9318 Project Report

Name: ZhiHao LI
Peng LIU
zID: z5144901
z5178833

Q1 viterbi algorithm

1. Initialize 2 [N x T] matrices, delta and record to mark down the probabilities and paths.

2. When we fixed the time stamp and the next state, loop over all the previous states to compute each probability with certain time stamp and next state.

tempProbability = delta[sp][ts-1] ¥times transition[sp][sn] ¥times emission[sn][Obeserve[ts]]

3. Then we record the probability and previous state in delta and record. If there is a larger probability, then we update both values of delta and records.

4. When we finish the iteration, we can get max probability in the last time stamp, which means we can know the final state with the max probability.

5. Hence, by using this state value, we can backtrack the previous state values based on the record matrix.

Q2 top k viterbi algorithm

1. Similar to Q1, we initialize 2 [T x N x K] matrices, delta and record to mark down all k probabilities and paths. However, if we strictly follow Q1's solution, we will miss some paths which cross a state more than once.

2. Hence, we make a copy of record matrix into the rank matrix, which will mark down if a state is crossed by a path more than once. For example, in toy example, "red yellow blue green" have [3, 0, 0, 1, 2, 4] and [3, 0, 0, 0, 2, 4], which state 2 has been crossed by twice, hence the rank in state 2 is 0 and 1, which will lead the algorithm to compute different path.

3. Same as Q1, we fixed next state and time stamp to iterate all previous states in k times, which we will get N x K probabilities. Then we compute the most large K probabilities and get their state value and rank value.

4. Now we have the max probabilities state and rank, we can backtrack the previous states just like Q1. Finally we return the wanted list and finish Q2.

Q3 advanced decoding

Utilizing Additive smoothing. Let's change smooth = 1 to be smooth = δ (0< δ < 1).

the formula is : $P_{Add-1(w_i|w_{i-1})} = \frac{c(w_{i-1}w_i)+\delta}{c(w_{i-1})+\delta V}$