National University of Computer and Emerging Sciences, Lahore Campus

	Course Name:	Artificial Intelligence	Course Code:	CS 401
	Program:	BS(CS)	Semester:	Fall 2017
	Duration:	150 Minutes	Total Points:	60
	Paper Date:	Wednesday, Dec 13 2017	Weight	50%
	Section:	ALL	Page(s):	10
	Exam Type:	Final		

Student : Name:_____ Roll No._____ Section:_____

Instruction/Notes:	Students might use a hand written A4 sized cheat sheet during exam but sharing of cheat sheet during examination is not allowed (i.e. every student is allowed to have his own copy of the sheet).
	Please Solve all questions on the question paper If needed you might attach rough sheets at the end.

Problem 1 [Short Questions]   [3 + 3 + 3 + 3 + 3 Points]

a) I have black-box implementation of a **min-heap /priority queue** that can manage the priority on the basis of a real valued attribute such that **smaller values are preferred over larger values**.
For each of the following search algorithms, what must be the **basis of priority** if I want to use the priority queue for managing list of open nodes (typically called a fringe) during the execution of the algorithm.

BFS

DFS

Uniform Cost Search

Greedy Best First

A* Search

b) Which one of the following tasks are examples of classification, Regression or none of these.

Estimating weight of a rice grain

- ○ Classification
- ○ Regression
- ○ None

Separating a rice grain as broken or not broken

- ○ Classification
- ○ Regression
- ○ None

Tracking the trajectory of an object in a visual sequence

- ○ Classification
- ○ Regression
- ○ None

Recognizing handwritten characters

- ○ Classification
- ○ Regression
- ○ None

Finding shortest path from a source node to a destination node

- ○ Classification
- ○ Regression
- ○ None

Computing heuristic value of a game state.

- ○ Classification
- ○ Regression
- ○ None

c) Give a very short description and use of each of the following
Training Data

Validation Data

Test Data

d) Given that the following probability distributions over **independent binary valued variables A, B, and C,** compute their joint distribution.

Pr{A=0} = 0.1	Pr{B=0} = 0.9	Pr{C= 0} = 0.3

e) Given the following Joint distribution over three variables A, B and C compute their marginal distributions and decide if the three variables are independent.

A	B	C	P
0	0	0	0.001
0	0	1	0.009
0	1	0	0.009
0	1	1	0.081
1	0	0	0.009
1	0	1	0081
1	1	0	0.081
1	1	1	0.729

Problem 2 [Search Algorithms]   [2 + 2 + 4 + 4 Points]

Answer the following questions about the search space shown in the figure below, **S** is source and **G** is destination. The arrows indicate the moves and the numbers by the arrows give the step-cost associated with a move. Assume that any ties are settled alphabetically during insertion in the fringe/open list. For the questions that ask for a path, please give your answers in the form S − A − D − G. for path.
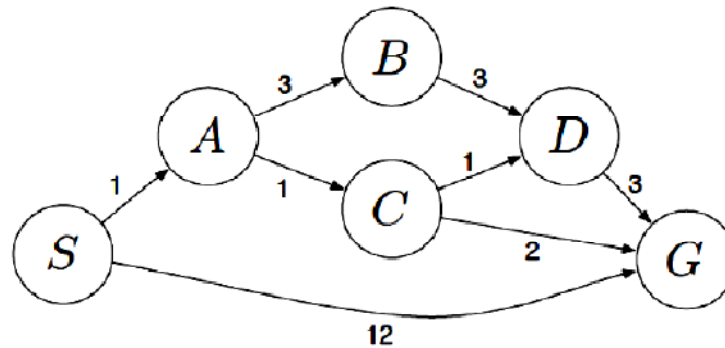
Part a) What is the minimum and maximum number of nodes expanded by DFS if it is used to find the path from S to G in the above search space? Please also specify the nodes that will be expanded.

Part b) What is the minimum and maximum number of nodes expanded by DFS if it is used to find the path from S to G in the above search space? Please also specify the nodes that will be expanded.

Part c) Give a non-trivial heuristic function for which the A* algorithm will find the optimal path from S to G. Specify the path that would be found by A* graph search algorithm if it uses your heuristic function? Also mention the nodes in the fringe after the path found?

Part d) Give a non-trivial heuristic function for which the A* algorithm will find a **non-optimal** path from **S** to **G**. Specify the path that would be found by A* graph search algorithm if it uses your heuristic function? Also mention the nodes in the fringe after the path found?

Problem 3 [Naive Bayes]   [4 + 6 Points]

To classify an email as SPAM or HAM It has been decided to use the Naive Bayes along with unigram features (presence of a word in the email). Further, it has been decided that the words with length larger than 3 will be used for classification

Hence, if an email with **K** key words **w₁, w₂, ... ,wₖ** having length greater than 3 then its label is computed using

$$\text{class}= \arg\max_{c \in C} P(c|d) = \arg\max_{c \in C} P(c) \prod_{i=1}^{K} P(w_i|c).$$

Where **arg max** means that the class with maximum probability will be the predicted class.

The probabilities **P(wi |c)** have already been estimated and given in following table.

W	P(W|SPAM)	P(W|HAM)
note	1/6	1/8
self	1/4	1/4
become	1/4	1/12
perfect	1/8	1/12

Part a) For each of the following cases, compute the label of the email given below

EMAIL TEXT : it is a self note to become perfect

if P(Y = SPAM) is 0.3

if P(Y = HAM) is 0.4

Given the following five emails as a training set:

Training Set

(SPAM) dear sir, I write to you in hope of recovering my gold watch. sir please return my watch.
(SPAM) dear customer, please retry.
(HAM) hey, lunch at 12?
(HAM) fine, watch it tomorrow night
(HAM) dear baig, I am planning to join FAST-LAHORE.

Part b) Compute the estimates of following parameters needed by Naive Bayes. Use Laplace smoothing with K=2 in your estimates. Ignoring all words of length less than 3, and ignoring all non-alphabetical characters.

i) P(W = sir | Y = HAM)

ii) P(W = dear | Y = SPAM)
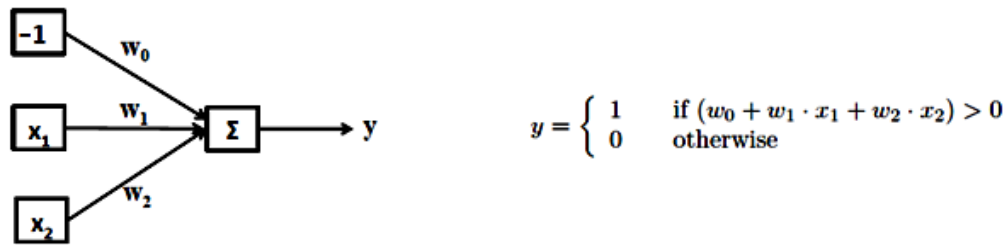
iii) P(Y = HAM)

Problem 4 [Perceptron Learning]   [4 + 4 Points]

Part a) Consider the perceptron with first input always set to -1 and the two features X_1 and X_2 are binary. The output **y** is also binary and is computed using a threshold as shown below

$$y = \begin{cases} 1 & \text{if } (w_0 + w_1 \cdot x_1 + w_2 \cdot x_2) > 0 \\ 0 & \text{otherwise} \end{cases}$$

- o   Specify Weights so that the above perceptron represents a logical AND gate

- o   Specify Weights so that the above perceptron represents a logical OR gate

- o   Specify Weights so that the above perceptron represents a logical NAND gate

- o   Specify Weights so that the above perceptron represents a logical XOR gate

Part b) Derive a weight update rule for a perceptron that uses **tanh** as the activation function.

Problem 5. Decision Tree   [4 +1 Points]

A candy manufacturer interviews a customer on his willingness to eat a candy of a particular color or flavor. Following table shows the collected responses: To predict the Edibility from the color and flavor of the candy it has been decided to build a decision tree.

Color	Flavor	Edibility
Red	Grape	Yes
Red	Cherry	Yes
Green	Grape	Yes
Green	Cherry	No
Blue	Grape	No
Blue	Cherry	No

Part a) Use the above data to learn a decision tree for predicting edibility.

Part b) Using your decision tree, what would you predict for the edibility of a blue, blueberry flavored candy?

Problem 6. Games [1 + 1 + 1 + 5 +2 Points]

Consider the game tree picture below where A-F represent some real values. Assume the nodes are explored from left to right and standard alpha beta pruning is used.
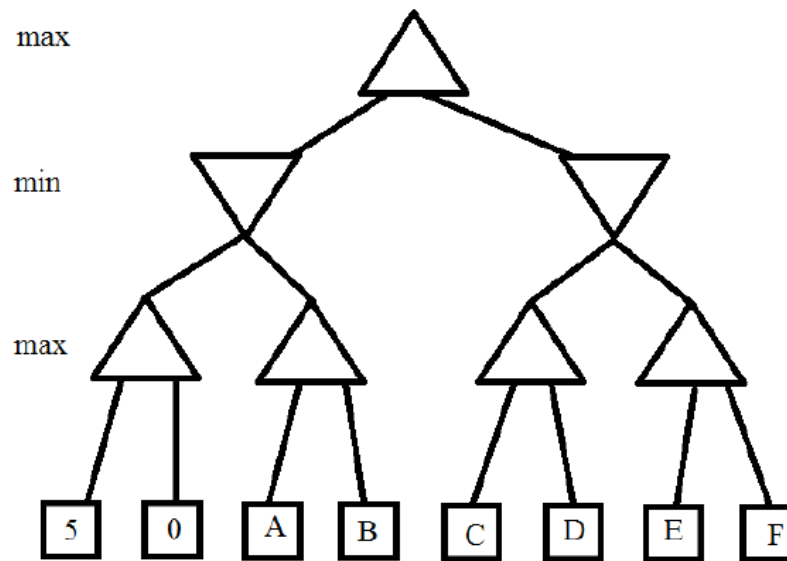
Part a) Give a value of A such that B is pruned.

Part b) [True or False] There are SOME values of A and B such that the sub-tree containing C and D is pruned? If true specify the values and if false give a reason

Part c) Assuming that B = 5 and A = 5, give a value of C and D such that the sub-tree containing E and F is pruned.

Part d) For what values of A, B, C, D, E, and F a none of the nodes will be pruned by the alpha-beta pruning algorithm. **(More than one possible answers)**

Part e) For what values of A, B, C, D, E, and F a maximum of the nodes will be pruned by the alpha-beta pruning algorithm Specify all pruned nodes for the values you specify and also mark all pruned nodes. **(More than one possible solution)**