1 Chocolate Factory


1.1

Imagine that you work at ACME Chocolate Factory, confectioners extraordinaire. Your job is to keep an eye on the conveyor belt, watching the chocolates as they come out of the press one at a time.

Suppose that ACME makes two types of chocolates: ones with almonds and ones without. For the first few problems, assume you can tell with 100% accuracy what the chocolate contains. In the control room, there is a lever that switches the almond control on and off. When the conveyor is turned on at the beginning of the day, there is a 50% chance that the almond lever is on, and a 50% chance that it is off. As soon as the conveyor belt is turned on, it starts making a piece of candy.

Unfortunately, someone has let a monkey loose in the control room, and it has locked the door and started the conveyor belt. The lever cannot be moved while a piece of candy is being made. Between pieces, however, there is a 30% chance that the monkey switches the lever to the other position (i.e. turns almonds on if it was off, or off if it was on). Draw a Markov Model that represents the situation and give the prior distribution on the states as well as the transition matrix.

**Solution:** There are two states: one in which the conveyor belt is producing plain chocolates (P) and another in which it is producing almond chocolates (A).

![Markov Model Diagram](image)

The prior distribution is: \( P(P) = P(A) = 1/2 \) and the transition probabilities are:

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>0.7</td>
<td>0.3</td>
</tr>
<tr>
<td>P</td>
<td>0.3</td>
<td>0.7</td>
</tr>
</tbody>
</table>

1.2

Now assume that there is a coconut lever as well, so that there are four types of candy: Plain, Almond, Coconut, and Almond+Coconut. Again, there is to 50% chance of the lever being on at the beginning of the day, and the chance of the monkey switching the state of the second lever between candies is also 30%. Assume that the switching of the levers is independent of each other. Now Draw a model for production of all four types of chocolate.

**Solution:** Now we have four states: P: plain, A: almond, C: coconut, B: both almond and coconut.
Because the levers are independent, we have a uniform prior on the states: \( P(A) = P(B) = P(C) = P(P) = 1/4 \). The state transition probabilities are (the probability of both levers being flipped is 0.09, the probability of one being flipped is 0.7*0.3=0.21, and the probability of neither being flipped is 0.49):

\[
\begin{array}{cccc}
  & P & A & C & B \\
  P & 0.49 & 0.21 & 0.21 & 0.09 \\
  A & 0.21 & 0.49 & 0.09 & 0.21 \\
  C & 0.21 & 0.09 & 0.49 & 0.21 \\
  B & 0.09 & 0.21 & 0.21 & 0.49 \\
\end{array}
\]

1.3
What is the probability that the machine will produce, in order: \{Plain, Almond, Almond, Almond+Coconut\}?  
**Solution:** \( P(P, A, A, B) = P(P)P(A|P)P(A|A)P(B|A) = 0.25 \times 0.21 \times 0.49 \times 0.21 = 0.0054 \).

1.4
Now assume that you can’t tell what’s inside the chocolate candy, only that the chocolate is light or dark. Use the following table for \( P(\text{color} \mid \text{stuff inside}) \):

<table>
<thead>
<tr>
<th>Inside</th>
<th>( P(\text{Color} = \text{light}) )</th>
<th>( P(\text{Color} = \text{dark}) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Plain</td>
<td>0.1</td>
<td>0.9</td>
</tr>
<tr>
<td>Almond</td>
<td>0.3</td>
<td>0.7</td>
</tr>
<tr>
<td>Coconut</td>
<td>0.8</td>
<td>0.2</td>
</tr>
<tr>
<td>Almond+Coconut</td>
<td>0.9</td>
<td>0.1</td>
</tr>
</tbody>
</table>

Assume that the color of one chocolate is independent of the next, given the filling. If the first two chocolates were dark and light, respectively, what is the smoothed distribution over the state at time 1 and the smoothed distribution over the state at time 2?  
**Solution:** We want to compute \( P(S_1 \mid E_1 = d, E_2 = l) \) and \( P(S_2 \mid E_1 = d, E_2 = l) \):
First, the forward messages are  
\( f_v[0] = [.25, .25, .25, .25]^T \).
\[
f v[1] = \alpha \begin{bmatrix}
.9 & 0 & 0 & 0 \\
0 & .7 & 0 & 0 \\
0 & 0 & .2 & 0 \\
0 & 0 & 0 & .1 \\
\end{bmatrix} \quad f v[0] = \alpha [0.225, .175, 0.05, 0.025]^T = [0.4737, 0.3684, 0.1053, 0.0526]^T \\
f v[2] = \alpha \begin{bmatrix}
.1 & 0 & 0 & 0 \\
0 & .3 & 0 & 0 \\
0 & 0 & .8 & 0 \\
0 & 0 & 0 & .9 \\
\end{bmatrix} \quad f v[1] = [0.0780, 0.2091, 0.3623, 0.3505]^T.
\]

Now, the backward messages:
\[
b[3] = [1, 1, 1, 1]^T \\
b[2] = \begin{bmatrix}
.49 & .21 & .21 & .09 \\
.21 & .49 & .09 & .21 \\
.21 & .09 & .49 & .21 \\
.09 & .21 & .21 & .49 \\
\end{bmatrix} \quad b[3] = [0.3610, 0.4290, 0.6290, 0.6810]^T.
\]

Finally, we have:
\[
P(S_1|E_1 = d, E_2 = l) = \alpha f v[1]b[2] = [0.3967, 0.3666, 0.1536, 0.0831]^T \\
P(S_2|E_1 = d, E_2 = l) = \alpha f v[2]b[3] = [0.0780, 0.2091, 0.3623, 0.3505]^T.
\]
2 Submarines and PRM Solutions

You are performing surveillance, trying to decide which destination in a harbor a particular submarine is headed toward. At each time step, the situation can be characterized by the following variables:

- **destination** Which destination the submarine is headed to.
- **location** The submarine’s current location
- **observation** Your noisy observation (via underwater sensing) of the submarine’s location
- **action** The submarine’s actions (speed and steering)

In addition, there is a variable, **type**, which encodes the type of the submarine (which is useful to know, because it affects the sub’s speed and maneuverability).

(a) (6 pts) Draw a dynamic Bayesian network diagram that describes this system. Only the **observation** variable is directly observable. Show how the values of the variables at the current time step depend on their values in the previous time step. Your model should be able to encode these relationships:

- Submarines tend to stick with the same destination, and don’t frequently change which one they’re aiming at.
- The choice of action depends on relative position of the submarine and its destination.

\[
\text{Parents}(\text{destination}_{t+1}) = \{\text{destination}_t\} \\
\text{Parents}(\text{location}_{t+1}) = \{\text{location}_t, \text{action}_t\} \\
\text{Parents}(\text{observation}_{t+1}) = \{\text{location}_{t+1}\} \\
\text{Parents}(\text{action}_{t+1}) = \{\text{destination}_{t+1}, \text{location}_{t+1}\} \\
\]

(b) (2 pts) What would you change in your diagram in order to model the idea that some types of submarines have different ranges and that the distance to a location might affect its selection as a destination?

\[
\text{Parents}(\text{destination}_{t+1}) = \{\text{destination}_t, \text{type}, \text{location}_{t+1}\} \\
\]

(c) (3 pts) If you have made three observations, \(O_1\), \(O_2\), and \(O_3\), give an expression for the probability distribution over the destination at step 3:

\[
\Pr(D_3|O_1, O_2, O_3) \\
\]

using only probabilities stored in the CPTs of the network (in part (a)).
\[
Pr(D_3 | O_1, O_2, O_3) = \frac{\sum_{L_1, L_2, A_1, A_2, A_3, D_1, D_2} Pr(D_3, O_1, O_2, O_3)}{\sum_{L_1, L_2, A_1, A_2, A_3, D_1, D_2} Pr(O_1, O_2, O_3)} \\
= \frac{\sum Pr(D_3, O_1, O_2, O_3)}{\sum Pr(O_1, O_2, O_3)} = \frac{\sum Pr(DBN)}{\sum Pr(DBN)} \\
DBN = \left\{ \begin{array}{l} Pr(D_3 | D_2) \times Pr(D_2 | D_1) \times Pr(D_1) \\
Pr(L_3 | L_2, A_2) \times Pr(L_2 | L_1, A_1) \times Pr(L_1) \\
Pr(O_2 | L_3) \times Pr(O_2 | L_2) \times Pr(O_1 | L_1) \\
Pr(A_3 | D_3, L_3) \times Pr(A_2 | D_2, L_2) \times Pr(A_1 | D_1, L_2) \end{array} \right\}
\]

3. (10 points) Now we'll consider modeling a problem using probabilistic relational models. Here is our world:

- A car’s speed depends on the size of its engine and the mood of its driver.
- A person’s mood depends on his bank balance.
- A person’s bank balance depends on his employer.

(a) (3 pts) Draw a relational probabilistic model (or show in rules) the structure of this model. What are the classes? What are their simple attributes? What are their complex attributes? How are they related?

<table>
<thead>
<tr>
<th>Simple</th>
</tr>
</thead>
<tbody>
<tr>
<td>Speed(car)</td>
</tr>
<tr>
<td>EngineSize(car)</td>
</tr>
<tr>
<td>Mood(person)</td>
</tr>
<tr>
<td>BankBalance(person)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Complex</th>
</tr>
</thead>
<tbody>
<tr>
<td>Owner(car):person</td>
</tr>
<tr>
<td>Employer(person):person</td>
</tr>
</tbody>
</table>

\[
\text{Parents}(\text{Speed(car)}) = \{\text{EngineSize(car)}, \text{Mood(Owner(car))}\} \\
\text{Parents}(\text{Mood(person)}) = \{\text{BankBalance(person)}\} \\
\text{Parents}(\text{BankBalance(person)}) = \{\text{Employer(person)}\}
\]

(b) (3 pts) Assume the following world (Camaro1 and Pinto2 are cars):

- owner(Camaro1) = John
- owner(Pinto2) = Mary

Draw the Bayesian network associated with the instantiation of your model in this world.
Parents(Speed(Camaro1)) = \{\text{EngineSize(Camaro1), Mood(John)}\}
Parents(Speed(Pinto2)) = \{\text{EngineSize(Pinto2), Mood(Mary)}\}

Parents(Mood(John)) = \{\text{BankBalance(John)}\}
Parents(Mood(Mary)) = \{\text{BankBalance(Mary)}\}

Parents(BankBalance(John)) = \{\text{Employer(John)}\}
Parents(BankBalance(Mary)) = \{\text{Employer(Mary)}\}

(c) (2 pts) Say what would have to change in your model if the mood of the car’s owner also depended on how comfortable the seats were?

Parents(Mood(person)) = \{\text{BankBalance(person), Seats(\text{CarOf(person)})}\}

(d) (1 pt) Are the speeds of Camaro1 and Pinto2 independent, assuming we don’t know who John and Mary work for?

Yes

(e) (1 pt) Are the speeds of Camaro1 and Pinto2 independent given that they both work for Yoyodyne?

No