Interacting with Intelligence

Phil Cohen

Chief Scientist, AI and SVP, Advanced Technology
Are we stuck in a conversational local minimum?

Why?  How to get out

Voice AI

- Semantic Parsing
- Dialogue/plan-reasoning

Semantic Parser

Plan-based Reasoning Engine

Collaborative Dialogue

Meaning Representation

Knowledge Graphs encoding facts, actions, ...
Current State of the Practice -- NLU

Statistical “intent” classification ~ Propositional Content

- Easy to train simple classifiers; nice toolkits
- Limited expressiveness to date
- Challenges
  - Multi-“intent”, multi-domain, multi-utterance turns
  - Yes/no questions
  - Anaphora
  - Complex constructions
  - Compositionality
  - Speech act types expressing intentions other than commands/actions
- “Light is better here”
Semantic Parser

An old idea is new again

- Process that maps a sentence to a representation of its semantics, i.e., its meaning
- Meaning is represented as a **logical form**, a logical language including:
  - **Entities**, such as objects in the domain, events, variables, tuples,
  - **Relations** such as predicate/argument structures, types, and
  - **Operators**, including conjunction, quantification, superlatives, comparatives, aggregation, sequence, conditional, variables etc.

- Logical forms are **compositional** – meaning of whole is a function of meaning of parts
- Logical forms can be
  - Vague, with pronouns and referential expressions resolved later via context
  - Fused with LFs from other modalities
  - Input to learning and inference
  - Input to dialogue management subsystems
  - Mapped to backend data sources,
  - Executed to retrieve data or invoke APIs,
Query

How long are the rivers that flow through each state that borders California?

Answer

<table>
<thead>
<tr>
<th>Length</th>
<th>River</th>
<th>State</th>
</tr>
</thead>
<tbody>
<tr>
<td>805</td>
<td>gila</td>
<td>arizona</td>
</tr>
<tr>
<td>1670</td>
<td>snake</td>
<td>oregon</td>
</tr>
<tr>
<td>1953</td>
<td>columbia</td>
<td>oregon</td>
</tr>
<tr>
<td>2333</td>
<td>colorado</td>
<td>arizona</td>
</tr>
<tr>
<td>2333</td>
<td>colorado</td>
<td>nevada</td>
</tr>
</tbody>
</table>

Logical Form:

answer([A,B,C],
  (len(B,A),river(B), traverse(B,C),state(C),
   next_to(C, california), state(california)))
<table>
<thead>
<tr>
<th>Linguistic construction</th>
<th>Geo</th>
</tr>
</thead>
<tbody>
<tr>
<td>Relative clauses</td>
<td><em>that</em> border California</td>
</tr>
<tr>
<td>Comparatives</td>
<td><em>taller</em> than Mount Shasta</td>
</tr>
<tr>
<td>Superlatives</td>
<td><em>tallest</em> mountain</td>
</tr>
<tr>
<td>Negation</td>
<td>does not border California</td>
</tr>
<tr>
<td>Anaphora (pronouns)</td>
<td><em>their</em> population which <em>one</em> ...</td>
</tr>
<tr>
<td>Ordinals</td>
<td>What is the capital of the second one.</td>
</tr>
<tr>
<td>Ordinals + superlatives</td>
<td>What is the capital of the second largest state?</td>
</tr>
<tr>
<td>Ellipsis</td>
<td><em>Which is taller than shasta?</em></td>
</tr>
<tr>
<td>Quantifiers</td>
<td><em>the capital of each</em> state that borders California</td>
</tr>
<tr>
<td>Yes/No Questions</td>
<td>Does Pennsylvania have a larger population than Ohio?</td>
</tr>
<tr>
<td>Conjunction</td>
<td><em>what is the capital and population of the states that border both Ohio and New York</em></td>
</tr>
</tbody>
</table>
More complex queries

Query:

How tall is the highest mountain in the state with the smallest population

<table>
<thead>
<tr>
<th>Elevation</th>
<th>Mountain</th>
<th>State</th>
<th>Population</th>
</tr>
</thead>
<tbody>
<tr>
<td>6194</td>
<td>mckinley</td>
<td>alaska</td>
<td>401800</td>
</tr>
</tbody>
</table>

answer([A,B,C,D],
  (highest(B,
    (smallest(D,(state(C),
      population(C,D)))),
    mountain(B),loc(B,C)),
    elevation(B,A))))
More question-answering

More Examples

» What is the capital of each state that borders Ohio?
  • answer([A,B],[capital(A),state(B),loc(A,B),
next_to(B,stateid(ohio)), state(stateid(ohio))])

Yes/No Q’s, comparatives

» Is the tallest mountain in california higher than the tallest
  mountain in Alaska?
  answer(A,[highest(B,[mountain(B),loc(B,stateid(california))],
state(stateid(california))]),
highest(C,[mountain(C),loc(C,stateid(alaska))],
state(stateid(alaska))]),
higher(B,C)->A=yes;A=no)
Other use cases

Complex Commands

*Set the temperature to 70 degrees at 8am every other tuesday*
*Find the latest version of this file and send it to phil and john*
*Close all the shades except in the kitchen*

Standing Orders

*When Bob replies to this message, send his reply to my team*
*Turn off the sprinklers for twenty four hours whenever we get one inch of rain*

Prohibitions

*Don’t schedule meetings at lunch time*

Assertions:

*I don’t like text messages when I’m in a meeting*
# Hybrid Symbolic/Statistical Semantic Parser

<table>
<thead>
<tr>
<th>Priority</th>
<th>Development/Training</th>
<th>Test</th>
<th>Arbitrated</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Symbolic (S)</td>
<td>84.2%</td>
<td>94.3%</td>
</tr>
<tr>
<td></td>
<td>Statistical (St)</td>
<td>91.0%</td>
<td></td>
</tr>
<tr>
<td>Top Priority</td>
<td>N= 3853</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>S: 61.6%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>St: 81.9%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other Priority</td>
<td>Statistical</td>
<td></td>
<td>85.5%</td>
</tr>
<tr>
<td></td>
<td>N= 3963</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>S: 61.6%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>St: 81.9%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>N= 883</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Challenges

- Where do the predicates come from?
- Relationship to Ontologies/KBs?
- What are the LF constructions?
- How to map various linguistic constructions to LF constructions?
- How to map those predicates to backend APIs/DBs?
- Tooling
- Deep Learning
- Data collection
Dialogue

- Mostly undeveloped in current virtual assistants

- Finite-state, system driven, speech act sequences, authored, rigid

- Slot-filling (GUS (Bobrow et al., 1977), DARPA’s Air Travel Information System (ATIS), 1990+)

- Task-oriented (Grosz, 1977)


- Most current research and practical dialogue systems are Slot Fillers.
  » Why?
  - Easy to implement, uses statistical “intent” classification, learn policies for responding
  - “Light is better here”,

- Slot-filling is not easily extended to other types of dialogues; should be a special case

- What do we want, and how do we get there?
Collaborative dialogue driven by plan recognition

Please note:
• Contextual reference to song
• Referring to group as “their” and “they”
• System infers purpose of question is for user to see next concert in region
• System offers to buy tickets
• System informs user that plan will fail (venue is sold out)
• System finds another method to see concert
• System offers to buy tickets for user
• System proactively provides seating chart to get information needed to purchase ticket
Common Patterns of Plan Inference

Intent / Plan recognition

- Request(\text{Action}) \rightarrow \text{Action}_1 \rightarrow \text{Effect}_2 \rightarrow \text{Action}_2 \ldots \rightarrow \text{Effect}_n \text{ (top goal)}

- Req(\text{Time of Next concert}) \rightarrow (\text{Date/Loc}) \rightarrow \text{Go to Loc} \rightarrow \text{Attend Concert} \rightarrow \text{Hear musical group}

- Check preconditions

- On failure, find another plan to achieve goal
  - Find substitute objects for goals, edit plan, confirm
  - Other operations possible -- system plans to overcome obstacles; system plans to achieve goals

- Common precondition failures --
  - business must be open to conduct commerce in person \rightarrow \text{find one that is open}
  - Item is in stock \rightarrow \text{find another business selling item, (e.g., concert venue)}
  - Cannot arrive on time for an appointment \rightarrow \text{make another appointment; change vendor}
Use of Knowledge Base

- **Coldplay** is a musical group
  - “Adventure of a lifetime” is a song by Coldplay
- Musical groups play multiple live concerts
- Each live concert takes place at a venue
- Venues sell tickets for concert events
- **Precondition**: To attend an event, one must have a ticket
- There is a limited number of tickets
- Concerts are Entertainment Events, which are Events
- Events take place at times and locations
- People like to view events
- **Precondition**: To view an event, the viewer must be at the same place at same time
- Viewer must travel to the venue

... 

Need not be first principles reasoning.