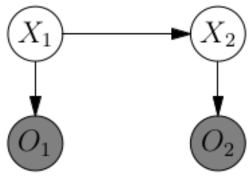
Consider the following Hidden Markov Model.



X_1	$\Pr(X_1)$
0	0.3
1	0.7

X_t	X_{t+1}	$\Pr(X_{t+1} \mid X_t)$
0	0	0.4
0	1	0.6
1	0	0.8
1	1	0.2

X_t	O_t	$\Pr(O_t \mid X_t)$
0	A	0.9
0	B	0.1
1	A	0.5
1	B	0.5

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.

Solution. We must compute $\Pr(X_2, O_1 = A, O_2 = B)$ for each value of X_2. That is, we compute the two probabilities $\Pr(X_2 = 0, O_1 = A, O_2 = B)$ and $\Pr(X_2 = 1, O_1 = A, O_2 = B)$. We do this by first computing $\Pr(X_2 = 0, x_1, O_1 = A, O_2 = B)$ for each $x_1 \in \{0, 1\}$ (the possible values for the random variable X_1), and then we sum over x_1 to get the marginal probability $\Pr(X_2 = 0, O_1 = A, O_2 = B)$.

$$\Pr(X_2, O_1 = A, O_2 = B) = \Sigma_{x_1} \Pr(X_2, x_1, O_1 = A, O_2 = B).$$

Recall, by the definition of conditional probability (or the chain rule), we have

$$\begin{aligned}
\Pr(X_2, x_1, O_1 = A, O_2 = B) &= \Pr(x_1)\Pr(X_2, O_1 = A, O_2 = B \mid x_1) \\
&= \Pr(x_1)P(O_1 = A \mid x_1)\Pr(X_2, O_2 = B \mid x_1, O_1 = A) \\
&= \Pr(x_1)P(O_1 = A \mid x_1)\Pr(X_2 \mid x_1, O_1 = A)\Pr(O_2 = B \mid X_2, x_1, O_1 = A)
\end{aligned}$$

Now we use the structure of the Bayes net to infer some conditional independences and simplify the above express. Specifically, O_2 is independent of O_1 given X_2 and O_2 is independent of X_1 given X_2. Also, X_2 is independent of O_1 given X_1. Therefore, $\Pr(X_2, x_1, O_1 = A, O_2 = B) = \Pr(x_1) \Pr(O_1 = A \mid x_1)\Pr(X_2 \mid x_1) \Pr(O_2 = B \mid X_2)$. We now use the data in the tables above to compute

$$\begin{aligned}
\Pr(X_2 = 0, X_1 = 0, O_1 = A, O_2 = B) &= \Pr(X_1 = 0)\Pr(O_1 = A \mid X_1 = 0)\Pr(X_2 = 0 \mid X_1 = 0)\Pr(O_2 = B \mid X_2 = 0) \\
&= 0.3 \cdot 0.9 \cdot 0.4 \cdot 0.1, \text{ and}
\end{aligned}$$

$$\begin{aligned}
\Pr(X_2 = 0, X_1 = 1, O_1 = A, O_2 = B) &= \Pr(X_1 = 1)\Pr(O_1 = A \mid X_1 = 1)\Pr(X_2 = 0 \mid X_1 = 1)\Pr(O_2 = B \mid X_2 = 0) \\
&= 0.7 \cdot 0.5 \cdot 0.8 \cdot 0.1
\end{aligned}$$

Therefore, $\Pr(X_2 = 0, O_1 = A, O_2 = B) = 0.3 \cdot 0.9 \cdot 0.4 \cdot 0.1 + 0.7 \cdot 0.5 \cdot 0.8 \cdot 0.1 = 0.1\ (0.3 \cdot 0.9 \cdot 0.4 + 0.7 \cdot 0.5 \cdot 0.8) = 0.1\ (0.108 + 0.280) = 0.0388$, Similarly,

$$\begin{aligned}
\Pr(X_2 = 1, X_1 = 0, O_1 = A, O_2 = B) &= \Pr(X_1 = 0)\Pr(O_1 = A \mid X_1 = 0)\Pr(X_2 = 1 \mid X_1 = 0)\Pr(O_2 = B \mid X_2 = 1) \\
&= 0.3 \cdot 0.9 \cdot 0.6 \cdot 0.5, \text{ and}
\end{aligned}$$

$$\begin{aligned}
\Pr(X_2 = 1, X_1 = 1, O_1 = A, O_2 = B) &= \Pr(X_1 = 1)\Pr(O_1 = A \mid X_1 = 1)\Pr(X_2 = 1 \mid X_1 = 1)\Pr(O_2 = B \mid X_2 = 1) \\
&= 0.7 \cdot 0.5 \cdot 0.2 \cdot 0.5
\end{aligned}$$

Therefore, $\Pr(X_2 = 1, O_1 = A, O_2 = B) = 0.3 \cdot 0.9 \cdot 0.6 \cdot 0.5 + 0.7 \cdot 0.5 \cdot 0.2 \cdot 0.5 = 0.5\ (0.3 \cdot 0.9 \cdot 0.6 + 0.7 \cdot 0.5 \cdot 0.2) = 0.5\ (0.162 + 0.07) = 0.1160$, So the distribution $\Pr(X_2, O_1 = A, O_2 = B)$ is given by the following table:

X_2	$\Pr(X_2, O_1 = A, O_2 = B)$
0	0.0388
1	0.1160