Started on	Wednesday, 24 July 2024, 9:32 PM
State	Finished
Completed on	Wednesday, 31 July 2024, 10:23 PM
Time taken	7 days
Marks	13/13
Grade	**10** out of 10 (**100**%)

Question 1

Correct

Mark 1 out of 1

A Bayesian network models a domain

Select one or more:

- [ ] a. as a joint probability distribution of observable variables.
- [x] b. as a joint probability distribution of random variables. Note that not all of them must be observable! If some of their values are hidden, this just complicates the training process.
- [ ] c. as a mapping from inputs to outputs.

Your answer is correct.

Correct

Marks for this submission: 1/1.

Question **2**

Correct

Mark 3 out of 3

Bayesian networks use a graph representation for modelling a problem. What is true about this graph representation?

The **graph** is

- ◉ acyclic and directed.
- ○ is directed, and the underlying undirected graph is acyclic.

Mark 1 out of 1

The **nodes** of the graph correspond to

- ◉ random variables.
- ○ observable random variables.
- ○ random variable values.

Mark 1 out of 1

A **link** in the graph between nodes A and B models that

- ○ values of A **influence** the probability of values of B
- ○ certain values of A **causally imply** values of B
- ◉ values of A **directly influence** the probability of values of B

Mark 1 out of 1

Correct

Marks for this submission: 3/3.

Question 3

Correct

Mark 2 out of 2

Special features characterizing a **Naive Bayes Classifier** are:

- [ ] They only have one child node, accepting as parent all other variables.
- [ ] They assume that all but one variable are independent.
- [x] They assume that all but one variable are conditionally independent.
- [x] They only have one root node serving as parent for all other variables.

Mark 1 out of 1

When do we talk about a **mixed Bayesian net**? The network both features

- ( ) polytree and non-polytree sub-networks.
- (•) continuous and discrete variables.
- ( ) dependent and independent variables.

Mark 1 out of 1

Correct

Marks for this submission: 2/2.

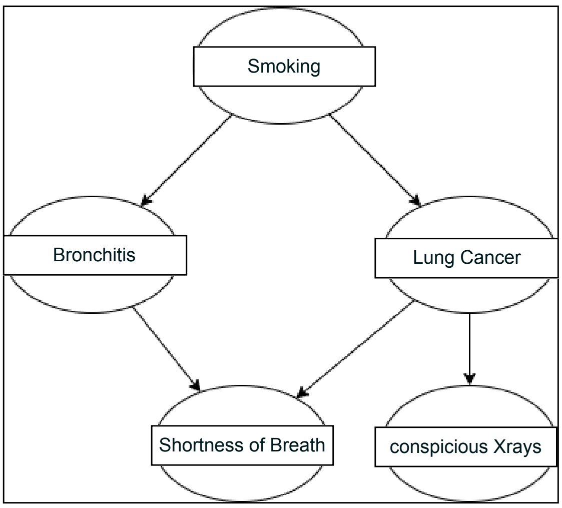
Question 4

Correct

Mark 1 out of 1

Let's construct a Bayesian network from some prior information about our domain.

Assume you are given the following information by doctors: *The goal is to be able to diagnose lung cancer. We know that **lung cancer** more often occurs given that a **person smokes**; the same holds for **Bronchitis**. A symptom for both of these diseases is **shortness of breath**. Independently of the patient's observation, the effects of lung cancer on the tissue can sometimes be spotted directly in **X-ray** images.*



Your answer is correct.

Correct

Marks for this submission: 1/1.

Question 5

Correct

Mark 1 out of 1

Given the Bayesian net from the previous question, which of the following statements are true?

Select one or more:

- [x] a. Given the lung cancer value, occurrence of conspicious Xrays is conditionally independent from shortness of breath.
- [x] b. Given the smoking value, lung cancer and bronchitis are conditionally independent.
- [ ] c. The only observable variables are Xrays and Shortness of Breath.
- [x] d. The directed graph is free of cycles.
- [x] e. There is no variable (absolutely) independent of all others.
- [ ] f. Smoking is independent of all other variables.
- [ ] g. Given the bronchitis value, lung cancer and shortness of breath are conditionally independent.
- [ ] h. This is a polytree Bayesian network.

Your answer is correct.

Correct

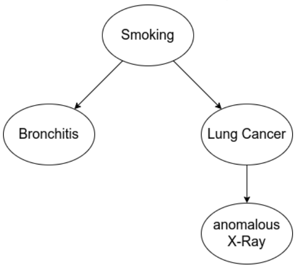
Marks for this submission: 1/1.

Question 6
Complete
Not graded

[2024-08-01] **QUESTION WAS MADE OPTIONAL:** This question is optional. **You can submit the quiz even if Moodle shows answers of this question as incorrect.** This was done after never ending technical issues with the capability of the underlying solver Maxima to correctly identify algebraic equivalence of different probability formula formulations.
If you still want to practice the probability formulations (which I would highly recommend), unfold the solutions here:

▶ show solutions to formulas ...

Let's consider a simplified version of the Bayesian network for lung cancer detection by removing the short of breath attribute. This leaves us with the binary variables Smoking (S), Bronchitis (B), Lung Cancer (L), and anomalous X-Ray (X).



You were given the following joint probabilities, obtained, e.g., from data or estimation:

Joint Probabilities

[%]

		s	s	¬s	¬s
		l	¬l	l	¬l
b	x	0.9	0.07	0.012	0.0796
b	¬x	0.6	3.43	0.008	3.9004
¬b	x	2.7	0.21	0.228	1.5124
¬b	¬x	1.8	10.29	0.152	74.1076

As python dict:
{"x": {"l": {"b": {"s": 0.9, "not_s": 0.012},
"not_b": {"s": 2.7, "not_s": 0.228}},
"not_l": {"b": {"s": 0.07, "not_s": 0.0796},
"not_b": {"s": 0.21, "not_s": 1.5124}}},
"not_x": {"l": {"b": {"s": 0.6, "not_s": 0.008},
"not_b": {"s": 1.8, "not_s": 0.152}},
"not_l": {"b": {"s": 3.43, "not_s": 3.9004},
"not_b": {"s": 10.29, "not_s": 74.1076}}}}

1. Conditional Probability Tables

Now first calculate the conditional probability tables:

P(s)
0.20

	P(b | S)	P(l | S)
s	0.25	0.3
¬s	0.05	0.005

	P(x | L)
l	0.60

	P(x | L)
¬l	0.02

2. Inference formulas

Recall that Bayesian networks simplify queries to the joint probability a lot, as they reduce atomic joint probabilities to a product of conditional probabilities: $P(X_1, \ldots, X_n) = \prod_i P(X_i \mid \text{Parents}(X_i))$. To turn any query into this easy-to-query formulation that allows to directly look up values of the conditional probability tables, the following steps are needed:

1. (Bayes: If a conditional probability is queried: Use Bayes' theorem $P(X \mid Y) = \frac{P(X,Y)}{P(Y)}$ to reformulate this only using joint probabilities/probability distributions.)
2. Marginalization: Turn into a sum of atomic probabilities/distributions by using marginalization to account for all non-constrained random variables.
3. Chain rule + independence: For atomic probability/distribution holds: $P(X_1, \ldots, X_n) = \prod_i P(x_i \mid \text{Parents}(X_i))$.

Let's use this now to formulate some inference formulas. Please use the notation indicated in the notation hints below.

Formulas (I): What is the formula for getting an atomic joint probability in terms of conditional probabilities from the tables (see notation hints below)?

$P(x, l, b, s) =$ Prob_X_given_L(x,l) * Prob_L_given_S(l,s) * Prob_B_giver

Your last answer was interpreted as follows:

$$P^{X|L}(x \mid l) \cdot P^{L|S}(l \mid s) \cdot P^{B|S}(b \mid s) \cdot P^{S}(s)$$

The variables found in your answer were: $[b, l, s, x]$

(Note: Here you cannot just write Prob(1,1,1,1) as answer ;-) Try to formulate as product of conditional probabilities!)

Formulas (II): Using the marginalization trick, we can now also determine a formula for the probability of having lung cancer and observing anomalous X-ray images (this is the unnormalized version of $P(L \mid X) = \frac{1}{P(X)} P(L, X)$). In particular, when anomalies are found, what is, in terms of probabilities from the tables, are the formulas for

- *(Example)* $P(b, l) = \sum_s \sum_x P(x \mid l)P(l \mid s)P(b \mid s)P(s) = \sum_s P(l \mid s)P(b \mid s)P(s) \underbrace{\sum_x P(x \mid l)}_{=1} = \sum_s P(l \mid s)P(b \mid s)P(s)$

  written as $P(l \mid s)P(b \mid s)P(s) + P(l \mid \neg s)P(b \mid \neg s)P(\neg s) =$ *Prob_L_given_S(l,s)*Prob_B_given_S(b,s)*Prob_S(s) + Prob_L_given_S(l,not_s)*Prob_B_given_S(b,not_s)*Prob_S(not_s)*
- $P(l, x) =$ Prob_X_given_L(x,l) * (Prob_L_given_S(l,s)*Prob_S(s) + P

  Your last answer was interpreted as follows:

  $$P^{X|L}(x \mid l) \cdot \left(P^{L|S}(l \mid s) \cdot P^{S}(s) + P^{L|S}(l \mid \neg s) \cdot P^{S}(\neg s)\right)$$

  The variables found in your answer were: $[l, \neg s, s, x]$
- $P(s, x) =$ Prob_S(s) * (Prob_X_given_L(x,l)*Prob_L_given_S(l,s) + P

  Your last answer was interpreted as follows:

  $$P^{S}(s) \cdot \left(P^{X|L}(x \mid l) \cdot P^{L|S}(l \mid s) + P^{X|L}(x \mid \neg l) \cdot P^{L|S}(\neg l \mid s)\right)$$

  The variables found in your answer were: $[l, \neg l, s, x]$

Notation hints: In order to allow you to enter any valid formulation of the same term as answer, the question checking is done by the algebraic checker Maxima. This also allows you to check through different ways to express your formulas. Please respect the following when formulating your answer:

- Simplify as much as possible:
  - Remove all parts that sum to 1.
  - None of the final formulas requires functions different from the ones above (i.e., absolute probabilities of single values
- Make sure to provide the formula for a probability value, not a probability distribution (i.e., do not use A in the formula, but its values a and not_a).
- The valid functions you may use for specifying (absolute and conditional) probabilities are:
  - Functions (Mind the order of arguments in the functions! These are normal Maxima functions, i.e., the order of parameters must not be changed!):
    - $P(x, l, b, s) =$ *Prob(x,l,b,s)*
    - $P(a) = P^{A}(a) =$ Prob_A(a): *Prob_X(x)*= $P(x)$, *Prob_L(l)*= $P(l)$, *Prob_B(b)*= $P(l)$, *Prob_S(s)* = $P(s)$
    - $P(a \mid b) = P^{A|B}(a \mid b) =$ Prob_A_given_B(a,b): *Prob_X_given_L(x, l)*= $P(x \mid l)$, *Prob_L_given_S(l, s)* = $P(l \mid s)$, *Prob_B_given_S(b, s)* = $P(b \mid s)$
  - Symbols: *x, not_x, l, not_l, b, not_b, s, not_s*

    Note that the symbols are actual placeholders that are not assigned a value. In your formula, you should only use those placeholders that are asked for in the question description. The others must be avoided, or instantiated by actual values that the system can calculate with (use 0 for false and 1 for true).

3. Inference Calculation

Given the numbers from the table your formulas, what is (giving 4 decimal digits)

$P(x) =$ 5.717 %

$P(l \mid x) =$ 67.2268 %

$P(s \mid x) =$ 67.9271 %

Remark: Why use this function notation? Note that, e.g., $P(L \mid X)$ and $P(S \mid X)$ are both functions on tuples of binary (in this case truth) variable values, but are very different density functions. Humans can directly derive which functions are meant from the name of the variables, but this must be made explicit when programming them, e.g., by adding them to the function names or as additional parameters.
Implementing a full parser for probability notation is out of scope of this exercise ;-)

Table entries:
\(P(s)\):
Your answer is **correct**.
Marks for this submission: 0/0.

\(P(b\mid s)\):
Your answer is **correct**.
Marks for this submission: 0/0.

\(P(b\mid \neg s)\):
Your answer is **correct**.
Marks for this submission: 0/0.

\(P(l\mid s)\):
Your answer is **correct**.
Marks for this submission: 0/0.

\(P(l\mid\neg s)\):
Your answer is **correct**.
Marks for this submission: 0/0.

\(P(x\mid l)\):
Your answer is **correct**.
Marks for this submission: 0/0.

\(P(x\mid\neg l)\):
Your answer is **correct**.
Marks for this submission: 0/0.

Inference formulas:
\(P(x,l,b,s)\):
Your answer is **correct**.
Marks for this submission: 0/0.

\(P(l,x)\):
Your answer is **correct**.
You can also write \(P(\neg s) =\)Prob_S(not_s) as \((1-P(s))=\) (1-Prob_S(s)).

Marks for this submission: 0/0.

\(P(s,x)\):
Your answer is **correct**.
You can also write \(P(\neg l\mid s) =\)Prob_L_given_S(not_l,s) as \((1-P(l\mid s))=\) (1-Prob_L_given_S(l,s)).

Marks for this submission: 0/0.

Inference calculation:
$P(x)$:
Your answer is **correct**.
Marks for this submission: 0/0.

$P(l\mid x)$:
Your answer is **correct**.
Marks for this submission: 0/0.

$P(s\mid x)$:
Your answer is **correct**.
Marks for this submission: 0/0.

Question 7

Correct

Mark 5 out of 5

Consider again the example given before about a lung cancer detector, with joint probability values known/assessable. Now assume that we have no prior knowledge about a sensible setup of links of the Bayesian network, but must (semi-)automatically derive that from the data.
Recall that we have seen an algorithm that allows to do so step-by-step for the simple case where the joint probabilities are known. The idea is to select for each newly attached node a minimal set of parents s.t. $P(X_n \mid X_1,\dots, X_{n-1}) = P(X_n \mid \text{Parents}(X_n))$.

We apply this algorithm to the order B, X, L, S of variables.
Let's go through the claims that need to be checked for building the structure step-by-step:

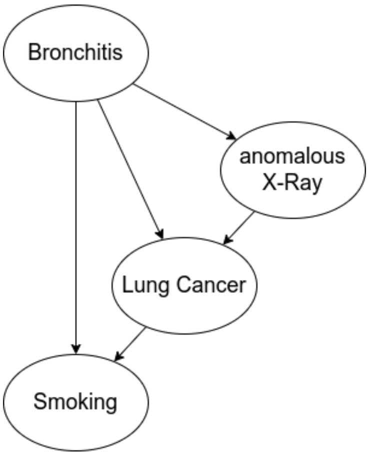
1. Claim 1: B is a parent of X, i.e., $P(X)\neq P(X\mid b)$ or $P(X)\neq P(X\mid \neg b)$.
2. Claim 2a: B is a parent of L, i.e., $P(L\mid X)\neq P(L\mid b,X)$ or $P(L\mid X)\neq P(L\mid \neg b,X)$.
3. Claim 2b: X is a parent of L, i.e., $P(L\mid B)\neq P(L\mid B,x)$ or $P(L\mid B)\neq P(L\mid B,\neg x)$.
4. Claim 3a: B is a parent of S, i.e., $P(S\mid X,L)\neq P(S\mid X,L,b)$ or $P(S\mid X,L)\neq P(S\mid X,L,\neg b)$.
5. Claim 3b: X is a parent of S, i.e., $P(S\mid L,B)\neq P(S\mid x,L,B)$ or $P(S\mid L,B)\neq P(S\mid \neg x,L,B)$.
6. Claim 3c: L is a parent of S, i.e., $P(S\mid X,B)\neq P(S\mid X,l,B)$ or $P(S\mid X,B)\neq P(S\mid X, \neg l,B)$.

Which of the following is true about the resulting Bayesian net?

Select one or more:

- [ ] a. We obtain a naive Bayesian net.
- [ ] b. The resulting net is a polytree.
- [x] c. B is parent to all its successors X, L, S.
- [ ] d. X and L are siblings, i.e., conditionally independent given B.
- [x] e. S only has parents B and L, i.e., it is conditionally independent of X given B and L.

Your answer is correct.
The resulting tree is unfortunately far from being a naive Bayesian net or a polytree Bayesian net:



Correct
Marks for this submission: 5/5.

◄ 03. Quiz - Hypothesis Testing

Jump to...

05. Quiz - Bayesian Learning ►