Advance AI  
Ch14: Solution

Nishank Singhal

**14.1** We have a bag of three biased coins a, b, and c with probabilities of coming up heads of 20%, 60%, and 80%, respectively. One coin is drawn randomly from the bag (with equal likelihood of drawing each of the three coins), and then the coin is flipped three times to generate the outcomes X1, X2, and X3.

**a**. Draw the Bayesian network corresponding to this setup and define the necessary CPTs.

**b**. Calculate which coin was most likely to have been drawn from the bag if the observed flips come out heads twice and tails once.

Solution :-  
a) /---> X1

/ |

/ |

/ v

Coin ----> X2

\ |

\ |

\ v

\---> X3

The necessary conditional probability tables (CPTs) are:

1. The prior probability distribution over the three coins:
   1. P(Coin = a) = 1/3
   2. P(Coin = b) = 1/3
   3. P(Coin = c) = 1/3
2. The probability distribution of the outcomes given the coin that was chosen:
   1. P(X1 = H | Coin = a) = 0.2 P(X1 = T | Coin = a) = 0.8
   2. P(X2 = H | Coin = a) = 0.2 P(X2 = T | Coin = a) = 0.8
   3. P(X3 = H | Coin = a) = 0.2 P(X3 = T | Coin = a) = 0.8
   4. P(X1 = H | Coin = b) = 0.6 P(X1 = T | Coin = b) = 0.4
   5. P(X2 = H | Coin = b) = 0.6 P(X2 = T | Coin = b) = 0.4
   6. P(X3 = H | Coin = b) = 0.6 P(X3 = T | Coin = b) = 0.4
   7. P(X1 = H | Coin = c) = 0.8 P(X1 = T | Coin = c) = 0.2
   8. P(X2 = H | Coin = c) = 0.8 P(X2 = T | Coin = c) = 0.2
   9. P(X3 = H | Coin = c) = 0.8 P(X3 = T | Coin = c) = 0.2

b) To calculate which coin was most likely to have been drawn from the bag given the observed flips X1=H, X2=H, and X3=T, we need to use Bayes' theorem:

P(C|X1=H, X2=H, X3=T) = P(X1=H, X2=H, X3=T|C) \* P(C) / P(X1=H, X2=H, X3=T)

∝ P(2heads,1tails|C)P(C)  
∝ P(2 heads, 1 tails|C)

where in the second line we observe that the constant of proportionality 1/P (2 heads, 1 tails) is independent of C, and in the last we observe that P(C) is also independent of the value of C since it is, by hypothesis, equal to 1/3.

Here, **P(C)** is the prior probability of choosing each coin, which is 1/3 for each coin. **P(X1=H, X2=H, X3=T|C)** is the likelihood of the observed flips given each coin, which is determined by multiplying the probabilities of getting H, H, and T given the bias of the coin.  
For coin A, the likelihood is:  
P(X1=H, X2=H, X3=T|A) = 0.2 \* 0.2 \* 0.8 = 0.032  
For coin B, the likelihood is:  
P(X1=H, X2=H, X3=T|B) = 0.6 \* 0.6 \* 0.4 = 0.144  
For coin C, the likelihood is:  
P(X1=H, X2=H, X3=T|C) = 0.8 \* 0.8 \* 0.2 = 0.128

Note that since the CPTs for each coin are the same, we would get the same probability above for any ordering of 2 heads and 1 tail. Since there are three such orderings, we have

P(2heads, 1tails|C = a) = 3 × 0.032 = 0.096   
P(2heads, 1tails|C = b) = 3 × 0.144 = 0.432  
P(2heads, 1tails|C = c) = 3 × 0.128= 0.384

showing that coin b is most likely to have been drawn.

Alternatively, one could directly compute the value of P(C|2 heads, 1 tails).

14.4) Consider the Bayesian network in Figure 14.2.

Diagram

Description automatically generated

a) If no evidence is observed, are Burglary and Earthquake independent? Prove this from the numerical semantics and from the topological semantics.  
b) If we observe Alarm = true , are Burglary and Earthquake independent? Justify your answer by calculating whether the probabilities involved satisfy the definition of conditional independence.

Solution:  
a) Topological Semantics : As burglary and earthquake are not connected to the alarm in the graphic, we may treat them as distinct events. Numerical Semantics : According to the graphic, the scenario can be written as

P(B,E) = P(B)P(E),

where B stands for burglary and E for earthquake.

P(B,E) = P(B|parents(B)) p(E|parents(E)) =P(B) P(E) =0.1 x 0.2 = 0.02

approximately equal to 0 Therefore, we can say that Burglary and Earthquake independent.

b)

P (A) = P (A|B,E) \*P (B) \* P (E) + P (A|B, ¬E) \* P (B) \* P (¬E) + P (A|¬B,E) \* P (¬B) \* P (E) + P (A|¬B, ¬E)P (¬B)P (¬E)

* 0.95 \* 0.001 \* 0.002 + 0.94 \* 0.001 \* 0.998 + 0.29 \* 0.999 \* 0.002 + 0.001 \* 0.999 \* 0.998
* 0.0000019 + 0.00093812 + 0.00057942 + 0.00098901
* 0.00251

P (B|A) = P (A|B)P (B) / P (A) = P (A|B,E)P (B)P (E)+ P (A|B, ¬E)P (B)P (¬E) /P (A)

* 0.95 \* 0.001 \* 0.002 + 0.94 \* 0.001 \* 0.998 / 0.00251
* 0.00094002 / 0.00251
* 0.37451

P (E|A) = P (A, E) / P (A) = P (A, E|B) \* P (B) + P (A, E|¬B) \* P (¬B) / P (A)

* P (A|B,E) \* P (B) \* P (E)+ P (A|¬B,E) \* P (¬B) \* P (E) / P (A)
* 0.95 \* 0.001 \* 0.002 + 0.29 \* 0.999 \* 0.002 / 0.00251
* 0.00058132 / 0.00251
* Therefore,
* P (E|A) = 0.2316
* P (B|A)P(E|A) = 0.37451 × 0.2316 = 0.0867

P (B,E|A) = P (A|B,E)P (B,E) / P (A) = P (A|B,E)P (B)P (E) / P (A)

* 0.95 \* 0.001 \* 0.002 / 0.00251
* 0.000757

So, P (B,E|A) not equal to P (B|A)P(E|A), which implies that they are not independent when alarm = true.

**14.11** In your local nuclear power station, there is an alarm that senses when a temperature gauge exceeds a given threshold. The gauge measures the temperature of the core. Consider the Boolean variables A (alarm sounds), FA (alarm is faulty), and FG (gauge is faulty) and the multivalued nodes G (gauge reading) and T (actual core temperature).

1. **Draw a Bayesian network for this domain, given that the gauge is more likely to fail when the core temperature gets too high.**
2. **Is your network a polytree? Why or why not?**
3. **Suppose there are just two possible actual and measured temperatures, normal and high;**

**the probability that the gauge gives the correct temperature is x when it is working, but**

**y when it is faulty. Give the conditional probability table associated with G.**

1. **Suppose the alarm works correctly unless it is faulty, in which case it never sounds.**

**Give the conditional probability table associated with A.**

**Solution:**a) The key aspects are: the failure nodes are parents of the sensor nodes, and the temperature node is a parent of both the gauge and the gauge failure node. It is exactly this kind of correlation that makes it difficult for humans to understand what is happening in complex systems with unreliable sensors.

**Diagram

Description automatically generated**

b) No matter which way the student draws the network, it should not be a polytree because of the fact that the temperature influences the gauge in two ways.

c) The CPT for G is shown below. Students should pay careful attention to the semantics of FG,

| T=n | T=h |

--------|---------|---------|

G=f, FG | 1-y | y |

--------|---------|---------|

G=f, ~FG| y | 1-y |

--------|---------|---------|

G=t, FG | x | 1-x |

--------|---------|---------|

G=t, ~FG| 1-x | x |

d) The CPT for A is as follows:

| FA=t, FG=t | FA=t, FG=f | FA=f, FG=t | FA=f, FG=f |

-------- |-----------|-----------|-----------|-----------|

T=n, G=n | 1 | 1 | 0 | 0 |

-------- |-----------|-----------|-----------|-----------|

T=n, G=h | 1 | 0 | 0 | 1 |

--------|-----------|-----------|-----------|-----------|

T=h, G=n | 0 | 0 | 1 | 1 |

--------|-----------|-----------|-----------|-----------|

T=h, G=h | 0 | 1 | 1 | 0 |