G53KRR exercise on Bayesian networks.

This is exercise 2 after Chapter 12 in Brachman and Levesque's book:

Consider the following example: *Metastatic cancer is a possible cause of a brain tumor and is also an explanation for an increased total serum calcium. In turn, either of these could cause a patient to fall into occasional coma. Severe headache could also be explained by a brain tumor.*

(a) Represent these causal links in a belief network. Let a stand for 'metastatic cancer', b for 'increased total serum calcium', c for 'brain tumor', d for 'occasional coma', and e for 'severe headaches'.

(b) Give an example of an independence assumption that is implicit in this network.

(c) Suppose the following probabilities are given: $Pr(a) = 0.2, Pr(b|a) = 0.8, Pr(b|\neg a) = 0.2, Pr(c|a) = 0.2, Pr(c|\neg a) = 0.05, Pr(e|c) = 0.8, Pr(e|\neg c) = 0.6, Pr(d|b \wedge c) = 0.8, Pr(d|b \wedge \neg c) = 0.8, Pr(d|\neg b \wedge c) = 0.8, Pr(d|\neg b \wedge \neg c) = 0.05$ and assume that it is also given that some patient is suffering from severe headaches but has not fallen into a coma. Calculate joint probabilities for the eight remaining possibilities (that is, according to whether a, b, and c are true or false).

(d) According to the numbers given, the a priori probability that the patient has metastatic cancer is 0.2. Given that the patient is suffering from severe headaches but has not fallen into a coma, are we now more or less inclined to believe that the patient has cancer? Explain.

Answers :

(a) Sorry for an ascii drawing. The main thing here is that arcs go from cause (e.g. brain tumor) to effect (e.g. headache). Other layouts like the one I did on the board are OK too.

```
a ------> b
|         |
|         |
|         |
v         v
c ------> d
|
|
|
v
e
```

(b) Examples are:

- $Pr(c \mid a \wedge b) = Pr(c \mid a)$, $Pr(c \mid \neg a \wedge b) = Pr(c \mid \neg a)$ etc.
- $Pr(d \mid a \wedge b \wedge c) = Pr(d \mid b \wedge c)$
- $Pr(e \mid a \wedge b \wedge c \wedge d) = Pr(e \mid c)$

(c) I spell out the computation of the probability of the first conjunction in more detail, after that I will skip the chain rule and use the negation rule without mentioning it.

1. $Pr(a \wedge b \wedge c \wedge \neg d \wedge e) =$ (using the normal chain rule)
   $Pr(a) \cdot Pr(b \mid a) \cdot Pr(c \mid a \wedge b) \cdot Pr(\neg d \mid a \wedge b \wedge c) \cdot Pr(e \mid a \wedge b \wedge c \wedge \neg d) =$ (substituting conditional probabilities using independence assumptions of the network)
   $Pr(a) \cdot Pr(b \mid a) \cdot Pr(c \mid a) \cdot Pr(\neg d \mid b \wedge c) \cdot Pr(e \mid c) =$
   (using the negation rule $Pr(\neg d \mid b \wedge c) = 1 - Pr(d \mid b \wedge c)$)
   $Pr(a) \cdot Pr(b \mid a) \cdot Pr(c \mid a) \cdot (1 - Pr(d \mid b \wedge c)) \cdot Pr(e \mid c) =$
   $0.2 \cdot 0.8 \cdot 0.2 \cdot 0.2 \cdot 0.8 = 0.00512$
2. $Pr(a \wedge b \wedge \neg c \wedge \neg d \wedge e) =$
   $Pr(a) \cdot Pr(b \mid a) \cdot (1 - Pr(c \mid a)) \cdot (1 - Pr(d \mid b \wedge \neg c)) \cdot Pr(e \mid \neg c) =$
   $0.2 \cdot 0.8 \cdot 0.8 \cdot 0.2 \cdot 0.6 = 0.01536$
3. $Pr(a \wedge \neg b \wedge c \wedge \neg d \wedge e) =$
   $Pr(a) \cdot (1 - Pr(b \mid a)) \cdot Pr(c \mid a) \cdot (1 - Pr(d \mid \neg b \wedge c)) \cdot Pr(e \mid c) =$
   $0.2 \cdot 0.2 \cdot 0.2 \cdot 0.2 \cdot 0.8 = 0.00128$
4. $Pr(a \wedge \neg b \wedge \neg c \wedge \neg d \wedge e) =$
   $Pr(a) \cdot (1 - Pr(b \mid a)) \cdot (1 - Pr(c \mid a)) \cdot (1 - Pr(d \mid \neg b \wedge \neg c)) \cdot Pr(e \mid \neg c) =$
   $0.2 \cdot 0.2 \cdot 0.8 \cdot 0.95 \cdot 0.6 = 0.01824$
5. $Pr(\neg a \wedge b \wedge c \wedge \neg d \wedge e) =$
   $(1 - Pr(a)) \cdot Pr(b \mid \neg a) \cdot Pr(c \mid \neg a) \cdot (1 - Pr(d \mid b \wedge c)) \cdot Pr(e \mid c) =$
   $0.8 \cdot 0.2 \cdot 0.05 \cdot 0.2 \cdot 0.8 = 0.00128$
6. $Pr(\neg a \wedge b \wedge \neg c \wedge \neg d \wedge e) =$
   $(1 - Pr(a)) \cdot Pr(b \mid \neg a) \cdot (1 - Pr(c \mid \neg a)) \cdot (1 - Pr(d \mid b \wedge \neg c)) \cdot Pr(e \mid \neg c) =$
   $0.8 \cdot 0.2 \cdot 0.95 \cdot 0.2 \cdot 0.6 = 0.01824$
7. $Pr(\neg a \wedge \neg b \wedge c \wedge \neg d \wedge e) =$
   $(1 - Pr(a)) \cdot (1 - Pr(b \mid \neg a)) \cdot Pr(c \mid \neg a) \cdot (1 - Pr(d \mid \neg b \wedge c)) \cdot Pr(e \mid c) =$
   $0.8 \cdot 0.8 \cdot 0.05 \cdot 0.2 \cdot 0.8 = 0.00512$
8. $Pr(\neg a \wedge \neg b \wedge \neg c \wedge \neg d \wedge e) =$
   $(1 - Pr(a)) \cdot (1 - Pr(b \mid \neg a)) \cdot (1 - Pr(c \mid \neg a)) \cdot (1 - Pr(d \mid b \wedge \neg c)) \cdot Pr(e \mid \neg c) =$
   $0.8 \cdot 0.8 \cdot 0.95 \cdot 0.95 \cdot 0.6 = 0.34656$

(d) We are asked whether $Pr(a \mid \neg d \wedge e)$ is greater or smaller than $Pr(a)$.

$Pr(a \mid \neg d \wedge e) = Pr(a \wedge \neg d \wedge e)/Pr(\neg d \wedge e)$ (conditional probability definition). We need to compute $Pr(a \wedge \neg d \wedge e)$ and $Pr(\neg d \wedge e)$, and to do that we use the probabilities we computed above. They describe all 8 possible states of the world given that $\neg d$ and e are true, and they are all disjoint. We are using $Pr(X) = Pr(X \wedge Y) + Pr(X \wedge \neg Y)$, or that the probability of the union of disjoint events equals to the sum of probabilities of those events.

So $Pr(a \wedge \neg d \wedge e) = Pr(a \wedge b \wedge c \wedge \neg d \wedge e) + Pr(a \wedge b \wedge \neg c \wedge \neg d \wedge e) + Pr(a \wedge \neg b \wedge c \wedge \neg d \wedge e) + Pr(a \wedge \neg b \wedge \neg c \wedge \neg d \wedge e)$ and $Pr(\neg d \wedge e)$ is the sum of all 8 numbers above.

$Pr(a \wedge \neg d \wedge e) = 0.04$

$Pr(\neg d \wedge e) = 0.04 + 0.00128 + 0.01824 + 0.00512 + 0.34656 = 0.4112$

$Pr(a \mid \neg d \wedge e) = 0.04/0.4112$ which is approximately 0.1. So the probability got smaller.