

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

Homework 4

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.

1.

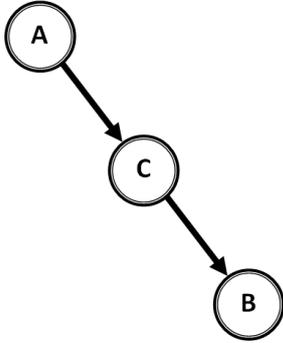


Figure 1.1

- (a) Is A independent of B?
- (b) is A conditionally independent of B given C?

2.

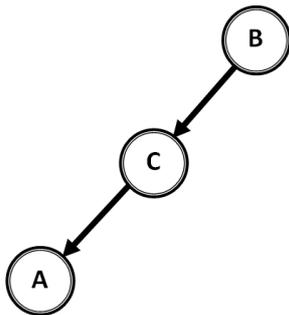


Figure 1.2

- (a) Is A independent of B?
- (b) is A conditionally independent of B given C?

3.

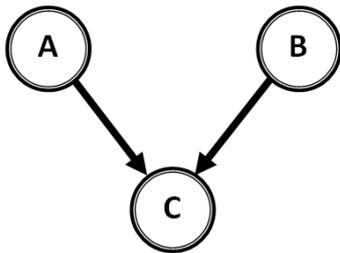


Figure 1.3

- (a) Is A independent of B?
- (b) is A conditionally independent of B given C?

4.

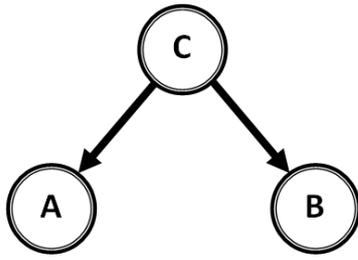


Figure 1.4

- (a) Is A independent of B?
- (b) is A conditionally independent of B given C?

5.

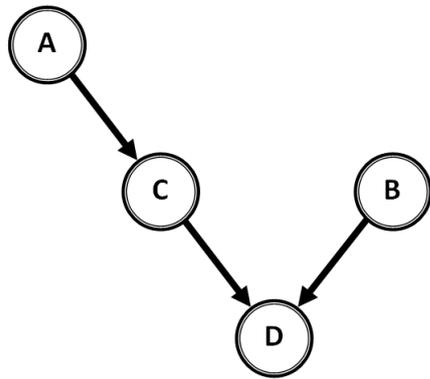
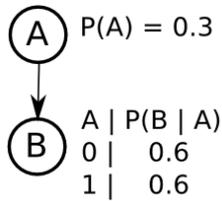


Figure 1.5

- (a) Is A independent of B?
- (b) Is A conditionally independent of B given C?
- (c) Is A conditionally independent of B given D?

Problem 2: Practice with Bayes Nets (20 points)

Part 1: A strange Bayes net (10 points)



- In one or two sentences, describe what is strange about this Bayes net.
- In a short paragraph, state whether or not this Bayes net is a valid representation and explain your answer.
- Redraw this Bayes net in order to fix the issue you described in part **a**

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.

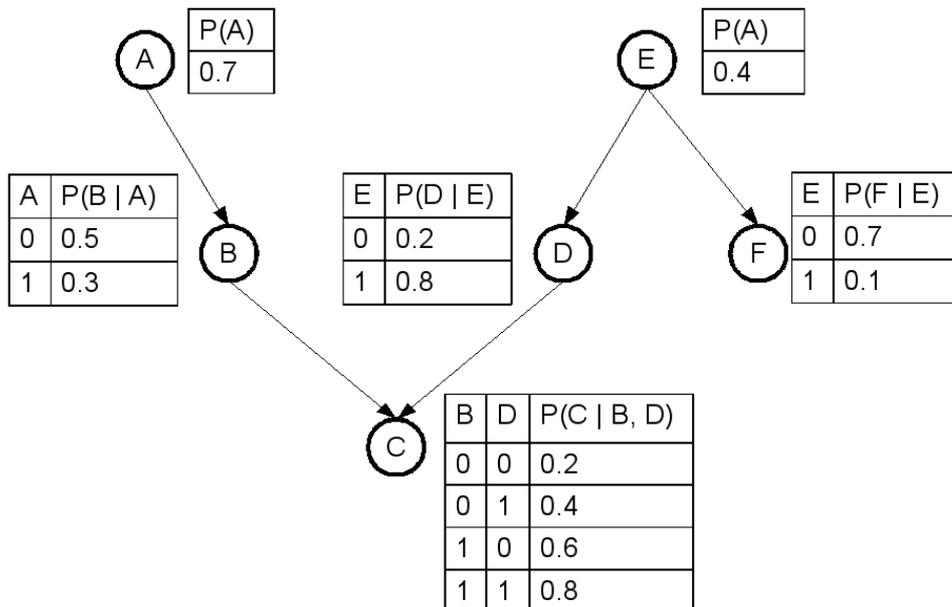
Let $P(F)$ be the probability you win.

A and C are independent.

D and E are independent.

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:

- $P(B, D)$ (5 points)
- $P(B, D | C)$ (5 points)
- $P(A | C)$ (5 points)
- $P(C, F | D)$ (5 points)
- $P(F | \neg A, C)$ (5 points)

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)

- What is the probability EngineCrank cranks? (Hint: use the Query button)
- What is the CPT (conditional probability table) for Lights? (Hint: use the Edit Function button)
- 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)

Part 2: Charge (8 points)

- What is the probability the Alternator is OK?
- What is the probability the Alternator is OK given that Charge is Low?
- What is the probability the Alternator is OK given that there is a Leak?
- What is the probability the Alternator is OK given that Charge is Low and there is a Leak?

Part 3: A bit trickier (11 points)

- What is the probability that Charge is Good, given that EngineCrank is NoCrank?
- What is the probability that Lights is NoLight, given that EngineCrank is NoCrank?
- What is the probability that Charge is Good and that Lights is NoLight, given that EngineCrank is NoCrank?

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	+	-	+	+	+

- Calculate the conditional entropy, $H(\text{HasArrhythmia}|\text{Sex} = \text{Female})$.
- Under the attribute selection measure $\frac{\text{Gain}^2(S,A)}{\text{Cost}(A)}$, what would be the first split in the tree? Assume that $\text{Cost}(\text{Sex}) = \text{Cost}(\text{Over60}) = 1$, $\text{Cost}(\text{Diabetic}) = 3$, $\text{Cost}(\text{HighPulse}) = 2$, $\text{Cost}(\text{AbnormalEKG}) = 5$.
- 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 $\text{Age} = k$ by dividing the continuously-valued data points into two groups, $\text{Age} \geq k$ and $\text{Age} < k$. What division might we choose, based on information gain?