Information Extraction

Wei Xu

(many slides from Greg Durrett, Luheng He, Emma Strubell)
This Lecture

- How do we represent information for information extraction?
- Semantic role labeling / abstract meaning representation
- Relation extraction
- Slot filling
- Open Information Extraction
IE: The Big Picture

- How do we represent information? What do we extract?
  - Semantic roles
  - Abstract meaning representation
  - Slot fillers
  - Entity-relation-entity triples (fixed ontology or open)
Semantic Role Labeling
Semantic Role Labeling

- Find out 5W in text — “who did what to whom, when and where”
- Identify predicate, disambiguate it, identify that predicate’s arguments

Figure from He et al. (2017)
In 1950 Alan M. Turing **published** "Computing machinery and intelligence" in Mind, in which he **proposed** that machines could be **tested** for intelligence using questions and answers.

<table>
<thead>
<tr>
<th>Predicate</th>
<th>Question</th>
<th>Answer</th>
</tr>
</thead>
<tbody>
<tr>
<td>published</td>
<td>1 Who published something?</td>
<td>Alan M. Turing</td>
</tr>
<tr>
<td></td>
<td>2 What was published?</td>
<td>&quot;Computing Machinery and Intelligence&quot;</td>
</tr>
<tr>
<td></td>
<td>3 When was something published?</td>
<td>In 1950</td>
</tr>
<tr>
<td>proposed</td>
<td>4 Who proposed something?</td>
<td>Alan M. Turing</td>
</tr>
<tr>
<td></td>
<td>5 What did someone propose?</td>
<td>that machines could be tested for intelligent using questions and answers</td>
</tr>
<tr>
<td></td>
<td>6 When did someone propose something?</td>
<td>In 1950</td>
</tr>
<tr>
<td>tested</td>
<td>7 What can be tested?</td>
<td>machines</td>
</tr>
<tr>
<td></td>
<td>8 What can something be tested for?</td>
<td>intelligence</td>
</tr>
<tr>
<td></td>
<td>9 How can something be tested?</td>
<td>using questions and answers</td>
</tr>
<tr>
<td>using</td>
<td>10 What was being used?</td>
<td>questions and answers</td>
</tr>
<tr>
<td></td>
<td>11 Why was something being used?</td>
<td>tested for intelligence</td>
</tr>
</tbody>
</table>

Figure from FitzGerald et al. (2018)
The robot *broke* my favorite mug with a wrench.

My mug *broke* into pieces immediately.
The robot *broke* my favorite mug with a wrench.

My mug *broke* into pieces immediately.

Figure from He et al. (2017)
The robot *broke* my favorite mug with a wrench.

My mug *broke* into pieces immediately.

Figure from He et al. (2017)
The robot *broke* my favorite mug with a wrench.

Frame: `break.01`

<table>
<thead>
<tr>
<th>role</th>
<th>description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ARG0</td>
<td>breaker</td>
</tr>
<tr>
<td>ARG1</td>
<td>thing broken</td>
</tr>
<tr>
<td>ARG2</td>
<td>instrument</td>
</tr>
<tr>
<td>ARG3</td>
<td>pieces (final state)</td>
</tr>
<tr>
<td>ARG4</td>
<td>broken away from what?</td>
</tr>
</tbody>
</table>

My mug *broke* into pieces immediately.

Frame: `break.01`

<table>
<thead>
<tr>
<th>role</th>
<th>description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ARG0</td>
<td>thing broken</td>
</tr>
<tr>
<td>ARG1</td>
<td>pieces</td>
</tr>
<tr>
<td>ARG2</td>
<td>instrument</td>
</tr>
<tr>
<td>ARG3</td>
<td>broken away from what?</td>
</tr>
</tbody>
</table>
The Proposition Bank (PropBank)

Core roles:
Verb-specific roles (ARG0-ARG5) defined in frame files

<table>
<thead>
<tr>
<th>Frame: break.01</th>
</tr>
</thead>
<tbody>
<tr>
<td>role</td>
</tr>
<tr>
<td>ARG0</td>
</tr>
<tr>
<td>ARG1</td>
</tr>
<tr>
<td>ARG2</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Frame: buy.01</th>
</tr>
</thead>
<tbody>
<tr>
<td>role</td>
</tr>
<tr>
<td>ARG0</td>
</tr>
<tr>
<td>ARG1</td>
</tr>
<tr>
<td>ARG2</td>
</tr>
<tr>
<td>ARG3</td>
</tr>
<tr>
<td>ARG4</td>
</tr>
</tbody>
</table>

Adjunct roles:
(ARGM-) shared across verbs

<table>
<thead>
<tr>
<th>role</th>
<th>description</th>
</tr>
</thead>
<tbody>
<tr>
<td>TMP</td>
<td>temporal</td>
</tr>
<tr>
<td>LOC</td>
<td>location</td>
</tr>
<tr>
<td>MNR</td>
<td>manner</td>
</tr>
<tr>
<td>DIR</td>
<td>direction</td>
</tr>
<tr>
<td>CAU</td>
<td>cause</td>
</tr>
<tr>
<td>PRP</td>
<td>purpose</td>
</tr>
</tbody>
</table>

Annotated on top of the Penn Treebank Syntax

PropBank Annotation Guidelines, Bonial et al., 2010

Figure from He et al. (2017)
Semantic Role Labeling

- Identify predicates (*love*) using a classifier (not shown)
- Identify ARG0, ARG1, etc. as a tagging task with a BiLSTM conditioned on *love*
- Other systems incorporate syntax, joint predicate-argument finding

Figure from He et al. (2017)
Residual / Highway Connections

Non-linearity:
\[ F(c_{t-1}, h_{t-1}, x_t) \]

Output to the next layer:
\[ h_t \]

Shortcut:
\[ x_t \]

Input from the previous layer:
\[ h_{t-1} \]

Recurrent input:
\[ h_{t-1} \]

New output:
\[ h_t + x_t \]

Gated highway network:
\[ r_t \circ h_t + (1 - r_t) \circ x_t \]
\[ r_t = \sigma(f(h_{t-1}, x_t)) \]

References:
- Deep Residual Networks, Kaiming He, ICML 2016 Tutorial
- Training Very Deep Networks, Srivastava et al., 2015

Figure from He et al. (2017)
SRL Systems

(syntax-based)

Pipeline Systems
- sentence, predicate
- syntactic features
- argument id.
- candidate argument spans
- labeling
- labeled arguments
- ILP/DP
- prediction

Punyakanok et al., 2008
Täckström et al., 2015
FitzGerald et al., 2015

End-to-end Systems
- sentence, predicate
- context window features
- Deep BiLSTM + CRF layer
- BIO sequence
- Viterbi
- prediction

Collobert et al., 2011
Zhou and Xu, 2015
Wang et. al, 2015

The System on previous slide
- Deep BiLSTM
- BIO sequence
- prediction

He et al., 2017

Figure from He et al. (2017)
10 Years of PropBank SRL

Figure from Strubell et al. (2018)
Can we combine the two approaches — incorporate syntactic information into neural networks?

Multi-task learning with related tasks, e.g., part-of-speech tagging, dependency parsing ...

Syntactically-informed self-attention: use the Transformer to encore the sentence; in one head, token attends to its likely syntactic parents; in next layer, tokens observe all other parents.
Recall: Transformer (multi-head self-attention)  
Figure from Strubell et al. (2018)
Syntactically-Informed Self-Attention

[Dozat and Manning 2017]
Linguistically-Informed Self-Attention

Figure from Strubell et al. (2018)
Linguistically-Informed Self-Attention

Figure from Strubell et al. (2018)
Linguistically-Informed Self-Attention

Figure from Strubell et al. (2018)
Linguistically-Informed Self-Attention
Figure from Strubell et al. (2018)
Linguistically-Informed Self-Attention

Figure from Strubell et al. (2018)
Linguistically-Informed Self-Attention

Figure from Strubell et al. (2018)
Linguistically-Informed Self-Attention

Layer J
Multi-head self-attention + feed forward

Layer p
Syntactically-informed self-attention

Layer r
Multi-head self-attention + feed forward

Bilinear

predicates

args

B-ARG₀

Nobel committee awards Strickland who advanced optics

NLP

VBZ/PRED

NN

VBN/PRED

NNP

NN

WP

NN

NNP

WP

VBN/PRED

VBZ/PRED

NN

VBN/PRED

NN

VBZ/PRED

NN

VBZ/PRED

NN

VBZ/PRED

NN

VBZ/PRED

NN

VBZ/PRED

NN

VBZ/PRED

NN

VBZ/PRED

NN

VBZ/PRED

NN

VBZ/PRED

NN

VBZ/PRED

NN

VBZ/PRED

NN

VBZ/PRED

NN

VBZ/PRED

NN

VBZ/PRED

NN

VBZ/PRED

NN

VBZ/PRED

NN

VBZ/PRED

NN

VBZ/PRED

NN

VBZ/PRED

NN

VBZ/PRED

NN

VBZ/PRED

NN

VBZ/PRED

NN

VBZ/PRED

NN

VBZ/PRED

NN

VBZ/PRED

NN

VBZ/PRED

NN

VBZ/PRED

NN

VBZ/PRED

NN

VBZ/PRED

NN

VBZ/PRED

NN

VBZ/PRED

NN

VBZ/PRED

NN

VBZ/PRED

NN

VBZ/PRED

NN

VBZ/PRED

NN

VBZ/PRED

NN

VBZ/PRED

NN

VBZ/PRED

NN

VBZ/PRED

NN

VBZ/PRED

NN

VBZ/PRED

NN

VBZ/PRED

NN

VBZ/PRED

NN

VBZ/PRED

NN

VBZ/PRED

NN

VBZ/PRED

NN

VBZ/PRED

NN

VBZ/PRED

NN

VBZ/PRED

NN

VBZ/PRED

NN

VBZ/PRED

NN

VBZ/PRED

NN

VBZ/PRED

NN

VBZ/PRED

NN

VBZ/PRED

NN

VBZ/PRED

NN

VBZ/PRED

NN

VBZ/PRED

NN

VBZ/PRED

NN

VBZ/PRED

NN

VBZ/PRED

NN

VBZ/PRED

NN

VBZ/PRED

NN

VBZ/PRED

NN

VBZ/PRED

NN

VBZ/PRED

NN

VBZ/PRED

NN

VBZ/PRED

NN

VBZ/PRED

NN

VBZ/PRED

NN

VBZ/PRED

NN

VBZ/PRED

NN

VBZ/PRED

NN

VBZ/PRED

NN

VBZ/PRED

NN

VBZ/PRED

NN

VBZ/PRED

NN

VBZ/PRED

NN

VBZ/PRED

NN

VBZ/PRED

NN

VBZ/PRED

NN

VBZ/PRED

NN

VBZ/PRED

NN

VBZ/PRED

NN

VBZ/PRED

NN

VBZ/PRED

NN

VBZ/PRED

NN

VBZ/PRED

NN

VBZ/PRED

NN

VBZ/PRED

NN

VBZ/PRED

NN

VBZ/PRED

NN

VBZ/PRED

NN

VBZ/PRED

NN

VBZ/PRED

NN

VBZ/PRED

NN

VBZ/PRED

NN

VBZ/PRED

NN

VBZ/PRED

NN

VBZ/PRED

NN

VBZ/PRED

NN

VBZ/PRED

NN

VBZ/PRED

NN

VBZ/PRED

NN

VBZ/PRED

NN

VBZ/PRED

NN

VBZ/PRED

NN

VBZ/PRED

NN

VBZ/PRED

NN

VBZ/PRED

NN

VBZ/PRED

NN

VBZ/PRED

NN

VBZ/PRED

NN

VBZ/PRED

NN

VBZ/PRED

NN

VBZ/PRED

NN

VBZ/PRED

NN

VBZ/PRED

NN

VBZ/PRED

NN

VBZ/PRED

NN

VBZ/PRED

NN

VBZ/PRED

NN

VBZ/PRED

NN

VBZ/PRED

NN

VBZ/PRED

NN

VBZ/PRED

NN

VBZ/PRED

NN

VBZ/PRED

NN

VBZ/PRED

NN

VBZ/PRED

NN

VBZ/PRED

NN

VBZ/PRED

NN

VBZ/PRED

NN

VBZ/PRED

NN

VBZ/PRED

NN

VBZ/PRED

NN

VBZ/PRED

NN

VBZ/PRED

NN

VBZ/PRED

NN

VBZ/PRED

NN

VBZ/PRED

NN

VBZ/PRED

NN

VBZ/PRED

NN

VBZ/PRED

NN

VBZ/PRED

NN

VBZ/PRED

NN

VBZ/PRED

NN

VBZ/PRED

NN

VBZ/PRED

NN

VBZ/PRED

NN

VBZ/PRED

NN

VBZ/PRED

NN

VBZ/PRED

NN

VBZ/PRED

NN

VBZ/PRED
Linguistically-Informed Self-Attention

Layer $r$
- Bilinear
- Predicates
- Args

Layer $p$
- Multi-head self-attention + feed forward
- Syntactically-informed self-attention

Layer $J$
- Multi-head self-attention + feed forward

Text:
- Nobel committee awards Strickland who advanced optics
Linguistically-Informed Self-Attention

Layer J
Multi-head self-attention + feed forward

Layer p
Syntactically-informed self-attention

Layer r
Multi-head self-attention + feed forward

Nobel award committee for Strickland who advanced optics.
Linguistically-Informed Self-Attention

Bilinear
predicates
args

Layer J

Multi-head self-attention + feed forward

Layer p

Syntactically-informed self-attention

Layer r

Multi-head self-attention + feed forward

Nobel committee awards Strickland who advanced optics
**Linguistically-Informed Self-Attention**

Layer $r$  
**Multi-head self-attention + feed forward**

Layer $p$  
**Syntactically-informed self-attention**

Layer $J$  
**Multi-head self-attention + feed forward**

**Bilinear**

**predicates**

**args**

Nobel committee awards Strickland who advanced optics
Linguistically-Informed Self-Attention

Bilinear

categories

Layer J

Multi-head self-attention + feed forward

Layer p

Syntactically-informed self-attention

Layer r

Multi-head self-attention + feed forward

Nobel committee awards Strickland who advanced optics
Linguistically-Informed Self-Attention

Layer r
Multi-head self-attention + feed forward

Layer p
Syntactically-informed self-attention

Layer J
Multi-head self-attention + feed forward

Bilinear
predicates
args

Linguistically-Informed Self-Aff(on
SRL

arg predicates Bilinear

Nobel committee awards Strickland who advanced optics

Syntactically-informed self-attention

committee awards Strickland who advanced optics

Syntactically-informed self-attention

committee awards Strickland who advanced optics

Syntactically-informed self-attention

committee awards Strickland who advanced optics
## Linguistically-Informed Self-Attention

<table>
<thead>
<tr>
<th></th>
<th>GloVe</th>
<th>ELMo</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>in-domain (dev)</td>
<td>in-domain (dev)</td>
</tr>
<tr>
<td>He et al. 2017</td>
<td>81.5</td>
<td>---</td>
</tr>
<tr>
<td>He et al. 2018</td>
<td>81.6</td>
<td>85.3</td>
</tr>
<tr>
<td>SA</td>
<td>82.39</td>
<td>85.26</td>
</tr>
<tr>
<td><strong>LISA</strong></td>
<td><strong>82.24</strong></td>
<td><strong>85.35</strong></td>
</tr>
<tr>
<td>+D&amp;M</td>
<td>83.58</td>
<td>85.17</td>
</tr>
<tr>
<td>+Gold</td>
<td>86.81</td>
<td>87.63</td>
</tr>
</tbody>
</table>

Strubell et al. (2018)
Why SRL is difficult? or NLP in general

- Syntactic Alternation

The robot *plays* piano.

ARG0: player
ARG2: instrument

The cafe is *playing* my favorite song.

ARG0: player
ARG1: thing performed

The music *plays* softly.

ARG1: thing performed

ARGM-MNR
Why SRL is difficult? or NLP in general

- Prepositional Phrase (PP) Attachment

I *eat* [pasta] [with delight].

ARG0

eater

ARG1

meal

ARGM-MNR

manner

I *eat* [pasta with broccoli].

ARG0

eater

ARG1

meal
Why SRL is difficult? or NLP in general

- Long Dependencies

We *flew* to Chicago.

We remember the nice view *flying* to Chicago.

We remember John and Mary *flying* to Chicago.
Why SRL is difficult? or NLP in general

- Even harder for out-of-domain data

“Dip chicken breasts into eggs to coat”

Active, Ser133-phosphorylated CREB effects transcription of CRE-dependent genes via interaction with the 265-kDa ...
Abstract Meaning Representation
Abstract Meaning Representation

- Graph-structured annotation
- Nodes are variables labeled by concepts; edges are semantic relations

“Bob ate four cakes that he bought.”

(x2 / eat-01
  :ARG0 (x1 / person
    :name (n / name
      :op1 "Bob")
    :wiki "Bob_X")
  :ARG1 (x4 / cake
    :quant 4
    :ARG1-of (x7 / buy-01
      :ARG0 x1)))

Figure from Gruzitis et al. (2014)
Abstract Meaning Representation

- Superset of SRL: full sentence analyses, contains coreference and multi-word expressions as well
- F1 scores in the 60s: hard!
- So comprehensive that it’s hard to predict, but still doesn’t handle tense or some other things...

Banarescu et al. (2014)
Summarization with AMR

- Merge AMRs across multiple sentences
- Summarization = subgraph extraction

Sentence A: I saw Joe’s dog, which was running in the garden.
Sentence B: The dog was chasing a cat.

Summary: Joe’s dog was chasing a cat in the garden.

Liu et al. (2015)
Slot Filling
Most conservative, narrow form of IE

- **Magnitude**: 7.3
- **Time**: Sunday
- **Epicenter**: 103 kms (64 miles) southeast of the city of As-Sulaymaniyah

US Geological Survey initially said the quake was of a magnitude 7.2, before revising it to 7.3.

**Speaker**: Alan Clark
**Title**: “Gender Roles in the Holy Roman Empire”
**Location**: Allagher Center Main Auditorium

Old work: HMMs, later CRFs trained per role
**Slot Filling: MUC**

### Template

<table>
<thead>
<tr>
<th>SELLER</th>
<th>BUSINESS</th>
<th>ACQUIRED</th>
<th>PURCHASER</th>
</tr>
</thead>
<tbody>
<tr>
<td>CSR Limited</td>
<td>Oil and Gas</td>
<td>Delhi Fund</td>
<td>Esso Inc.</td>
</tr>
</tbody>
</table>

### Document

[S CSR] has said that [S it] has sold [S its] [B oil interests] held in [A Delhi Fund]. [P Esso Inc.] did not disclose how much [P they] paid for [A Dehli].

- Key aspect: need to combine information across multiple mentions of an entity using coreference

---

Haghighi and Klein (2010)
Extract code-related concepts in software engineering forums, but not everything that looks like a concept is a concept (e.g., *windows*).

```java
ArrayList<String> list = new ArrayList<String>();
list.add("http://www.google.com");
exchange.getOut().setHeader("endpoints", list);
```

and, inside camel route i want to iterate through this list

A variable name in this context

Relation Extraction
Relation Extraction

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

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

- Use NER-like system to identify entity spans, classify relations between entity pairs with a classifier

- Systems can be feature-based or neural, look at surface words, 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)

\[
\begin{align*}
Y & \text{ is a } X \\
X \text{ such as [list]} & \text{ cities such as Berlin, Paris, and London.} \\
\text{other } X \text{ including } Y & \text{ other cities including Berlin}
\end{align*}
\]

- 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]
Distant Supervision

- Learn decently accurate classifiers for ~100 Freebase relations
- Could be used to crawl the web and expand our knowledge base

<table>
<thead>
<tr>
<th>Relation name</th>
<th>100 instances</th>
<th>1000 instances</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Syn</td>
<td>Lex</td>
</tr>
<tr>
<td>/film/director/film</td>
<td>0.49</td>
<td>0.43</td>
</tr>
<tr>
<td>/film/writer/film</td>
<td>0.70</td>
<td>0.60</td>
</tr>
<tr>
<td>/geography/river/basin_countries</td>
<td>0.65</td>
<td>0.64</td>
</tr>
<tr>
<td>/location/country/administrative.divisions</td>
<td>0.68</td>
<td>0.59</td>
</tr>
<tr>
<td>/location/location/contains</td>
<td>0.81</td>
<td>0.89</td>
</tr>
<tr>
<td>/location/us_county/country_seat</td>
<td>0.51</td>
<td>0.51</td>
</tr>
<tr>
<td>/music/artist/origin</td>
<td>0.64</td>
<td>0.66</td>
</tr>
<tr>
<td>/people/deceased_person/place_of_death</td>
<td>0.80</td>
<td>0.79</td>
</tr>
<tr>
<td>/people/person/nationality</td>
<td>0.61</td>
<td>0.70</td>
</tr>
<tr>
<td>/people/person/place_of_birth</td>
<td>0.78</td>
<td>0.77</td>
</tr>
<tr>
<td><strong>Average</strong></td>
<td><strong>0.67</strong></td>
<td><strong>0.66</strong></td>
</tr>
</tbody>
</table>

Mintz et al. (2009)
Distant Supervision

- Inherently have noise in training data, need special methods (e.g., multi-instance learning) to handle false positives AND false negatives.

```
<table>
<thead>
<tr>
<th>Entity 1</th>
<th>Entity 2</th>
<th>Relation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Thailand</td>
<td>Bangkok</td>
<td>/location/country/capital</td>
</tr>
</tbody>
</table>
```

Sentences mentioning the two entities:

1. *Bangkok* is the most populous city of Thailand.
2. *Bangkok* grew rapidly during the 1960s through the 1980s and now exerts a significant impact among Thailand’s politics, economy, education, media and modern society.
3. The nation of *Thailand* is about to get its very first visit ever from a president this weekend, President Obama, so the American Embassy in *Bangkok* is understandably very excited right now.

Figure from Jiang et al. (2016)
Instead of labels on each individual instance, the learner only observes labels on bags of instances.

A bag is labeled negative, if all the examples in it are negative.

A bag is labeled positive, if there is at least one positive example.

Xu et al. (2013), Pershina et al. (2014), Tabassum (2016)
Multi-instance Learning

Xu et al. (2013), Pershina et al. (2014), Tabassum (2016)
Open IE
Open Information Extraction

- “Open”ness — want to be able to extract all kinds of information from open-domain text
- Acquire commonsense knowledge just from “reading” about it, but need to process lots of text (“machine reading”)
- Typically no fixed relation inventory
Extract positive examples of (e, r, e) triples via parsing and heuristics

Train a Naive Bayes classifier to filter triples from raw text: uses features on POS tags, lexical features, stopwords, etc.

Barack Obama, 44th president of the United States, was born on August 4, 1961 in Honolulu

=> Barack Obama, was born in, Honolulu

80x faster than running a parser (which was slow in 2007...)

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)
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
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.
Takeaways

- SRL/AMR: handle a bunch of phenomena, but more or less like syntax++ in terms of what they represent

- Relation extraction: can collect data with distant supervision, use this to expand knowledge bases

- Slot filling: tied to a specific ontology, but gives fine-grained information

- 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

- Many, many applications and techniques