

ID:14311

NAME: Hidayat

Note: Attempt all Questions.

Question No. 1:

(15)

- a. Explain Part of Speech Tagging (POS) and explain POS tag ambiguity with two examples. (03)

POS Tagging:

It is a process of converting a sentence to forms – list of words, list of tuples (where each tuple is having a form (*word, tag*)). The tag in case of is a part-of-speech tag, and signifies whether the word is a noun, adjective, verb, and so on.

Default tagging: is a basic step for the part-of-speech tagging. It is performed using the Default Tagger class. The Default Tagger class takes 'tag' as a single argument. **NN** is the tag for a singular noun. Default Tagger is most useful when it gets to work with most common part-of-speech tag. That's why a noun tag is recommended.

POS tag ambiguity:

Common parts of speech in English are noun, verb, adjective, adverb, etc.

The POS tagging problem is **to determine the POS tag for a particular instance of a word**. The main problem with POS tagging is **ambiguity**.

For Example

POS Tag Ambiguity

In English : I bank₁ on the bank₂ on the river bank₃ for my transactions.

Bank₁ is verb, the other two banks are noun

b. State difference between open vs. closed classes. (02)

✚ **Open class (CONTENT/LEXICAL)**

- Lexical words deal with content and vocabulary.
- They have concrete meaning that goes beyond their function in a sentence.

✚ **Closed class (Grammatical/Function)**

- Grammatical words deal with the formation of sentences.
- They have ambiguous meaning and serve to express grammatical relationships with other words within a sentence.

c. Apply Viterbi Algorithm on the below given bigram and lexical probabilities; (10)

Initial Probabilities	
Noun	$\frac{1}{3}$
Verb	0
Other	$\frac{1}{3}$

Bigram Probabilities			
	Noun	Verb	Other
Noun	$\frac{1}{4}$	$\frac{1}{4}$	0
Verb	$\frac{1}{4}$	0	$\frac{1}{4}$
Other	$\frac{1}{3}$	0	$\frac{1}{3}$

Lexical Probabilities					
	O1=time	O2=flies	O3=like	O4=an	O5=arrow
Noun	$\frac{1}{5}$	$\frac{1}{5}$	0	0	$\frac{1}{5}$
Verb	$\frac{1}{5}$	$\frac{2}{5}$	$\frac{1}{5}$	0	0
Other	0	0	$\frac{1}{5}$	$\frac{2}{5}$	0

Hi	P1(h1)	P2(h2)	P3(h3)	P4(h4)	P5(h5)
Noun	$1/3 \times 1/5 = 1/15$	$1/5 \times 1/4 \times 1/15 = 1/300$	$0 \times 1/150 = 0$	0	$1/5 \times 1/4 \times 1/16875$
Verb	0	$2/5 \times 1/4 \times 1/15 = 1/150$	$1/5 \times 1/4 \times 1/150 = 1/3000$	0	0
other	0	0	$1/5 \times 1/3 \times 1/150 = 1/2250$	$1/3 \times 2/5 \times 1/2250 = 1/16875$	0

Question No. 2:

(05)

Apply Bayesian theorem over the below given string:

^John got many NLP books. ^He found them all very interesting.

Where for lexical probabilities assume John=0.5, got=0.3, many=0.2, NLP=0.1 and books=0.

POS Tags:

^N V A N N. ^N V N A R A.

Recording Numbers:

	^	N	V	A	R	.
^	0	2	0	0	0	0
N	0	1	2	1	0	1
V	0	1	0	1	0	0
A	0	1	0	0	1	1
R	0	0	0	1	0	0
.	1	0	0	0	0	0

Bigram Probability:

	^	N	V	A	R	.
^	0	1	0	0	0	0
N	0	1/5	2/5	1/5	0	1/5

V	0	1/2	0	1/2	0	0
A	0	1/3	0	0	1/3	1/3
R	0	0	0	1	0	0
.	1	0	0	0	0	0

Where **lexical probability** is given i.e.

John=0.5, got=0.3, many=0.2, NLP=0.1 and books=0.

Good Luck ☺