Artificial Intelligence: Representation and Problem Solving
15-381, Fall 2009

Homework 4 SOLUTIONS

DUE: Wednesday, October 28th, 2009, 3:00 P.M.

REMINDER: Due at the beginning of the class on the day it is due.

[**Maximum points: 100**]

Handin instructions

- Bring the typeset or handwritten answers to the rest of the questions to class. Answers *MUST* be legible for credit. Make sure that both teammates' names are on the handin.
- Late homework is due by 3:00 P.M. on the day they are due. Please bring late homeworks to GHC 7027 (slide it under the door if Heather Carney is not there).

Guidelines

- You can work in a group of two people. This group does not need to be the same group as for the other homeworks. You only need to turn in one writeup per group, but you need to include the andrew ID's of the group members.
- If you have any questions about this assignment, contact the instructors at 15381-instructors@cs.cmu.edu

Problem 1: Independence (15 points)

Answer each part of each question with 1-2 sentences. These can be a bit tricky, so be careful! It is a good idea to check your answers with numerical examples.
For some of these questions, the solution is "not necessarily" rather than just "no." For example, in part 1.a, one could set the probabilities for A, B and C such that P(A,B)=P(A)P(B). However the structure of the Bayes net does not guarantee their independence.

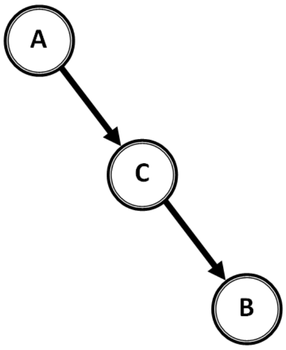
1. 

Figure 1.1

   (a) Is A independent of B?
   Not necessarily
   (b) is A conditionally independent of B given C?
   Yes

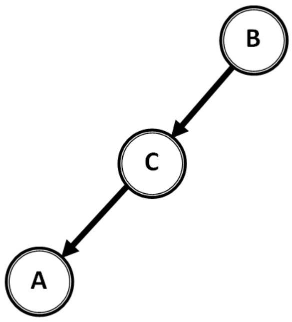
2. 

Figure 1.2

   (a) Is A independent of B?
   Not necessarily
   (b) is A conditionally independent of B given C?
   Yes

3.

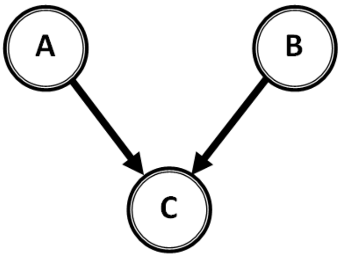


Figure 1.3

(a) Is A independent of B?
Yes

(b) is A conditionally independent of B given C?
Not necessarily. Given C, knowing A can "explain away" B. For example, if your alarm is going off, knowing that there was an earthquake makes it less likely that there was a burglary.

4.

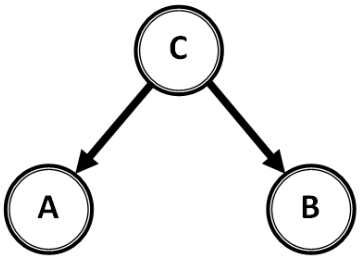


Figure 1.4

(a) Is A independent of B?
Not necessarily

(b) is A conditionally independent of B given C?
Yes

5.

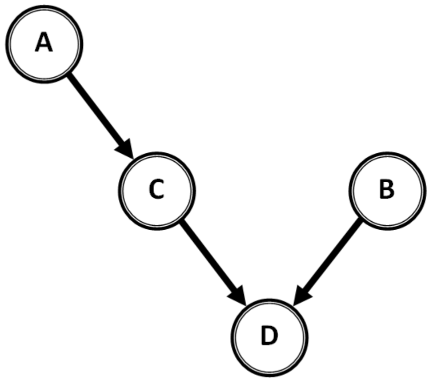


Figure 1.5

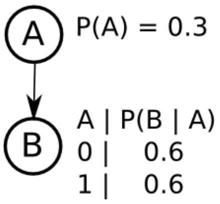
(a) Is A independent of B?
Yes

(b) Is A conditionally independent of B given C?
Yes

(c) Is A conditionally independent of B given D?
Not necessarily, for the same reason given in 3b

Problem 2: Practice with Bayes Nets (20 points)

Part 1: A strange Bayes net (10 points)



a. In one or two sentences, describe what is strange about this Bayes net.
   A and B are independent, even though the Bayes net implies that they are conditionally dependent.

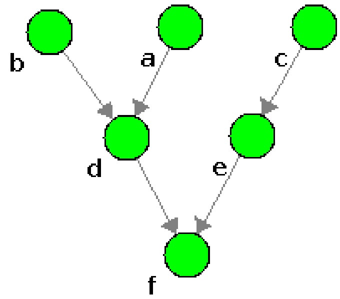
b. In a short paragraph, state whether or not this Bayes net is a valid representation and explain your answer.
   This is a valid Bayes net, though it is more complicated than it needs to be. The CPT for B really only needs 1 row. Credit was given if you claimed this is not a valid Bayes net because the edge demonstrates a dependence.

c. Redraw this Bayes net in order to fix the issue you described in part **a**
   The same Bayes net, just without the arrow from A to B

Part 2: Pokerface (10 points)

Given the following clues, draw the simplest (fewest edges) Bayes net that relates A,B,C,D,E, and F.

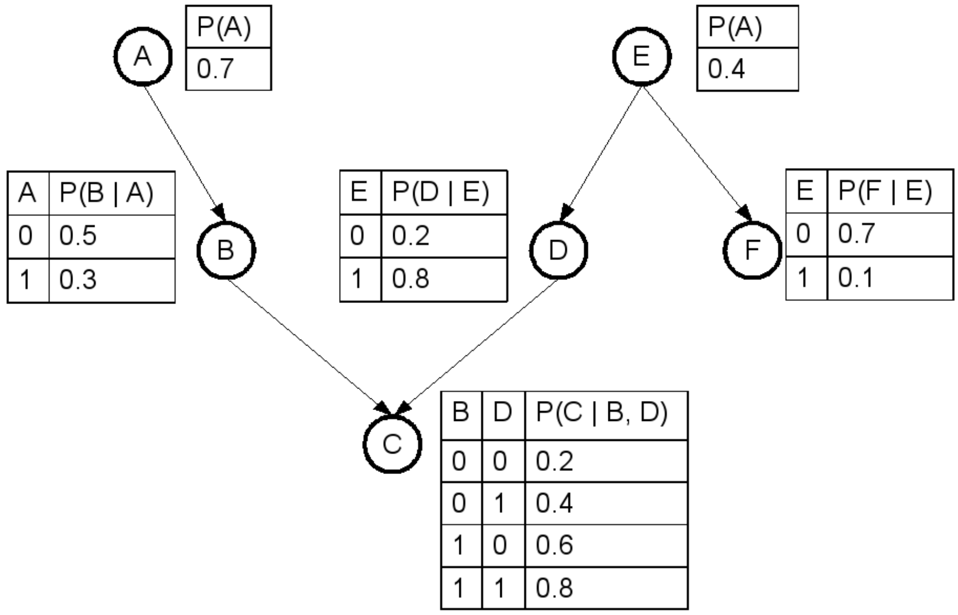
Let P(A) be the probability you are good at poker.
Let P(B) be the probability you are dealt good cards.
Let P(C) be the probability you are good at bluffing.
Let P(D) be the probability you have the best hand.
Let P(E) be the probability the other player folds.
Let P(F) be the probability you win.
A and C are independent.
D and E are independent.



This Bayes net satisfies the requirements

Problem 3: Calculations (25 points)

For each of the following parts, you must show all your work for credit. **Please give each answer with 3 significant digits**



Given this Bayes net, calculate each of the following:

a. P(B, D) (5 points)

$P(B,D)$

$P(B)P(D)$

$(P(B|A)P(A) + P(B|\neg A)P(\neg A))(P(D|E)P(E) + P(D|\neg E)P(\neg E))$

$(.21 + .15)(.32 + .12)$

$(.36)(.44)$

$.158$

b. P(B, D | C) (5 points)

P(B,D|C) = P(B,C,D) / P(C)

P(B,C,D) = P(C | B,D)P(B | D)P(D)
= P(C | B,D)P(B)P(D)
= (.8)(.36)(.44)
= .127
P(C)
= P(C| ¬B,¬D)P(¬B)p(¬D) + P(C| ¬B,D)P(¬B)p(D) + P(C|B,¬D)P(B)p(¬D) + P(C|B,D)P(B)p(D)
= .432

P(B,D|C)= P(B,C,D) / P(C) = .127 / .432 = .294

c. P(A | C) (5 points)

P(A|C) = P(A,C) / P(C)
P(C) = .432 (from part b)
P(A,C) = P(C | A)P(A)
P(C | A) = P(C | A, ¬B, ¬D)P(¬B|A)P(¬D)
+ P(C | A, ¬B, D)P(¬B|A)P(D)
+ P(C | A, B, ¬D)P(B|A)P(¬D)
+ P(C | A, B, ¬D)P(B|A)P(D)
=(.2)(.7)(.56)+(.4)(.7)(.44)+(.6)(.3)(.56)+(.8)(.3)(.44)
=.408
P(A,C) = P(C | A)P(A) = (.408)(.7) = .286
P(A|C) = P(A,C) / P(C) = (.286) / (.432) = .662 (Or .661 depending on rounding)

d. P(C, F | D) (5 points)

P(C, F | D) = P(C | D, F)P(F | D) = P(C | D)P(F | D)

P(C | D) = P(C| ¬B,D)P(¬B)+P(C|B,D)P(B)
= (.4)(.64) + (.8)(.36) = .544

P(E | D) = P(D | E)P(E) / P(D) = (.8)(.4)/.44 = .727
P(F | D) = P(F | D, E)P(E | D) + P(F | D, ¬E)P(¬E | D)
= (.1)(.727) + (.7)(.273) = .264

P(C, F | D) = .544*.264 = .144

e. P(F | ¬A, C) (5 points)

Note: .42 is WRONG. You can't get rid of the ¬A.

$P(F \mid \neg A, C) = P(F \mid \neg A, C, E)P(E \mid \neg A,C) + P(F \mid \neg A, C, \neg E)P(\neg E \mid \neg A,C)$

$P(E \mid \neg A,C) = P(C \mid \neg A, E)P(E \mid \neg A) / P(C \mid \neg A) = P(C \mid \neg A, E)P(E) / P(C \mid \neg A)$

$$\begin{aligned} P(C \mid \neg A) = {} & P(C \mid \neg B, \neg D)P(\neg B \mid \neg A)P(\neg D) \\ & + P(C \mid \neg B, D)P(\neg B \mid \neg A)P(D) \\ & + P(C \mid B, \neg D)P(B \mid \neg A)P(\neg D) \\ & + P(C \mid B, \neg D)P(B \mid \neg A)P(D) \\ & =(.2)(.5)(.56)+(.4)(.5)(.44)+(.6)(.5)(.56)+(.8)(.5)(.44) \\ & =.488 \end{aligned}$$

$$\begin{aligned} P(C \mid \neg A, E) = {} & P(C \mid \neg B, \neg D)P(\neg B \mid \neg A)P(\neg D \mid E) \\ & + P(C \mid \neg B, D)P(\neg B \mid \neg A)P(D \mid E) \\ & + P(C \mid B, \neg D)P(B \mid \neg A)P(\neg D \mid E) \\ & + P(C \mid B, \neg D)P(B \mid \neg A)P(D \mid E) \\ & =(.2)(.5)(.2)+(.4)(.5)(.8)+(.6)(.5)(.2)+(.8)(.5)(.8) \\ & =.56 \end{aligned}$$

$P(E \mid \neg A,C) = (.56)(.4) / (.488) = .459$

$P(F \mid \neg A, C) = (.1)(.459) + (.7)(.541) = .425$

Problem 4: JavaBayes (25 points)

Download JavaBayes from www.cs.cmu.edu/~javabayes
Run JavaBayes by going into the Classes directory and typing java JavaBayes
Load an example by going to File→Open, and then choosing an example from the Examples directory.

For this problem load the car-starts problem from the Examples/CarStarts directory. **Please express answers with 4 significant digits**

Part 1: Warm up (6 points)

a. What is the probability EngineCranks cranks? (Hint: use the Query button)
.9329

b. What is the CPT (conditional probability table) for Lights? (Hint: use the Edit Function button)
P(Lights | BatteryPower=Good) = 0.9
P(Lights | BatteryPower=Poor) = 0.0

c. What is the probability BatteryPower is Good given that the Radio is Dead? (Hint: use the Observe button to observe that the radio is dead, and then use the Query button)
.5889

Part 2: Charge (8 points)

a. What is the probability the Alternator is OK?
.9995

b. What is the probability the Alternator is OK given that Charge is Low?
.9532

c. What is the probability the Alternator is OK given that there is a Leak?
.9995

d. What is the probability the Alternator is OK given that Charge is Low and there is a Leak?
.9997

Part 3: A bit trickier (11 points)

a. What is the probability that Charge is Good, given that EngineCranks is NoCrank?
.9361

b. What is the probability that Lights is NoLight, given that EngineCranks is NoCrank?
.8743

c. What is the probability that Charge is Good and that Lights is NoLight, given that EngineCranks is NoCrank?
P(Charge=Good,Lights=NoLight | EngineCranks=NoCrank)
=P(Charge=Good | Lights=NoLight, EngineCranks=NoCrank)
$*$P(Lights=NoLight | EngineCranks=NoCrank)
=(.9269)(.8743)
=.8104

Problem 5: Decision Trees (15 points)

Use the following (fictional) dataset for this problem. It is the records of 12 hypothetical patients, with attributes Sex, age Over 60, Diabetic, high Pulse rate, abnormal EKG; and classification HasArrhythmia.

Patient	Sex	Over60	Diabetic	Pulse	EKG	HasArrhythmia
1	M	+	+	-	-	-
2	M	-	-	+	+	+
3	M	-	+	+	-	-
4	M	+	-	-	+	+
5	M	+	+	+	-	+
6	M	-	+	+	-	+
7	F	-	-	+	-	-
8	F	+	+	+	+	+
9	F	-	+	-	+	+
10	F	+	-	-	-	-
11	F	+	+	-	-	-
12	F	+	-	+	+	+

a. Calculate the conditional entropy, $H(HasArrhythmia|Sex = Female)$.
P(HasArrhythmia | Sex=Female) = 0.5
H(HasArrhythmia | Sex=Female)
=-((0.5)log$_2$(0.5) + (0.5)log$_2$(0.5))
=1

b. Under the attribute selection measure $\frac{Gain^2(S,A)}{Cost(A)}$, what would be the first split in the tree? Assume that $Cost(Sex) = Cost(Over60) = 1$, $Cost(Diabetic) = 3$, $Cost(HighPulse) = 2$, $Cost(AbnormalEKG) = 5$.
The first split would be AbnormalEKG

c. Suppose that, on a different set of patients, we knew their exact ages. Ages of positive examples are: {40, 60, 62, 64, 70, 74, 75, 82} and negative examples are: {33, 35, 42, 45, 49, 52, 58, 59, 80}. Suppose that all other attributes in the data set are poor predictors, so we want to split the tree at $Age = k$ by dividing the continuously-valued data points into two groups, $Age \geq k$ and $Age < k$. What division might we choose, based on information gain?
k = 60