ECE368: Probabilistic Reasoning

Lab 2 – Part II: Hidden Markov Model

Name:	Seunghyun Cho	**Student Number:**	1001681325

You should hand in: 1) A scanned .pdf version of this sheet with your answers (file size should be under 2 MB); 2) one Python file inference.py that contains your code. The files should be uploaded to Quercus.

1. (a) Write down the formulas of the forward-backward algorithm to compute the marginal distribution $p(\mathbf{z}_i|(\hat{x}_0, \hat{y}_0), \ldots, (\hat{x}_{N-1}, \hat{y}_{N-1}))$ for $i = 0, 1, \ldots, N-1$. Your answer should contain the initializations of the forward and backward messages, the recursion relations of the messages, and the computation of the marginal distribution based on the messages. (1 **pt**)

$$\alpha(z_n) = P(X_n|z_n) \sum_{z_{n-1}} \alpha(z_{n-1}) P(z_n|z_{n-1})$$

$$\beta(z_n) = \sum_{z_{n+1}} \beta(z_{n+1}) P(X_{n+1}|z_{n+1}) P(z_{n+1}|z_n)$$

$$P(z_n|X) = \frac{\alpha(z_n)\beta(z_n)}{\sum_{z_n} \alpha(z_n)\beta(z_n)}$$

These are normalized after computed

Initialize $\alpha_{init} = \text{prior\_distribution}(z)$ $\begin{cases} 1 & \text{if } z.\text{action} = \text{'stay'} \\ 0 & \text{else} \end{cases}$

$\beta_{init} = 1$

(b) After you run the forward-backward algorithm on the data in test.txt, write down the obtained marginal distribution of the state at $i = 99$ (the last time step), i.e., $p(\mathbf{z}_{99}|(\hat{x}_0, \hat{y}_0), \ldots, (\hat{x}_{99}, \hat{y}_{99}))$. Only include states with non-zero probability in your answer. (2 **pt**)

(11, 0, 'stay') 0.81026
(11, 0, 'right') 0.1796
(10, 1, 'down') 0.010128

2. Modify your forward-backward algorithm so that it can handle missing observations. After you run the modified forward-backward algorithm on the data in test_missing.txt, write down the obtained marginal distribution of the state at $i = 30$, i.e., $p(\mathbf{z}_{30}|(\hat{x}_0, \hat{y}_0), \ldots, (\hat{x}_{99}, \hat{y}_{99}))$. Only include states with non-zero probability in your answer. (1 **pt**)

(6, 7, 'right') 0.91304
(5, 7, 'stay') 0.043478
(5, 7, 'right') 0.043478

3. (a) Write down the formulas of the Viterbi algorithm using $\mathbf{z}_i$ and $(\hat{x}_i, \hat{y}_i), i = 0, 1, \ldots, N-1$. Your answer should contain the initialization of the messages and the recursion of the messages in the Viterbi algorithm. (1 **pt**)

$$W_1(Z_1) = \ln P(Z_1) + \ln P(X_1 | Z_1)$$

$$W_n(Z_n) = \ln P(X_n | Z_n) + \max_{Z_{n-1}} \{ \ln P(Z_n | Z_{n-1}) + W_{n-1}(Z_{n-1}) \}$$

$$Z_n^* = \arg\max_{Z_n} W_N(Z_N)$$

(b) After you run the Viterbi algorithm on the data in test_missing.txt, write down the last 10 hidden states of the most likely sequence (i.e., $i = 90, 91, 92, \ldots, 99$) based on the MAP estimate. (3 **pt**)

i	x	y	action
90	11	5	down
91	11	6	down
92	11	7	down
93	11	7	stay
94	11	7	stay
95	10	7	left
96	9	7	left
97	8	7	left
98	7	7	left

continued

99	6	7	left

4. Compute and compare the error probabilities of $\{\tilde{\mathbf{z}}_i\}$ and $\{\check{\mathbf{z}}_i\}$ using the data in test_missing.txt. The error probability of $\{\tilde{\mathbf{z}}_i\}$ is $\boxed{3\%}$. The error probability of $\{\check{\mathbf{z}}_i\}$ is $\boxed{2\%}$. (1 **pt**)

5. Is sequence $\{\check{\mathbf{z}}_i\}$ a valid sequence? If not, please find a small segment $\check{\mathbf{z}}_i, \check{\mathbf{z}}_{i+1}$ that violates the transition model for some time step i. You answer should specify the value of i as well as the corresponding states $\check{\mathbf{z}}_i, \check{\mathbf{z}}_{i+1}$. (1 **pt**)

No,

i	(x,y)	action
64	(3,7)	stay
65	(2,7)	stay

Violates