LING 388: Language and Computers

Sandiway Fong

Lecture 2
Administrivia

• Lab Exercises Today

• Homework 1 is out today
  – *due next Monday by midnight*
  – *I’ll go over the homework in class next Tuesday*
Example from Last Time

• Ambiguity
  – *Where can I see the bus stop?*
  – *stop*: verb or part of the noun-noun compound *bus stop*
  – Context (Discourse or situation)

http://www.clker.com
Example from Last Time

**Stanford Parser**

Please enter a sentence to be parsed:

where can I see the bus stop?

Language: [English] Sample Sentence Parse

Your query

where can I see the bus stop?

Tagging

where/WRB can/MD I/FRP see/VB the/DT bus/NN stop/NN ?/.  

Parse

(ROOT (SBARQ (WHADVP (WRB where)) (SQ (MD can) (NP (FRP I)) (VP (VB see) (NP (DT the) (NN bus) (NN stop)))) (.))))

http://nlp.stanford.edu:8080/parser/
### Example from Last Time

• **POS Tagging:**

  **Your query**

  *where can I see the bus stop?*

  **Tagging**

  *where/WRB can/MD I/PRP see/VB the/DT bus/NN stop/NN ?/.*

<table>
<thead>
<tr>
<th>Number</th>
<th>Tag</th>
<th>Description</th>
<th>Tag</th>
<th>Description</th>
<th>Tag</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>CC</td>
<td>Coordinating conjunction</td>
<td>13</td>
<td>NNS</td>
<td>25</td>
<td>TO</td>
</tr>
<tr>
<td>2</td>
<td>CD</td>
<td>Cardinal number</td>
<td>14</td>
<td>NNP</td>
<td>26</td>
<td>UH Interjection</td>
</tr>
<tr>
<td>3</td>
<td>DT</td>
<td>Determiner</td>
<td>15</td>
<td>NNPS</td>
<td>27</td>
<td>VB Verb, base form</td>
</tr>
<tr>
<td>4</td>
<td>EX</td>
<td>Existential <em>there</em></td>
<td>16</td>
<td>PDT</td>
<td>28</td>
<td>VBD Verb, past tense</td>
</tr>
<tr>
<td>5</td>
<td>FW</td>
<td>Foreign word</td>
<td>17</td>
<td>POS</td>
<td>29</td>
<td>VBG Verb, gerund or present participle</td>
</tr>
<tr>
<td>6</td>
<td>IN</td>
<td>Preposition or subordinating</td>
<td>18</td>
<td>PRP</td>
<td>30</td>
<td>VBN Verb, past participle</td>
</tr>
<tr>
<td>7</td>
<td>JJ</td>
<td>Adjective</td>
<td>19</td>
<td>PRPS</td>
<td>31</td>
<td>VBP Verb, non-3rd person singular present</td>
</tr>
<tr>
<td>8</td>
<td>JJR</td>
<td>Adjective, comparative</td>
<td>20</td>
<td>RB</td>
<td>32</td>
<td>VBZ Verb, 3rd person singular present</td>
</tr>
<tr>
<td>9</td>
<td>JJS</td>
<td>Adjective, superlative</td>
<td>21</td>
<td>RBR</td>
<td>33</td>
<td>WDT Wh-determiner</td>
</tr>
<tr>
<td>10</td>
<td>LS</td>
<td>List item marker</td>
<td>22</td>
<td>RBS</td>
<td>34</td>
<td>WP Wh-pronoun</td>
</tr>
<tr>
<td>11</td>
<td>MD</td>
<td>Modal</td>
<td>23</td>
<td>RP</td>
<td>35</td>
<td>WPS Possessive wh-pronoun</td>
</tr>
<tr>
<td>12</td>
<td>NN</td>
<td>Noun, singular or mass</td>
<td>24</td>
<td>SYM</td>
<td>36</td>
<td>WRB Wh-adverb</td>
</tr>
</tbody>
</table>
Example from Last Time

• S-EXP Parse:
  1. (PHRASE S-EXP₁ .. S-EXPₙ)  e.g. (VP (VB see) (NP ... )
  2. (POS WORD)               e.g. (NN bus)

Tagset (Penn Treebank)
http://www.ldc.upenn.edu/Catalog/docs/LDC95T7/cl93.html
Example from Last Time

• Phrase tagset:

<table>
<thead>
<tr>
<th></th>
<th>Tag</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>ADJP</td>
<td>Adjective phrase</td>
</tr>
<tr>
<td>2</td>
<td>ADVP</td>
<td>Adverb phrase</td>
</tr>
<tr>
<td>3</td>
<td>NP</td>
<td>Noun phrase</td>
</tr>
<tr>
<td>4</td>
<td>PP</td>
<td>Prepositional phrase</td>
</tr>
<tr>
<td>5</td>
<td>S</td>
<td>Simple declarative clause</td>
</tr>
<tr>
<td>6</td>
<td>SBAR</td>
<td>Clause introduced by subordinating conjunction or 0 (see below)</td>
</tr>
<tr>
<td>7</td>
<td>SBARQ</td>
<td>Direct question introduced by wh-word or wh-phrase</td>
</tr>
<tr>
<td>8</td>
<td>SINV</td>
<td>Declarative sentence with subject-aux inversion</td>
</tr>
<tr>
<td>9</td>
<td>SQ</td>
<td>Subconstituent of SBARQ excluding wh-word or wh-phrase</td>
</tr>
<tr>
<td>10</td>
<td>VP</td>
<td>Verb phrase</td>
</tr>
<tr>
<td>11</td>
<td>WHADVP</td>
<td>wh-adverb phrase</td>
</tr>
<tr>
<td>12</td>
<td>WHNP</td>
<td>wh-noun phrase</td>
</tr>
<tr>
<td>13</td>
<td>WHPP</td>
<td>wh-prepositional phrase</td>
</tr>
<tr>
<td>14</td>
<td>X</td>
<td>Constituent of unknown or uncertain category</td>
</tr>
</tbody>
</table>

http://www.uniml.com/pmwiki
Example from Last Time

- Typed dependencies between words in a sentence form a graph:
  - relation(Word₁, Word₂) – Word₂ modifies Word₁
  - Example: nsubj(see, I) – “I modifies see: I is the nominal subject of see”

Typed dependencies

- advmod(see-4, where-1)
- aux(see-4, can-2)
- nsubj(see-4, I-3)
- root(ROOT-0, see-4)
- det(stop-7, the-5)
- nn(stop-7, bus-6)
- dobj(see-4, stop-7)

7 relations = 7 arcs in the graph

Explanation of the relations used can be found in http://nlp.stanford.edu/software/dependencies_manual.pdf
Example from Last Time

- Comparison (parse vs. typed dependency graph):
Exercise 1

• Consider the syntactically ambiguous sentence:
  – *I prodded the boy with a stick*.

• What is the ambiguity here?
• How can that ambiguity be expressed?
• Which interpretation does the Stanford Parser prefer?
Exercise 2

• Compare with:
  – *The boy with a stick prodded me.*

• Why is this sentence not similarly ambiguous?
• Does the Stanford Parser give the right parse?
Exercise 3

• Compare the Stanford parser interpretation for the two sentences:
  – *I prodded the boy with a stick.*
  – *I saw the boy with a telescope.*

• What is the difference?
• Can you force the parser to change structure by changing the object of the preposition?
Exercise 4

• Contrast the following two sentences:
  – John is too stubborn to talk to.
  – John is too stubborn to talk to Bill.

• What is missing from the output of the Stanford Parser?
Exercise 4

Stanford Parser

Please enter a sentence to be parsed:
John is too stubborn to talk to.

Language: English  Sample Sentence  Parse

Your query
John is too stubborn to talk to.

Tagging
John/NNP is/VBZ too/RB stubborn/JJ to/TO talk/VB to/TO ./.

Parse
(Root
  (S
    (NP (NNP John))
    (VP (VBZ is)
      (ADJP (RB too) (JJ stubborn))
    )
  )
  (S
    (VP (TO to)
      (VP (VB talk)
        (PP (TO to)))))
  (.
 .)))

Typed dependencies
nsubj(stubborn-4, John-1)
cop(stubborn-4, is-2)
advmod(stubborn-4, too-3)
root(ROOT-0, stubborn-4)
aux(talk-6, to-5)
xcomp(stubborn-4, talk-6)
prep(talk-6, to-7)

Typed dependencies, collapsed
nsubj(stubborn-4, John-1)
xsubj(talk-6, John-1)
cop(stubborn-4, is-2)
advmod(stubborn-4, too-3)
root(ROOT-0, stubborn-4)
aux(talk-6, to-5)
xcomp(stubborn-4, talk-6)
prep(talk-6, to-7)

Statistics
Tokens: 8
Time: 0.038 s
Exercise 4

Stanford Parser

Please enter a sentence to be parsed:
John is too stubborn to talk to Bill.

Your query
John is too stubborn to talk to Bill.

Tagging
John/NNP is/VBZ too/RB stubborn/JJ to/TO talk/VB to/TO Bill/NNP ./.

Parse
(ROOT
  (S
    (NP (NNP John))
    (VP (VBZ is)
      (ADJP (RB too) (JJ stubborn))
      (S
        (VP (TO to)
          (VP (VB talk)
            (PP (TO to)
              (NP (NNP Bill)))))))))

Typed dependencies
nsbj(stubborn-4, John-1)
cop(stubborn-4, is-2)
advm(stubborn-4, too-3)
root(ROOT-0, stubborn-4)
aux(talk-6, to-5)
xcomp(stubborn-4, talk-6)
prep(talk-6, to-7)
pobj(to-7, Bill-8)

Typed dependencies, collapsed
nsbj(stubborn-4, John-1)
xsubj(talk-6, John-1)
cop(stubborn-4, is-2)
advm(stubborn-4, too-3)
root(ROOT-0, stubborn-4)
aux(talk-6, to-5)
xcomp(stubborn-4, talk-6)
prep_to(talk-6, Bill-8)

Statistics
Tokens: 9
Time: 0.061 s
Homework 1

• Remember the syllabus described last time …
  – submit by email to your TA Ben Martin
    bamartin@email.arizona.edu
  – next Monday (by midnight)
  – Subject of email should read: LING 388 Homework 1
  – Put your name at the top of the file
  – Collect all answers in one file
    • plain text or PDF formats only please (not .docx)
    • should you choose to include any screen snapshots or
      illustrations to support your answer, they should not be in
      separate attachments – make them part of your single
      document
Homework 1

Examine the Stanford Parser output on these two sentences:

1. *Which car did Mary like?*
2. *Which car did Mary like the convertible?*

Questions:

1. (4pts) Does the typed dependencies output offer any advantage over the parse output for the 1st sentence?
2. (4pts) What is wrong with the 2nd sentence?
3. (4pts) Is it possible to deduce from the Stanford Parser output that the 2nd sentence is ungrammatical?
Homework 1 Review

• Recall the ambiguous example:
  – *Where can I see the bus stop?*
  the Stanford parser analyses “bus stop” preferentially as a noun-noun compound:
    (NP (DT the) (NN bus) (NN stop))
  4. (4pts) Give a (question) sentence ending in “... *the bus stop?*” where the Stanford parser analyses “stop” as a verb.
  5. (4pts) What syntactic situations would force a parser to decide to analyze “stop” in “... *the bus stop?*” as a noun (vs. a verb)?
Homework Resources

• Tagset (Penn Treebank)
  – Parts of speech (POS)
  – https://www.ling.upenn.edu/courses/Fall_2003/ling001/penn_treebank_pos.html
  – More detailed, including syntactic labels
  – http://www.ldc.upenn.edu/Catalog/docs/LDC95T7/cl93.html

• Dependencies