

CMSC421 Fall 2005 Homework 4

December 6, 2005

1 Problem 1: Russell & Norvig exercise 13.5 (10 pts)

- a. $\binom{52}{5} = 2,598,960$
- b. $1/2,598,960 = 3.85E - 7$
- c. $4/2,598,960 = 1/649,740 = 1.539E - 6$
 $(13 \times 48)/2,598,960 \approx 2.41E - 4$

2 Problem 2: Russell & Norvig exercise 13.8 (10 pts)

We are given the following information:

$$P(\text{positive}|\text{disease}) = 0.99$$

$$P(\text{negative}|\sim \text{disease}) = 0.99$$

$$P(\text{disease}) = 0.0001$$

and the test result is *positive*. What the patient is concerned about is $P(\text{disease}|\text{positive})$. Roughly speaking, the reason it is a good thing that the disease is rare is that $P(\text{disease}|\text{positive})$ is proportional to $P(\text{disease})$, so a lower prior for *disease* will mean a lower value for $P(\text{disease}|\text{positive})$. If 10,000 people take the test, we expect 1 to actually have the disease, and most likely test positive, while the rest do not have the disease, but 1% of them (about 100 people) will test positive anyway, so $P(\text{disease}|\text{positive})$ will be about 1 in 100. More precisely, using the normalization equation:

$$\begin{aligned} P(\text{disease}|\text{positive}) &= \frac{P(\text{positive}|\text{disease})P(\text{disease})}{P(\text{positive}|\text{disease})P(\text{disease}) + P(\text{positive}|\sim \text{disease})P(\sim \text{disease})} \\ &= \frac{0.99 \times 0.0001}{0.99 \times 0.0001 + (1 - 0.99) \times (1 - 0.0001)} \\ &= .009804 \end{aligned}$$

3 Problem 3: Bayesian Networks (15 pts)

a Compute $P(A)$ in two ways:

(a) By generating the *entire joint distribution* over these variables and explicitly summing the appropriate entries.

S	G	A	$P(S) \cdot P(G S) \cdot P(A S)$	$P(S, A, G)$
t	t	t	$0.3 \cdot 0.7 \cdot 0.6$	0.126
t	t	f	$0.3 \cdot 0.7 \cdot 0.4$	0.084
t	f	t	$0.3 \cdot 0.3 \cdot 0.6$	0.054
t	f	f	$0.3 \cdot 0.3 \cdot 0.4$	0.036
f	t	t	$0.7 \cdot 0.2 \cdot 0.3$	0.042
f	t	f	$0.7 \cdot 0.2 \cdot 0.7$	0.098
f	f	t	$0.7 \cdot 0.8 \cdot 0.3$	0.168
f	f	f	$0.7 \cdot 0.8 \cdot 0.7$	0.392

$$\begin{aligned}
 P(a) &= P(s, a, g) + P(s, a, \bar{g}) + P(\bar{s}, a, g) + P(\bar{s}, a, \bar{g}) \\
 &= 0.39 \\
 P(\bar{a}) &= P(s, \bar{a}, g) + P(s, \bar{a}, \bar{g}) + P(\bar{s}, \bar{a}, g) + P(\bar{s}, \bar{a}, \bar{g}) \\
 &= 0.61
 \end{aligned}$$

(b) Using the variable elimination algorithm

One way:

$$\begin{aligned}
 P(A) &= \sum_{S, G} P(S) \cdot P(A | S) \cdot P(G | S) \\
 &= \sum_S P(S) P(A | S) \cdot F_G(S)
 \end{aligned}$$

where

$$F_G(S) = \sum_G P(G | S) = \frac{S}{t \mid 1.0} \frac{F_G(S)}{f \mid 1.0}$$

$$\begin{aligned}
 P(A) &= \sum_S P(S) \cdot P(A | S) \cdot F_G(S) \\
 &= \frac{A}{t \mid 0.3 \cdot 0.6 + 0.7 \cdot 0.3 = 0.39} \frac{F_S(A)}{f \mid 0.3 \cdot 0.4 + 0.7 \cdot 0.7 = 0.61}
 \end{aligned}$$

b Using conditional independence, compute $P(\bar{g}, a | s)$ and $P(\bar{g}, a | \bar{s})$. Then use Bayes rule to compute $P(s | \bar{g}, a)$.

$$\begin{aligned}
P(\bar{g}, a \mid s) &= P(\bar{g} \mid a, s) \cdot P(a \mid s) \\
&= P(\bar{g} \mid s) \cdot P(a \mid s) \\
&= 0.3 \cdot 0.6 = 0.18
\end{aligned}$$

similarly for $\bar{s}$: $P(\bar{g}, a \mid \bar{s}) = 0.24$

Using Bayes rule:

$$P(s \mid \bar{g}, a) = \frac{P(\bar{g}, a \mid s)P(s)}{P(\bar{g}, a)} = 0.243$$

- c The enterprising cmisc421 student notices that there are two types of people that drive SUVs, people from California (C) and people with large families (F). After collecting some statistics, the student arrives at the following form for the Bayesian network:

Using the chain rule from probability compute the probability $P(\bar{g}, a, s, c, \bar{f})$.

$$P(\bar{g}, a, s, c, \bar{f}) = P(\bar{g} \mid s)P(a \mid s)P(s \mid c, \bar{f})P(c)P(\bar{f}) = 0.02205$$

- d Without explicitly generating the entire probability distribution, compute $P(s)$ and $P(s \mid c)$ and $P(s \mid \bar{c})$. (Hint: consider different values of F.)

$$\begin{aligned}
P(S \mid C) &= \sum_F P(S \mid C, F)P(F) \\
P(S) &= \sum_C P(S \mid C)P(C)
\end{aligned}$$

- e Show how you can compute the value of $P(G)$ in the complete distribution without doing any additional work (i.e., based solely on the work you have already done). Explain.

$$P(G) = \sum_S P(G \mid S)P(S)$$

- f Using the rules for determining when two variables are independent of each other given evidence, answer the following (T/F) for the BN above:

I(C,G) is false
 I(A,G) is false
 I(C,F) is true
 I(C,F | S) is false
 I(C,F | A) is false
 I(F,A | S) is true

4 Problem 4: Bayesian Networks (10 pts)

- a Suppose we have the query $P(G)$ and the following elimination ordering: first A, then B, C, D, E, F. Show what the variable elimination algorithm will do on this graph by listing for each variable eliminated, the new factor generated and the factors used to generate the new factor. The factors should be written in the form $f_i(X_1, \dots, X_k)$ for some set of variables $X_1, \dots, X_k$.

var eliminated	factors used	factor generated
A	$P(C A), P(D A, B), P(A)$	$f_1(B, C, D)$
B	$f_1(B, C, D), P(B)$	$f_2(C, D)$
C	$P(E C, D), f_2(C, D)$	$f_3(D, E)$
D	$P(D G), f_3(D, E)$	$f_4(G, E)$
E	$P(F E, G) f_4(G, E)$	$f_5(F, G)$
F	$f_5(F, G)$	$f_6(G)$

- b Elimination ordering matters to the complexity of inference. Assume all variables have the same domain size $m > 2$. Find an elimination ordering for the same query whose time complexity as a function of m is worse than the ordering given above. As in the above problem, show what the factors generated at each step.

var eliminated	factors used	factor generated
E	$P(E C, D), P(F E, G)$	$f_1(C, D, F, G)$
B	$P(B), P(D A, B)$	$f_2(A, D)$
A	$P(A), P(C A), f_2(A, D)$	$f_3(C, D)$
D	$f_1(C, D, F, G), f_3(C, D), P(G D)$	$f_4(C, F, G)$
C	$f_4(C, F, G)$	$f_5(F, G)$
F	$f_5(F, G)$	$f_6(G)$

There are a number of other possibilities.