Dependency Parsing

Instructor: Yoav Artzi

Slides adapted from Dan Klein, Luke Zettlemoyer, Chris Manning, and Dan Jurafsky, and David Weiss
Overview

• The parsing problem
• Methods
  – Transition-based parsing
• Evaluation
• Projectivity
Parse Trees

- Part-of-speech Tagging:
  - Word classes
- Parsing:
  - From words to phrases to sentences
  - Relations between words
- Two views
  - Dependency
  - Constituency
Dependency Parsing

• Dependency structure shows which words depend on (modify or are arguments of) which other words.

The boy put the tortoise on the rug
Constituency (Phrase Structure) Parsing

• Phrase structure organizes words into nested constituents
• Linguists can, and do, argue about details
• Lots of ambiguity

```
new art critics write reviews with computers
```
Dependency Structure

• Syntactic structure consists of:
  – Lexical items
  – Binary asymmetric relations \(\rightarrow\) dependencies

Dependencies are typed with name of grammatical relation
Dependency Structure

- Syntactic structure consists of:
  - Lexical items
  - Binary asymmetric relations → dependencies
Dependency Structure

- Syntactic structure consists of:
  - Lexical items
  - Binary asymmetric relations \(\rightarrow\) dependencies

Dependencies form a tree
Dependency Structure

• Syntactic structure consists of:
  – Lexical items
  – Binary asymmetric relations $\rightarrow$ dependencies

Dependencies form a tree
Let’s Parse

John saw Mary

He said that the boy who was wearing the blue shirt with the white pockets has left the building.
Methods for Dependency Parsing

- Dynamic programming
  - Eisner (1996): $O(n^3)$
- Graph algorithms
  - McDonald et al. (2005): score edges independently using classifier and use maximum spanning tree
- Constraint satisfaction
  - Start with all edges, eliminate based on hard constraints
- “Deterministic parsing”
  - Left-to-right, each choice is done with a classifier

```
jumped
    nsubj
      boy
        det
          the
            det
              the
                det
                  fence
        amod
          little
            the
        prep
          over
```

What are the sources of information for dependency parsing?

1. Bilexical affinities
   - [issues ➔ the] is plausible
2. Dependency distance
   - mostly with nearby words
3. Intervening material
   - Dependencies rarely span intervening verbs or punctuation
4. Valency of heads
   - How many dependents on which side are usual for a head?

ROOT Discussion of the outstanding issues was completed .
MaltParse (Nivre et al. 2008)

- Greedy transition-based parser
- Each decision: how to attach each word as we encounter it
  - If you are familiar: like shift-reduce parser
- Select each action with a classifier
- The parser has:
  - a stack $\sigma$, written with the top to the right
    - which starts with the ROOT symbol
  - a buffer $\beta$, written with the top to the left
    - which starts with the input sentence
  - a set of dependency arcs $A$
    - which starts off empty
  - a set of actions
Arc-standard Dependency Parsing

Start: $\sigma = [\text{ROOT}], \beta = w_1, \ldots, w_n, A = \emptyset$

- **Shift** $\sigma, w_i|\beta, A \rightarrow \sigma|w_i, \beta, A$
- **Left-Arc$_r$** $\sigma|w_i, w_j|\beta, A \rightarrow \sigma, w_j|\beta, A \cup \{r(w_j,w_i)\}$
- **Right-Arc$_r$** $\sigma|w_i, w_j|\beta, A \rightarrow \sigma, w_i|\beta, A \cup \{r(w_i,w_j)\}$

Finish: $\beta = \emptyset$

ROOT Joe likes Marry