CS 440: Introduction to AI

Homework 4

Due: November 9 11:59PM, 2010

Your answers must be concise and clear. Explain sufficiently that we can easily determine what you understand. We will give more points for a brief interesting discussion with no answer than for a bluffing answer.

Please email your solution to the TA at cs440ta@cs.illinois.edu.

1 Probability

1. Consider a joint probability distribution $P(X, Y, Z)$ over variables X, Y, and Z. Assume that the probability of each outcome is non-zero. Is the following equation always true? Prove its correctness or incorrectness.

$$P(X \mid Y, Z) = \frac{P(Y|X,Z)P(X|Z)}{P(Y|Z)}$$

2. After your yearly checkup, the doctor has bad news and good news. The bad news is that you tested positive for a serious disease and that the test is 95% accurate (i.e., the probability of testing positive when you have the disease is 0.95 and the probability of testing negative when you don't have the disease is 0.95). The good news is that this is a rare disease, striking only 1 in 100,000 individuals. Why is it good news that the disease is rare? What are the chances that you actually have the disease?

2 Bayesian Networks

1. Our research lab is having problems with his autonomous car and wants to model the system using a Bayesian network. The objective is to avoid obstacles by reducing the speed of the car and turning to drive around the obstacle. The variable **ObstaclePresent (OP)** denotes the presence of an obstacle in the route. If an obstacle is present it is very likely that the obstacle detection system will detect this obstacle. **ObstacleDetected (OD)** determines whether and obstacle was detected or not. Besides obstacles, the autonomous driver has to pay attention in cars coming from behind. The variable **CarBehind (CB)** indicates has value true if there is a car behind our autonomous car. The variable **CarDetected (CD)** indicates that a car was detected behind our car, and it is likely to be true if there is actually a car behind.

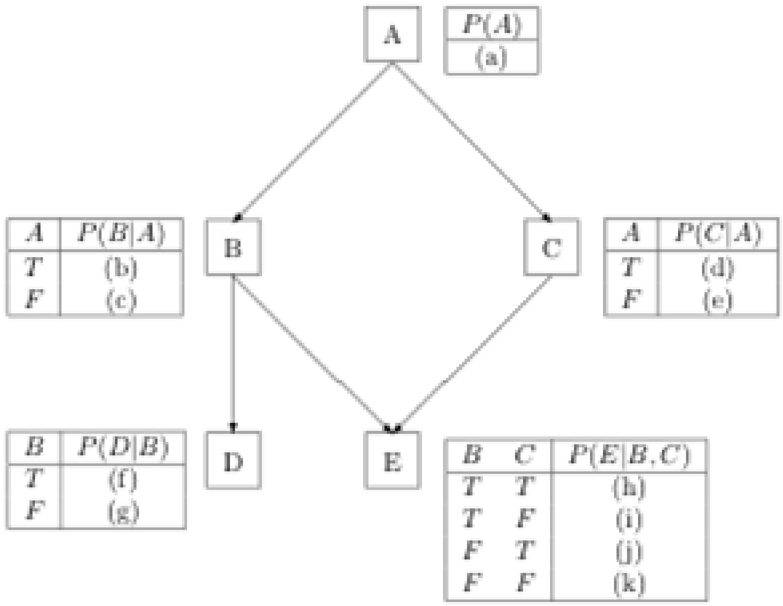
   If we detect an obstacle the driver is likely to decide to turn around, as well as it is likely to decide to reduce the speed. However, the driver is also likely to decide not to reduce the speed if it detects a car behind. **DecideToTurn (DTT)** indicates that the driver decided to turn and **DecideToReduce (DTR)** indicates that the driver decided to reduce the speed.

   Currently, one of the big problems of our autonomous driver is the communication between the decision components and the actuators. Even if the driver decides to reduce the speed there is chance that the car will not reduce the speed and even if the driver decides to turn there is a chance that the car will not turn. There is also a chance that the car will reduce speed or turn without a decision from the driver. **ReducedSpeed (RS)** indicates that the car actually reduced its speed while **Turned (T)** indicates whether the car actually turned or not. It is also important to remember that the car is likely to fail if it attempts to turn while driving in high speeds.

   (a) With the eight boolean variables, build a Bayesian Network so that the minmal number of parameters are needed to represent the conditional probabilities.

   (b) How many parameters are needed to represent the conditional probabilities in (a)?

2. Given the following Bayesian network of Boolean variables:



and the joint probability table

		A				$\neg A$			
		B		$\neg B$		B		$\neg B$	
		C	$\neg C$	C	$\neg C$	C	$\neg C$	C	$\neg C$
D	E	0.0024	0.0504	0.0252	0.0196	0.01152	0.06912	0.03024	0.00672
	$\neg E$	0.0216	0.0056	0.0168	0.0784	0.10368	0.00768	0.02016	0.02688
$\neg D$	E	0.0036	0.0756	0.0108	0.0084	0.01728	0.10368	0.01296	0.00288
	$\neg E$	0.0324	0.0084	0.0072	0.0336	0.15552	0.01152	0.00864	0.01152

i. Compute the probabilities (a) to (k) showing your work.

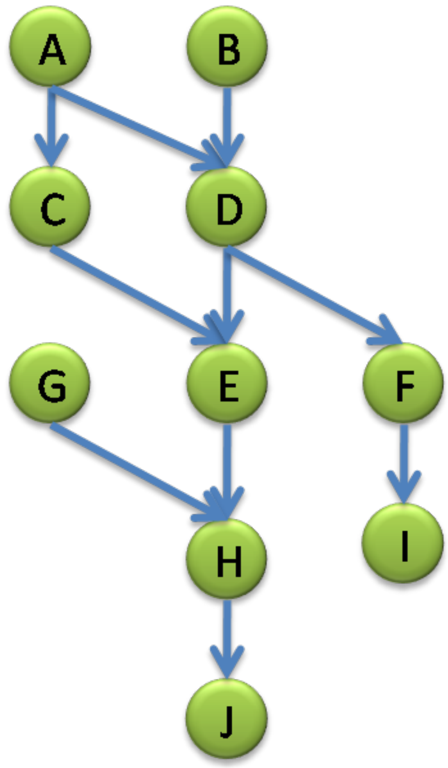
ii. Calculate $P(D = T \mid C = T)$ (or $P(d \mid c)$) in terms of the probabilities (a) to (k). (Use the letters a to k and not the number that they represent).

iii. Use the results of (i) and (ii) to obtain a value for the probability.

iv. Use the values from the joint probability table to check that (ii) and (iii) are correct.

v. As you did in (ii), calculate $P(A = F, B = T \mid D = T)$ (or $P(\neg a, b \mid d)$) in terms of the probabilities (a) to (k). (Use the letters a to k and not the number that they represent).

3. Given the Bayesian Network



(a) Which nodes constitute the Markov blanket of node E? To which nodes E is conditionally independent given E's Markov blanket?

(b) For each of the following conditional independences, decide whether True or False. ($I(A; B \mid C)$ is True if A and B are conditionally independent given C. Otherwise, it is False.)

i. $I(E; J \mid D)$

ii. $I(E; I \mid D)$

iii. $I(E; A \mid D)$

iv. $I(G; I \mid E)$

v. $I(A; F \mid B)$