COMP-424: Artificial intelligence

Homework 3

Due on *myCourses* on March 25, 2019 at 9:00pm.
Head TA (for grading inquiries): Alika Utepova <alika.utepova@mail.mcgill.ca>

General instructions.

- This is an <u>individual</u> assignment. You can discuss solutions with your classmates, but should only exchange information orally, or else if in writing through the discussion board on *myCourses*. All other forms of written exchange are prohibited.
- Unless otherwise mentioned, the only sources you should need to answer these questions are your course notes, the textbook, and the links provided. Any other source used should be acknowledged with proper referencing style in your submitted solution.
- Submit a <u>single</u> pdf document containing all your pages of your written solution on your McGill's *myCourses* account. You can scan-in hand-written pages. If necessary, learn how to combine many pdf files into one.
- You may solve the questions by hand or by writing a program, but if you write a program, you must not rely on existing implementations, and must do it from scratch (and must submit your code along with the pdf).

Question 1: Designing a Bayesian Network

Toby the cat is not having a good day. His sister Lucy ate all the food provided by their owner, Claire, so he has to find a way to feed himself. He can try to catch a squirrel outside, and if he succeeds, he is going to eat it. If Toby is tired, he is less likely to catch the squirrel. Toby can also try to steal Claire's sandwich, which takes less effort, so it does not depend on whether he is tired. However even if he succeeds, he might not get to eat it (for example Claire may be quick enough to snatch it back). Finally, if Toby manages to eat at least something, he might feel happy, despite all the events. Though if Toby is tired, he is less likely to feel happy in general.

Consider the Boolean variables: H (happy), E (eats at least one item), C (catches squirrel), S (steals sandwich) and T (tired).

a. Draw a Bayesian network for this domain. Only include the Boolean variables listed above, so your network should have 5 nodes.
b. Is your network a polytree? Why or why not? (A polytree is a graph that has no directed or undirected cycles.)
c. Suppose the probability that Toby catches the squirrel is x when he is tired, and y when he is not tired. Give the conditional probability table associated with C.
d. Suppose that if Toby catches the squirrel, he will eat it with probability 1, and if he successfully steals a sandwich, he will eat it with probability 0.4. If he fails at both hunting and stealing, then he will not eat anything. Give the conditional probability table associated with E.
e. Suppose Toby is happy. Write down the expression for the probability that Toby is tired, in terms of the various conditional probabilities in the network.

Question 2: Inference in Bayesian Networks

Consider the following Bayesian Network

R: Rush Hour
B: Bad Weather
A: Accident
T: Traffic Jam
S: Sirens
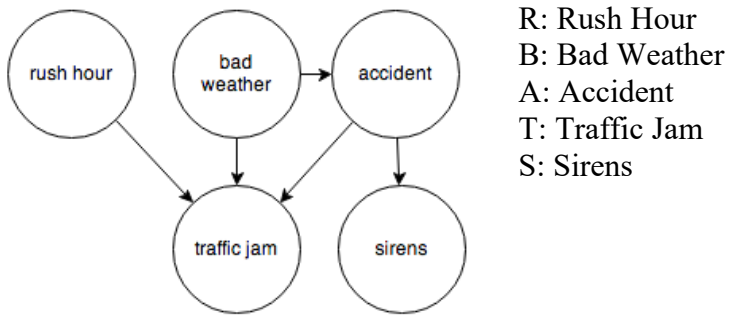
We will denote random variables with capital letters (e.g., R), and the binary outcomes with lowercase letters (e.g., r, and ¬r).

The network has the following parameters:

P(b)=0.3
P(r)=0.15

$P(t|r, b,a) = 0.98$
$P(t|r, \neg b,a) = 0.9$
$P(t|r,b,\neg a) = 0.88$
$P(t|r,\neg b,\neg a) = 0.85$
$P(t|\neg r,b,a) = 0.5$
$P(t|\neg r,b,\neg a) = 0.4$
$P(t|\neg r,\neg b,a) = 0.6$
$P(t|\neg r,\neg b,\neg a) = 0.05$

$P(s|a) = 0.9$
$P(s|\neg a) = 0.2$

$P(a|b) = 0.6$
$P(a|\neg b) = 0.3$

Compute the following terms using basic axioms of probability and the conditional independence properties encoded in the above graph. You can use Bayes Ball properties to simplify the computation, if applicable.

a. P(a, r)
b. P(b, ¬a)
c. P(b|s)

Question 3: Variable Elimination

For the graph above, compute the MAP result of querying P(T|b) using variable elimination with the following order: S, A, R, T.

Clearly explain each step. For each of the intermediate factors created, explain what probabilistic function it represents.

Question 4: Learning with Bayesian Networks

Consider the following Bayesian network. Assume that the variables are distributed according to Bernoulli distributions.
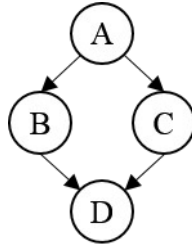
a. We are given the following dataset with 146 samples, from which we will estimate the parameters of the model.

A	B	C	D	# Instances
0	0	0	0	2
0	0	0	1	4
0	0	1	0	18
0	0	1	1	3
0	1	0	0	14
0	1	0	1	2
0	1	1	0	32
0	1	1	1	10
1	0	0	0	1
1	0	0	1	0
1	0	1	0	3
1	0	1	1	24
1	1	0	0	0
1	1	0	1	9
1	1	1	0	14
1	1	1	1	10

  i.   Enumerate the parameters that must be learned. Specify the parameter name and the probability that it represents (i.e., for each parameter, write something in the form, $\theta_X = Pr(X)$.

  ii. Give the maximum likelihood estimate for each parameter.

  iii. Give the maximum a posterior estimate for each parameter after applying Laplace smoothing.

b. Assume that in addition to the data in the table above, you are given the following incomplete data instances:

	A	B	C	D
S1	1	?	1	0
S2	1	1	0	?

We will apply the (soft) EM algorithm on these instances. Initialize the model using your parameter estimates from part a subpart ii (i.e., use the MLE).

i.   Show the computation of the first E-step, providing the weights for each possible assignment of the incomplete data for each sample.

ii.   What are the parameters obtained for the first M-step? Weight each of the samples from the original dataset and the two new samples equally (i.e., you now have 148 samples).

iii. Show the computation of the second E-step.