CS471- Homework 5

Due date: March 25nd at 11:59pm EST

Introduction

The objective of this assignment is to get familiar with Probability and Bayesian Networks.

Submission

Your homework must be typed and must contain your name and Purdue ID. To submit your assignment, log into `data.cs.purdue.edu` (physically go to the lab or use ssh remotely) and follow these steps:

1. To ssh use the command: `ssh username@data.cs.purdue.edu`
2. Make a directory named `username-hw5` (all letters in lower case)
3. Copy your PDF and code inside it. To do it remotely use the comand from your computer:
   `scp ./path/to/your-file.pdf username@data.cs.purdue.edu:./remote/path/from-home-dir/`
4. Go to the directory containing `username-hw5` (e.g., if the files are in /homes/aporco/aporco-hw5, go to /homes/aporco), and execute the following command:

   `turnin -c cs471 -p hw5 username-hw5`

   (e.g. Aldo would use: `turnin -c cs471 -p hw5 aporco-hw5` to submit his work)
5. To overwrite an old submission, simply execute this command again.
6. To verify the contents of your submission, execute the following command:

   `turnin -v -c cs471 -p hw5`

Required files

You will need to submit 1 file:

- The PDF containing your typed answers.

Problem Set

Problem 1: Give a joint distribution for Boolean random variables A, B, and C for each scenario. Give a brief intuitive interpretation of the variables. The notation $i(x, y)$ means that x and y are independent.
a. $i(A, B)$, $i(A, C)$, and $i(B, C)$.
b. $i(A, B)$ and $i(A, C)$, but not $i(B, C)$.
c. $i(A, B)$ and $i(A, C)$, but not $i(A, B \wedge C)$.

Problem 2: Given the full joint distribution shown in Figure 13.3, calculate the following:
a. $P(toothache)$
b. $\mathbf{P}(Cavity)$
c. $\mathbf{P}(Toothache \mid cavity)$
d. $\mathbf{P}(Cavity \mid toothache \vee catch)$

Problem 3: Suppose you are a witness to a nighttime hit-and-run accident involving a taxi in Athens. All taxis in Athens are blue or green. You swear, under oath, that the taxi was blue. Extensive testing shows that, under the dim lighting conditions, discrimination between blue and green is 70% reliable.
a. Is it possible to calculate the most likely color for the taxi? (Hint: distinguish carefully between the proposition that the taxi is blue and the proposition that it appears blue.)
b. What if you know that 8 out of 10 Athenian taxis are green?

Problem 4: We have a bag of three biased coins a, b, and c with probabilities of coming up heads of 40%, 40%, and 80%, respectively. One coin is drawn randomly from the bag (with equal likelihood of drawing each of the three coins), and then the coin is flipped three times to generate the outcomes X_1, X_2, and X_3.
a. Draw the Bayesian network corresponding to this setup and define the necessary CPTs.
b. Calculate which coin was most likely to have been drawn from the bag if the observed flips come out heads twice and tails once.

Problem 5: In your local nuclear power station, there is an alarm that senses when a temperature gauge exceeds a given threshold. The gauge measures the temperature of the core. Consider the Boolean variables A (alarm sounds), F_A (alarm is faulty), and F_G (gauge is faulty) and the multivalued nodes G (gauge reading) and T (actual core temperature).
a. Draw a Bayesian network for this domain, given that the gauge is more likely to fail when the core temperature gets too high.
b. Is your network a polytree? Why or why not?
c. Suppose there are just two possible actual and measured temperatures, normal and high; the probability that the gauge gives the correct temperature is x when it is working, but y when it is faulty. Give the conditional probability table associated with G.
d. Suppose the alarm works correctly unless it is faulty, in which case it never sounds. Give the conditional probability table associated with A.
e. Suppose the alarm and gauge are working and the alarm sounds. Calculate an expression for the probability that the temperature of the core is too high, in terms of the various conditional probabilities in the network.