Exercise 2: TDT4171 Artificial Intelligence Methods

Jim Totland

February 2021

1 Hidden Markov Model

a)

Let $X_t \in \{0, 1\}$ represent the state varibale, where $X_t = 0$ corresponds to no fish in the lake at time t, and $X_t = 1$ corresponds to fish in the lake at time t. Let also $\mathbf{e}_t \in \{0, 1\}$ represent the evidence at time t, where $\mathbf{e}_t = 0$ corresponds to no birds nearby at time t and $\mathbf{e}_t = 1$ corresponds to birds nearby at time t. We assume that X_t is a Markov process whose state is not directly observable, and that $P(\mathbf{e}_t|X_t = x_t, \ldots, X_0 = x_0) = P(\mathbf{e}_t|X_t = x_t)$. The hidden Markov model can be fully specified by the transition matrix and the sensor model. The transition matrix is given by

$$\mathbf{T} = \begin{pmatrix} 0.7 & 0.3 \\ 0.2 & 0.8 \end{pmatrix} \tag{1}$$

where $\mathbf{T}_{ij} = P(X_{t+1} = j - 1|X_t = i - 1)$. The sensor model can be represented by two matrices,

$$\mathbf{O}_0 = \begin{pmatrix} 0.8 & 0 \\ 0 & 0.25 \end{pmatrix} \quad \mathbf{O}_1 = \begin{pmatrix} 0.2 & 0 \\ 0 & 0.75 \end{pmatrix} \tag{2}$$

where $\mathbf{O}_i$ corresponds to $\mathbf{e}_t = i$ and contains on the j'th diagonal entry $P(\mathbf{e}_t|X_t = j - 1)$.

b)

The results from my implementation are given in the table below

t	1	2	3	4	5	6
$P(X_t = true \mid e_{1:t})$	0.8209	0.9020	0.4852	0.8165	0.4313	0.7997
$P(X_t = false \mid e_{1:t})$	0.1791	0.0980	0.5148	0.1835	0.5687	0.2003

$$P(X_t|e_{1:t}) = \langle 0.654, 0.346 \rangle. \tag{3}$$

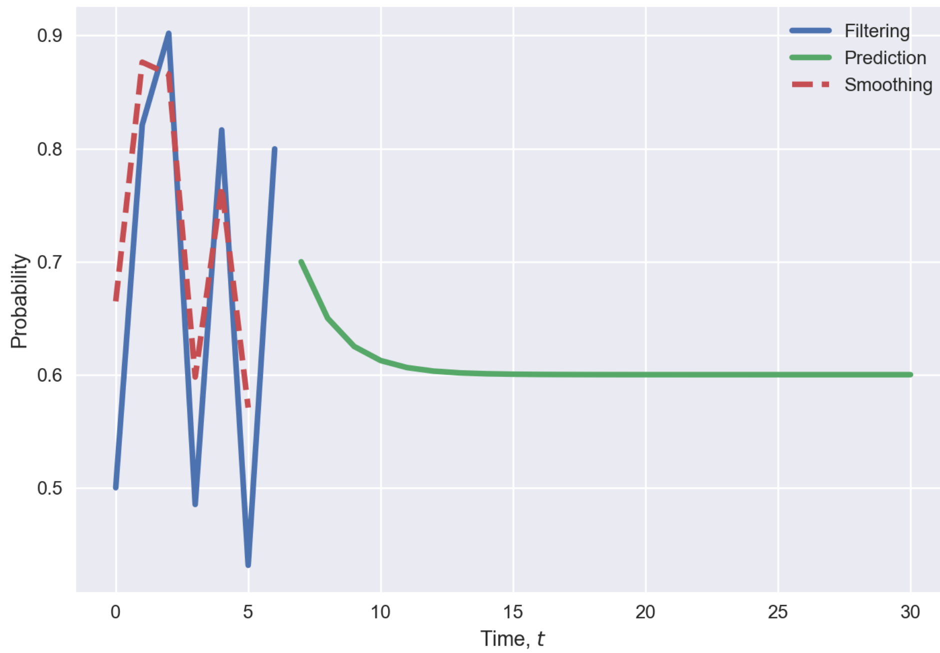


Figure 1: The results from filtering (blue), prediction (green) and smoothing (green) plotted against time, t.

This operation is called 'filtering' and calculates X's distribution at time t, based on the evidence up to that point in time. The results of the filtering can also be viewed graphically as the blue graph in figure 1.

c)

Some of the results from my implementation are given in the table below:

t	7	8	9	$\cdots$	29	30
$P(X_t = true \mid e_{1:6})$	0.6999	0.6499	0.6250	$\cdots$	0.6000	0.6000
$P(X_t = false \mid e_{1:6})$	0.3001	0.3501	0.3750	$\cdots$	0.4000	0.4000

This operation is called 'prediction' and estimated the future distributions of our state variable, given past and current evidence. The results are illustrated graphically as the green graph in figure 1. From the data and the graph, we clearly see that $\mathbf{P}(X_t|e_{1:6})$ for $t > 6$ approaches a fixed point $\langle 0.6, 0.4 \rangle$. This is called the stationary distribution of X_t.

d)

The results from my implementation are given below:

t	0	1	2	3	4	5
$P(X_t = true \mid e_{1:6})$	0.6649	0.8764	0.8658	0.5979	0.7666	0.5708
$P(X_t = false \mid e_{1:6})$	0.3351	0.1236	0.1342	0.4021	0.2334	0.4292

This operation is called 'smoothing' and estimates a more accurate distribution of the state variable in previous states, given new evidence. The results are plotted in 1 as the red, dashed graph.

e)

This operation is called 'most likely sequence', and yields the sequence of previous states that is most likely to produce the state at time t. The results from my implementations are found in the tables below. For each value of X_t, the most likely path to that state, $[x_1, \ldots, x_{t-1}]$ is given. The most likely value of X_t is indicated in bold typeface.

<u>$t = 1$:</u>

	Most likely sequence
$\boldsymbol{X_1 = true}$	
$X_1 = false$	

<u>$t = 2$:</u>

	Most likely sequence
$\boldsymbol{X_2 = true}$	$[true]$
$X_2 = false$	$[true]$

<u>$t = 3$:</u>

	Most likely sequence
$\boldsymbol{X_3 = true}$	$[true, true]$
$X_3 = false$	$[true, true]$

<u>$t = 4$:</u>

	Most likely sequence
$\boldsymbol{X_4 = true}$	$[true, true, true]$
$X_4 = false$	$[true, true, false]$

<u>$t = 5$:</u>

	Most likely sequence
$\boldsymbol{X_5 = true}$	$[true, true, true, true]$
$X_5 = false$	$[true, true, false, true]$

<u>$t = 6$:</u>

	Most likely sequence
$\boldsymbol{X_6 = true}$	$[true, true, true, true, true]$
$X_6 = false$	$[true, true, false, true, false]$

$\boxed{2}$ Dynamic Bayesian Networks

$AT_t \in \{\text{true} \Leftrightarrow \text{animal tracks}, \ \text{false} \Leftrightarrow \text{no animal tracks}\}$

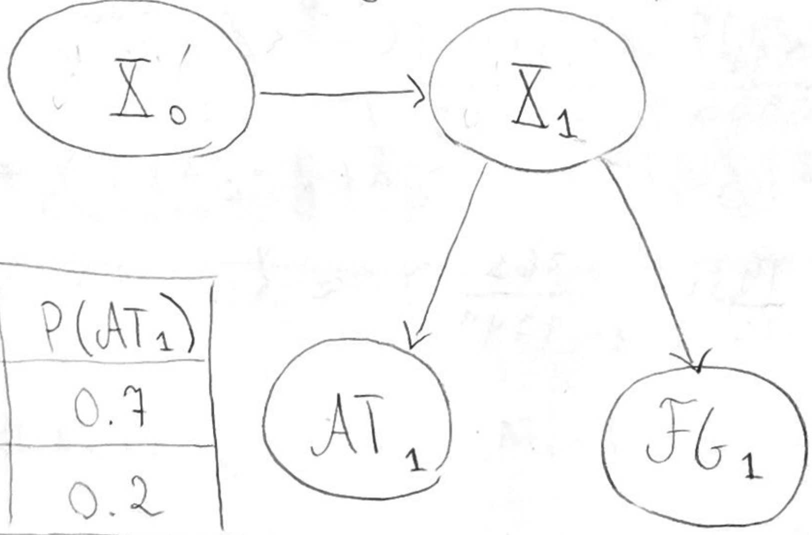
$FG_t \in \{\text{true} \Leftrightarrow \text{food gone}, \ \text{false} \Leftrightarrow \text{food not gone}\}$

$X_t \in \{\text{true} \Leftrightarrow \text{animals nearby}, \ \text{false} \Leftrightarrow \text{no animals nearby}\}$

a)

$P(X_0)$
0.7

X_0	$P(X_1)$
t	0.8
f	0.3



X_1	$P(AT_1)$
t	0.7
f	0.2

X_1	$P(FG_1)$
t	0.3
f	0.1

Above, the prior, transition and sensor model are specified for the DBN. I have defined $AT_t :=$ AnimalTracks$_t$, $FG_t :=$ FoodGone$_t$ and $X_t :=$ AnimalsNearby t.

b)

$$P(X_1 \mid e_1) = P(X_1 \mid \{AT_1 = t, FG_1 = t\})$$

$$= \alpha\, P(\{AT_1 = t, FG_1 = t\} \mid X_1) \sum_{x_0} P(X_1 \mid x_0) \cdot P(x_0)$$

$$= \alpha \langle 0.7 \cdot 0.3,\ 0.2 \cdot 0.1 \rangle \left[\langle 0.8, 0.2 \rangle \cdot 0.7 + \langle 0.3, 0.7 \rangle \cdot 0.3 \right]$$

$$\boxed{AT_t \perp\!\!\!\perp FG_t \mid X_t}$$

$$= \alpha \cdot \langle 0.21, 0.02 \rangle \cdot \langle 0.65, 0.35 \rangle$$

$$= \alpha \langle 0.1365, 0.007 \rangle$$

$$\Rightarrow f_1 := \langle 0.9512.., 0.049 \rangle = \left\langle \frac{39}{41}, \frac{2}{41} \right\rangle$$

b)

ii) $P(X_2 | e_{1:2}) = \alpha \, P(\{AT_2 = f, FG_2 = t\} | X_2)$

$\cdot \sum_{x_1} P(X_2 | x_1) \cdot P(x_1 | e_1)$

$(AT_t \perp\!\!\!\perp FG_t | X_t)$

$= \alpha \, P(AT_2 = f | X_2) \cdot P(FG_2 = t | X_2)$

$\cdot \left[\langle 0.8, 0.2 \rangle \cdot \frac{39}{41} + \langle 0.3, 0.7 \rangle \cdot \frac{2}{41} \right]$

$= \alpha \, \langle 0.3, 0.8 \rangle \langle 0.3, 0.1 \rangle \left[\langle \frac{159}{205}, \frac{46}{205} \rangle \right]$

$= \alpha \, \langle \frac{1431}{20500}, \frac{92}{5125} \rangle$

$\Rightarrow \quad f_{1:2} = \langle \frac{1431}{1799}, \frac{368}{1799} \rangle \approx \langle 0.7954, 0.2046 \rangle$

iii) $P(X_3 | e_{1:3}) = \alpha \, P(\{AT_3 = f, FG_3 = f\} | X_3)$

$\cdot \sum_{x_2} P(X_3 | x_2) \cdot P(x_2 | e_2)$

$= \alpha \, P(AT_3 = f | X_3) \cdot P(FG_3 = f | X_3) \cdot \left[\langle 0.8, 0.2 \rangle \cdot \frac{1431}{1799} \right.$

$\left. + \langle 0.3, 0.7 \rangle \cdot \frac{368}{1799} \right]$

$$= \alpha \langle 0.3, 0.8 \rangle \langle 0.7, 0.9 \rangle \left[\left\langle \frac{6276}{8995}, \frac{2719}{8995} \right\rangle \right]$$

$$= \alpha \left\langle \frac{4707}{32125}, 0.21764 \right\rangle$$

$$\Rightarrow \quad f_{1:3} = \langle 0.4024, 0.5976 \rangle$$

Explaination

$$\boxed{\begin{aligned} P(AT_3 = f \mid X_3) &= P(AT_3 = f, X_3) \cdot P(X_3) \\ &= \langle P(AT_3 = f, X_3 = t) \cdot P(X_3 = t), P(AT_3 = f, X_3 = f) \cdot P(X_3 = f) \rangle \\ &= \langle P(AT_3 = f \mid X_3 = t), P(AT_3 = f \mid X_3 = f) \rangle \end{aligned}}$$

iv) $P(X_4 \mid e_{1:4}) = \alpha P(\{AT_4 = t, FG_4 = f\} \mid X_4) \cdot \sum_{x_3} P(X_4 \mid x_3) \cdot P(x_3 \mid e_3)$

$$= \alpha \langle 0.7, 0.2 \rangle \langle 0.7, 0.9 \rangle \cdot \left[\langle 0.8, 0.2 \rangle \cdot \frac{4707}{32125} + \langle 0.3, 0.7 \rangle \cdot 0.21764 \right]$$

$$= \alpha \left\langle \frac{49}{100}, \frac{9}{50} \right\rangle [\langle 0.18251, 0.18165 \rangle]$$

$$= \alpha \langle 0.08943, 0.032697 \rangle$$

$$\Rightarrow \quad f_{1:4} = \langle 0.7323, 0.2677 \rangle$$

c)

i) $P(X_5 | e_{1:4}) = \sum_{x_4} P(X_5 | x_4) \cdot P(x_4 | e_{1:4})$

$= \langle 0.8, 0.2 \rangle \cdot 0.7323 + \langle 0.3, 0.7 \rangle \cdot 0.2677$

from c)

$= \underline{\underline{\langle 0.66615, 0.33385 \rangle}}$

ii) $P(X_6 | e_{1:4}) = \sum_{x_5} P(X_6 | x_5) \cdot P(x_5 | e_{1:4})$

$= \langle 0.8, 0.2 \rangle \cdot 0.66615 + \langle 0.3, 0.7 \rangle \cdot 0.33385$

$= \underline{\underline{\langle 0.633075, 0.366925 \rangle}}$

iii) $P(X_7 | e_{1:4}) = \sum_{x_6} P(X_7 | x_6) \cdot P(x_6 | e_{1:4})$

$= \langle 0.8, 0.2 \rangle \cdot 0.633075 + \langle 0.3, 0.7 \rangle \cdot 0.366925$

$= \underline{\underline{\langle 0.6165375, 0.3834625 \rangle}}$

iv)

$$P(X_8|e_{1:4}) = \sum_{x_7} P(X_8|x_7) \cdot P(x_7|e_{1:4})$$

$$= \langle 0.8, 0.2 \rangle \cdot 0.6165375 + \langle 0.3, 0.7 \rangle \cdot 0.3834625$$

$$= \underline{\underline{\langle 0.60826875, 0.39173125 \rangle}}$$

d)

Assume $P(X_k|e_{1:4}) = \langle 0.6, 0.4 \rangle$ for some $k \geq 4$. Then

$$P(X_{k+1}|e_{1:4}) = \langle 0.8, 0.2 \rangle \cdot 0.6 + \langle 0.3, 0.7 \rangle \cdot 0.4$$

$$= \underline{\underline{\langle 0.6, 0.4 \rangle}}$$

$\Rightarrow \langle 0.6, 0.4 \rangle$ is the stationary distribution e)

$$P(X_0 | e_{1:4}) = \alpha P(X_0) \cdot \underbrace{P(e_{1:4} | X_0)}_{b_{1:4}}$$

Must calculate:

$$b_{1:4} = \sum_{x_1} P(e_1 | x_1) \cdot P(e_{2:4} | x_1) \cdot P(x_1 | X_0)$$

Must go "backwards" because of second factor

i)

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

$$b_{5:4} := \mathbb{1}$$

$$b_{4:4} = P(e_{4:} | X_3) = \sum_{x_4} P(e_4 | x_4) \cdot \mathbb{1} \cdot P(x_4 | X_3)$$

$$= \sum_{x_4} P(AT_4 = t | x_4) \cdot P(FG_4 = f | x_4) \cdot \mathbb{1} \cdot P(x_4 | X_3)$$

conditionally independent, given X_4

$$= 0.7 \cdot 0.7 \cdot \langle 0.8, 0.3 \rangle + 0.2 \cdot 0.9 \cdot \langle 0.2, 0.7 \rangle$$

$$= \left\langle \frac{107}{250}, \frac{273}{1000} \right\rangle \Rightarrow P(X_3 | e_{1:4}) = \alpha \langle 0.1722272, 0.1631448 \rangle$$

$$\Rightarrow P(X_3 | e_{1:4}) = \underline{\underline{\langle 0.5135, 0.4865 \rangle}}$$

ii)

$$P(X_2|e_{1:4}) = \alpha f_{1:2} \times b_{3:4} \quad . \quad f_{1:2} = \left\langle \frac{1431}{1799}, \frac{368}{1799} \right\rangle$$

$$b_{3:4} = P(e_{3:4}|X_2) = \sum_{x_3} P(e_3|x_3) \cdot P(e_4|x_3) \cdot P(x_3|X_2)$$

$$= 0.3 \cdot 0.7 \cdot \frac{107}{250} \langle 0.8, 0.3 \rangle + 0.8 \cdot 0.9 \cdot \frac{273}{1000} \cdot \langle 0.2, 0.7 \rangle$$

$$= \left\langle \frac{6951}{62500}, 0.164556 \right\rangle$$

$$\Rightarrow P(X_2|e_{1:4}) = \alpha \langle 0.08847, 0.03366 \rangle$$

$$\Rightarrow P(X_2|e_{1:4}) = \underline{\underline{\langle 0.7244, 0.2756 \rangle}}$$

iii) $P(X_1|e_{1:4}) = \alpha f_1 \times b_{2:4}$. $f_1 = \left\langle \frac{39}{41}, \frac{2}{41} \right\rangle$

$$b_{2:4} = P(e_{2:4}|X_1) = \sum_{x_2} P(e_2|x_2) \cdot P(e_3|x_2) \cdot P(x_2|X_1)$$

$$= 0.3 \cdot 0.3 \cdot \frac{6951}{62500} \langle 0.8, 0.3 \rangle + 0.8 \cdot 0.1 \cdot 0.164556 \cdot \langle 0.2, 0.7 \rangle$$

$$= \langle 0.010640, 0.012218 \rangle$$

$$\Rightarrow P(X_1|e_{1:4}) = \alpha \langle 0.010121, 0.000596 \rangle$$

$$\Rightarrow \quad P(X_1 | e_{1:4}) = \underline{\underline{\langle 0.9444, 0.0556 \rangle}}$$

iv)

$$P(X_0 | e_{1:4}) = \alpha P(X_0) \times b_{1:4}. \quad P(X_0) = \langle 0.7, 0.3 \rangle$$

$$b_{1:4} = P(e_{1:4} | X_1) = \sum_{x_1} P(e_1 | x_1) \cdot P(e_{2:4} | x_1) P(x_1 | X_0)$$

$$= 0.7 \cdot 0.3 \cdot 0.010640 \cdot \langle 0.8, 0.3 \rangle + 0.2 \cdot 0.1 \cdot 0.012218 \cdot \langle 0.2, 0.7 \rangle$$

$$= \langle 1.8364 \cdot 10^{-3}, 8.4137 \cdot 10^{-4} \rangle$$

$$\Rightarrow \quad P(X_0 | e_{1:4}) = \alpha \langle 1.28548 \cdot 10^{-3}, 2.52411 \cdot 10^{-4} \rangle$$

$$\Rightarrow \quad P(X_0 | e_{1:4}) = \underline{\underline{\langle 0.8359, 0.1641 \rangle}}$$