United International University
Department of Computer Science and Engineering
CSI 341: Artificial Intelligence, Final Exam, Spring 2018
Total Marks: 40, Time: 2:00 Hours

Answer all the questions. Marks are indicated in the right side of each question.

1. (a) Consider the following server-allocation problem at Fig. 1. There are 3 servers (S_i's), and 5 process requests (P_i's) for the servers. Each server must be used to run exactly 1 process, and all the servers must have different processes running. Design the problem instance as a formal **Constraint Satisfaction Problem (CSP)** with suitable **variables**, their exact **domains**, and associated **constraints**. The edges at the graph denote suitable process - server compatibility. [3]

Figure 1: Server-Process compatibility for q. 1a

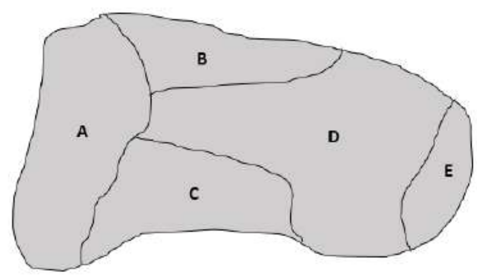


Figure 2: Map for q. 1b

(b) Suppose you have to color the map at Fig. 2 with 3 colors ({*Red, Green, Blue*}) so that no two adjacent regions have the same color. Constructing this problem as a **CSP**, show the steps followed by the **backtracking search** with the **minimum remaining values** heuristic to solve it. [3]

(c) What should be a suitable objective function to use with the **greedy hill-climbing search** for solving a map-coloring problem? Will it be a maximization or a minimization problem then? [2]

2. (a) You are provided with the following data collected from a survey done on a single day at a smart-phone sales center. 14 young men, 8 young women, 7 old men, and 1 old woman bought Nokia phone. 8 young men, 12 young women, 5 old men, and 3 old women bought Samsung phone. 14 young men, 9 young women, 11 old men, and 8 old women bought iphone. Answer the following. [3 + 1.5 + 1.5]

   i. Construct a **Full Joint Probability Distribution** table, defining three random variables: *A(Age)*, *G(Gender)*, and *S(Smartphone bought)*.

   ii. Calculate the (conditional) probability of an individual uniformly selected from this population not to buy iphone if he / she is of old age, from your table.

   iii. What is the expected number of *Samsung* users among 30 randomly selected users from this population?

(b) Company A supplies 40% of the computers sold and is late 5% of the time. Company B supplies 30% of the computers sold and is late 3% of the time. Company C supplies another 30% and is late 2.5% of the time. A computer arrives late; what is the probability that it came from Company A or Company B? [3]

3. (a) Consider the **Bayesian network** at Fig. 3 with all Boolean random variables. Calculate the number of probability distributions, and the minimum number of probabilities required for this network. [2]

(b) For the **Bayesian network** at Fig. 4 with all Boolean random variables, determine the following.

i. Probability of grass being wet, not raining, and sprinkler being on. [1]

ii. Probability of sprinkler being turned off. [2]

iii. Probability that it had rained, if you have observed that the grass is wet. [3]

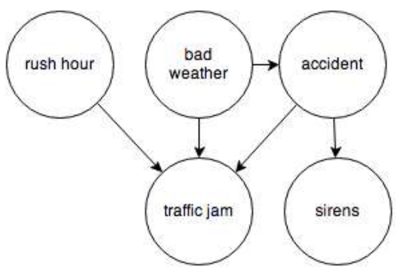
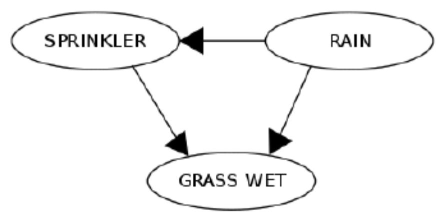


Figure 3: Bayesian network for q. 3a

	SPRINKLER	
RAIN	T	F
F	0.4	0.6
T	0.01	0.99



RAIN	
T	F
0.2	0.8

		GRASS WET	
SPRINKLER	RAIN	T	F
F	F	0.0	1.0
F	T	0.8	0.2
T	F	0.9	0.1
T	T	0.99	0.01

Figure 4: Bayesian network for q. 3b

4. Suppose that you have collected the following data from the weather conditions of a golf course, and its effect on playing golf and the net profit earned by the nearby restaurants. The features are: *Outlook, Temperature, Humidity,* and *Windy*; the outputs are *Play Golf* and *Profit*.

Outlook	Temperature	Humidity	Windy	Play Golf	Profit
Rainy	Hot	High	False	No	2000
Overcast	Hot	High	False	Yes	5500
Sunny	Mild	High	True	Yes	4500
Sunny	Cool	Normal	True	No	3000
Overcast	Cool	Normal	True	Yes	5800
Rainy	Mild	High	False	No	3500
Rainy	Cool	Normal	False	Yes	6200
Sunny	Mild	Normal	False	Yes	6500
Rainy	Mild	Normal	True	Yes	5900
Overcast	Hot	Normal	False	Yes	4800

(a) You are trying to predict the *Profit* for some new data. Should it be a classification or a regression task? Explain briefly. [1]

(b) Using the **unweighted *k*-NN algorithm** with k = 5, predict both the *Play Golf* and *Profit* values for the following data: <Overcast, Hot, High, True>. Use Hamming distance as the distance function. [3]

(c) Determine whether a golf game should be played or not using a **Naive Bayes Classifier**, if you have recorded the following feature values: <Sunny, Hot, Normal, True>. [3]

(d) The **Naive Bayes classifier** assumes conditional independence of all the features given labels - explain briefly. [1]

5. You are doing a research on daily temperature shifts for a city to analyze global warming. You have divided the city temperature in three categories: high, moderate, and low. Today is a high temperature day. The transition matrix for change in temperature each day has been estimated as follows.

Next Day → Today↓	High	Moderate	Low
High	0.7	0.2	0.1
Moderate	0.3	0.5	0.2
Low	0.2	0.4	0.4

(a) Modeling the scenario as a **Markov model**, determine the probability of the day temperature being low after two days. [3]

(b) Determine the probability of the temperature being high in the long-run (stationary distribution). [3]

(c) What will the result of question 5b be if today's temperature had been low? [1]