Syntactic Processing: Parts-of-Speech Tagging

CSE354 - Spring 2020
Task

- Syntactic Processing
  - Parts-of-Speech Tagging

how?

- Machine learning:
  - Logistic regression
Parts-of-Speech

Open Class:

Nouns, Verbs, Adjectives, Adverbs
Parts-of-Speech

Open Class:

Nouns, Verbs, Adjectives, Adverbs

Function words:

Determiners, conjunctions, pronouns, prepositions
# Parts-of-Speech: The Penn Treebank Tagset

<table>
<thead>
<tr>
<th></th>
<th>Description</th>
<th></th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. CC</td>
<td>Coordinating conjunction</td>
<td>25. TO</td>
<td>to</td>
</tr>
<tr>
<td>2. CD</td>
<td>Cardinal number</td>
<td>26. UH</td>
<td>Interjection</td>
</tr>
<tr>
<td>3. DT</td>
<td>Determiner</td>
<td>27. VB</td>
<td>Verb, base form</td>
</tr>
<tr>
<td>4. EX</td>
<td>Existential <em>there</em></td>
<td>28. VBD</td>
<td>Verb, past tense</td>
</tr>
<tr>
<td>5. FW</td>
<td>Foreign word</td>
<td>29. VBG</td>
<td>Verb, gerund/present participle</td>
</tr>
<tr>
<td>6. IN</td>
<td>Preposition/subordinating conjunction</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. JJ</td>
<td>Adjective</td>
<td>30. VBN</td>
<td>Verb, past participle</td>
</tr>
<tr>
<td>8. JJR</td>
<td>Adjective, comparative</td>
<td>31. VBP</td>
<td>Verb, non-3rd ps. sing. present</td>
</tr>
<tr>
<td>9. JJS</td>
<td>Adjective, superlative</td>
<td>32. VBZ</td>
<td>Verb, 3rd ps. sing. present</td>
</tr>
<tr>
<td>10. LS</td>
<td>List item marker</td>
<td>33. WDT</td>
<td><em>wh</em>-determiner</td>
</tr>
<tr>
<td>11. MD</td>
<td>Modal</td>
<td>34. WP</td>
<td><em>wh</em>-pronoun</td>
</tr>
<tr>
<td>12. NN</td>
<td>Noun, singular or mass</td>
<td>35. WP$</td>
<td>Possessive <em>wh</em>-pronoun</td>
</tr>
<tr>
<td>13. NNS</td>
<td>Noun, plural</td>
<td>36. WRB</td>
<td><em>wh</em>-adverb</td>
</tr>
<tr>
<td>14. NNP</td>
<td>Proper noun, singular</td>
<td>37. #</td>
<td>Punctuation</td>
</tr>
<tr>
<td>15. NNPS</td>
<td>Proper noun, plural</td>
<td>38. $</td>
<td>Dollar sign</td>
</tr>
<tr>
<td>16. PDT</td>
<td>Preposition</td>
<td>39. .</td>
<td>Sentence-final punctuation</td>
</tr>
<tr>
<td>17. POS</td>
<td>Possessive ending</td>
<td>40. ,</td>
<td>Comma</td>
</tr>
<tr>
<td>18. PRP</td>
<td>Personal pronoun</td>
<td>41. :</td>
<td>Colon, semi-colon</td>
</tr>
<tr>
<td>19. PPS</td>
<td>Possessive pronoun</td>
<td>42. (</td>
<td>Left bracket character</td>
</tr>
<tr>
<td>20. RB</td>
<td>Adverb</td>
<td>43. )</td>
<td>Right bracket character</td>
</tr>
<tr>
<td>21. RBR</td>
<td>Adverb, comparative</td>
<td>44. &quot;</td>
<td>Straight double quote</td>
</tr>
<tr>
<td>22. RBS</td>
<td>Adverb, superlative</td>
<td>45. '</td>
<td>Left open single quote</td>
</tr>
<tr>
<td>23. RP</td>
<td>Particle</td>
<td>46. &quot;</td>
<td>Left open double quote</td>
</tr>
<tr>
<td>24. SYM</td>
<td>Symbol (mathematical or scientific)</td>
<td>47. '</td>
<td>Right close single quote</td>
</tr>
<tr>
<td></td>
<td></td>
<td>48. &quot;</td>
<td>Right close double quote</td>
</tr>
</tbody>
</table>
# Parts-of-Speech: Social Media Tagset (Gimpel et al., 2010)

## Nominal, Nominal + Verbal

<table>
<thead>
<tr>
<th>Tag</th>
<th>Description</th>
<th>Examples</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>common noun (NN, NNS)</td>
<td>books someone</td>
<td>13.7</td>
</tr>
<tr>
<td>O</td>
<td>pronoun (personal/WH; not possessive; PRP, WP)</td>
<td>it you u mee</td>
<td>6.8</td>
</tr>
<tr>
<td>S</td>
<td>nominal + possessive</td>
<td>books’ someone’s</td>
<td>0.1</td>
</tr>
<tr>
<td>P</td>
<td>proper noun (NNP, NNPS)</td>
<td>lebron usa iPad</td>
<td>6.4</td>
</tr>
<tr>
<td>Z</td>
<td>proper noun + possessive</td>
<td>America’s</td>
<td>0.2</td>
</tr>
<tr>
<td>L</td>
<td>nominal + verbal</td>
<td>he’s book’ll iono (= I don’t know)</td>
<td>1.6</td>
</tr>
</tbody>
</table>

## Other open-class words

| V   | verb incl. copula, auxiliaries (V*, MD) | might gonna ought couldn’t is eats | 15.1 |
| A   | adjective (J*) | good fav lil | 5.1 |
| R   | adverb (R*, WRB) | 2 (i.e., too) | 4.6 |
| I   | interjection (UH) | lol haha FTW yea right | 2.6 |

## Other closed-class words

| D   | determiner (WDT, DT, WP$, PRP$) | the teh its it’s | 6.5 |
| P   | pre- or postposition, or subordinating conjunction (IN, TO) | while to for 2 (i.e., to) 4 (i.e., for) | 8.7 |
| &   | coordinating conjunction (CC) | and n & + BUT | 1.7 |
| T   | verb particle (RP) | out off Up UP | 0.6 |
| X   | existential there, predeterminers (EX, PDT) | both | 0.1 |
| Y   | X + verbal | there’s all’s | 0.0 |

## Miscellaneous

<table>
<thead>
<tr>
<th>Tag</th>
<th>Description</th>
<th>Examples</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>$</td>
<td>punctuation (#, $, ' ', (,, ), , , ; , ` , '')</td>
<td>2010 four 9:30</td>
<td>1.5</td>
</tr>
<tr>
<td></td>
<td>other abbreviations, foreign words, possessive endings, symbols, garbage (FW, POS, SYM, LS)</td>
<td>illy (I love you) wby (what about you)’s</td>
<td>1.1</td>
</tr>
<tr>
<td>@</td>
<td>at-mention (indicates another user as a recipient of a tweet)</td>
<td>@nObama</td>
<td>4.9</td>
</tr>
<tr>
<td>~</td>
<td>discourse marker, indications of continuation of a message across multiple tweets</td>
<td>RT and : in retweet construction RT @user : hello</td>
<td>3.4</td>
</tr>
<tr>
<td>U</td>
<td>URL or email address</td>
<td><a href="http://bit.ly/xyz">http://bit.ly/xyz</a></td>
<td>1.6</td>
</tr>
<tr>
<td>E</td>
<td>emoticon</td>
<td>:-</td>
<td>:(&lt;3 o...O</td>
</tr>
</tbody>
</table>
POS Tagging: Applications

- Resolving ambiguity (speech: “lead”)
- Shallow searching: find noun phrases
- Speed up parsing
- Use as feature (or in place of word)
POS Tagging: Applications

- Resolving ambiguity (speech: “lead”)
- Shallow searching: find noun phrases
- Speed up parsing
- Use as feature (or in place of word)

For this course:

- An introduction to language-based classification (logistic regression)
- Understand what modern deep learning methods are dealing with implicitly.
Window-based POS Tagging

The book looks brief so I am happy.
Window-based POS Tagging

The book looks brief so I am happy.
The book looks brief so I am happy.
Window-based POS Tagging

The *book* looks *brief* so I am happy.

* D  N  ?
The book looks brief so I am happy.
Window-based POS Tagging

The book looks brief so I am happy.
Window-based POS Tagging

The book looks brief so I am happy.
Window-based POS Tagging

The book looks brief so I am happy.

D N V ?

window size of 3
Window-based POS Tagging

The book *looks* brief so I am happy.

D N V ?
Window-based POS Tagging

The book looks brief so I am happy.

$P(pos_i = 'N'|word_i = "brief") = 0.3$
Window-based POS Tagging

The book looks brief so I am happy.

P(pos$_i$ = ‘N’|word$_i$ = “brief”) = 0.3
P(pos$_i$ = ‘V’|word$_i$ = “brief”) = 0.4
P(pos$_i$ = ‘A’|word$_i$ = “brief”) = 0.3
Window-based POS Tagging

The book looks brief so I am happy.

\[
\begin{align*}
P(p_i='N'|w_i=brief) &= .30 \\
P(p_i='V'|w_i=brief) &= .40 \\
P(p_i='A'|w_i=brief) &= .30
\end{align*}
\]
Window-based POS Tagging

The book looks brief so I am happy.

The book looks brief so I am happy.

P(p_i='N'|w_i=brief,w_{i-1}=looks,w_{i+1}=so) = ??

P(p_i='V'|w_i=brief,w_{i-1}=looks,w_{i+1}=so) = ??

P(p_i='A'|w_i=brief,w_{i-1}=looks,w_{i+1}=so) = ??
The book looks brief so I am happy.

The given text: The book looks brief so I am happy.

Window size of 3

\[ P(p_i='N'|w_i=brieff,w_{i-1}=looks,w_{i+1}=so) = 0.005 \]
\[ P(p_i='V'|w_i=brieff,w_{i-1}=looks,w_{i+1}=so) = 0.005 \]
\[ P(p_i='A'|w_i=brieff,w_{i-1}=looks,w_{i+1}=so) = 0.99 \]
The book *looks* brief so I am happy.

Window size of 3

P($p_i$='N'|$w_i$=brief,$w_{i-1}$=looks,$w_{i+1}$=so) = 0.3
P($p_i$='V'|$w_i$=brief,$w_{i-1}$=looks,$w_{i+1}$=so) = 0.4
P($p_i$='A'|$w_i$=brief,$w_{i-1}$=looks,$w_{i+1}$=so) = 0.3

More likely, because we haven’t seen this context before.
Window-based POS Tagging

More likely, because we haven’t seen this context before.

The book **looks** brief so I am happy.

\[
P(p_i='N'|w_i=brief, w_{i-1}=looks, w_{i+1}=so) = .3
\]

\[
P(p_i='V'|w_i=brief, w_{i-1}=looks, w_{i+1}=so) = .4
\]

\[
P(p_i='A'|w_i=brief, w_{i-1}=looks, w_{i+1}=so) = .3
\]
The book looks brief so I am happy.

\[ P(p_i='N'|w_i=\text{brief}, w_{i-1}=\text{looks}, w_{i+1}=\text{so}) = 0.3 \]
\[ P(p_i='V'|w_i=\text{brief}, w_{i-1}=\text{looks}, w_{i+1}=\text{so}) = 0.4 \]
\[ P(p_i='A'|w_i=\text{brief}, w_{i-1}=\text{looks}, w_{i+1}=\text{so}) = 0.3 \]
The book looks brief so I am happy.

$P(p_i='N'|w_i=brief,w_{i-1}=looks,w_{i+1}=so) = .3$

$P(p_i='V'|w_i=brief,w_{i-1}=looks,w_{i+1}=so) = .4$

$P(p_i='A'|w_i=brief,w_{i-1}=looks,w_{i+1}=so) = .3$
Sequential Model

The book looks brief so I am happy.

P(p_i='N'|p_{i-1}=V) = 0.4
P(p_i='V'|p_{i-1}=V) = 0.10
P(p_i='A'|p_{i-1}=V) = 0.4

window size of 3

sequence order of 1
Sequential Model

The book looks brief so I am happy.

P(p_i='N'|p_{i-1}=V,w_i=brief) = .3
P(p_i='V'|p_{i-1}=V,w_i=brief) = .05
P(p_i='A'|p_{i-1}=V,w_i=brief) = .65
Sequence modeling

-- Tasks that in which a current label is dependent on previous labels within a sequence.

More generally: tasks that can leverage the order of words.

Most basic example: Language Modeling

-- Predicting the next word given previous.