Chatbots

- Can model as machine translation, but need to endow with diversity, add consistency among answers, ...

Li et al. (2016)

Task-oriented dialogue

- Involves both generation and language understanding
- Dialogue state: reflects any information about the conversation (e.g., search history)
- User utterance -> update dialogue state -> take action (e.g., query the restaurant database) -> say something

Young et al. (2013)
Find me a good sushi restaurant in Chelsea

restaurant_type <- sushi
location <- Chelsea
curr_result <- execute_search()
Sushi Seki Chelsea is a sushi restaurant in Chelsea with 4.4 stars on Google

How expensive is it?
get_value(cost, curr_result)
Entrees are around $30 each

This Lecture

- How do we represent information for information extraction?
- Relation extraction
- Slot filling
- Open Information Extraction

Semantic Representations

- “World” is a set of entities and predicates

<table>
<thead>
<tr>
<th>person</th>
<th>president</th>
<th>stab</th>
</tr>
</thead>
<tbody>
<tr>
<td>Brutus</td>
<td>Obama</td>
<td>Brutus, Caesar</td>
</tr>
<tr>
<td>Caesar</td>
<td>Bush</td>
<td>...</td>
</tr>
<tr>
<td>Obama</td>
<td>Bush</td>
<td>...</td>
</tr>
<tr>
<td>Bush</td>
<td>...</td>
<td>...</td>
</tr>
</tbody>
</table>

- Statements are logical expressions that evaluate to true or false

Brutus stabs Caesar => true
Caesar was stabbed => true

Example credit: Asad Sayeed
Semantic Representations

- Brutus stabs Caesar
  \( \text{stab} \text{Brutus, Caesar} \)
- Brutus stabbed Caesar with a knife
  \( \text{stab} \text{Brutus, Caesar, instrument=knife} \)
- Brutus stabbed Caesar with a knife in the agora
  \( \text{stab} \text{Brutus, Caesar, instrument=knife, location=agora} \)
- Brutus stabbed Caesar with a knife in the agora on the Ides of March
  
  Example credit: Asad Sayeed
(At least) Two Solutions

- Entity-relation-entity triples: focus on entities and their relations (note that prominent events can still be entities)
  (Barack Obama, presidentOf, United States)

- Slot filling: specific ontology, populate information in a predefined way

Entity-Relation-Entity Pairs

- Entity-relation-entity triples: focus on entities and their relations (note that prominent events can still be entities)
  (Barack Obama, presidentOf, United States)

- Slot filling: specific ontology, populate information in a predefined way

Entity-Relation-Entity Pairs

- Represent semantics as relationships between entities; relationships are drawn from a fixed ontology

<table>
<thead>
<tr>
<th>Entity-Relation-Entity Pairs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Barack Obama, presidentOf, United States</td>
</tr>
<tr>
<td>(Barack Obama, signed, the Affordable Care act)</td>
</tr>
<tr>
<td>(Several prominent Republicans, denounce, the new law)</td>
</tr>
</tbody>
</table>

Table 5: Sample facts of YAGO

<table>
<thead>
<tr>
<th>Entity</th>
<th>Type</th>
<th>Subclass</th>
</tr>
</thead>
<tbody>
<tr>
<td>Zidane</td>
<td>TYPE+SUBCLASS</td>
<td>football player</td>
</tr>
<tr>
<td>Zidane</td>
<td>TYPE</td>
<td>Person from Marseille</td>
</tr>
<tr>
<td>Zidane</td>
<td>TYPE</td>
<td>Legion d’honneur recipient</td>
</tr>
<tr>
<td>Zidane</td>
<td>BORNINYEAR</td>
<td>1972</td>
</tr>
<tr>
<td>“Paris”</td>
<td>FAMILYNAMEOF</td>
<td>Priscilla Paris</td>
</tr>
<tr>
<td>“Paris”</td>
<td>GIVENNAMEOF</td>
<td>Paris Hilton</td>
</tr>
<tr>
<td>“Paris”</td>
<td>MEANS</td>
<td>Paris, France</td>
</tr>
<tr>
<td>“Paris”</td>
<td>MEANS</td>
<td>Paris, Texas</td>
</tr>
<tr>
<td>Paris, France</td>
<td>LOCATEDIN</td>
<td>France</td>
</tr>
<tr>
<td>Paris, France</td>
<td>TYPE+SUBCLASS</td>
<td>capital</td>
</tr>
<tr>
<td>Paris, France</td>
<td>TYPE</td>
<td>Eurovision host city</td>
</tr>
<tr>
<td>Paris, France</td>
<td>ESTABLISHEDIN</td>
<td>-300</td>
</tr>
</tbody>
</table>

Suchanek et al. (2007)

Entity-Relation-Entity Pairs

- Can easy query about relations in the knowledge base
  when was Barack Obama born? \( \lambda x. \text{born}(\text{Barack}_\text{Obama}, x) \)
  how many children does Barack Obama have?
  \( \text{sizeof}(\lambda x. \text{isParent}(x, \text{Barack}_\text{Obama})) \)
  how old was Barack Obama when he became president?
  — no timeOfBecomingPresident relation
  how many Wimbledon victories has Serena Williams had?
  — Wimbledon are listed, but no isWimbledon predicate

Open IE

- Entity-relation-entity triples aren’t necessarily grounded in an ontology

- Extract strings and let a downstream system figure it out

Barack Obama signed the Affordable Care act on Tuesday. He gave a speech later that afternoon on how the act would help the American people. Several prominent Republicans were quick to denounce the new law.

(Barack Obama, signed, the Affordable Care act)
(Several prominent Republicans, denounce, the new law)
Slot Filling

- Represent information about a particular event like an earthquake

- **magnitude** | **time**

> Indian Express — A massive earthquake of *magnitude 7.3* struck Iraq on Sunday, 103 kms (64 miles) southeast of the city of As-Sulaymaniyah, the US Geological Survey said, reports Reuters. US Geological Survey initially said the quake was of a magnitude 7.2, before revising it to 7.3. **epicenter**

IE: The Big Picture

- How do we represent information? What do we extract?
  - Entity-relation-entity triples (fixed ontology or open)
  - Slot fillers

- Where does that information come from? (closed vs. open IE)
  - Closed: limited set of documents, domain-specific
  - Open: try to use lots of information (the whole Internet)

Relation Extraction

- Extract entity-relation-entity triples from a fixed inventory

- **Located_In**

> During the war in *Iraq*, *American journalists* were sometimes caught in the line of fire

- Pipelined classifiers looking at surface level, syntactic features (dependency paths), semantic roles

- Problem: limited data for scaling to big ontologies

ACE (2003-2005)
Hearst Patterns

- Syntactic patterns especially for finding hypernym-hyponym pairs ("is a" relations)
  - Y is a X  
    Berlin is a city
  - X such as [list]  
    cities such as Berlin, Paris, and London.
  - other X including Y  
    other cities including Berlin
- Totally unsupervised way of harvesting world knowledge for tasks like parsing and coreference (Bansal and Klein, 2011-2012)

Hearst (1992)

Distant Supervision

- Lots of relations in our knowledge base already (e.g., 23,000 film-director relations); use these to bootstrap more training data
  - If two entities in a relation appear in the same sentence, assume the sentence expresses the relation
    [Steven Spielberg]'s film [Saving Private Ryan] is loosely based on the brothers' story
    Allison co-produced the Academy Award-winning [Saving Private Ryan], directed by [Steven Spielberg]

Mintz et al. (2009)

Distant Supervision

- Learn decently accurate classifiers for ~100 Freebase relations

<table>
<thead>
<tr>
<th>Relation name</th>
<th>100 instances</th>
<th>1000 instances</th>
</tr>
</thead>
<tbody>
<tr>
<td>/film/director/film</td>
<td>0.49</td>
<td>0.49</td>
</tr>
<tr>
<td>/film/writer/film</td>
<td>0.70</td>
<td>0.70</td>
</tr>
<tr>
<td>/geography/river/basin.countries</td>
<td>0.65</td>
<td>0.65</td>
</tr>
<tr>
<td>/location/country/administrative_divisions</td>
<td>0.68</td>
<td>0.68</td>
</tr>
<tr>
<td>/location/location/contains</td>
<td>0.81</td>
<td>0.81</td>
</tr>
<tr>
<td>/location/us.county/counties</td>
<td>0.51</td>
<td>0.51</td>
</tr>
<tr>
<td>/music/artist/origins</td>
<td>0.64</td>
<td>0.64</td>
</tr>
<tr>
<td>/people/deceased.person/place.of.death</td>
<td>0.80</td>
<td>0.80</td>
</tr>
<tr>
<td>/people/person/nationality</td>
<td>0.61</td>
<td>0.61</td>
</tr>
<tr>
<td>/people/person/place.of.birth</td>
<td>0.78</td>
<td>0.78</td>
</tr>
<tr>
<td><strong>Average</strong></td>
<td><strong>0.67</strong></td>
<td><strong>0.67</strong></td>
</tr>
</tbody>
</table>

Mintz et al. (2009)

Slot Filling
Slot Filling

- Extract a fixed set of roles from a relatively ordered text like a seminar announcement

  Speaker: [Alan Clark]_Speaker
  [“Gender Roles in the Holy Roman Empire”]_Title
  [Allagher Center Main Auditorium]_Location

  This talk will discuss...

- Old work: HMMs, later CRFs trained per role

Freitag and McCallum (2000)

Slot Filling: MUC

- Extract a fixed set of roles from a relatively ordered text like a seminar announcement

  Speaker: [Alan Clark]_Speaker
  [“Gender Roles in the Holy Roman Empire”]_Title
  [Allagher Center Main Auditorium]_Location

  This talk will discuss...

- Old work: HMMs, later CRFs trained per role

Haghighi and Klein (2010)

Slot Filling: Forums

- Extract product occurrences in cybercrime forums, but not everything that looks like a product is a product

  TITLE: [ buy ] Backconnect bot
  BODY: Looking for a solid backconnect bot .
  If you know of anyone who codes them please let me know

  (a) File 0-initiator4856

  TITLE: Exploit cleaning ?
  BODY: Have some Exploits i need fud .

  (b) File 0-initiator10815

Portnoff et al. (2017), Durrett et al. (2017)

Open IE + IR

- Can retrieve additional information about specific events

- If we’re uncertain about extractions, fetch another article to improve confidence

current belief

latest extraction

Narasimhan et al. (2016)
Can retrieve additional information about specific events
If we’re uncertain about extractions, fetch another article to improve confidence

Use reinforcement learning to send queries about specific things

Typically no fixed relation inventory

“Open”ness — want to be able to extract all kinds of information from open-domain text

“Machine reading the web” — acquire commonsense knowledge just from reading about it, but need to process lots of text
TextRunner

- Supervised system
- Extract positive examples of (e, r, e) triples via parsing and heuristics
- Train a Naive Bayes classifier to filter pairs from raw text: uses features on POS tags, lexical features, stopwords, etc.

Barack Obama, 44th president of the United States, was born in Honolulu

=> Barack Obama, was born in Honolulu

- 80x faster than running a parser
- Use multiple instances of extractions to assign probability to a relation

Banko et al. (2007)

Exploiting Redundancy

- 9M web pages / 133M sentences
- 2.2 tuples extracted per sentence, filter based on probabilities

- Concrete: definitely true
- Abstract: possibly true but underspecified
- Hard to evaluate: can assess precision of extracted facts, but how do we know recall?

Banko et al. (2007)

ReVerb

- More constraints: open relations have to begin with verb, end with preposition, be contiguous (e.g., was born on)
- Extract more meaningful relations, particularly with light verbs

ReVerb

- For each verb, identify the longest sequence of words following the verb that satisfy a POS regex (V.*P) and which satisfy heuristic lexical constraints on specificity
- Find the nearest arguments on either side of the relation
- Annotators labeled relations in 500 documents to assess recall

Fader et al. (2011)
**NELL**

- Entity typing/resolution + relation classification to read facts about things, combine with logical inference as well
- Coupling constraints: types of arguments to relations must match the relation extracted

```
zooInCity(CincinnaZoo, CincinnaCity)
The CincinnaZoo is located north of downtown Cincinna City
```

**QA from Open IE**

**Takeaways**

- Relation extraction: well-defined task for specific relations, can collect data with distant supervision
- Slot filling: tied to a specific ontology, can be complex and needs annotated data
- Open IE: extracts lots of things, but hard to know how good or useful they are
  - Can combine with standard question answering
  - Add new facts to knowledge bases