

Maria Frey
LING492B
2 April 2021

Lab 4 Questions

Question 1:

The probability of the ‘time flies quickly’ is 0.00054. The number of non-zero ways of tagging this sequence is the number of words to the power of the number of tags, so 3^3 , which is 27.

Question 2:

The best tag sequence for the word sequence is: [#, N, V, R, #]. This means that ‘time’ is a noun that is the subject of the verb ‘flies’ and ‘quickly’ is an adverb modifying the verb. So, the meaning of the sentence is that time, an object, flies in a quick manner.

Question 3:

The probability of ‘quickly time flies’ is 0.00024. The parameters that would need to be changed to make ‘quickly time flies’ more likely than ‘time flies quickly’ would be the transition probabilities. In the current transition probabilities, an adverb is not very likely given ‘#’ and a noun is not likely given an adverb. Therefore, increasing $P(R\#)$ and $P(N|R)$ would make ‘quickly time flies’ more likely than ‘time flies quickly’.

Question 4:

In addition to adding *swat* to every state’s emission, the changes that I made were to all of the emission probabilities given V and N (highlighted below):

```
B2 = {'N':{'swat':0.0, 'time':0.4, 'flies':0.59, 'quickly':0.005, '#':0.0}, 'V':{'swat':0.52, 'time':0.15, 'flies':0.3, 'quickly':0.03, '#':0.0}, 'R':{'swat':0.0, 'time':0.01, 'flies':0.01, 'quickly':0.98, '#':0.0}, '#':{'swat':0.0, 'time':0.0, 'flies':0.0, 'quickly':0.0, '#':1.0}}  
['#', 'V', 'N', 'R', '#']
```

Question 5:

The probability of ‘swat flies quickly’ being a VNR tag sequence is higher than ‘time flies quickly’ being a VNR because the highest emission probability for ‘time’ is an N not a V. So, it comes down to the emission probability of (‘time’|V) and (‘swat’|V) and the probability of time being a verb is much lower than the probability of *swat* being a verb. The tag sequence generated for ‘flies time’ was actually NN not NV because the transition probability (N|N) is pretty high and the probabilities (‘time’|N) and (‘flies’|N) are also high, so the highest tag sequence is NN. So, ‘flies *swat*’ as NV is more likely than ‘flies time’.