Natural Language Processing

Lecture 18b: Semantic Roles
You should already have been convinced that grammatical structure is an important aspect of language.

Now we are discussing semantics or meaning.

Up until today, we have talked about meaning as something that individual words have (whether in isolation or in context).

So far today, we have talked about representing the meanings of propositions/sentences in meaning representation languages.

Now, we are going to discuss an enhancement to this view, the notion that individual noun phrases can be characterized as having roles relative to a predicate or frame.
• Noah built an ark out of gopher wood.
• He loaded two of every animal onto the ark.
• Noah piloted the ark into stormy weather.
• When the skies cleared, all rejoiced.
• Noah₁ built an ark₂ out of gopher wood.
• He₁ loaded two of every animal onto the ark₂.
• Noah₁ piloted the ark₂ into stormy weather.
• When the skies₃ cleared, all₄ rejoiced.
Paraphrase

• Noah built an ark out of gopher wood.
• An ark was built by Noah. It was made from gopher wood.
• Noah constructed an ark with wood from a gopher tree.
• Using gopher wood, Noah managed to put together an ark.
• Noah built an ark.
• ...

Traditional Semantic Roles

• In the linguistics literature, one sees a number of common terms for semantic roles
  – Agent
  – Patient
  – Theme
  – Force
  – Experiencer
  – Stimulus
  – Recipient
  – Source
  – Goal
  – etc.

• These have their place, and are useful to know if you want to understand what a semantic role is, but are not widely used in NLP

• In NLP, we tend to use finer-grained (and sometimes cryptically named) semantic role labels
Traditional Semantic Roles

- **David** *threw* **the midterms** from **Pausch Bridge** to **the hillside below**.
  - **David**—agent
  - **the midterms**—theme
  - **Pausch Bridge**—source
  - **the hillside below**—goal
Neo-Davidsonian Representation

• David *threw* the midterms from Pausch Bridge to the hillside below
  – $\text{THROW}(\text{David, midterms, PauschBridge, hillside})$
  – $\exists e \ \text{THROW}(e) \land \text{AGENT}(e, \text{David}) \land \text{THEME}(e, \text{midterms})$
    $\land \text{SOURCE}(e, \text{PauschBridge}) \land \text{GOAL}(e, \text{hillside})$

• The midterms were *thrown* from Pausch Bridge
  – $\text{THROW}(\text{midterms, PauschBridge})$
  – $\exists e \ \text{THROW}(e) \land \text{THEME}(e, \text{midterms}) \land \text{SOURCE}(e, \text{PauschBridge})$
Semantic Role Labeling

**Input:** a sentence, paragraph, or document

**Output:** for each predicate*, labeled spans identifying each of its arguments.

*Predicates are sometimes identified in the input, sometimes not.
Predicates

• Noah built an ark out of gopher wood.
• An ark was built by Noah. It was made from gopher wood.
• Noah constructed an ark with wood from a gopher tree.
• Using gopher wood, Noah managed to put together an ark.
Predicates and Arguments

• Noah built an ark out of gopher wood.
• An ark was built by Noah. It was made from gopher wood.
• Noah constructed an ark with wood from a gopher tree.
• Using gopher wood, Noah managed to put together an ark.
Breaking, Eating, Opening

- John broke the window.
- The window broke.
- John is always breaking things.
- The broken window testified to John’s malfeasance.
- Eat!
- We ate dinner.
- We already ate.
- The pies were eaten up quickly.
- Our gluttony was complete.
- Open up!
- Someone left the door open.
- John opens the window at night.
Introducing PropBank

- Corpus (PTB) with propositions annotated
  - Predicates (verbs)
  - Arguments (semantic roles)
- Semantic roles are Arg0, Arg1, etc., each with a description
  - Arg0 is typically the most agent-like argument
  - Labels for other arguments are somewhat arbitrary
“Agree” in PropBank

- **arg0**: agreeer
- **arg1**: proposition
- **arg2**: other entity agreeing
- The group agreed it wouldn’t make an offer.
- Usually John agrees with Mary on everything
“Fall (move downward)” in PropBank

- **arg1**: logical subject, patient, thing falling
- **arg2**: extent, amount fallen
- **arg3**: starting point
- **arg4**: ending point
- **argM-loc**: medium
- **Sales** fell to **$251.2 million** from **$278.8 million**.
- **The average junk bond** fell by **4.2%**.
- **The meteor** fell through **the atmosphere**, crashing into **Cambridge**.
Figure 20.16 Parse tree for a PropBank sentence, showing the PropBank argument labels. The dotted line shows the path feature NP↑S↓VP↓VBD for ARG0, the NP-SBJ constituent the San Francisco Examiner.
FrameNet

- A **frame** is a schematic representation of a situation involving various participants, and other conceptual roles
- In FrameNet, frames—not verbs—are first-class citizens
  - To a first approximation, verbs that relate to the same situation belong to the same frame
  - Roles are given fine-grained labels that are specific to the frame, but not the verb
  - Frames can center around words other than verbs
### change_position_on_a_scale

<table>
<thead>
<tr>
<th>Core roles</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>ATTRIBUTE</td>
<td>scalar property that the ITEM possesses</td>
</tr>
<tr>
<td>DIFFERENCE</td>
<td>distance by which an ITEM changes its position</td>
</tr>
<tr>
<td>FINAL_STATE</td>
<td>ITEM’s state after the change</td>
</tr>
<tr>
<td>FINAL_VALUE</td>
<td>position on the scale where ITEM ends up</td>
</tr>
<tr>
<td>INITIAL_STATE</td>
<td>ITEM’s state before the change</td>
</tr>
<tr>
<td>INITIAL_VALUE</td>
<td>position on the scale from which the ITEM moves</td>
</tr>
<tr>
<td>ITEM</td>
<td>entity that has a position on the scale</td>
</tr>
<tr>
<td>VALUE_RANGE</td>
<td>portion of the scale along which values of ATTRIBUTE fluctuate</td>
</tr>
</tbody>
</table>

| Some non-core roles ...                                                   |                                                                                           |
| DURATION                                                                  | length of time over which the change occurs                                               |
| SPEED                                                                     | rate of change of the value                                                               |
| GROUP                                                                     | the group in which an ITEM changes the value of an ATTRIBUTE                              |
• **Verbs**: advance, climb, decline, decrease, diminish, dip, double, drop, dwindle, edge, explode, fall, fluctuate, gain, grow, increase, jump, move, mushroom, plummet, reach, rise, rocket, shift, skyrocket, slide, soar, swell, swing, triple, tumble

• **Nouns**: decline, decrease, escalation, explosion, fall, fluctuation, gain, growth, hike, increase, rise, shift, tumble

• **Adverb**: increasingly
Demo

https://framenet.icsi.berkeley.edu/fndrupal/
How Can We Build an SRL System?

(1) Parse

(2) For each predicate word in the parse:
   For each node in the parse:
      Classify the node with respect to the predicate