CSE 5525 Artificial Intelligence II

Quiz #6: Hidden Markov Models

Wei Xu, Ohio State University

Your Name: ______________________ OSU Username: _______________

At time t, Peter is in some state X_t. The two states Peter alternates between are saving the world (denoted as S) and being a CSE student (denoted as C). Let the evidence E_t be whether or not Peter is seen in the CSE labs at time t.

The transition probabilities are provided in the following table (left), where the row corresponds to X_{t-1} and the column to X_t. For example, $P(X_t = S|X_{t-1} = C) = 0.4$.

X_{t-1}	X_t	$P(X_t\|X_{t-1})$
C	C	0.6
C	S	0.4
S	C	0.2
S	S	0.8

X_t	E_t	$P(E_t\|X_t)$
C	true	0.7
C	false	0.3
S	true	0.1
S	false	0.9

The model for evidence E_t is provided in the following table, where the row corresponds to X_t and the column to E_t. For example, $P(E_t = false|X_t = S) = 0.9$.

Questions:

1) Assume we have current belief $B(X_t) = P(X_t|e_{i:t})$, how to compute for passage of time?

$$B'(X_{t+1}) = P(X_{t+1}|e_{1:t}) = \sum_{X_t=\{S,C\}} P(X_{t+1}|X_t)\underbrace{P(X_t : e_{1:t})}_{B(X_t)}$$

How to update the belief with the observation of evidence (forward algorithm)?

$$B(X_{t+1}) = P(X_{t+1}|e_{1:t+1}) \propto_{X_{t+1}} B'(X_{t+1})P(e_{t+1}|X_{t+1})$$

2) Let the initial beliefs be $P(X_0 = S) = P(X_0 = C) = 0.5$. Fill in the following table for $t = 1$, first computing for passage of time, and then for observation of evidence $E_1 = true$. Finally, normalize the values from the observation column to get the beliefs. Round to three decimal places.

$P(X_1|X_0=S)P(X_0=S|e_0) + P(X_1|X_0=C)P(X_0=C|e_0)$ $B'(X_1)P(E_1=true|X_1)$

X_1	passage of time $B'(X_1)$	update with evidence $E_1 = true$	$B(X_1)$
S	$0.8\times0.5+0.4\times0.5=0.6$	$0.6\times0.1=0.06$	$\frac{0.06}{0.06+0.28}=0.176$
C	$0.2\times0.5+0.6\times0.5=0.4$	$0.4\times0.7=0.28$	$\frac{0.28}{0.06+0.28}=0.824$

3) Repeat for $t = 2$, with the observation of evidence $E_2 = false$. When using any previous value for computations, use their rounded value. Round to three decimal places.

$P(X_2|X_1=S)B(X_1=S) + P(X_2|X_1=C)B(X_1=C)$

X_2	passage of time $B'(X_2)$	update with evidence $E_2 = false$	$B(X_2)$
S	$0.8\times0.176+0.4\times0.824=0.470$	$0.470\times0.9=0.423$	$\frac{0.423}{0.423+0.159}=0.727$
C	$0.2\times0.176+0.6\times0.824=0.530$	$0.530\times0.3=0.159$	$\frac{0.159}{0.423+0.159}=0.273$

4) Assume now we are using a particle filter with 3 particles to apprximate our belief instead of using exact inference as in 1 – 3). Image we have just applied transition model sampling (passage of time) from state X_0 to X_1, and now have the set of particles S, S, C. What is our belief about X_1 before considering noisy evidence?

X_1	passage of time $B'(X_1)$
S	$\frac{2}{3}$
C	$\frac{1}{3}$

5) Now assume we receive evidence $E_1 = true$. What is the weight for each particle, and what is our belief now about X_1 (before weighted re-sampling)?

particle	weight
S	0.1
S	0.1
C	0.7

X_1	after observation $B(X_1)$
S	$\frac{0.1+0.1}{0.1+0.1+0.7}=0.222$
C	$\frac{0.7}{0.1+0.1+0.7}=0.778$

6) Will performing weighted re-sampling on these weighted particles to obtain our three new particle representation for X_1 cause our belief to change?

Yes, sampling is an approximation and three unweighted particles will not be able to represent 0.222, 0.778 (i.e. only can do 0, 1 / 0.333, 0.667 / 0.667, 0.333 / 1, 0)

7) Recall from the lecture that $m_t[x_t] = max_{x_{1:t-1}} P(x_{1:t-1}, x_t, e_{1:t})$, which is the probability of the most likely path that ends at x_t, considering the path up to t and the evidence up to t. Use **Viterbi algorithm** to compute $m_1[S_1]$, $m_1[C_1]$, $m_2[S_2]$, $m_2[C_2]$ for the sequence of evidence $E_1 = true$, $E_2 = false$. Define $m_0[S] = m_0[C] = 0.5$. Use exact numbers in your calculations and answers. What was Peter most likely doing at time $t = 1$ and at time $t = 2$?

$$M_t[X_t] = \max_{X_{1:t-1}} P(X_{1:t-1}, X_t, e_{1:t})$$

$$= P(e_t|X_t) \cdot \max_{X_{t-1}} P(X_t|X_{t-1})\, M_{t-1}[X_{t-1}]$$

$$m_1[S_1] = P(E_1=\text{true}|X_1=S) \cdot \max_{X_{t-1}} \begin{cases} P(X_1=S|X_0=S)\, M_0[X_0=S] \\ P(X_1=S|X_0=C)\, M_0[X_0=C] \end{cases}$$

$$= 0.1 \times \max(0.8 \times 0.5,\ 0.4 \times 0.5)$$

$$= 0.04$$

$$m_1[C_1] = P(E_1=\text{true}|X_1=C) \cdot \max \begin{cases} P(X_1=C|X_0=S)\, M_0[S] \\ P(X_1=C|X_0=C)\, M_0[C] \end{cases}$$

$$= 0.7 \times \max(0.2 \times 0.5,\ 0.6 \times 0.5)$$

$$= 0.21$$