Bayesian networks

Note Title

We start with some basic probability theory, then discuss Bayesian networks.

1. Independence

Events A and B are independent if

$$P(A \wedge B) = P(A)\,P(B)$$

e.g. suppose $P(A) = 0.5$, $P(B) = 0.7$, $P(C) = 0.3$
$P(A \wedge B) = 0.35$ $P(A \wedge C) = 0.2$

exercise: are A and B independent?
are A and C independent?

2. Conditional probability

The conditional probability of A given B is

$$P(A|B) = \frac{P(A \wedge B)}{P(B)}$$

exercise: Using above probabilities for A, B, C, compute:
i) $P(A|B)$
ii) $P(A|C)$

3. Bayes' rule

Bayes' rule tells us how to reverse a conditional probability:

$$P(A|B) = \frac{P(B|A)\,P(A)}{P(B)}$$

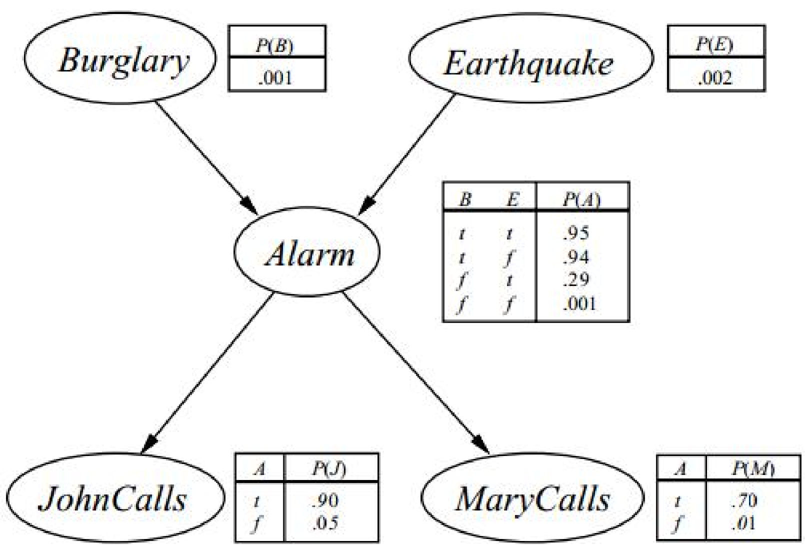
Exercise: Using above probabilities, compute:

i) $P(B|A)$

ii) $P(C|A)$

4. Bayesian network

Example from the textbook (fig 14.2):



Informal description: Arrows show direct influence of one variable on another.

Why are Bayesian networks interesting and important?
Two reasons:

1. They can accurately represent many real problems
2. Efficient algorithms exist for computing with them
   - Specifically, a problem with N binary variables requires 2^N numbers to describe its probability distribution. But a Bayesian network needs only $O(N 2^K)$ operations to compute answers to typical problems. (Here, K is the maximum number of parents of any node.)

Formal definition: A Bayesian network is

- a directed acyclic graph
- each node is a random variable X_i
- each X_i has prob. dist. $P(X_i \mid Parents(X_i))$
- the joint distribution is the product of these conditionals

the distribution of all variables simultaneously: $p(X_1, X_2, \dots X_N)$.

Example: Fig 14.2 in textbook (reproduced above)

Exercise: Write down the joint distribution of all variables in this network

Solution: $P(B)\,P(E)\,P(A \mid B, E)\,P(J \mid A)\,P(M \mid A)$

Further exercises (solutions below): Compute the following:

1. $P(B)$, $P(\neg B)$
2. $P(E)$, $P(\neg E)$
3. $P(A \mid B, E)$ $\quad$ $P(\neg A \mid B, E)$
   $P(A \mid B, \neg E)$ $\quad$ $P(\neg A \mid B, \neg E)$
   $P(A \mid \neg B, E)$ $\quad$ $P(\neg A \mid \neg B, E)$
   $P(A \mid \neg B, \neg E)$ $\quad$ $P(\neg A \mid \neg B, \neg E)$
4. $P(A, E \mid B)$, $P(A, \neg E \mid B)$
5. $P(A \mid B)$, $P(\neg A \mid B)$
6. $P(J \mid B)$, $P(\neg J \mid B)$
7. $P(J \mid \neg B)$, $P(\neg J \mid \neg B)$
8. $P(A \mid \neg B)$, $P(\neg A \mid \neg B)$
9. $P(B \mid J)$, $P(\neg B \mid J)$

(In class, we do only 9, assuming 1–8 are done already)

Solutions to
Example calculations for the alarm network, Fig 14.2 p512.
(All calculations to 3 significant figures.)

read directly off the figure

① $P(B) = 0.001$ $\qquad P(\neg B) = 0.999$

② $P(E) = 0.002$ $\qquad P(\neg E) = 0.998$

③
$$P(A \mid B, E) = 0.95 \qquad P(\neg A \mid B, E) = 0.05$$
$$P(A \mid B, \neg E) = 0.94 \qquad P(\neg A \mid B, \neg E) = 0.06$$
$$P(A \mid \neg B, E) = 0.29 \qquad P(\neg A \mid \neg B, E) = 0.71$$
$$P(A \mid \neg B, \neg E) = 0.001 \qquad P(\neg A \mid \neg B, \neg E) = 0.999$$

④
$$P(A, E \mid B) = \underbrace{P(A \mid E, B)}_{=0.95 \text{ from above}} \; \underbrace{P(E \mid B)}_{}$$

follows from definition of conditional probability, with a "$|B$" added everywhere: $P(X, Y) = P(X \mid Y)\, P(Y)$, so $P(X, Y \mid B) = P(X \mid Y, B)\, P(Y \mid B)$

look at the network – E and B are independent – so $P(E \mid B) = P(E) = 0.002$

$$= 0.95 \times 0.002$$
$$= 0.0019$$

$$P(A, \neg E \mid B) = P(A \mid \neg E, B)\, P(\neg E \mid B)$$ (same reasoning as above)
$$= 0.94 \times 0.998$$
$$0.938$$

(5) $P(A|B) = P(A, E|B) + P(A, \neg E|B)$]– sum over all possibilities of other relevant variables

$= 0.0019 + 0.938$]– from (4)

$= 0.940$

$P(\neg A|B) = 1 - P(A|B)$]– by definition

$= 0.060$

(6) $P(J|B) = P(J, A|B) + P(J, \neg A|B)$]– sum over all possibilities of other relevant variables

$= P(J|A, B)\,P(A|B) + P(J|\neg A, B)\,P(\neg A|B)$] (from definition of cond prob, as in (4))

using conditional independence in the network

$= P(J|A)\,P(A|B) + P(J|\neg A)\,P(\neg A|B)$

$= 0.9 \times 0.940 + 0.05 \times 0.060$

$= 0.849$

$P(\neg J|B) = 1 - P(J|B) = 1 - 0.849$

$= 0.151$

(7) $P(J \mid \neg B) =$ [all reasoning is analagous to (6) and is omitted]

$$= P(J, A \mid \neg B) + P(J, \neg A \mid \neg B)$$
$$= P(J \mid A)\,\underbrace{P(A \mid \neg B)}\, + P(J \mid \neg A)\,\underbrace{P(\neg A \mid \neg B)}$$

$$P(A \mid \neg B) = P(A, E \mid \neg B) + P(A, \neg E \mid \neg B)$$
$$= P(A \mid E, \neg B)P(E) + P(A \mid \neg E, \neg B)P(\neg E)$$
$$= 0.29 \times 0.002 + 0.001 \times 0.998$$
$$= 0.00158$$

$$P(\neg A \mid \neg B) = 1 - P(A \mid \neg B) = 1 - 0.00158$$
$$= 0.998$$

$$= 0.9 \times 0.00158 + 0.05 \times 0.998$$
$$= 0.0513$$

$$P(\neg J \mid \neg B) = 1 - P(J \mid \neg B) = 1 - 0.0513$$
$$= 0.949$$

(8)
$$
\begin{aligned}
P(A \mid \neg B) &= P(A, E \mid \neg B) + P(A, \neg E \mid \neg B) \\
&= P(A \mid E, \neg B)\, P(E) + P(A \mid \neg E, \neg B)\, P(\neg E) \\
&= 0.29 \times 0.002 + 0.001 \times 0.998 \\
&= 0.00158
\end{aligned}
$$

same reasoning as (4) and (5) combined

$$
\begin{aligned}
P(\neg A \mid \neg B) &= 1 - P(A \mid \neg B) \\
&= 0.998
\end{aligned}
$$

(9)
$$
P(B \mid J) = \underbrace{P(J \mid B)}_{\text{know from (6)}}\ \underbrace{P(B)}_{\text{know from (1)}} \Big/ \underbrace{P(J)}
$$
(Bayes' rule)

Could calculate $P(J)$ explicitly, but this is a lengthy calculation. We will calculate it implicitly, later. Leave it as $P(J)$ for now.

$$
\begin{aligned}
&= 0.849 \times 0.001 \,/\, P(J) \\
&= 0.000849 \,/\, P(J)
\end{aligned}
$$

$$
P(\neg B \mid J) = \underbrace{P(J \mid \neg B)}_{\text{know from (7)}}\ \underbrace{P(\neg B)}_{\text{know from (1)}} \Big/ \underbrace{P(J)}_{\text{leave as } P(J) \text{ for now}}
$$
(Bayes' rule)

$$
\begin{aligned}
&= 0.0513 \times 0.999 \,/\, P(J) \\
&= 0.0512 \,/\, P(J)
\end{aligned}
$$

Now, we know that

$$P(B|J) + P(\neg B|J) = 1$$

Thus, we have:

$$0.000849 / P(J) + 0.0512 / P(J) = 1$$

Rearrange: $P(J) = 0.000849 + 0.0512$
$= 0.0521$ ← thus, P(J) has been calculated !!

Finally, we have

$$P(B|J) = 0.000849 / P(J) = 0.000849 / 0.0521$$
$$= 0.0163$$