NTNU
Kunnskap for en bedre verden

TDT4171 - Artificial Intelligence Methods

Assignment 2 - Dynamic models

Sander Lindberg

Februar, 2021

1 Hidden Markov Model

1.1 Problems

Problem a

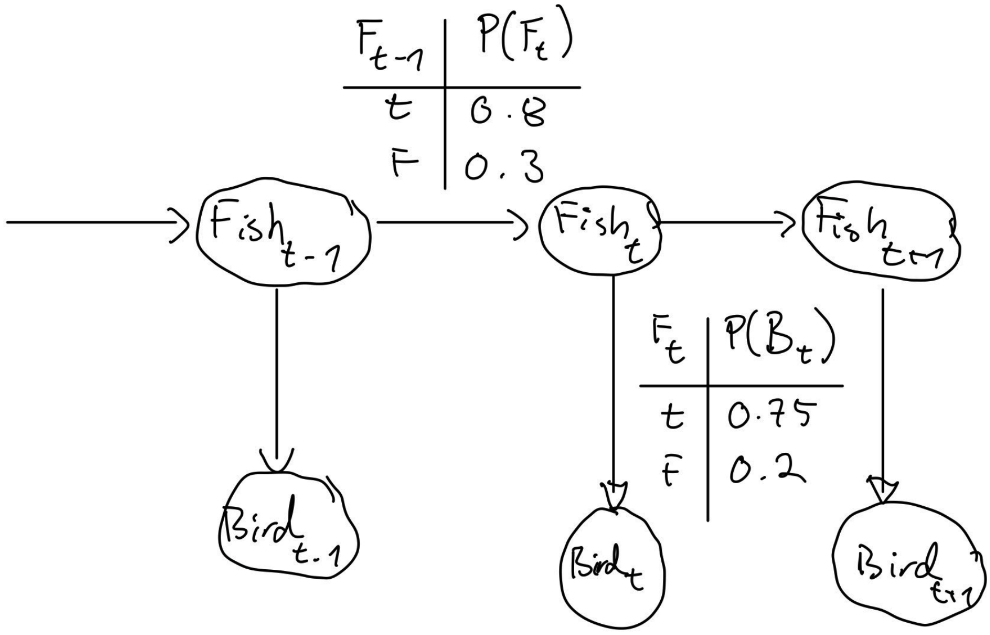

Figure 1: Markov model for the information in task 1a

$Fish\ nearby$	$\overline{Fish\ nearby}$
0.5	0.5

Table 1: Prior probabilities

	$Fish\ nearby$	$\overline{Fish\ nearby}$
$Fish\ nearby$	0.8	0.2
$\overline{Fish\ nearby}$	0.3	0.7

Table 2: Transition probabilities

	$Birdsnearby$	$\overline{Birds\ nearby}$
$Fish$	0.75	0.25
$\overline{Fish}$	0.2	0.8

Table 3: Emission probabilities

Problem b

See the delivered code for implementation.

This operation is filtering. Filtering is used to make a prediction, based on evidence from the present and past. So for example in this subtask, when calculating $P(X_2|e_{1:2})$, we are predicting whether or not there are fish nearby on day 2, given the evidence of birds nearby on day 1 and day 2.

Problem c

See the delivered code for implementation.

```
          #######################
          #       Task 1b       #
          #######################
P(X_t|e_1:t): [Fish nearby, No fish nearby]
-------------------------------------------

P(X_1|e_1:1): [0.82089552 0.17910448]
P(X_2|e_1:2): [0.90197069 0.09802931]
P(X_3|e_1:3): [0.48518523 0.51481477]
P(X_4|e_1:4): [0.81645924 0.18354076]
P(X_5|e_1:5): [0.43134895 0.56865105]
P(X_6|e_1:6): [0.79970863 0.20029137]
```

Figure 2: Output from code implementation in task 1b

```
          #######################
          #       Task 1c       #
          #######################
P(X_t|e_1:6): [Fish nearby, No fish nearby]
-------------------------------------------

P(X_7|e_1:6): [0.69985432 0.30014568]
P(X_8|e_1:6): [0.64992716 0.35007284]
P(X_9|e_1:6): [0.62496358 0.37503642]
P(X_10|e_1:6): [0.61248179 0.38751821]
P(X_11|e_1:6): [0.60624089 0.39375911]
P(X_12|e_1:6): [0.60312045 0.39687955]
P(X_13|e_1:6): [0.60156022 0.39843978]
P(X_14|e_1:6): [0.60078011 0.39921989]
P(X_15|e_1:6): [0.60039006 0.39960994]
P(X_16|e_1:6): [0.60019503 0.39980497]
P(X_17|e_1:6): [0.60009751 0.39990249]
P(X_18|e_1:6): [0.60004876 0.39995124]
P(X_19|e_1:6): [0.60002438 0.39997562]
P(X_20|e_1:6): [0.60001219 0.39998781]
P(X_21|e_1:6): [0.60000609 0.39999391]
P(X_22|e_1:6): [0.60000305 0.39999695]
P(X_23|e_1:6): [0.60000152 0.39999848]
P(X_24|e_1:6): [0.60000076 0.39999924]
P(X_25|e_1:6): [0.60000038 0.39999962]
P(X_26|e_1:6): [0.60000019 0.39999981]
P(X_27|e_1:6): [0.6000001 0.3999999]
P(X_28|e_1:6): [0.60000005 0.39999995]
P(X_29|e_1:6): [0.60000002 0.39999998]
P(X_30|e_1:6): [0.60000001 0.39999999]
```

Figure 3: Output from code implementation in task 1c

This operation is prediction. Prediction is used to make a prediction about the future, based based purely on evidence from the past. So for example in this subtask, when calculating $P(X_7|e_{1:4})$, we are predicting whether or not there are fish nearby on day 5, given the evidence of birds nearby on day $1-6$, but we have no evidence for day 7. This result is then used for predicting day 8, which is used to predict day 9 and so on. The distribution in this subtask is converging toward a 60% chance that there are fish nearby on day t and 40% that there are no fish nearby. This "convergence probabilities" are called the **stationary distribution** of the Markov chain.

Problem d

See the delivered code for implementation.

```
        #######################
        #       Task 1d       #
        #######################
P(X_t|e_1:6): [Fish nearby, No fish nearby]
-------------------------------------------

P(X_0|e_1:6): [0.66485218 0.33514782]
P(X_1|e_1:6): [0.87640731 0.12359269]
P(X_2|e_1:6): [0.86578657 0.13421343]
P(X_3|e_1:6): [0.59792735 0.40207265]
P(X_4|e_1:6): [0.76663731 0.23336269]
P(X_5|e_1:6): [0.57082582 0.42917418]
```

Figure 4: Output from code implementation in task 1d

This operation is smoothing. Smoothing is used to make a prediction for day t given evidence up to the present.

Problem e

See the delivered code for implementation.

```
        #######################
        #       Task 1e       #
        #######################
        Most likely sequence (1-6)
-------------------------------------------

            Fish: [0.82089552]
            Fish: [0.49253731]
            No fish: [0.11820896]
            Fish: [0.05910448]
            No fish: [0.01418507]
            Fish: [0.00709254]
```

Figure 5: Output from code implementation in task 1e

This operation is most likely sequence. Most likely sequence is as the name suggests, the most likely sequence of events which has generated this sequence of evidence. In this subtask, the most likely sequence for fish nearby or no fish nearby are: Fish, Fish, No fish, Fish, No fish, Fish.

2 Dynamic Bayesian Network

2.1 Problems

Problem a

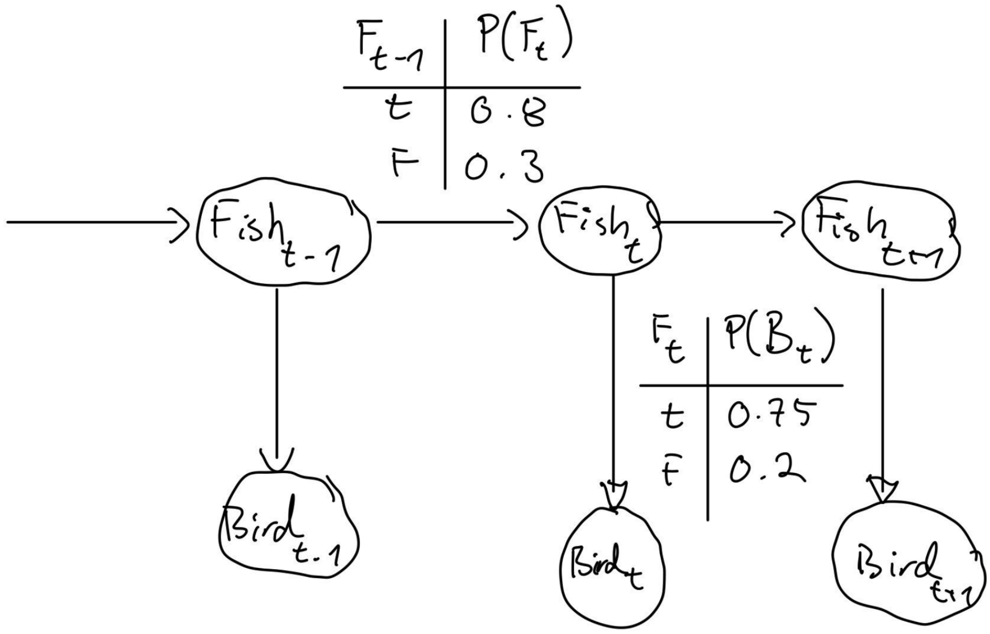

Figure 6: Dynamic Bayesian network for the information in task 2a

Animals nearby	$\overline{\textit{Animals nearby}}$
0.7	0.3

Table 4: Prior probabilities

	Animals nearby	$\overline{\textit{Animals nearby}}$
Animals nearby	0.8	0.2
$\overline{\textit{Animals nearby}}$	0.3	0.7

Table 5: Transition probabilities

	Animal tracks		$\overline{\textit{Animal tracks}}$	
	Food	$\overline{\textit{Food}}$	*Food*	$\overline{\textit{Food}}$
Animalsnearby	0.49	0.21	0.21	0.09
$\overline{\textit{Animalsnearby}}$	0.18	0.02	0.72	0.08

Table 6: Emission probabilities

Problem b

Compute

$$P(X_t|e_{1:t}), \text{ for } t = 1, 2, 3, 4$$

I'll begin to introduce the observation matrices, derived from table 6:

$$O_1 = \begin{bmatrix} 0.21 & 0 \\ 0 & 0.02 \end{bmatrix} \; O_2 = \begin{bmatrix} 0.09 & 0 \\ 0 & 0.08 \end{bmatrix} \; O_3 = \begin{bmatrix} 0.21 & 0 \\ 0 & 0.72 \end{bmatrix} \; O_4 = \begin{bmatrix} 0.49 & 0 \\ 0 & 0.18 \end{bmatrix}$$

For all calculations below, I will use the general method given below

$$\begin{aligned} P(X_{t+1}|e_{1:t+1}) &= \alpha \cdot P(e_{t+1}|X_{t+1}) \cdot \sum_{X_t} P(X_{t+1}|x_t) \cdot P(x_t|e_{1:t}) \\ &= \alpha \cdot P(e_{t+1}|X_{t+1}) \cdot T^T \cdot P(x_t|e_{1:t}) \end{aligned}$$

Where T^T is the transpose of the transition matrix, given in table 5.

$$\begin{aligned} P(X_1|e_{1:1}) &= \alpha \cdot P(e_1|X_1) \cdot P(X_0) \\ &= \alpha \cdot O_1 \cdot T^T \cdot P(X_0) \\ &= \alpha \cdot \begin{bmatrix} 0.21 & 0 \\ 0 & 0.02 \end{bmatrix} \cdot \begin{bmatrix} 0.8 & 0.3 \\ 0.2 & 0.7 \end{bmatrix} \cdot \begin{bmatrix} 0.7 \\ 0.3 \end{bmatrix} \\ &= \alpha \cdot \begin{bmatrix} 0.168 & 0.063 \\ 0.004 & 0.014 \end{bmatrix} \cdot \begin{bmatrix} 0.7 \\ 0.3 \end{bmatrix} \\ &= \alpha \cdot \begin{bmatrix} 0.1365 \\ 0.007 \end{bmatrix} = \begin{bmatrix} 0.9512 \\ 0.04878 \end{bmatrix} \end{aligned}$$

$$\begin{aligned} P(X_2|e_{1:2}) &= \alpha \cdot P(e_2|X_2) \cdot P(X_2|e_1) \\ &= \alpha \cdot O_2 \cdot T^T \cdot P(X_2|e_1) \\ &= \alpha \cdot \begin{bmatrix} 0.09 & 0 \\ 0 & 0.08 \end{bmatrix} \cdot \begin{bmatrix} 0.8 & 0.3 \\ 0.2 & 0.7 \end{bmatrix} \cdot \begin{bmatrix} 0.9512 \\ 0.04878 \end{bmatrix} \\ &= \alpha \cdot \begin{bmatrix} 0.072 & 0.027 \\ 0.016 & 0.056 \end{bmatrix} \cdot \begin{bmatrix} 0.9512 \\ 0.04878 \end{bmatrix} \\ &= \alpha \cdot \begin{bmatrix} 0.0698 \\ 0.01795 \end{bmatrix} = \begin{bmatrix} 0.7954 \\ 0.2045 \end{bmatrix} \end{aligned}$$

$$\begin{aligned} P(X_3|e_{1:3}) &= \alpha \cdot P(e_3|X_3) \cdot P(X_3|e_2) \\ &= \alpha \cdot O_3 \cdot T^T \cdot P(X_3|e_2) \\ &= \alpha \cdot \begin{bmatrix} 0.21 & 0 \\ 0 & 0.72 \end{bmatrix} \cdot \begin{bmatrix} 0.8 & 0.3 \\ 0.2 & 0.7 \end{bmatrix} \cdot \begin{bmatrix} 0.7954 \\ 0.2045 \end{bmatrix} \\ &= \alpha \cdot \begin{bmatrix} 0.168 & 0.063 \\ 0.144 & 0.504 \end{bmatrix} \cdot \begin{bmatrix} 0.7954 \\ 0.2045 \end{bmatrix} \\ &= \alpha \cdot \begin{bmatrix} 0.1456 \\ 0.2176 \end{bmatrix} = \begin{bmatrix} 0.4023 \\ 0.5976 \end{bmatrix} \end{aligned}$$

$$\begin{aligned} P(X_4|e_{1:4}) &= \alpha \cdot P(e_4|X_4) \cdot P(X_4|e_3) \\ &= \alpha \cdot O_4 \cdot T^T \cdot P(X_4|e_3) \\ &= \alpha \cdot \begin{bmatrix} 0.49 & 0 \\ 0 & 0.18 \end{bmatrix} \cdot \begin{bmatrix} 0.8 & 0.3 \\ 0.2 & 0.7 \end{bmatrix} \cdot \begin{bmatrix} 0.4023 \\ 0.5976 \end{bmatrix} \\ &= \alpha \cdot \begin{bmatrix} 0.392 & 0.147 \\ 0.036 & 0.126 \end{bmatrix} \cdot \begin{bmatrix} 0.4023 \\ 0.5976 \end{bmatrix} \\ &= \alpha \cdot \begin{bmatrix} 0.2456 \\ 0.0898 \end{bmatrix} = \begin{bmatrix} 0.7323 \\ 0.2677 \end{bmatrix} \end{aligned}$$

Problem c

Compute

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

I will again use the formula above, with one modification; the observation matrix O is the identity matrix $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$. As multiplying with this is just like multiplying with 1, I will not write it in the expressions.

$$\begin{aligned} P(X_5|e_{1:4}) &= \alpha \cdot T^T \cdot P(X_4|e_{1:4}) \\ &= \alpha \cdot \begin{bmatrix} 0.8 & 0.3 \\ 0.2 & 0.7 \end{bmatrix} \cdot \begin{bmatrix} 0.7323 \\ 0.2677 \end{bmatrix} \\ &= \begin{bmatrix} 0.66615 \\ 0.33385 \end{bmatrix} \end{aligned}$$

$$\begin{aligned} P(X_6|e_{1:4}) &= \alpha \cdot T^T \cdot P(X_5|e_{1:4}) \\ &= \alpha \cdot \begin{bmatrix} 0.8 & 0.3 \\ 0.2 & 0.7 \end{bmatrix} \cdot \begin{bmatrix} 0.66615 \\ 0.33385 \end{bmatrix} \\ &= \begin{bmatrix} 0.63305 \\ 0.36695 \end{bmatrix} \end{aligned}$$

$$\begin{aligned} P(X_7|e_{1:4}) &= \alpha \cdot T^T \cdot P(X_6|e_{1:4}) \\ &= \alpha \cdot \begin{bmatrix} 0.8 & 0.3 \\ 0.2 & 0.7 \end{bmatrix} \cdot \begin{bmatrix} 0.63305 \\ 0.36695 \end{bmatrix} \\ &= \begin{bmatrix} 0.616525 \\ 0.383565 \end{bmatrix} \end{aligned}$$

$$\begin{aligned} P(X_8|e_{1:4}) &= \alpha \cdot T^T \cdot P(X_7|e_{1:4}) \\ &= \alpha \cdot \begin{bmatrix} 0.8 & 0.3 \\ 0.2 & 0.7 \end{bmatrix} \cdot \begin{bmatrix} 0.616525 \\ 0.383565 \end{bmatrix} \\ &= \begin{bmatrix} 0.608295 \\ 0.3918005 \end{bmatrix} \end{aligned}$$

Problem d

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

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

Since I know that $\lim_{t\to\infty} P(X_t|e_{1:4}) = \begin{bmatrix} 0.6 \\ 0.4 \end{bmatrix}$, I know that $P(X_t|e_{1:4}) = \begin{bmatrix} 0.6 \\ 0.4 \end{bmatrix}$. I can use this to calculate $P(X_{t+1}|e_{1:4})$ and see that the result is also $\begin{bmatrix} 0.6 \\ 0.4 \end{bmatrix}$. Using the same method as in problem c:

$$
\begin{aligned}
P(X_{t+1}|e_{1:4}) &= T^T \dot{P}(X_t|e_{1:4}) \\
&= \begin{bmatrix} 0.8 & 0.3 \\ 0.2 & 0.7 \end{bmatrix} \cdot \begin{bmatrix} 0.6 \\ 0.4 \end{bmatrix} \\
&= \begin{bmatrix} 0.8 \cdot 0.6 + 0.3 \cdot 0.4 \\ 0.2 \cdot 0.6 + 0.7 \cdot 0.4 \end{bmatrix} \\
&= \begin{bmatrix} 0.6 \\ 0.4 \end{bmatrix}
\end{aligned}
$$

Problem e

Compute

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

In this subtask, I'll use the forward-backward algorithm for smoothing. I'll start by listing all "f's" and the starting "b":

$$f_0 = \begin{bmatrix} 0.7 \\ 0.3 \end{bmatrix} \quad f_{1:1} = \begin{bmatrix} 0.9512 \\ 0.04878 \end{bmatrix} \quad f_{1:2} = \begin{bmatrix} 0.7954 \\ 0.2045 \end{bmatrix} \quad f_{1:3} = \begin{bmatrix} 0.4023 \\ 0.5976 \end{bmatrix} \quad f_{1:4} = \begin{bmatrix} 0.7323 \\ 0.2677 \end{bmatrix} \quad b_{5:4} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad b_{4:4} = \begin{bmatrix} 0.7323 \\ 0.2677 \end{bmatrix}$$

I'll use the two formulas:

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

$$b_{k+1:t} = P(e_{k+1:t}|X_k) = T \cdot O_k \cdot b_{k+2:t} \tag{2}$$

Assuming here that the k in the formula are the t given in the task and $t = 4$.

$$
\begin{aligned}
P(X_3|e_{1:4}) &= \alpha \cdot f_{1:3} \cdot b_{4:4} \\
&= \alpha \cdot \begin{bmatrix} 0.9512 \\ 0.04878 \end{bmatrix} \cdot \begin{bmatrix} 0.7323 \\ 0.2677 \end{bmatrix} \\
&= \alpha \cdot \begin{bmatrix} 0.2846 \\ 0.1599 \end{bmatrix} = \begin{bmatrix} 0.64818 \\ 0.35185 \end{bmatrix}
\end{aligned}
$$

$$
\begin{aligned}
b_{3:4} &= \begin{bmatrix} 0.8 & 0.2 \\ 0.3 & 0.7 \end{bmatrix} \cdot \begin{bmatrix} 0.21 & 0 \\ 0 & 0.72 \end{bmatrix} \cdot \begin{bmatrix} 0.7323 \\ 0.2677 \end{bmatrix} \\
&= \begin{bmatrix} 0.168 & 0.144 \\ 0.063 & 0.504 \end{bmatrix} \cdot \begin{bmatrix} 0.7323 \\ 0.2677 \end{bmatrix} \\
&= \begin{bmatrix} 0.161572 \\ 0.1810557 \end{bmatrix}
\end{aligned}
$$

$$
\begin{aligned}
P(X_2|e_{1:4}) &= \alpha \cdot f_{1:2} \cdot b_{3:4} \\
&= \alpha \cdot \begin{bmatrix} 0.7954 \\ 0.2045 \end{bmatrix} \cdot \begin{bmatrix} 0.161572 \\ 0.1810557 \end{bmatrix} \\
&= \alpha \cdot \begin{bmatrix} 0.128517 \\ 0.03702 \end{bmatrix} = \begin{bmatrix} 0.77636 \\ 0.22364 \end{bmatrix}
\end{aligned}
$$

$$
\begin{aligned}
b_{2:4} &= \begin{bmatrix} 0.8 & 0.2 \\ 0.3 & 0.7 \end{bmatrix} \cdot \begin{bmatrix} 0.09 & 0 \\ 0 & 0.08 \end{bmatrix} \cdot \begin{bmatrix} 0.161572 \\ 0.1810557 \end{bmatrix} \\
&= \begin{bmatrix} 0.072 & 0.016 \\ 0.027 & 0.056 \end{bmatrix} \cdot \begin{bmatrix} 0.161572 \\ 0.1810557 \end{bmatrix} \\
&= \begin{bmatrix} 0.01483104 \\ 0.01461442 \end{bmatrix}
\end{aligned}
$$

$$
\begin{aligned}
P(X_1|e_{1:4}) &= \alpha \cdot f_{1:1} \cdot b_{2:4} \\
&= \alpha \cdot \begin{bmatrix} 0.9512 \\ 0.04878 \end{bmatrix} \cdot \begin{bmatrix} 0.01483104 \\ 0.01461442 \end{bmatrix} \\
&= \alpha \cdot \begin{bmatrix} 0.0141073 \\ 7.13e-4 \end{bmatrix} = \begin{bmatrix} 0.9521 \\ 0.0479 \end{bmatrix}
\end{aligned}
$$

$$
\begin{aligned}
b_{1:4} &= \begin{bmatrix} 0.8 & 0.2 \\ 0.3 & 0.7 \end{bmatrix} \cdot \begin{bmatrix} 0.21 & 0 \\ 0 & 0.02 \end{bmatrix} \cdot \begin{bmatrix} 0.01483104 \\ 0.01461442 \end{bmatrix} \\
&= \begin{bmatrix} 0.168 & 0.004 \\ 0.063 & 0.014 \end{bmatrix} \cdot \begin{bmatrix} 0.01483104 \\ 0.01461442 \end{bmatrix} \\
&= \begin{bmatrix} 2.55e-3 \\ 1.4e-3 \end{bmatrix}
\end{aligned}
$$

$$
\begin{aligned}
P(X_0|e_{1:4}) &= \alpha \cdot f_0 \cdot b_{1:4} \\
&= \alpha \cdot \begin{bmatrix} 0.7 \\ 0.3 \end{bmatrix} \cdot \begin{bmatrix} 2.55e-3 \\ 1.4e-3 \end{bmatrix} \\
&= \alpha \cdot \begin{bmatrix} 1.785e-3 \\ 4.2e-4 \end{bmatrix} = \begin{bmatrix} 0.8095 \\ 0.1905 \end{bmatrix}
\end{aligned}
$$