Kunnskap for en bedre verden

Department of Computer Science

TDT4171 Metoder i kunstig intelligens

Assignment 2 - Probabilistic Reasoning over Time

Jostein Hjortland Tysse

Table of Contents

1 Hidden Markov Model

a) Modeling

To formulate the problem as a Hidden Markov Model, we need the transition probabilities, observation probabilities, and the initial state probabilities.

Let F_t be the state variable for "There is fish nearby on day t", and B_t be the evidence variable "There are birds nearby on day t".

F_0	$P(F_0)$
$true$	0.5
$false$	0.5

Table 1: Initial Probabilities Table

F_{t-1}	$P(F_t)$
$true$	0.8
$false$	0.3

Table 2: Transition Model Table

F_t	$P(B_t)$
$true$	0.75
$false$	0.2

Table 3: Sensor Model Table

The Transition matrix contains the probabilities T_{ij} for transitioning from state i to j.

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

We have six observations that can be modeled as matrices for convenient calculations. O_i represents the i-th observation.

$$O_1, O_2, O_3, O_4, O_5, O_6 = \begin{bmatrix} 0.75 & 0 \\ 0 & 0.2 \end{bmatrix}, \begin{bmatrix} 0.75 & 0 \\ 0 & 0.2 \end{bmatrix}, \begin{bmatrix} 0.25 & 0 \\ 0 & 0.8 \end{bmatrix}, \begin{bmatrix} 0.75 & 0 \\ 0 & 0.2 \end{bmatrix}, \begin{bmatrix} 0.25 & 0 \\ 0 & 0.8 \end{bmatrix}, \begin{bmatrix} 0.75 & 0 \\ 0 & 0.2 \end{bmatrix}$$

b) Filtering

Compute

$$P(X_t|e_{1:t}), \text{ for } t = 1, \ldots, 6$$

This calculation gives the probability distribution over state t, given all evidence $1, \ldots, t$. This operation is known as *filtering* [1]. To perform filtering, we can use the forward equation from Russel and Norvig [1]. On matrix form, the equation can be stated as

$$f_{1:t+1} = \alpha O_{t+1} T^T f_{1:t}$$

In Figure 1, we can see a plot of the results from the calculations of the Python script. See Figure 6 for more exact numbers.

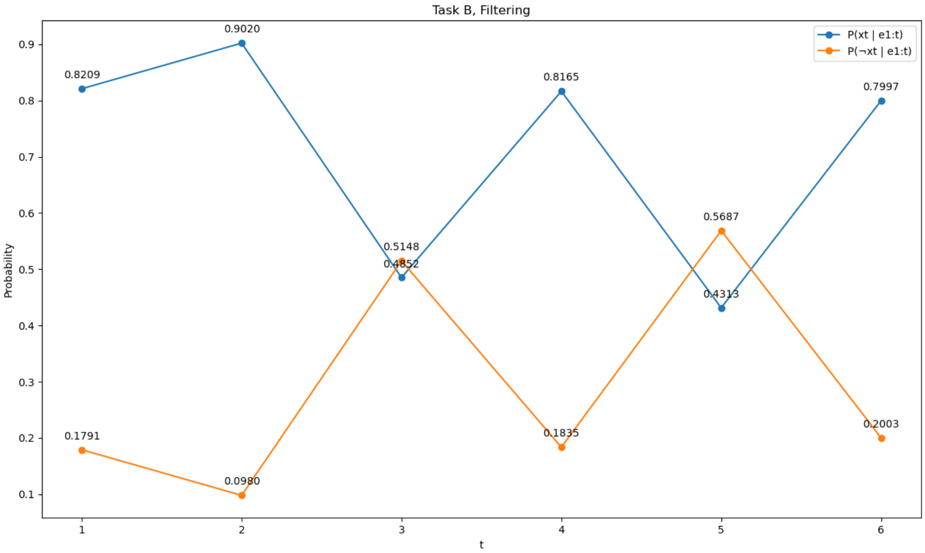

Figure 1: Filtering from 1 to 6

c) Prediction

[H] Compute

$$P(X_t|e_{1:6}),\ \text{for}\ t = 7, \ldots, 30$$

This calculations gives the probability distributions for future states. Calculating future probabilities is known as *prediction* [1]. To calculate the prediction, we can use the filtering approach, just without adding new evidence. As we see from the plot in Figure 2 and the values in Figure 6, the probability distribution converges towards $\langle 0.6, 0.4 \rangle$. This means that as t increases, the probability for fish, assuming we get no more evidence, will get closer and closer to 0.60. If we look at T^T as the stochastic matrix for the problem, the steady state is given by the eigenvector for the eigenvalue $\lambda = 1$. Calculating the eigenvector gives $\begin{bmatrix} 1.5 \\ 1 \end{bmatrix}$, so we have $\alpha \begin{bmatrix} 1.5 \\ 1 \end{bmatrix} = \begin{bmatrix} 0.6 \\ 0.4 \end{bmatrix}$.

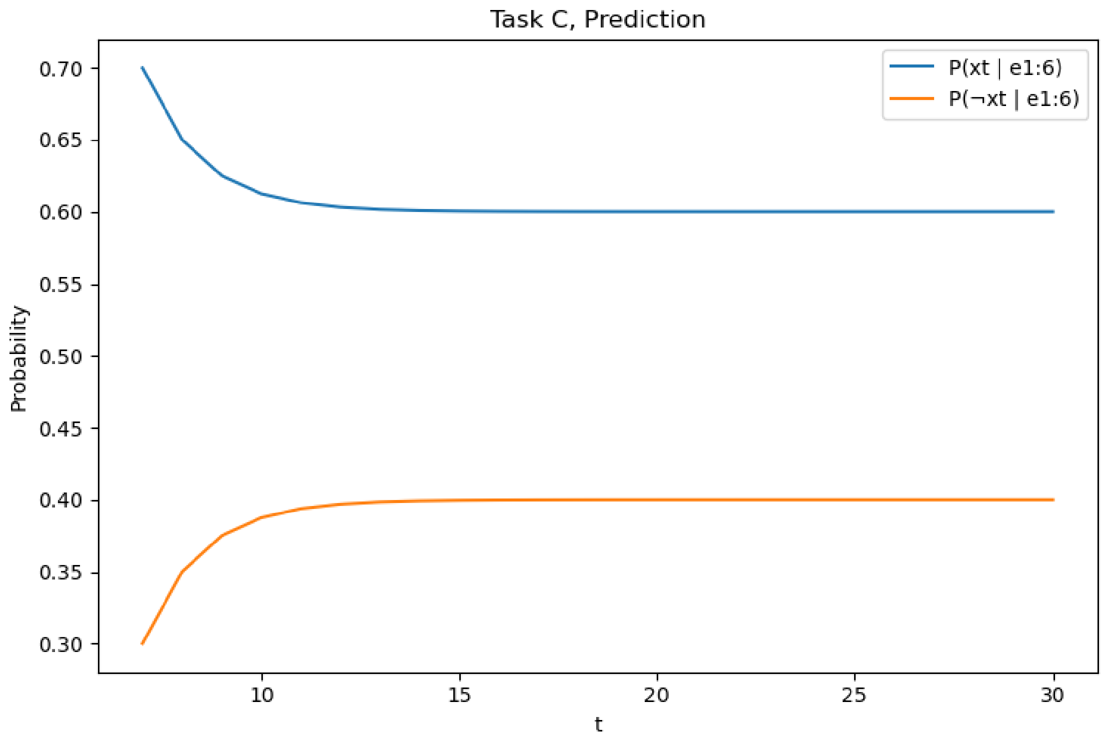

Figure 2: Prediction for t from 7 to 30

d) Smoothing

Compute

$$P(X_t|e_{1:6}), \text{ for } t = 0, \ldots, 5$$

This calculation computes probability distributions over past states, given all evidence to date. Since we have more evidence, we will get a better estimate than we could at the time. The operation is known as *smooting* [1]. To compute the smoothed probabilities, we use a combination of the forward an backwards algorithms frrom Russel and Norvig. The backwords algorithm is given as

$$b_{k+1:t} = TO_{k+1}b_{k+2:t}$$

The smoothing can then be calculated by

$$P(X_k|e_{1:t}) = \alpha f_{1:k} \times b_{k+1:t}$$

We can see the results from smoothing in Figure 3, and more details in Figure 6.

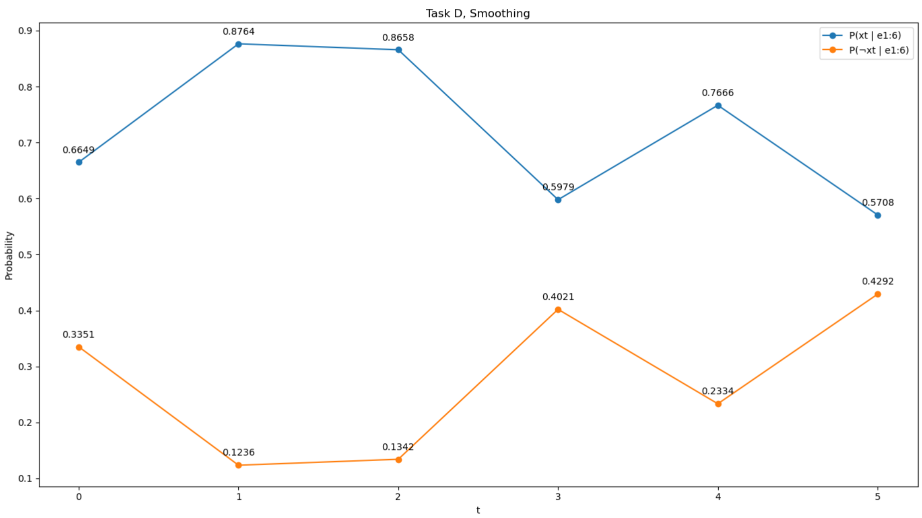

Figure 3: Smoothing from 0 to 5

e) Viterbi

Compute

$$\text{argmax } P(x_1, \ldots, x_{t-1}, X_t | e_{1:t}), \text{ for } t = 1, \ldots, 6$$

This calculations gives the sequence of states most likely to have generated the observed sequence. This is known as finding the *most likely explanation* [1]. To solve the problem, we can use the Viterbi algorithm

Figure 4 shows the calculation graph for the Viterbi algorithm. We can find the paths by using the graph, and follow the bold arrows backwards from the desired state. (As an example, the boxes for True are marked as bold because the most likely sequence for fish at day t is that it was fish in the water every day.)

All the probabilities and paths for all t can be found in Figure 5.

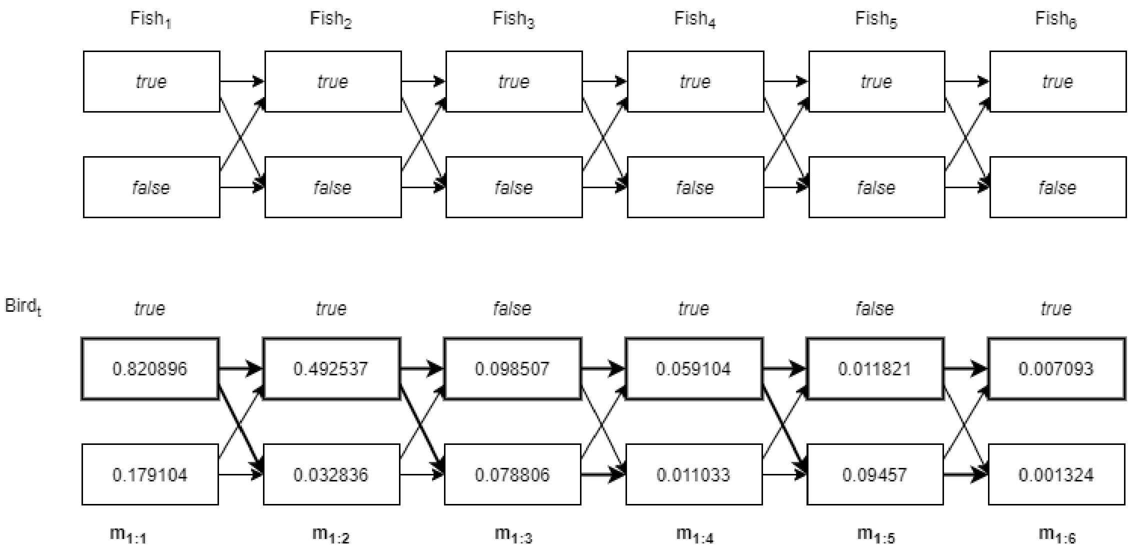

Figure 4: Viterbi algorithm graph

```
Viterbi:
t = 1
Probabilities:
----------
|0.820896|
|0.179104|
----------
Most Likely Paths:
X_t = True: [0]    X_t = False: [1]

t = 2
Probabilities:
-------------------
|0.820896|0.492537|
|0.179104|0.032836|
-------------------
Most Likely Paths:
X_t = True: [0, 0]    X_t = False: [0, 1]

t = 3
Probabilities:
----------------------------
|0.820896|0.492537|0.098507|
|0.179104|0.032836|0.078806|
----------------------------
Most Likely Paths:
X_t = True: [0, 0, 0]    X_t = False: [0, 0, 1]

t = 4
Probabilities:
-------------------------------------
|0.820896|0.492537|0.098507|0.059104|
|0.179104|0.032836|0.078806|0.011033|
-------------------------------------
Most Likely Paths:
X_t = True: [0, 0, 0, 0]    X_t = False: [0, 0, 1, 1]

t = 5
Probabilities:
----------------------------------------------
|0.820896|0.492537|0.098507|0.059104|0.011821|
|0.179104|0.032836|0.078806|0.011033|0.009457|
----------------------------------------------
Most Likely Paths:
X_t = True: [0, 0, 0, 0, 0]    X_t = False: [0, 0, 0, 0, 1]

t = 6
Probabilities:
-------------------------------------------------------
|0.820896|0.492537|0.098507|0.059104|0.011821|0.007093|
|0.179104|0.032836|0.078806|0.011033|0.009457|0.001324|
-------------------------------------------------------
Most Likely Paths:
X_t = True: [0, 0, 0, 0, 0, 0]    X_t = False: [0, 0, 0, 0, 1, 1]
```

Figure 5: Viterbi algorithm all t

```
Filtering, t from 1 to 6
[0.8208955223880596, 0.9019706922688226, 0.48518523354966764, 0.8164592380824628, 0.43134895280917424,
0.7997086306722755]

Prediction from 7 to 30:
[0.6998543153361378, 0.649927157668069, 0.6249635788340345, 0.6124817894170173, 0.6062408947085086, 0.6
031204473542544, 0.6015602236771272, 0.6007801118385636, 0.6003900559192819, 0.600195027959641, 0.60009
75139798206, 0.6000487569899103, 0.6000243784949552, 0.6000121892474776, 0.6000060946237389, 0.60000304
73118695, 0.6000015236559347, 0.6000007618279674, 0.6000003809139837, 0.6000001904569918, 0.60000009522
8496, 0.600000047614248, 0.600000023807124, 0.600000011903562]

Smoothing:
[0.6648521781099611, 0.8764073126577231, 0.8657865721231751, 0.5979273517549293, 0.766637312312745, 0.5
708258150005808]
```

Figure 6: Console output, task b,c,d.

2 Dynamic Bayesian Network

a) Modeling

The transition model, sensor model and prior probabilities are shown in Figure 7. Since all variables are Boolean, we don't need to represent more states, since $P(\neg X) = 1 - P(X)$. The network may continue infinitely (depending on the number of observations), but will always have the same transition and probability models for $t-1$ to t as shown in the figure from 0 to 1. The prior probability $P(A_0)$ only holds for the initial guess at time $t = 0$.

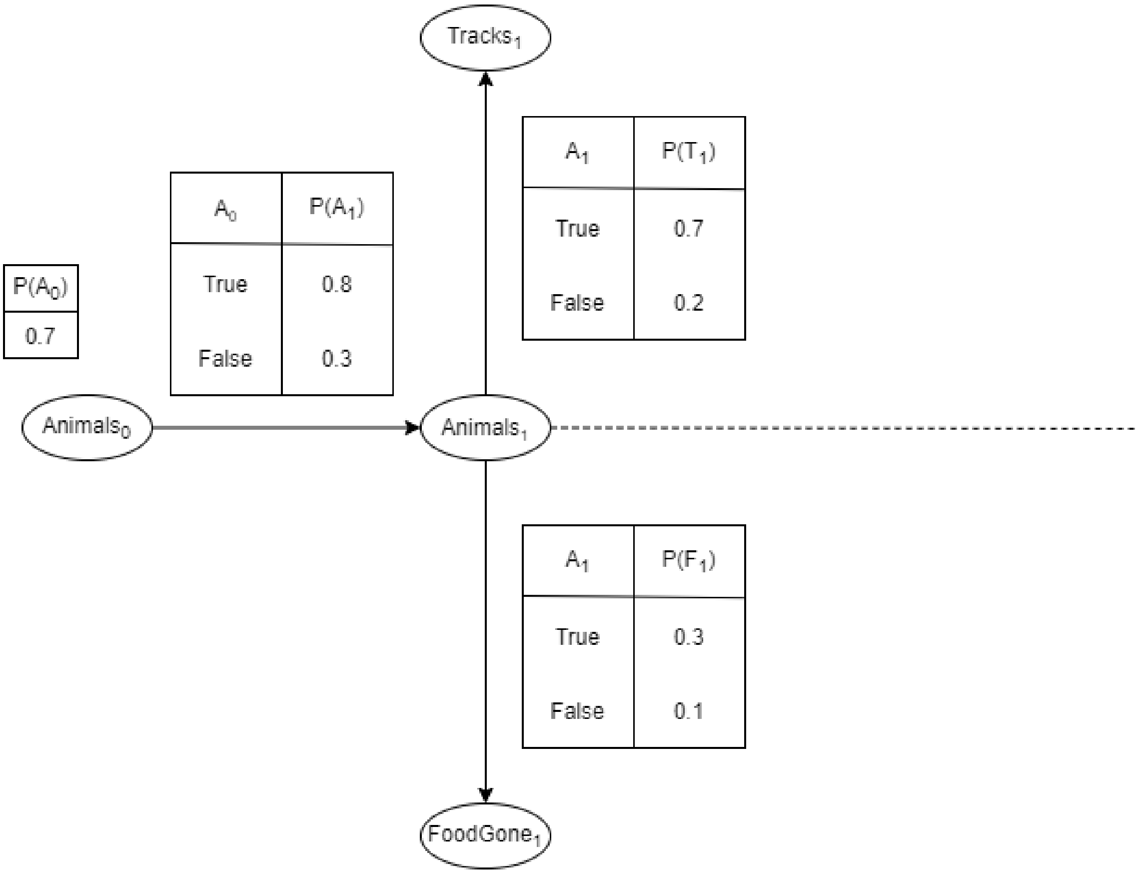

Figure 7: DBN Model with probabilities

According to Russel and Norvig [1], every discrete variable DBN can be represented as a Hidden Markov Model. Let's try to do this, as it will enable us to use the matrix equations to calculate the probabilities for the next parts.

During these parts, I might use $P(X_t)$ and $P(A_t)$ interchangeably, both represent the same thing: the probability distribution for $AnimalsNearby_t$.

Let the transition matrix for $AnimalsNearby$ be given as

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

Let O_{Ti} and O_{Fi} be the observations for $TracksNearby_i$ and $FoodGone_i$, respectively. If we i.e. observe both tracks and that the food is gone on day i, we can represent this at matrices

$$O_{Ti}, O_{Fi} = \begin{bmatrix}0.7 & 0\\ 0 & 0.2\end{bmatrix}, \begin{bmatrix}0.3 & 0\\ 0 & 0.1\end{bmatrix}$$

Let e_i be the evidence on day i. We have two sensors, so we can say that $e_i = O_{Ti} \wedge O_{Fi}$

Let X_i be the probability distributrion for $AnimalsNearby$ on day i.

The probabilities of observing the evidence given the state of animals is then

$$P(e_i|X_i) = P(O_{Ti} \wedge O_{Fi}|X_i)$$

We are given that $(AnimalTracks_t \perp\!\!\!\perp FoodGone_t|AnimalsNearby_t)$.

This implies that

$$P(e_i|X_i) = O_{Ti} \cdot O_{Fi}$$

We are now ready to begin with the calculations. Please note that I have calculated the following values on a hand calculator, so the accuracy of the values will vary due to rounding.

b) Filtering

Compute

$$P(X_t|e_{1:t}), \text{ for } t = 1, \ldots, 4$$

To compute this, we will use the forward equation.

The calculated values are shown in Table 4. All calculations can be found in the following pages.

t	$P(X_t\|e_{1:t})$
1	$\langle 0.95122, 0.04878 \rangle$
2	$\langle 0.79476, 0.20524 \rangle$
3	$\langle 0.40196, 0.59804 \rangle$
4	$\langle 0.73211, 0.26789 \rangle$

Table 4: Filtering probabilities

$e_1 = \{T, F\}$ $e_2 = \{\neg T, F\}$

$e_3 = \{\neg T, \neg F\}$ $e_4 = \{T, \neg F\}$

$T \perp\!\!\!\perp F \mid A$

$t = 1$

$$P(X_1 \mid e_{1:1}) = \alpha\, P(e_{1:1} \mid X_1) \cdot \pi$$

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

$$= \alpha \begin{pmatrix} 0.168 & 0.063 \\ 0.004 & 0.014 \end{pmatrix} \begin{pmatrix} 0.7 \\ 0.3 \end{pmatrix} = \alpha \begin{pmatrix} 0.1365 \\ 0.007 \end{pmatrix}$$

$$\approx \begin{pmatrix} 0.95122 \\ 0.04878 \end{pmatrix}$$

$\underline{t=2}$ $\quad f_{1:1} = \begin{pmatrix} 0.95122 \\ 0.04878 \end{pmatrix}$

$e_2 = \{\neg T,\ F\}$

$$P(X_2 \mid e_{1:2}) = f_{1:2} = \alpha\, O_{T2} \cdot O_{F2} \cdot A^T \cdot f_{1:1}$$

$$= \alpha \begin{pmatrix} 0.3 & 0 \\ 0 & 0.8 \end{pmatrix} \begin{pmatrix} 0.3 & 0 \\ 0 & 0.1 \end{pmatrix} \cdot A \cdot f_{1:1}$$

$$= \alpha \begin{pmatrix} 0.09 & 0 \\ 0 & 0.08 \end{pmatrix} \begin{pmatrix} 0.8 & 0.3 \\ 0.2 & 0.7 \end{pmatrix} \cdot f_{1:1}$$

$$= \alpha \begin{pmatrix} 0.072 & 0.021 \\ 0.016 & 0.056 \end{pmatrix} \begin{pmatrix} 0.95122 \\ 0.04878 \end{pmatrix} \approx \alpha \begin{pmatrix} 0.0695122 \\ 0.0179512 \end{pmatrix}$$

$$\approx \underline{\underline{\begin{pmatrix} 0.79476 \\ 0.20524 \end{pmatrix}}}$$

$\underline{t=3}$ $\qquad f_{1:2} = \begin{pmatrix} 0,79476 \\ 0,20524 \end{pmatrix}$

$e_3 = \{\neg T, \neg F\}$

$$P(X_3 \mid e_{1:3}) = f_{1:3} = \alpha\, O_{T_3} \cdot O_{F_3} \cdot A^T \cdot f_{1:2}$$

$$= \alpha \begin{pmatrix} 0,3 & 0 \\ 0 & 0,8 \end{pmatrix} \begin{pmatrix} 0,7 & 0 \\ 0 & 0,9 \end{pmatrix} \cdot A \cdot f_{1:2}$$

$$= \alpha \begin{pmatrix} 0,21 & 0 \\ 0 & 0,72 \end{pmatrix} \begin{pmatrix} 0,8 & 0,3 \\ 0,2 & 0,7 \end{pmatrix} \cdot f_{1:2}$$

$$= \alpha \begin{pmatrix} 0.168 & 0.063 \\ 0.144 & 0.504 \end{pmatrix} \begin{pmatrix} 0,79476 \\ 0.20524 \end{pmatrix} \approx \alpha \begin{pmatrix} 0,1464498 \\ 0.2178864 \end{pmatrix}$$

$$\approx \underline{\underline{\begin{pmatrix} 0,40196 \\ 0,59804 \end{pmatrix}}}$$

172 0385
163 264496

$\underline{t=4}$ $\qquad f_{1:3} = \begin{pmatrix} 0,40196 \\ 0,59804 \end{pmatrix}$

$e_4 = \{T, \neg F\}$

$$P(X_4 \mid e_{1:4}) = f_{1:4} = \alpha O_{T4} \cdot O_{F4} \cdot A^T \cdot f_{1:3}$$

$$= \alpha \begin{pmatrix} 0,7 & 0 \\ 0 & 0,2 \end{pmatrix} \begin{pmatrix} 0,7 & 0 \\ 0 & 0,9 \end{pmatrix} \cdot A \cdot f_{1:3}$$

$$= \alpha \begin{pmatrix} 0,49 & 0 \\ 0 & 0,18 \end{pmatrix} \begin{pmatrix} 0,8 & 0,3 \\ 0,2 & 0,7 \end{pmatrix} \cdot f_{1:3}$$

$$= \alpha \begin{pmatrix} 0.392 & 0,147 \\ 0.036 & 0,126 \end{pmatrix} \begin{pmatrix} 0,40196 \\ 0,59804 \end{pmatrix} = \alpha \begin{pmatrix} 0.2454802 \\ 0.0898236 \end{pmatrix}$$

$$\approx \underline{\underline{\begin{pmatrix} 0,73211 \\ 0,26789 \end{pmatrix}}}$$

c) Prediction

Compute

$$P(X_t|e_{1:4}), \text{ for } t = 5, \ldots, 8$$

To compute the prediction, we can start with the forward message for $t = 4$, and multiply with the transition matrix. Then we use the result to multiply again with the transition matrix, and so on.

The results are given in Table 5, and the calculations can be found in the following pages.

t	$P(X_t\|e_{1:4})$
5	$\langle 0.666055, 0.333945 \rangle$
6	$\langle 0.6330275, 0.3669725 \rangle$
4	$\langle 0.61651375, 0.38348625 \rangle$
8	$\langle 0.608256875, 0.39174315 \rangle$

Table 5: Prediction Probabilities

$$P(X_t \mid e_{1:4}), \quad t = 5, 6, 7, 8$$

$t = 5$

$$f_{1:4} = \begin{pmatrix} 0.73211 \\ 0.26799 \end{pmatrix}$$

$$P(X_5 \mid e_{1:4}) = A^T \cdot f_{1:4}$$

$$= \begin{pmatrix} 0.8 & 0.3 \\ 0.2 & 0.7 \end{pmatrix} \begin{pmatrix} 0.73211 \\ 0.26789 \end{pmatrix} = \begin{pmatrix} 0.666055 \\ 0.333945 \end{pmatrix}$$

$\underline{t = 6}$

$$P(X_6 \mid e_{1:4}) = \begin{pmatrix} 0.8 & 0.3 \\ 0.2 & 0.7 \end{pmatrix} \begin{pmatrix} 0.666055 \\ 0.333945 \end{pmatrix}$$

$$= \begin{pmatrix} 0.6330275 \\ 0.3669725 \end{pmatrix}$$

$\underline{t=7}$

$$P(X_7 \mid e_{1:4}) = \begin{pmatrix} 0{,}8 & 0{,}3 \\ 0{,}2 & 0{,}7 \end{pmatrix} \begin{pmatrix} 0{,}6330275 \\ 0{,}3669725 \end{pmatrix}$$

$$= \begin{pmatrix} 0{,}61651375 \\ 0{,}38348625 \end{pmatrix}$$

$t=8$

$$P(X_8 \mid e_{1:4}) = \begin{pmatrix} 0{,}8 & 0{,}3 \\ 0{,}2 & 0{,}7 \end{pmatrix} \begin{pmatrix} 0{,}61651375 \\ 0{,}38348625 \end{pmatrix}$$

$$= \begin{pmatrix} 0{,}608256875 \\ 0{,}391743125 \end{pmatrix}$$

d) Proof of convergence

Let x be a column vector representing the probability distribution at a given time. To find the distribution for the next timestep, using only the transition model A, we can do the calculation $A^T x$. A^T is a stochastic matrix where the columns sum to one. When t goes to infinity, the initial distribution does not matter, and the probabilities are given only by the matrix. A stochastic matrix always has $\lambda = 1$ as an eigenvalue, and the eigenvector for this value will give the steady state for the Markov process.

After calculating the eigenvector, we find that the steady state will be $\begin{bmatrix} 0.6 \\ 0.4 \end{bmatrix}$, and we have verified that

$$\lim_{t\to\infty} P(X_t|e_{1:4}) = \langle 0.60, 0.40 \rangle$$

The next two pages shows the calculations for finding eigenvalues and the vector.

$$A^T = \begin{pmatrix} 0,8 & 0,3 \\ 0,2 & 0,7 \end{pmatrix}$$

$$\begin{vmatrix} 0.8-\lambda & 0,3 \\ 0,2 & 0,7-\lambda \end{vmatrix} = (0,8-\lambda)(0,7-\lambda) - 0.06 = 0$$

$$\lambda^2 - 1,5\lambda + 0,56 - 0.06 = 0$$

$$\lambda^2 - 1,5\lambda + 0,5 = 0$$

$$\lambda = \frac{1,5 \pm \sqrt{2,24 - 2}}{2} = \frac{1,5 \pm 0,5}{2}$$

$$\lambda_1 = 1 \quad , \quad \lambda_2 = 0,5$$

$$\begin{pmatrix} 0,8-1 & 0,3 \\ 0,2 & 0,7-1 \end{pmatrix} \sim \begin{pmatrix} -0,2 & 0,3 \\ 0,2 & -0,3 \end{pmatrix}$$

$$\sim \begin{pmatrix} -2 & 3 \\ 0 & 0 \end{pmatrix} = \vec{0}$$

$$x = \begin{pmatrix} 3 \\ 2 \end{pmatrix}$$

$$\alpha \cdot \begin{pmatrix} 3 \\ 2 \end{pmatrix} = \underline{\underline{\begin{pmatrix} 0,6 \\ 0.4 \end{pmatrix}}}$$

e) Smoothing

Compute

$$P(X_t|e_{1:4}), \text{ for } t = 0, \dots, 3$$

To calculate the smoothing, we will use both forward and backwards messages. The following pages show how I first calculate the backwards messages, and then multiply them with the forward messages from task a) to get the smoothing values. Table 6 shows the calculated smoothed probabilities.

t	$P(X_t|e_{1:4})$
0	$\langle 0.8358720341, 0.1641279659 \rangle$
1	$\langle 0.94439, 0.05561 \rangle$
2	$\langle 0.7235388, 0.2764612 \rangle$
3	$\langle 0.51308, 0.48692 \rangle$

Table 6: Smoothing Probabilities

$P(X_t \mid e_{1:4})$, $t = 0,1,2,3$

SMOOTH

$$P(X_t \mid e_{1:4}) = \alpha \, f_{1:t} \times b_{t+1:4}$$

$$b_{k+1:t} = T O_{k+1} b_{k+2:t}$$

First, calculate Backward.

$$A = \begin{pmatrix} 0.8 & 0.2 \\ 0.3 & 0.7 \end{pmatrix}$$

$$b_{5:4} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

$$b_{4:4} = A \cdot O_{T4} \cdot O_{F4} \cdot b_{5:4}$$

$$= A \cdot \begin{pmatrix} 0.7 & 0 \\ 0 & 0.2 \end{pmatrix} \begin{pmatrix} 0.7 & 0 \\ 0 & 0.9 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

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

$$= \begin{pmatrix} 0{,}392 & 0.036 \\ 0.147 & 0{,}126 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0{,}428 \\ 0{,}273 \end{pmatrix}$$

$$b_{3:4} = A \cdot O_{T3} \cdot O_{F3} \cdot b_{4:4}$$

$$= A \cdot \begin{pmatrix} 0,3 & 0 \\ 0 & 0,8 \end{pmatrix} \begin{pmatrix} 0.7 & 0 \\ 0 & 0.9 \end{pmatrix} \begin{pmatrix} 0,428 \\ 0,273 \end{pmatrix}$$

$$= \begin{pmatrix} 0.8 & 0.2 \\ 0,3 & 0.7 \end{pmatrix} \cdot \begin{pmatrix} 0.21 & 0 \\ 0 & 0.72 \end{pmatrix} \begin{pmatrix} 0,428 \\ 0,273 \end{pmatrix}$$

$$= \begin{pmatrix} 0,168 & 0,144 \\ 0.063 & 0,504 \end{pmatrix} \begin{pmatrix} 0,428 \\ 0,273 \end{pmatrix} = \begin{pmatrix} 0.111216 \\ 0,164556 \end{pmatrix}$$

$$b_{2:4} = A \cdot O_{T2} \cdot O_{F2} \cdot b_{3:4}$$

$$= A \cdot \begin{pmatrix} 0{,}3 & 0 \\ 0 & 0{,}8 \end{pmatrix} \begin{pmatrix} 0.3 & 0 \\ 0 & 0{,}1 \end{pmatrix} \begin{pmatrix} 0.111216 \\ 0{,}164556 \end{pmatrix}$$

$$= \begin{pmatrix} 0.8 & 0.2 \\ 0{,}3 & 0{,}7 \end{pmatrix} \cdot \begin{pmatrix} 0{,}09 & 0 \\ 0 & 0.08 \end{pmatrix} \begin{pmatrix} 0.111216 \\ 0{,}164556 \end{pmatrix}$$

$$= \begin{pmatrix} 0{,}072 & 0{,}016 \\ 0.027 & 0.056 \end{pmatrix} \begin{pmatrix} 0.111216 \\ 0{,}164556 \end{pmatrix} \simeq \begin{pmatrix} 0.0106404 \\ 0{,}012217968 \end{pmatrix}$$

$$b_{1:4} = A \cdot O_{T1} \cdot O_{F1} \cdot b_{2:4}$$

$$= A \cdot \begin{pmatrix} 0.7 & 0 \\ 0 & 0.2 \end{pmatrix} \begin{pmatrix} 0.3 & 0 \\ 0 & 0.1 \end{pmatrix} \begin{pmatrix} 0.0106404 \\ 0.012217968 \end{pmatrix}$$

$$= \begin{pmatrix} 0.8 & 0.2 \\ 0.3 & 0.7 \end{pmatrix} \cdot \begin{pmatrix} 0.21 & 0 \\ 0 & 0.02 \end{pmatrix} \begin{pmatrix} 0.0106404 \\ 0.012217968 \end{pmatrix}$$

$$= \begin{pmatrix} 0.168 & 0.004 \\ 0.063 & 0.014 \end{pmatrix} \begin{pmatrix} 0.0106404 \\ 0.012217968 \end{pmatrix} = \begin{pmatrix} 0.00183645907 2 \\ 0.000841396752 \end{pmatrix}$$

$$t = 0$$

$$P(X_0 \mid e_{1:4}) = \alpha \cdot \pi \times b_{1:4}$$

$$= \alpha \cdot \begin{pmatrix} 0{,}7 \\ 0{,}3 \end{pmatrix} \times \begin{pmatrix} 0{,}00183645907 2 \\ 0{,}000841396750 \end{pmatrix}$$

$$= \alpha \begin{pmatrix} 0{,}00128552135 \\ 0{,}0002524190256 \end{pmatrix}$$

$$= \begin{pmatrix} 0{,}8358720341 \\ 0{,}1641279659 \end{pmatrix}$$

$t = 1$

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

$$= \alpha \cdot \begin{pmatrix} 0.95122 \\ 0.04878 \end{pmatrix} \begin{pmatrix} 0.0106404 \\ 0.012217968 \end{pmatrix}$$

$$\simeq \alpha \begin{pmatrix} 0.010121361129 \\ 0.0005959924479 \end{pmatrix}$$

$$= \begin{pmatrix} 0.94439 \\ 0.05561 \end{pmatrix}$$

$t = 2$

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

$$= \alpha \cdot \begin{pmatrix} 0.79476 \\ 0.20524 \end{pmatrix} \times \begin{pmatrix} 0.111216 \\ 0.164556 \end{pmatrix}$$

$$\approx \alpha \begin{pmatrix} 0.08839002816 \\ 0.03377347344 \end{pmatrix}$$

$$\approx \begin{pmatrix} 0.7235388 \\ 0.2764612 \end{pmatrix}$$

$t = 3$

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

$$= \alpha \begin{pmatrix} 0.40196 \\ 0.59804 \end{pmatrix} \times \begin{pmatrix} 0.428 \\ 0.273 \end{pmatrix}$$

$$\approx \alpha \begin{pmatrix} 0.17203888 \\ 0.16326492 \end{pmatrix}$$

$$\approx \begin{pmatrix} 0.51308 \\ 0.48692 \end{pmatrix}$$

References

[1] Stuart Russel and Peter Norvig. *Artificial Intelligence, A Modern Approach*. 3. Pearson Education, inc., 2010.