Assignment 4 - Solution

1. Calculate Conditional Probabilities (20pts)

   a. \(P(\text{toothache}) = 0.2\)
   b. \(P(\text{catch}) = 0.34\)
   c. \(P(\text{toothache} | \text{catch}) = \frac{0.108 + 0.016}{0.108 + 0.072 + 0.016 + 0.144} = 0.3647\)
   d. \(P(\text{cavity} | \text{toothache} \lor \text{catch}) = \frac{0.108 + 0.012 + 0.072}{0.108 + 0.012 + 0.016 + 0.064 + 0.072 + 0.144} = 0.4615\)

2. Draw Bayesian Network with CPTs (20pts)

   a. [Diagram of Bayesian Network]

   b. Probability of drawing a coin randomly

<table>
<thead>
<tr>
<th></th>
<th>(P(\text{coin}))</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>(\frac{1}{3})</td>
</tr>
<tr>
<td>b</td>
<td>(\frac{1}{3})</td>
</tr>
<tr>
<td>c</td>
<td>(\frac{1}{3})</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Head</th>
<th>Tail</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>0.4</td>
<td>0.6</td>
</tr>
<tr>
<td>b</td>
<td>0.5</td>
<td>0.5</td>
</tr>
<tr>
<td>c</td>
<td>0.6</td>
<td>0.4</td>
</tr>
</tbody>
</table>
Using Bayes Rule,
\[ P(C | 2H, 1T) = \frac{P(2H, 1T | C) P(C)}{P(2H, 1T)} \]

Since, \( P(C) \) is equal for all values of \( C \) and \( P(2H, 1T) \) is not dependent on \( C \), we can ignore them leading to

\[ P(C | 2H, 1T) \propto P(2H, 1T | C) \]

Since \( X1, X2 \) and \( X3 \) are conditionally independent given \( C \), the RHS from above can be written as:

\[ = P(H | C) P(H | C) P(T | C) \]

First calculate the probability to get head twice and tail once for each coin:

\[ P(2H, 1T | a) = 3 \times (0.4 \times 0.4 \times 0.6) = 0.288 \]

\[ P(2H, 1T | b) = 3 \times (0.5 \times 0.5 \times 0.5) = 0.375 \]

\[ P(2H, 1T | c) = 3 \times (0.6 \times 0.6 \times 0.4) = 0.432 \]

Since each flip is an independent event, we multiply the probabilities of getting head twice and tail once. Also, we need to multiply the probability by 3, because there are 3 different sequences getting head twice and tail once. Therefore, coin \( c \) has the highest probability to get head twice and tail once.
3. Decision Tree Learning (60pts)

Step 1, Calculate Information Gain
{Alt: 0.0, Bar: 0.0, Fri: 0.0207, Pat: 0.5409, Price: 0.1957, Rain: 0.0, Res: 0.0207, Type: 0.0, Est: 0.2075}
Attribute Chosen: Pat

Step 2, Pat: None
All same class examples, result is No

Step 3, Pat: Some
All same class examples, result is Yes

Step 4, Pat: Full
Calculate information gain
{Alt: 0.1092, Bar: 0.0, Fri: 0.1092, Price: 0.2516, Rain: 0.1092, Res: 0.2516, Type: 0.2516, Est: 0.2516}
Attribute chosen: Est (there is tie among Price, Res, Type, and Est)

Step 5, Est: >60
All same class examples, result is No

Step 6, Est: 30-60
Calculate information gain
\{Alt: 0.0, Bar: 1.0, Fri: 1.0, Price: 0.0, Rain: 0.0, Res: 0.0, Type: 1.0\}
Attribute chosen: Bar (There is tie among Bar, Fri, and Type)

Step 7, Bar: No
All same class examples, result is No

Step 8, Bar: Yes
All same class examples, result is Yes

Step 9, Est: 10-30
Calculate information gain
\{Alt: 0.0, Bar: 1.0, Fri: 0.0, Price: 1.0, Rain: 0.0, Res: 1.0, Type: 1.0\}
Attribute chosen: Price (There is tie of Bar, Price, Res, and Type)

Step 10, Price: $
All same class examples, result is Yes

Step 11, Price: $$$
All same class examples, result is No

Step 12, Price: $$
no example, no branch

Step 13, Est: 0-10
no example, no branch