Learning from Natural Language Interactions

Percy Liang

Smart Customers Workshop — June 2, 2017
Natural language interfaces

1. Alexa (Amazon Echo)
2. Cortana (Windows 10 Phone)
3. Siri (iPhone)
4. Google Now (Android)
Language as programs

What is the largest city in Europe by population?
Language as programs

*What is the largest city in Europe by population?*

semantic parsing

Cities
What is the largest city in Europe by population?
Language as programs

What is the largest city in Europe by population?

semantic parsing

Cities ContainedBy(Europe)
Language as programs

What is the largest city in Europe by population?

semantic parsing

Cities ∩ ContainedBy(Europe)
Language as programs

What is the largest city in Europe by population?

semantic parsing

Cities \cap ContainedBy(Europe) \quad Population

[database]
Language as programs

What is the largest city in Europe by population?

semantic parsing

\[\text{argmax}(\text{Cities} \cap \text{ContainedBy(Europe)}, \text{Population})\]
What is the largest city in Europe by population?

**Language as programs**

Semantic parsing:

\[
\text{argmax}(\text{Cities} \cap \text{ContainedBy(Europe)}, \text{Population})
\]

Execute:

Istanbul
Language as programs

Remind me to buy milk after my last meeting on Monday.
Language as programs

Remind me to buy milk after my last meeting on Monday.

semantic parsing

Add(Buy(Milk), argmax(Meetings ∩ HasDate(2016-07-18), EndTime))
Language as programs

Remind me to buy milk after my last meeting on Monday.

semantic parsing

Add(Buy(Milk), argmax(Meetings \cap HasDate(2016-07-18), EndTime))

execute

[reminder added]
Language as programs

- [sentence]
- semantic parsing
- [program]
- execute
- [behavior]
No recourse on failures

“Shut up my music”
That’s not very nice.

“Make my music be quiet”
Sorry, I can’t help you with that, Kerber.

“Turn my music off
goddamnit”
Playing all songs…
How do we develop systems that improve with use?
Language games

Wittgenstein (1953):

Language derives its meaning from use.
SHRDLURN

- sees goal
- perform actions
- has language
- no language

[with Sida Wang, Chris Manning; 2016]
SHRDLURN

sees goal
has language

remove red

perform actions
no language

[with Sida Wang, Chris Manning; 2016]
SHRDLURN

- remove red
- add(hascolor(red))
- add(hascolor(brown))
- remove(hascolor(red))
- remove(hascolor(brown))

sees goal
perform actions

has language
no language

[with Sida Wang, Chris Manning; 2016]
SHRDLURN

remove red
add(hascolor(red))
add(hascolor(brown))
remove(hascolor(red))
remove(hascolor(brown))

sees goal
perform actions

has language
no language

[with Sida Wang, Chris Manning; 2016]
Real users (June-Sept 2016)

10,000+ utterances

add brown on the top unless the rightmost
pick up blue blocks
+ 1 2 3 4 5 r
Not the brown block!, The orange block!
add blo 1 bro
rem ora blo
add blo 6 pin
add blo 134 bl
去掉最后一一块
在蓝色块上面加一层橙色块
smaz 1 a 3 jednou
retire les blocs bleus

move all blocks but middle
- 1 br - 4 br - 6 br
一番奥にオレンジを置く, 一番右の赤を消す
add red one on the first
lift 1 3 5
add one orange block on top of each orange
去掉 蓝色 方块
smaz 1 a 2 a 3 a 5
quita el bloque marrón
quita el primer bloque por la derecha
drop orange not left not right
add brown on all blue in line 2 in line 3
Add x x o x o x red block
只保留桔黄色的方块
quitar cubo rojo, quitar ultimo cubo rojo

System learns language from scratch through interaction!
Voxelurn

add two chairs 5 spaces apart

[with Sida Wang, Sam Ginn, Chris Manning; 2017]
add two chairs 5 spaces apart
Voxelurn

**add two chairs 5 spaces apart**

Learn from definitions using grammar induction:

**add chair, move 5 left, add chair**

```plaintext
(:s (:blkr (:s (:loop (number 3) (:s (: add brown here) (:for (call adj top this) (: select)))) (:loop (number 3) (:for (call adj bot this) (: select))))) (:loop (number 3) (:for (call adj left this) (: select))))) (:s (:s (:s (:blkr (:s (:loop (number 3) (:s (: add brown here) (:for (call adj top this) (: select)))) (:loop (number 3) (:for (call adj bot this) (: select))))) (:loop (number 3) (:for (call adj left this) (: select))))) (:s (:s (:s (:s (:blkr (:s (:loop (number 3) (:s (: add brown here) (:for (call adj top this) (: select)))) (:loop (number 3) (:for (call adj bot this) (: select))))) (:loop (number 3) (:for (call adj left this) (: select)))) (:blkr (:s (:loop (number 3) (:s (: add brown here) (:for (call adj top this) (: select)))) (:loop (number 3) (:for (call adj bot this) (: select)))))) (:loop (number 3) (:for (call adj right this) (: select)))))) (:blkr (:s (:loop (number 3) (:s (: add brown here) (:for (call adj top this) (: select)))) (:loop (number 3) (:for (call adj bot this) (: select)))))))
```

[with Sida Wang, Sam Ginn, Chris Manning; 2017]
Voxelurn

Learn from definitions using grammar induction:

add chair, move 5 left, add chair

(add two chairs 5 spaces apart)
Voxelurn

Learn from definitions using grammar induction:

*add 4 legs, add chair base, add chair back, ...*

(with Sida Wang, Sam Ginn, Chris Manning; 2017)
Voxelurn

3 days, 42 users, 230 structures, 64075 utterances, 2495 definitions

*add chair, move 5 left, add chair*

*add 4 legs, add chair base, add chair back, ...*
Collaborative dialogue

<table>
<thead>
<tr>
<th>Name</th>
<th>Company</th>
<th>Time</th>
<th>Location</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kathy</td>
<td>TRT Holdings</td>
<td>afternoon</td>
<td>indoor</td>
</tr>
<tr>
<td>Jason</td>
<td>Dollar General</td>
<td>afternoon</td>
<td>indoor</td>
</tr>
<tr>
<td>Johnny</td>
<td>TRT Holdings</td>
<td>afternoon</td>
<td>outdoor</td>
</tr>
<tr>
<td>Frank</td>
<td>SFN Group</td>
<td>afternoon</td>
<td>indoor</td>
</tr>
<tr>
<td>Catherine</td>
<td>Dollar General</td>
<td>afternoon</td>
<td>indoor</td>
</tr>
<tr>
<td>Catherine</td>
<td>Weis Markets</td>
<td>afternoon</td>
<td>indoor</td>
</tr>
<tr>
<td>Kathleen</td>
<td>TRT Holdings</td>
<td>morning</td>
<td>indoor</td>
</tr>
<tr>
<td>Lori</td>
<td>TRT Holdings</td>
<td>afternoon</td>
<td>indoor</td>
</tr>
<tr>
<td>Frank</td>
<td>L&amp;L Hawaiian Barbecue</td>
<td>afternoon</td>
<td>outdoor</td>
</tr>
</tbody>
</table>

---

**Names and Companies:**

- Justin: New Era Tickets
- Gloria: L&L Hawaiian Barbecue
- Kathleen: Advance Auto Paris
- Justin: Dollar General
- Anna: Arctic Cat
- Steven: Dollar General
- Wayne: R.J. Corman Railroad
- Alexander: R.J. Corman Railroad

**Dialog:**

- **User:** Hi
- **System:** Hello
- **User:** 1 of my morning likes the outdoors
- **System:** All of my friends prefer morning
- **User:** And all like indoor except one
- **System:** Kathleen?
- **User:** SELECT (Kathleen, TRT Holdings, morning, indoor)
- **System:** SELECT (Kathleen, TRT Holdings, morning, indoor)
Collaborative dialogue

combine structured knowledge + unstructured text

[with He He, Anusha Balakrishnan, Mihail Eric; 2017]
Web agents in the World of Bits

**Query:** Can you book a flight from San Francisco to New York?

**Query:** What are the top rated place to eat Korean food in SF?

**Query:** What is the monthly payment for $2000 with a term of 2 years at today's rates?

**Query:** How much is an iPhone 7 Plus in Indiana?

reinforcement learning + open-domain realism
Web agents in the World of Bits

1. Input
   - Image
   - DOM
   - Query

2. State Features
   - CNN
   - Feature

3. Action Representation
   - LocalCNN
   - Sample
   - Sample
   - GlobalCNN
   - Sample

4. Action
   - Mouse
   - Mouse buttons
   - Keyboard
Summary

• Learn immediately from user interaction

• Collaborative dialogue / problem solving

• Carry out open-ended actions on the web