

PART 1 – Tagging: Tag these two sentences, using the PTB tagset below. Don't bother with punctuation, but tag all other tokens.

WASHINGTON -- Pope Francis , the spiritual leader of 1.2 billion Catholics , challenged Congress and by extension the mightiest nation in history on Thursday to break out of its cycle of polarization and paralysis to finally use its power to heal the `` open wounds '' of a planet torn by hatred , greed , poverty and pollution .



Juliana Hausser @Jul_hausser · Sep 14

Thought we had off for the pope visit, idk wtf I'm doing

Tag	Description	Example	Tag	Description	Example
CC	coordin. conjunction	<i>and, but, or</i>	SYM	symbol	<i>+, %, &</i>
CD	cardinal number	<i>one, two</i>	TO	“to”	<i>to</i>
DT	determiner	<i>a, the</i>	UH	interjection	<i>ah, oops</i>
EX	existential ‘there’	<i>there</i>	VB	verb base form	<i>eat</i>
FW	foreign word	<i>mea culpa</i>	VBD	verb past tense	<i>ate</i>
IN	preposition/sub-conj	<i>of, in, by</i>	VBG	verb gerund	<i>eating</i>
JJ	adjective	<i>yellow</i>	VBN	verb past participle	<i>eaten</i>
JJR	adj., comparative	<i>bigger</i>	VBP	verb non-3sg pres	<i>eat</i>
JJS	adj., superlative	<i>wildest</i>	VBZ	verb 3sg pres	<i>eats</i>
LS	list item marker	<i>1, 2, One</i>	WDT	wh-determiner	<i>which, that</i>
MD	modal	<i>can, should</i>	WP	wh-pronoun	<i>what, who</i>
NN	noun, sing. or mass	<i>llama</i>	WP\$	possessive wh-	<i>whose</i>
NNS	noun, plural	<i>llamas</i>	WRB	wh-adverb	<i>how, where</i>
NNP	proper noun, sing.	<i>IBM</i>	\$	dollar sign	<i>\$</i>
NNPS	proper noun, plural	<i>Carolinas</i>	#	pound sign	<i>#</i>
PDT	predeterminer	<i>all, both</i>	“	left quote	<i>‘ or “</i>
POS	possessive ending	<i>'s</i>	”	right quote	<i>’ or ”</i>
PRP	personal pronoun	<i>I, you, he</i>	(left parenthesis	<i>[, (, {, <</i>
PRP\$	possessive pronoun	<i>your, one’s</i>)	right parenthesis	<i>],), }, ></i>
RB	adverb	<i>quickly, never</i>	,	comma	<i>,</i>
RBR	adverb, comparative	<i>faster</i>	.	sentence-final punc	<i>. ! ?</i>
RBS	adverb, superlative	<i>fastest</i>	:	mid-sentence punc	<i>: ; ... --</i>
RP	particle	<i>up, off</i>			

Figure 9.1 Penn Treebank part-of-speech tags (including punctuation).

PART 2 – Phrase extraction: In the text, bracket all instances of noun phrases, of length 2 or higher, that are identified by one of the Justeson&Katz patterns. Recall that they are defined over coarse tags (Adj, Noun, Prep.) as follows. (*Noun* does not include pronouns, but does include both common and proper nouns.)

Adj* Noun+ (including: AN, NN, NNN, ANN, AAN, etc.)

Adj* Noun+ Prep Adj* Noun+ (including: NPN ANPN ANPAN etc.)

PART 3 — Better patterns: The point of phrase extraction is to identify terms that may be meaningful for users, or for classification features or other purposes. **Propose two new patterns**, one of which must include verbs in some way. For each, give examples of phrases it would match, and justify why the pattern is useful. The examples could be from the text examples here, or they could be new ones you make up.