1. **What is probability of ‘time flies quickly’? How many tag sequences are contributing to this probability (that is, how many non-zero ways of tagging ‘time flies quickly’ are there according to this HMM)?**
   Probability of ‘time flies quickly’: 0.00072656460425
   There are 27 tag sequences contributing to this probability according to this HMM. This number is all possible combinations of # X X X #, where X can be N, V, or R, or 3 to the power of 3. This is because each tag (other than #) can transition to all other tags (other than #).
2. **What is the best tag sequence for ‘time flies quickly’? What meaning does this tag sequence correspond to?**
   Best tag sequence: # N V R #, meaning this sentence is most likely to have noun, verb, adverb structure (‘time’ is most likely to be a noun, ‘flies’ is most likely to be a verb, and ‘quickly' is most likely to be an adverb in this sequence).
3. **What’s the probability of ‘quickly time flies’? What parameters would you need to change to make ‘quickly time flies’ more likely than ‘time flies quickly’ (don’t actually change the parameters in your code)? Be specific!**
   Probability of ‘quickly time flies’: 0.0006319457220000002
   We would have to change the transition probabilities in our dictionary to make this sequence more likely than ‘time flies quickly.’ Specifically, we would have to increase the probability of the following transitions to be higher than their counterparts:
   - # transitioning to R (to make the probability of starting with adverb highest)
   - R transitioning to N (to make theprobability of noun following adverb highest)
4. **Modify matrix B to add a new word ‘swat’ to the lexicon. Assume that ‘swat’ can only be a verb, and be sure your matrix defines proper conditional probability distributions. Set your probabilities so that the most likely tag sequence for ‘swat flies quickly’ is ‘V N R’. (hint: you will need to change probabilities in multiple rows). Show this matrix in your question write-up, and also encode it into the script as a dictionary ‘B2’. Once you define B2 at the top of your code, you can uncomment the third print statement in the main code to tag ‘swat flies quickly’ and make sure you have defined B2 correctly.**
   B2:

	time	flies	quickly	swat	#
N	0.70	0.25	0.05	0	0
V	0.28	0.30	0.02	0.10	0
R	0.005	0.005	0.99	0	0
#	0	0	0	0	1

5. **In your new HMM from Q4, ‘V N R’ is the most likely tag sequence for ‘swat flies quickly’. How is this probability different from the probability of the same tag sequence for ‘time flies quickly’ (i.e. what parameters are relevant to this comparison)? Which is higher? Explain why the answer to this question**

determines the answer to the following question: - Which is higher P(‘N V’, ‘flies time’) or P(‘N V’, ‘flies swat’)

The probability of ‘V N R’ for ‘swat flies quickly’ is higher than the probability of ‘V N R’ for ‘time flies quickly. This is because ‘time’ is more likely to emit ‘N’ (0.70) rather than ‘V’ (0.28) in our B2 matrix, while ‘swat’ is most likely to emit ‘V’ in our B2 matrix. Using the same parameters of emission probabilities, we can determine that P(‘N V’, ‘flies swat’) is higher than P(‘N V’, ‘flies time’). Again, this is due to the fact that ‘swat’ is most likely to emit ‘V’ in our B2 matrix while ‘time’ is less likely to emit ‘V’ as opposed to ‘N’ in our B2 matrix.