1 Working with HMMs (20 points)

You have taken an interest in a new company call *facebrochure*. You have the following model of the company: If it is profitable during any particular quarter, it has a 80% chance of remaining profitable in the subsequent quarter, and if it is *not* profitable, there is a 60% chance of it remaining unprofitable in the next quarter. The company isn’t yet listed on the stock exchange, so they are not required to post quarterly earnings (no ground truth information about the state). However, you believe that if the company is profitable, there is 60% chance that a favorable article will appear in the newspaper at the end of the quarter. If the company is not profitable, there is only a 40% chance of a favorable article.

Assume that the company is known to be profitable at quarter 0, and that two favorable articles have appeared in each of the two subsequent quarters.

For the following parts to this question, you should show all of your work.

a) Compute the probability of each path through the state space. (5 points)
b) Use the forward algorithm to compute the distribution over states at time 2 and verify that this is consistent with your answer to the previous part. (5 points)
c) Use the backward algorithm and your answer to the previous part to compute the smoothed distribution over states at time 1. Verify that this is consistent with your computation of the path probabilities. (5 points)
d) Use the Viterbi algorithm to compute the highest probability path and verify that this is consistent with your answer to the first part of the this question. (5 points)
2 Extreme Effects of Smoothing (10 points)

Consider an HMM with two states with $P(s|s) = 1.0 = P(\overline{s}|\overline{s}) = 1.0$ (time indices are implicit), and two observations, $O \in \{o, \overline{o}\}$. Assume a known initial state distribution at time step 0 of $[0.5, 0.5]$. Provide the following:

1. The model observation probabilities: $P(O|S)$
2. Observations at times steps 1 and 2: $O_1, O_2$

To guarantee the following (plugging in your chosen values for $O_1$ and $O_2$):

1. $P(S_1 = s|O_1) > 0$
2. $P(S_1 = s|O_1, O_2) = 0$

Show the calculations to support your claim that both conditions are satisfied.
3 Counting Paths (10 points)

Consider an HMM with $n$ states and $m$ observations. Assume that for all $i, j$, $P(s_i|s_j) > 0$ (time indices are implicit), and that for all $i, k$, $P(o_k|s_i) > 0$.

1. For a known initial distribution and $T$ observations, how many possible paths through the state space are there? (5 points)

2. Given the same assumptions in the previous part, derive an expression for the number of paths that pass through a state at a particular time step $t$. Is this the same for all states and all time steps? (5 points)

For each of above questions, be sure to provide a clear mathematical justification for your answer.


4 Counting Paths by Dynamic Programming (20 points)

For this question, we are interested in counting the number of paths through the state space with non-zero probability. As in the previous question, assume $n$ states, $m$ observations, and $T$ time steps, but unlike the previous question, it may be the case that $P(s_i | s_j) = 0$ for one or more $i, j$, and it may be the case that $P(o_k | s_i) = 0$ for one or more $i, k$.

Drawing upon concepts from the Viterbi algorithm and/or the Forward-Backward algorithm, provide pseudocode for an algorithm that takes as input:

- An initial state distribution over states at time $t = 0$,
- A sequence of observations $O_1 \ldots O_T$,

and counts the number of paths with non-zero probability that pass through any state at any time. You should also provide a brief, English explanation of how your algorithm differs from the Viterbi algorithm and/or the Forward-Backward algorithm, and an explanation of why it is computing the right thing.