CSCI 3202 – Introduction to AI Name_____
Instructor: Hoenigman
Final Exam – Fall 2015

You are allowed to use one 8.5x11 page of notes for this exam. The notes can be written front and back in whatever font size your prefer. Anything other than permitted materials, such as texts from friends, Internet searches, email, etc is not permitted and is considered cheating.

If you are caught cheating, it guarantees you an F in the class, a date with the Honor Code council, and my animosity for all time.

So, please sign: "On my honor as a University of Colorado student, I promise not to cheat on this test by giving or receiving help from my peers, or engaging in any other activities that are not allowed under the rules of the exam."

_____

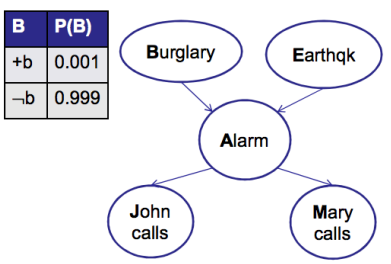
Now, on to the exam…

Problems

1. Consider the following data from the word HMM in Assignment 8, where the left column is the correct letter (hidden state) and the right column is the observation.

   Using this sample of 10 data points, calculate P(state = 'c'|'f').

a	a
c	c
c	f
o	o
u	u
n	n
t	f
_	_
v	f
f	i



B	P(B)
+b	0.001
¬b	0.999

E	P(E)
+e	0.002
¬e	0.998

B	E	A	P(A|B,E)
+b	+e	+a	0.95
+b	+e	¬a	0.05
+b	¬e	+a	0.94
+b	¬e	¬a	0.06
¬b	+e	+a	0.29
¬b	+e	¬a	0.71
¬b	¬e	+a	0.001
¬b	¬e	¬a	0.999

A	J	P(J|A)
+a	+j	0.9
+a	¬j	0.1
¬a	+j	0.05
¬a	¬j	0.95

A	M	P(M|A)
+a	+m	0.7
+a	¬m	0.3
¬a	+m	0.01
¬a	¬m	0.99

(Note: ¬ is false, and + is true.)

2. Using the Bayes network shown here, calculate

   a. P(¬b, +a)

   b. P(+a | +j)

   c. P(¬b | +m)

3. Circle T or F for each of the following statements about the Bayes net in Question 2?
    a. T/F P(J | A) = P(J)
    b. T/F P(J | A, M) = P(J | A)
    c. T/F P(J | A, B) = P(J | A)
    d. T/F P(A | E) = P(A | B, E)

4. Assume the following hypothetical situation: earlier today I walked down to the coffee cart in the engineering lobby and I asked 3 people how they were feeling and whether they were healthy or had a cold. Each person described how they felt over the last 5 days, including the current day, and their state on that day.

	Person 1	Feeling	Person 2	Feeling	Person 3	Feeling
Day 1	Healthy	Good	Healthy	Good	Healthy	Good
Day 2	Cold	Tired	Healthy	Good	Healthy	Good
Day 3	Cold	Tired	Healthy	Tired	Cold	Good
Day 4	Cold	Tired	Healthy	Tired	Cold	Tired
Day 5 (current)	Healthy	Tired	Healthy	Good	Cold	Tired

    a. Use this data to formulate an HMM for determining whether a person is Healthy or has a Cold from whether they feel Tired or Good. You can assume that initially, everyone is healthy.

b. Use your HMM and the Viterbi algorithm to answer the following question: If I walk up to a random person in the lobby and they say that over the last three days, including today, they have felt Good, Tired, and Tired, which of the following is true:
    a. It is more likely that they have a Cold.
    b. It is more likely that they're Healthy.
    c. It is equally likely that they have a Cold or they're Healthy.

(Show your calculations.)

5. A gambler goes to a casino with $50 and decides to play Roulette because he doesn't know much about gambling. In this Roulette, if he places $25 on red and the marble lands on red, then he wins $25. The same is true for black; if he places $25 on black and the marble lands on black, then he wins $25. However, if the marble lands on the opposite color than the one he selected, he loses his $25 bet.

   The gambler chooses to place $25 on red on each spin, always. If red comes up, he wins $25. If black comes up, he loses $25. He will exit the game when he has $75 or $0. The odds of winning are 50%.

   a. What are the four states that the gambler can be in?

   b. What are the transition probabilities for each of the states? Hint: the transition matrix is 4x4.

   c. The gambler starts with $50. What is the probability that he has $25 after one spin, two spins, and three spins?

6. Assume you are given the following data set with a *feature vector: <F1, F2>*, and a class label of – or + to indicate that the data point belongs in the – or + class. For this data set, if the initial weight vector is *<-0.25, -1.82>*, what will the weight vector be after one iteration through all three data points?

Data ID	Class Label	F1	F2
1	-	-1.35	0.89
2	+	-1.6	-0.93
3	+	-1.8	0.25

7. Consider the following data set for six bicycles comprised of three input attributes (*Type, Manufacturer, and Price*). The binary output is whether you buy the bike, Yes or No. Calculate the information gain for each variable and determine the most important variable in your decision to purchase a bicycle.

Example	Type	Manufacturer	Price	Buy (Y/N)
$\mathbf{X_1}$	Road	Specialized	Low	No
$\mathbf{X_2}$	Road	Specialized	High	No
$\mathbf{X_3}$	Mountain	Light Speed	Low	No
$\mathbf{X_4}$	Road	Light Speed	High	Yes
$\mathbf{X_5}$	Road	Light Speed	Low	Yes
$\mathbf{X_6}$	Mountain	Light Speed	High	Yes