



**VIT**<sup>®</sup>

Vellore Institute of Technology  
(Deemed to be University under section 3 of the UGC Act 1956)

Reg. No. :

24MCS1017

மத்தீர் பாதை

Final Assessment Test(FAT) - Nov/Dec 2024

Programme	M.Tech./Ph.D.	Semester	Fall Semester 2024-25
Course Code	MCSE501L	Faculty Name	Prof. Rajakumar R
Course Title	Data Structures and Algorithms	Slot	A2+TA2
		Class Nbr	CH2024250103180
Time	3 hours	Max. Marks	100

**General Instructions**

- Write only Register Number in the Question Paper where space is provided (right-side at the top) & do not write any other details.

**Course Outcomes**

- 1.Understand and analyze the space and time complexity of the algorithms.
2. Identification of suitable data structure for a given problem.
3. Implementation of graph algorithms in various real-life applications.
4. Implementation of heaps and trees for querying and searching.
5. Use of basic data structures in advanced data structure operations.
6. Use of searching and sorting in various real-life applications.

**Section - I**  
**Answer all Questions (10 × 10 Marks)**

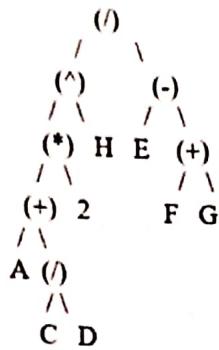
\*M - Marks

Q.No	Question	*M	CO	BL
01.	<pre>function Fn1(n):     i = 1     while i &lt;= n:         print(i)         i = i * 2  function Fn2(n):     if n == 1 then         return 1     else         return n + Fn2(n - 1)</pre> <p>You are given two functions namely, Fn1 and Fn2. Perform the following operations:</p> <ol style="list-style-type: none"> <li>a. Analyze the time complexity of the function Fn1 and explain how the number of iterations grows with respect to n in the function Fn1. (5 marks)</li> <li>b. Analyze the time complexity of the function Fn2 using back substitution method, where the recurrence relation is <math>T(n)=T(n-1) + O(1)</math>. (5 marks)</li> </ol>	10	1	4

02.	<p>A web browser maintains a history of visited web pages using a stack. Each time a user visits a new page, the page's URL is pushed onto the stack. When the user clicks the back button, the most recent page is popped from the stack, and the browser navigates to that page.</p> <ol style="list-style-type: none"> <li>Construct a pseudo-code for array-based implementation of the push and pop operations along with checking stack emptiness for the browser's back functionality (5 Marks)</li> <li>Add a "viewHistory" operation that generates a list of URLs representing the browser's navigation history up to the current page and analyze the space complexity involved in maintaining the stack of URLs (5 Marks)</li> </ol>	10	2	6
03.	<p>A university is developing a student management system to handle student records using a singly linked list. Each node in the linked list represents a student's record and contains the following information:</p> <ul style="list-style-type: none"> <li>• Student ID (a unique identifier)</li> <li>• Name (the student's name)</li> <li>• Age (the student's age)</li> <li>• Course (the course the student is enrolled in)</li> <li>• Next (a pointer to the next student's record in the list)</li> </ul> <p>The university wants to implement the following features:</p> <ol style="list-style-type: none"> <li>Priority Insertion: If a new student is added with an ID that is lower than any existing student IDs, they should be inserted at the beginning of the list. If the ID is higher, they should be inserted at the end of the list. Construct a pseudocode for this priority insertion. (5 marks)</li> <li>Update Feature: If a student record is found by ID, the system should allow for updating the student's name and course without removing them from the list. Develop a pseudocode for this update feature. (5 marks)</li> </ol>	10	2	6
04.	<p>A bakery receives the following delivery times for their pastry orders (in minutes): [30, 10, 40, 20, 50, 70, 60]</p> <ol style="list-style-type: none"> <li>Develop an algorithm for the Cocktail Sort that sorts an array of delivery times in ascending order and analyze the time complexity. (5 marks)</li> <li>Apply the Cocktail Sort algorithm to the provided dataset and demonstrate the sorting process. Show the state of the array after each pass, detailing how the elements are reordered step by step. (5 marks)</li> </ol>	10	5,6	3
05.	<p>A tech company has developed a new search engine designed to quickly retrieve product information from a sorted database containing millions of items. The search engine needs to implement an efficient searching algorithm to allow users to find product details using their unique product IDs. To achieve this, the company decides to use Exponential Search due to its efficiency in handling large sorted datasets. The product IDs in the database are as follows:</p> <p>[1, 5, 9, 14, 23, 38, 47, 56, 73, 82, 97, 105, 128, 150, 175, 200, 250, 300, 350, 400]</p> <p>The company receives a request from a user to find the details of the product with ID 175. Employ exponential search to locate this product efficiently by writing the process and show the steps taken during the search process.</p>	10	5,6	3

06.

10 2,4 2



Using the given expression tree, perform the following:

- Construct a pseudocode for preorder and post order traversal of the expression tree (5 marks)
- Perform a preorder and post order traversals of the expression tree and list the sequence of nodes as they would be visited in each traversal. (5 marks)

07. A sports organization is organizing a knockout tournament for a popular game. The tournament follows a single-elimination format, where players compete in pairs, and the winner of each match advances to the next round until a champion is determined. To facilitate the management of the tournament, the organization decides to use a tournament tree to keep track of the winners of each match efficiently.

The initial dataset of players along with their ranks (Player Number-Rank) is as follows:

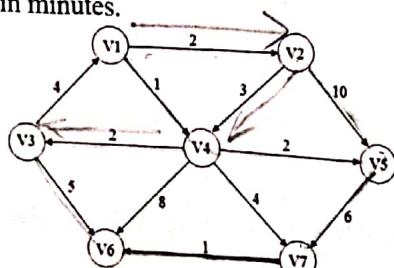
- Players: P1-3, P2-5, P3-6, P4-4, P5-20, P6-8, P7-2, P8-9

The organization needs to build a tournament tree based on the results of the matches, with the following initial results:

- Match 1: P1 vs P2
- Match 2: P3 vs P4
- Match 3: P5 vs P6
- Match 4: P7 vs P8

- Construct a minimum winner tree based on the provided player data and match results. Clearly illustrate the structure of the tree, indicating each match and its winner. (5 marks)
- After the first round of matches, explain how to retrieve the overall winner of the tournament. Describe the steps taken to determine the champion from the tournament tree. (5 marks)

08. A logistics company is optimizing its delivery routes between various warehouses in a city. They need to determine the shortest path for their delivery trucks to minimize travel time. The company represents the city as a weighted directed graph, where, Vertices represent warehouses and Edges represent the roads between the warehouses, with weights indicating the travel time in minutes.



- Using Dijkstra's algorithm, determine the shortest path from warehouse V1 to all other warehouses in the given graph by showing step by step illustration of Shortest Path Calculation. (8 marks)
- List the shortest distance to each warehouse from V1. (2 marks)

09. A telecommunications company is planning to lay down fiber optic cables to connect various communication towers in a region. The company wants to ensure that the network is efficient and cost-effective, meaning they want to connect all towers with the minimum total length of cables. The towers and the potential connections between them can be represented as a weighted undirected graph, where, Vertices represent the communication towers and Edges represent the possible connections between towers, with weights indicating the distance to lay cables. The company has identified the following potential connections between the towers (with distances in kilometers):
- 
- a. Apply the Reverse-Delete algorithm to construct the Minimum Spanning Tree (MST) for the given graph. Describe each step involved in the process, including which edges are removed and which are retained. (8 marks)
- b. Calculate the total weight of the Minimum Spanning Tree obtained from the Reverse-Delete algorithm. (2 marks)
10. A media streaming company is looking to optimize the storage of its audio files by compressing the data. To do this, they plan to use Huffman Coding, and the company has analyzed the frequencies of different sound samples (in kilobytes) used in their files and wants to compress the data efficiently based on these frequencies. The following table shows the sound samples and their respective frequencies:
- | Sound Sample | Frequency (in KB) |
|--------------|-------------------|
| A            | 45                |
| B            | 13                |
| C            | 12                |
| D            | 16                |
| E            | 9                 |
| F            | 5                 |
- a. Use a Min-Heap to construct the Huffman Tree for the given sound samples and their frequencies. Show each step of the process, detailing how the tree is built by merging nodes and calculating new frequencies. (7 marks)
- b. After constructing the Huffman Tree, assign binary Huffman codes to each sound sample. Provide the Huffman code for each sound sample. (3 marks)

**BL-Bloom's Taxonomy Levels - (1.Remembering, 2.Understanding, 3.Applying, 4.Analysing, 5.Evaluating, 6.Creating)**





**VIT®**

**Vellore Institute of Technology**  
(Deemed to be University under section 3 of the UGC Act, 1956)

Reg. No. :

24MCS1017

தெர்டாரிய

**Final Assessment Test(FAT) - Nov/Dec 2024**

Programme	M.Tech.	Semester	Fall Semester 2024-25
Course Code	MCSE503L	Faculty Name	Prof. Thanikachalam V
Course Title	Computer Architecture and Organisation	Slot	C2+TC2
Time	3 hours	Class Nbr	CH2024250103182
		Max. Marks	100

**General Instructions**

- Write only Register Number in the Question Paper where space is provided (right-side at the top) & do not write any other details.

**Course Outcomes**

1. Outline the developments in the evolution of computer architectures and parallel programming paradigms.
2. Comprehend the various programming languages and libraries for parallel computing Platforms.
3. Use of profiling tools to analyze the performance of applications by interpreting the given data.
4. Evaluate efficiency trade-offs among alternative parallel computing architectures for an efficient parallel application design.
5. Develop parallel programs using OpenMP and CUDA and analyze performance parameters such as speed-up, and efficiency for parallel programs against serial programs.

**Section - I**

**Answer all Questions (10 × 10 Marks)**

\*M - Marks

Q.No	Question	*M	CO	BL															
01.	<p>Consider the execution of a program that results in the execution of 2 million instructions on a 400-MHz processor. The program consists of four major types of instructions. The instruction mix and the CPI for each instruction type are given in the table.</p> <table border="1"> <thead> <tr> <th>Instruction Type</th><th>CPI</th><th>Instruction Mix( % )</th></tr> </thead> <tbody> <tr> <td>Arithmetic and Logic</td><td>1</td><td>60</td></tr> <tr> <td>Load/store with cache Hit</td><td>2</td><td>18</td></tr> <tr> <td>Branch</td><td>4</td><td>12</td></tr> <tr> <td>Memory reference with cache miss</td><td>8</td><td>10</td></tr> </tbody> </table> <p>a. Calculate the MIPS rate. [ 5 Marks ]  b. Calculate the CPU execution Time. [ 5 Marks ]</p>	Instruction Type	CPI	Instruction Mix( % )	Arithmetic and Logic	1	60	Load/store with cache Hit	2	18	Branch	4	12	Memory reference with cache miss	8	10	10	1	3
Instruction Type	CPI	Instruction Mix( % )																	
Arithmetic and Logic	1	60																	
Load/store with cache Hit	2	18																	
Branch	4	12																	
Memory reference with cache miss	8	10																	
02.	<p>a. Write an Assembly language program with zero address instructions to evaluate the following arithmetic instruction. [ 5 Marks ]</p> $X = \{ (a - b) + c * (d * e - f) \} / \{ (g + h * k) \}$ <p>b. When implementing a business application, how do multicore and multiprocessor systems differ in terms of Communication, Scalability and Resource Sharing? Identify the benefits and limitations of each system in these aspects. [ 5 Marks ]</p>	10	1	3															

03.	Consider a computer with the following characteristics: 1 Mbyte of main memory, 1 byte of word size, 16 bytes of block size, and 64 Kbytes of cache size. a. For the main memory addresses of CABBE, give the corresponding tag, cache line address, and word offsets for a direct-mapped cache. [ 5 Marks ] b. For the main memory addresses of F0010 give the corresponding tag, cache set, and offset values for a two-way set-associative cache. [ 5 Marks ]	10	1	3
04.	a. Consider a 2-way set associative cache memory with 4 sets and a total of 8 cache blocks (0-7) and a main memory with 128 blocks (0-127). What memory blocks will be present in the cache after the following sequence of memory block references ( 0, 5, 3, 9, 7, 0, 16, 55) use LRU as a replacement policy, assuming that the cache is initially empty? [ 5 Marks ] b. Consider a Direct Mapped Cache with 8 cache blocks (numbered 0-7). If the memory block requests are in the following order 3,5,2,8,0,63,9,16,20,17,25,18,30,24,2,63,5,82,17,24. What memory blocks will be present in the cache after the above sequence of memory block references? [ 5 marks ]	10	1	3
05.	Consider a 5-stage pipelined processor, where each stage needs one cycle for all instructions given below.  Load R1, [1000] Load R3, 5(R2) MUL R4, R1, R3 DIV R5, R1, R4 SUB R6, R4, R5  a. Identify all types of data dependencies (potential data hazards) in the given Instructions. Specify if the data dependence is RAW, WAW, or WAR. [ 5 Marks ] b. Show the number of cycles needed to execute these instructions using operand forwarding. [5 Marks ]	10	4	3
06.	a. Write an openMP program to Calculate the sum of the first N numbers (e.g., $1+2+3+\dots+N$ ) using multiple threads and schedule clauses. [5 Marks ] b. Using openMP, Implement a program to find the maximum value in an array by dividing the array into sections, where each section is processed by a separate thread. Use the openMP sections directive to split the array and calculate the maximum in each section, then compare the results to find the overall maximum. [ 5 Marks ]	10	2	3
07.	You are developing a CUDA-based application for an e-commerce platform that needs to process massive sales data. Each product has a cost and a quantity sold, stored in two large vectors. To determine the total revenue generated from each product, you need to perform element-wise multiplication of the cost and quantity vectors, where each element of the result represents the revenue for a single product. a. Write a C Program for the above case. [4 Marks] b. Write a CUDA program for the above case. [ 6 Marks ]	10	5	3
08.	Discuss how the combination of Intel Trace Analyzer and Collector (ITAC) and Intel VTune Profiler (VTune Amplifier XE) is utilized to identify and address performance bottlenecks in a parallel QuickSort application. Provide examples of specific insights each tool can offer and explain how these insights can lead to optimizations in the QuickSort algorithm.	10	3	3

09. Analyze the various strategies for low-power design in electronic systems, with a particular focus on smart wearable devices such as fitness trackers. Outline at least four specific strategies and illustrate how they are implemented in these devices. [  $4 \times 2.5 = 10$  Marks ]

10 | 5 | 3

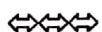
10. The given code will run on the Tomasulo architecture, with an initial setup (Cycle 0).

10 | 4 | 3

MNEMONICS		LATENCY
LD	F0, 10(R2)	2
ADD	F10, F8, F0	1
DIV	F2, F10, F6	5
LD	F4, 0(R3)	2
ADD	F12, F4, F2	1

- a. Name the 3 Stages of Tomasulo Algorithm. [ 3 Marks ]  
 b. Show what happens during the execution of the code for the first four (4) cycles, starting with Cycle 1. You are recommended to use the Tomasulo\_table\_template. [ 7 Marks ]

**BL-Bloom's Taxonomy Levels - (1.Remembering, 2.Understanding, 3.Applying, 4.Analysing, 5.Evaluating, 6.Creating)**





Final Assessment Test(FAT) - Nov/Dec 2024

Programme	M.Tech.	Semester	Fall Semester 2024-25
Course Code	MCSE504L	Faculty Name	Prof. Manimegalai
Course Title	Operating Systems	Slot	D2+TD2
		Class Nbr	CH2024250103184
Time	3 hours	Max. Marks	100

**General Instructions**

- Write only Register Number in the Question Paper where space is provided (right-side at the top) & do not write any other details.

**Course Outcomes**

- Understand the fundamental operating system abstractions, including processes, threads, semaphores, and file systems.
- Implement scheduling, devising and addressing synchronization issues.
- Gain an understanding of memory management tasks.
- Develop real-time working prototypes and different small-scale and medium-scale embedded systems.
- Comprehend the basics of virtualization and differentiate types of virtualization.

**Section - I**

Answer all Questions (7 × 10 Marks)

\*M - Marks

Q.No	Question	*M	CO	BL
01.	<p>Assume that you are responsible for building a banking system where multiple transactions, such as deposits and withdrawals, must be processed concurrently. The system needs to fork a new process for each transaction request from clients. The parent process remains responsible for listening to new transaction requests, while each child process handles one transaction. After processing the transaction, the child process must terminate cleanly without leaving zombie processes.</p> <p>In the course of the testing phase, your team observes that in few cases, the bank balance is updated incorrectly when multiple transactions are processed simultaneously, and orphaned processes start accumulating due to improper handling of child process termination. Considering these points,</p> <p>a) Write a program using the fork() system call to simulate the following scenario: [5 Marks]</p> <ul style="list-style-type: none"> <li>• The parent process listens for transaction requests.</li> <li>• For each transaction (deposit/withdrawal), the parent forks a new child process.</li> <li>• The child process handles a transaction (e.g., deposit or withdraw an amount in Rs.) and updates the balance, then terminates.</li> </ul> <p>b) Explain how the fork() system call works in this scenario, detailing the roles of the parent and child processes and their interaction in managing transactions. Additionally, describe the system calls required to prevent zombie processes and their importance in maintaining proper process management. [5 Marks]</p>	10	1	3

02.	<p>In a museum exhibit with limited capacity, visitors are allowed to view the artifacts simultaneously, while the <b>curator</b> occasionally performs maintenance tasks. Multiple visitors can be inside the exhibit at the same time, but when the curator performs maintenance, no visitors are allowed to enter and they stand in a queue and wait till the maintenance is over. Two semaphores are used to manage access: <b>VisitorAccess</b> is initialized to 1 and controls the entry of visitors, and <b>CuratorAccess</b> is also initialized to 1 and controls when the curator can perform maintenance.</p> <p>a) Illustrate with explanation the pseudocode for <b>Process VisitorEntry</b>, which allows multiple visitors to enter and view the exhibit, while ensuring that the curator does not begin maintenance while visitors are present. [6 Marks]</p> <p>b) Illustrate with explanation the pseudocode for <b>Process CuratorMaintenance</b>, which allows the curator to perform maintenance when no visitors are inside. The curator should wait until the exhibit is empty before beginning maintenance, and no visitors should be allowed to enter during the maintenance period. [4 Marks]</p>	10	1	4
03.	<p>Imagine, you are managing the resource allocation for a large hospital where critical resources such as ventilators, MRI machines and operating rooms are limited hence shared among multiple patients. The hospital system tracks patient resource requests and allocations, and it has run into a halt issue where multiple patients are waiting indefinitely for resources.</p> <p>The hospital has 2 Ventilators, 3 MRI Machines and 2 Operating Rooms. The current resource allocation and requests from 5 patients are as follows:</p> <p>Patient A allocated with one MRI machine and has requested for one ventilator. Patient B allocated with one ventilator and has requested for one MRI machine and one operating room. Patient C allocated with one operating room and one ventilator and has requested for one MRI machine. Patient D allocated with one MRI machine and has requested for one operating room. Patient E allocated with none and has requested for one ventilator and one MRI machine.</p> <p>a) Construct a Resource Allocation Graph (RAG) for the above scenario, considering multiple instances of each resource type. [4 Marks]</p> <p>b) Using a suitable algorithm, find the order of the patients for the treatment, ensuring no patient is untreated. [6 Marks]</p>	10	2	3
04.	<p>In a segmented memory management system, a process is divided into four segments with the following properties:</p> <ul style="list-style-type: none"> <li>• Segment 0 (Code): Start Address = 8000, End Address = 9000</li> <li>• Segment 1 (Data): Start Address = 9001, End Address = 9500</li> <li>• Segment 2 (Heap): Start Address = 9501, End Address = 11000</li> <li>• Segment 3 (Stack): Start Address = 11001, End Address = 12000</li> </ul> <p>A program attempts to access the following references:</p> <ol style="list-style-type: none"> <li>1. Segment 0, Offset 500</li> <li>2. Segment 1, Offset 450</li> <li>3. Segment 2, Offset 1600</li> <li>4. Segment 3, Offset 750</li> <li>5. Segment 1, Offset 600</li> </ol> <p>a) For each logical address, determine whether it is valid or invalid based on the segment properties. [5 Marks]</p> <p>b) Generate the logical addresses represented in 16-bits for the above references. [5 Marks]</p>	10	3	1

05.	<p>The tourism development authority "ABCD" manages a comprehensive digital platform that stores and organizes a wide range of tourism-related data. This data includes both small files, such as itineraries, brochures, and travel guides, as well as large media files like high-resolution images and promotional videos of various tourist destinations. As the system needs to handle diverse file sizes, it is crucial to ensure that the file storage method chosen is efficient in terms of both space utilization and data access speed. Additionally, the platform requires smooth management and minimal fragmentation to avoid performance bottlenecks, especially as files are frequently accessed, updated, or expanded when new content is added. Given these requirements explain the following with diagrams:</p> <p>a) Which file allocation methods would you recommend to balance efficient storage of small files with fast access to large media files while minimizing fragmentation? Justify your claim with proper reasons. [5 Marks]</p> <p>b) How can these methods ensure the system remains scalable and agile as the volume of tourism data grows? [5 Marks]</p>	10	3	1
06.	<p>In a real-time operating system, tasks have specific timing constraints that must be met to ensure system reliability and performance. Consider the following tasks with their respective periods and execution times:</p> <p>Task T1: Period = 30ms, Execution Time = 5ms      Task T2: Period = 5ms, Execution Time = 2ms      Task T3: Period = 10ms, Execution Time = 2ms</p> <p>a) Calculate the CPU utilization for the set of tasks and determine if the system is schedulable under Rate Monotonic Scheduling. [5 Marks]</p> <p>b) Create a Gantt chart showing the scheduling of these tasks over one complete hyper-period. Are there any idle times within the hyper-period, and if so, how do they occur and impact the schedulability of the task set? [5 Marks]</p>	10	4	5
07.	<p>As the IT administrator for a large school, you are responsible for managing a complex infrastructure that supports various applications across multiple departments. Currently, the school relies on separate physical servers for critical tasks such as managing student records, hosting the school's website, running learning management systems (LMS), and handling day-to-day administrative tasks. While this setup has been functional, maintaining and managing these individual servers has become increasingly costly and cumbersome, especially as the school's digital demands continue to grow.</p> <p>With the growing number of applications and increasing demands on IT resources, the current infrastructure presents several challenges. Each physical server consumes space, power, and cooling resources, leading to higher operational costs. Moreover, managing multiple isolated servers results in inefficiencies, underutilized hardware, and a heavier administrative workload. As the school expands and adopts new technologies, it is becoming clear that a more efficient solution is necessary to meet these evolving needs.</p> <p>a) Device a technique that you would adapt to abstract hardware of a single computer into several different execution environments to improve efficiency and reduce operational costs? Explain the method with the help of necessary diagrams [5 Marks]</p> <p>b) What are the key differences between Type 1 and Type 2 hypervisors, and which would be more suitable for the school's IT environment? [5 Marks]</p>	10	5	2

**Section - II**  
**Answer all Questions (2 × 15 Marks)**

\*M - Marks

Q.No	Question	*M	CO	BL
08.	<p>Consider a system with a single CPU core and four processes A, B, C and D.</p> <ul style="list-style-type: none"> <li>i. Process A arrives at <math>t = 0</math>, and runs on the CPU for 8-time units before it finishes.</li> <li>ii. Process B arrives at <math>t = 4</math>, and requires an initial CPU time of 3 units, after which it blocks to perform I/O for 3 time units. After returning from I/O wait, it executes for a further 5 units of CPU time before terminating.</li> <li>iii. Process C arrives at <math>t = 6</math>, and runs for 2 units of time on the CPU before terminating.</li> <li>iv. Process D arrives at <math>t = 8</math>, and requires an initial CPU time of 2 units, after which it blocks to perform I/O for 2 time units. After returning from I/O wait, it executes for a further 3 units of CPU time before terminating.</li> </ul> <p>Compare the below scheduling algorithms and comment on the better algorithm based on the average waiting time for this case</p> <ul style="list-style-type: none"> <li>a) Non Preemptive SJF Versus Preemptive SJF [7 Marks]</li> <li>b) Non Preemptive Priority Scheduling Versus Preemptive Priority Scheduling (Consider the priority sequence as D&gt;B&gt;C&gt;A) [8 Marks]</li> </ul> <p>NB: Only the size of the current CPU burst (excluding the time spent for waiting on I/O) is considered as the “job size” in these schedulers.</p>	15	2	5
09.	<p>You are managing the memory for the Income Tax Department's application, which processes tax returns and manages taxpayer data. The system uses a page size of 4 KB and has a total memory capacity of 16 KB.</p> <p>During peak filing season where all the frames are used, the application accesses the following pages in the order below:</p> <p>1, 2, 3, 4, 1, 2, 5, 1, 2, 3, 4, 5, 1, 4, 5</p> <ul style="list-style-type: none"> <li>a) Evaluate FIFO, LRU and Optimal page replacement algorithms to determine which is the most efficient for managing memory under this workload. [9 Marks]</li> <li>b) In case of FIFO, if the number of frames is increased by one, then compare the performance with the previous case. What is the phenomenon named if increasing the frames reduces the performance? Comment on this phenomenon. [6 Marks]</li> </ul>	15	3	5

**BL-Bloom's Taxonomy Levels - (1.Remembering, 2.Understanding, 3.Applying, 4.Analysing, 5.Evaluating, 6.Creating)**





**VIT**<sup>®</sup>

Vellore Institute of Technology  
(Deemed to be University under section 3 of the UGC Act, 1956)

Reg. No. :

24MCS1017

மத்தீர் ஜாவாடாபு

### Final Assessment Test(FAT) - Nov/Dec 2024

Programme	M.Tech.	Semester	Fall Semester 2024-25
Course Code	MCSE506L	Faculty Name	Prof. Balasundaram A
Course Title	Database Systems	Slot	E2+TE2
Time	3 hours	Class Nbr	CH2024250103186
		Max. Marks	100

#### General Instructions

- Write only Register Number in the Question Paper where space is provided (right-side at the top) & do not write any other details.

#### Course Outcomes

- CO1. Design and implement a database depending on the business requirements, considering various design issues
- CO2. Understand the concepts of Indexing, Query optimization, transaction management, concurrency control, and recovery mechanisms
- CO3. Learn to apply parallel and distributed databases in Real-time scenarios
- CO4. Categorize and design the structured, semi-structured, and unstructured databases
- CO5. Characterize the database threats and their countermeasures

#### Section - I Answer all Questions (6 × 15 Marks)

\*M - Marks

Q.No	Question	*M	CO	BL
01.	<p>i. Assume that you are planning to maintain the data of an online movie reservation system. Justify with suitable reasons why a relational database model based approach would be best suited for this case compared to the conventional way of maintaining the data in files. (5 marks)</p> <p>ii. Consider the following ER diagram representing a Football league database. Note that a player can stay in only one team during a season.</p> <pre>     erDiagram         {             string matchId;             string stadiumName;             number score;             string playerName;             number age;             string date;             string teamName;             string result;             string ranking;              match {                 string matchId;                 string stadiumName;                 string date;             }             score {                 number score;                 string result;             }             name {                 string playerName;                 number age;             }              match }o--o{ player : "played"             score }o--o{ player : "teamPlayed"             player }o--o{ team : "playerOf"             player }o--o{ team : "seasonScore"             team }o--o{ ranking : "ranking"         }     </pre> <p>Identify the cardinality ratio and map this ER diagram to relational model. (8 marks) Mention the primary key and foreign keys for all the possible relations. (2 marks)</p>	15	1	3



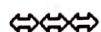
04.	<p>i. Consider the below XML file named Students.xml. Design the equivalent XSD file to check whether the XML is well formed or not. (6 marks)</p>	15	4	3
	<pre>&lt;students&gt;   &lt;student&gt;     &lt;name&gt;Rick Grimes&lt;/name&gt;     &lt;age&gt;35&lt;/age&gt;     &lt;subject&gt;Maths&lt;/subject&gt;     &lt;gender&gt;Male&lt;/gender&gt;   &lt;/student&gt;   &lt;student&gt;     &lt;name&gt;Daryl Dixon &lt;/name&gt;     &lt;age&gt;33&lt;/age&gt;     &lt;subject&gt;Science&lt;/subject&gt;     &lt;gender&gt;Male&lt;/gender&gt;   &lt;/student&gt;   &lt;student&gt;     &lt;name&gt;Maggie&lt;/name&gt;     &lt;age&gt;36&lt;/age&gt;     &lt;subject&gt;Arts&lt;/subject&gt;     &lt;gender&gt;Female&lt;/gender&gt;   &lt;/student&gt; &lt;/students&gt;</pre> <p>ii. Consider a relation <b>BranchDetails</b> (<u>BranchID</u>, Location, State, MgrID, CustomerCount). Transform the relational data into its equivalent XML representations for the following scenarios:</p> <ol style="list-style-type: none"> <li>Select the BranchID and ManagerID for Chennai Branch. (2 marks)</li> <li>Create an XML document with BranchID, Location and CustomerCount sorted by CustomerCount. (2 marks)</li> <li>Create an XML document with BranchID, Location and CustomerCount where the customer count is &gt; 100000. (2 marks)</li> <li>Create an XML document with BranchID, Location and CustomerCount for the branches located in the state “TamilNadu”. (3 marks)</li> </ol>			
05.	<p>Assume you are working as a database administrator in a software development company. You are tasked with designing database for a social media platform managing user-generated content, like posts, comments, likes, and user profiles.</p> <ol style="list-style-type: none"> <li>Explain how a NoSQL database could work effectively in this scenario. Your answer must include at least three key features of NoSQL (6 marks).</li> <li>Also explain how CAP (Consistency, Availability, Partition Tolerance) theorem will be applied for the given scenario in order to optimize both user experience and data integrity of the social media platform (5 marks).</li> <li>Discuss in general about the potential challenges and limitations that exist in using NoSQL for real world applications. (4 marks)</li> </ol>	15	5	2
06.	<p>Consider a relation CUSTOMERS (CUSTID, CUSTNAME, DEPT, DATEJOINED, MOBILE, EMAIL).</p> <ol style="list-style-type: none"> <li>Demonstrate with a query and its relevant explanation detailing how SQL Injection can be performed over the CUSTOMERS relation and how the CUSTOMERS data get exposed due to SQL injection. (5 marks)</li> <li>Discuss in detail how Discretionary Access Control is provided at account level with the customers relation as an example. (5 marks)</li> <li>Create a VIEW over the CUSTOMERS table and demonstrate a scenario with necessary explanation how VIEWS enhance the security aspect in database systems.(5 marks)</li> </ol>	15	5	4

**Section - II**  
**Answer all Questions (1 × 10 Marks)**

\*M - Marks

Q.No	Question	*M	CO	BL
07.	<p>A University is planning to install new Wi-Fi hotspots across the different academic buildings and hostels in the university to improve internet access. Each building has information such as buildingID, type (hostel or academic), Location (geometry).</p> <p>i. Write SQL queries using spatial database functions to locate all the academic buildings within 500 meters of Hostel block A (Id=Hostel_A_1). (3 marks)</p> <p>ii. Explain how spatial indexing could be used to improve the performance of above query. (3 marks)</p> <p>iii. Suppose the Building table gets modified to store building image along with buildingID, Location. Explain how content-based image retrieval is used for fetching the details of specific building given its image as input. ( 4 marks)</p>	10	4	4

**BL-Bloom's Taxonomy Levels - (1.Remembering, 2.Understanding, 3.Applying, 4.Analysing, 5.Evaluating, 6.Creating)**





**VIT®**

Vellore Institute of Technology  
(Deemed to be University under section 3 of the UGC Act, 1956)

Reg. No. :

24MCS1017

Final Assessment Test(FAT) - Nov/Dec 2024

Programme	M.Tech./Ph.D.	Semester	Fall Semester 2024-25
Course Code	MCSE616L	Faculty Name	Prof. Sellam V
Course Title	Data Visualization	Slot	F2
		Class Nbr	CH2024250103237
Time	3 hours	Max. Marks	100

#### General Instructions

- Write only Register Number in the Question Paper where space is provided (right-side at the top) & do not write any other details.

#### Course Outcomes

1. Analyze the different data types, visualization types to bring out the insight.
2. Relate the visualization towards the problem based on the dataset to analyze and bring out valuable insight on large dataset.
3. Design visualization dashboard to support the decision making on large scale data.
4. Demonstrate the analysis of large dataset using various visualization techniques and tools.

#### Section - I

Answer all Questions (10 × 10 Marks)

\*M - Marks

Q.No	Question	*M	CO	BL
01.	<p>FinViz Analytics has created a dashboard for three user types: a) Individual investors b) Financial analysts, c) Portfolio managers. Each user needs insights from Market Data (stock prices, volumes, indices, etc.), Customer Portfolio Data (asset holdings, allocation, returns), and Transaction Data (trade type, price, volume, fees).</p> <p>i. Exploit one key task for each user related to tracking stock performance, sector trend analysis, or portfolio risk management. (3 Marks)</p> <p>ii. Suggest two effective dashboard features for each user. (3 Marks)</p> <p>iii. Design a middle-level data abstraction for each user type to group stocks by performance and summarize sectors. (4 Marks)</p>	10	1,2	1
02.	<p>As a housing board analyst, use charts to visualise key data and explain each chart's significance for:</p> <p>i. Show resident distribution by age, family size, income, and compare satisfaction with maintenance, amenities, and safety across developments. (4 Marks)</p> <p>ii. Track rent prices, income percentage spent on housing, and trends in occupancy and vacancy rates, linking these to economic conditions. (3 Marks)</p> <p>iii. Policy Impact: Assess the effectiveness of housing policies by comparing stability metrics (e.g., eviction rates) for low-income families before and after implementation. (3 Marks)</p>	10	1,2,3	3

03. Exploit using vector-based visualization system for an agricultural organization that needs to display various metrics related to crop health and environmental conditions. 10 3,4 4

Loc ID	Latitude	Longitude	Date	Soil Moisture (%)	Sunlight Hours (hrs)	Temperature (°C)	Crop Yield (tons/ha)
1	34.0522	-118.2437	2024-08-01	30	8	25	2.5
1	34.0522	-118.2437	2024-08-02	28	7	26	2.8
2	34.0522	-118.2537	2024-08-01	35	10	24	3.0
2	34.0522	-118.2537	2024-08-02	33	9	25	3.2
3	34.0622	-118.2437	2024-08-01	40	6	23	2.0
3	34.0622	-118.2437	2024-08-02	38	5	24	2.1
4	34.0622	-118.2537	2024-08-01	25	12	27	3.5
4	34.0622	-118.2537	2024-08-02	23	11	28	3.7

Apply appropriate vector visualisation techniques to analyse the following and justify your answer:

- i) Soil Moisture and Crop Yield ( 3 Marks)
- ii) Sunlight Hours and Temperature Data (3 Marks)
- iii) Overall Crop Health and Performance (4 Marks)

04. A logistics company needs a dashboard to monitor last year's delivery data, assuming monthly data for 12 months, covering total deliveries, average delivery time (in hours), and on-time delivery rate (as a percentage). The goal is to optimize routes and enhance customer satisfaction. 10 2,4 3
- i. Design a strategy to visualize: (6 marks)
- Delivery Volume: Show the total deliveries over 12 months to identify trends and seasonality.
  - Average Delivery Time: Track the average delivery time over 12 months to assess efficiency and peak periods.
  - On-Time Rate: Display the on-time delivery rate over 12 months to monitor service quality and improvement areas.
- ii. Discuss potential challenges in visualizing time series data for 12 months of logistics data and suggest solutions. (4 marks)

05.	Perform hierarchical clustering on a retail customer dataset (1,000 customers) with features like income, age, spending score, purchases, and product preferences to identify customer segments for marketing. i. Describe the clustering method, justify the linkage choice, and explain how to construct a dendrogram to determine the optimal number of clusters. (4 Marks) ii. Discuss how to evaluate cluster stability, define reliability metrics, and profile segments with key metrics. (4 Marks) iii. Propose two targeted marketing strategies for each segment, including success metrics. (2 Marks)	10	1,5	5																																																																																																																		
06.	The company monitors monthly sales data across five distinct regions throughout the year. Your objective is to create a dual-visualization approach to analyse regional sales performance over time. Use both a heat map and an interactive line chart to reveal trends, patterns, and anomalies in the data. (10 Marks)	10	3,5	5																																																																																																																		
07.	<table border="1"> <thead> <tr> <th>Region</th><th>Jan</th><th>Feb</th><th>Mar</th><th>Apr</th><th>May</th><th>June</th><th>July</th><th>Aug</th><th>Sep</th><th>Oct</th><th>Nov</th><th>Dec</th></tr> </thead> <tbody> <tr> <td>North</td><td>1000</td><td>1200</td><td>1100</td><td>1050</td><td>1150</td><td>1250</td><td>1300</td><td>1400</td><td>1350</td><td>1200</td><td>1100</td><td>1400</td></tr> <tr> <td>South</td><td>950</td><td>1000</td><td>1050</td><td>980</td><td>1150</td><td>1100</td><td>1200</td><td>1300</td><td>1250</td><td>1150</td><td>1050</td><td>1300</td></tr> <tr> <td>East</td><td>850</td><td>900</td><td>950</td><td>910</td><td>1000</td><td>950</td><td>1100</td><td>1200</td><td>1150</td><td>1050</td><td>1000</td><td>1150</td></tr> <tr> <td>West</td><td>800</td><td>850</td><td>900</td><td>870</td><td>950</td><td>900</td><td>950</td><td>1000</td><td>980</td><td>900</td><td>850</td><td>1000</td></tr> <tr> <td>Central</td><td>1050</td><td>1100</td><td>1200</td><td>1150</td><td>1250</td><td>1300</td><td>1350</td><td>1400</td><td>1300</td><td>1250</td><td>1200</td><td>1450</td></tr> </tbody> </table> <p>Analyse a community health initiative that examines various health-related features across different neighbourhoods in a city:</p> <table border="1"> <thead> <tr> <th>Feature Type</th><th>Location (X,Y)</th><th>Population Density (people/sq.m)</th><th>Access to healthcare services</th><th>Prevalence of chronic diseases (%)</th><th>Public Health Programs Available</th></tr> </thead> <tbody> <tr> <td>Neighbourhood A</td><td>(10, 20)</td><td>5000</td><td>YES</td><td>15</td><td>YES</td></tr> <tr> <td>Neighbourhood B</td><td>(15, 25)</td><td>3500</td><td>NO</td><td>25</td><td>NO</td></tr> <tr> <td>Neighbourhood C</td><td>(20, 30)</td><td>4200</td><td>YES</td><td>10</td><td>YES</td></tr> <tr> <td>Neighbourhood D</td><td>(5, 10)</td><td>8000</td><td>NO</td><td>30</td><td>NO</td></tr> <tr> <td>Neighbourhood E</td><td>(25, 15)</td><td>2000</td><td>YES</td><td>20</td><td>YES</td></tr> </tbody> </table> <p>i. Propose an icon-based geometric projection technique for visualizing health features. Specify icons for each feature type, a color scheme for prevalence of chronic diseases, and how icon size reflects population density. Justify how your design communicates spatial relationships and health disparities. (5 Marks) ii. Discuss how pixel-based visualization can be utilized to assess health service access and inform public health strategies. Consider insights obtained from access to healthcare services and prevalence of chronic diseases. (5 Marks)</p>	Region	Jan	Feb	Mar	Apr	May	June	July	Aug	Sep	Oct	Nov	Dec	North	1000	1200	1100	1050	1150	1250	1300	1400	1350	1200	1100	1400	South	950	1000	1050	980	1150	1100	1200	1300	1250	1150	1050	1300	East	850	900	950	910	1000	950	1100	1200	1150	1050	1000	1150	West	800	850	900	870	950	900	950	1000	980	900	850	1000	Central	1050	1100	1200	1150	1250	1300	1350	1400	1300	1250	1200	1450	Feature Type	Location (X,Y)	Population Density (people/sq.m)	Access to healthcare services	Prevalence of chronic diseases (%)	Public Health Programs Available	Neighbourhood A	(10, 20)	5000	YES	15	YES	Neighbourhood B	(15, 25)	3500	NO	25	NO	Neighbourhood C	(20, 30)	4200	YES	10	YES	Neighbourhood D	(5, 10)	8000	NO	30	NO	Neighbourhood E	(25, 15)	2000	YES	20	YES	10	6,7	5
Region	Jan	Feb	Mar	Apr	May	June	July	Aug	Sep	Oct	Nov	Dec																																																																																																										
North	1000	1200	1100	1050	1150	1250	1300	1400	1350	1200	1100	1400																																																																																																										
South	950	1000	1050	980	1150	1100	1200	1300	1250	1150	1050	1300																																																																																																										
East	850	900	950	910	1000	950	1100	1200	1150	1050	1000	1150																																																																																																										
West	800	850	900	870	950	900	950	1000	980	900	850	1000																																																																																																										
Central	1050	1100	1200	1150	1250	1300	1350	1400	1300	1250	1200	1450																																																																																																										
Feature Type	Location (X,Y)	Population Density (people/sq.m)	Access to healthcare services	Prevalence of chronic diseases (%)	Public Health Programs Available																																																																																																																	
Neighbourhood A	(10, 20)	5000	YES	15	YES																																																																																																																	
Neighbourhood B	(15, 25)	3500	NO	25	NO																																																																																																																	
Neighbourhood C	(20, 30)	4200	YES	10	YES																																																																																																																	
Neighbourhood D	(5, 10)	8000	NO	30	NO																																																																																																																	
Neighbourhood E	(25, 15)	2000	YES	20	YES																																																																																																																	

08.	Analyze a dataset of 6 species in a coastal ecosystem with variables: habitat type, population, water quality, size, and threat level. i. Visualize the relationships between population, water quality, size, and threat level. Justify the chosen method's effectiveness in showing correlations. (5 Marks) ii. Analyze the relationships between species, habitat type, and threat level. Compare the results. (5 Marks)	10	1,4	3
09.	Analyse student dataset which has 35 records includes ID, department, GPA, age, gender, course, and attendance using Tableau. a. Propose 5 advanced visualizations to analyse the link between demographics and academic performance, detailing their purpose and insights. (5 Marks) b. Detail how to leverage Tableau's interactive filters, parameters, and dashboard actions to facilitate advanced demographic segmentation, allowing for dynamic cross-filtering between multiple visualizations to uncover deeper insights. (5 Marks)	10	5,6	1
10.	A dataset of gym members with attributes: ID, Workout Type, Fitness Level, Attendance Rate, Age Group, and Gender. i. Create a Tableau visualisation using Fitness Level, Attendance Rate, and Gender, and explain your design choices. (4 Marks) ii. Based on the visualization, answer: a. What is the correlation between fitness levels and attendance rates? b. Are there gender-based performance differences by workout type? c. What recommendations could improve engagement for low-performing groups?	10	5,6,7	6

**BL-Bloom's Taxonomy Levels - (1.Remembering, 2.Understanding, 3.Applying, 4.Analysing, 5.Evaluating, 6.Creating)**





ST/TECH/2024

**LAB FAT – Nov 2024**

Programme : M.Tech.  
Course : Technical Report Writing  
Faculty : Dr. I. Ajit  
Time : 1 ½ hours

Semester : Fall 2024  
Code : MENG 501 P  
Slot : L3+L4+L23+L24  
Max. Marks : 50

**Answer ALL the questions**

**50 Marks**

1. Explain the purpose of a *review of literature* in a research project. Discuss how it helps in identifying research gaps, analyzing existing methodologies, and building a strong foundation for the study. Provide examples to illustrate your answer. 10
2. Describe the structure of a research paper, highlighting the purpose and key components of each section (e.g., Abstract, Introduction, Literature Review, Methodology, Results, Discussion, Conclusion, and References). Explain how a well-organized structure contributes to the clarity and effectiveness of the research presentation. 10
3. Proofread the following paragraph and rewrite it correctly with proper punctuation and spelling:  

Post-traumatic stress disorder (PTSD) is a severe anxiety disorder that effects millions of people around the world. Individuals can develop PTSD after experiencing any event that results in psychological trauma. Symptoms of ptsd involve flashbacks to the traumatic event, nightmares, obsessive behavior, anger, insomnia, difficulty concentrating, and hypervigilance. Individuals Who suffer from PTSD can experience significant difficulties in social relationships, have lower selfesteem, and have trouble maintaining employment.

10
4. Write an essay consisting of a few technical jargons on any ONE of the following topics. Restrict your essay to 200 words. Underline all the technical vocabulary you have used in it.  
  1. Digital Marketing
  2. Impact of Technology on Communication
  3. Technology and Negative Effects10
5. What is the primary purpose of using a questionnaire in research, and how does it help in collecting and analyzing data effectively? Provide examples of scenarios where a questionnaire is the most suitable data collection tool. 10