1 Basic computations

Consider the following Hidden Markov Model.

\[
\begin{array}{c}
X_1 \\
\downarrow
\end{array}
\begin{array}{c}
X_2 \\
\downarrow
\end{array}
\begin{array}{c}
O_2 \\
\downarrow
\end{array}
\begin{array}{c|c}
X_1 & Pr(X_1) \\
0 & 0.3 \\
1 & 0.7 \\
\end{array}
\begin{array}{c|c|c}
X_t & X_{t+1} & Pr(X_{t+1}|X_t) \\
0 & 0 & 0.4 \\
0 & 1 & 0.6 \\
1 & 0 & 0.8 \\
1 & 1 & 0.2 \\
\end{array}
\begin{array}{c|c|c}
X_t & O_t & Pr(O_t|X_t) \\
0 & A & 0.9 \\
0 & B & 0.1 \\
1 & A & 0.5 \\
1 & B & 0.5 \\
\end{array}
\]

Suppose that \(O_1 = A\) and \(O_2 = B\) is observed.

Use the Forward algorithm to compute the probability distribution \(Pr(X_2, O_1 = A, O_2 = B)\). Show your work. You do not need to evaluate arithmetic expressions involving only numbers.

\[
\begin{array}{c|c}
X_1 & Pr(X_1, O_1 = A) \\
0 & 0.3 \cdot 0.9 \\
1 & 0.7 \cdot 0.5 \\
\end{array}
\]

Use the Viterbi algorithm to compute the maximum probability sequence \(X_1, X_2\). Show your work.

\[
\begin{array}{c|c}
X_1 & Pr(X_1, O_1 = A) \\
0 & 0.3 \cdot 0.9 \\
1 & 0.7 \cdot 0.5 \\
\end{array}
\]

\[
\begin{array}{c|c}
X_2 & \max_{x_1} Pr(X_1 = x_1, X_2, O_1 = A, O_2 = B) \\
0 & 0.1 \cdot \max(0.4 \cdot (0.3 \cdot 0.9), 0.8 \cdot (0.7 \cdot 0.5)) = 0.1 \cdot \max(0.108, 0.28) = 0.028 \\
1 & 0.5 \cdot \max(0.6 \cdot (0.3 \cdot 0.9), 0.2 \cdot (0.7 \cdot 0.5)) = 0.5 \cdot \max(0.162, 0.07) = 0.081 \\
\end{array}
\]

\[
\begin{array}{c|c}
\arg\max & \\
X_1 = 1 & X_1 = 0 \\
\end{array}
\]

Thus, in the maximum probability sequence, \(X_2 = 1\) and \(X_1 = 0\).

True or false: Variable elimination is generally more accurate than the Forward algorithm. Explain your answer.

They both perform exact inference.
2 Tracking a Jabberwock

You have been put in charge of a Jabberwock for your friend Lewis. The Jabberwock is kept in a large tugley wood which is conveniently divided into an $N \times N$ grid. It wanders freely around the $N^2$ possible cells. At each time step $t = 1, 2, 3, \ldots$, the Jabberwock is in some cell $X_t \in \{1, \ldots, N\}^2$, and it moves to cell $X_{t+1}$ randomly as follows: with probability $1 - \epsilon$, it chooses one of the (up to 4) valid neighboring cells uniformly at random; with probability $\epsilon$, it uses its magical powers to teleport to a random cell uniformly at random among the $N^2$ possibilities (it might teleport to the same cell). Suppose $\epsilon = \frac{1}{2}$, $N = 10$ and that the Jabberwock always starts in $X_1 = (1, 1)$.

1. Compute the probability that the Jabberwock will be in $X_2 = (2, 1)$ at time step 2. What about $\Pr(X_2 = (4, 4))$?

$\Pr(X_2 = (2, 1)) = \frac{1}{2} \cdot \frac{1}{2} + \frac{1}{2} \cdot \frac{1}{100} = 0.255$

$\Pr(X_2 = (4, 4)) = \frac{1}{2} \cdot \frac{1}{100} = 0.005$

At each time step $t$, you don’t see $X_t$ but see $E_t$, which is the row that the Jabberwock is in; that is, if $X_t = (r, c)$, then $E_t = r$. You still know that $X_1 = (1, 1)$.

2. Suppose we see that $E_1 = 1$, $E_2 = 2$, $E_3 = 10$. Fill in the following table with the distribution over $X_t$ after each time step, taking into consideration the evidence. Your answer should be concise. *Hint:* you should not need to do any heavy calculations.

<table>
<thead>
<tr>
<th>$t$</th>
<th>$\Pr(X_t, e_{1:t-1})$</th>
<th>$\Pr(X_t, e_{1:t})$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>(1,1) : 1.0, (others) : 0.0</td>
<td>(1,1) : 1.0, (others): 0.0</td>
</tr>
<tr>
<td>2</td>
<td>(1,2), (2,1): 51/200, (others) : 1/200</td>
<td>(2,1): 51/200, (2,2+): 1/200</td>
</tr>
</tbody>
</table>