Collab Worksheet 7

CS440/ECE448, Spring 2021

Week of 3/26 - 3/31, 2021

Question 1

Ron the Con is a skilled con artist who works Green Street. After showing the audience a fair coin, Ron will either continue to use the fair coin (with probability $1 - P$) or secretly exchange it for an unfair coin (with probability P). If he flips the fair coin, it comes up heads with probability 50%. If he flips the unfair coin, it comes up heads with probability 100%. After every flip, he either continues to use the same coin (with probability $1 - P$), or he secretly switches coins (with probability P).

(a) (1 point) This process can be represented as a hidden Markov model, where the state variable is $Y_t = 1$ for the unfair coin, $Y_t = 0$ for the fair coin, and the evidence variable is $X_t = 1$ for heads, $X_t = 0$ for tails. Write the conditional probability table $p(X_t|Y_t)$.

Solution:

	$P(X_t\|Y_t)$	
Y_t	0	1
0	1/2	1/2
1	0	1

(b) (2 points) For this part of the problem only, assume that after every flip, Ron either continues to use the same coin (with probability $1-P$), or he secretly switches coins (with probability P). Write the conditional probability table $p(Y_t|Y_{t-1})$.

Solution:

	$P(Y_t\|Y_{t-1})$	
$Y_yt - 1$	0	1
0	$1 - P$	P
1	P	$1 - P$

(c) (2 points) For this part of the problem only, assume that after every flip, Ron NEVER changes the coin. In other words, with probability P, he chooses the unfair coin on the first flip, and then continues to use the unfair coin for every flip thereafter; with probability $1 - P$, he chooses the fair coin on the first flip, then he continues to use the fair coin for every flip thereafter. In terms of P, how many heads in a row would you need to see before you conclude that the coin is unfair?

Solution: We decide "unfair" if $P > (1-P)(0.5)^n$, thus

$$\ln(P) > \ln(1-P) + n\ln(0.5)$$
$$\ln(P) - \ln(1-P) > -n\ln(2)$$
$$-\frac{\ln(P) - \ln(1-P)}{\ln(2)} < n$$

Question 2

Consider a hidden Markov model (HMM) whose hidden variable denotes part of speech (POS), $Y_t \in \{N, V\}$ where N =noun, V =verb, the initial state probability is $P(Y_1 = N) = 0.8$, and the transition probabilities are $P(Y_t = N|Y_{t-1} = N) = 0.1$ and $P(Y_t = V|Y_{t-1} = V) = 0.1$. Suppose we have the observation probability matrix given in the following table:

X_t	rose	bill	likes
$P(X_t\|Y_t = N)$	0.4	0.4	0.2
$P(X_t\|Y_t = V)$	0.2	0.2	0.6

You are given the sentence "bill rose." You want to figure out whether each of these two words, "bill" and "rose", is being used as a noun or a verb.

(a) List the four possible combinations of (Y_1, Y_2). For each possible combination, give $P(Y_1, X_1 = \text{bill}, Y_2, X_2 = \text{rose})$. Give your results as mathematical expressions in terms of constants; do not simplify.

Solution:

$P(Y_1, X_1 = \text{bill}, Y_2, X_2 = \text{rose})$	$Y_2 = N$	$Y_2 = V$
$Y_1 = N$	(0.8)(0.4)(0.1)(0.4)	(0.8)(0.4)(0.9)(0.2)
$Y_1 = V$	(0.2)(0.2)(0.9)(0.4)	(0.2)(0.2)(0.1)(0.2)

(b) Let $\mathcal{D}$ be the event $(X_1 = \text{bill}, X_2 = \text{rose})$. Find $P(Y_2 = V|\mathcal{D})$.

Solution:

$$\begin{aligned} P(\mathcal{D}, Y_2 = V) &= P(Y_1 = N, Y_2 = V, \mathcal{D}) + P(Y_1 = V, Y_2 = V, \mathcal{D}) \\ &= (0.8)(0.4)(0.9)(0.2) + (0.2)(0.2)(0.1)(0.2) \\ P(\mathcal{D}, Y_2 = N) &= P(Y_1 = N, Y_2 = N, \mathcal{D}) + P(Y_1 = V, Y_2 = N, \mathcal{D}) \\ &= (0.8)(0.4)(0.1)(0.4) + (0.2)(0.2)(0.9)(0.4) \end{aligned}$$

Dividing the first row by the sum of the two rows, we get

$$P(Y_2 = V|\mathcal{D}) = \frac{(0.8)(0.4)(0.9)(0.2) + (0.2)(0.2)(0.1)(0.2)}{(0.8)(0.4)(0.9)(0.2) + (0.2)(0.2)(0.1)(0.2) + (0.8)(0.4)(0.1)(0.4) + (0.2)(0.2)(0.9)(0.4)}$$

(c) Use the Viterbi algorithm to find the most likely state sequence for this sentence.

Solution:

- To find the backpointer from $Y_2 = N$, we find the maximum among the two possibilities $P(Y_1 = N, Y_2 = N, \mathcal{D})$ and $P(Y_1 = V, Y_2 = N, \mathcal{D})$. The larger of the two is $P(Y_1 = V, Y_2 = N, \mathcal{D}) = (0.2)(0.2)(0.9)(0.4)$, so the backpointer from $Y_2 = N$ points to $Y_1 = V$.

- To find the backpointer from $Y_2 = V$, we find the maximum among the two possibilities $P(Y_1 = N, Y_2 = V, \mathcal{D})$ and $P(Y_1 = V, Y_2 = V, \mathcal{D})$. The larger of the two is $P(Y_1 = N, Y_2 = V, \mathcal{D}) = (0.8)(0.4)(0.9)(0.2)$, so the backpointer from $Y_2 = V$ points to $Y_1 = N$.

- To find the best terminal state, then, we find the maximum among the two possibilities $P(Y_1 = V, Y_2 = N, \mathcal{D})$ and $P(Y_1 = N, Y_2 = V, \mathcal{D})$. The larger of the two is $P(Y_1 = N, Y_2 = V, \mathcal{D}) = (0.8)(0.4)(0.9)(0.2)$, so the maximum likelihood state sequence is $(Y_1, Y_2) = (N, V)$.