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:

$$B_t(X_t) = \frac{(i) \max_{x_{t-1}} \sum_{x_t} (ii) \max_{x_t} \sum_{x_{t-1}} (iii) \sum_{x_t} (iv) \sum_{x_t} B_{t-1}(x_{t-1})}{1}$$

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

(i) $\max_{x_{t-1}}$ $\sum_{x_t}$ $\max_{x_t}$ $\sum_{x_{t-1}}$ $1$

(ii) $p(X_t|x_{t-1})$ $p(X_t|x_{t-1})p(X_t|a_t)$ $p(X_t)$ $p(X_t|x_{t-1}, a_t)$ $1$

$\bigcirc$ None of the above combinations is correct

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

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

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) $\max_{a_t} \prod_{t=1}^{T}$ $\sum_{t=1}^{T}$ $\min_t$ $1$

(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) $U_T$ $\frac{1}{U_t}$ $\frac{1}{U_T}$ $U_t$ $1$

$\bigcirc$ None of the above combinations is correct

(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:

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

(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.
<table>
<thead>
<tr>
<th>Method</th>
<th>Most accurate</th>
<th>Least accurate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exact inference</td>
<td>○ 1</td>
<td>○ 2</td>
</tr>
<tr>
<td>Particle filtering with no resampling</td>
<td>○ 1</td>
<td>○ 2</td>
</tr>
<tr>
<td>Particle filtering with resampling before every time elapse</td>
<td>○ 1</td>
<td>○ 2</td>
</tr>
<tr>
<td>Particle filtering with resampling before every other time elapse</td>
<td>○ 1</td>
<td>○ 2</td>
</tr>
</tbody>
</table>
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 individual 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:

<table>
<thead>
<tr>
<th>individual $i$</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
<th>16</th>
<th>17</th>
<th>18</th>
<th>19</th>
</tr>
</thead>
<tbody>
<tr>
<td>first observation $X_v^{(i)}$</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>second observation $X_a^{(i)}$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>individual’s type $Y^{(i)}$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

The superscript ($i$) denotes that the datum is the $i$th 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 naive 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 $q$?

$$p_v = \quad p_a = \quad q =$$

(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).

$$P(Y^{(20)} = 1 | X_v^{(20)} = 1, X_a^{(20)} = 1) =$$
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:

<table>
<thead>
<tr>
<th>individual $i$</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
<th>16</th>
<th>17</th>
<th>18</th>
<th>19</th>
</tr>
</thead>
<tbody>
<tr>
<td>first observation $X_v^{(i)}$</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>second observation $X_a^{(i)}$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>individual’s type $Y^{(i)}$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>bus $j$</th>
<th>0</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>bus type $Z^{(j)}$</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

(b) You can model this using a variant of naïve bayes, where now 9 consecutive labels $Y^{(i)}, \ldots, Y^{(i+8)}$ 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 $r$?

$q_0 = \quad q_1 = \quad r = \quad$
(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).

\[
P(Y^{(20)} = 1, X_v^{(20)} = 1, X_a^{(20)} = 1, Y^{(19)} = 1, Y^{(18)} = 1) = \\
\]