

# Assignment 4

## Question 1

From the question, we know that:

$$P(\text{positive test results} \mid \text{having the Rett syndrome disease}) = 0.99$$

$$P(\text{having the Rett syndrome disease}) = 0.0001$$

So, using Bayes' theorem, we have that:

$$\begin{aligned} P(\text{positive test results}) &= P(\text{positive test results} \mid \text{having the Rett syndrome disease}) \cdot P(\text{having the Rett syndrome disease}) + P(\text{positive test results} \mid \text{not having the Rett syndrome disease}) \cdot P(\text{not having the Rett syndrome disease}) \\ &= 0.99 \cdot 0.0001 + 0.01 \cdot 0.9999 = 0.010098 \end{aligned}$$

$$\begin{aligned} P(\text{having the Rett syndrome disease} \mid \text{positive test results}) &= \frac{P(\text{positive test results} \mid \text{having the Rett syndrome disease}) \cdot P(\text{having the Rett syndrome disease})}{P(\text{positive test results})} \\ &= 0.99 \cdot 0.0001 / 0.010098 = 0.0098 = 0.98\% \end{aligned}$$

So the chances that this patient does actually have the disease is 0.98%.

## Question 2

Structure 1:

$$P(A,B,C,D) = P(A) \cdot P(B \mid A) \cdot P(C) \cdot P(D \mid A,C)$$

So the minimum number of parameters required to fully specify the distribution is:

$$1+2+1+2 \cdot 2=8$$

Structure 2:

$$P(A,B,C,D) = P(A) \cdot P(B \mid A) \cdot P(C \mid A) \cdot P(D \mid A,C)$$

So the minimum number of parameters required to fully specify the distribution is:

$$1+2+2+2 \cdot 2=9$$

Structure 3:

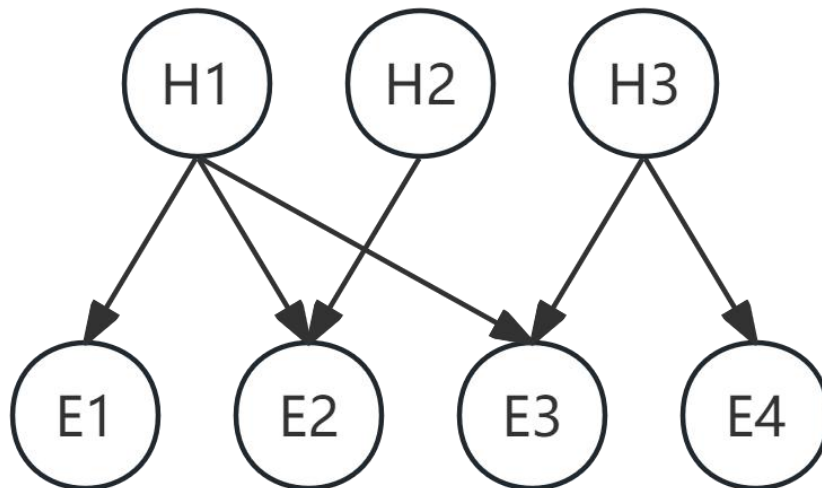
$$P(A,B,C,D) = P(A) \cdot P(B | A) \cdot P(C | A) \cdot P(D | A)$$

So the minimum number of parameters required to fully specify the distribution is:

$$1+2+2+2=7$$

### Question 3

1. The Bayesian network for this problem is as follows:



2.

$$P(H1, H2, H3, E1, E2, E3, E4)$$

$$= P(H1) \cdot P(H2) \cdot P(H3) \cdot P(E1 | H1) \cdot P(E2 | H1, H2) \cdot P(E3 | H1, H3) \cdot P(E4 | H3)$$

3. The number of independent parameters that is required to describe this joint distribution is:

$$1+1+1+2+2*2+2*2+2=15$$

4.

If no conditional independence is assumed, the joint distribution is a single table over all 7 binary variables, requiring:  $2^7 - 1 = 127$  parameters.

5.

E4 depends only on H3, so observing E4 provides direct evidence about H3.

This does not give any new information about H1 or H2 since they are marginally independent of H3 and E4.

6.

E2=True: E2 depends on H1 and H2, so this provides information about H1 and H2.

Since E2 does not depend on H3, this observation does not affect H3.

E4=True: As in (5), E4 provides information about H3. However, since H3 is marginally independent of H1 and H2, observing E4 does not change our beliefs about H1 or H2.

Therefore, observing E4=true tells us about H3, but does not influence our beliefs about H1 or H2 (no additional information gained about them).

## Question 4

From the Bayesian network, we know that:

$$P(e | w1) = \sum_{w2, w3} P(e | w2, w3) \cdot P(w2 | w1) \cdot P(w3 | w1)$$

By calculating based on the different values of w2 and w3, we can obtain that:

$$\begin{aligned} P(e | w1) &= P(e | w2, w3) \cdot P(w2 | w1) \cdot P(w3 | w1) \\ &\quad + P(e | w2, \overline{w3}) \cdot P(w2 | w1) \cdot P(\overline{w3} | w1) \\ &\quad + P(e | \overline{w2}, w3) \cdot P(\overline{w2} | w1) \cdot P(w3 | w1) \\ &\quad + P(e | \overline{w2}, \overline{w3}) \cdot P(\overline{w2} | w1) \cdot P(\overline{w3} | w1) \\ &= 0.3 \times 0.2 \times 0.7 \\ &\quad + 0.25 \times 0.2 \times (1 - 0.7) \\ &\quad + 0.1 \times (1 - 0.2) \times 0.7 \\ &\quad + 0.35 \times (1 - 0.2) \times (1 - 0.7) \\ &= 0.042 + 0.015 + 0.056 + 0.084 \\ &= 0.197 \end{aligned}$$

## Question 5

The probability of system restart P(R) is computed by:

Eliminating S:  $P(D) = \sum_S P(S) \cdot P(D | S)$

Eliminating D:  $P(W) = \sum_D P(D) \cdot P(W | D)$

Eliminating W:  $P(P) = \sum_W P(W) \cdot P(P | W)$

Eliminating P:  $P(R) = \sum_P P(P) \cdot P(R | P)$

So the final expression of P(R) is:

$$P(R) = \sum_P \left[ \sum_W \left[ \sum_D \left[ \sum_S P(S) \cdot P(D | S) \right] \cdot P(W | D) \right] \cdot P(P | W) \right] \cdot P(R | P)$$

## Question 6

$$P(E | S = t) = \frac{P(E, S = t)}{P(S = t)} \quad (1)$$

From the Bayesian network, we know that:

$$P(E, S, W, D) = P(S) \cdot P(E) \cdot P(W | S, E) \cdot P(D | W)$$

$$\text{So, } P(E, S = t) = P(S = t) \cdot P(E) \cdot \sum_W P(W | S = t, E) \cdot [\sum_D P(D | W)]$$

$\sum_D P(D | W)$  is the sum of marginal probabilities and is always 1, so we have:

$$\begin{aligned} P(E, S = t) &= P(S = t) \cdot P(E) \cdot \sum_W P(W | S = t, E) \\ &= f_1(E) \cdot f_2(S = t) \cdot \sum_W f_3(E, S = t, W) \end{aligned}$$

Marginalize  $P(E, S = t)$  over  $E$ , we have:

$$\begin{aligned} P(S = t) &= \sum_E P(E, S = t) \\ &= f_2(S = t) \cdot \sum_E f_1(E) \cdot \sum_W f_3(E, S = t, W) \end{aligned}$$

Combine the formulas of  $P(E, S = t)$  and  $P(S = t)$ , we have that:

$$P(E | S = t) = \frac{f_1(E) \cdot \sum_W f_3(E, S = t, W)}{\sum_E f_1(E) \cdot \sum_W f_3(E, S = t, W)}$$

Substituting offered CPT, we have:

$$\begin{aligned} P(E = t | S = t) &= \frac{f_1(E = t) \cdot \sum_W f_3(E = t, S = t, W)}{\sum_E f_1(E) \cdot \sum_W f_3(E, S = t, W)} \\ &= \frac{0.8 \times (0.4 + 0.7)}{0.8 \times (0.4 + 0.7) + 0.2 \times (0.1 + 0.8)} \\ &= \frac{0.88}{1.06} = 0.83 \end{aligned}$$

Similarly, we have:

$$\begin{aligned} P(E = f | S = t) &= \frac{f_1(E = f) \cdot \sum_W f_3(E = f, S = t, W)}{\sum_E f_1(E) \cdot \sum_W f_3(E, S = t, W)} \\ &= \frac{0.2 \times (0.1 + 0.8)}{0.8 \times (0.4 + 0.7) + 0.2 \times (0.1 + 0.8)} \\ &= \frac{0.18}{1.06} = 0.17 \end{aligned}$$

## Question 7

### Case1: Income.

Fuzzy variable: Low, Medium, High income.

The division of income does not have a clear boundary, so it is suitable to use fuzzy

variables.

### **Case2: Speed.**

Fuzzy variable: Slow, Normal, Fast.

Although speed is a continuous and precise quantity, human perception of speed is usually vague. Therefore, it is appropriate to use fuzzy variables.

### **Case3: A TV show.**

Fuzzy variable: Boring, Neutral, Somewhat interesting, Very interesting

The perception of TV programs is subjective and lacks a clear standard, so it is suitable for using fuzzy variables.

### **Case4: A meal.**

Fuzzy variable: Disgusting, Neutral, Tasty, Delicious

Humans have a vague perception of the taste, smell, and other properties of food, so it is suitable to use fuzzy variables.

### **Case5: Traffic Light.**

There is no need for fuzzy variables. Red, yellow, and green are discrete and distinct categories with no fuzzy boundaries. The ambiguity not only makes no sense but may also lead to decision errors.

## **Question 8**

(1) Max-min composition

$$A1-S1: \max\{\min(0.6, 0.8), \min(0.2, 0.7), \min(0.8, 0.9)\} = \max\{0.6, 0.2, 0.8\} = 0.8$$

$$A1-S2: \max\{\min(0.6, 0.7), \min(0.2, 0.8), \min(0.8, 0.6)\} = \max\{0.6, 0.2, 0.6\} = 0.6$$

$$A1-S3: \max\{\min(0.6, 0.3), \min(0.2, 0.9), \min(0.8, 0.2)\} = \max\{0.3, 0.2, 0.2\} = 0.3$$

$$A2-S1: \max\{\min(0.1, 0.8), \min(0.5, 0.7), \min(0.7, 0.9)\} = \max\{0.1, 0.5, 0.7\} = 0.7$$

$$A2-S2: \max\{\min(0.1, 0.7), \min(0.5, 0.8), \min(0.7, 0.6)\} = \max\{0.1, 0.5, 0.6\} = 0.6$$

$$A2-S3: \max\{\min(0.1, 0.3), \min(0.5, 0.9), \min(0.7, 0.2)\} = \max\{0.1, 0.5, 0.2\} = 0.5$$

$$A3-S1: \max\{\min(0.9, 0.8), \min(0.4, 0.7), \min(0.3, 0.9)\} = \max\{0.8, 0.4, 0.3\} = 0.8$$

$$A3-S2: \max\{\min(0.9, 0.7), \min(0.4, 0.8), \min(0.3, 0.6)\} = \max\{0.7, 0.4, 0.3\} = 0.7$$

$$A3-S3: \max\{\min(0.9, 0.3), \min(0.4, 0.9), \min(0.3, 0.2)\} = \max\{0.3, 0.4, 0.2\} = 0.4$$

So the association of animals with different symptoms is:

$$R \circ T = \begin{bmatrix} 0.8 & 0.6 & 0.3 \\ 0.7 & 0.6 & 0.5 \\ 0.8 & 0.7 & 0.4 \end{bmatrix}$$

(2) Max-product composition

$$A1-S1: \max\{0.6*0.8, 0.2*0.7, 0.8*0.9\} = \max\{0.48, 0.14, 0.72\}=0.72$$

$$A1-S2: \max\{0.6*0.7, 0.2*0.8, 0.8*0.6\} = \max\{0.42, 0.16, 0.48\}=0.48$$

$$A1-S3: \max\{0.6*0.3, 0.2*0.9, 0.8*0.2\} = \max\{0.18, 0.18, 0.16\}=0.18$$

$$A2-S1: \max\{0.1*0.8, 0.5*0.7, 0.7*0.9\} = \max\{0.08, 0.35, 0.63\}=0.63$$

$$A2-S2: \max\{0.1*0.7, 0.5*0.8, 0.7*0.6\} = \max\{0.07, 0.40, 0.42\}=0.42$$

$$A2-S3: \max\{0.1*0.3, 0.5*0.9, 0.7*0.2\} = \max\{0.03, 0.45, 0.14\}=0.45$$

$$A3-S1: \max\{0.9*0.8, 0.4*0.7, 0.3*0.9\} = \max\{0.72, 0.28, 0.27\}=0.72$$

$$A3-S2: \max\{0.9*0.7, 0.4*0.8, 0.3*0.6\} = \max\{0.63, 0.32, 0.18\}=0.63$$

$$A3-S3: \max\{0.9*0.3, 0.4*0.9, 0.3*0.2\} = \max\{0.27, 0.36, 0.06\}=0.36$$

So the association of animals with different symptoms is:

$$R \circ T = \begin{bmatrix} 0.72 & 0.48 & 0.18 \\ 0.63 & 0.42 & 0.45 \\ 0.72 & 0.63 & 0.36 \end{bmatrix}$$