Learning Grounded Semantic Parsers

Luke Zettlemoyer
Computer Science & Engineering
University of Washington

joint work with Yoav Artzi, Nicholas FitzGerald,
Cynthia Matuszek, Liefeng Bo, Dieter Fox
Mapping Sentences to Meaning

Texas borders Kansas.

next-to(TEX, KAN)
What states border Texas?

\[ \lambda x. \text{state}(x) \land \text{next-to}(x, \text{TEX}) \]
Mapping Sentences to Meaning

What states border Texas?
\[ \lambda x. \text{state}(x) \land \text{next-to}(x, \text{TEX}) \]

Machine Learning Problem
Given: Many input, output pairs
Learn: A function that maps sentences to lambda-calculus expressions
A Challenge: Learning Hidden Structure

Approach 1. Fully annotated training examples (parse trees):
A Challenge: Learning Hidden Structure

Approach 1. Fully annotated training examples (parse trees):

Show me the latest flight from Boston to Seattle on Friday

\[
\text{argmax}(\lambda x. \text{flight}(x) \land \text{from}(x, \text{BOS}) \land \text{to}(x, \text{SEA}) \land \text{day}(x, \text{FRI}), \lambda y. \text{time}(y))
\]
A Challenge: Learning Hidden Structure

Recent approaches. Only require annotation of final meanings
A Challenge: Learning Hidden Structure

Recent approaches. Only require annotation of final meanings

Show me the latest flight from Boston to Seattle on Friday

\[
\text{argmax}(\lambda x. \text{flight}(x) \land \text{from}(x, \text{BOS}) \land \text{to}(x, \text{SEA}) \land \text{day}(x, \text{FRI}), \lambda y. \text{time}(y))
\]
A Challenge: Learning Hidden Structure

Recent approaches. Only require annotation of final meanings

Show me the latest flight from Boston to Seattle on Friday

Supervised semantic parsing with:
- Machine Translation
- Inducing Logic Programming
- Prob. CFG Parsing
- Prob. PDA
- Support Vector Machines
- Perceptron-style Learning
- Higher-order unification

\[
\text{argmax}(\lambda x. \text{flight}(x) \land \text{from}(x, \text{BOS}) \land \text{to}(x, \text{SEA}) \land \text{day}(x, FRI), \lambda y. \text{time}(y))
\]
A Challenge: Learning Latent Semantics

Recent approaches. Only require annotation of final meanings

Show me the latest flight from Boston to Seattle on Friday
A Challenge: Learning Latent Semantics

Goal. Annotate as little as possible!

Show me the latest flight from Boston to Seattle on Friday

Less supervision:
- Question-Answers Semi-supervision
- Confidence-driven Unsupervised Learning from Natural Instructions
- Learning from Conversations
- Learning from Robotic Demonstrations

[Clarke et al. 2010, Liang et al. 2011]
[Goldwasser et al. 2011]
[Goldwasser and Roth, 2011]
[Artzi and Zettlemoyer, 2011]
[Matuszek et al. 2012]
A Challenge: Learning Latent Semantics

Part 1: Bootstrapping from Conversations

**SYSTEM**  how can i help you ?

**USER**  i ' d like to go from richmond to atlantic city on june seventh

**SYSTEM**  leaving richmond on thursday, june seventh arriving in atlantic city leaving at what time ?

**USER**  eight a. m.

**SYSTEM**  leaving in the morning what airline would you prefer ?

**USER**  u. s. air

[conversation continues]
A Challenge: Learning Latent Semantics

Part 1: Bootstrapping from Conversations

**SYSTEM** how can i help you?

**USER** i 'd like to go from richmond to atlantic city on june seventh

**SYSTEM** leaving richmond on thursday, june seventh arriving in atlantic city leaving at what time?

**USER** eight a. m.

**SYSTEM** leaving in the morning what airline would you prefer?

**USER** u. s. air

[conversations continues]

Part 2: Learning in Visual Scenes

Show me the ones that are not blue
Background: CCG and Learning with Labeled Logical Forms
Parsing with CCGs

Lexicon

Lexical Entry

Text

Category

Syntax

λ

Combinators

Application

Composition

Type-shifting

[Steedman 96,00]
Probabilistic CCGs

Lexicon: \[ \Lambda = \begin{array}{c}
\text{Texas} \\
\text{NP} \\
\text{tex}
\end{array} \]

Parameters: \[ \theta \]
Probabilistic CCGs

Lexicon: \( \Lambda \)

\[
\Lambda = \begin{pmatrix}
\text{Texas} & \text{NP} & \text{tex}
\end{pmatrix}
\]

Parameters: \( \theta \)

Probability distribution: sentence \( x \), parse \( y \), logical form \( z \)

- Log-linear model:
  \[
P(y, z|x; \theta, \Lambda) = \frac{e^{\theta \cdot \phi(x, y, z)}}{\sum_{(y', z')} e^{\theta \cdot \phi(x, y', z')}}
\]

- Parsing:
  \[
f(x) = \arg \max_z p(z|x; \theta, \Lambda)
\]

where
  \[
p(z|x; \theta, \Lambda) = \sum_y p(y, z|x; \theta, \Lambda)
\]
We will learn a lexicon and parameters

Training Examples:

List flights to Singapore.
\[ \lambda x. \text{flight}(x) \land \text{to}(x, \text{SIN}) \]

CCG Lexicon:

Parameters:

, \ W
A Multilingual Learning Algorithm

Key challenge: learn from data with different natural languages and meaning representations

English, logical-form:

\[ \text{NL: what states border texas} \]
\[ \text{MR: } \lambda x. \text{state}(x) \land \text{next_to}(x, \text{tex}) \]

Turkish, functional query language:

\[ \text{NL: texas a siniri olan eyaletler nelerdir} \]
\[ \text{MR: answer(state(next_to_2(stateid \ tex)))} \]
A Multilingual Learning Algorithm

Key challenge: learn from data with different natural languages and meaning representations

English, logical-form:

NL: what states border texas  
MR: $\lambda x. \text{state}(x) \land \text{next_to}(x, \text{tex})$

Turkish, functional query language:

NL: texas a siniri olan eyaletler nelerdir  
MR: $\text{answer} (\text{state}(\text{next_to}_2(\text{stateid} \text{ tex})))$

Approach: use higher-order unification to recursively learn the meaning of smaller words and phrases  

[Kwiatkowski et al. 2010,2011]
### Example Learned Lexical Entries

<table>
<thead>
<tr>
<th>English</th>
<th>Turkish</th>
<th>Spanish</th>
</tr>
</thead>
<tbody>
<tr>
<td>population of $\vdash NP/NP : \lambda x. population(x)$</td>
<td>nedir $\vdash S\backslash NP\backslash(NP</td>
<td>NP) : \lambda f \lambda x. f(x)$</td>
</tr>
<tr>
<td>smallest $\vdash NP/(S</td>
<td>NP) : \lambda f \arg\min(y, f(y), size(y))$</td>
<td>sehir $\vdash S</td>
</tr>
<tr>
<td>what $\vdash S</td>
<td>NP/(S</td>
<td>NP) : \lambda f \lambda x. f(x)$</td>
</tr>
<tr>
<td>border $\vdash S</td>
<td>NP/NP : \lambda x \lambda y.next_to(y, x)$</td>
<td>siniri $\vdash S</td>
</tr>
<tr>
<td>state $\vdash S</td>
<td>NP : \lambda x. state(x)$</td>
<td>kac tane $\vdash S\backslash NP/(S</td>
</tr>
<tr>
<td>most $\vdash NP/(S</td>
<td>NP)/(S</td>
<td>NP)/(S</td>
</tr>
</tbody>
</table>

#### Japanese
- no $\vdash NP|NP/(NP|NP) : \lambda f \lambda x. f(x)$
- shuu $\vdash S|NP : \lambda x. state(x)$
- nan desu ka $\vdash S\backslash NP/(NP|NP) : \lambda f \lambda x. f(x)$
- wa $\vdash NP|NP/(NP|NP) : \lambda f \lambda x. f(x)$
- ikutsu $\vdash NP/(S|NP)/(S|NP/(S|NP)): \lambda f \lambda g. count(x, f(g(x)))$
- chiiki $\vdash NP\backslash NP : \lambda x. area(x)$

#### Spanish
- en $\vdash S|NP/NP: \lambda x \lambda y. loc(y, x)$
- que es la $\vdash S/\backslash NP/(NP|NP): \lambda f \lambda x. f(x)$
- pequena $\vdash NP\backslash(NP|NP): \lambda g \lambda f \arg\min(y, f(y), g(y))$
- estado $\vdash S\backslash NP : \lambda x. state(x)$
- mas $\vdash S\backslash(S|NP)/(S|NP)/(NP|NP/(S|NP)): \lambda f \lambda g \lambda h.\arg\max(x, h(x), f(g, x))$
- mayores $\vdash S\backslash NP/(S|NP) : \lambda f \lambda x. f(x) \land major(x)$
Bootstrapping Semantic Parsers from Conversations

[Artzi & Zettlemoyer, EMNLP 2011]
Task Oriented Dialog

**SYSTEM**  how can i help you ?

**USER**  i ' d like to go from richmond to atlantic city on june seventh

**SYSTEM**  leaving richmond on thursday, june seventh arriving in atlantic city leaving at what time ?

**USER**  eight a. m.

**SYSTEM**  leaving in the morning what airline would you prefer ?

**USER**  u. s. air

**SYSTEM**  i currently have you leaving richmond on thursday, june seventh in the morning arriving in atlantic city flying on u. s. air . is this correct ?

**USER**  yes

**SYSTEM**  o . k . then . where would you like to go next ?

**USER**  richmond

[conversation continues]
how can i help you?

i would like to fly from atlanta georgia to london england on september twenty fourth in the early evening i would like to return on october first departing from london in the late morning
how can i help you?

i would like to fly from atlanta georgia to london england on september twenty fourth in the early evening i would like to return on october first departing from london in the late morning
System Recovery

SYSTEM: how can i help you?

USER: i would like to fly from atlanta georgia to london england on september twenty fourth in the early evening i would like to return on october first departing from london in the late morning

SYSTEM: leaving what city?
how can i help you?

i would like to fly from atlanta georgia to london england on september twenty fourth in the early evening i would like to return on october first departing from london in the late morning

leaving what city?

atlanta georgia
System Recovery

**SYSTEM:** how can i help you?

**USER:** i would like to fly from atlanta georgia to london england on september twenty fourth in the early evening i would like to return on october first departing from london in the late morning

**SYSTEM:** leaving what city?

**USER:** atlanta georgia

**SYSTEM:** leaving atlanta. going to which city?
how can i help you?

i would like to fly from atlanta georgia to london england on september twenty fourth in the early evening. i would like to return on october first departing from london in the late morning.

leaving what city?

atlanta georgia

leaving atlanta. going to which city?

london
how can i help you?
i would like to fly from atlanta georgia to london england on september twenty fourth in the early evening i would like to return on october first departing from london in the late morning

leaving what city?

atlanta georgia

leaving atlanta. going to which city?
london

arriving in london england. what date would you like to depart atlanta?
how can i help you?
i would like to fly from atlanta georgia to london england on september twenty fourth in the early evening i would like to return on october first departing from london in the late morning
leaving what city?
atlanta georgia
leaving atlanta. going to which city?
london
arriving in london england. what date would you like to depart atlanta?
september twenty fourth in the early evening
USER: I would like to fly from Atlanta, Georgia to London, England on September 24th in the early evening. I would like to return on October first departing from London in the late morning.

SYSTEM: Leaving what city? (ASK: λx. from(fl, x))

USER: Atlanta, Georgia

SYSTEM: Leaving Atlanta. (CONFIRM: from(fl, ATL)) Going to which city? (ASK: λx. to(fl, x))

USER: London

[Conversation continues]
Automatically Annotated Logs

SYSTEM: how can i help you? (OPEN_TASK)

USER: i would like to fly from atlanta georgia to london england on september twenty fourth in the early evening i would like to return on october first departing from london in the late morning

SYSTEM: leaving what city? (ASK: \( \lambda x.\text{from}(fl, x) \))

USER: atlanta georgia

SYSTEM: leaving atlanta. (CONFIRM: \( \text{from}(fl, ATL) \)) going to which city? (ASK: \( \lambda x.\text{to}(fl, x) \))

USER: london

[conversation continues]
System utterance
USER: i would like to fly from atlanta georgia to london england on september twenty fourth in the early evening i would like to return on october first departing from london in the late morning

SYSTEM: leaving what city? (ASK: \( \lambda x. \text{from}(\text{fl}, x) \))

USER: atlanta georgia

SYSTEM: leaving atlanta. (CONFIRM: from(fl, ATL)) going to which city? (ASK: \( \lambda x. \text{to}(\text{fl}, x) \))

USER: london

[conversation continues]

System utterance

Natural Language
**SYSTEM:** how can i help you? **(OPEN_TASK)**

**USER:** i would like to fly from atlanta georgia to london england on september twenty fourth in the early evening i would like to return on october first departing from london in the late morning

**SYSTEM:** leaving what city? **(ASK: \( \lambda x. \text{from}(fl, x) \))**

**USER:** atlanta georgia

**SYSTEM:** leaving atlanta. **(CONFIRM from(fl, ATL))** going to which city? **(ASK: \( \lambda x. \text{to}(fl, x) \))

**USER:** london

[conversation continues]
how can i help you?  

i would like to fly from atlanta georgia to london england on september twenty fourth in the early evening i would like to return on october first departing from london in the late morning

leaving what city?  

leaving atlanta.  

going to which city?

[conversation continues]
Automatically Annotated Logs

**SYSTEM:** how can i help you? (OPEN_TASK)

**USER:** i would like to fly from atlanta georgia to london england on september twenty fourth in the early evening i would like to return on october first departing from london in the late morning

**SYSTEM:** leaving what city? (ASK: \( \lambda x. \text{from}(fl, x) \))

**USER:** atlanta georgia

**SYSTEM:** leaving atlanta. (CONFIRM: from(fl, ATL)) going to which city? (ASK: \( \lambda x. \text{to}(fl, x) \))

**USER:** london

[conversation continues]

System utterance

- **Natural Language**
- **Speech Act**
- **Logical Meaning**
- **Conversational Objects**
**SYSTEM:** how can i help you? *(OPEN_TASK)*

**USER:** i would like to fly from atlanta georgia to london england on september twenty fourth in the early evening i would like to return on october first departing from london in the late morning

**SYSTEM:** leaving what city? *(ASK: \(\lambda x. from(fl, x)\))*

**USER:** atlanta georgia

**SYSTEM:** leaving atlanta. *(CONFIRM: from(fl, ATL))* going to which city? *(ASK: \(\lambda x. to(fl, x)\))*

**USER:** london

[conversation continues]

System utterance
One Training Example

**SYSTEM:** how can i help you? *(OPEN_TASK)*

**USER:** i would like to fly from atlanta georgia to london england on september twenty fourth in the early evening i would like to return on october first departing from london in the late morning

**SYSTEM:** leaving what city? *(ASK: \( \lambda x. \text{from}(fl, x) \))*

**USER:** atlanta georgia

**SYSTEM:** leaving atlanta. *(CONFIRM: \text{from}(fl, ATL))* going to which city? *(ASK: \( \lambda x. \text{to}(fl, x) \))*

**USER:** london

[conversation continues]
Supervised Case

i want to fly from new york to los angeles

\[ \lambda x. \text{from}(x, \text{NYC}) \land \text{to}(x, \text{LA}) \]
Loss Signal Instead of Labels

\[ \mathcal{L} : \langle z, C \rangle \rightarrow \mathbb{R} \]

i want to fly from new york to los angeles
Loss Signal Instead of Labels

\[ \mathcal{L} : \langle z, C \rangle \rightarrow \mathbb{R} \]

i want to fly from new york to los angeles

\[ \lambda x. \text{from}(x, \text{NYC}) \land \text{to}(x, \text{LA}) \]
\[ \lambda x. \text{from}(x, \text{NYC}) \land \text{from}(x, \text{LA}) \]
\[ \lambda x. \text{from}(x, \text{LA}) \]
\[ \lambda x[]. \text{to}(x[0], \text{NYC}) \land \text{from}(x[1], \text{LA}) \]
\[ \ldots \]
Loss Signal Instead of Labels

\[ \mathcal{L} : \langle z, \mathcal{C} \rangle \rightarrow \mathbb{R} \]

i want to fly from new york to los angeles

1 -1.8 \( \lambda x. \text{from}(x, \text{NYC}) \land \text{to}(x, \text{LA}) \)

4 2.0 \( \lambda x. \text{from}(x, \text{NYC}) \land \text{from}(x, \text{LA}) \)

2 -0.9 \( \lambda x. \text{from}(x, \text{LA}) \)

3 1.5 \( \lambda x[]. \text{to}(x[0], \text{NYC}) \land \text{from}(x[1], \text{LA}) \)

\ldots
Computing Conversation Loss

1. Select the relevant subset of system utterances
2. Align the logical expression to content from selected system utterances
3. Score the alignment
Convergence Loss

**SYSTEM:** how can i help you *(OPEN_TASK)*

**USER:** i want to fly from new york to los angeles

**SYSTEM:** from new york. *(CONFIRM: from(fl, NYC)) to where? *(ASK: \(\lambda x.\text{to}(fl, x)\))

**USER:** los angeles

**SYSTEM:** flying to los angeles. *(CONFIRM: to(fl, LA)) where do you want to go from los angeles? ? *(ASK: \(\lambda x.\text{to}(fl, x)\& from(fl, LA))\)

**USER:** back to new york

**SYSTEM:** from los angeles to new york *(CONFIRM: from(fl2, LA)\& to(fl2, NYC))

**SYSTEM:** i have a direct flight from new york to los angeles for $480. ok? *(OFFER: direct(fl)\& from(fl, NYC)\& to(fl2, LA)\& cost(fl, 480))

[conversation continues]
SYSTEM: how can i help you (OPEN_TASK)

USER: i want to fly from new york to los angeles

SYSTEM: from new york. (CONFIRM: from(fl, NYC)) to where? (ASK: λx.to(fl, x))

USER: los angeles

SYSTEM: flying to los angeles. (CONFIRM: to(fl, LA)) where do you want to go from los angeles? ? (ASK: λx.to(fl2, x)∧from(fl2, LA))

USER: back to new york

SYSTEM: from los angeles to new york (CONFIRM: from(fl2, LA)∧to(fl2, NYC))

SYSTEM: i have a direct flight from new york to los angeles for $480. ok? (OFFER: direct(fl)∧from(fl, NYC)∧to(fl, LA)∧cost(fl, 480))

[conversation continues]
Extract Properties

SYSTEM: how can i help you (OPEN_TASK)

USER: i want to fly from new york to los angeles

SYSTEM: from new york. (CONFIRM: from(fl, NYC)) to where? (ASK: \(\lambda x.\text{to}(fl, x)\))

USER: los angeles

SYSTEM: flying to los angeles. (CONFIRM: to(fl, LA)) where do you want to go from los angeles? ? (ASK: \(\lambda x.\text{to}(fl2, x)\)∧from(fl2, LA))

USER: back to new york

SYSTEM: from los angeles to new york (CONFIRM: from(fl2, LA)∧to(fl2, NYC))

SYSTEM: i have a direct flight from new york to los angeles for $480. ok? (OFFER: direct(fl)∧from(fl, NYC)∧to(fl, LA)∧cost(fl, 480))

[conversation continues]
how can i help you (OPEN_TASK)

i want to fly from new york to los angeles

from new york. (CONFIRM: from(fl, NYC) to where? (ASK: λx.to(fl, x))

los angeles

flying to los angeles. (CONFIRM: to(fl, LA)) where do you want to go from los angeles? ? (ASK: λx.to(fl2, x)∧from(fl2, LA))

back to new york

from los angeles to new york (CONFIRM: from(fl2, LA)∧to(fl2, NYC))

i have a direct flight from new york to los angeles for $480. ok? (OFFER: ...)

[conversation continues]
Comparing Two Candidates

USER: i want to fly from new york to los angeles

\[
\lambda x. \text{from}(x, NYC) \land \text{to}(x, LA)
\]
Alignment: \(x \rightarrow fl\)
\[
Loss = -0.5 - 0.83 = -1.33
\]

\[
\lambda x[]. \text{from}(x[0], NYC) \land \text{from}(x[1], LA)
\]
Alignment: \(x[0] \rightarrow fl, x[1] \rightarrow fl2\)
\[
Loss = -0.5 - 0.16 = -0.66
\]

<table>
<thead>
<tr>
<th>Property</th>
<th>Loss</th>
</tr>
</thead>
<tbody>
<tr>
<td>(from, fl, NYC)</td>
<td>-0.5</td>
</tr>
<tr>
<td>(to, fl, LA)</td>
<td>-0.83</td>
</tr>
<tr>
<td>(from, fl2, LA)</td>
<td>-0.16</td>
</tr>
<tr>
<td>(to, fl2, NYC)</td>
<td>-0.16</td>
</tr>
</tbody>
</table>
Comparing Two Candidates

\[ \lambda x. \text{from}(x, \text{NYC}) \land \text{to}(x, \text{LA}) \]
Alignment: \( x \to fl \)

\[ \text{Loss} = -0.5 - 0.83 = -1.33 \]

<table>
<thead>
<tr>
<th>Property</th>
<th>Loss</th>
</tr>
</thead>
<tbody>
<tr>
<td>(from, fl, NYC)</td>
<td>-0.5</td>
</tr>
<tr>
<td>(to, fl, LA)</td>
<td>-0.83</td>
</tr>
<tr>
<td>(from, fl2, LA)</td>
<td>-0.16</td>
</tr>
<tr>
<td>(to, fl2, NYC)</td>
<td>-0.16</td>
</tr>
</tbody>
</table>

\[ \lambda x[]. \text{from}(x[0], \text{NYC}) \land \text{from}(x[1], \text{LA}) \]
Alignment: \( x[0] \to fl, x[1] \to fl2 \)

\[ \text{Loss} = -0.5 - 0.16 = -0.66 \]

USER: i want to fly from new york to los angeles
Learning Algorithm

• Online

• Loss-driven

• 2 steps (for each training example):
  
  – Lexical generation
  
  – Loss-sensitive parameter update [Singh-Miller and Collins 2007]
    
    • Parse using the current model
    
    • Split all parses into two sets using the loss function: optimal and non-optimal
    
    • Find model score violations between the sets
    
    • Do a perceptron update using these violations
**DARPA Communicator**

- Raw conversational logs [Walker et al. 2002]
- Annotated system utterances

<table>
<thead>
<tr>
<th></th>
<th>Lucent</th>
<th></th>
<th>BBN</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Train</td>
<td>Test</td>
<td>Train</td>
<td>Test</td>
</tr>
<tr>
<td>Conversations</td>
<td>144</td>
<td>70</td>
<td>71</td>
<td>91</td>
</tr>
<tr>
<td>Context Independent Utterances</td>
<td>208</td>
<td>96</td>
<td>67</td>
<td>67</td>
</tr>
</tbody>
</table>
Results: Ablation Tests

Development Set Results
Exact Match, F1 Measure

Without Conversational Loss    Without Domain Loss    Our Approach

Lucent

BBN
Results: Evaluation Data

F1 Measure

- No Conversation Baseline
- Our Approach
- Supervised

<table>
<thead>
<tr>
<th>Test Case</th>
<th>Lucent Exact Match</th>
<th>Lucent Partial Credit</th>
<th>BBN Exact Match</th>
<th>BBN Partial Credit</th>
</tr>
</thead>
<tbody>
<tr>
<td>F1 Measure</td>
<td>0.250</td>
<td>0.500</td>
<td>0.750</td>
<td>1.000</td>
</tr>
</tbody>
</table>
Towards a joint model of language and perception for object naming

Cynthia Matuszek, Nicholas FitzGerald, Liefeng Bo, Dieter Fox, Luke Zettlemoyer
Learning Attributes

Handle sentences about novel things

“These are the limes”

No longer assuming underlying concepts already exist!

Requires

- Perception models
  - Classifiers: green; round

- Language model
  - How words relate to these detectors

Need a joint model for learning these together!

“a lime”

\[ \lambda x. \text{green}(x) \land \text{round}(x) \]

[Matuszek, FitzGerald, Bo, Fox, Zettlemoyer ICML 12]
Data Collection

UW current effort: object reference

Data collected on Amazon Mechanical Turk
Record interactions, including speech, video and depth info
Ask Turkers to describe what the person is doing
“All of these are yellow toys.”

\[ \lambda x. \text{obj-color}(x, \text{color-yellow}) \]
What is the Parent Saying?

Watch the video, then describe what the parent is saying to the child, in complete sentences.

- Pretend you are a parent teaching a child about something.
- The question is: How does the parent describe this group of objects?

Your answer should be the sentence(s) the parent said while pointing to these things.

“All of these are yellow toys.”

\[ \lambda x. \text{obj-color}(x, \text{color-yellow}) \]
What is the Parent Saying?

Watch the video, then describe what the parent is saying to the child, in complete sentences.

- Pretend you are a parent teaching a child about something.
- The question is: How does the parent describe this group of objects?

Your answer should be the sentence(s) the parent said while pointing to these things.

“All of these are yellow toys.”

\( \lambda x. \) obj-color(\( x \), color-yellow)
Your answer should be the sentence(s) the parent said while pointing to these things.

“Here are some blue shapes.”

\[ \lambda x. \text{obj-color}(x, \text{color-blue}) \]
What is the Parent Saying?

Watch the video, then **describe what the parent is saying to the child**, in complete sentences.

- Pretend you are a parent teaching a child about something.
- The question is: **How does the parent describe this group of objects?**

Your answer should be the sentence(s) the parent said while pointing to these things.

"Here are some blue shapes."

\[ \lambda x. \text{obj-color}(x, \text{color-blue}) \]
Here are some blue shapes.

\[ \lambda x. \text{obj-color}(x, \text{color-blue}) \]
Your answer should be the sentence(s) the parent said while pointing to these things.

"This one’s an orange ball."

\[ \lambda x. \text{obj-color}(x, \text{color-orange}) \land \text{obj-shape}(x, \text{spheroid}) \]
How does the parent describe this group of objects?

“\[ \lambda x. \text{obj-color}(x, \text{color-orange}) \land \text{obj-shape}(x, \text{spheroid}) \]”
Describe the Circled Objects

Look at the image, then describe *only* the circled objects.

- Answer this question:
  
  *How would you describe the objects that are circled (to distinguish them from the rest)?*
  
  - Using complete English sentences
  - Describing the *objects themselves* (not their placement)
  - Click here to review instructions

Your answer should be the description of *only* those objects:

“These are all the green and yellow objects.”

\[ \lambda x. \text{obj-color}(x, \text{color-green}) \lor \text{obj-color}(x, \text{color-yellow}) \]
Describe the Circled Objects

Look at the image, then describe *only* the circled objects.

• Answer this question:
  How would you describe the objects that are circled (to distinguish them from the rest)?
  
  - Using complete English sentences
  - Describing the *objects themselves* (not their placement)
  - Click here to review instructions

Your answer should be the description of *only* those objects:

“This is everything but the blue rectangle.”

\( \lambda x. \neg (\text{obj-color}(x, \text{color-blue}) \land \text{obj-shape}(x, \text{shape-rect}))\)
A Joint Model for Naming Objects

These are the ones that are not blue
A Joint Model for Naming Objects

These are the ones that are not blue

$$\lambda x. \neg \text{color}(x, \text{blue})$$
A Joint Model for Naming Objects

These are the ones that are not blue

\[ \lambda x. \neg \text{color}(x, \text{blue}) \]

Semantic Parsing

Attribute Classification
A Joint Model for Naming Objects

These are the ones that are not blue

\[ \lambda x. \neg \text{color}(x, \text{blue}) \]

Semantic Parsing

Attribute Classification

Blue
A Joint Model for Naming Objects

These are the ones that are not blue

$$\lambda x. \neg \text{color}(x, \text{blue})$$
A Joint Model for Naming Objects

These are the ones that are not blue

\[ \lambda x. \neg \text{color}(x, \text{blue}) \]

Semantic Parsing

Attribute Classification

Blue  Green  Round
A Joint Model for Naming Objects

These are the ones that are not blue

\[ \lambda x. \neg \text{color}(x, \text{blue}) \]

Semantic Parsing

Attribute Classification

Blue Green Round
A Joint Model for Naming Objects

These are the ones that are not blue

\[ \lambda x. \neg \text{color}(x, \text{blue}) \]

Semantic Parsing

Attribute Classification

Blue  Green  Round  Broccoli

...  ...  ...  ...
A Joint Model for Naming Objects

These are the ones that are not blue

\[ \lambda x. \neg \text{color}(x, \text{blue}) \]

Semantic Parsing

Attribute Classification

Grounded Query

Blue  Green  Round  Broccoli
Need Uncertainty!

These are the ones that are not blue

\[ \lambda x. \neg \text{color}(x, \text{blue}) \]

Semantic Parsing

Attribute Classification

Grounded Query

\[ P(G, Z, A | I, X) \]
Need Uncertainty!

These are the ones that are not blue

\[ \lambda x. \neg \text{color}(x, \text{blue}) \]

Semantic Parsing

Grounded Query

P(G,Z,A|I,X)

Attribute Classification

Blue  Green  Round  Broccoli

-  ...  ...  .
Need Uncertainty!

These are the ones that are not blue

\[ \lambda x. \neg \text{color}(x, \text{blue}) \]

P(G,Z,A|I,X)

Semantic Parsing

Grounded Query

Attribute Classification

Blue  Green  Round  Broccoli

Grounded Query

P(G,Z,A|I,X)
Need Uncertainty!

These are the ones that are not blue

\[ \lambda x. \neg \text{color}(x, \text{blue}) \]

Semantic Parsing

Grounded Query

P(G,Z,A | I, X)

Attribute Classification

Blue
Green
Round
Broccoli
Need Uncertainty!

These are the ones that are not blue

\[ \lambda x. \neg \text{color}(x, \text{blue}) \]

Semantic Parsing

Grounded Query

\[ P(G, Z, A | I, X) \]

Attribute Classification

Blue  Green  Round  Broccoli
Need Uncertainty!

These are the ones that are not blue

$\lambda x. \neg \text{color}(x, \text{blue})$

Semantic Parsing

Grounded Query

$P(G, Z, A | I, X)$

Attribute Classification

Blue  Green  Round

Broccoli

Grounded Query
Need Uncertainty!

These are the ones that are not blue

\[ \lambda x. \neg \text{color}(x, \text{blue}) \]

Semantic Parsing

Grounded Query

\[ P(G, Z, A | I, X) = P(Z | X) \prod_{A \in A} P(A | I) \delta(Z(A) = G) \]
Need Uncertainty!

These are the ones that are not blue

\[ \lambda x. \neg \text{color}(x\text{, blue}) \]

Semantic Parsing

Grounded Query

\[ P(G, Z, A \mid I, X) = P(Z \mid X) \prod_{A \in A} P(A \mid I) \delta(Z(A) = G) \]

Parsing Model

Attribute Classification

Blue

Green

Round

Broccoli
Need Uncertainty!

These are the ones that are not blue

\[ \lambda x. \neg \text{color}(x, \text{blue}) \]

Semantic Parsing

Grounded Query

\[ P(G, Z, A | I, X) = P(Z | X) \prod_{A \in A} P(A | I) \delta(Z(A) = G) \]

Attribute Classification

Blue  Green  Round

Grounded Query

Parsing Model

Vision Model
Need Uncertainty!

These are the ones that are not blue

\[ \lambda x. \neg \text{color}(x, \text{blue}) \]

Semantic Parsing

Grounded Query

\[ P(G, Z, A | I, X) = P(Z | X) \prod_{A \in A} P(A | I) \delta(Z(A) = G) \]

Attribute Classification

Blue

Green

Round

Broccoli

Parsing Model

Vision Model

Grounding Query
We can do joint learning!

Language helps determine attribute relations

Language is ambiguous: “This is <new-word>.”

- New color attribute?
- New shape attribute?
- Synonym?
- No attribute at all
We can do joint learning!

Language helps determine attribute relations

Language is ambiguous: “This is <new-word>.”

New color attribute?

“This is pink.”

New shape attribute?

Synonym?

No attribute at all
We can do joint learning!

Language helps determine attribute relations

Language is ambiguous: “This is <new-word>.”

New color attribute? 
“This is pink.”

New shape attribute? 
“This is triangular.”

Synonym?

No attribute at all
We can do joint learning!

Language helps determine attribute relations

Language is ambiguous: “This is $<\text{new-word}>$."

New color attribute?
  “This is pink.”

New shape attribute?
  “This is triangular.”

Synonym?
  “This is red.”

No attribute at all
We can do joint learning!

Language helps determine attribute relations

Language is ambiguous: “This is <new-word>.”

New color attribute?
  “This is pink.”

New shape attribute?
  “This is triangular.”

Synonym?
  “This is red.”

No attribute at all
  “This is an object.”
Fully Supervised Learning

If we label sentence meaning, object groups, and alignment → Decomposes into two independent learning problems

Semantic Parsing

These are the ones that are not blue
\[ \lambda x. \neg \text{color}(x, \text{blue}) \]

... Those are the purple broccoli
\[ \lambda x. \text{color}(x, \text{purp}) \land \text{shape}(x, \text{broc}) \]

Existing Learning Algs

\[ P(Z|X), \text{ a supervised semantic parser} \]

Attribute Classification

Existing Learning Algs

\[ P(A|I), \text{ for all attributes A using logistic regression models} \]

\[ P(G,Z,A|I,X) = P(Z|X) \prod_{A \in A} P(A|I) \delta(Z(A)=G) \]
Unsupervised Learning

What if we don’t want to annotate data (at least not much)

Training Examples:

(These are the ones that are not blue, , , ) ...
Unsupervised Learning

What if we don’t want to annotate data (at least not much)

Training Examples:

(These are the ones that are not blue, , ) ...

Initialization:

• Train a small supervised model
  About 50 labeled scenes (sentence/logic and object/attributes)

• Add $N$ new, unknown attribute classifiers
  - Initialize to the uniform distribution
  - Pair with every unknown word/phrase

$$P(G,Z,A|I,X) = P(Z|X) \prod_{A \in A} P(A|I) \delta(Z(A)=G)$$
Unsupervised Learning

What if we don’t want to annotate data (at least not much)

Training Examples:

(These are the ones that are not blue, ) ...

\[
P(G,Z,A|I,X) = P(Z|X) \prod_{A \in A} P(A|I) \delta(Z(A)=G)
\]
Unsupervised Learning

What if we don’t want to annotate data (at least not much)

Training Examples:

(These are the ones that are not blue, ..., ) ...

Use an EM-style algorithm:

1. Compute latent \( P(Z|X) \) and \( P(A|I) \)
   - Fills in values consistent labeled grounded queries
2. Re-estimate parameters of parsing and vision models
   - Use expectations in place of labels to manage uncertainty

\[
P(G,Z,A|I,X) = P(Z|X) \prod_{A \in A} P(A|I) \delta(Z(A)=G)
\]
Unsupervised Learning

What if we don’t want to annotate data (at least not much)

Training Examples:

\[
P(G,Z,A|I,X) = P(Z|X) \prod_{A \in A} P(A|I) \delta(Z(A)=G)
\]

Use an EM-style algorithm:

1. Compute latent \( P(Z|X) \) and \( P(A|I) \)
   • Fills in values consistent labeled grounded queries
2. Re-estimate parameters of parsing and vision models
   • Use expectations in place of labels to manage uncertainty

Key algorithmic advances

• How to compute the expectations (Step 1) efficiently
• Use parameter estimation to find symbol / classifier alignment
Experimental Evaluation

142 scenes
6 colors and 6 shape attributes
~1,000 NL sentences from Mechanical Turk
Ground truth formulas and classifier assignments

- 20 splits into
  - 30% training items for initialization phase (3 colors, 3 shapes)
  - 55% training items for teaching phase (3 new colors, 3 new shapes)
  - 10% test cases with new colors+shapes
Experimental Evaluation

142 scenes
6 colors and 6 shape attributes
~1,000 NL sentences from Mechanical Turk
Ground truth formulas and classifier assignments

- 20 splits into
  - 30% training items for initialization phase (3 colors, 3 shapes)
  - 55% training items for teaching phase (3 new colors, 3 new shapes)
  - 10% test cases with new colors+shapes

- Precision = 0.85; Recall = 0.8
# Learned Classifiers

<table>
<thead>
<tr>
<th></th>
<th>blue</th>
<th>brown</th>
<th>orange</th>
<th>arch</th>
<th>rect</th>
<th>triangle</th>
<th>yellow</th>
<th>green</th>
<th>red</th>
<th>cylinder</th>
<th>cube</th>
<th>spheroid</th>
</tr>
</thead>
<tbody>
<tr>
<td>blue</td>
<td>1.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.26</td>
<td>0.16</td>
<td>0.17</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.24</td>
<td>0.28</td>
<td>0.00</td>
</tr>
<tr>
<td>brown</td>
<td>0.00</td>
<td>0.99</td>
<td>0.00</td>
<td>0.18</td>
<td>0.23</td>
<td>0.16</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.15</td>
<td>0.08</td>
<td>0.47</td>
</tr>
<tr>
<td>orange</td>
<td>0.00</td>
<td>0.00</td>
<td>0.95</td>
<td>0.18</td>
<td>0.19</td>
<td>0.27</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.26</td>
<td>0.42</td>
<td>0.00</td>
</tr>
<tr>
<td>arch</td>
<td>0.29</td>
<td>0.14</td>
<td>0.15</td>
<td>0.98</td>
<td>0.00</td>
<td>0.29</td>
<td>0.17</td>
<td>0.25</td>
<td>0.22</td>
<td>0.15</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>rect</td>
<td>0.26</td>
<td>0.29</td>
<td>0.24</td>
<td>0.09</td>
<td>0.86</td>
<td>0.13</td>
<td>0.36</td>
<td>0.18</td>
<td>0.31</td>
<td>0.10</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>triangle</td>
<td>0.19</td>
<td>0.22</td>
<td>0.08</td>
<td>0.06</td>
<td>0.08</td>
<td>0.76</td>
<td>0.05</td>
<td>0.15</td>
<td>0.13</td>
<td>0.24</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>NEW</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.20</td>
<td>0.20</td>
<td>0.00</td>
<td>0.99</td>
<td>0.00</td>
<td>0.00</td>
<td>0.18</td>
<td>0.07</td>
<td>0.33</td>
</tr>
<tr>
<td>NEW</td>
<td>0.00</td>
<td>0.03</td>
<td>0.00</td>
<td>0.21</td>
<td>0.12</td>
<td>0.28</td>
<td>0.00</td>
<td>1.00</td>
<td>0.00</td>
<td>0.23</td>
<td>0.22</td>
<td>0.15</td>
</tr>
<tr>
<td>NEW</td>
<td>0.00</td>
<td>0.00</td>
<td>0.07</td>
<td>0.23</td>
<td>0.26</td>
<td>0.09</td>
<td>0.00</td>
<td>0.00</td>
<td>0.93</td>
<td>0.20</td>
<td>0.17</td>
<td>0.24</td>
</tr>
<tr>
<td>NEW</td>
<td>0.52</td>
<td>0.46</td>
<td>0.51</td>
<td>0.04</td>
<td>0.07</td>
<td>0.37</td>
<td>0.23</td>
<td>0.39</td>
<td>0.30</td>
<td>0.90</td>
<td>0.86</td>
<td>0.28</td>
</tr>
<tr>
<td>NEW</td>
<td>0.39</td>
<td>0.43</td>
<td>0.44</td>
<td>0.42</td>
<td>0.38</td>
<td>0.43</td>
<td>0.45</td>
<td>0.46</td>
<td>0.40</td>
<td>0.47</td>
<td>0.44</td>
<td>0.45</td>
</tr>
<tr>
<td>NEW</td>
<td>0.12</td>
<td>0.27</td>
<td>0.17</td>
<td>0.10</td>
<td>0.05</td>
<td>0.08</td>
<td>0.34</td>
<td>0.13</td>
<td>0.28</td>
<td>0.23</td>
<td>0.37</td>
<td>0.87</td>
</tr>
</tbody>
</table>
Failure cases

Bad parses:

“This is a red, toy triangle.”

\[ \lambda x. \text{obj-shape}(x, \text{shape-triangle}) \land \text{obj-shape}(x, \text{shape-rect}) \]

Bad classification:

Cylinders (lengthwise) look like rectangles; cylinders (upright) look like cubes
Failure cases

Bad parses:

“This is a red, toy triangle.”

\[ \lambda x. \text{obj-shape}(x, \text{shape-triangle}) \land \text{obj-shape}(x, \text{shape-rect}) \]

Two different shapes instead of a shape and a color

Bad classification:

Cylinders (lengthwise) look like rectangles; cylinders (upright) look like cubes
Failure cases

Bad parses:

“This is a red, toy triangle.”

\[ \lambda x. \text{obj-shape}(x, \text{shape-triangle}) \land \text{obj-shape}(x, \text{shape-rect}) \]

Two different shapes instead of a shape and a color

Bad classification:

Cylinders (lengthwise) look like rectangles; cylinders (upright) look like cubes

Humans made the same errors, even with pre-segmentation view
2 Types of Grounded Learning

• Learning from conversation logs
  – Loss driven, requiring no annotations
  – First step towards autonomous, self-improving dialog systems

• Joint model of language and visual attributes
  – Full modeling pipeline from pixels to meaning
  – Can extend vocabulary by pairing new words with learned visual attributes
Current Work: Self-learning dialog systems

Goal: learn full dialog semantic parser the produces context-dependent analyses

USER: i would like to fly from atlanta georgia to london england on september twenty fourth in the early evening i would like to return on october first departing from london in the late morning

SYSTEM: leaving what city? (ASK:λx.from(fl, x))

USER: atlanta georgia.

SYSTEM: leaving atlanta. (CONF:from(fl, ATL)) going to which city? (ASK:λx.to(fl, x))

USER: london

SYSTEM: arriving in (CONF:to(fl, LON)) what date would to depart Atlanta? (ASK:λx.from(fl, ATL) ∧ depart-date(fl, x))

USER: september twenty fourth in the early evening

[conversation continues]
Future Work: Mapping Speech to Logical From

Uh, flights New York to Singapore, sure

ACCEPT: \( \lambda x. \text{flight}(x) \land \text{from}(x, \text{NYC}) \land \text{to}(x, \text{SIN}) \)
Future Work: 3D Object Reference Tasks

For Example: a scene from the Google 3D Warehouse

Many possible sentences:
The four red chairs are around the kitchen table.
There are two stools at the bar between the sink and the television.
...

Future: Scaling Up Parsers for Databases

• Bigger data sets!
  – Can we build an NL interface to Freebase?
  – Will need more efficient parsing algorithms and semi-supervised learning
Learning Grounded Semantic Parsers

Special thanks to the team!!!!

(not all pictured) Yoav Artzi, Nicholas FitzGerald, Tom Kwiatkowski, Cynthia Matuszek, Liefeng Bo, Dieter Fox, Adrienne Wang, Mark Yatskar

for more info: http://www.cs.washington.edu/homes/lsz/