TDT4171 Artificial Intelligence Methods
Assignment 2

Danilas Miscenko

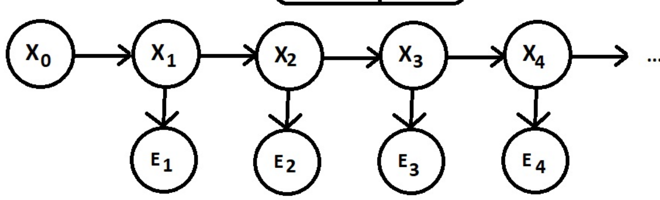
February 17, 2021

1 Hidden Markov Model

a) Formulate the information given above as a hidden Markov model, and provide the completeprobability tables for the model.

P(X0)	
x	0.5
¬x	0.5

$P(x_t \mid x_{t-1})$	0.8
$P(\neg x_t \mid x_{t-1})$	0.2
$P(x_t \mid \neg x_{t-1})$	0.3
$P(\neg x_t \mid \neg x_{t-1})$	0.7



$P(e_t \mid x_t)$	0.75
$P(\neg e_t \mid x_t)$	0.25
$P(e_t \mid \neg x_t)$	0.2
$P(\neg e_t \mid \neg x_t)$	0.8

Figure 1: Hidden Markov Model

b) Compute

$$P(X_t \mid e_{1:t}), \textbf{ for } t = 0, \ldots, 6$$

What kind of operation is this (filtering, prediction, smoothing, likelihood of the evidence, or most likely sequence)? Describe in words what kind of information this operation provides us.

This is a filtering operation. It provides us with a guess of the current state corrected by the evidence we can gather in, so the inference of the state probabilities given the evidence we know about the current state and all previous evidence.

```
Task 1 b)
For t=1
        Probability for P(X_1|e_1:1)=[0.78571429 0.21428571]
For t=2
        Probability for P(X_2|e_1:2)=[0.87125749 0.12874251]
For t=3
        Probability for P(X_3|e_1:3)=[0.41025213 0.58974787]
For t=4
        Probability for P(X_4|e_1:4)=[0.75382494 0.24617506]
For t=5
        Probability for P(X_5|e_1:5)=[0.34373906 0.65626094]
For t=6
        Probability for P(X_6|e_1:6)=[0.72829148 0.27170852]
```

Figure 2: Printout in terminal. Comments on the code in the code.

c) Compute

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

What kind of operation is this (filtering, prediction, smoothing, likelihood of the evidence, or most likely sequence)? Describe in words what kind of information this operation provides us. What happens to the distribution in Equation b) as it increases?

This is a prediction operation. Since we don't have any evidence for the next 23 states, we can only guess what the probabilities of those next 14 states are given the evidence that we have gotten.

```
Task 1 c)
For t=7
        Probability for P(X_7|e_1:4)=[0.66414574 0.33585426]
For t=8
        Probability for P(X_8|e_1:4)=[0.63207287 0.36792713]
For t=9
        Probability for P(X_9|e_1:4)=[0.61603643 0.38396357]
For t=10
        Probability for P(X_10|e_1:4)=[0.60801822 0.39198178]
For t=11
        Probability for P(X_11|e_1:4)=[0.60400911 0.39599089]
For t=12
        Probability for P(X_12|e_1:4)=[0.60200455 0.39799545]
For t=13
        Probability for P(X_13|e_1:4)=[0.60100228 0.39899772]
For t=14
        Probability for P(X_14|e_1:4)=[0.60050114 0.39949886]
For t=15
        Probability for P(X_15|e_1:4)=[0.60025057 0.39974943]
For t=16
        Probability for P(X_16|e_1:4)=[0.60012528 0.39987472]
For t=17
        Probability for P(X_17|e_1:4)=[0.60006264 0.39993736]
For t=18
        Probability for P(X_18|e_1:4)=[0.60003132 0.39996868]
For t=19
        Probability for P(X_19|e_1:4)=[0.60001566 0.39998434]
For t=20
        Probability for P(X_20|e_1:4)=[0.60000783 0.39999217]
For t=21
        Probability for P(X_21|e_1:4)=[0.60000392 0.39999608]
For t=22
        Probability for P(X_22|e_1:4)=[0.60000196 0.39999804]
For t=23
        Probability for P(X_23|e_1:4)=[0.60000098 0.39999902]
For t=24
        Probability for P(X_24|e_1:4)=[0.60000049 0.39999951]
For t=25
        Probability for P(X_25|e_1:4)=[0.60000024 0.39999976]
For t=26
        Probability for P(X_26|e_1:4)=[0.60000012 0.39999988]
For t=27
        Probability for P(X_27|e_1:4)=[0.60000006 0.39999994]
For t=28
        Probability for P(X_28|e_1:4)=[0.60000003 0.39999997]
For t=29
        Probability for P(X_29|e_1:4)=[0.60000002 0.39999998]
For t=30
        Probability for P(X_30|e_1:4)=[0.60000001 0.39999999]
```

Figure 3: Printout in terminal. Comments on the code in the code.

d) Compute

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

What kind of operation is this (filtering, prediction, smoothing, likelihood of the evidence, or most likely sequence)? Describe in words what kind of information this operation provides us.

This is a smoothing operation, and it provides us with the probabilities of past states given the evidence we have now, so using what we have now to sharpen our sences about the past.

e) Compute

$$\arg\max_{x1,\ldots,x_{t-1}} P(x_1, \ldots, x_{t-1}, X_t \mid e_{1:t}), \textbf{ for } t = 1, \ldots, 6$$

What kind of operation is this (filtering, prediction, smoothing, likelihood of the evidence, or most likely sequence)? Describe in words what kind of information this operation provides us.

This is a "most likely sequence" operation. It provides us with the most likely sequence of events that got us to the state we're currently in given the evidence we currently have.

2 Dynamic Bayesian Network

a) Formulate the information given above as a dynamic Bayesian network and provide the complete probability tables for the model.

The network consists of the starting node X_0 which has an arrow to the next node X_t. This node has an arrow to the node E_t, which is in fact two nodes, because we have two evidence factors, At_t (animal tracks at time t) and Fg_t (food gone at time t). A copy of the node X_t (together with it's children) is then appended as a child to X_t and the t is increased for both X and E, indicating the time t+1. This copying is then repeated ad infinitum.

Below is a temporal model of the Bayesian network. This model in my opinion is the best at showing the probability tables while being quite concise. The roundabout arrow indicates that the table points to a copy of itself as the time t increases.

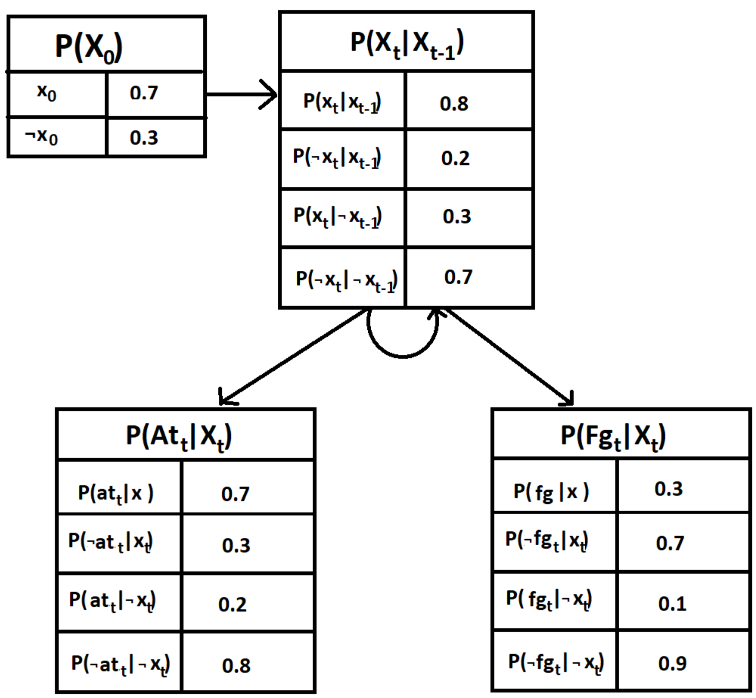


Figure 4: Dynamic Bayesian Network

b) Compute

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

For $t = 1$:

Calculating the prediction $P(X_1)$ first:

$$\begin{aligned}
P(X_1) &= \sum_{x0} P(X_1 \mid x_0) \cdot P(x_0) \\
&=< 0.8, 0.2 > \cdot\, 0.7+ < 0.3, 0.7 > \cdot\, 0.3 \\
&=< 0.56, 0.14 > + < 0.09, 0.21 > \\
P(X_1) &= \underline{< 0.65, 0.35 >}
\end{aligned}$$

Correcting the prediction with the evidence e1, which is {animal tracks, food gone}:

$$\begin{aligned}
P(X_1 \mid e1) &= \alpha P(e1 \mid X_1) \cdot P(X_1) \\
&= \alpha P(at_1, fg_1 \mid X_1) \cdot P(X_1)
\end{aligned}$$

Since $At_t \perp\!\!\!\perp Fg_t \mid X_t$, this can be written as:

$$\begin{aligned}
&= \alpha P(at_1 \mid X_1) \cdot P(fg_1 \mid X_1) \cdot P(X_1) \\
&= \alpha < 0.7 \cdot\ 0.3 \cdot\ 0.65, 0.2 \cdot\ 0.1 \cdot\ 0.35 > \\
&= \alpha < 0.1365, 0.007 > \\
&= \frac{1}{0.1365 + 0.007} \times < 0.1365, 0.007 > \\
P(X_1 \mid e1) &= \underline{\underline{< 0.951, 0.049 >}}
\end{aligned}$$

For $t = 2$:

Calculating the prediction $P(X_2 \mid e_1)$ first:

$$\begin{aligned}
P(X_2 \mid e_1) &= \sum_{x_1} P(X_2 \mid x_1) \cdot P(x_1 \mid e_1) \\
&=< 0.8, 0.2 > \cdot\, 0.951+ < 0.3, 0.7 > \cdot\, 0.007 \\
&=< 0.7608, 0.1902 > + < 0.0021, 0.0049 > \\
P(X_2 \mid e_1) &= \underline{< 0.7629, 0.1951 >}
\end{aligned}$$

Correcting the prediction with the evidence e2, which is {no animal tracks, food gone}:

$$\begin{aligned}
P(X_2 \mid e_{1:2}) &= \alpha P(e_2 \mid X_2) \cdot P(X_2 \mid e_1) \\
&= \alpha P(\neg at_2 \mid X_2) \cdot P(fg_2 \mid X_2) \cdot P(X_2 \mid e_1) \\
&= \alpha < 0.3 \cdot\ 0.3 \cdot\ 0.7629, 0.8 \cdot\ 0.1 \cdot\ 0.1951 > \\
&= \alpha < 0.068661, 0.015608 > \\
P(X_2 \mid e_{1:2}) &= \underline{\underline{< 0.814, 0.186 >}}
\end{aligned}$$

For $t = 3$:

Calculating the prediction $P(X_3 \mid e_{1:2})$ first:

$$\begin{aligned}
P(X_3 \mid e_{1:2}) &= \sum_{x_2} P(X_3 \mid x_2) \cdot P(x_2 \mid e_{1:2}) \\
&= < 0.8, 0.2 > \cdot 0.814 + < 0.3, 0.7 > \cdot 0.186 \\
&= < 0.6512, 0.1628 > + < 0.0558, 0.1302 > \\
P(X_3 \mid e_{1:2}) &= \underline{< 0.707, 0.293 >}
\end{aligned}$$

Correcting the prediction with the evidence e3, which is {no animal tracks, food not gone}:

$$\begin{aligned}
P(X_3 \mid e_{1:3}) &= \alpha P(e_3 \mid X_3) \cdot P(X_3 \mid e_{1:2}) \\
&= \alpha P(\neg at_3 \mid X_3) \cdot P(\neg fg_3 \mid X_3) \cdot P(X_3 \mid e_{1:2}) \\
&= \alpha < 0.3 \cdot\ 0.7 \cdot\ 0.707, 0.8 \cdot\ 0.9 \cdot\ 0.293 > \\
&= \alpha < 0.14847, 0.21096 > \\
P(X_3 \mid e_{1:3}) &= \underline{\underline{< 0.413, 0.587 >}}
\end{aligned}$$

For $t = 4$:

Calculating the prediction $P(X_4 \mid e_{1:3})$ first:

$$\begin{aligned}
P(X_4 \mid e_{1:3}) &= \sum_{x_3} P(X_4 \mid x_3) \cdot P(x_3 \mid e_{1:3}) \\
&= < 0.8, 0.2 > \cdot 0.413 + < 0.3, 0.7 > \cdot 0.587 \\
&= < 0.3304, 0.0826 > + < 0.1734, 0.4109 > \\
P(X_4 \mid e_{1:3}) &= \underline{< 0.5038, 0.4935 >}
\end{aligned}$$

Correcting the prediction with the evidence e4, which is {animal tracks, food not gone}:

$$\begin{aligned}
P(X_4 \mid e_{1:4}) &= \alpha P(e_4 \mid X_4) \cdot P(X_4 \mid e_{1:3}) \\
&= \alpha P(at_4 \mid X_4) \cdot P(\neg fg_4 \mid X_4) \cdot P(X_4 \mid e_{1:3}) \\
&= \alpha < 0.7 \cdot\ 0.7 \cdot\ 0.5038, 0.2 \cdot\ 0.9 \cdot\ 0.4935 > \\
&= \alpha < 0.246862, 0.08883 > \\
P(X_4 \mid e_{1:4}) &= \underline{\underline{< 0.735, 0.265 >}}
\end{aligned}$$

c) Compute

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

For $t = 5$:

$$\begin{aligned}
P(X_5 \mid e_{1:4}) &= \sum_{x_4} P(X_5 \mid x_4) \cdot P(x_4) \\
&= < 0.8, 0.2 > \cdot\, 0.735 + < 0.3, 0.7 > \cdot\, 0.265 \\
&= < 0.588, 0.147 > + < 0.0795, 0.1855 > \\
P(X_5 \mid e_{1:4}) &= \underline{\underline{< 0.6675, 0.3325 >}}
\end{aligned}$$

For $t = 6$:

$$\begin{aligned}
P(X_6 \mid e_{1:4}) &= \sum_{x_5} P(X_6 \mid x_5) \cdot P(x_5) \\
&= < 0.8, 0.2 > \cdot\, 0.6675 + < 0.3, 0.7 > \cdot\, 0.3325 \\
&= < 0.534, 0.1335 > + < 0.09975, 0.23275 > \\
P(X_6 \mid e_{1:4}) &= \underline{\underline{< 0.63375, 0.36625 >}}
\end{aligned}$$

For $t = 7$:

$$\begin{aligned}
P(X_7 \mid e_{1:4}) &= \sum_{x_6} P(X_7 \mid x_6) \cdot P(x_6) \\
&= < 0.8, 0.2 > \cdot\, 0.63375 + < 0.3, 0.7 > \cdot\, 0.36625 \\
&= < 0.507, 0.12675 > + < 0.109875, 0.256375 > \\
P(X_7 \mid e_{1:4}) &= \underline{\underline{< 0.616875, 0.383125 >}}
\end{aligned}$$

For $t = 8$:

$$\begin{aligned}
P(X_8 \mid e_{1:4}) &= \sum_{x_7} P(X_8 \mid x_7) \cdot P(x_7) \\
&= < 0.8, 0.2 > \cdot\, 0.616875 + < 0.3, 0.7 > \cdot\, 0.383125 \\
&= < 0.4935, 0.123375 > + < 0.1149375, 0.2681875 > \\
P(X_8 \mid e_{1:4}) &= \underline{\underline{< 0.6084375, 0.39155 >}}
\end{aligned}$$

d) By forecasting further and further into the future, you should see that the probability converges towards a fixed point. Verify that

$$\lim_{x \to \infty} P(X_t \mid e_{1:4}) = < 0.6, 0.4 >$$

Assuming that the values converge on $< 0.6, 0.4 >$ they shouldn't change if another step of prediction is performed with these values.

$$\begin{aligned}
P(X_{\infty+1} \mid e_{1:4}) &= \sum_{x_\infty} P(X_{\infty+1} \mid x_\infty) \cdot P(x_\infty)) \\
&=< 0.8, 0.2 > \cdot 0.6+ < 0.3, 0.7 > \cdot 0.4 \\
&=< 0.48, 0.12 > + < 0.12, 0.28 > \\
P(X_{\infty+1} \mid e_{1:4}) &= \underline{\underline{< 0.6, 0.4 >}}
\end{aligned}$$

e) Compute

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

Since this is backwards recursion, I will start from $t = 3$:

Forward part $P(X_3 \mid e_{1:3}) = \underline{< 0.413, 0.587 >}$

Backward part $\beta_{4:3} = P(e_{4:3} \mid X_3) = \underline{< 1, 1 >}$

$$\begin{aligned}
P(X_3 \mid e_{1:4}) &= \alpha P(X_3 \mid e_{1:3}) \cdot \beta_{4:3} \\
&= \alpha < 0.511 \cdot\ 1, 0.489 \cdot\ 1 > \\
&= \underline{\underline{< 0.511, 0.489 >}}
\end{aligned}$$

For $t = 2$:

Forward part $P(X_2 \mid e_{1:2}) = \underline{< 0.814, 0.186 >}$

Backward part $\beta_{3:2}$:

$$\begin{aligned}
P(e_{4:4} \mid X_2) &= \sum_{x_3} P(e_{3:4}, X_3 | X_2) \\
&= \sum_{x_3} P(e_{4:4} \mid x_3) \cdot P(e_3 \mid x_3) \cdot P(X_4 \mid x_3) \\
&= \sum_{x_3} \beta_{4:3} \cdot P(\neg at_3 \mid x_3) \cdot P(\neg fg_3 \mid x3) \cdot P(X_4 \mid x_3)) \\
&= (1 \cdot 0.3 \cdot 0.7 \cdot < 0.8, 0.2 >) + (1 \cdot 0.8 \cdot 0.9 \cdot < 0.3, 0.7 >) \\
&= 0.21 \cdot < 0.8, 0.2 > +0.72 \cdot < 0.3, 0.7 > \\
&=< 0.168, 0.042 > + < 0.216, 0.504 > \\
&= \underline{< 0.384, 0.546 >}
\end{aligned}$$

$$\begin{aligned}
P(X_2 \mid e_{1:4}) &= \alpha P(X_2 \mid e_{1:2}) \cdot \beta_{3:2} \\
&= \alpha < 0.814 \cdot\ 0.384, 0.186 \cdot\ 0.546 > \\
&= \alpha < 0.312576, 0.101556 > \\
&= \underline{\underline{< 0.755, 0.245 >}}
\end{aligned}$$

For $t = 1$:

Forward part $P(X_1 \mid e_{1:1}) = \underline{< 0.951, 0.049 >}$

Backward part $\beta_{2:1}$:

$$
\begin{aligned}
P(e_{3:4} \mid X_1) &= \sum_{x_2} P(e_{2:4}, x_2 \mid X_1) \\
&= \sum_{x_2} P(e_{3:4} \mid x_2 \cdot P(e_2) \mid x_2) \cdot P(X_3 \mid x_2) \\
&= \sum_{x_2} \beta_{3:2} \cdot P(\neg at_2 \mid x_2) \cdot P(fg_2 \mid x_2) \cdot P(X_3 \mid x_2) \\
&= (0.755 \cdot 0.3 \cdot\ 0.3 \cdot < 0.8, 0.2 >) + (0.245 \cdot 0.8 \cdot 0.1 \cdot < 0.3, 0.7 >) \\
&= 0.06795 \cdot < 0.8, 0.2 > + 0.0196 \cdot < 0.3, 0.7 > \\
&= < 0.01359, 0.006912 > + < 0.00588, 0.01372 > \\
&= \underline{< 0.01947, 0.020632 >}
\end{aligned}
$$

$$
\begin{aligned}
P(X_1 \mid e_{1:4}) &= \alpha P(X_1 \mid e_{1:1}) \cdot \beta_{2:1} \\
&= \alpha < 0.951 \cdot\ 0.01947, 0.049 \cdot\ 0.020632 > \\
&= \alpha < 0.0185159, 0.0010109 > \\
&= \underline{\underline{< 0.948, 0.052 >}}
\end{aligned}
$$

For $t = 0$

Forward part $P(X_0 \mid e_{0:0}) = \underline{< 0.65, 0.35 >}$

Backward part $\beta_{1:0}$:

$$
\begin{aligned}
P(e_{2:4} \mid X_0) &= \sum_{x_1} P(e_{1:4}, x_1 \mid X_0) \\
&= \sum_{x_1} P(e_{2:4} \mid x_2 \cdot P(e_1 \mid x_1) \cdot P(X_2 \mid x_1)) \\
&= \sum_{x_1} \beta_{2:1} \cdot P(at_1 \mid x_1) \cdot P(fg_1 \mid x_1) \cdot P(X_2 \mid x_1) \\
&= (0.01947 \cdot 0.7 \cdot 0.3 \cdot < 0.8, 0.2 >) + (0.020632 \cdot 0.2 \cdot 0.1 \cdot < 0.3, 0.7 >) \\
&= < 0.0032709, 0.081774 > + < 0.0001237, 0.0002888 > \\
&= \underline{< 0.0033946, 0.0820628 >}
\end{aligned}
$$

$$
\begin{aligned}
P(X_0 \mid e_{1:4}) &= \alpha P(X_0 \mid e_{0:0}) \cdot \beta_{1:0} \\
&= \alpha < 0.65 \cdot 0.0033946, 0.35 \cdot 0.0820628 > \\
&= \alpha < 0.0022064, 0.0287219 > \\
&= \underline{\underline{0.07, 0, 93}}
\end{aligned}
$$