

COMP-424: Artificial intelligence

Homework 3

Due on *myCourses* on March 16, 2019 at 9:00pm.

Head TA (for grading inquiries): Maxime Wabartha <maxime.wabartha@mail.mcgill.ca>

General instructions.

- This is an individual 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 single 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

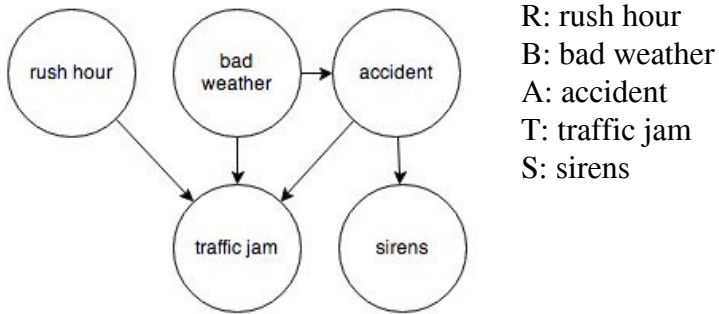
Marv the cat is having a bad day. His brother Harry ate all the food set out by their owner, Shannon, so Marv has to find a way to feed himself. He can try to catch a fish in the lake outside, and if he succeeds, he'll eat it. If it's a hot day, Marv will be sluggish, so he's less likely to catch a fish. Marv can also try to steal Shannon's sandwich, which does not depend on the outside temperature. However, even if he succeeds in stealing the sandwich, he might not get to eat it (for example, Shannon may notice and snatch it back). Finally, if Marv manages to eat at least something, he might feel content, in spite of everything. However, if it's hot out, he is less likely to feel content in general.

Consider the Boolean variables: H (it's a hot day), C (Marv is content), E (Marv eats at least one item), F (Marv catches a fish), S (Marv steals the sandwich).

- Draw a Bayesian network for this domain. Only include the Boolean variables listed above, so your network should have 5 nodes.
- Is your network a polytree? Why or why not? (NB: a polytree is a graph that has no directed or undirected cycles.)
- Suppose the probability that Marv catches the fish is x when it's hot, and y when it is not. Give the conditional probability table associated with F.
- Suppose that if Marv catches a fish, he will eat it with probability 1, and if he successfully steals the sandwich, he will eat it with probability 0.5. If he fails at both hunting and stealing, then he will not eat anything. Give the conditional probability table associated with E.
- Suppose Marv is content. Write down the expression for the probability that it is a hot day, in terms of the various conditional probabilities in the network.

Question 2: Inference in Bayesian Networks

Consider the following Bayesian Network



We denote random variables with capital letters (e.g., R for “rush hour”), and the binary outcomes with lowercase letters (e.g., r and $\neg r$ for “it is rush hour” and “it is not rush hour,” respectively).

The network has the following parameters:

$$P(b)=0.4$$

$$P(r)=0.2$$

$$P(tlr, b, a) = 0.98$$

$$P(tlr, \neg b, a) = 0.9$$

$$P(tlr, b, \neg a) = 0.88$$

$$P(tlr, \neg b, \neg a) = 0.85$$

$$P(tl \neg r, b, a) = 0.5$$

$$P(tl \neg r, b, \neg a) = 0.4$$

$$P(tl \neg r, \neg b, a) = 0.6$$

$$P(tl \neg r, \neg b, \neg a) = 0.05$$

$$P(s|a) = 0.95$$

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

$$P(a|b) = 0.7$$

$$P(a|\neg b) = 0.2$$

Compute the following terms using basic axioms of probability and the conditional independence properties encoded in the above graph.

a. $P(a, \neg r)$

b. $P(b, a)$

For the query $P(b|a)$:

c. Use Bayes Ball to determine the set of nodes that can be pruned from the graph.

d. Compute $P(b|a)$ using the simplifications determined in Part c.

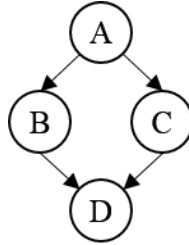
Question 3: Variable Elimination

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 144 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	31
0	1	1	1	10
1	0	0	0	1
1	0	0	1	0
1	0	1	0	3
1	0	1	1	23
1	1	0	0	0
1	1	0	1	9
1	1	1	0	14
1	1	1	1	10

- 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, $\theta_X = Pr(X)$).
- Give the maximum likelihood estimate for each parameter.
- Give the maximum *a posteriori* (MAP) 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	?
S2	1	1	0	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 146 samples).
- iii. Show the computation of the second E-step.