



- 1 a) The Bayesian network for this problem is shown in figure Figure 1.

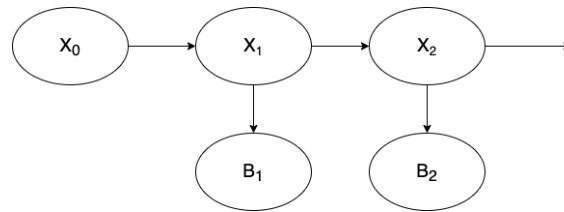


Figure 1: Bayesian network for the problem.

Here  $X_t$  is the state distribution on time  $t$  and  $B_t$  is the evidence distribution at time  $t$ .

The probability distributions are given by tables 1, 2 and 3.

	$P(X_t)$	
$X_{t-1}$	$t$	$f$
$t$	0.8	0.2
$f$	0.3	0.7

Table 1: Probability distribution for  $P(X_t|X_{t-1})$ .

	$P(B_t)$	
$X_t$	$t$	$f$
$t$	0.75	0.25
$f$	0.2	0.8

Table 2: Probability distribution for  $P(B_t|X_t)$ .

$P(X_0)$	
$t$	$f$
0.5	0.5

Table 3: Probability distribution for  $P(X_0)$ .

Formulation as a hidden Markov model:

$$\mathbf{T} = \begin{bmatrix} 0.8 & 0.2 \\ 0.3 & 0.7 \end{bmatrix}$$

$$\mathbf{P}(X_0) = \begin{bmatrix} 0.5 \\ 0.5 \end{bmatrix}$$

$\mathbf{T}$  is the transition model and  $\mathbf{P}(X_0)$  is the initial distribution.

The sensor model  $\mathbf{O}_t$  depends on the evidence given:

$$\mathbf{e}_t = \text{BirdsNearby} \implies \mathbf{O}_t = \begin{bmatrix} 0.75 & 0 \\ 0 & 0.2 \end{bmatrix}$$

$$\mathbf{e}_t = \text{NoBirdsNearby} \implies \mathbf{O}_t = \begin{bmatrix} 0.25 & 0 \\ 0 & 0.8 \end{bmatrix}$$

- b) This kind of operation is filtering, where we calculate the current belief state given all the evidence to date. The results are given in Figure 2. This operation gives us the probability distribution of the current state given all the evidence to date.

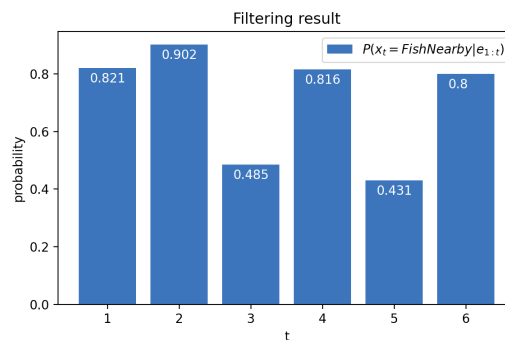


Figure 2: Filtering for  $t = 1, \dots, 6$ .

As we can see the filtering results vary with the evidence, when we observe birds nearby the probability of fish nearby is also high, while when there are no birds nearby the probability of fish nearby is lower, which makes sense from the transition and sensor model.

- c) This kind of operation is prediction, where we calculate the distribution of a future state, based on the previous evidence, but no new addition evidence are given. This is the same as filtering without adding new evidence. The results are given in Figure 3.

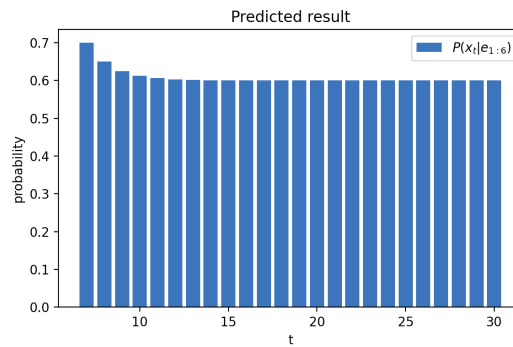


Figure 3: Prediction for  $t = 7, \dots, 30$ .

As we can see, the belief state converges towards  $[0.6 \ 0.4]^\top$  as  $t$  increases. This is because  $[0.6 \ 0.4]^\top$  is the eigenvector for the eigenvalue 1 in the matrix  $T^\top$ . This will be explained further in Task 2b.

- d) This kind of operation is smoothing, where we calculate the distribution of a past state, given all the evidence up to the present. The results are given in Figure 4.

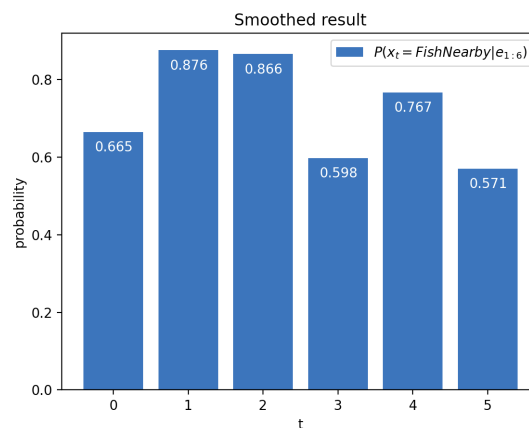


Figure 4: Smoothing for  $t = 0, \dots, 5$ .

As we can see the smoothed results gives a higher probability of fish nearby when there are no birds nearby than filtering. For example on day 3 the smoothed result is 0.598 while the filtered is 0.485, this is because the birds nearby on day 4 makes it more likely to be fish nearby on day 3 since fish nearby tends to persist.

- e) This kind of operation is most likely sequence, where we calculate the most like sequence of states given all the evidence up to the present. The results where the sequence given in Figure 5.

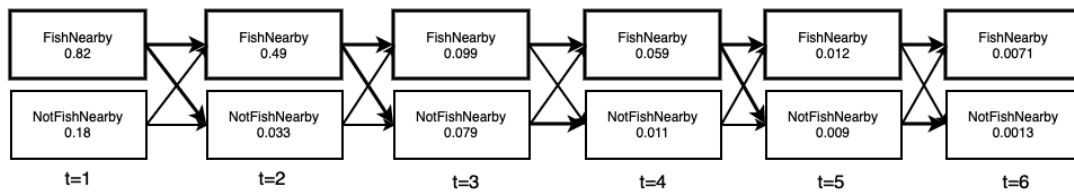
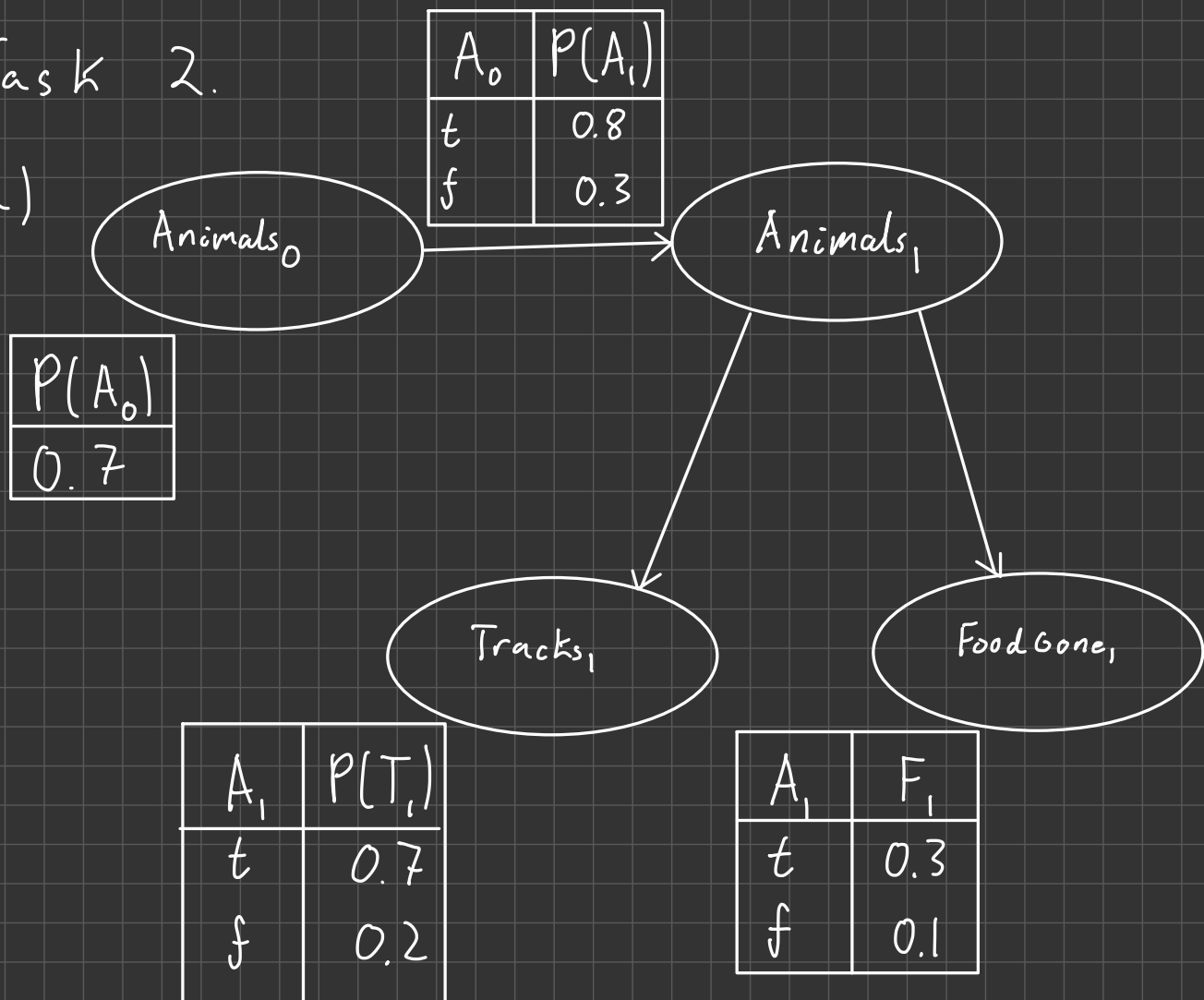


Figure 5: Viterbi sequence for  $t = 1, \dots, 6$ .

So even though we had evidence of no birds nearby on two of the days, the most likely sequence is that all days had fish nearby. This happens because the Viterbi algorithm takes into account the joint probabilities over all the time steps.

Task 2.

a)



A = Animals, T = Tracks, F = Food Gone.

b) From equation 15.5 in the book we have that:

$$P(X_{t+1} | e_{1:t+1}) = \alpha P(e_{t+1} | X_{t+1}) \sum_{x_t} P(X_{t+1} | x_t) P(x_t | e_{1:t})$$

In matrix-form we have:

$$P(X_{t+1} | e_{1:t+1}) = \alpha O_{t+1} T^T P(X_t | e_{1:t})$$

Which is a recursive equation with the base case:

$$P(X_0 | e_{1:0}) = P(X_0) = P(A_0) = \begin{bmatrix} 0.7 \\ 0.3 \end{bmatrix}$$

$T$  is the transition model defined as:

$$T = \begin{bmatrix} 0.8 & 0.2 \\ 0.3 & 0.7 \end{bmatrix}$$

$O_t$  is the sensor-model.

$$O_t = \begin{bmatrix} P(t_t \wedge f_t | a_t) & 0 \\ 0 & P(t_t \wedge f_t | \neg a_t) \end{bmatrix}$$

Since  $T_t$  and  $F_t$  is conditionally independent given  $A_t$  we have:

$$O_t = \begin{bmatrix} P(t_t | a_t) P(f_t | a_t) & 0 \\ 0 & P(t_t | \neg a_t) P(f_t | \neg a_t) \end{bmatrix}$$

Now we can start calculating:

$t=1$ :

$$P(X_1 | e_{1:1}) = \propto O_1 T^T P(X_0)$$

$$= \alpha \begin{bmatrix} 0.7 \cdot 0.3 & 0 \\ 0 & 0.2 \cdot 0.1 \end{bmatrix} \begin{bmatrix} 0.8 & 0.3 \\ 0.2 & 0.7 \end{bmatrix} \begin{bmatrix} 0.7 \\ 0.3 \end{bmatrix}$$

$$= \alpha \begin{bmatrix} 0.21 & 0 \\ 0 & 0.02 \end{bmatrix} \begin{bmatrix} 0.65 \\ 0.35 \end{bmatrix}$$

$$= \alpha \begin{bmatrix} 0.1365 \\ 0.007 \end{bmatrix} = \begin{bmatrix} 39/41 \\ 2/41 \end{bmatrix} = f_{1:1}$$

$t=2$ :

$$P(X_2 | e_{1:2}) = \alpha O_2^T P(X_1 | e_{1:1})$$

$$= \alpha \begin{bmatrix} 0.3 \cdot 0.3 & 0 \\ 0 & 0.8 \cdot 0.1 \end{bmatrix} \begin{bmatrix} 0.8 & 0.3 \\ 0.2 & 0.7 \end{bmatrix} \begin{bmatrix} 39/41 \\ 2/41 \end{bmatrix}$$

$$= \alpha \begin{bmatrix} 0.09 & 0 \\ 0 & 0.08 \end{bmatrix} \begin{bmatrix} 159/205 \\ 46/205 \end{bmatrix}$$

$$= \alpha \begin{bmatrix} 1431/20500 \\ 92/5125 \end{bmatrix} \approx \begin{bmatrix} 0.79544 \\ 0.20456 \end{bmatrix} = f_{1:2}$$

$t=3$ :

$$P(X_3 | e_{1:t}) = \alpha O_3^T P(X_2 | e_{1:2})$$

$$= \alpha \begin{bmatrix} 0.3 \cdot 0.7 & 0 \\ 0 & 0.8 \cdot 0.9 \end{bmatrix} \begin{bmatrix} 0.8 & 0.3 \\ 0.2 & 0.7 \end{bmatrix} \begin{bmatrix} 0.79544 \\ 0.20456 \end{bmatrix}$$

$$\begin{aligned}
&= \alpha \begin{bmatrix} 0.21 & 0 \\ 0 & 0.72 \end{bmatrix} \begin{bmatrix} 0.69772 \\ 0.30228 \end{bmatrix} \\
&= \alpha \begin{bmatrix} 0.14652 \\ 0.21764 \end{bmatrix} \approx \begin{bmatrix} 0.40235 \\ 0.59765 \end{bmatrix} = f_{1:3}
\end{aligned}$$

$t=4$ :

$$\begin{aligned}
P(X_4 | e_{1:4}) &= \alpha O_4 T^T P(X_3 | e_{1:3}) \\
&= \alpha \begin{bmatrix} 0.7 \cdot 0.7 & 0 \\ 0 & 0.2 \cdot 0.9 \end{bmatrix} \begin{bmatrix} 0.8 & 0.3 \\ 0.2 & 0.7 \end{bmatrix} \begin{bmatrix} 0.40235 \\ 0.59765 \end{bmatrix} \\
&= \alpha \begin{bmatrix} 0.49 & 0 \\ 0 & 0.18 \end{bmatrix} \begin{bmatrix} 0.501175 \\ 0.498825 \end{bmatrix} \\
&= \alpha \begin{bmatrix} 0.2458575 \\ 0.0897885 \end{bmatrix} \approx \begin{bmatrix} 0.732 \\ 0.268 \end{bmatrix} = f_{1:4}
\end{aligned}$$

c) Prediction is the same as filtering without adding new evidence:

$$P(X_{t+k+1} | e_{1:t}) = \sum_{x_{t+k}} P(X_{t+k+1} | x_{t+k}) P(x_{t+k} | e_{1:t})$$

We want to calculate this with  $t=4$  and  $k=0, 1, 2, 3$ .



We can rewrite this using matrices:

$$P(X_{t+k+1} | e_{1:t}) = T^T f_{1:t+k} = f_{1:t+k+1}$$

$k=0$ :

$$P(X_5 | e_{1:4}) = T^T \overset{\substack{\swarrow \text{gotten from previous task}}}{f_{1:4}}$$

$$= \begin{bmatrix} 0.8 & 0.3 \\ 0.2 & 0.7 \end{bmatrix} \begin{bmatrix} 0.732266 \\ 0.267733 \end{bmatrix}$$

$$\approx \begin{bmatrix} 0.66613 \\ 0.33387 \end{bmatrix} = f_{1:5}$$

$k=1$ :

$$P(X_6 | e_{1:4}) = T^T f_{1:5}$$

$$= \begin{bmatrix} 0.8 & 0.3 \\ 0.2 & 0.7 \end{bmatrix} \begin{bmatrix} 0.666133 \\ 0.333867 \end{bmatrix} \approx \begin{bmatrix} 0.63307 \\ 0.36693 \end{bmatrix} = f_{1:6}$$

$k=2$ :

$$P(X_7 | e_{1:4}) = T^T f_{1:6}$$

$$= \begin{bmatrix} 0.8 & 0.3 \\ 0.2 & 0.7 \end{bmatrix} \begin{bmatrix} 0.633066 \\ 0.366934 \end{bmatrix}$$

$$\approx \begin{bmatrix} 0.61653 \\ 0.38347 \end{bmatrix} = f_{1:7}$$

$k=3$ :

$$P(X_8 | e_{1:4}) = T^T f_{1:7}$$

$$= \begin{bmatrix} 0.8 & 0.3 \\ 0.2 & 0.7 \end{bmatrix} \begin{bmatrix} 0.616533 \\ 0.383467 \end{bmatrix} \approx \begin{bmatrix} 0.60827 \\ 0.39173 \end{bmatrix} = f_{1:8}$$

d)  $\lim_{t \rightarrow \infty} P(X_t | e_{1:4}) = \begin{bmatrix} 0.6 \\ 0.4 \end{bmatrix}$

$$P(X_t | e_{1:4}) = \begin{bmatrix} 0.6 \\ 0.4 \end{bmatrix}$$

$$\Rightarrow P(X_{t+1} | e_{1:4}) = T^T P(X_t | e_{1:4})$$

$$= \begin{bmatrix} 0.8 & 0.3 \\ 0.2 & 0.7 \end{bmatrix} \begin{bmatrix} 0.6 \\ 0.4 \end{bmatrix} = \begin{bmatrix} 0.6 \\ 0.4 \end{bmatrix}$$

Thus  $P(X_t | e_{1:4})$  has converged to  $[0.6 \ 0.4]^T$ .

Next I will demonstrate why it converges to  $[0.6 \ 0.4]^T$

We have that  $T^T$  is a regular stochastic matrix. Then  $T^T$  has a unique equilibrium vector  $q$ . For every vector  $x_0$  where the Markov-process starts, the Markovchain  $\{x_t\}$  will converge towards  $q$  when  $t \rightarrow \infty$ .

This  $q$  can be found by finding the eigenvector corresponding to the eigenvalue 1:

$$T^T - I_2 = \begin{bmatrix} 0.8 & 0.3 \\ 0.2 & 0.7 \end{bmatrix} - \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} -0.2 & 0.3 \\ 0.2 & -0.3 \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} -0.2 & 0.3 \\ 0.2 & -0.3 \end{bmatrix} q = 0$$

$$\begin{bmatrix} -0.2 & 0.3 \\ 0.2 & -0.3 \end{bmatrix} \sim \begin{bmatrix} -0.2 & 0.3 \\ 0 & 0 \end{bmatrix}$$

$$\Rightarrow -0.2q_1 + 0.3q_2 = 0$$

$$q_1 = 1.5q_2$$

$$\Rightarrow q = t \begin{bmatrix} 1.5 \\ 1 \end{bmatrix}$$

We want  $q$  to be a probability vector:

$$\Rightarrow q = \frac{1}{1+1.5} \begin{bmatrix} 1.5 \\ 1 \end{bmatrix} = \begin{bmatrix} 0.6 \\ 0.4 \end{bmatrix}.$$

$$\begin{aligned}
 e) \quad P(X_t | e_{1:4}) &= \alpha P(X_t | e_{1:t}) P(e_{t+1:4} | X_t) \\
 &= \alpha f_{1:t} \times b_{t+1:4} \quad (\text{pointwise multiplication})
 \end{aligned}$$

$$\begin{aligned}
 b_{k+1:t} &= P(e_{k+1:t} | X_k) \\
 &= \sum_{X_{k+1}} \underbrace{P(e_{k+1} | X_{k+1})}_{\text{sensor model}} \underbrace{P(e_{k+2:t} | X_{k+1})}_{b_{k+2:t}} \underbrace{P(X_{k+1} | X_k)}_{\text{transition model}} \\
 &= T O_{k+1} b_{k+2:t}
 \end{aligned}$$

The base case for this recursive process is

$$b_{t+1:t} = P(e_{t+1:t} | X_t) = P(\cdot | X_t) \mathbf{1} = \begin{bmatrix} 1 & 1 \end{bmatrix}^T$$

$t=3$ : ↙ have from previous task

$$P(X_3 | e_{1:4}) = \alpha f_{1:3} \times b_{4:4}$$

$$b_{4:4} = T O_4 b_{5:4}$$

$$= \begin{bmatrix} 0.8 & 0.2 \\ 0.3 & 0.7 \end{bmatrix} \begin{bmatrix} 0.49 & 0 \\ 0 & 0.18 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$= \begin{bmatrix} 0.8 & 0.2 \\ 0.3 & 0.7 \end{bmatrix} \begin{bmatrix} 0.21 \\ 0.72 \end{bmatrix} = \begin{bmatrix} 0.428 \\ 0.273 \end{bmatrix}$$

$$\begin{aligned}
 \Rightarrow P(X_3 | e_{1:4}) &= \alpha \begin{bmatrix} 0.40235191 \\ 0.59764809 \end{bmatrix} \times \begin{bmatrix} 0.428 \\ 0.273 \end{bmatrix} \\
 &= \alpha \begin{bmatrix} 0.1722066 \\ 0.1631579 \end{bmatrix} \\
 &\approx \begin{bmatrix} 0.513491 \\ 0.486509 \end{bmatrix}
 \end{aligned}$$

$t = 2:$

$$P(X_2 | e_{1:4}) = \alpha f_{1:2} \times b_{3:4}$$

$$\begin{aligned}
 b_{3:4} = T O_3 b_{4:4} &= \begin{bmatrix} 0.8 & 0.2 \\ 0.3 & 0.7 \end{bmatrix} \begin{bmatrix} 0.21 & 0 \\ 0 & 0.72 \end{bmatrix} \begin{bmatrix} 0.428 \\ 0.273 \end{bmatrix} \\
 &= \begin{bmatrix} 0.8 & 0.2 \\ 0.3 & 0.7 \end{bmatrix} \begin{bmatrix} 0.08988 \\ 0.19656 \end{bmatrix} \\
 &= \begin{bmatrix} 0.111216 \\ 0.164556 \end{bmatrix}
 \end{aligned}$$

$$\begin{aligned}
 \Rightarrow P(X_2 | e_{1:4}) &= \alpha \begin{bmatrix} 0.7954419 \\ 0.2045581 \end{bmatrix} \times \begin{bmatrix} 0.111216 \\ 0.164556 \end{bmatrix} \\
 &= \alpha \begin{bmatrix} 0.0884659 \\ 0.0336613 \end{bmatrix} \approx \begin{bmatrix} 0.724375 \\ 0.275625 \end{bmatrix}
 \end{aligned}$$

$t=1$ :

$$P(X_1 | e_{1:4}) = \alpha f_{1:1} \times b_{2:4}$$

$$b_{2:4} = T O_2 b_{3:4} = \begin{bmatrix} 0.8 & 0.2 \\ 0.3 & 0.7 \end{bmatrix} \begin{bmatrix} 0.09 & 0 \\ 0 & 0.08 \end{bmatrix} \begin{bmatrix} 0.111216 \\ 0.164556 \end{bmatrix}$$

$$= \begin{bmatrix} 0.8 & 0.2 \\ 0.3 & 0.7 \end{bmatrix} \begin{bmatrix} 0.01000944 \\ 0.01316448 \end{bmatrix} = \begin{bmatrix} 0.01064045 \\ 0.01221797 \end{bmatrix}$$

$$\Rightarrow P(X_1 | e_{1:4}) = \alpha \begin{bmatrix} 0.95121951 \\ 0.04878049 \end{bmatrix} \times \begin{bmatrix} 0.01064045 \\ 0.01221797 \end{bmatrix}$$
$$= \alpha \begin{bmatrix} 0.0101214 \\ 0.0005960 \end{bmatrix} \approx \begin{bmatrix} 0.94439 \\ 0.05561 \end{bmatrix}$$

$t=0$ :

$$P(X_0 | e_{1:4}) = \alpha f_{1:0} \times b_{1:4}$$

$$b_{1:4} = T O_1 b_{2:4} = \begin{bmatrix} 0.8 & 0.2 \\ 0.3 & 0.7 \end{bmatrix} \begin{bmatrix} 0.21 & 0 \\ 0 & 0.02 \end{bmatrix} \begin{bmatrix} 0.0106405 \\ 0.0122180 \end{bmatrix}$$

$$= \begin{bmatrix} 0.8 & 0.2 \\ 0.3 & 0.7 \end{bmatrix} \begin{bmatrix} 0.0022345 \\ 0.00024436 \end{bmatrix} = \begin{bmatrix} 0.001836472 \\ 0.000841402 \end{bmatrix}$$

$$\Rightarrow P(X_0 | e_{1:4}) = \alpha \begin{bmatrix} 0.7 \\ 0.3 \end{bmatrix} \times \begin{bmatrix} 0.001836472 \\ 0.000841402 \end{bmatrix}$$

$$= \propto \begin{bmatrix} 0.00128553 \\ 0.00025242 \end{bmatrix} \approx \begin{bmatrix} 0.83587 \\ 0.16413 \end{bmatrix}$$