CS 422 – Assignment 1

Due: 24th Sep 2022, 11:59 PM

Instructions:

1. **Please provide typed answers in a word document or in pdf. Please do not provide handwritten answers.**
2. **This is individual assignment, so no collaboration is allowed.**

Question 1. (5 pts)

A game is played with rolling three fair six-sided dice. The result of the game is decided based on the sum of the outcome of the three dices. The player wins $1 if the total is 2,7,12 or 15, he loses $1 if the total is 9 or 18. For the remaining cases he gets to play one more time. But this time he wins $2 if the total is 5,9,10 or 13 and loses $1 for the remaining cases. Will you play this game? What is the winning probability?

Question 2. (5 points)

There is a trip being planned for 450 people by using a transportation company. The transportation company has 11 large buses, each of which can hold 40 people and 8 small buses, each of which can hold 30 people. The rental cost for a large bus is 800 and for a small bus is 600. If there are only 14 drivers available , calculate how many buses of each type should be used for the trip with the least possible cost? Provide the details of the method employed and explain if there are any drawbacks to the approach?

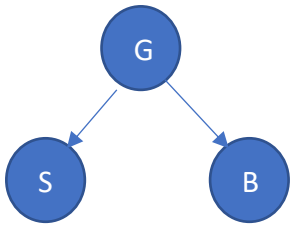
Question 3. (5 pts)

In this problem, you will be tasked with learning the local probability tables of a belief network. There are three variables Genre, G , rating of Sonia, S and rating of Bob, B.

$G \in$ {drama - d, comedy - c, thriller - t}

$S \in$ {1, 2, 3, 4, 5}

$B \in$ {1, 2, 3, 4, 5}



Please find the local probability tables (showing only non-zero values) of the above Bayesian Network given the following 10 data points:

1. d, 5, 4 (indicating drama movie, where rating of Sonia and Bob were 5 and 4 respectively)
2. c, 3, 5

3. t, 5, 3
4. d, 4, 4
5. t, 5, 4
6. d, 5, 5
7. c, 3, 4
8. c, 4, 5
9. t, 5, 2
10. t, 4, 3

Question 4. (10 pts)

Credit card companies lose millions of dollars due to frauds. You will help a company build Fraud detection system. Here is some of the information provided by the credit card company:

Travelling:2% of transactions are fraudulent
Not travelling: 0.2% are fraudulent,

8% transactions happen when cardholder travelling

If card holder not travelling, 10% of fraudulent transactions are foreign purchases, whereas 1% of legitimate transactions are foreign purchases.

If cardholder is travelling, 90% of transactions are foreign purchases regardless of legitimacy of transactions

Purchases over internet are more likely to be fraudulent and this is true for card holders without a computer

60% population owns a computer and for those cardholders
1% of legitimate transactions are done over the internet
Percentage increases to 2% for fraudulent transactions
For those who don't own computer
0.1% of legitimate transactions are over internet
1.1% are fraudulent transactions

Credit card company does not know whether the card holder owns a computer, but can guess based on whether recent transactions involve purchase of computer related accessories

10% of those who own computer purchase computer related accessory using their credit card

0.1% of those who don't own any computer purchase a computer related accessory

Based on the above information, please answer the following questions:

1. What is the Bayes Network corresponding to the above description? List down the nodes and links between the nodes. [2 points]

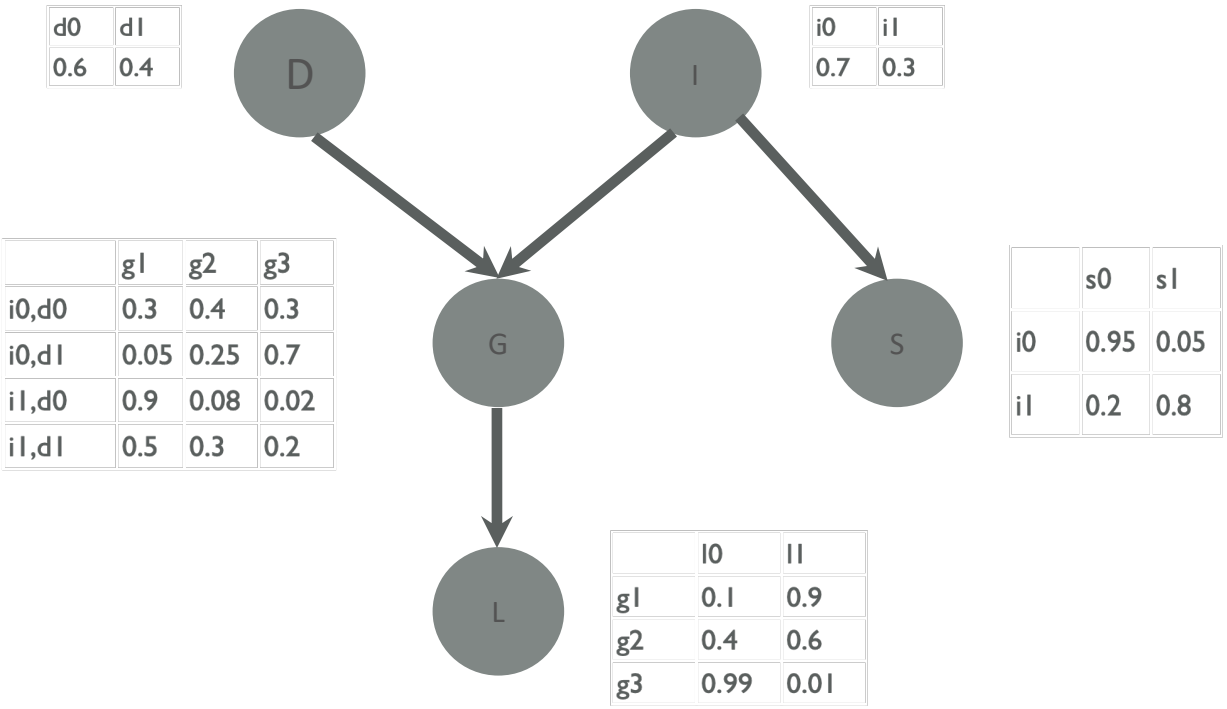
2. What is the prior probability (i.e., before we search for previous computer related purchases and before we verify whether it is a foreign and/or an Internet purchase) that the current transaction is a fraud? What is the probability that the current transaction is a fraud once we have verified that it is not a foreign transaction, but an Internet purchase and that the cardholder did not purchase computer related accessories in the past week? Use likelihood weighting approach with 100,000 samples. Please provide 10 samples generated. How different is the answer compared to if variable elimination is used?[4 points]

3. After computing those probabilities, the fraud detection system raises a flag and recommends that the cardholder be called to confirm the transaction. An agent calls at the domicile of the cardholder but she is not home. Her spouse confirms that she is currently out of town on a business trip. How does the probability of a fraud change based on this new piece of information? [4 points]

Question 5. (10 pts)

You are given the following Bayes network with the conditional probabilities.

Please show how to compute the most probable explanation for S = s0 and G = g0 using variable elimination (forward and backward pass)? Provide the final answer.



d0	d1
0.6	0.4

i0	i1
0.7	0.3

	g1	g2	g3
i0,d0	0.3	0.4	0.3
i0,d1	0.05	0.25	0.7
i1,d0	0.9	0.08	0.02
i1,d1	0.5	0.3	0.2

	s0	s1
i0	0.95	0.05
i1	0.2	0.8

	l0	l1
g1	0.1	0.9
g2	0.4	0.6
g3	0.99	0.01

Question 6. (10 pts)

Taxi drivers in Singapore can pick up customers from any location that is not on highways. However, they require guidance on where to pick up customers when they do not have

customers on board and there are no bookings. In this question, we address this guidance problem.

There are four locations: L1, L2, L3 and L4 from where taxi drivers can pick up and drop off customers. At any decision epoch, the chances of the taxi driver picking up a customer from different locations are provided in Table 1 below:

Location	Chance of finding customer
L1	0.2
L2	0.6
L3	0.4
L4	0.7

Table 1

Once the taxi driver picks up a customer, the customer determines the destination. Observed probabilities (from past data) of a customer starting from a source location and going to a destination location:

Source→ Destination	Probability
L1 → L2	0.4
L1 → L3	0.2
L1 → L4	0.4
L2 → L1	0.6
L2 → L3	0.4
L3 → L1	0.6
L3 → L4	0.4
L4 → L1	0.4
L4 → L2	0.6

Table 2

Just to avoid any confusion, first row of Table 2 below the heading row indicates that a customer picked up from L1 gets dropped off at L2 with 0.4 probability.

Travel fare between different locations are as follows:

Source→ Destination	Fare
L1 → L2	25$
L1 → L3	12$
L1 → L4	6$
L2 → L1	10$
L2 → L3	8$
L3 → L1	15$
L3 → L4	9$
L4 → L1	11$
L4 → L2	5$

Cost of travelling between locations is as follows:

Source→ Destination	Fare
L1 → L2	1$

L1 → L3	1.5$
L1 → L4	1.25$
L2 → L1	1$
L2 → L3	1.75$
L3 → L1	1.5$
L3 → L4	1.2$
L4 → L1	1.25$
L4 → L2	1.00$

Travel between any other locations costs 1$.

The taxi driver can either pickup from the current location or move to another location. Pickup corresponds to picking up a customer (if one is found) and dropping them of at their destination. Pickup action succeeds with probabilities specified in *Table 1* and if the taxi picks up a customer, destination location is determined by the probabilities in *Table 2*. When Pickup action fails, the taxi remains in its current location. Move to another location is always successful and taxi moves to the desired location with probability 1.

Both pickup and move actions take one-time step each. Please provide the following:

(a) **An MDP model that guides the taxi driver on "move and pickup customers"**. MDP is the tuple <S, A, P, R>, i.e., the states, actions, transition probability matrices (for different actions) and reward functions.

(b) **Solve the MDP model using policy iteration and discount factor of 0.95**. Provide:
   a. your code; and
   b. show two steps of value iteration in the solution.

(c) Solve the MDP model using linear programming formulation and discount factor of 0.95. Provide**:**
   a. Your LP formulation
   b. Your code