Methods in Artificial
Inteligence
TDT4171
Spring 2021

Norwegian University of Science
and Technology
Department for Computer Science

Jørgen Rosager
Exercise 2

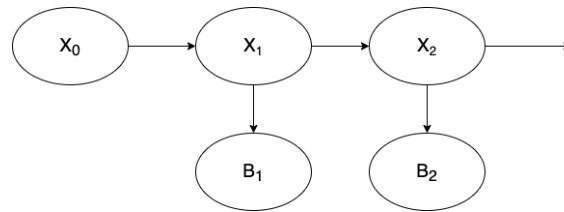1 **a)** The Bayesian network for this problem is shown in figure Figure 1.

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 be 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:

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

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

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.
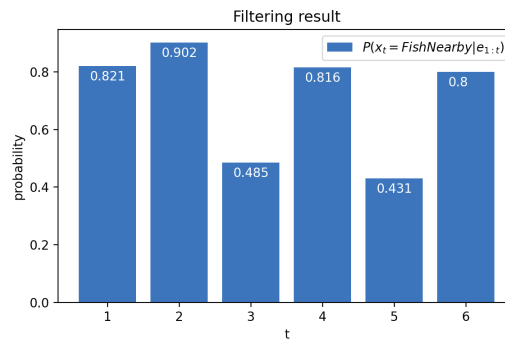

Figure 2: Filtering for $t = 1, ..., 6$.

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.
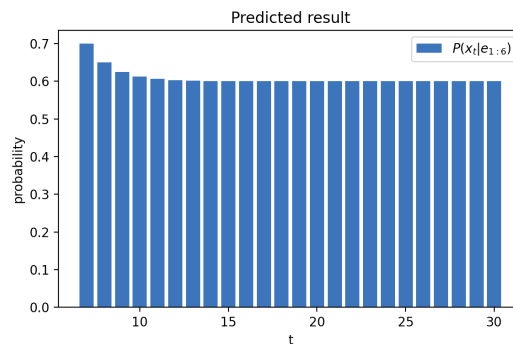

Figure 3: Prediction for $t = 7, ..., 30$.

As we can see, the belief state converges towards $\begin{bmatrix} 0.6 & 0.4 \end{bmatrix}^\top$ as t increases. This is because $\begin{bmatrix} 0.6 & 0.4 \end{bmatrix}^\top$ 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.
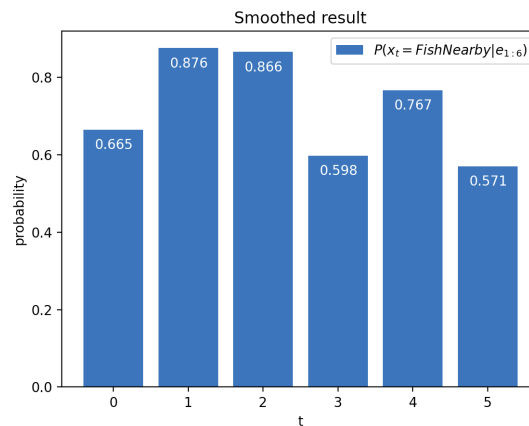

Figure 4: Smoothing for $t = 0, ..., 5$.

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.
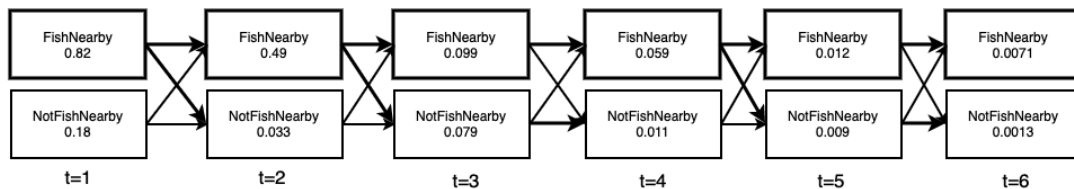

Figure 5: Viterbi sequence for $t = 1, ..., 6$.