

COMP 3225

Natural Language Processing

Parts of Speech Tagging

Stuart E. Middleton

sem03@soton.ac.uk

University of Southampton

Copyright University of Southampton 2021.

Content for internal use at University of Southampton only.

Slides may include content publicly shared for education purposes via <https://web.stanford.edu/~jurafsky/slp3/>

Overview

- Introduction to Parts of Speech (POS)
- Tagsets
- POS Tagging
- <break - discussion point>
- Hidden Markov Model (HMM) POS tagger

Introduction to Parts of Speech (POS)

- What is Parts of Speech?
 - The concept of **Parts-of-Speech (POS)** probably originated from Dionysius Thrax of Alexandria around 100 BC attempting to summarize Greek linguistic knowledge
 - Thrax defined eight POS which we still use today
 - Noun, Verb, Pronoun, Preposition, Adverb, Conjunction, Participle, Article
 - POS can also be called **word classes**, **morphological classes** or **lexical tags**
 - POS are generally assigned to individual words or morphemes
 - Labelling POS is called **POS Tagging**
- Proper names
 - A proper name is called a **Named Entity**, and can be a multi-word phrase (e.g. New York)
 - Labelling named entities is called **Named Entity Recognition (NER)**
 - Named entity types include Location, Person, Organization etc.

Introduction to Parts of Speech (POS)

- Why are POS and Named Entities useful?
 - POS gives us clues to neighboring words and syntactic structure
 - Nouns are preceded by determiners and adjectives ... the river
 - Verbs are preceded by nouns ... swimming in the river
 - Verbs have dependency links to nouns swimming ^{dobj} → the river
 - POS tagging a key aspect of parsing natural language
 - NER is important to many natural language understanding tasks such as question answering, stance detection and information extraction
- Sequence labelling tasks
 - Input >> X >> Sequence of words
 - Output >> Y >> Sequence of labels
 - $\text{len}(X) == \text{len}(Y)$
- Both POS tagging and NER can be formulated as a sequence labelling task

Introduction to Parts of Speech (POS)

	Tag	Description	Example
Open Class	ADJ	Adjective: noun modifiers describing properties	<i>red, young, awesome</i>
	ADV	Adverb: verb modifiers of time, place, manner	<i>very, slowly, home, yesterday</i>
	NOUN	words for persons, places, things, etc.	<i>algorithm, cat, mango, beauty</i>
	VERB	words for actions and processes	<i>draw, provide, go</i>
	PROPN	Proper noun: name of a person, organization, place, etc..	<i>Regina, IBM, Colorado</i>
	INTJ	Interjection: exclamation, greeting, yes/no response, etc.	<i>oh, um, yes, hello</i>
Closed Class Words	ADP	Adposition (Preposition/Postposition): marks a noun's spacial, temporal, or other relation	<i>in, on, by under</i>
	AUX	Auxiliary: helping verb marking tense, aspect, mood, etc.,	<i>can, may, should, are</i>
	CCONJ	Coordinating Conjunction: joins two phrases/clauses	<i>and, or, but</i>
	DET	Determiner: marks noun phrase properties	<i>a, an, the, this</i>
	NUM	Numeral	<i>one, two, first, second</i>
	PART	Particle: a preposition-like form used together with a verb	<i>up, down, on, off, in, out, at, by</i>
	PRON	Pronoun: a shorthand for referring to an entity or event	<i>she, who, I, others</i>
Other	SCONJ	Subordinating Conjunction: joins a main clause with a subordinate clause such as a sentential complement	<i>that, which</i>
	PUNCT	Punctuation	<i>; , ()</i>
	SYM	Symbols like \$ or emoji	<i>\$, %</i>
	X	Other	<i>asdf, qwfg</i>

The 17 parts of speech in the Universal Dependencies tagset for English

Introduction to Parts of Speech (POS)

	Tag	Description	Example
Open Class	ADJ	Adjective: noun modifiers describing properties	<i>red, young, awesome</i>
	ADV	Adverb: verb modifiers of time, place, manner	<i>very, slowly, home, yesterday</i>
	NOUN	words for persons, places, things, etc.	<i>algorithm, cat, mango, beauty</i>
	VERB	words for actions and processes	<i>draw, provide, go</i>
	PROPN	Proper noun: name of a person, organization, place, etc..	<i>Regina, IBM, Colorado</i>
	INTJ	Interjection: exclamation, greeting, yes/no response, etc.	<i>oh, um, yes, hello</i>
Closed Class Words	ADP	Adposition (Preposition/Postposition): marks a noun's spacial, temporal, or other relation	<i>in, on, by under</i>
	AUX	Auxiliary: helping verb marking tense, aspect, mood, etc.,	<i>can, may, should, are</i>
	CCONJ	Coordinating Conjunction: joins two phrases/clauses	<i>and, or, but</i>
	DET	Determiner: marks noun phrase properties	<i>a, an, the, this</i>
	NUM	Numeral	<i>one, two, first, second</i>
	PART	Particle: a preposition-like form used together with a verb	<i>up, down, on, off, in, out, at, by</i>
	PRON	Pronoun: a shorthand for referring to an entity or event	<i>she, who, I, others</i>
Other	SCONJ	Subordinating Conjunction: joins a main clause with a subordinate clause such as a sentential complement	<i>that, which</i>
	PUNCT	Punctuation	<i>; , ()</i>
	SYM	Symbols like \$ or emoji	<i>\$, %</i>
	X	Other	<i>asdf, qwfg</i>

- **Closed Classes** - fixed membership
 - Typically **function words** used for structuring grammar (of, it, and, you)
- **Open Classes** - open membership
 - **Noun** (including **proper noun**), **verb**, **adjective**, **adverb**, **interjection**

Introduction to Parts of Speech (POS)

	Tag	Description	Example
Open Class	ADJ	Adjective: noun modifiers describing properties	<i>red, young, awesome</i>
	ADV	Adverb: verb modifiers of time, place, manner	<i>very, slowly, home, yesterday</i>
	NOUN	words for persons, places, things, etc.	<i>algorithm, cat, mango, beauty</i>
	VERB	words for actions and processes	<i>draw, provide, go</i>
	PROPN	Proper noun: name of a person, organization, place, etc..	<i>Regina, IBM, Colorado</i>
	INTJ	Interjection: exclamation, greeting, yes/no response, etc.	<i>oh, um, yes, hello</i>
Closed Class Words	ADP	Adposition (Preposition/Postposition): marks a noun's spacial, temporal, or other relation	<i>in, on, by under</i>
	AUX	Auxiliary: helping verb marking tense, aspect, mood, etc.,	<i>can, may, should, are</i>
	CCONJ	Coordinating Conjunction: joins two phrases/clauses	<i>and, or, but</i>
	DET	Determiner: marks noun phrase properties	<i>a, an, the, this</i>
	NUM	Numeral	<i>one, two, first, second</i>
	PART	Particle: a preposition-like form used together with a verb	<i>up, down, on, off, in, out, at, by</i>
	PRON	Pronoun: a shorthand for referring to an entity or event	<i>she, who, I, others</i>
Other	SCONJ	Subordinating Conjunction: joins a main clause with a subordinate clause such as a sentential complement	<i>that, which</i>
	PUNCT	Punctuation	<i>; , ()</i>
	SYM	Symbols like \$ or emoji	<i>\$, %</i>
	X	Other	<i>asdf, qwfg</i>

- **Nouns**

- **Common Noun** - concrete terms, abstractions and verb-like terms
 - **Count Noun** - occur in the singular and plural (goat/goats, relationship/relationships, and irregular ones like sheep/sheep)
 - **Mass Noun** - homogeneous groups (snow, salt)
- **Proper Noun** - names of things

Introduction to Parts of Speech (POS)

	Tag	Description	Example
Open Class	ADJ	Adjective: noun modifiers describing properties	<i>red, young, awesome</i>
	ADV	Adverb: verb modifiers of time, place, manner	<i>very, slowly, home, yesterday</i>
	NOUN	words for persons, places, things, etc.	<i>algorithm, cat, mango, beauty</i>
	VERB	words for actions and processes	<i>draw, provide, go</i>
	PROPN	Proper noun: name of a person, organization, place, etc..	<i>Regina, IBM, Colorado</i>
	INTJ	Interjection: exclamation, greeting, yes/no response, etc.	<i>oh, um, yes, hello</i>
Closed Class Words	ADP	Adposition (Preposition/Postposition): marks a noun's spacial, temporal, or other relation	<i>in, on, by under</i>
	AUX	Auxiliary: helping verb marking tense, aspect, mood, etc.,	<i>can, may, should, are</i>
	CCONJ	Coordinating Conjunction: joins two phrases/clauses	<i>and, or, but</i>
	DET	Determiner: marks noun phrase properties	<i>a, an, the, this</i>
	NUM	Numeral	<i>one, two, first, second</i>
	PART	Particle: a preposition-like form used together with a verb	<i>up, down, on, off, in, out, at, by</i>
	PRON	Pronoun: a shorthand for referring to an entity or event	<i>she, who, I, others</i>
Other	SCONJ	Subordinating Conjunction: joins a main clause with a subordinate clause such as a sentential complement	<i>that, which</i>
	PUNCT	Punctuation	<i>; , ()</i>
	SYM	Symbols like \$ or emoji	<i>\$, %</i>
	X	Other	<i>asdf, qwfg</i>

- **Verbs** - actions and processes
- **Adjective** - properties or qualities of nouns

Introduction to Parts of Speech (POS)

	Tag	Description	Example
Open Class	ADJ	Adjective: noun modifiers describing properties	<i>red, young, awesome</i>
	ADV	Adverb: verb modifiers of time, place, manner	<i>very, slowly, home, yesterday</i>
	NOUN	words for persons, places, things, etc.	<i>algorithm, cat, mango, beauty</i>
	VERB	words for actions and processes	<i>draw, provide, go</i>
	PROPN	Proper noun: name of a person, organization, place, etc..	<i>Regina, IBM, Colorado</i>
	INTJ	Interjection: exclamation, greeting, yes/no response, etc.	<i>oh, um, yes, hello</i>
Closed Class Words	ADP	Adposition (Preposition/Postposition): marks a noun's spacial, temporal, or other relation	<i>in, on, by under</i>
	AUX	Auxiliary: helping verb marking tense, aspect, mood, etc.,	<i>can, may, should, are</i>
	CCONJ	Coordinating Conjunction: joins two phrases/clauses	<i>and, or, but</i>
	DET	Determiner: marks noun phrase properties	<i>a, an, the, this</i>
	NUM	Numeral	<i>one, two, first, second</i>
	PART	Particle: a preposition-like form used together with a verb	<i>up, down, on, off, in, out, at, by</i>
	PRON	Pronoun: a shorthand for referring to an entity or event	<i>she, who, I, others</i>
Other	SCONJ	Subordinating Conjunction: joins a main clause with a subordinate clause such as a sentential complement	<i>that, which</i>
	PUNCT	Punctuation	<i>; , ()</i>
	SYM	Symbols like \$ or emoji	<i>\$, %</i>
	X	Other	<i>asdf, qwfg</i>

- **Adverb - modifier words**
 - **Locative Adverb** - direction or location of some action (here, downhill)
 - **Degree Adverbs** - extent of some action (extremely, very, somewhat)
 - **Manner Adverbs** - manner of some action (slowly, slinkily, delicately)
 - **Temporal Adverbs** - time of action or event (yesterday, Monday)

Introduction to Parts of Speech (POS)

	Tag	Description	Example
Open Class	ADJ	Adjective: noun modifiers describing properties	<i>red, young, awesome</i>
	ADV	Adverb: verb modifiers of time, place, manner	<i>very, slowly, home, yesterday</i>
	NOUN	words for persons, places, things, etc.	<i>algorithm, cat, mango, beauty</i>
	VERB	words for actions and processes	<i>draw, provide, go</i>
	PROPN	Proper noun: name of a person, organization, place, etc..	<i>Regina, IBM, Colorado</i>
	INTJ	Interjection: exclamation, greeting, yes/no response, etc.	<i>oh, um, yes, hello</i>
Closed Class Words	ADP	Adposition (Preposition/Postposition): marks a noun's spacial, temporal, or other relation	<i>in, on, by under</i>
	AUX	Auxiliary: helping verb marking tense, aspect, mood, etc.,	<i>can, may, should, are</i>
	CCONJ	Coordinating Conjunction: joins two phrases/clauses	<i>and, or, but</i>
	DET	Determiner: marks noun phrase properties	<i>a, an, the, this</i>
	NUM	Numeral	<i>one, two, first, second</i>
	PART	Particle: a preposition-like form used together with a verb	<i>up, down, on, off, in, out, at, by</i>
	PRON	Pronoun: a shorthand for referring to an entity or event	<i>she, who, I, others</i>
Other	SCONJ	Subordinating Conjunction: joins a main clause with a subordinate clause such as a sentential complement	<i>that, which</i>
	PUNCT	Punctuation	<i>; , ()</i>
	SYM	Symbols like \$ or emoji	<i>\$, %</i>
	X	Other	<i>asdf, qwfg</i>

- **Interjection** - exclamation (oh, hey, alas), greetings (hello, goodbye) and question responses (yes, no, uh-huh)

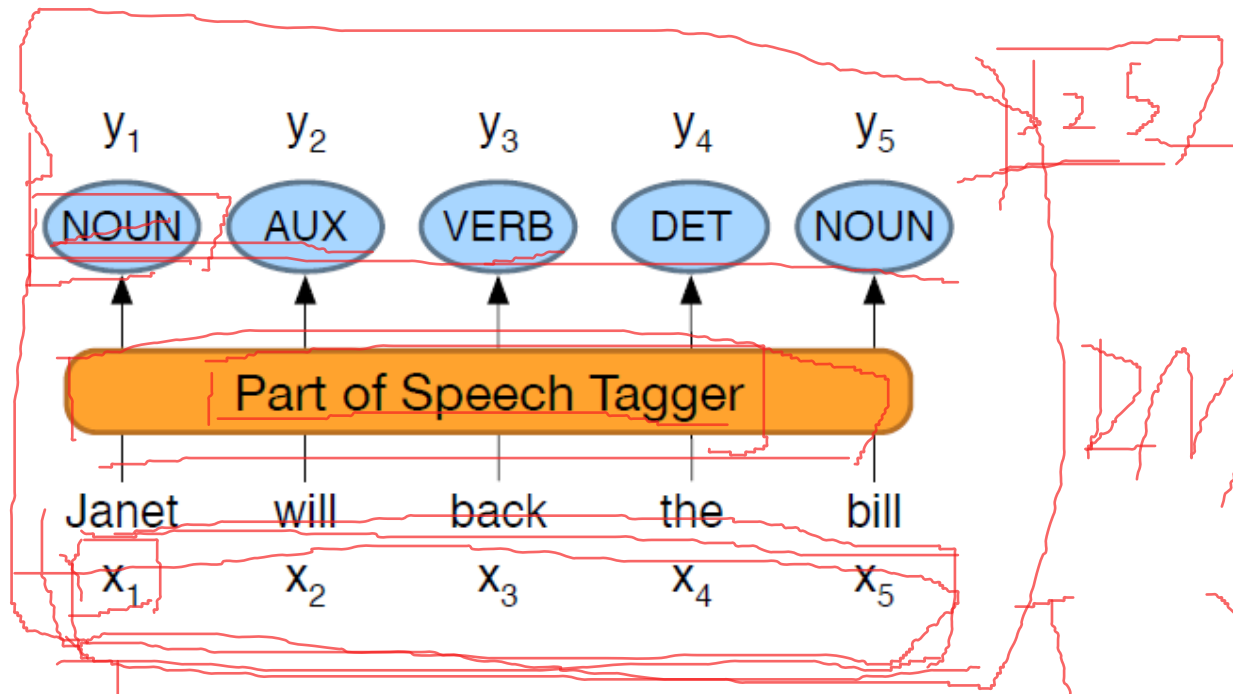
Tagsets

- A list of POS labels is called a **Tagset**
- Tagsets come in different shapes and sizes
 - Penn Treebank (45 labels)
 - Brown Corpus (87 labels)
 - C7 Tagset (146 labels)
- Penn Treebank

Tag	Description	Example	Tag	Description	Example	Tag	Description	Example
CC	coord. conj.	<i>and, but, or</i>	NNP	proper noun, sing.	<i>IBM</i>	TO	“to”	<i>to</i>
CD	cardinal number	<i>one, two</i>	NNPS	proper noun, plu.	<i>Carolinas</i>	UH	interjection	<i>ah, oops</i>
DT	determiner	<i>a, the</i>	NNS	noun, plural	<i>llamas</i>	VB	verb base	<i>eat</i>
EX	existential ‘there’	<i>there</i>	PDT	predeterminer	<i>all, both</i>	VBD	verb past tense	<i>ate</i>
FW	foreign word	<i>mea culpa</i>	POS	possessive ending	<i>’s</i>	VBG	verb gerund	<i>eating</i>
IN	preposition/ subordin-conj	<i>of, in, by</i>	PRP	personal pronoun	<i>I, you, he</i>	VBN	verb past partici- ple	<i>eaten</i>
JJ	adjective	<i>yellow</i>	PRP\$	possess. pronoun	<i>your, one’s</i>	VBP	verb non-3sg-pr	<i>eat</i>
JJR	comparative adj	<i>bigger</i>	RB	adverb	<i>quickly</i>	VBZ	verb 3sg pres	<i>eats</i>
JJS	superlative adj	<i>wildest</i>	RBR	comparative adv	<i>faster</i>	WDT	wh-determ.	<i>which, that</i>
LS	list item marker	<i>I, 2, One</i>	RBS	superlatv. adv	<i>fastest</i>	WP	wh-pronoun	<i>what, who</i>
MD	modal	<i>can, should</i>	RP	particle	<i>up, off</i>	WP\$	wh-possess.	<i>whose</i>
NN	sing or mass noun	<i>llama</i>	SYM	symbol	<i>+, %, &</i>	WRB	wh-adverb	<i>how, where</i>

POS Tagging

- POS tagging is the process of assigning a POS tag to each word in a text
 - Input sequence $\gg X \gg x_1; x_2; \dots; x_n$ of (tokenized) words
 - Output sequence $\gg Y \gg y_1; y_2; \dots; y_n$ of POS tags
 - Each output y_i corresponding exactly to one input x_i



POS Tagging

- POS Tagging is a disambiguation task
 - Words are ambiguous - they have more than one possible POS
 - Goal is to find the correct tag for any given situation
- Example
 - 'book' can be a verb (book that flight) or a noun (hand me that book).
- There are many ways you can build a POS tagger
 - Rule-based taggers
 - Set of possible POS for a word >> hand-crafted disambiguation rules
 - EngCG tagger >> 56,000 entries each with a set of morphological and syntactic features
 - Transformation-based taggers
 - Supervised learning of tagging rules + small set of hand-crafted templates
 - Brill tagger
 - Hidden Markov Model (HMM) >> this lecture
 - Conditional Random Fields (CRF) >> NER lecture

Break

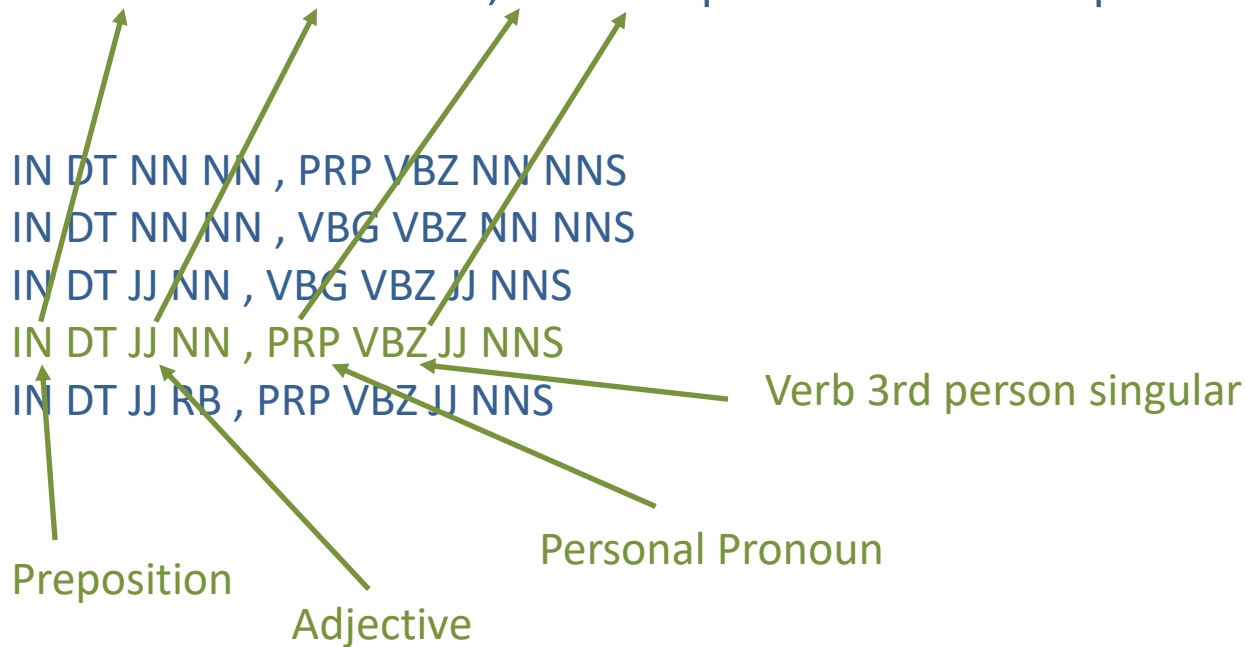
- Panopto Quiz - discussion point
- What would be the PENN Treebank POS tags for the following sentence:
"At the same time, it develops multinational operations"

IN DT NN NN , PRP VBZ NN NNS
 IN DT NN NN , VBG VBZ NN NNS
 IN DT JJ NN , VBG VBZ JJ NNS
 IN DT JJ NN , PRP VBZ JJ NNS
 IN DT JJ RB , PRP VBZ JJ NNS

Tag	Description	Example	Tag	Description	Example	Tag	Description	Example
CC	coord. conj.	<i>and, but, or</i>	NNP	proper noun, sing.	<i>IBM</i>	TO	“to”	<i>to</i>
CD	cardinal number	<i>one, two</i>	NNPS	proper noun, plu.	<i>Carolinas</i>	UH	interjection	<i>ah, oops</i>
DT	determiner	<i>a, the</i>	NNS	noun, plural	<i>llamas</i>	VB	verb base	<i>eat</i>
EX	existential ‘there’	<i>there</i>	PDT	predeterminer	<i>all, both</i>	VBD	verb past tense	<i>ate</i>
FW	foreign word	<i>mea culpa</i>	POS	possessive ending	<i>’s</i>	VBG	verb gerund	<i>eating</i>
IN	preposition/ subordin-conj	<i>of, in, by</i>	PRP	personal pronoun	<i>I, you, he</i>	VBN	verb past partici- ple	<i>eaten</i>
JJ	adjective	<i>yellow</i>	PRP\$	possess. pronoun	<i>your, one’s</i>	VBP	verb non-3sg-pr	<i>eat</i>
JJR	comparative adj	<i>bigger</i>	RB	adverb	<i>quickly</i>	VBZ	verb 3sg pres	<i>eats</i>
JJS	superlative adj	<i>wildest</i>	RBR	comparative adv	<i>faster</i>	WDT	wh-determ.	<i>which, that</i>
LS	list item marker	<i>1, 2, One</i>	RBS	superlatv. adv	<i>fastest</i>	WP	wh-pronoun	<i>what, who</i>
MD	modal	<i>can, should</i>	RP	particle	<i>up, off</i>	WP\$	wh-possess.	<i>whose</i>
NN	sing or mass noun	<i>llama</i>	SYM	symbol	<i>+, %, &</i>	WRB	wh-adverb	<i>how, where</i>

Break

- Panopto Quiz - discussion point
- What would be the PENN Treebank POS tags for the following sentence:
"At the same time, it develops multinational operations"

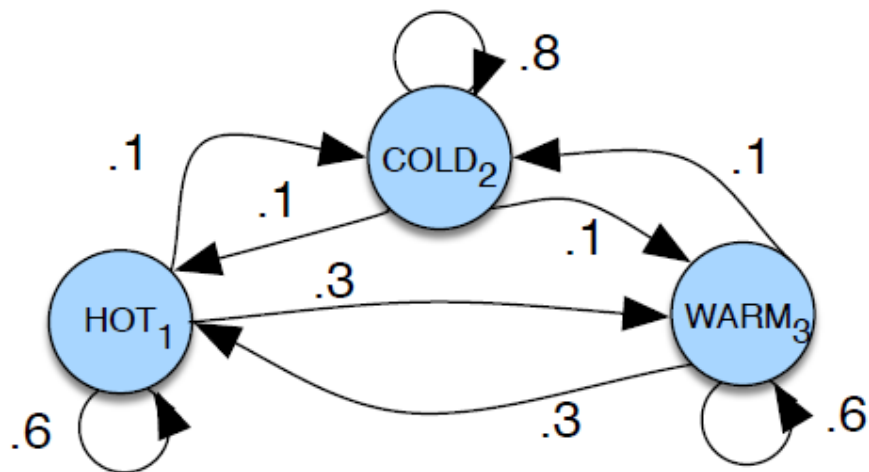


Hidden Markov Model (HMM) POS tagger

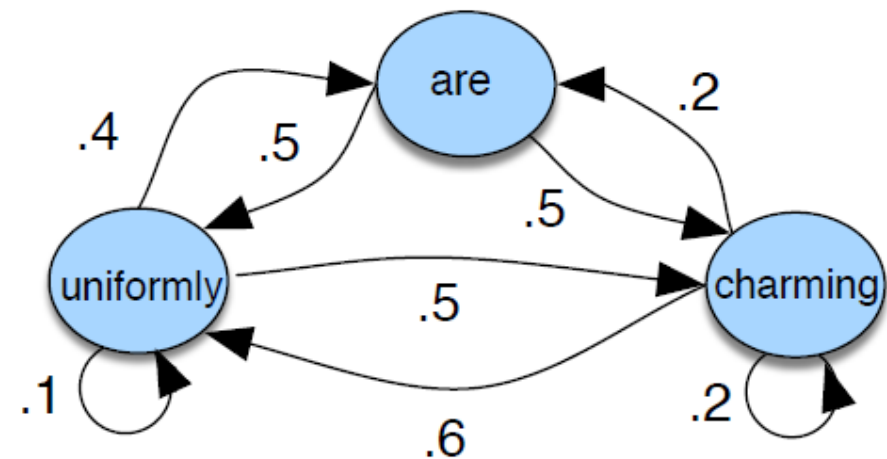
- Markov chain
 - Model of the probability of a next state given the current state
 - Very strong assumption >> future depends only on current state (not past)
 - State = set of variables (words)
- Sequence of state variables = q_1, q_2, \dots, q_i

Markov Assumption: $P(q_i = a | q_1 \dots q_{i-1}) = P(q_i = a | q_{i-1})$

- Prob of state q_i depends only on state q_{i-1}



(a)



(b)

Hidden Markov Model (HMM) POS tagger

- Markov chain
 - Model of the probability of a next state given the current state
 - Very strong assumption >> future depends only on current state (not past)
 - State = set of variables (words)
- Sequence of state variables = q_1, q_2, \dots, q_i

Markov Assumption: $P(q_i = a | q_1 \dots q_{i-1}) = P(q_i = a | q_{i-1})$

- **Markov Model** >> all events (words and tags) must be observed

$Q = q_1 q_2 \dots q_N$	a set of N states
$A = a_{11} a_{12} \dots a_{N1} \dots a_{NN}$	a transition probability matrix A , each a_{ij} representing the probability of moving from state i to state j , s.t. $\sum_{j=1}^n a_{ij} = 1 \quad \forall i$
$\pi = \pi_1, \pi_2, \dots, \pi_N$	an initial probability distribution over states. π_i is the probability that the Markov chain will start in state i . Some states j may have $\pi_j = 0$, meaning that they cannot be initial states. Also, $\sum_{i=1}^n \pi_i = 1$

Hidden Markov Model (HMM) POS tagger

- **Hidden Markov Model** >> POS tags are hidden states, which we must infer from observed words

Markov Assumption: $P(q_i|q_1, \dots, q_{i-1}) = P(q_i|q_{i-1})$

Output Independence: $P(o_i|q_1, \dots, q_i, \dots, q_T, o_1, \dots, o_i, \dots, o_T) = P(o_i|q_i)$

$Q = q_1 q_2 \dots q_N$	a set of N states
$A = a_{11} \dots a_{ij} \dots a_{NN}$	a transition probability matrix A , each a_{ij} representing the probability of moving from state i to state j , s.t. $\sum_{j=1}^N a_{ij} = 1 \quad \forall i$
$O = o_1 o_2 \dots o_T$	a sequence of T observations , each one drawn from a vocabulary $V = v_1, v_2, \dots, v_V$
$B = b_i(o_t)$	a sequence of observation likelihoods , also called emission probabilities , each expressing the probability of an observation o_t being generated from a state q_i
$\pi = \pi_1, \pi_2, \dots, \pi_N$	an initial probability distribution over states. π_i is the probability that the Markov chain will start in state i . Some states j may have $\pi_j = 0$, meaning that they cannot be initial states. Also, $\sum_{i=1}^n \pi_i = 1$

Hidden Markov Model (HMM) POS tagger



- Compute A matrix (tag transition probability) by counting
- WSJ corpus >> MD occurs 13124 times, and is followed by VB 10471 times

$$P(t_i|t_{i-1}) = \frac{C(t_{i-1}, t_i)}{C(t_{i-1})} \quad P(VB|MD) = \frac{C(MD, VB)}{C(MD)} = \frac{10471}{13124} = .80$$

- Compute B matrix (word observation probability) by counting
- WSJ corpus >> Of the 13124 occurrences of MD, it is associated with the word 'will' 4046 times

$$P(w_i|t_i) = \frac{C(t_i, w_i)}{C(t_i)} \quad P(will|MD) = \frac{C(MD, will)}{C(MD)} = \frac{4046}{13124} = .31$$

- If we were going to generate a MD tag, how likely is it that this tag will have the word 'will'?

Hidden Markov Model (HMM) POS tagger

- Decoding = Given a model and a sequence of observations ($w_1 \dots w_n$), find the most probable sequence of states ($t_1 \dots t_n$)

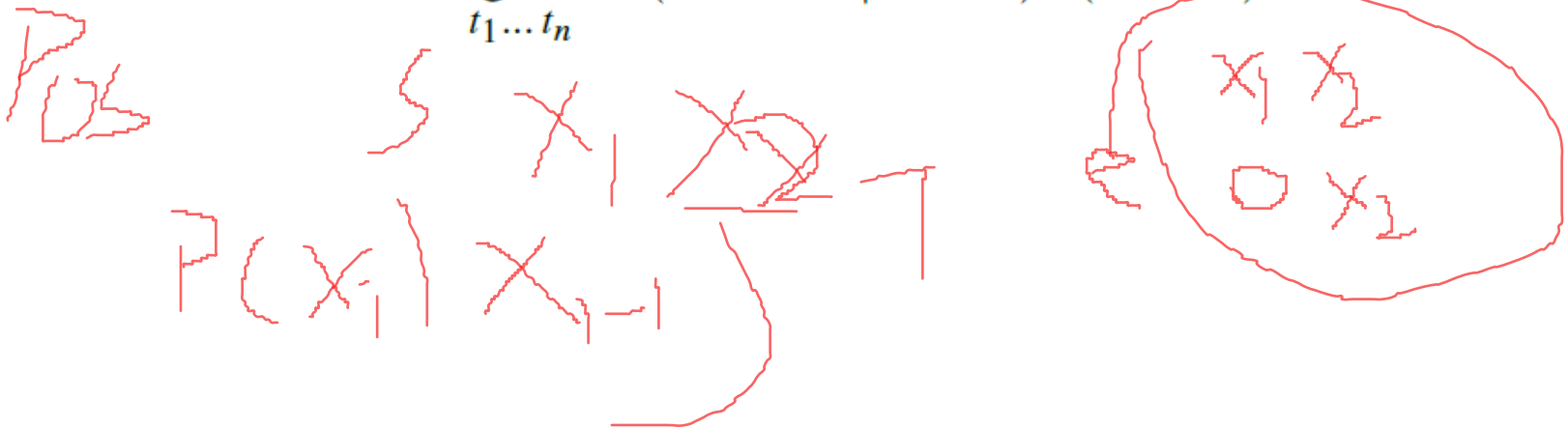
$$\hat{t}_{1:n} = \operatorname{argmax}_{t_1 \dots t_n} P(t_1 \dots t_n | w_1 \dots w_n)$$

- Apply Bayes' rule

$$P(x|y) = \frac{P(y|x)P(x)}{P(y)} \quad \hat{t}_{1:n} = \operatorname{argmax}_{t_1 \dots t_n} \frac{P(w_1 \dots w_n | t_1 \dots t_n) P(t_1 \dots t_n)}{P(w_1 \dots w_n)}$$

- Probability of seq of words does not vary by tag seq (so drop it)

$$\hat{t}_{1:n} = \operatorname{argmax}_{t_1 \dots t_n} P(w_1 \dots w_n | t_1 \dots t_n) P(t_1 \dots t_n)$$



Hidden Markov Model (HMM) POS tagger

- Decoding (from last slide)

$$\hat{t}_{1:n} = \operatorname{argmax}_{t_1 \dots t_n} P(w_1 \dots w_n | t_1 \dots t_n) P(t_1 \dots t_n)$$

- Assumptions

- Probability of word depends only on tag (independent on neighbours)
- Probability of tag depends only on previous tag (bigram)

$$P(w_1 \dots w_n | t_1 \dots t_n) \approx \prod_{i=1}^n P(w_i | t_i)$$

Matrix B

$$P(t_1 \dots t_n) \approx \prod_{i=1}^n P(t_i | t_{i-1})$$

Matrix A

- Decoding (using matrix A and B)

$$\hat{t}_{1:n} = \operatorname{argmax}_{t_1 \dots t_n} P(t_1 \dots t_n | w_1 \dots w_n) \approx \operatorname{argmax}_{t_1 \dots t_n} \prod_{i=1}^n \overbrace{P(w_i | t_i)}^{\text{emission}} \overbrace{P(t_i | t_{i-1})}^{\text{transition}}$$

Hidden Markov Model (HMM) POS tagger

- The Viterbi Algorithm is an efficient algorithm for HMM decoding using dynamic programming

```
function VITERBI(observations of len  $T$ , state-graph of len  $N$ ) returns best-path, path-prob

create a path probability matrix viterbi[ $N, T$ ]
for each state  $s$  from 1 to  $N$  do                                ; initialization step
     $viterbi[s, 1] \leftarrow \pi_s * b_s(o_1)$ 
     $backpointer[s, 1] \leftarrow 0$ 
for each time step  $t$  from 2 to  $T$  do                            ; recursion step
    for each state  $s$  from 1 to  $N$  do
         $viterbi[s, t] \leftarrow \max_{s'=1}^N viterbi[s', t-1] * a_{s', s} * b_s(o_t)$ 
         $backpointer[s, t] \leftarrow \operatorname{argmax}_{s'=1}^N viterbi[s', t-1] * a_{s', s} * b_s(o_t)$ 
     $bestpathprob \leftarrow \max_{s=1}^N viterbi[s, T]$                         ; termination step
     $bestpathpointer \leftarrow \operatorname{argmax}_{s=1}^N viterbi[s, T]$             ; termination step
     $bestpath \leftarrow$  the path starting at state  $bestpathpointer$ , that follows  $backpointer[]$  to states back in time
return bestpath, bestpathprob
```

- Exercise >> Run through worked example of the Viterbi algorithm from 3rd edition (online) chapter 8 of course text

Required Reading

- Sequence Labelling for Parts of Speech
 - Jurafsky and Martin, Speech and Language Processing, 3rd edition (online)
>> chapter 8

Questions

- Panopto Quiz - 1 minute brainstorm for interactive questions
Please write down in Panopto quiz in **1 minute** two or three questions that you would like to have answered at the next interactive session.

Do it **right now** while its fresh.

Take a screen shot of your questions and **bring them with you** at the interactive session so you have something to ask.