

ELL881 : Quiz 1 (Solutions)

February 2, 2023

Question 1:

Word	Frequency
try	3
do	2
not	2
there	1
is	1
no	1
or	1
you	1
must	1

$$P_{GT}^*(anakin) = N_1/N = 6/13$$

$$\begin{aligned} P_{GT}^*(not) &= \frac{c^*}{N} \\ &= \frac{(3 * N_3/N_2)}{N} \\ &= \frac{3 * 1}{2 * 13} \\ &= \frac{3}{26} \end{aligned}$$

Question 2:

Refer to the algorithm in the lecture slides. We are constructing a DP table of size $N * T$ and performing nested iterations.

Time Complexity: $O(N^2T)$

Space Complexity: $O(NT)$

Question 3:

(**From slides:**) Binarization is crucial for the cubic time CFG parsing. Without binarization, you don't get parsing cubic in the length of the sentence and in the number of nonterminals in the grammar.

Question 4:

Method	Task Type	Model Type	Training
Logistic Regression	Classification	Discriminative	MLE/SGD (both accepted)
Naive Bayes	Classification	Generative	MLE
HMMs	Tagging	Generative	MLE

Question 5:

1. $[!@%*]$
2. $[/ d/]$ and $[/[A-Za-z]/]$
3. $\wedge(!.*skywalker.*)\$ \mid \wedge(. *skywalker.* @)\$$