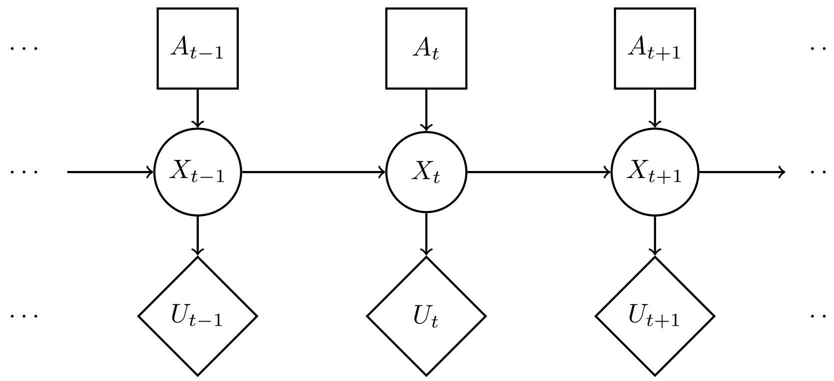
Exam Prep 10 Solutions

Q1. Planning ahead with HMMs

Pacman is tired of using HMMs to estimate the location of ghosts. He wants to use HMMs to plan what actions to take in order to maximize his utility. Pacman uses the HMM (drawn to the right) of length T to model the planning problem. In the HMM, $X_{1:T}$ is the sequence of hidden states of Pacman's world, $A_{1:T}$ are actions Pacman can take, and U_t is the utility Pacman receives at the particular hidden state X_t. Notice that there are no evidence variables, and utilities are not discounted.



(a) The belief at time t is defined as $B_t(X_t) = p(X_t|a_{1:t})$. The forward algorithm update has the following form:

$$B_t(X_t) = \underline{\quad\textbf{(i)}\quad}\ \underline{\quad\textbf{(ii)}\quad}\ B_{t-1}(x_{t-1}).$$

Complete the expression by choosing the option that fills in each blank.

(i)	○ $\max_{x_{t-1}}$	● $\sum_{x_{t-1}}$	○ $\max_{x_t}$	○ $\sum_{x_t}$	○ 1
(ii)	○ $p(X_t\|x_{t-1})$	○ $p(X_t\|x_{t-1})p(X_t\|a_t)$	○ $p(X_t)$	● $p(X_t\|x_{t-1}, a_t)$	○ 1

○ None of the above combinations is correct

$$\begin{aligned} B_t(X_t) &= p(X_t|a_{1:t}) \\ &= \sum_{x_{t-1}} p(X_t|x_{t-1}, a_t)p(x_{t-1}|a_{1:t-1}) \\ &= \sum_{x_{t-1}} p(X_t|x_{t-1}, a_t)B_{t-1}(x_{t-1}) \end{aligned}$$

(b) Pacman would like to take actions $A_{1:T}$ that maximizes the expected sum of utilities, which has the following form:

$$\text{MEU}_{1:T} = \underline{\ \textbf{(i)}\ }\ \underline{\ \textbf{(ii)}\ }\ \underline{\ \textbf{(iii)}\ }\ \underline{\ \textbf{(iv)}\ }\ \underline{\ \textbf{(v)}\ }$$

Complete the expression by choosing the option that fills in each blank.

(i)	● $\max_{a_{1:T}}$	○ $\max_{a_T}$	○ $\sum_{a_{1:T}}$	○ $\sum_{a_T}$	○ 1
(ii)	○ $\max_t$	○ $\prod_{t=1}^{T}$	● $\sum_{t=1}^{T}$	○ $\min_t$	○ 1
(iii)	○ $\sum_{x_t, a_t}$	● $\sum_{x_t}$	○ $\sum_{a_t}$	○ $\sum_{x_T}$	○ 1
(iv)	○ $p(x_t\|x_{t-1}, a_t)$	○ $p(x_t)$	● $B_t(x_t)$	○ $B_T(x_T)$	○ 1
(v)	○ U_T	○ $\frac{1}{U_t}$	○ $\frac{1}{U_T}$	● U_t	○ 1

○ None of the above combinations is correct

$$\text{MEU}_{1:T} = \max_{a_{1:T}} \sum_{t=1}^{T} \sum_{x_t} B_t(x_t) U_t(x_t)$$

(c) A greedy ghost now offers to tell Pacman the values of some of the hidden states. Pacman needs your help to figure out if the ghost's information is useful. Assume that the transition function $p(x_t|x_{t-1}, a_t)$ is not deterministic. **With respect to the utility** U_t, mark all that can be True:

■ $\text{VPI}(X_{t-1}|X_{t-2}) > 0$ ☐ $\text{VPI}(X_{t-2}|X_{t-1}) > 0$ ■ $\text{VPI}(X_{t-1}|X_{t-2}) = 0$ ■ $\text{VPI}(X_{t-2}|X_{t-1}) = 0$ ☐ None of the above

It is always possible that VPI $= 0$. Can guarantee $\text{VPI}(E|e)$ is not greater than 0 if E is independent of parents(U) given e.

(d) Pacman notices that calculating the beliefs under this model is very slow using exact inference. He therefore decides to try out various particle filter methods to speed up inference. Order the following methods by how accurate their estimate of $B_T(X_T)$ is? If different methods give an equivalently accurate estimate, mark them as the same number.

	Most accurate			Least accurate
Exact inference	● 1	○ 2	○ 3	○ 4
Particle filtering with no resampling	○ 1	● 2	○ 3	○ 4
Particle filtering with resampling before every time elapse	○ 1	○ 2	○ 3	● 4
Particle filtering with resampling before every other time elapse	○ 1	○ 2	● 3	○ 4

Exact inference will always be more accurate than using a particle filter. When comparing the particle filter resampling approaches, notice that because there are no observations, each particle will have weight 1. Therefore resampling when particle weights are 1 could lead to particles being lost and hence prove bad.

Q2. Naive Bayes: Pacman or Ghost?

You are standing by an exit as either Pacmen or ghosts come out of it. Every time someone comes out, you get two observations: a visual one and an auditory one, denoted by the random variables X_v and X_a, respectively. The visual observation informs you that the individual is either a Pacman ($X_v = 1$) or a ghost ($X_v = 0$). The auditory observation X_a is defined analogously. Your observations are a noisy measurement of the individual's true type, which is denoted by Y. After the indiviual comes out, you find out what they really are: either a Pacman ($Y = 1$) or a ghost ($Y = 0$). You have logged your observations and the true types of the first 20 individuals:

individual i	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
first observation $X_v^{(i)}$	0	0	1	0	1	0	0	1	1	1	0	1	1	0	1	1	1	0	0	0
second observation $X_a^{(i)}$	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0
individual's type $Y^{(i)}$	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	0	0	0

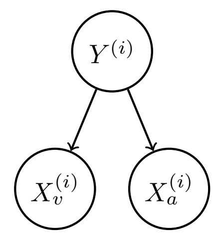
The superscript (i) denotes that the datum is the ith one. Now, the individual with $i = 20$ comes out, and you want to predict the individual's type $Y^{(20)}$ given that you observed $X_v^{(20)} = 1$ and $X_a^{(20)} = 1$.

(a) Assume that the types are independent, and that the observations are independent conditioned on the type. You can model this using naïve Bayes, with $X_v^{(i)}$ and $X_a^{(i)}$ as the features and $Y^{(i)}$ as the labels. Assume the probability distributions take on the following form:

$$P(X_v^{(i)} = x_v | Y^{(i)} = y) = \begin{cases} p_v & \text{if } x_v = y \\ 1 - p_v & \text{if } x_v \neq y \end{cases}$$

$$P(X_a^{(i)} = x_a | Y^{(i)} = y) = \begin{cases} p_a & \text{if } x_a = y \\ 1 - p_a & \text{if } x_a \neq y \end{cases}$$

$$P(Y^{(i)} = 1) = q$$



for $p_v, p_a, q \in [0, 1]$ and $i \in \mathbb{N}$.

(i) What's the maximum likelihood estimate of p_v, p_a and q?

$p_v =$ $\frac{4}{5}$ $\qquad$ $p_a =$ $\frac{3}{5}$ $\qquad$ $q =$ $\frac{1}{2}$

To estimate q, we count 10 $Y = 1$ and 10 $Y = 0$ in the data. For p_v, we have $p_v = 8/10$ cases where $X_v = 1$ given $Y = 1$ and $1 - p_v = 2/10$ cases where $X_v = 1$ given $Y = 0$. So $p_v = 4/5$. For p_a, we have $p_a = 2/10$ cases where $X_a = 1$ given $Y = 1$ and $1 - p_v = 0/10$ cases where $X_v = 1$ given $Y = 0$. The average of $2/10$ and 1 is $3/5$.

(ii) What is the probability that the next individual is Pacman given your observations? Express your answer in terms of the parameters p_v, p_a and q (you might not need all of them).

$$P(Y^{(20)} = 1 | X_v^{(20)} = 1, X_a^{(20)} = 1) = \frac{p_v p_a q}{p_v p_a q + (1 - p_v)(1 - p_a)(1 - q)}$$

The joint distribution $P(Y = 1, X_v = 1, X_a = 1) = p_v p_a q$. For the denominator, we need to sum out over Y, that is, we need $P(Y = 1, X_v = 1, X_a = 1) + P(Y = 0, X_v = 1, X_a = 1)$.

Now, assume that you are given additional information: you are told that the individuals are actually coming out of a bus that just arrived, and each bus carries *exactly* 9 individuals. Unlike before, the types of every 9 consecutive individuals are *conditionally* independent given the bus type, which is denoted by Z. Only after all of the 9 individuals have walked out, you find out the bus type: one that carries mostly Pacmans ($Z = 1$) or one that carries mostly ghosts ($Z = 0$). Thus, you only know the bus type in which the first 18 individuals came in:

individual i	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
first observation $X_v^{(i)}$	0	0	1	0	1	0	0	1	1	1	0	1	1	0	1	1	1	0	0	0
second observation $X_a^{(i)}$	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0
individual's type $Y^{(i)}$	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	0	0	0

bus j	0	1	
bus type $Z^{(j)}$	0	1	

(b) You can model this using a variant of naïve bayes, where now 9 consecutive labels $Y^{(i)}, \dots, Y^{(i+8)}$ are *conditionally* independent given the bus type $Z^{(j)}$, for bus j and individual $i = 9j$. Assume the probability distributions take on the following form:

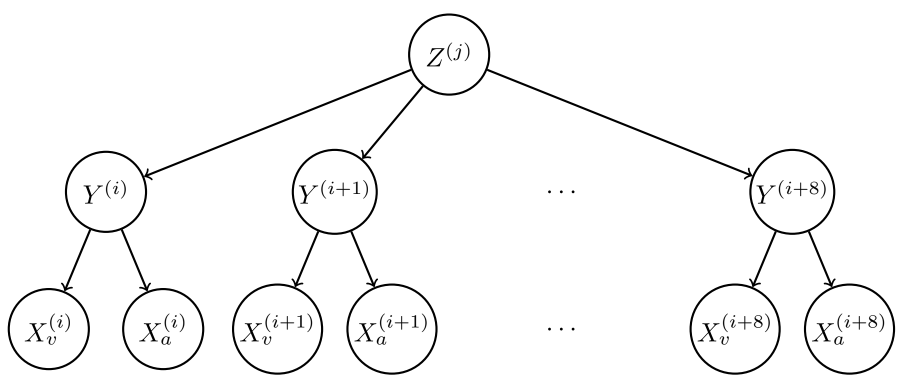
$$P(X_v^{(i)} = x_v | Y^{(i)} = y) = \begin{cases} p_v & \text{if } x_v = y \\ 1 - p_v & \text{if } x_v \neq y \end{cases}$$

$$P(X_a^{(i)} = x_a | Y^{(i)} = y) = \begin{cases} p_a & \text{if } x_a = y \\ 1 - p_a & \text{if } x_a \neq y \end{cases}$$

$$P(Y^{(i)} = 1 | Z^{(j)} = z) = \begin{cases} q_0 & \text{if } z = 0 \\ q_1 & \text{if } z = 1 \end{cases}$$

$$P(Z^{(j)} = 1) = r$$

for $p, q_0, q_1, r \in [0, 1]$ and $i, j \in \mathbb{N}$.



(i) What's the maximum likelihood estimate of q_0, q_1 and r?

$q_0 =$ $\frac{2}{9}$ $\qquad q_1 =$ $\frac{8}{9}$ $\qquad r =$ $\frac{1}{2}$

For r, we've seen one ghost bus and one pacman bus, so $r = 1/2$. For q_0, we're finding $P(Y = 1|Z = 0)$, which is 2/9. For q_1, we're finding $P(Y = 1|Z = 1)$, which is 8/9.

(ii) Compute the following joint probability. Simplify your answer as much as possible and express it in terms of the parameters p_v, p_a, q_0, q_1 and r (you might not need all of them).

$P(Y^{(20)} = 1, X_v^{(20)} = 1, X_a^{(20)} = 1, Y^{(19)} = 1, Y^{(18)} = 1) =$ $\underline{\quad p_a p_v[q_0^3(1-r) + q_1^3 r] \quad}$

$$
\begin{aligned}
&P(Y^{(20)} = 1, X_v^{(20)} = 1, X_a^{(20)} = 1, Y^{(19)} = 1, Y^{(18)} = 1) \\
&= \sum_z P(Y^{(20)} = 1|Z^{(2)} = z)P(Z^{(2)} = z)P(X_v^{(20)} = 1|Y^{(20)} = 1)P(X_a^{(20)} = 1|Y^{(20)} = 1) \\
&\qquad P(Y^{(19)} = 1|Z^{(2)} = z)P(Y^{(18)} = 1|Z^{(2)} = z) \\
&= q_0(1-r)p_a p_v q_0 q_0 + q_1 r p_a p_v q_1 q_1 \\
&= p_a p_v[q_0^3(1-r) + q_1^3 r]
\end{aligned}
$$