Natural Language Processing

Lecture 23—11/17/2015
Jim Martin

Today

- More Semantics
  - Review/Finish up compositional semantics
- Semantic role labeling
- Quiz review

Semantic Analysis

- Semantic analysis is the process of taking in some linguistic input and assigning a meaning representation to it.
  - There a lot of different ways to do this that make more or less (or no) use of syntax
  - We're going to start with the idea that syntax does matter
    - The compositional rule-to-rule approach

Meaning

$$\exists e,x,y,z Giving(e) \wedge Giver(e,x) \wedge Given(e,y) \wedge Givee(e,z)$$

The *giving*, *giver*, *given*, *givee* predicates get their meaning from the set of facts that are encoded in some knowledge-base.

Problem

$$\exists e,x,y,z Giving(e) \wedge Giver(e,x) \wedge Given(e,y) \wedge Givee(e,z)$$

Unfortunately, this approach is flawed in two serious ways.

1. Doesn't take into account word senses. So the meaning of *given a book* and *give a cold* can't be same.
2. There's no easy way to capture the similarity of givers, takers, senders, holders, closers… etc.

Solutions

$$\exists e,x,y,z Giving_{99}(e) \wedge Giver_{99}(e,x) \wedge Given_{99}(e,y) \wedge Givee_{99}(e,z)$$

1. Use distinct word senses for the predicates and their roles.
2. Generalize the roles to capture similarities across roles for different words

Semantic Roles

- In our semantics examples, we used various FOL predicates to capture various aspects of events, including the notion of *roles*
  - Havers, takers, givers, servers, etc.
  - All specific to each verb/predicate.
- Thematic roles
  - Thematic roles are semantic generalizations over the specific roles that occur with specific verbs.
  - I.e. Takers, givers, eaters, makers, doers, killers, all have something in common
    - -er
    - They' re all the agents of the actions
  - *We can generalize across other roles as well to come up with a small finite set of such roles*

Thematic Roles

Thematic Role	Example
AGENT	*The waiter* spilled the soup.
EXPERIENCER	*John* has a headache.
FORCE	*The wind* blows debris from the mall into our yards.
THEME	Only after Benjamin Franklin broke *the ice*...
RESULT	The French government has built a *regulation-size baseball diamond*...
CONTENT	Mona asked *"You met Mary Ann at a supermarket?"*
INSTRUMENT	He turned to poaching catfish, stunning them *with a shocking device*...
BENEFICIARY	Whenever Ann Callahan makes hotel reservations *for her boss*...
SOURCE	I flew in *from Boston*.
GOAL	I drove *to Portland*.

Thematic Roles

- Takes some of the work away from the verbs.
  - It' s not the case that every verb is unique and has to completely specify how all of its arguments behave
  - Provides a locus for organizing semantic processing
  - It permits us to distinguish near surface-level semantics from deeper semantics

Linking

- Thematic roles, syntactic categories and their positions in larger syntactic structures are all intertwined in complicated ways. For example…
  - AGENTS often appear as grammatical subjects
  - In a `VP->V NP` rule, the NP is often a THEME
- So how might we go about studying/ investigating these ideas?
  - Get a corpus
  - Do some annotation

Resources

- For parsing we had TreeBanks
- For lexical semantics we have WordNets
- So for thematic roles....

Resources

- There are 2 major English resources out there with thematic-role-like data
  - PropBank
    - Layered on the Penn TreeBank
      - Small number (25ish) labels
      - For each semantic predicate, identify the constituents in the tree that are arguments to that predicate and
      - Label each with its appropriate role
    - Many domain-specific variants
  - FrameNet
    - Based on a theory of semantics known as frame semantics.
      - Large number of frame-specific labels

Propbank Example

- *Cover* (as in smear)
  - Arg0 (agent: the causer of the covering)
  - Arg1 (theme: "thing covered")
  - Arg2 (covering: "stuff being smeared")
  - *[McAdams and crew] covered [ ] with [checked linoleum].*

Propbank

- Arg0 and Arg1 roughly correspond to the notions of agent and theme
  - Causer and thing most directly effected
- The remaining args are verb specific
  - So there really aren't a small finite set of roles
  - Arg3 for "cover" isn't the same as the Arg3 for "give"...

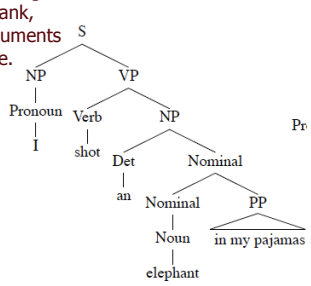
Problems

- What exactly is a role?
- What's the right set of roles?
- Are such roles universal across languages?
- Are these roles atomic?
  - I.e. Agents
    - Animate, Volitional, Direct causers, etc
- Can we automatically label syntactic constituents with thematic roles?
  - Semantic Role Labeling (next)

Semantic Role Labeling

Semantic Role Labeling

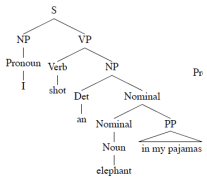
- SRL is the task of automatically (1) identifying and (2) labeling the thematic roles (propbank, framenet, etc.) to the arguments to each verb in a sentence.



Pr

Two Tasks

- Given a sentence and parse
  - For each verb in a sentence
    - For each constituent
      - Decide if it is an argument to that verb
      - And if it is an argument, determine what kind.



Pr

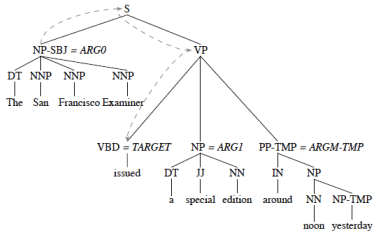
General Approach

- Supervised machine learning using a resource like Propbank as a training set
  1. Train a binary classifier to do the "Is this an argument task".
  2. Train an multi-class classifier to further classify the particular role type.
- In both cases, features are extracted from the syntactic parse and lexical items

Features

- Lots of the features have been used...
  - One of the most important is the "path" feature

NP↑S↓VP↓VBD



Quiz Questions