



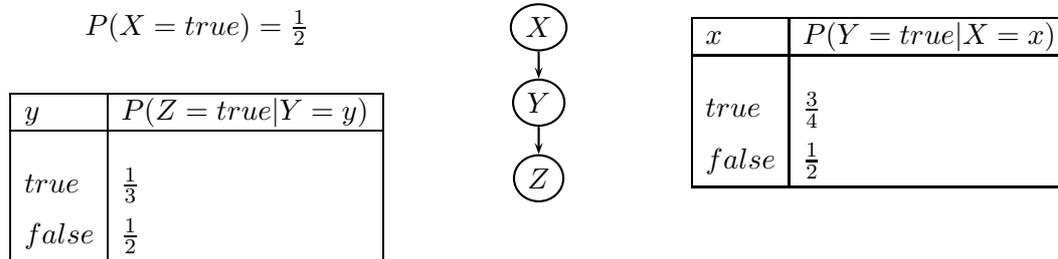
10th Theoretical Assignment in  
 Artificial Intelligence (SS 2005)  
**Solutions**

– This exercise sheet will be discussed in a plenary tutorial on –  
 – wednesday, july 13th 2005 between 16-18 h –  
 – in lecture hall HS002 –

**Exercise 10.1**

(50 P)

Consider the following Bayesian network with three Boolean random variables  $X, Y, Z$ .



Using the Bayesian network above, answer the following questions.

- (a) Compute  $P(Y = \text{false} | X = \text{true})$ : the conditional probability that  $Y$  is false given that  $X$  is true. (5 P)
- (b) Compute  $P(Y = \text{false} | X = \text{false})$ : the conditional probability that  $Y$  is false given that  $X$  is false. (5 P)
- (c) Compute  $P(X = \text{true}, Y = \text{false})$ : the probability that  $X$  is true and  $Y$  is false. (5 P)
- (d) Compute  $P(X = \text{false}, Y = \text{false})$ : the probability that  $X$  is false and  $Y$  is false. (5 P)
- (e) Compute  $P(Y = \text{false})$ : the probability that  $Y$  is false. (5 P)
- (f) Compute  $P(X = \text{true} | Y = \text{false})$ : the conditional probability that  $X$  is true given that  $Y$  is false. (5 P)
- (g) Compute  $P(X = \text{true}, Y = \text{false}, Z = \text{true})$ : the probability that  $X$  and  $Z$  are true but  $Y$  is false. (5 P)
- (h) Compute  $P(X = \text{true}, Z = \text{true})$ : the probability that  $X$  and  $Z$  are true. (5 P)
- (i) Compute  $P(Y = \text{false} | X = \text{true}, Z = \text{true})$ : the probability that  $Y$  is false given that  $X$  and  $Z$  are true. (5 P)
- (j) Does the equation  $P(Y = \text{false} | X = \text{true}, Z = \text{true}) = P(Y = \text{false} | X = \text{true})$  hold? (5 P)

---

**Solution:**

---

- (a)  $P(Y = false|X = true) = 1 - P(Y = true|X = true) = 1 - \frac{3}{4} = \frac{1}{4}$  (or, 0.25).  $P(Y = true|X = true)$  was computed by looking in the given table.
- (b)  $P(Y = false|X = false) = 1 - P(Y = true|X = false) = 1 - \frac{1}{2} = \frac{1}{2}$  (or, 0.5).  $P(Y = true|X = false)$  was computed by looking in the given table.
- (c)  $P(X = true, Y = false) = P(Y = false|X = true)P(X = true) = (\frac{1}{4})(\frac{1}{2}) = \frac{1}{8}$  (or, 0.125) using the value from part (a). The equation  $P(X = true, Y = false) = P(Y = false|X = true)P(X = true)$  is the product rule.
- (d)  $P(X = false, Y = false) = P(Y = false|X = false)P(X = false) = (\frac{1}{2})(\frac{1}{2}) = \frac{1}{4}$  (or, 0.25) using the value from part (b). The first equation is the product rule.
- (e)  $P(Y = false) = P(X = true, Y = false) + P(X = false, Y = false) = \frac{1}{8} + \frac{1}{4} = \frac{3}{8}$  (or, 0.375) using the values from parts (c) and (d). The first equation holds since the set of points (samples) where  $Y = false$  is the disjoint union of the set of points where  $Y = false$  and  $X = true$  and the set of points where  $Y = false$  and  $X = false$ .
- (f)  $P(X = true|Y = false) = \frac{P(X=true, Y=false)}{P(Y=false)} = \frac{1/8}{3/8} = \frac{1}{3}$  (or, 0.3333...) using the values from parts (c) and (e). The first equation is the definition of conditional probability.
- (g)  $P(X = true, Y = false, Z = true) = P(Z = true|Y = false)P(Y = false, X = true) = (\frac{1}{2})(\frac{1}{8}) = \frac{1}{16}$  (or 0.0625) using the value from the conditional probability table for  $Z$  and the value computed in part (c). The first equation holds using the product rule and the fact that we are using a Bayesian network. In particular,  $P(Z = true|Y = false, X = true) = P(Z = true|Y = false)$  holds since  $X$  and  $Y$  are ancestors of  $Z$ , but  $Y$  is the only parent of  $Z$ .
- (h)  $P(X = true, Z = true) = P(X = true, Z = true, Y = false) + P(X = true, Z = true, Y = true) = \frac{1}{16} + P(Z = true|Y = true)P(Y = true, X = true)$ . We know  $P(Z = true|Y = true) = \frac{1}{3}$  by looking at the table for  $Z$  in the network. We compute  $P(Y = true, X = true) = P(Y = true|X = true)P(X = true) = (\frac{3}{4})(\frac{1}{2}) = \frac{3}{8}$ . Hence  $P(X = true, Z = true) = \frac{1}{16} + (\frac{1}{3})(\frac{3}{8}) = \frac{1}{16} + \frac{1}{8} = \frac{3}{16}$  (or 0.1875)
- (i)  $P(Y = false|X = true, Z = true) = \frac{P(Y=false, X=true, Z=true)}{P(X=true, Z=true)} = \frac{1/16}{3/16} = \frac{1}{3}$  (or, 0.33333...) using the values computed in parts (g) and (h).
- (j) No, the equation does not hold.  $P(Y = false|X = true) = \frac{1}{4}$  by part (a), but  $P(Y = false|X = true, Z = true) = \frac{1}{3}$  by part (i).

### Exercise 10.2

(30 P)

Putting together a decision tree is all a matter of choosing which attribute to test at each node in the tree. In the following exercise, you should calculate the information gain for each attribute of the restaurant domain illustrated in Russell/Norvig.

1. To work out the information gain you, firstly, need to calculate the entropy for the set of examples  $S$ . (15 P)
2. For each attribute, use the entropy of  $S$  to calculate the information gain. (15 P)

(Note: You should take  $0 \log_2(0) = 0$ )

---

**Solution:**

---

### 10.2.1 Entropy

$$\text{Alt=YES } \text{info}[3, 3] = \text{entropy}\left(\frac{3}{6}, \frac{3}{6}\right) = -\frac{3}{6}\log_2\left(\frac{3}{6}\right) - \frac{3}{6}\log_2\left(\frac{3}{6}\right) = \frac{1}{2} + \frac{1}{2} = 1$$

$$\text{Alt=NO } \text{info}[3, 3] = \text{entropy}\left(\frac{3}{6}, \frac{3}{6}\right) = -\frac{3}{6}\log_2\left(\frac{3}{6}\right) - \frac{3}{6}\log_2\left(\frac{3}{6}\right) = \frac{1}{2} + \frac{1}{2} = 1$$

$$\text{Bar=YES } \text{info}[3, 3] = \text{entropy}\left(\frac{3}{6}, \frac{3}{6}\right) = -\frac{3}{6}\log_2\left(\frac{3}{6}\right) - \frac{3}{6}\log_2\left(\frac{3}{6}\right) = \frac{1}{2} + \frac{1}{2} = 1$$

$$\text{Bar=NO } \text{info}[3, 3] = \text{entropy}\left(\frac{3}{6}, \frac{3}{6}\right) = -\frac{3}{6}\log_2\left(\frac{3}{6}\right) - \frac{3}{6}\log_2\left(\frac{3}{6}\right) = \frac{1}{2} + \frac{1}{2} = 1$$

$$\text{Fri=YES } \text{info}[2, 3] = \text{entropy}\left(\frac{2}{5}, \frac{3}{5}\right) = -\frac{2}{5}\log_2\left(\frac{2}{5}\right) - \frac{3}{5}\log_2\left(\frac{3}{5}\right) = 0,53 + 0,44 = 0,97$$

$$\text{Fri=NO } \text{info}[4, 3] = \text{entropy}\left(\frac{4}{7}, \frac{3}{7}\right) = -\frac{4}{7}\log_2\left(\frac{4}{7}\right) - \frac{3}{7}\log_2\left(\frac{3}{7}\right) = 0,46 + 0,52 = 0,98$$

$$\text{Hun=YES } \text{info}[5, 2] = \text{entropy}\left(\frac{5}{7}, \frac{2}{7}\right) = -\frac{5}{7}\log_2\left(\frac{5}{7}\right) - \frac{2}{7}\log_2\left(\frac{2}{7}\right) = 0,35 + 0,52 = 0,87$$

$$\text{Hun=NO } \text{info}[1, 4] = \text{entropy}\left(\frac{1}{5}, \frac{4}{5}\right) = -\frac{1}{5}\log_2\left(\frac{1}{5}\right) - \frac{4}{5}\log_2\left(\frac{4}{5}\right) = 0,46 + 0,25 = 0,71$$

$$\text{Pat=NONE } \text{info}[0, 2] = \text{entropy}\left(\frac{0}{2}, \frac{2}{2}\right) = 0$$

$$\text{Pat=SOME } \text{info}[4, 0] = \text{entropy}\left(\frac{4}{4}, \frac{0}{4}\right) = 0$$

$$\text{Pat=FULL } \text{info}[2, 4] = \text{entropy}\left(\frac{2}{6}, \frac{4}{6}\right) = -\frac{2}{6}\log_2\left(\frac{2}{6}\right) - \frac{4}{6}\log_2\left(\frac{4}{6}\right) = 0,53 + 0,39 = 0,99$$

$$\text{Price=\$ } \text{info}[3, 4] = \text{entropy}\left(\frac{3}{7}, \frac{4}{7}\right) = -\frac{3}{7}\log_2\left(\frac{3}{7}\right) - \frac{4}{7}\log_2\left(\frac{4}{7}\right) = 0,52 + 0,46 = 0,98$$

$$\text{Price==\$\$ } \text{info}[2, 0] = \text{entropy}\left(\frac{2}{2}, \frac{0}{2}\right) = 0$$

$$\text{Price==$$$ } \text{info}[1, 2] = \text{entropy}\left(\frac{1}{3}, \frac{2}{3}\right) = -\frac{1}{3}\log_2\left(\frac{1}{3}\right) - \frac{2}{3}\log_2\left(\frac{2}{3}\right) = 0,53 + 0,39 = 0,92$$

$$\text{Rain=YES } \text{info}[3, 2] = \text{entropy}\left(\frac{3}{5}, \frac{2}{5}\right) = -\frac{3}{5}\log_2\left(\frac{3}{5}\right) - \frac{2}{5}\log_2\left(\frac{2}{5}\right) = 0,44 + 0,53 = 0,97$$

$$\text{Rain=NO } \text{info}[3, 4] = \text{entropy}\left(\frac{3}{7}, \frac{4}{7}\right) = -\frac{3}{7}\log_2\left(\frac{3}{7}\right) - \frac{4}{7}\log_2\left(\frac{4}{7}\right) = 0,98$$

$$\text{Res=YES } \text{info}[3, 2] = \text{entropy}\left(\frac{3}{5}, \frac{2}{5}\right) = 0,97$$

$$\text{Res=NO } \text{info}[3, 4] = \text{entropy}\left(\frac{3}{7}, \frac{4}{7}\right) = -\frac{3}{7}\log_2\left(\frac{3}{7}\right) - \frac{4}{7}\log_2\left(\frac{4}{7}\right) = 0,98$$

$$\text{Type=FRENCH } \text{info}[1, 1] = \text{entropy}\left(\frac{1}{2}, \frac{1}{2}\right) = 1$$

$$\text{Type=ITALIEN } \text{info}[1, 1] = \text{entropy}\left(\frac{1}{2}, \frac{1}{2}\right) = 1$$

$$\text{Type=BURGER } \text{info}[2, 2] = \text{entropy}\left(\frac{2}{4}, \frac{2}{4}\right) = 1$$

$$\text{Type=THAI } \text{info}[2, 2] = \text{entropy}\left(\frac{2}{4}, \frac{2}{4}\right) = 1$$

$$\text{Est=0-10 } \text{info}[4, 2] = \text{entropy}\left(\frac{4}{6}, \frac{2}{6}\right) = -\frac{4}{6}\log_2\left(\frac{4}{6}\right) - \frac{2}{6}\log_2\left(\frac{2}{6}\right) = 0,39 + 0,53 = 0,92$$

$$\text{Est=10-30 } \text{info}[1, 1] = \text{entropy}\left(\frac{1}{2}, \frac{1}{2}\right) = 1$$

$$\text{Est=30-60 } \text{info}[1, 1] = \text{entropy}\left(\frac{1}{2}, \frac{1}{2}\right) = 1$$

$$\text{Est=60 } \text{info}[0, 2] = \text{entropy}\left(\frac{0}{2}, \frac{2}{2}\right) = 0$$

### 10.2.2 Information Gain

- $gain(Alt) = info[6, 6] - info([3, 3], [3, 3]) = 1 - (\frac{6}{12} + \frac{6}{12}) = 0 \text{ bits}$
- $gain(Bar) = info[6, 6] - info([3, 3], [3, 3]) = 1 - (\frac{6}{12} + \frac{6}{12}) = 0 \text{ bits}$
- $gain(Fri) = info[6, 6] - info([2, 3], [4, 3]) = 1 - (\frac{5}{12}0,97 + \frac{7}{12}0,98) = 0,02 \text{ bits}$
- $gain(Hun) = info[6, 6] - info([5, 2][1, 4]) = 1 - (\frac{7}{12}0,87 + \frac{5}{12}0,71) = 0,2 \text{ bits}$
- $gain(Pat) = info[6, 6] - info([0, 2], [4, 0], [2, 4]) = 1 - (\frac{6}{12}0,92) = 0,54 \text{ bits}$
- $gain(Price) = info[6, 6] - info([3, 4], [2, 0], [1, 2]) = 1 - (\frac{7}{12}0,98 + \frac{2}{12}0 + \frac{3}{12}0,92) = 0,2 \text{ bits}$
- $gain(Rain) = info[6, 6] - info([3, 2], [3, 4]) = 1 - (\frac{5}{12}0,97 + \frac{7}{12}0,98) = 0,02 \text{ bits}$
- $gain(Res) = info[6, 6] - info([3, 2], [3, 4]) = 1 - (\frac{5}{12}0,97 + \frac{7}{12}0,98) = 0,02 \text{ bits}$
- $gain(Type) = info[6, 6] - info([1, 1], [1, 1], [2, 2], [2, 2]) = 1 - (\frac{2}{12} + \frac{2}{12} + \frac{4}{12} + \frac{4}{12}) = 0 \text{ bits}$
- $gain(Est) = info[6, 6] - info([4, 2], [1, 1], [1, 1], [0, 2]) = 1 - (\frac{6}{12}0,92 + \frac{2}{12} + \frac{2}{12}) = 0,21 \text{ bits}$

### Exercise 10.3 (20 P)

Indicate the decision trees for the following boolean functions:

1.  $A \wedge \neg B$  (5 P)
2.  $A \vee (B \wedge C)$  (5 P)
3. (10 P)

A	B	C	f(A,B,C)
0	0	0	1
0	0	1	1
0	1	0	1
0	1	1	1
1	0	0	0
1	0	1	1
1	1	0	0
1	1	1	1

**Solution:**

