COURSE CODE AI233AI

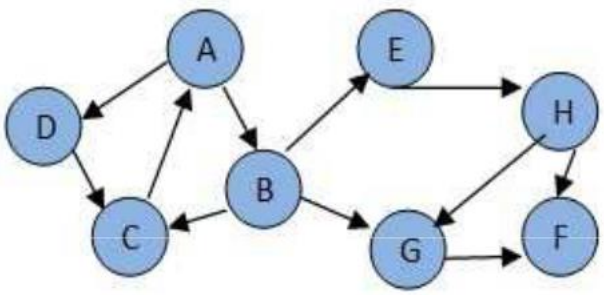
RV COLLEGE OF ENGINEERING
Autonomous Institution affiliated to VTU
III Semester B.E. April -2023 Examinations
DEPARTMENT OF ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING
FUNDAMENTALS OF DATA STRUCTURES AND DATA ANALYSIS
(2022 SCHEME-MODEL PAPER)

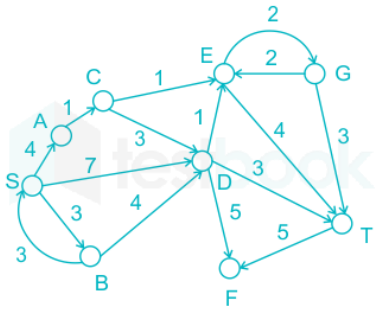
Time: 03 Hours ***Maximum Marks: 100***

Instructions to candidates:

1. Answer all questions from Part A. Part A questions should be answered in first three pages of the answer book only.
2. Answer FIVE full questions from Part B. In Part B, question number 2 is compulsory. Answer any one full question from 3 and 4, 5 and 6, 7 and 8, and 9 and 10.

PART-A (Objective type for one or two marks)
(True & false and match the following questions are not permitted)

1	1.1	You are asked to check the given infix arithmetic expression has balanced pairs of parenthesis, how do you make use of the stack to do this task?	02M	L2	CO1
	1.2	Liked lists have time advantage over arrays during insertion and deletion of data elements, justify.	02M	L2	CO1
	1.3	Consider the following graph. If there is ever a decision between multiple neighbor nodes in the BFS or DFS algorithms, assume we always choose the letter closest to the beginning of the alphabet first.  In what order will the nodes be visited using a Breadth First Search and a Depth First Search?	02M	L3	CO1
	1.4	In a complete k-ary tree, every internal node has exactly k children or no child. The number of leaves in such a tree with n internal nodes is_____	01M	L2	CO1
	1.5	When do we say an hash function is a perfect hash function? Is h(x)=x mod m, perfect hash function?	02M	L1	CO1
	1.6	You are asked to develop a route map between cities, which data structure is suitable for this purpose and why?	01M	L2	CO2
	1.7	Write the conditions to declare the Queue full state of the Circular Queue.	01M	L1	CO1
	1.8	Write a code snippet to free the memory allocated to all the nodes of a singly-linked list which has starting node address as START.	02M	L2	CO1

	1.9	Which shortest path will be discovered by Dijkstra's algorithm between node S to node T?  	01M	L2	CO1
	1.10	By considering an example of E-commernce, give any one application of data mining.	02M	L2	CO3
	1.11	Regression allows researchers to predict or explain the variation in one variable based on another variable, justify this statement with a suitable example.	02M	L3	CO3
	1.12	Write the importance of evaluation phase in the CRISP-DM process.	02M	L1	CO3

PART-B (Maximum subdivisions is limited to 4 in each question)

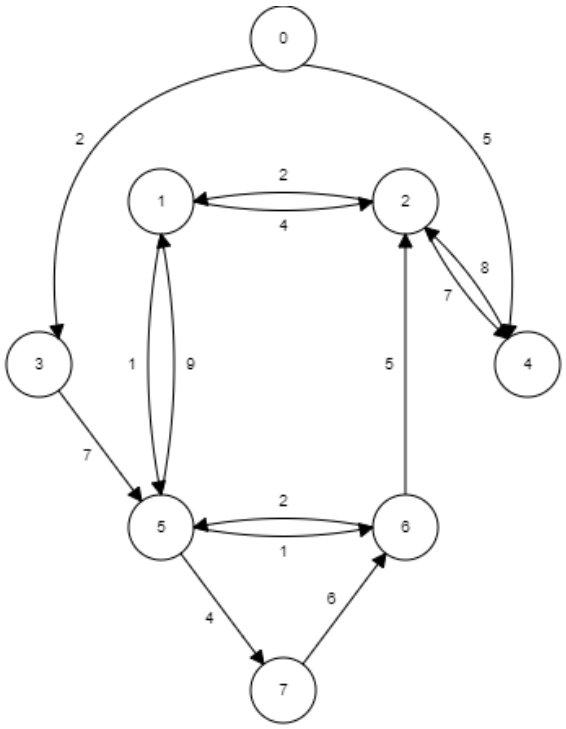
		UNIT-I			

2	a	Write a complete C program to create a singly linked list with the following node structure. After creating a list with 'n' nodes, delete data node which is in K^{th} position.
		
		struct node
		{
		int data;
		struct node *link;
		};
			06M	L3	CO1

	b	Discuss the following data structure concepts; i. Pointers to arrays ii. Passing addresses to functions	04M	L1	CO1
	c	Justify the following with some illustrative examples; i. Insertion and Deletion in a linked list is having better time complexity than insertion and deletion in arrays. ii. Dynamic memory allocation is an overhead for some situations	06M	L3	CO1

		UNIT-II			
3	a	Prove that the height of a binary tree with 'n' internal nodes is at least $\log_2(n + 1)$ and at most $n - 1$.	04M	L2	CO1
	b	Write a C Program to create a Binary Search Tree of Names of Cities along with population number and displaying the Tree using all three traversals.	06M	L3	CO2
	c	Given the following use cases, how and which data structure you finalize? Also mention the operations you implement on these data structure. 1. Storing logs and history of a software system. 2. Online Order Management for an eatery	06M	L3	CO2
		OR			

4	a	You are asked to check whether the given string is a palindrome or not using a character Stack. Write a C program to perform the same. **(Example: MALAYALAM is a palindrome)**	06M	L3	CO2
	b	Write an algorithm to create an expression tree for a valid postfix expression. Show the construction of an expression tree by considering **abc+*d/**	06M	L3	CO2
	c	Write a note on Heap. How does Heap differ from a binary tree?	04M	L1	CO1

		UNIT-III			
5	a	Write Dijkstra's algorithm to find the single source shortest path distances to all the nodes in a Graph. Give the output of Dijkstra's algorithm for the graph in Figure 5a, and assume the source vertex as 0.  Figure 5a	06M	L3	CO1
	b	Write a C program to demonstrate the working of hashing by division, by assuming m as the maximum length of the hash table. Your program can accept any number of input values of type integer.	06M	L3	CO1
	c	Explain the concept of Graph Isomorphism by considering an example of Social Networks.	04M	L2	CO2
		OR			

6	a	Given the following C function, which is written for BFS traversal of a graph, with source vertex as v. Correct the logic if there are any errors, and write the corrected version. Assume proper initializations of adj[][], visited[], front, and rear before the function was called.
		
		void bfs(int v)
		{
		for (i = 1; i <= n; i++)
		if (adj[v][i] && visited[i])
		queue[rear++] = i;
			06M	L3	CO1

		if (front <= rear)
		{
		visited[queue[front]] = 1;
		printf("Node visited = %d\n", queue[front]);
		bfs(adj[v][i]);
		}
		}			

	b	Given the input {4371, 1323, 6173, 4199, 4344, 9679, 1989}, a fixed table size of 10, and a hash function h(x) = x mod 10, show the working of; **a. Hashing with Open addressing and b. Hashing with Quadratic probing.** Write algorithm for implementing Quadratic probing technique for closed hashing.	06M	L3	CO1
	c	Graphs with negative weights are needed sometimes for better modeling the application scenario. Think of any such scenario, and illustrate the same.	04M	L2	CO2

		UNIT-IV			
7	a	Write a block diagram depicting various phases of the CRISP-DM Process and briefly explain the importance of all stages.	08M	L2	CO3
	b	Data driven decision making is an important component in data analytic thinking. By consideting an example of movie recommendation services in OTT platforms like Netflix, discuss the possible approach followed by highlighting at least 4 important points.	08M	L3	CO3
		OR			
8	a	Give suitable real-world business examples for the following; 1. Classificaion 2. Regression 3. Profiling 4. Clustering	08M	L2	CO3
	b	Data driven decision making is an important component in data analytic thinking. By consideting an example of product recommendation services in E-commerce platforms like Amazon, discuss the possible approach followed by highlighting at least 4 important points.	08M	L3	CO3

		UNIT-V			
9	a	Discuss your understanding on the terms Prediction and Prediction Models. Give suitable examples for the same.	08M	L2	CO4
	b	By considering the Mushroom dataset, discuss the process of attribute selection using information gain.	08M	L3	CO4
		OR			
10	a	Discuss your understanding on the terms Induction and Deduction. Give suitable examples for the same.	08M	L2	CO4
	b	By considering the example of people classification, discuss the process of supervised segmentation with Tree-structured models.	08M	L3	CO4

Course Outcomes: After completing the course, the students will be able to

CO1	Apply the knowledge of data structures in providing solutions to some software development requirements.
CO2	Identify appropriate data structures and understand requirements in solving some problems of industry and society
CO3	Perform data analysis of some real-world scientific/business use cases and present the analysis results.
CO4	Use data analysis tools to illustrate the principles of data interpretation, statistical analysis, and graphical visualizations of the datasets.
CO5	Appraise data structures and analysis knowledge to build a successful career as an AIML engineer, work in teams, and communicate their ideas effectively.