Artificial Intelligence (521495A), Spring 2022
Exercise 2 : Probability and Bayesian networks
Solutions

This handout contains the example solutions for the problems 1-4.

Problem 1. Let’s have a joint probability table of variables ”well-prepared” (W, +w: yes, -w: no) and ”passing the exam” (E, +e: yes, -e: no), when studying students’ success in the course.

W	E	$P(W,E)$
+w	+e	0.45
+w	-e	0.10
-w	+e	0.05
-w	-e	0.40

Calculate the following conditional probabilities

(a) $P(+w \mid +e) = P(+w, +e)/P(+e) = 0.45/0.5 = 0.9$

(b) $P(-w \mid +e) = 1 - P(+w \mid +e) = 1 - 0.9 = 0.1$

(c) $P(+e \mid -w) = P(+e, -w)/P(-w) = 0.05/0.45 \approx 0.11$

(d) $P(W \mid +e) = \alpha P(+e, W) = \alpha < 0.45, 0.05 >=< 0.9, 0.1 >$

Problem 2. A patient takes a cancer test and the result is positive. The test returns a correct positive 97% of the cases and correct negative 95% of the cases. Furthermore 0.004 of the population have the cancer.

Does the patient have cancer or not? Use Bayes rule to find out.

Let use random variables with values

$+c$: cancer
$-c$: no cancer
$+t$: positive test
$-t$: negative test

The known probabilities are

$P(+c) = 0.004$
$P(+t \mid +c) = 0.97$
$P(+t \mid -c) = 1 - P(-t \mid -c) = 1 - 0.95 = 0.05$
$P(-c) = 1 - P(c) = 1 - 0.004 = 0.996$

Use Bayes rule

$$\begin{aligned} P(+c \mid +t) &= \frac{P(+t| + c)P(+c)}{P(+t)} \\ &= \frac{P(+t| + c)P(+c)}{P(+t \mid +c)P(+c) + P(+t \mid -c)P(-c)} \\ &= \frac{0.97 * 0.004}{0.97 * 0.004 + 0.05 * 0.996} \approx 0.07 \end{aligned}$$

$$P(-c \mid +t) = 1 - P(+c \mid +t) \approx 0.93$$

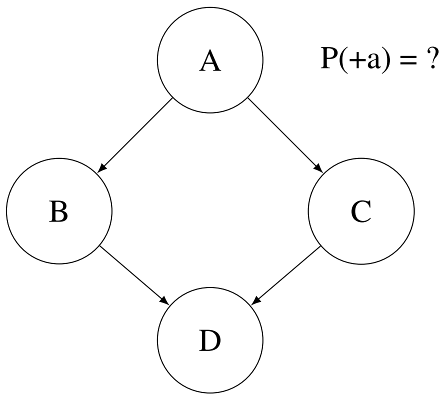
The patient does not have cancer with high probability (93%).

Problem 3. Consider the Bayesian network below. Asssuming that $P(+a,+b,+c,+d) = 1/24$ and $P(-a,-b,-c,-d) = 1/36$.

Calculate the following probabilities

(a) $P(+a)$

(b) $P(+b \mid +a)$



$P(+b \mid -a) = 1/2$
$P(+b \mid +a) =?$

$P(+c \mid +a) = 1/4$
$P(+c \mid -a) = 3/4$

$P(+d \mid -b,-c) = 1/3$
$P(+d \mid +b,+c) = 3/4$

Using chain rule, conditional independencies of variables, and the known probabilities, the joint probabilities can be calculated as follows:

(a) $P(-a,-b,-c,-d) = P(-d \mid -a,-b,-c)P(-c \mid -a,-b)P(-b \mid -a)P(-a) = P(-d \mid -b,-c)P(-c \mid -a)P(-b \mid -a)P(-a) = (1-1/3)(1-3/4)(1-1/2)P(-a) = 1/36$

$P(-a) = 12/36 = 1/3$ and $P(+a) = 2/3$

(a) $P(+a,+b,+c,+d) = P(+d \mid +a,+b,+c)P(+c \mid +a,+b)P(+b \mid +a)P(+a) = P(+d \mid +b,+c)P(+c \mid +a)P(+b \mid +a)P(+a) = (3/4)(1/4)P(+b \mid +a)(2/3) = 1/24$

$P(+b \mid +a) = 48/144 = 1/3.$

Problem 4. Consider a Bayesian network for medical diagnosis where having heart problems (H, +h: yes, -h: no) depends on blood pressure (B, +b: high, -b: low). Furthermore, blood pressure depends on doing exercises (E, +e: yes, -e: no) and having healthy diet (D, +d: yes, -d: no). The probability tables are given below.

(a) Construct a Bayesian network for given problem

(b) Formulate the equation for joint probability distiribution $P(E,D,B,H)$ using the network structure (i.e., the conditional independencies)

(c) Calculate the probability of $P(+e,-d,+b,+h)$

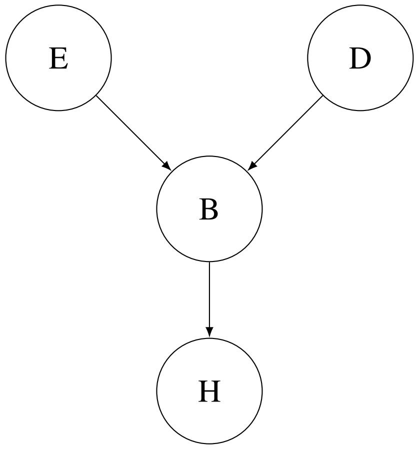
E	$P(E)$
+e	0.5
-e	0.5

D	$P(D)$
+d	0.75
-d	0.25

E	D	B	$P(B \mid E,D)$
+e	+d	+b	0.05
+e	+d	-b	0.95
+e	-d	+b	0.55
+e	-d	-b	0.45
-e	+d	+b	0.45
-e	+d	-b	0.55
-e	-d	+b	0.90
-e	-d	-b	0.10

B	H	$P(H \mid B)$
+b	+h	0.75
+b	-h	0.25
-b	+h	0.05
-b	-h	0.95

(a) The bayes net



(b) $P(E,D,B,H) = P(H \mid B)P(B \mid E,D)P(E)P(D)$

(c) $P(+e,-d,+b,+h) = P(+h \mid +b)P(+b \mid +e,-d)P(+e)P(-d) = 0.75 * 0.55 * 0.5 * 0.25 \approx 0.0516$