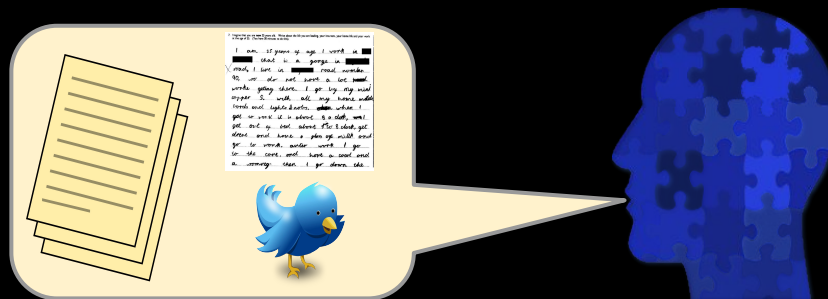# Syntactic Processing: Parts-of-Speech Tagging

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 2
The Penn Treebank POS tagset.

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

Parts-of-Speech:
Social Media Tagset
(Gimpel et al., 2010)

Tag	Description	Examples	%
Nominal, Nominal + Verbal			
N	common noun (NN, NNS)	books someone	13.7
O	pronoun (personal/WH; not possessive; PRP, WP)	it you u meeee	6.8
S	nominal + possessive	books' someone's	0.1
^	proper noun (NNP, NNPS)	lebron usa iPad	6.4
Z	proper noun + possessive	America's	0.2
L	nominal + verbal	he's book'll iono (= *I don't know*)	1.6
M	proper noun + verbal	Mark'll	0.0

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
!	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

Twitter/online-specific			
#	hashtag (indicates topic/category for tweet)	#acl	1.0
@	at-mention (indicates another user as a recipient of a tweet)	@BarackObama	4.9
~	discourse marker, indications of continuation of a message across multiple tweets	RT and : in retweet construction RT @user : hello	3.4
U	URL or email address	http://bit.ly/xyz	1.6
E	emoticon	:-) :b (: <3 o_O	1.0
Miscellaneous			
$	numeral (CD)	2010 four 9:30	1.5
,	punctuation (#, $, '', (,) , , . :, ``)	!!! .... ?!?	11.6
G	other abbreviations, foreign words, possessive endings, symbols, garbage (FW, POS, SYM, LS)	ily (*I love you*) wby (*what about you*) 's ♫ --> awesome...I'm	1.1

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 .

D

Window-based POS Tagging

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

D   N

Window-based POS Tagging

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

D   N   ?

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 .

D N V A

Window-based POS Tagging

The book looks brief so I am happy .

↓ ↓ ↓ ↓

D N V ?

Window-based POS Tagging

window size
of 3

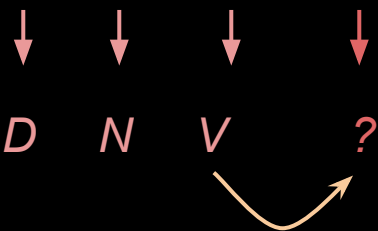*The book looks brief so I am happy .*

↓   ↓   ↓   ↓

D   N   V   ?

Window-based POS Tagging

window size of 3

The book looks brief so I am happy .

D N V ?

Window-based POS Tagging

window size
of 3

The book looks brief so I am happy .

D   N   V   ?

$P(pos_i = \text{'N'} | word_i = \text{"brief"}) = 0.3$

Window-based POS Tagging

window size of 3

The book looks brief so I am happy .

D   N   V   ?

$P(pos_i = \text{'N'}|word_i = \text{"brief"}) = 0.3$

$P(pos_i = \text{'V'}|word_i = \text{"brief"}) = 0.4$

$P(pos_i = \text{'A'}|word_i = \text{"brief"}) = 0.3$

Window-based POS Tagging

window size of 3

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

↓ ↓ ↓ ↓

D   N   V   ?

$P(p_i=\text{'N'}|w_i=brief) = .30$
$P(p_i=\text{'V'}|w_i=brief) = .40$
$P(p_i=\text{'A'}|w_i=brief) = .30$

Window-based POS Tagging

window size
of 3

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

D   N   V   ?

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

Window-based POS Tagging

window size of 3

ideal result

The book looks brief so I am happy .

D   N   V   ?

$P(p_i=\text{'N'}|w_i=brief,w_{i-1}=looks,w_{i+1}=so) = .005$

$P(p_i=\text{'V'}|w_i=brief,w_{i-1}=looks,w_{i+1}=so) = .005$

$P(p_i=\text{'A'}|w_i=brief,w_{i-1}=looks,w_{i+1}=so) = .99$

Window-based POS Tagging

window size of 3

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

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

D N V ?

$P(p_i = \text{'N'} | w_i = brief, w_{i-1} = looks, w_{i+1} = so) = .3$
$P(p_i = \text{'V'} | w_i = brief, w_{i-1} = looks, w_{i+1} = so) = .4$
$P(p_i = \text{'A'} | w_i = brief, w_{i-1} = looks, w_{i+1} = so) = .3$

Window-based POS Tagging

window size of 3

The book looks brief so I am happy .

D   N   V   ?

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

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

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

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

Sequential Model

window size of 3

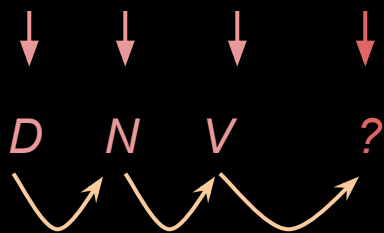*The book looks brief so I am happy .*

D   N   V   ?

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

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

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

sequence order of 1

Sequential Model

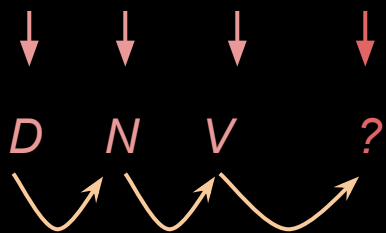window size
of 3

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

D   N   V   ?

$$P(p_i = 'N' | w_i = brief, w_{i-1} = looks, w_{i+1} = so) = .3$$
$$P(p_i = 'V' | w_i = brief, w_{i-1} = looks, w_{i+1} = so) = .4$$
$$P(p_i = 'A' | w_i = brief, w_{i-1} = looks, w_{i+1} = so) = .3$$

sequence
order of 1

Sequential Model

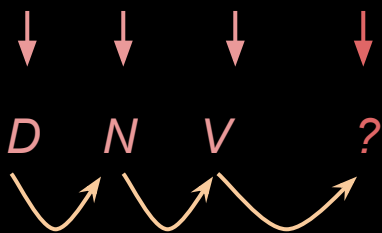window size of 3

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

D   N   V   ?

$P(p_i='N'|p_{i-1}=V) = .4$
$P(p_i='V'|p_{i-1}=V) = .10$
$P(p_i='A'|p_{i-1}=V) = .4$

sequence order of 1

Sequential Model

window size of 3

The book looks brief so I am happy .

D   N   V   ?

sequence order of 1

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

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.