NLP
Introduction to NLP

Dependency Parsing 3
## Evaluation of Dependency Parsing

- **Attachment Score** (Buchholz & Marsi 2006)
  - \# correct deps/\# deps (attached to the right head)
  - Unlabeled dependency accuracy (UAS)
  - Labeled dependency accuracy (LAS)

<table>
<thead>
<tr>
<th>#</th>
<th>Word</th>
<th>POS</th>
<th>Word</th>
<th>POS</th>
<th>Label</th>
<th>#</th>
<th>Word</th>
<th>POS</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Unionized</td>
<td>Unionized</td>
<td>VBN</td>
<td>VBN</td>
<td>-</td>
<td>2</td>
<td>NMOD</td>
<td>-</td>
</tr>
<tr>
<td>2</td>
<td>workers</td>
<td>workers</td>
<td>NNS</td>
<td>NNS</td>
<td>-</td>
<td>3</td>
<td>SBJ</td>
<td>-</td>
</tr>
<tr>
<td>3</td>
<td>are</td>
<td>are</td>
<td>VBP</td>
<td>VBP</td>
<td>-</td>
<td>0</td>
<td>ROOT</td>
<td>-</td>
</tr>
<tr>
<td>4</td>
<td>usually</td>
<td>usually</td>
<td>RB</td>
<td>RB</td>
<td>-</td>
<td>3</td>
<td>TMP</td>
<td>-</td>
</tr>
<tr>
<td>5</td>
<td>better</td>
<td>better</td>
<td>RBR</td>
<td>RBR</td>
<td>-</td>
<td>4</td>
<td>ADV</td>
<td>-</td>
</tr>
<tr>
<td>6</td>
<td>paid</td>
<td>paid</td>
<td>VBN</td>
<td>VBN</td>
<td>-</td>
<td>5</td>
<td>AMOD</td>
<td>-</td>
</tr>
<tr>
<td>7</td>
<td>than</td>
<td>than</td>
<td>IN</td>
<td>IN</td>
<td>-</td>
<td>5</td>
<td>AMOD</td>
<td>-</td>
</tr>
<tr>
<td>8</td>
<td>their</td>
<td>their</td>
<td>PRP$</td>
<td>PRP$</td>
<td>-</td>
<td>10</td>
<td>NMOD</td>
<td>-</td>
</tr>
<tr>
<td>9</td>
<td>non-union</td>
<td>non-union</td>
<td>JJ</td>
<td>JJ</td>
<td>-</td>
<td>10</td>
<td>NMOD</td>
<td>-</td>
</tr>
<tr>
<td>10</td>
<td>counterparts</td>
<td>counterparts</td>
<td>NNS</td>
<td>NNS</td>
<td>-</td>
<td>7</td>
<td>PMOD</td>
<td>-</td>
</tr>
</tbody>
</table>
Complexity

- Projective (CKY) $O(n^5)$
- Projective (Eisner) $O(n^3)$
- Non-projective (MST - Chu-Liu-Edmonds) $O(n^2)$
- Projective (Malt) $O(n)$
Use in Information Extraction

Figure 1: The dependency tree of the sentence “The results demonstrated that KaiC interacts rhythmically with KaiA, KaiB, and SasA.”

1. KaiC - nsubj - interacts - prep_with - SasA
2. KaiC - nsubj - interacts - prep_with - SasA - conj_and - KaiA
3. KaiC - nsubj - interacts - prep_with - SasA - conj_and - KaiB
4. SasA - conj_and - KaiA
5. SasA - conj_and - KaiB

1. PROTX1 - nsubj - interacts - prep_with - PROTX2
2. PROTX1 - nsubj - interacts - prep_with - PROTX0 - conj_and - PROTX2
3. PROTX1 - nsubj - interacts - prep_with - PROTX0 - conj_and - PROTX2
4. PROTX1 - conj_and - PROTX2
5. PROTX1 - conj_and - PROTX2
6. PROTX1 - conj_and - PROTX0 - conj_and - PROTX2

[Erkan et al. 2007]
Dependency Kernels

$S_1 = \text{Protesters seized several pumping stations, holding 127 Shell workers hostage.}$

$S_2 = \text{Troops recently have raided churches, warning ministers to stop preaching.}$

Figure 1: Sentences as dependency graphs.

<table>
<thead>
<tr>
<th>Relation Instance</th>
<th>Shortest Path in Undirected Dependency Graph</th>
</tr>
</thead>
<tbody>
<tr>
<td>$S_1$: protesters AT stations</td>
<td>protesters $\rightarrow$ seized $\rightarrow$ stations</td>
</tr>
<tr>
<td>$S_1$: workers AT stations</td>
<td>workers $\rightarrow$ holding $\rightarrow$ protesters $\rightarrow$ seized $\leftarrow$ stations</td>
</tr>
<tr>
<td>$S_2$: troops AT churches</td>
<td>troops $\rightarrow$ raided $\rightarrow$ churches</td>
</tr>
<tr>
<td>$S_2$: ministers AT churches</td>
<td>ministers $\rightarrow$ warning $\rightarrow$ troops $\rightarrow$ raided $\leftarrow$ churches</td>
</tr>
</tbody>
</table>

Table 1: Shortest Path representation of relations.

[Bunescu and Mooney 2005]
External Links

• [http://ilk.uvt.nl/conll/](http://ilk.uvt.nl/conll/)  
  – CONLL-X Shared task
  – Prague Dependency Treebank
• [http://nextens.uvt.nl/depparse-wiki/SharedTaskWebsite](http://nextens.uvt.nl/depparse-wiki/SharedTaskWebsite)
• [http://nextens.uvt.nl/depparse-wiki/DataOverview](http://nextens.uvt.nl/depparse-wiki/DataOverview)
• [http://maltparser.org/](http://maltparser.org/)  
  – Joakim Nivre’s Maltparser
• [http://www.cs.ualberta.ca/~lindek/minipar.htm](http://www.cs.ualberta.ca/~lindek/minipar.htm)  
  – Dekang Lin’s Minipar
• [http://www.link.cs.cmu.edu/link/](http://www.link.cs.cmu.edu/link/)  
  – Daniel Sleator and Davy Temperley’s Link parser
Notes

• The original versions of MSTParser and MaltParser from 2007 achieve about 81% accuracy
  – Highest in Japanese (91-92%)
  – Lowest in Arabic and Turkish (63-67%)

• Non-projective parsing is harder than projective parsing
NLP