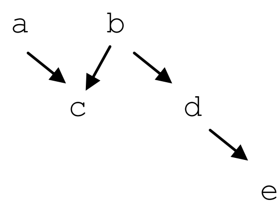
Ivan Bratko

PROLOG PROGRAMMING FOR ARTIFICIAL INTELLIGENCE, 4TH EDN.
ADDITIONAL EXERCISES IN BAYESIAN NETWORKS

Note: These exercises are supplementary to those in the book: I. Bratko, Prolog Programming for Artificial Intelligence, 4th edn., Pearson Education 2011.

Problem

Let a Bayesian network have the following structure:



This structure is also defined by the relation **parent/2** below. The relation **p/3** defines the probabilities in this network.

parent(a,c). parent(b,c). parent(b,d). parent(d,e).

```
p(a, 0.1).                   % Probability of a is 0.1
p(b, 0.1).
p(c, [a,b], 0.9).            % Conditional probability p(c|ab) = 0.9
p(c, [not a, b], 0.6).       % p(c|~ab) = 0.6
p(c, [a, not b], 0.8).
p(c, [not a, not b], 0.3).
p(d, [b], 0.9).
p(d, [not b], 0.1).
p(e, [d], 0.1).
p(e, [not d], 0.9).
```

For the questions a-c estimate, without numerical calculation, which of the two probabilities is greater than the other. Briefly justify your answers.

(a) p(c) or p(c|d)?

(b) p(a|c) or p(b|c)?

(c) p(a|c) or p(a|ce)?

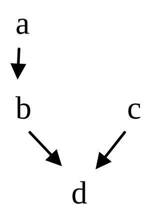
(d) Calculate (numerically) the probability p(e).

Answers

(a) p(c) < p(c | d), because d increases probability of b which in turn increases probability of c
(b) p(a | c) > p(b | c), because both a and b are a priori equally likely, and a is more likely (than b) to have caused c
(c) p(a | c) < p(a | c e), because if e then d is less likely, and so is b; since c is given, its cause a or b is needed, and since b is less likely in the case of e, a must be more likely
(d) p(e) = 0.756

Problem

Consider the following Bayesian network:



Let both a and c be very rare events, and let all the links in the network correspond to the causality between events. So p(b | a) >> p(b |~a). Also d is is much more likely when b and/or c happen. Determine the relation (<, >, or =) between the following probabilities:

(a) p(a) : p(a | c)
(b) p(a) : p(a | d)
(c) p(a | d) : p(a | c d)
(d) p(d | b c) : p(d | a b c)

Problem

Consider the Bayesian network:

$$
\begin{tikzcd}
 & a \arrow[dl] \arrow[dr] & \\
b \arrow[d] & & d \\
c & &
\end{tikzcd}
$$

Assume there are all the appropriate (conditional) probabilities given for this network.
(a) Derive the formula for computing the conditional probability

p(c | a) = …

for this network in terms of the probabilities in the network
(b) Give the formula for computing p(a | b) for this network.

Problem
The Prolog program below defines a Bayesian network. Predicate parent/2 defines the structure of the network. Predicate p/3 defines the corresponding probabilities.

parent(a,c). parent(b,c). parent(b,d).

p(a, 0.1). % p(a) = 0.1
p(b, 0.1).
p(c, [a,b], 0.9). % p(c | a and b) = 0.9
p(c, [not a, b], 0.1). % p(c | not a and b) = 0.1
p(c, [a, not b], 0.9).
p(c, [not a, not b], 0.1).
p(d, [b], 0.9).
p(d, [not b], 0.1).

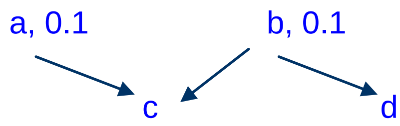
Estimate without calculation which probability is higher
(a) p(b) or p(b | d)? Explain why.

(b) p(a) or p(a | d)? Explain why.

(c) Calculate the probability p(a | b c).

Answers

Diagram helps:



a	b	p(c)
1	1	0.9
0	1	0.1
1	0	0.9
0	1	0.1

b	p(d)
1	0.9
0	0.1

(a) p(b) < p(b | d) because from conditional probabilities for d it follows that b causes d; so knowing d has happened, b is more likely

(b) p(a) = p(a | d) because a is independent of d (d is a's non-descendant)

(c) Conditional probability table for c shows that c does not depend on b at all. Therefore:

p(a | b and c) = p(a | c)

Using Bayes rule:

p(a | c) = p(a) * p(c | a) / p(c) = 0.1 * 0.9 / 0.18 = 0.5

So: p(a | b and c) = 0.5

p(c) = p(a)*p(c | a) + p(not a) * p(c | not a) = 0.1*0.9 + 0.9*0.1 = 0.18

Problem

Let there be an alarm device in a house against burglary. We consider three events that can possibly happen: burglary, storm, and alarm. Alarm may be caused by burglary or storm. We also have the following data: the probability of burglary is 0.001, the probability of storm is 0.01. Alarm cannot happen on its own (without burglary or storm). Burglary alone causes the alarm to go on in 90% of the cases. Storm alone causes the alarm to go on in 1% of the cases. Burglary and storm are probabilistically independent, and they independently cause alarm.

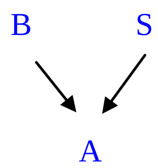
(a) Draw a Bayesian network for this problem, so that the network reflects the above stated causality relations.

(b) Specify approximate values of all the needed conditional and unconditional probabilities for this Bayes network.

(c) Specify exact values of all the needed conditional and unconditional probabilities for this Bayes network.

Answers

(a)



(b) p(B) = 0.001, p(S) = 0.01

B	S	p(A)
F	F	0.00
F	T	0.01
T	F	0.90
T	T	0.90

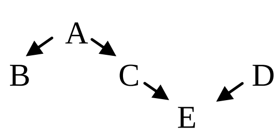
(c) Everything the same as in (b) except:

p(A | BS) = p(A caused by B) + (1-p(A caused by B) * p(A|S) =

0.90 + 0.10*0.01 = 0.901

Problem

Consider a Bayesian network with nodes A, B, C, D and E. The nodes represent events (that is, their possible values are just true or false). The structure of the network is:



The structure of the network enables the simplification of some conditional probabilities. For example, $p(E \mid A \wedge C) = p(E \mid C)$. Here, the condition part $A \wedge C$ was simplified into C. In the conditional probabilities below, how can the condition parts be simplified as much as possible by considering the structure of this Bayesian network?

(a) $p(C \mid \sim A \wedge B \wedge E)$

(b) $p(C \mid \sim A \wedge B \wedge D \wedge E)$
(c) $p(C \mid \sim B \wedge D \wedge E)$
(d) Express the conditional probability $p(B \mid C)$ in terms of probabilities that must be given by the definition of this network.

Answers
(a) $p(C \mid \sim A \wedge E)$
(b) $p(C \mid \sim A \wedge D \wedge E)$
(c) $p(C \mid \sim B \wedge D \wedge E)$ (not possible to simplify!)
(d) $p(B \mid C) = p(BC) / p(C) = (p(BC \mid A)p(A) + p(BC|\sim A) p(\sim A)) / p(C) =$
$= (p(A) p(B \mid A) p(C \mid A) + p(\sim A) p(B|\sim A) p(C|\sim A)) / (p(A) p(C \mid A) + p(\sim A) p(C|\sim A))$

Problem

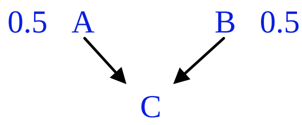
(a) Let A, B and C be Boolean variables with possible values 0 and 1, and there is the relation between these variables such that C = XOR(A,B) (that is C = (A+B) mod 2). Draw a Bayesian network and define the corresponding probabilities for this network which correspond to this relation. Let the links in the network indicate that A and B are the causes of C. Let prior probabilities of A and B be 0.5.

(b) Add to this network two additional nodes, D and K, and the corresponding links and probabilities so that this new network represents the following situation. We are testing in a written exam whether a student knows the operation XOR. In the exam problem, the student is given the values of A and B, an is asked to calculate XOR(A,B). The student's answer is D. D should ideally be equal C. But D may be different from C in the case the student does not know the topic. Even if the student knows the topic, the answer may be incorrect due to the student's silly mistake. The variable K = 1 if the student knows the XOR operation, otherwise K = 0. If the student knows the operation then his/her answer D will be correct in 99% of the cases. If the student does not know XOR, then the answer D will be chosen completely randomly with equal probabilities of 0 and 1. Let the links in the Bayesian network indicate that C and K are causes of D. The prior probability of K is 0.5.

(c) Let A=0, B=1 and D=1. Estimate what is, under this evidence, approximate probability that the student knows the topic: $p(K \mid \sim A\ B\ D)$.

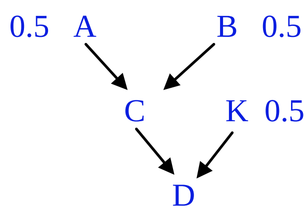
Answers

(a)



A	B	p(C)
0	0	0.0
1	0	1.0
0	1	1.0
1	1	0.0

(b)



C	K	p(D)
0	0	0.5
1	0	0.5
0	1	0.01
1	1	0.99

(c) $p(K \mid \sim A\ B\ D) =$

$p(K \mid C\ D) = p(K)\ p(CD \mid K) / p(C\ D) \approx 0.5 * 0.5 / 0.375 = 0.6667$

$p(C\ D \mid K) = p(C \mid K)\ p(D \mid C\ K) \approx 0.5 * 1.0$

$p(C\ D) = p(K) * p(CD \mid K) + p(\sim K) * p(CD \mid \sim K) \approx 0.5 * 0.5 + 0.5 * 0.5 * 0.5 = 0.375$

Here, the calculation was simplified by assuming that if the student knows the topic he/she never makes a mistake. More precise calculation would give $p(K \mid \sim A\ B\ D) = 0.6644295$