Computing with Natural Language

Percy Liang

ACL Workshop on Semantic Parsing - June 15, 2014

Stanford University
Paleobiology

The transition between the Pleistocene and Holocene was a dramatic event that had a profound impact on the evolution of life on Earth. During the Pleistocene, mass extinction events were common, but the Holocene saw a recovery of biodiversity. This transition is marked by the end of the last glacial period, which ended around 11,700 years ago. The Holocene, characterized by a warmer climate, saw the emergence of modern human populations and the development of agriculture.

The Pleistocene-Holocene transition is marked by a significant shift in the Earth's climate, with alternating periods of glacial and interglacial phases. During the glacial phases, sea levels were lower due to the increased volume of ice, which allowed for the expansion of land masses. This resulted in the formation of extensive continental shelves and the exposure of large areas of the ocean floor. The Holocene, on the other hand, is characterized by a rise in sea levels, which has had significant implications for coastal ecosystems and human societies.

The transition between these two periods is a critical time for understanding the evolution of life on Earth. It is during this time that many key developments occurred, including the emergence of modern humans and the development of agriculture. The study of this transition is crucial for understanding the impact of climate change on the evolution of life and for developing strategies to mitigate future environmental challenges.

References:
- paleobiology.org
- journal.plosone.org
Paleobiology

The interaction between the frigid oceanic waters beginning 14 million years ago is responsible for the transformation of continental shelves into mature permafrost environments. These permafrost environments have been studied in various regions, including the Rocky Mountains, the Canadian Arctic, and the Siberian Shelf. The rock layers have been altered due to the tectonic activity of the North American Plate, which has caused the movement of tectonic plates. The result is a unique set of geological structures that have been preserved for millions of years.

Paleobiology

paleobiodb.org

[PaleoDeepDive (Shanan Peters, Chris Ré)]
Where was the last American Mastodon found?
Where was the last American Mastadon found?

How long do species tend to exist before going extinct?
Where was the last American Mastadon found?

How long do species tend to exist before going extinct?

**Goal:** help scientists answer macro-questions

**Challenge:** requires computation / aggregation
Question answering via semantic parsing

*Where was the last American Mastadon found?*
Question answering via semantic parsing

Where was the last American Mastodon found?

semantic parsing

LocationOf.argmax(Type.Occurrence \sqcap Genus.Mammut, Period)
Question answering via semantic parsing

Where was the last American Mastadon found?

semantic parsing

\[ \text{LocationOf}. \text{argmax}(\text{Type.Occurrence} \sqcap \text{Genus.Mammut}, \text{Period}) \]

execute

New Mexico
Question answering via semantic parsing

*Where was the last American Mastadon found?*

- semantic parsing
- execute
- New Mexico
Email assistant via semantic parsing

Send a reminder to all authors who haven’t sent an abstract.
Send a reminder to all authors who haven’t sent an abstract.

∀x ∈ (Author ⊓ ¬Sent.Subject.Abstract) : Remind(x)
Email assistant via semantic parsing

Send a reminder to all authors who haven’t sent an abstract.

∀x ∈ (Author ⊓ ¬Sent.Subject.Abstract) : Remind(x)

execute

[5 emails sent]
Email assistant via semantic parsing

Send a reminder to all authors who haven’t sent an abstract.

semantic parsing

execute

[5 emails sent]
Semantic parsing

[utterance: user input]

semantic parsing

[program]

execute

[behavior: user output]

Programs affect the world
Outline

- **Semantic parsing in 5 minutes**
- A closer look at the elements
  - Knowledge base incompleteness
  - Lexical coverage
  - Search over logical forms
  - Learning via bootstrapping
  - Leveraging denotations ("grounding")
  - Datasets
- Final remarks
Framework

people who have lived in Chicago

Type.Person \sqcap \text{PlacesLived.Location.Chicago}

\{BarackObama, MichelleObama, \ldots\}
World: Freebase

100M entities (nodes) 1B assertions (edges)

MichelleObama
- Gender: Female
- PlacesLived: Event21
- Spouse: BarackObama

BarackObama
- DateOfBirth: 1961.08.04
- Profession: Politician
- PlaceOfBirth: Honolulu
- ContainedBy: UnitedStates
- ContainedBy: Hawaii

Event3
- Type: Person

Event8
- StartDate: 1992.10.03
- ContainedBy: UnitedStates

Event21
- Location: Chicago

UnitedStates
- ContainedBy: Hawaii

Hawaii
- Type: USState

UnitedStates
- Type: Country

Hawaii
- Type: City

Chicago
- Location: Event3

Honolulu
- Type: City
Logical forms

Type.Person ⊓ PlacesLived.Location.Chicago
Logical forms

Type.Person ⊓ PlacesLived.Location.Chicago

[Li, 2013]

From Liang, 2013
Logical forms

Type.Person \n PlacesLived.Location.Chicago

[Link to Liang, 2013]
Logical forms

Type: Person \sqcap PlacesLived: Location: Chicago

[Diagram of logical forms with entities and relationships such as Barack Obama, Michelle Obama, Chicago, Honolulu, etc.]
Framework

$x$ people who have lived in Chicago

parameters

$\theta$ $\zeta$ Type.Person $\sqcap$ PlacesLived.Location.Chicago

world

$w$ $y$ $\{BarackObama,MichelleObama,\ldots\}$
Derivations

Derivation: construction of logical form given utterance

Type.Location ⊓ R[PlaceOfBirth].BarackObama

where

Type.Location

was

R[PlaceOfBirth].BarackObama

born

Obama

where

BarackObama

R[PlaceOfBirth]
Derivations

**Derivation**: construction of logical form given utterance

```
Type.Location \sqcap R[PlaceOfBirth].BarackObama
```

```
Type.Location  was  R[PlaceOfBirth].BarackObama  ?

where
```

```
lexicon
```

```
BarackObama  R[PlaceOfBirth]

lexicon
```

```
Obama  born
```

```
lexicon
```
Derivations

**Derivation**: construction of logical form given utterance

\[
\text{Type.Location} \sqcap R[\text{PlaceOfBirth}].\text{BarackObama}
\]

\[
\text{Type.Location} \quad \text{was} \quad R[\text{PlaceOfBirth}].\text{BarackObama} \quad ?
\]

\[
\text{lexicon} \quad \text{where} \quad \text{lexicon}
\]

\[
\text{lexicon} \quad \text{lexicon}
\]

\[
\text{lexicon} \quad \text{lexicon}
\]

\[
\text{lexicon} \quad \text{lexicon}
\]

\[
\text{lexicon} \quad \text{lexicon}
\]

\[
\text{lexicon} \quad \text{lexicon}
\]

\[
\text{lexicon} \quad \text{lexicon}
\]

\[
\text{lexicon} \quad \text{lexicon}
\]

\[
\text{lexicon} \quad \text{lexicon}
\]

\[
\text{lexicon} \quad \text{lexicon}
\]

\[
\text{lexicon} \quad \text{lexicon}
\]

\[
\text{lexicon} \quad \text{lexicon}
\]

\[
\text{lexicon} \quad \text{lexicon}
\]

\[
\text{lexicon} \quad \text{lexicon}
\]

\[
\text{lexicon} \quad \text{lexicon}
\]

\[
\text{lexicon} \quad \text{lexicon}
\]

\[
\text{lexicon} \quad \text{lexicon}
\]

\[
\text{lexicon} \quad \text{lexicon}
\]

\[
\text{lexicon} \quad \text{lexicon}
\]

\[
\text{lexicon} \quad \text{lexicon}
\]

\[
\text{lexicon} \quad \text{lexicon}
\]

\[
\text{lexicon} \quad \text{lexicon}
\]

\[
\text{lexicon} \quad \text{lexicon}
\]

\[
\text{lexicon} \quad \text{lexicon}
\]

\[
\text{lexicon} \quad \text{lexicon}
\]

\[
\text{lexicon} \quad \text{lexicon}
\]

\[
\text{lexicon} \quad \text{lexicon}
\]

\[
\text{lexicon} \quad \text{lexicon}
\]

\[
\text{lexicon} \quad \text{lexicon}
\]

\[
\text{lexicon} \quad \text{lexicon}
\]

\[
\text{lexicon} \quad \text{lexicon}
\]

\[
\text{lexicon} \quad \text{lexicon}
\]

\[
\text{lexicon} \quad \text{lexicon}
\]

\[
\text{lexicon} \quad \text{lexicon}
\]

\[
\text{lexicon} \quad \text{lexicon}
\]

\[
\text{lexicon} \quad \text{lexicon}
\]

\[
\text{lexicon} \quad \text{lexicon}
\]

\[
\text{lexicon} \quad \text{lexicon}
\]

\[
\text{lexicon} \quad \text{lexicon}
\]

\[
\text{lexicon} \quad \text{lexicon}
\]

\[
\text{lexicon} \quad \text{lexicon}
\]

\[
\text{lexicon} \quad \text{lexicon}
\]

\[
\text{lexicon} \quad \text{lexicon}
\]

\[
\text{lexicon} \quad \text{lexicon}
\]

\[
\text{lexicon} \quad \text{lexicon}
\]

\[
\text{lexicon} \quad \text{lexicon}
\]

\[
\text{lexicon} \quad \text{lexicon}
\]

\[
\text{lexicon} \quad \text{lexicon}
\]

\[
\text{lexicon} \quad \text{lexicon}
\]

\[
\text{lexicon} \quad \text{lexicon}
\]

\[
\text{lexicon} \quad \text{lexicon}
\]

\[
\text{lexicon} \quad \text{lexicon}
\]

\[
\text{lexicon} \quad \text{lexicon}
\]

\[
\text{lexicon} \quad \text{lexicon}
\]

\[
\text{lexicon} \quad \text{lexicon}
\]

\[
\text{lexicon} \quad \text{lexicon}
\]

\[
\text{lexicon} \quad \text{lexicon}
\]

\[
\text{lexicon} \quad \text{lexicon}
\]

\[
\text{lexicon} \quad \text{lexicon}
\]

\[
\text{lexicon} \quad \text{lexicon}
\]

\[
\text{lexicon} \quad \text{lexicon}
\]

\[
\text{lexicon} \quad \text{lexicon}
\]

\[
\text{lexicon} \quad \text{lexicon}
\]

\[
\text{lexicon} \quad \text{lexicon}
\]

\[
\text{lexicon} \quad \text{lexicon}
\]

\[
\text{lexicon} \quad \text{lexicon}
\]

\[
\text{lexicon} \quad \text{lexicon}
\]

\[
\text{lexicon} \quad \text{lexicon}
\]

\[
\text{lexicon} \quad \text{lexicon}
\]

\[
\text{lexicon} \quad \text{lexicon}
\]

\[
\text{lexicon} \quad \text{lexicon}
\]

\[
\text{lexicon} \quad \text{lexicon}
\]

\[
\text{lexicon} \quad \text{lexicon}
\]

\[
\text{lexicon} \quad \text{lexicon}
\]

\[
\text{lexicon} \quad \text{lexicon}
\]

\[
\text{lexicon} \quad \text{lexicon}
\]
**Derivations**

**Derivation:** construction of logical form given utterance

\[
\text{Type.Location} \cap R[\text{PlaceOfBirth}].\text{BarackObama}
\]

\[
\text{intersect}
\]

\[
\text{Type.Location} \quad \text{was} \quad R[\text{PlaceOfBirth}].\text{BarackObama} \quad ?
\]

\[
\text{join}
\]

\[
\text{lexicon}
\]

\[
\text{where}
\]

\[
\text{BarackObama} \quad R[\text{PlaceOfBirth}]
\]

\[
\text{lexicon}
\]

\[
\text{lexicon}
\]

\[
\text{Obama} \quad \text{born}
\]
Grammar

utterance \rightarrow \text{Grammar} \rightarrow \text{derivation 1}
\rightarrow \text{derivation 2}
\rightarrow \ldots
Grammar

A Really Dumb Grammar

(lexicon) Obama ⇒ Unary : BarackObama
(lexicon) born ⇒ Binary : PlaceOfBirth
...
(join) Unary : u Binary : b ⇒ Unary : b.u
(intersect) Unary : u Unary : v ⇒ Unary : u ∩ v
Many possible derivations!

*Where was Obama born?*
Many possible derivations!

Where was Obama born?

\[ \downarrow \quad ? \]

set of candidate derivations \( \mathcal{D}(x) \)
Many possible derivations!

Where was Obama born?

set of candidate derivations $\mathcal{D}(x)$
Many possible derivations!

Where was Obama born?

? 12

set of candidate derivations $\mathcal{D}(x)$
\( x: \text{utterance} \)

\( d: \text{derivation} \)

Feature vector \( \phi(x, d) \in \mathbb{R}^f \):
Feature vector $\phi(x, d) \in \mathbb{R}^f$:

- apply join 1
- apply intersect 1
- apply lexicon 3
- skipped VBD-AUX 1
- skipped NN 0
- \textit{born} maps to \texttt{PlaceOfBirth} 1
- \textit{born} maps to \texttt{PlacesLived.Location} 0
- alignmentScore 1.52
- denotation-size=1 1
- ...
- ...
Scoring derivations

Feature vector: $\phi(x, d) = [1.3, 2, 0, 1, 0, 0, \ldots] \in \mathbb{R}^F$

Parameter vector: $\theta = [1.2, -2.7, 3.4, \ldots] \in \mathbb{R}^F$

Scoring function:

$$\text{Score}_\theta(x, d) = \phi(x, d) \cdot \theta$$
Log-linear model

Candidate derivations: $D(x)$
Log-linear model

Candidate derivations: $D(x)$

Model: distribution over derivations $d$ given utterance $x$

$$p(d \mid x, \theta) = \frac{\exp(\text{Score}_\theta(x,d))}{\sum_{d' \in D(x)} \exp(\text{Score}_\theta(x,d'))}$$
Learning

Training data:

- What’s Bulgaria’s capital?
  - Sofia
- When was Walmart started?
  - 1962
- What movies has Tom Cruise been in?
  - TopGun, VanillaSky, ...
  - ...
Learning

Training data:

What’s Bulgaria’s capital?
Sofia

When was Walmart started?
1962

What movies has Tom Cruise been in?
TopGun, VanillaSky, ...

... 

Objective: Maximum likelihood

$$\arg \max_\theta \sum_{i=1}^n \log p_\theta(y^{(i)} \mid x^{(i)})$$
Learning

Training data:

What’s Bulgaria’s capital?
Sofia
When was Walmart started?
1962
What movies has Tom Cruise been in?
TopGun, VanillaSky, ...
...

Objective: Maximum likelihood

$$\arg\max_{\theta} \sum_{i=1}^{n} \log p_{\theta}(y^{(i)} | x^{(i)})$$

Algorithm:

AdaGrad (SGD with per-feature step size)
Training intuition

Where did Mozart tupress?

Vienna
Training intuition

Where did Mozart tupress?

PlaceOfBirth.Mozart
PlaceOfDeath.Mozart
PlaceOfMarriage.Mozart

Vienna
Training intuition

Where did Mozart tupress?

PlaceOfBirth.Mozart ⇒ Salzburg
PlaceOfDeath.Mozart ⇒ Vienna
PlaceOfMarriage.Mozart ⇒ Vienna

Vienna
Training intuition

*Where did Mozart tupress?*

PlaceOfBirth.Mozart $\Rightarrow$ Salzburg

PlaceOfDeath.Mozart $\Rightarrow$ Vienna

PlaceOfMarriage.Mozart $\Rightarrow$ Vienna

Vienna
Training intuition

Where did Mozart typress?

PlaceOfBirth.Mozart ⇒ Salzburg
PlaceOfDeath.Mozart ⇒ Vienna
PlaceOfMarriage.Mozart ⇒ Vienna

Vienna

Where did William Hogarth typress?
Training intuition

Where did Mozart tupress?

PlaceOfBirth.Mozart ⇒ Salzburg
PlaceOfDeath.Mozart ⇒ Vienna
PlaceOfMarriage.Mozart ⇒ Vienna

Vienna

Where did William Hogarth tupress?

PlaceOfBirth.WilliamHogarth
PlaceOfDeath.WilliamHogarth
PlaceOfMarriage.WilliamHogarth

London
Training intuition

Where did Mozart tupress?

PlaceOfBirth.Mozart ⇒ Salzburg
PlaceOfDeath.Mozart ⇒ Vienna
PlaceOfMarriage.Mozart ⇒ Vienna

Vienna

Where did William Hogarth tupress?

PlaceOfBirth.WilliamHogarth ⇒ London
PlaceOfDeath.WilliamHogarth ⇒ London
PlaceOfMarriage.WilliamHogarth ⇒ Paddington

London
Training intuition

Where did Mozart tupress?

PlaceOfBirth.Mozart ⇒ Salzburg
PlaceOfDeath.Mozart ⇒ Vienna
PlaceOfMarriage.Mozart ⇒ Vienna

Vienna

Where did William Hogarth tupress?

PlaceOfBirth.WilliamHogarth ⇒ London
PlaceOfDeath.WilliamHogarth ⇒ London
PlaceOfMarriage.WilliamHogarth ⇒ Paddington

London
Training intuition

*Where did Mozart tupress?*

- **PlaceOfBirth.Mozart** $\Rightarrow$ Salzburg
- **PlaceOfDeath.Mozart** $\Rightarrow$ Vienna
- **PlaceOfMarriage.Mozart** $\Rightarrow$ Vienna

**Vienna**

*Where did William Hogarth tupress?*

- **PlaceOfBirth.WilliamHogarth** $\Rightarrow$ London
- **PlaceOfDeath.WilliamHogarth** $\Rightarrow$ London
- **PlaceOfMarriage.WilliamHogarth** $\Rightarrow$ Paddington

**London**
Outline

• Semantic parsing in 5 minutes

• **A closer look at the elements**
  – Knowledge base incompleteness
  – Lexical coverage
  – Search over logical forms
  – Learning via bootstrapping
  – Leveraging denotations ("grounding")
  – Datasets

• Final remarks
Outline

• Semantic parsing in 5 minutes
• A closer look at the elements
  – Knowledge base incompleteness
  – Lexical coverage
  – Search over logical forms
  – Learning via bootstrapping
  – Leveraging denotations ("grounding")
  – Datasets
• Final remarks
Challenge: incomplete knowledge base

What are the longest hiking trails in Baltimore?

Data Source

hiking trails in Baltimore
- Avalon Super Loop
- Patapsco Valley State Park
- Gunpowder Falls State Park
- Union Mills Hike
- Greenbury Point
...
Fewer than 10% general web questions can be answered via Freebase
Semantic parsing on the web

Input:

- query $x$
  - hiking trails near Baltimore
- web page $w$
Semantic parsing on the web

Input:

- query: hiking trails near Baltimore
- web page: [Pasupat & Liang, 2014]
Semantic parsing on the web

Input:
• query: hiking trails near Baltimore
• web page: [Pasupat & Liang, 2014]
Semantic parsing on the web

Input:

- query $x$
  
  *hiking trails near Baltimore*

- web page $w$

Output:

- list of entities $y$

  [Avalon Super Loop, Patapsco Valley State Park, ...]
Logical forms: XPath expressions

\[ z = /html[1]/body[1]/table[2]/tr/td[1] \]
Framework

hiking trails near Baltimore

\( x \)
Framework

hiking trails near Baltimore

(|Z| ≈ 8500)
Framework

hiking trails near Baltimore

(|Z| \approx 8500)

Model

Generation

/html[1]/body[1]/table[2]/tr/td[1]
hiking trails near Baltimore

$h_{|Z| \approx 8500}$

[Avalon Super Loop, Patapsco Valley State Park, ...]
Outline

- Semantic parsing in 5 minutes
- A closer look at the elements
  - Knowledge base incompleteness
  - **Lexical coverage**
  - Search over logical forms
  - Learning via bootstrapping
  - Leveraging denotations (“grounding”)
  - Datasets
- Final remarks
Challenge: lexical coverage

born ⇒ Type.City, PeopleBornHere, Profession.Lawyer, ...
Solution: alignment

Open information extraction on ClueWeb09:

(Barack Obama, was born in, Honolulu)
(Albert Einstein, was born in, Ulm)
(Barack Obama, lived in, Chicago)

... 15M triples ...
Solution: alignment

Open information extraction on ClueWeb09:

(Barack Obama, was born in, Honolulu)
(Albert Einstein, was born in, Ulm)
(Barack Obama, lived in, Chicago)
... 15M triples ...

Freebase:

(BarackObama, PlaceOfBirth, Honolulu)
(Albert Einstein, PlaceOfBirth, Ulm)
(BarackObama, PlacesLived.Location, Chicago)
... 400M triples ...
Match text and Freebase predicates

grew up in \rightarrow DateOfBirth
born in \rightarrow PlaceOfBirth
married in \rightarrow Marriage.StartDate
born in \rightarrow PlacesLived.Location

Similar schema matching / alignment ideas [Cai & Yates, 2013, Fader et. al, 2013, Yao & van Durme, 2014; etc.]
Challenge: variability in language

*What is the currency in the US?*
Challenge: variability in language

What is the currency in the US?

What money do they use in the states?

How do you pay in America?

What’s the currency of the US?

What money is accepted in the United States?

What money to take to the US?

...
A solution: paraphrasing

How many people live in Seattle?

What is the population of Seattle?

PopulationOf(Seattle)

850,000

paraphrase

What is the population of Seattle?

Convert to a text-only problem
Challenge: “sub-lexical compositionality”

grandmother

\( \lambda x. \text{Gender.Female} \sqcap \text{Parent.Parent}.x \)

mayor

\( \lambda x. \text{GovtPositionsHeld}.(\text{Title.Mayor} \sqcap \text{OfficeOfJurisdiction}.x) \)
Challenge: “sub-lexical compositionality”

grandmother

$\lambda x. \text{Gender.Female} \sqcap \text{Parent.Parent}.x$

mayor

$\lambda x. \text{GovtPositionsHeld.}(\text{Title.Mayor} \sqcap \text{OfficeOfJurisdiction}.x)$

presidents who have served two non-consecutive terms

[requires higher-order quantification]

presidents who were previously vice-presidents

[anaphora]

every other president

[weird quantification anaphora]
Outline

• Semantic parsing in 5 minutes
• A closer look at the elements
  – Knowledge base incompleteness
  – Lexical coverage
  – **Search over logical forms**
  – Learning via bootstrapping
  – Leveraging denotations ("grounding")
  – Datasets
• Final remarks
Many possible derivations!

Where was Obama born?

- A Really Dumb Grammar
  (lexicon) \( Obama \Rightarrow \) Unary : BarackObama
  (lexicon) \( born \Rightarrow \) Binary : PlaceOfBirth
  ...
  (join) Unary : \( u \) Binary : \( b \) \( \Rightarrow \) Unary : \( b.u \)
  (intersect) Unary : \( u \) Unary : \( v \) \( \Rightarrow \) Unary : \( u \cap v \)

set of candidate derivations \( \mathcal{D}(x) \)
Bridging

Type. University

alignment

Which college did Obama go to?

Barack Obama

alignment

Did Obama go to college?
Bridging: use neighboring predicates / type constraints
Bridging: use neighboring predicates / type constraints

Start building from parts with more certainty
Bridging to nowhere

Search logical forms based on "prior":

What countries in the world speak Arabic?
Bridging to nowhere

Search logical forms based on "prior":

What countries in the world speak Arabic?

ArabicAlphabet ArabicLang

[Berant & Liang, 2014]
Bridging to nowhere

Search logical forms based on "prior":

What countries in the world speak Arabic?

ArabicAlphabet

ArabicLang

LangSpoken.ArabicLang

LangFamily.Arabic

[Berant & Liang, 2014]
Bridging to nowhere

Search logical forms based on ”prior”:

What countries in the world speak Arabic?

ArabicAlphabet

ArabicLang

LangSpoken.ArabicLang

LangFamily.Arabic

Type.Country △ LangSpoken.ArabicLang

Count(Type.Country △ LangSpoken.ArabicLang)
Bridging to nowhere

Search logical forms based on ”prior”:

What countries in the world speak Arabic?

Start building from parts with more certainty

[Berant & Liang, 2014]
For what fraction of utterances was a candidate logical form correct?

[Berant et al., 2013] Paraphrasing

[Bar chart showing comparison between [Berant et al., 2013] and Paraphrasing]
Overapproximation via simple grammars

- Modeling correct derivations requires complex rules
Overapproximation via simple grammars

- Modeling correct derivations requires complex rules
- Simple rules generate overapproximation of good derivations
Overapproximation via simple grammars

- Modeling correct derivations requires complex rules
- Simple rules generate overapproximation of good derivations
- Hard grammar rules $\Rightarrow$ soft/overlapping features

- Hard grammar rules $\Rightarrow$ soft/overlapping features
Outline

• Semantic parsing in 5 minutes

• A closer look at the elements
  – Knowledge base incompleteness
  – Lexical coverage
  – Search over logical forms
  – **Learning via bootstrapping**
  – Leveraging denotations ("grounding")
  – Datasets

• Final remarks
Bootstrapping from easy examples

Iteration 1

Example 1  Example 2  Example 3  Example 4  Example 5

...  ...  ...  ...  ...

40
Bootstrapping from easy examples

Iteration 2

Example 1  Example 2  Example 3  Example 4  Example 5

... ... ... ... ...
Bootstrapping from easy examples

Iteration 3

<table>
<thead>
<tr>
<th>Example 1</th>
<th>Example 2</th>
<th>Example 3</th>
<th>Example 4</th>
<th>Example 5</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
</tbody>
</table>

...
Bootstrapping from easy examples

Iteration 4

Example 1  Example 2  Example 3  Example 4  Example 5

...  ...  ...  ...  ...
Bootstrapping from easy examples

On GeoQuery [Liang et al., 2011]:

![Graph showing the percentage of train examples over iterations]

- % train examples
- iteration
- Bootstrapping from easy examples
- On GeoQuery [Liang et al., 2011]:
Outline

• Semantic parsing in 5 minutes
• A closer look at the elements
  – Knowledge base incompleteness
  – Lexical coverage
  – Search over logical forms
  – Learning via bootstrapping
  – **Leveraging denotations** ("grounding")
  – Datasets
• Final remarks
Feature vector $\phi(x, d) \in \mathbb{R}^f$:
Feature vector $\phi(x, d) \in \mathbb{R}^f$:

- apply join: 1
- apply intersect: 1
- apply lexicon: 3
- skipped VBD-AUX: 1
- skipped NN: 0
- *born* maps to *PlaceOfBirth*: 1
- *born* maps to *PlacesLived.Location*: 0
- alignmentScore: 1.52
- denotation-size=1: 1
- ...
- ...
Denotation features for entity extraction

/html[1]/body[1]/table[2]/tr/td[1]

> hiking trails near Baltimore
  Avalon Super Loop
  Patapsco Valley State Park
  Gunpowder Falls State Park
  Rachel Carson Conservation Park
  Union Mills Hike
  ...

/html[1]/body[1]/div[2]/a

> hiking trails near Baltimore
  Home
  About Baltimore Tour
  Pricing
  Contact
  Online Support
  ...

Impact of denotation features

Free917
Impact of denotation features

Free917

WebQuestions
Impact of denotation features

![Bar chart showing the impact of denotation features](chart.png)
Impact of denotation features

Working with denotations actually provides more information than just logical forms
Outline

• Semantic parsing in 5 minutes
• A closer look at the elements
  – Knowledge base incompleteness
  – Lexical coverage
  – Search over logical forms
  – Learning via bootstrapping
  – Leveraging denotations ("grounding")
  – Datasets
• Final remarks
Dataset collection

Obtain naturally occurring questions (inputs)
Dataset collection

Obtain naturally occurring questions (inputs)

Strategy: breadth-first search over Google Suggest graph
Dataset collection

Obtain naturally occurring questions (inputs)

**Strategy:** breadth-first search over Google Suggest graph

*Where was Barack Obama born?*
Dataset collection

Obtain naturally occurring questions (inputs)

Strategy: breadth-first search over Google Suggest graph

Where was Barack Obama born?

Where was _ born?

Google Suggest

Barack Obama
Lady Gaga
Steve Jobs
Dataset collection

Obtain naturally occurring questions (inputs)

Strategy: breadth-first search over Google Suggest graph

Where was Barack Obama born?

Where was born?

Google Suggest
Barack Obama
Lady Gaga
Steve Jobs

Where was Steve Jobs born?
Dataset collection

Obtain naturally occurring questions (inputs)

Strategy: breadth-first search over Google Suggest graph

Where was Barack Obama born?

Where was _ born?

Google Suggest: Barack Obama, Lady Gaga, Steve Jobs

Where was Steve Jobs born?

Where was Steve Jobs _?

Google Suggest: born, raised, on the Forbes list
Dataset collection

Obtain naturally occurring questions (inputs)

Strategy: breadth-first search over Google Suggest graph

Where was Barack Obama born?

Where was _ born?

Where was Steve Jobs born?

Where was Steve Jobs _?

Where was Steve Jobs raised?
Dataset collection

Obtain naturally occurring questions (inputs)

Strategy: breadth-first search over Google Suggest graph

Where was Barack Obama born?

Where was _ born?

Google Suggest

Barack Obama
Lady Gaga
Steve Jobs

Where was Steve Jobs born?

Where was Steve Jobs _?

Google Suggest

born
raised
on the Forbes list

Where was Steve Jobs raised?

...

AMT annotation ⇒ 6.6K question/answer pairs
Question answering on WebQuestions dataset (6K questions) [Berant et al., 2013]

what did obama study in school
where to fly into bali
what was tupac name in juice
Question answering on WebQuestions dataset (6K questions) [Berant et al., 2013]

what did obama study in school

where to fly into bali

what was tupac name in juice
OpenWeb dataset

airlines of italy
natural causes of global warming
lsu football coaches
bf3 submachine guns
badminton tournaments
foods high in dha
technical colleges in south carolina
songs on glee season 5
singers who use auto tune
san francisco radio stations
Results on OpenWeb

Baseline
(Most frequent extraction predicates)

[Pasupat & Liang, 2014]

10.3

40.5

(Most frequent extraction predicates)
A new dataset?

compositional AND open-domain
A new dataset?

compositional AND open-domain

How old are presidents when they take office on average?
A new dataset?

compositional AND open-domain

How old are presidents when they take office on average?

52
A new dataset?

compositional AND open-domain

How old are presidents when they take office on average?
Other tasks

Playing computer games [Branavan et al., 2010, 2011]

Following navigational instructions [Tellex et. al 2011; Chen et. al 2012; Artzi et. al 2013]

Understanding visual world [Matuszek et. al; 2012, Krishnamurthy & Kollar, 2013]

Solving algebra word problems [Kushman et. al, 2013]
Outline

- Semantic parsing in 5 minutes
- A closer look at the elements
  - Knowledge base incompleteness
  - Lexical coverage
  - Search over logical forms
  - Learning via bootstrapping
  - Leveraging denotations (“grounding”)
  - Datasets

- Final remarks
SHRDLU [1971]
Person: Pick up a big red block.

Computer: OK.
**SHRDLU [1971]**

**Person:** Pick up a big red block.

**Computer:** OK.

**Person:** Grasp the pyramid.

**Computer:** I don’t understand which pyramid you mean.
**Person**: Pick up a big red block.

**Computer**: OK.

**Person**: Grasp the pyramid.

**Computer**: I don’t understand which pyramid you mean.

**Person** (changing their mind): Find a block which is taller than the one you are holding and put it into the box.

**Computer**: By ”it”, I assume you mean the block which is taller than the one I am holding.
Person: Pick up a big red block.
Computer: OK.
Person: Grasp the pyramid.
Computer: I don’t understand which pyramid you mean.
Person (changing their mind): Find a block which is taller than the one you are holding and put it into the box.
Computer: By ”it”, I assume you mean the block which is taller than the one I am holding.
Computer: OK.
Person: What does the box contain?
Computer: The blue pyramid and the blue block.
Person: What is the pyramid supported by?
Computer: The box.
Takeaway 1/3: denotations

[utterance: user input]

semantic parsing

[intermediate semantic representations (text? logical forms? vectors?)]

execute

[denotation: user output]

Semantic representations as a means to an end
Takeaway 2/3: learning and search

Challenging search / learning problem
Takeaway 2/3: learning and search

Challenging search / learning problem

Non-convex optimization
Takeaway 2/3: learning and search

Challenging search / learning problem

- Non-convex optimization
- Exponential search space
Takeaway 2/3: learning and search

Challenging search / learning problem

- Non-convex optimization

- Exponential search space

Need to create better abstractions for people to work on the core search/learning issues
Takeaway 3/3: data and users

Semantic parsing provides utility to users

Users provide get back realistic datasets

How long do species tend to exist before going extinct?

Semantic parsing is useful
Code and data online

http://www-nlp.stanford.edu/software/sempre/
Code and data online

http://www-nlp.stanford.edu/software/sempre/

Collaborators

Jonathan Berant (post-doc)
Andrew Chou (masters)
Roy Frostig (Ph.D.)
Panupong Pasupat (Ph.D.)
Code and data online

http://www-nlp.stanford.edu/software/sempre/

Collaborators

Jonathan Berant (post-doc)
Andrew Chou (masters)
Roy Frostig (Ph.D.)
Panupong Pasupat (Ph.D.)

Thank you!