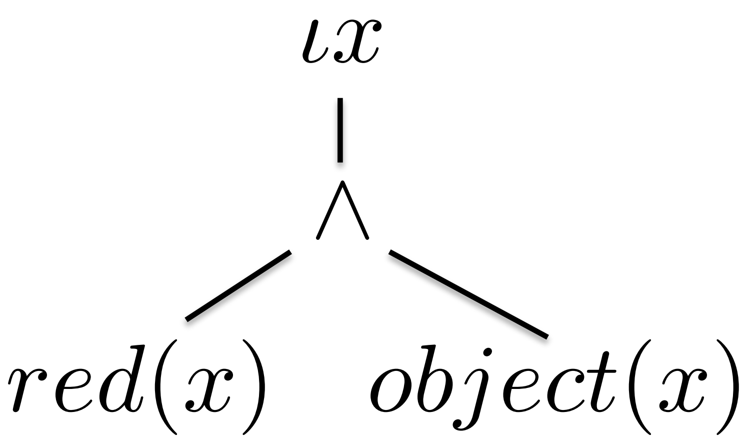
Head Predicate and Coverage

Coordination Child Coverage



$$[\wedge; SUB, ALL]$$

Situated Features



Head Predicate and Coverage

Coordination Child Coverage

Coordination Relative Cov.

Overview

Space of Referring Expressions

Probabilistic Model

Learning

- Data
- Algorithm

Experiments

Results

Conclusion

Learning – Data

$$\{(S_i, G_i, Z_i) : i = 1 \ldots n\}$$

Learning – Data

$$\{(S_i, G_i, Z_i) : i = 1 \ldots n\}$$

World State

```
obj1: red, sphere, apple
obj2: brown, triangle
obj3: yellow, fries
…
…
```

Learning – Data

$$\{(S_i, G_i, Z_i) : i = 1 \dots n\}$$

Target Set

Learning – Data

$$\{(S_i, G_i, Z_i) : i = 1 \ldots n\}$$

Labeled Logical Forms

$$\iota x.(yellow(x) \lor orange(x) \lor red(x) \lor green(x)) \land object(x) \land plu(x)$$
$$\iota x.(yellow(x) \lor orange(x) \lor red(x) \lor green(x)) \land object(x) \land plu(x)$$
$$\iota x.(yellow(x) \lor orange(x) \lor red(x) \lor green(x)) \land object(x) \land plu(x)$$
$$\iota x.\neg(brown(x) \lor blue(x)) \land object(x) \land plu(x)$$
$$\iota x.\neg(brown(x) \lor blue(x)) \land object(x) \land plu(x)$$
$$\mathcal{E}x.\neg(brown(x) \lor blue(x)) \land object(x) \land sg(x))$$
$$\cdots$$
$$\cdots$$

Learning – Data

$$\{(S_i, G_i, Z_i) : i = 1 \ldots n\}$$

<u>Empirical Distribution:</u> Q_i

$\hat{Q}(z \mid S_i, G_i)$

0.1	$\iota x.(yellow(x) \vee orange(x) \vee red(x) \vee green(x)) \wedge object(x) \wedge plu(x)$
0.3	$\iota x.\neg(brown(x) \vee blue(x)) \wedge object(x) \wedge plu(x)$
0.2	$\mathcal{E}x.\neg(brown(x) \vee blue(x)) \wedge object(x) \wedge sg(x))$
$\ldots$	$\ldots$

Learning Algorithm

- Online
- Stochastic Gradient Descent

Learning Algorithm

For $t = 1 \ldots T, i = 1 \ldots n$:

Step 1: (Update Global Model)

a. Compute the stochastic gradient
b. Update the parameters

Step 2: (Update Pruning Model)

For $j = 1 \ldots M$

a. Construct a set of positive and negative examples
b. Compute mini-batch stochastic gradient
c. Update complexity-j pruning parameters

Learning Algorithm

For $t = 1 \dots T, i = 1 \dots n$:

Step 1: (Update Global Model)

a. Compute the stochastic gradient

$$\Delta_\theta \leftarrow E_{Q_i(z|S_i,G_i)}[\phi_i(z)] - E_{\hat{P}(z|G_i,S_i;\theta,\Pi)}[\phi_i(z)]$$

b. Update the parameters

$$\gamma \leftarrow \frac{\alpha_0}{1+c\times\tau} \text{ where } \tau = i + t \times n$$
$$\theta \leftarrow \theta + \gamma\Delta_\theta$$

Step 2: (Update Pruning Model)

Learning Algorithm

For $t = 1 \dots T, i = 1 \dots n$:

Step 1: (Update Global Model)

Step 2: (Update Pruning Model)

For $j = 1 \dots M$

a. Construct a set of positive and negative examples

$$\mathcal{D}^+ \leftarrow \bigcup_{z \in Z_i} SUB(j, z).$$
$$\mathcal{D}^- \leftarrow \mathcal{A}_j \setminus \mathcal{D}^+$$

b. Compute mini-batch stochastic gradient

c. Update complexity-j pruning parameters

Learning Algorithm

For $t = 1 \ldots T, i = 1 \ldots n$:

Step 1: (Update Global Model)

Step 2: (Update Pruning Model)

For $j = 1 \ldots M$

a. Construct a set of positive and negative examples

b. Compute mini-batch stochastic gradient

$$\Delta_{\Pi_j} \leftarrow \frac{1}{|\mathcal{D}^+|} \sum_{z \in \mathcal{D}^+} (1 - P_j(z \mid S_i, G_i; \Pi_j))\phi_i(z) \\ - \frac{1}{|\mathcal{D}^-|} \sum_{z \in \mathcal{D}^-} P_j(z \mid S_i, G; \Pi j)\phi_i(z)$$

c. Update complexity-j pruning parameters

$$\Pi_j \leftarrow \Pi_j + \gamma \Delta_{\Pi_j}$$

Overview

Space of Referring Expressions

Probabilistic Model

Learning

Experiments

Results

Conclusion

Data Collection

“Please pick up ____________________”

- 269 scenes
  20 expressions / scene
  5380 expressions
- Data Split
  - Training:
    - 196 scenes (3920 exps)
    - Labeled semi-automatically
  - Dev:
    - 20 scenes (400 exps)
    - Hand-Labeled
  - Test:
    - 43 scenes (860 exps)
    - Hand-Labeled

Training Data

- Semi-Automatic Labeling
  - Trained semantic parser on initialization set
    - 10 scenes, 100 sentence-expression pairs
  - Hand engineered lexicon
  - 95% precision, 70% recall
  - Labeled 196 scenes (3920 exps)
  - Use scenes with at least 15 successful labels
- Total training set:
  - 141 scenes, 2587 expressions

Related Work

- Most previous systems are deterministic
  [Dale and Reiter 1995], [van Deempter 2002], [Gardent 2002], [Horacek 2004], [Gatt and van Deempter 2007], [Areces et al. 2008], [Ren et al. 2010], [Krahmer and van Deempter 2012], [van Deempter et al. 2012], …
- Learning to refer to a single objects
  - Compare state of the art approach
  - Visual Objects Algorithm [Mitchell et al 2013]
- Learning to refer to sets of objects
  - Requires more complex logical expressions
  - Present the first learning results + ablations

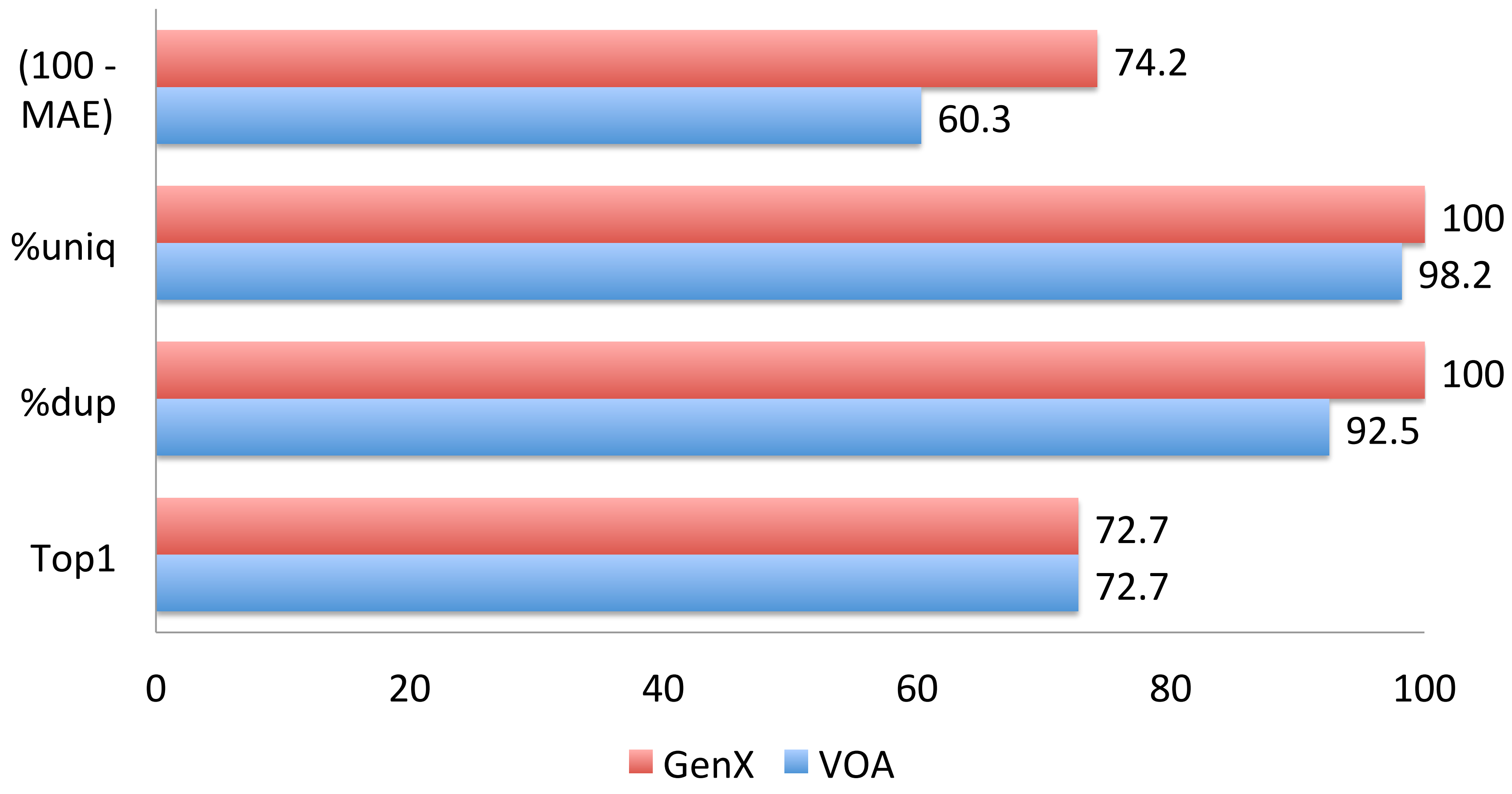
Evaluation

- Mean absolute error (100 - MAE)

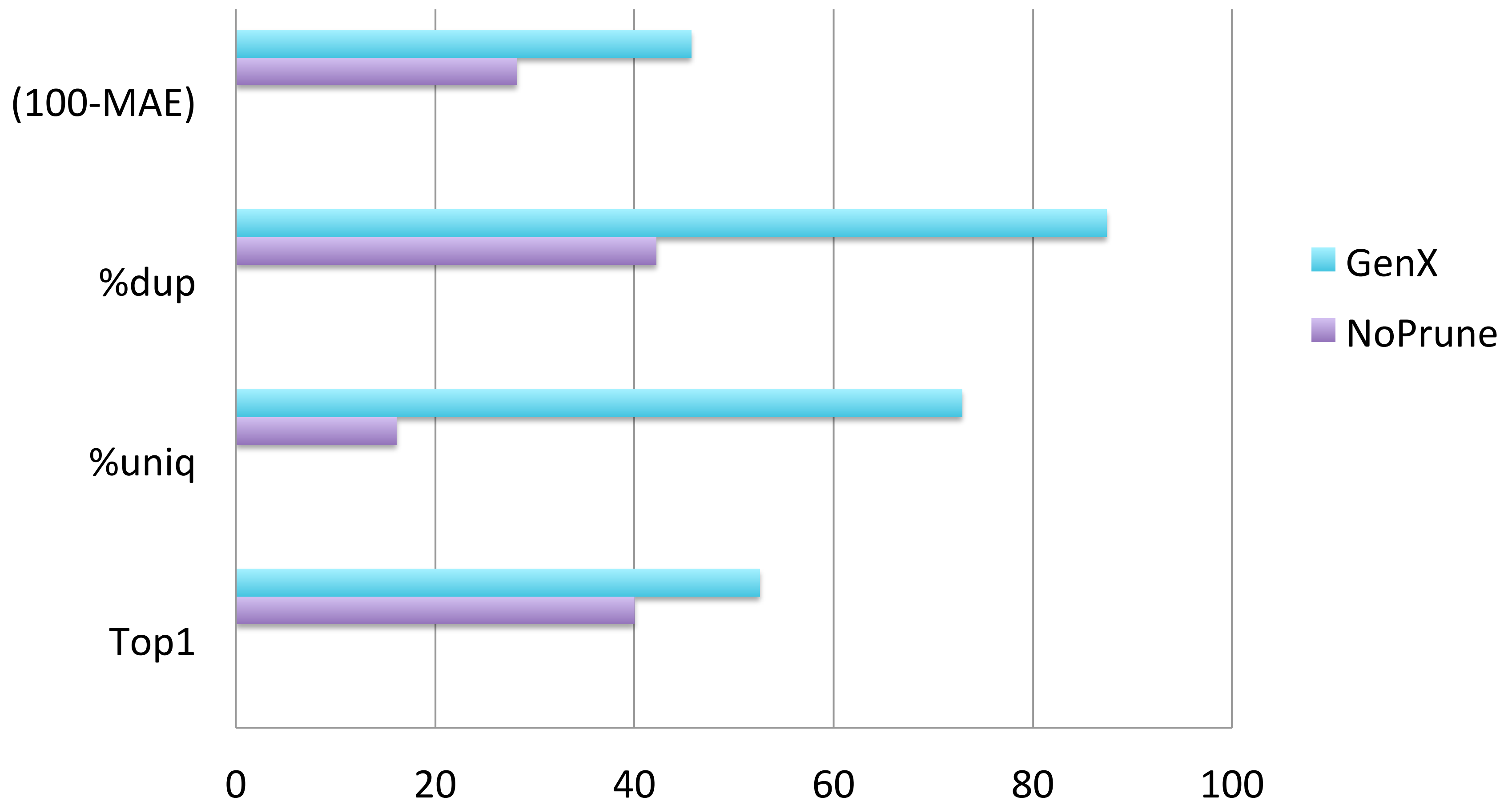
$$MAE = \frac{1}{2n}\sum_{i=1}^{n}\sum_{z\in\mathcal{Z}}|P(z \mid S_i, G_i) - Q(z \mid S_i, G_i)|$$

- Coverage
  - All logical forms: $\%_{dup}$
  - Unique logical forms: $\%_{uniq}$
- Top-1 Accuracy

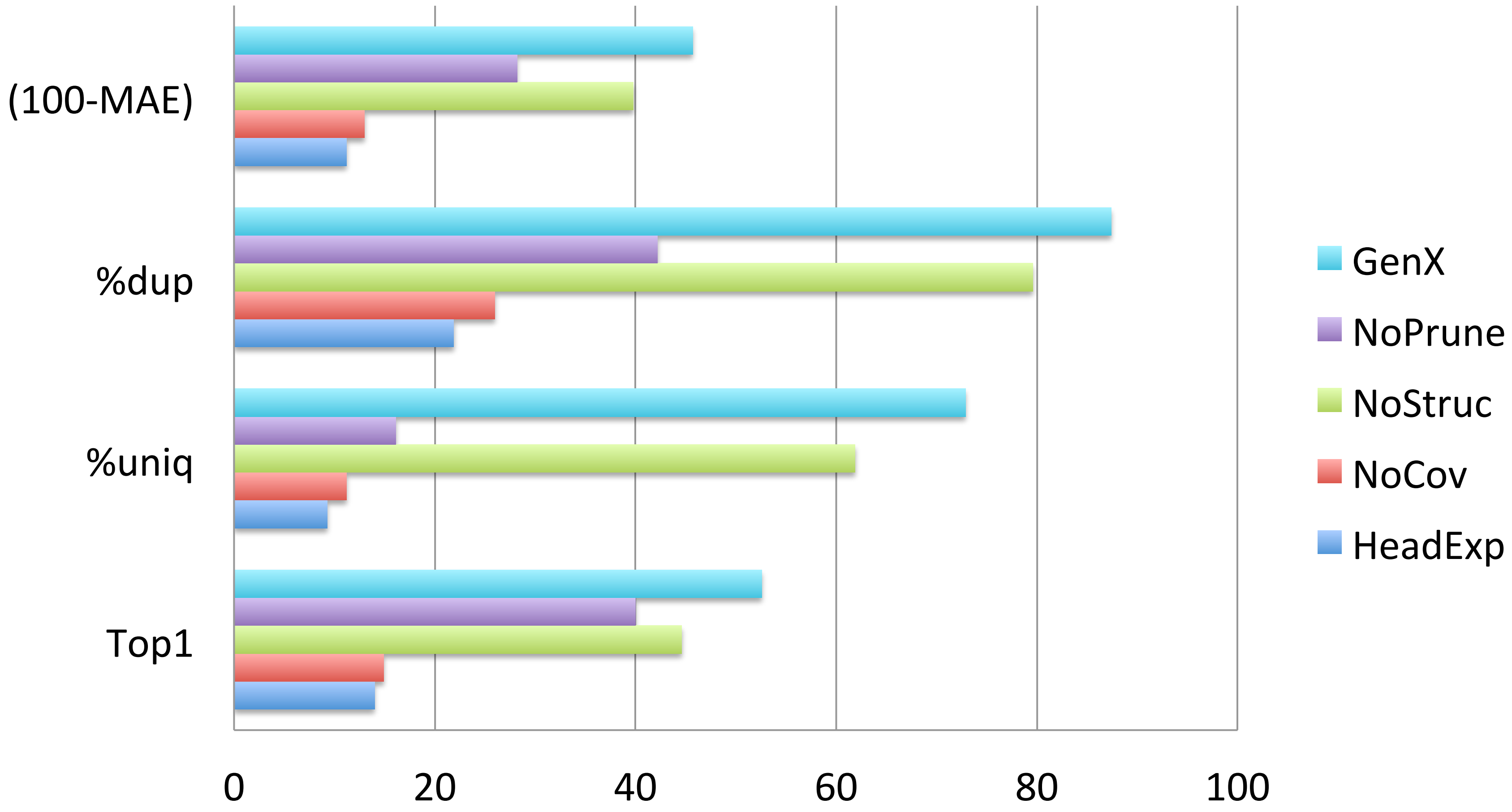
Results – Single Objects



Results – Object Sets



Results – Object Sets



Qualitative Results

Q	$\hat{P}$	z
.750	.320	$\iota(\lambda x.object(x) \wedge (yellow(x) \vee red(x)))$
	.114	$\iota(\lambda x.lego(x)) \cup \iota(\lambda x.red(x) \wedge apple(x))$
	.114	$\iota(\lambda x.yellow(x) \wedge lego(x))) \cup \iota(\lambda x.apple(x))$
	.044	$\iota(\lambda x.lego(x) \vee (red(x) \wedge apple(x)))$
	.044	$\iota(\lambda x.(yellow(x) \wedge lego(x)) \vee apple(x))$
	.036	$\iota(\lambda x.lego(x)) \cup \iota(\lambda x.red(x) \wedge sphere(x))$
	.026	$\iota(\lambda x.red(x) \wedge lego(x)) \cup \iota(\lambda x.red(x) \wedge sphere(x))$
.050	.021	$\iota(\lambda x.(lego(x) \wedge yellow(x)) \vee (red(x) \wedge apple(x)))$
	.017	$\iota(\lambda x.(lego(x) \wedge yellow(x)) \vee (red(x) \wedge sphere(x)))$
	.014	$\iota(\lambda x.yellow(x) \wedge lego(x)) \cup \iota(\lambda x.red(x) \wedge sphere(x))$

.100	.010	$\iota(\lambda x.yellow(x) \wedge object(x)) \cup \iota(\lambda x.apple(x))$
.050	.007	$\iota(\lambda x.yellow(x) \wedge object(x)) \cup \iota(\lambda x.red(x) \wedge sphere(x))$
.050	.005	$\iota(\lambda x.yellow(x) \wedge object(x)) \cup \iota(\lambda x.red(x) \wedge object(x))$

Conclusion

- Referring Expression Generation as Density Estimation
  - Global Model
  - Learned Pruning Model
- First results on density estimation for set reference
- State-of-the-art results on single objects

Future Work

- Full joint approach to REG

Future Work

- Full joint approach to REG

Physical Scene

Logical Form (LF)

$$\iota x.(green(x) \vee red(x)) \wedge sphere(x)$$

Sentence

``The green and red spheres.''

Future Work

- Joint approach to REG
- Extend to more general referring expressions

"the second jar to the left of the middle shelf"

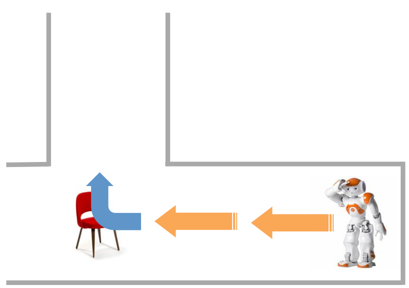
Future Work

- Joint approach to REG
- Extend to more general referring expressions
- Extend to other grounded language problems

States

Abbr.	Capital	Pop.
AL	Montgomery	3.9
AK	Juneau	0.4
AZ	Phoenix	2.7
WA	Olympia	4.1
NY	Albany	17.5
IL	Springfield	11.4

``What is the largest state?''



``Move to the chair and turn right.''

Questions?

http://nfitz.net

Data: Available Now
Code: Available Soon