Lesson

Bayesian Networks

2

In this Class
- Conditional probability.
- Conditional Probability Tables (CPTs).
- Dependencies and Markov property.
- Building CPTs with software.

Duration: 70 minutes

Conditional probability.

In the previous class we talked about the topology of the Bayesian networks, the nodes the states (values) of the nodes and how and when those nodes can connect with each other. Once the topology of the Bayesian network is defined, the next thing to do in row is to quantify the relationships between connected nodes. This is done by specifying a conditional probability distribution for each node or else to build the conditional probability table (CPT).

In a Bayesian network every node must have a CPT, associated with it. Conditional probabilities represent likelihoods based on prior information or past experience. A conditional probability is stated mathematically as \( P(x|p_1, p_2, \ldots, p_n) \), which means the probability of variable \( X \) in state \( x \) given parent \( P_1 \) in state \( p_1 \), parent \( P_2 \) in state \( p_2 \), ..., and parent \( P_n \) in state \( p_n \). For each node, CPT depicts all the possible combinations of values of the parent nodes. Each possible combination is called an Instantiation of the parent set. For each such instantiation of parent values there is a row in CPT that describes the probability that the child will take each of its values. As an example recall the Sinus problem from the previous class which for your convenience is shown again in Figure 1.

\[ \text{Figure 1} \]
Bayesian Networks

For the sake of simplicity we replaced the full names of the nodes with their first letter so Flu became Fl, Allergy became A, Sinus became S, Tiredness became Ti, Headache became H.

Let’s suppose now that a medical survey gives us the following data.

\[
P(A) = 0.3 \quad P(F) = 0.4
\]

\[
P(S | A^F) = 0.8 \quad P(S | A^F) = 0.5 \quad P(S | A^F) = 0.4 \quad P(S | A^F) = 0.1
\]

\[
P(H | S) = 0.7 \quad P(H | S) = 0.1 \quad P(Ti | S) = 0.8 \quad P(Ti | S) = 0.1
\]

One possible graphical representation of the conditional probability tables for each node for the model shown in Figure 1 is shown in Figure 2.

<table>
<thead>
<tr>
<th>(P(F=\text{true}))</th>
<th>(P(F=\text{false}))</th>
<th>Prior probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.4</td>
<td>0.6</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>(P(A=\text{true}))</th>
<th>(P(A=\text{false}))</th>
<th>Prior probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.3</td>
<td>0.7</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>(S)</th>
<th>(P(Ti=\text{true}))</th>
<th>(P(Ti=\text{false}))</th>
</tr>
</thead>
<tbody>
<tr>
<td>(T)</td>
<td>0.8</td>
<td>0.2</td>
</tr>
<tr>
<td>(F)</td>
<td>0.1</td>
<td>0.9</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>(F)</th>
<th>(A)</th>
<th>(P(S=\text{true}))</th>
<th>(P(S=\text{false}))</th>
</tr>
</thead>
<tbody>
<tr>
<td>(T)</td>
<td>(T)</td>
<td>0.8</td>
<td>0.2</td>
</tr>
<tr>
<td>(T)</td>
<td>(F)</td>
<td>0.5</td>
<td>0.5</td>
</tr>
<tr>
<td>(F)</td>
<td>(T)</td>
<td>0.4</td>
<td>0.6</td>
</tr>
<tr>
<td>(F)</td>
<td>(F)</td>
<td>0.1</td>
<td>0.9</td>
</tr>
</tbody>
</table>

Let’s take for example the CPT for node \(H\). The upper left column header is labeled \(S\) as this is the name of the parent node. The upper left column header is always labeled with the name of the parent node, i.e. the name of the node that has causal influence over the node in question. In this case, the node has only one parent, so there is only one column labeled \(S\) on that side of the table.
Bayesian Networks

but there may be more as you can see for node S. On the right side of that column there are more columns. These columns are connected with the node the CPT is associated with. The headers of those columns contain the name of the node followed by all the possible state names i.e. values. In this case there are two columns in the right side of the column S, the first one is labeled \( P(H=T) \) as this column contains all the probabilities for node H when it takes the value T (True). In a similar manner, the right most column is labeled \( P(H=F) \) as this column holds all the probabilities for node H when it takes the value F (False). Notice that the probabilities in each row sum up to 1. A more compact graphical explanation of CPT for node H you can see in Figure 3. The CPT of node H can answer questions such as “Given that the patient has sinus problems, what is the probability to have headache?” That question corresponds to the second row of the table, and the answer is 0.7 or 70%.

Roots with no parents, i.e. root nodes, also have an associated CPT, although it is degenerate, containing only one row representing its prior probabilities. In our example, the prior for the patient having Flu is given as 0.4, indicating that 40% of the population in the medical survey had Flu, while 30% had some kind of allergy.

Clearly someone can see that if a node has many parents or if the parents can take a large number of values, the CPT can get very large. In fact, the size of the CPT is exponential in the number of parents. Thus, for Boolean networks a variable with \( n \) parents requires a CPT with \( 2^{n+1} \) probabilities.

![Figure 3](image)

**Dependencies**

In a Bayesian network, there are no direct dependencies in the system being modeled which are not already explicitly shown via arcs. In our Sinus example there is no way for Allergy to influence Headache except by way of causing Sinus or not. That means that there is no hidden connection between Allergy and Headache, every independency suggested by the lack of an arc is real in the system. This direct dependency is known as the Markov property. Bayesian networks that have the Markov property are known as Independence-maps (i-maps)

On the other hand, although the lack of arcs, suggest independencies, the presence of arcs does not correspond necessarily to real dependencies in the system. The CPTs may be parameterized with such values so to negate any dependencies, so modeling a fully connected Bayesian network
Bayesian Networks

someone could represent any joint probability distribution over the variables. Nevertheless, while modeling Bayesian networks we should prefer minimal model and in particular, minimal I-maps. If every arc in a Bayesian network corresponds to real dependencies in the system, then the Bayesian network is called Dependence-map (D-map). If a D-map is I-map in the same time then the network is said to be a perfect map.

Setting the CPTs with software.

To set the CPTs for a Bayesian network using Belief and Decision tool, just do the following steps:

1. Start the application.
2. Click on File menu and load the file Exercise_1.xml that you created in Class 1. The main window should look like the one shown in Figure 4.

3. In the toolbar click the button Modify Probability Table.
4. Click on node Fl. The frame dialog Probability Table for Fl will open as shown in Figure 5, with prior probabilities set to default 0.5. Note that Fl is a root node so CPT contains only one row representing its prior probabilities.
5. Replace value 0.5 in the leftmost text box by typing 0.4. Notice that in the right text box the probability value is changing to 0.6 as the probabilities should sum up to 1.
6. Click OK to save the probability table.
Bayesian Networks

7. Having still selected the button **Modify Probability Table** in the toolbar, click on node **A**, to open the frame dialog **Probability Table for A**.
8. Make sure all the probabilities in the text boxes are filled as shown in **Figure 6**.

9. Click **OK**.
10. Click on node **S**, to open the frame dialog **Probability Table for S**.
11. Make sure all the probabilities in the text boxes are filled as shown in **Figure 7**.
12. Click **OK**.
13. Click on node **Ti**, to open the frame dialog **Probability Table for Ti**.
14. Make sure all the probabilities in the text boxes are filled as shown in **Figure 8**.
15. Click **OK**.
16. Click on node **H**, to open the frame dialog **Probability Table for H**.
17. Make sure all the probabilities in the text boxes are filled as shown in **Figure 9**.
18. Click OK.

19. You can view or change any time the CPT of any node by clicking on Modify Probability Table in toolbar and then by clicking on the node to open the frame dialog.

20. Save the modified Bayesian network in a file with name Exercise_2.xml.