 VIT Vellore Institute of Technology <small>(Deemed to be University under section 3 of the UGC Act, 1956)</small>	Final Assessment Test – Fall 2024-25 Semester - Nov 2024 KEY	
	Course code : BCSE302L	Slot : A1+TA1
	Course Title : Database Systems	Time: Three Hours
	Class Number(s) : VL2024250101532	Max. Marks: 100
	Faculty Name : R.Rajeshkannan	School: SCOPE
Emp Id : 12411	Mobile No.: 9789007360	

KEEPING MOBILE PHONE/ANY ELECTRONIC GADGETS, EVEN IN 'OFF' POSITION IS TREATED AS EXAM MALPRACTICE

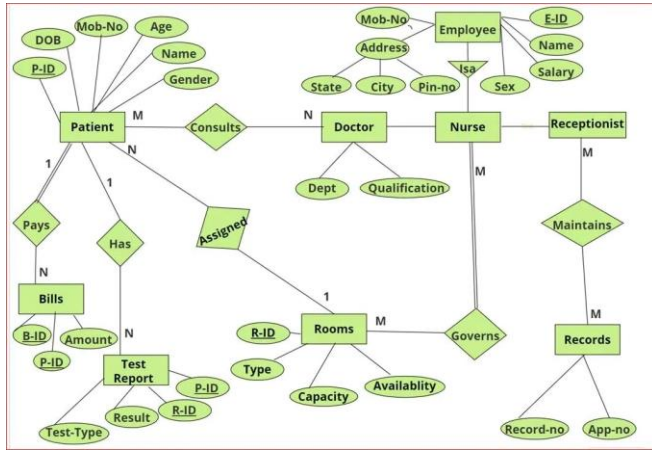
General Instructions if any:

1. Non Programmable calculator is permitted : YES /NO
2. Reference tables are permitted : YES / NO (if Yes, Please specify:
)

**Answer ALL Questions
(10 X 10 = 100 Marks)**

Sl. No:	Questions	Module No	Marks	CO	BL	E/A/T
1.	Discuss three-schema architecture used in database management. What are the key components, and how do they interact with each other? Why do we need mappings between schema levels?	1	10	CO1	L1	E
	Explain Three-Schema Architecture Internal Schema: Conceptual Schema: External Schema: Need for Mappings between Schema Levels Mappings between schema levels are essential for several reasons: Data Abstraction: Data Independence: <ul style="list-style-type: none"> o Logical Data Independence: o Physical Data Independence: Consistency and Integrity:					
2.	Design an Entity-Relationship Diagram (ER Diagram) for a Hospital Management System (HMS) that enables the registration and management of patient information, including demographics, medical history, and current condition, while featuring a doctor directory that maintains details such as specialties, contact information, and availability. The system should assist in room management by identifying available and occupied rooms, and handle nurse information,	2	10	CO2	L3	A

including departmental assignments and shift schedules. Additionally, it must provide functionality for generating, storing, and retrieving patient test reports with detailed results and dates, while managing records through unique identification numbers for easy reference. Billing functionality is crucial for handling billing information, including bill generation and payment tracking, and the system should support receptionist functions for retrieving patient data and scheduling appointments. Your task is to create an ER diagram that includes relevant entities, attributes and relationships, indicating cardinality and participation constraints.



3. Discuss the concepts of generalization and specialization in database design, using the extended entity-relationship model as a framework. Provide examples to illustrate these concepts and analyze the various types of constraints, including disjointness and completeness, and their impact on subclass relationships and the overall superclass structure.

Generalization and specialization are two fundamental concepts in database design, particularly in the context of the extended entity-relationship (EER) model.
 Explain
 Generalization
 Specialization
 Explain Constraints in the Extended Entity-Relationship Model
 Disjointness Constraint
 Completeness Constraint

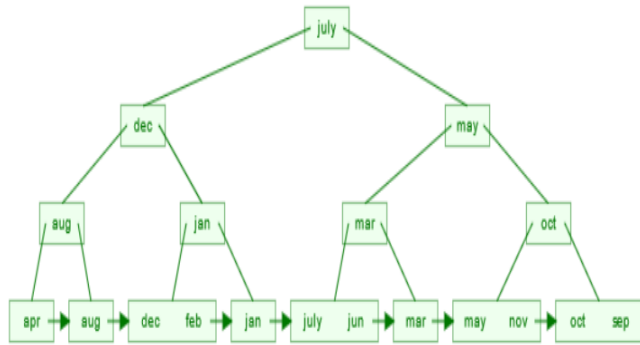
4. Consider relation $R(X,Y,Z,W,V)$ with the following functional dependencies:
 $X \rightarrow Y$
 $Y \rightarrow Z$
 $Z \rightarrow W$
 $W \rightarrow X$
 The decomposed Relations are $R_1(X,Y,Z)$ and $R_2(Z,W,V)$. Find if all the original dependencies are preserved or not.

	<p>Steps are required $F1 = X \rightarrow YZ, Y \rightarrow ZY, Z \rightarrow XY$ $F2 = Z \rightarrow W, W \rightarrow Z$ Verify equivalent $(F1 \cup F2) = F$</p> <ul style="list-style-type: none"> • Every functional dependency of $F1 \cup F2$ is a member of F • Every functional dependency of F is a member of $F1 \cup F2$ <p>All the original dependencies are preserved</p>					
5.	<p>Consider the relation $R(P,Q,R,S,T,U,V,W,X,Y)$ with the following set of functional dependencies:</p> <p>$PQ \rightarrow R$ $P \rightarrow ST$ $Q \rightarrow UV$ $U \rightarrow WX$ $S \rightarrow Y$</p> <p>Normalize the schema into Third Normal Form (3NF). Provide the resulting relations and explain the decomposition process for any violations encountered.</p>	3	10	CO2	L3	T
	<p>candidate key PQ</p> <p>Final 3NF Table</p> <p>$R1(P, S, T)$ $R2(Q, U, V)$ $R3(U, W, X)$ $R4(S, Y)$ $R5(PQR)$</p>					
6.	<p>You are the database administrator for an e-commerce platform that processes customer orders and manages payment transactions. When users transfer funds between their accounts, the system ensures that funds are only deducted after the user confirms the transaction. Simultaneously, the platform allows customers to place orders, reflecting inventory changes in real-time to prevent overselling of items. To safeguard against system failures, the application</p>	5	10	CO4	L2	E

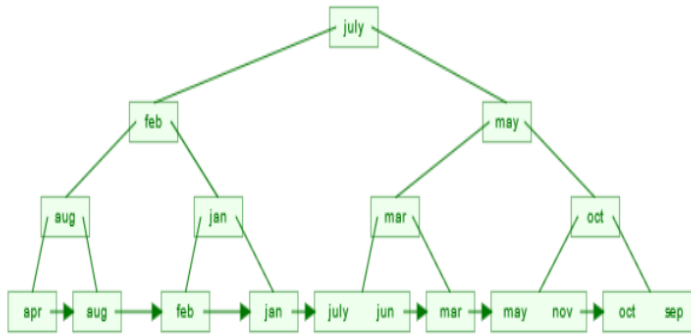
	<p>creates backup copies of database pages for quick recovery. This layered database recovery approach aims to maintain data integrity and reliability while providing a seamless user experience, even in the event of a transaction error or system crash.</p> <p>i. Identify and describe the approach taken for handling fund transfers and explain why this method is suitable for this scenario.</p> <p>ii. Explain the backup strategy implemented for recovering the database. How does this method enhance the recovery process, and what are its limitations?</p>					
	<p>Explain in Detail</p> <p>Deferred Update: In the fund transfer scenario, changes are not applied to the database until the transaction is committed. If the system crashes after deducting Rs1,000 from Account A but before the transaction is committed, the changes will not be saved. Upon recovery, the system will roll back to the last consistent state, meaning Account A retains its original balance, preventing any unintended loss.</p> <p>Backup and Restore: The platform also uses shadow paging to maintain a backup of the database pages. This technique helps ensure that the database can quickly recover to a consistent state in the event of a system failure.</p>					
7.	<p>Identify and justify how the following concurrency control problems affect data integrity and consistency in a multi-user environment.</p> <p>S1: r1(x);r2(x);w1(x);r3(x);w2(x);c2;r4(x);w3(x);c3;w1(y);c1;</p> <p>S2: r1(x);w1(x);r2(x);w2(x);c2;r3(x);c3;w1(y);c1;</p> <p>S3: r1(x); w1(x); r2(x); w2(x); r1(x);</p>	5	10	CO4	L4	A
	<p>Justification needed</p> <p>S1: Lost Update Problem:</p> <p>S2. Dirty Read Problem:</p>					

	S3: Unrepeatable Read Problem:					
8.	Explain the CAP theorem and its implications for distributed system databases. Analyze how the CAP theorem influences the database design choices of an online shopping platform, focusing on the trade-offs between consistency, availability, and partition tolerance.	7	10	CO5	L1	E
	<ul style="list-style-type: none"> Summarize how the CAP theorem informs the design and operational choices of the online shopping platform. Reflect on the practical implications of balancing consistency, availability, and partition tolerance in real-world distributed systems. 					
9.	<p>Show the B+ tree that would be formed by inserting the month names (12 strings) into a tree in their normal calendar sequence (jan, feb, mar, apr, may, jun, jul, aug, sep, oct, nov, dec). After the construction, delete the entries for dec and mar.</p> <p style="text-align: center;">(OR)</p> <p>Insert the elements 15,20,25 and 30 into a hash table of size 10. Assume that the $h1(k) = 2K+3$. The hash value is calculated from the formula $h1(k) \% \text{table size}$. The Tables do not need to be resized. If an element cannot be inserted successfully, state why. You need to display it up to the final table.</p> <ol style="list-style-type: none"> Perform quadratic probing Perform Double Hashing by using the secondary hash function $h2(k) = 3K+1$ 	4	10	CO3	L3	T
	<p>Evaluate this question based on the student's assumption</p> <p>B+ Tree</p>					

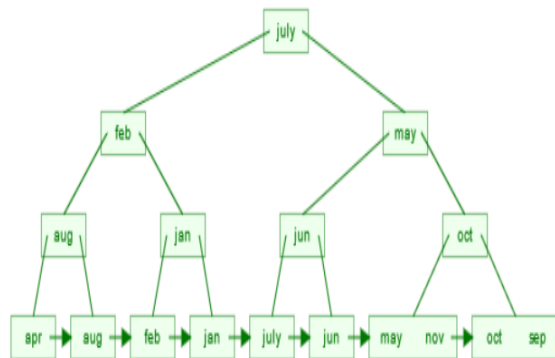
If Assume the order of the B+ Tree 3.



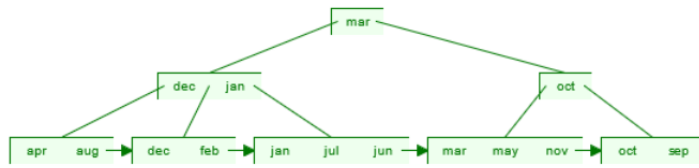
Delete After Dec



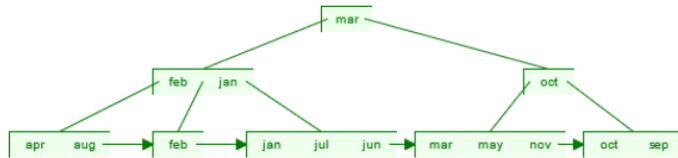
Delete After Mar



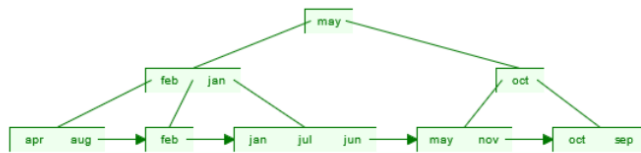
If Assume the order of the B+ Tree 4.



Delete After Dec



Delete After Mar



Quadratic probing

Quadratic probing is a collision resolution strategy where if a collision occurs at position i , the next probe is performed at $i + 1^2$, $i + 2^2$, $i + 3^2$,

Insert 15:

- $h_1(15) = (2(15) + 3) \bmod 10 = (30 + 3) \bmod 10 = 33 \bmod 10 = 3$
- **Insert 15** at index 3.

Insert 20:

- $h_1(20) = (2(20) + 3) \bmod 10 = (40 + 3) \bmod 10 = 43 \bmod 10 = 3$
- **Collision** at index 3 (already occupied by 15).
- Use quadratic probing:
 - $3 + 1^2 = 3 + 1 = 4$.
 - **Insert 20** at index 4

Insert 25:

- $h_1(25) = (2(25)+3) \bmod 10 = (50+3) \bmod 10 = 53 \bmod 10 = 3$
- **Collision** at index 3 (occupied by 15).
- Use quadratic probing:
 - $3 + 1^2 = 3 + 1 = 4$ (occupied by 20).
 - $3 + 2^2 = 3 + 4 = 7$.
 - **Insert 25** at index 7

Insert 30:

- $h_1(30) = (2(30)+3) \bmod 10 = (60+3) \bmod 10 = 63 \bmod 10 = 3$
- **Collision** at index 3 (occupied by 15).
- Use quadratic probing:
 - $3 + 1^2 = 3 + 1 = 4$ (occupied by 20).
 - $3 + 2^2 = 3 + 4 = 7$ (occupied by 25).
 - $3 + 3^2 = 3+9 = 12 \bmod 10 = 2$.
 - **Insert 30** at index 2.

```
Index: 0  1  2  3  4  5  6  7  8  9
Value: -  - 30 15 20 -  - 25 -  -
```

Double hashing

$$h_1(k) = 2K+3$$

$$h_2(k) = 3K+1$$

$$H(k) = h_1(k) + i \cdot h_2(k) \bmod 10 \text{ (table size)}$$

Insert 15:

- $h_1(15) = (2(15)+3) \bmod 10 = 33 \bmod 10 = 3$.
- **Insert 15** at index 3

Insert 20:

- $h_1(20) = (2(20)+3) \bmod 10 = (40+3) \bmod 10 = 43 \bmod 10 = 3$.
- **Collision** at index 3 (occupied by 15).

Use secondary hash function $h_2(20) = (3(20)+1) \bmod 10 = (60+1) \bmod 10 = 61 \bmod 10 = 1$.

$$H(k) = h_1(k) + i \cdot h_2(k) \bmod 10 \text{ (table size)}$$

$$3 + 1 \cdot 1 = 3 + 1 = 4. \quad (i = 1, \text{ next iteration})$$

Insert 20 at index 4.

Continue the process above for the next numbers

The final table after double hashing

```
Index: 0  1  2  3  4  5  6  7  8  9
Value: -  -  - 15 20 30 -  -  - 25
```

10.

i. Discuss different types of Two-Phase Locking (2PL) algorithms with an example.

Consider the following two transactions and schedule (time goes from top to bottom).

T1	T2
Read A	
Write A	
	Read A
	Read B
	Commit
Read B	
Write B	
Commit	

ii. Demonstrate how Two-Phase Locking (2PL) ensures a conflict-serializable schedule for the transactions mentioned above, using the notation $L_i[x]$ to indicate that transaction i acquires a lock on element x and $U_i[x]$ to indicate that transaction i releases its lock on x .

iii. Illustrate how using locks without 2PL can result in a schedule that is not conflict-serializable.

(OR)

Consider the following transactions with their respective timestamps:

T1: Timestamp = 150

T2: Timestamp = 250

T3: Timestamp = 350

Given the schedule of operations:

S: R1(X); R2(Y); W1(Z); R3(Y); W2(X); R1(Y);
W3(Z)

6

10

CO4

L3

T

	<p>i. Apply the Timestamp Ordering Protocol to determine the order of transaction execution.</p> <p>ii. Identify which operations are allowed and which are aborted based on the protocol rules.</p> <p>iii. Explain the reasons for any transaction aborts in the schedule and discuss how this affects the overall consistency of the transactions.</p>																										
	<p>i.</p> <p>Basic 2PL: Locks are acquired during the growing phase and released during the shrinking phase. This approach can lead to deadlocks and does not guarantee that transactions will be serializable.</p> <ul style="list-style-type: none"> • Conservative 2PL: Transactions acquire all necessary locks before starting their operations, ensuring that no deadlocks occur. However, this can lead to reduced concurrency. • Strict 2PL: Locks on all resources are held until the transaction is committed or aborted, ensuring full serializability but potentially leading to lower performance due to increased contention for locks. <p>ii.</p> <table border="1" data-bbox="245 1084 665 2098"> <thead> <tr> <th data-bbox="245 1084 418 1151">T1</th> <th data-bbox="418 1084 665 1151">T2</th> </tr> </thead> <tbody> <tr> <td data-bbox="245 1151 418 1196">L1(A)</td> <td data-bbox="418 1151 665 2098" rowspan="10">L2(A) – Block</td> </tr> <tr> <td data-bbox="245 1196 418 1240">R1(A)</td> </tr> <tr> <td data-bbox="245 1240 418 1285">W1(A)</td> </tr> <tr> <td data-bbox="245 1285 418 1330"></td> </tr> <tr> <td data-bbox="245 1330 418 1375"></td> </tr> <tr> <td data-bbox="245 1375 418 1420"></td> </tr> <tr> <td data-bbox="245 1420 418 1464">L1(B)</td> </tr> <tr> <td data-bbox="245 1464 418 1509">R1(B)</td> </tr> <tr> <td data-bbox="245 1509 418 1554">W1(B)</td> </tr> <tr> <td data-bbox="245 1554 418 1599">U1(A)</td> </tr> <tr> <td data-bbox="245 1599 418 1644">U1(B)</td> <td data-bbox="418 1599 665 2098" rowspan="3">L2[A] – Granted</td> </tr> <tr> <td data-bbox="245 1644 418 1688">C1</td> </tr> <tr> <td data-bbox="245 1688 418 1733"></td> </tr> <tr> <td data-bbox="245 1733 418 1778"></td> <td data-bbox="418 1733 665 1778">R2(A)</td> </tr> <tr> <td data-bbox="245 1778 418 1823"></td> <td data-bbox="418 1778 665 1823">L2(B)</td> </tr> </tbody> </table>	T1	T2	L1(A)	L2(A) – Block	R1(A)	W1(A)				L1(B)	R1(B)	W1(B)	U1(A)	U1(B)	L2[A] – Granted	C1			R2(A)		L2(B)					
T1	T2																										
L1(A)	L2(A) – Block																										
R1(A)																											
W1(A)																											
L1(B)																											
R1(B)																											
W1(B)																											
U1(A)																											
U1(B)	L2[A] – Granted																										
C1																											
	R2(A)																										
	L2(B)																										

R2(B)

U2(A)

U2(B)

C2

FINAL ORDER

T 1-> T2

iii.

T1	T2
L1(A)	
R1(A)	
W1(A)	
U1(A)	
	L2(A)
	R2(A)
	U2(A)
	L2(B)
	R2(B)
	U2(B)
	C2
L1(B)	
R1(B)	
W1(B)	
U1(B)	
C1	

Final order

T1 -> T2

i)

Final Order of Execution

- The overall order reflecting successful execution is T1 → T2 → T3

ii)

The allowed operations based on the Timestamp Ordering Protocol are:

1. **T1**: R1(X)
2. **T2**: R2(Y)
3. **T1**: W1(Z)
4. **T3**: R3(Y)
5. **T1**: R1(Y)

Aborted Operations:

- **T2**: W2(X) (T2 is aborted because it cannot write to X after T1 has read it).
- **T3**: W3(Z) (T3 is aborted because it cannot write to Z after T1 has written it).

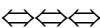
iii. Reasons for Transaction Aborts and Impact on Consistency

1. Transaction T2 Aborted (W2(X)):

- **Reason**: **T2** attempts to write to **X**, but **T1** has already read **X** earlier. According to the **Timestamp Ordering Protocol**, **T2** cannot write to **X** since **T1**'s read was earlier, and the transaction order is violated.
- **Impact on Consistency**: Aborting **T2** ensures that the read-write order is respected, maintaining consistency. The data for **X** will be consistent because **T1** has already read it before **T2** attempted to write.

2. Transaction T3 Aborted (W3(Z)):

- **Reason**: **T3** tries to write to **Z**, but **T1** has already written to **Z**. According to the **Timestamp Ordering Protocol**, **T3**'s write operation conflicts with **T1**'s earlier write, and **T3** cannot proceed.
- **Impact on Consistency**: Aborting **T3** ensures that **Z** is not overwritten in an inconsistent order. **T3** would have overwritten **Z** after **T1** had already written it, which would violate consistency.

	<p>The final execution reflects only the successful operations while noting that T2 and T3 had to be aborted due to conflicts with T1 based on the Timestamp Ordering Protocol. This ensures that the schedule remains serializable, preserving data consistency.</p> <ul style="list-style-type: none"> The overall order reflecting successful execution is T1 → T2 → T3 					
						

In the above 10 Questions, any 2 questions should have “**Either- Or**” choice. However, the questions can be from the same or different modules. Ensure the same CO is maintained for both questions. Also, ensure the same difficulty level (both the questions should be Average (‘A’) or both should be Tough (‘T’)) is maintained.

Guidelines for the preparation of the question paper

- Module No.: Type the module number as given in the syllabus, like 1, 2, 3, 4, 5.
- Difficulty Level: Easiness of the question in the grade Easy, Average, Tough (HOT), or E/A/T. The recommended proportion of the questions should be 25%, 50%, and 25%, respectively, to ensure a normal distribution.
- Fill out the appropriate BL: **Bloom’s Taxonomy Levels (1: Remembering, 2: Understanding, 3: Applying, 4: Analysing, 5: Evaluating, 6: Creating)**
- Fill out the appropriate COs (course outcomes) for each question, as it is vital for further mark analysis. CO: Course Outcomes (refer to the syllabus for course outcomes 1–6)
- The question paper should provide 20% of the Max Marks as choice for the students. This is mandatory for all UG and PG level courses.
- Under each question, if required, sub-sections can be included.
- Ensure that you have followed the approved syllabus.
- Set the question paper for the full syllabus, covering all the modules uniformly.
- Use the given template for the preparation of the question paper.
- General Instructions indicate the instructions for the whole of the question paper.
- Specific instructions can also be given for each section.
- Give the question number continuously, starting with '1', and do not break continuity when going to the next section.
- A detailed scheme of evaluation and an explicit key are mandatory along with the question paper. Without the key, the question paper will not be accepted. The same will be made accessible to students immediately after the completion of the examination for reference and possible re-evaluation.
- If your relative is appearing for this examination, please decline the QP setter appointment offer.
- All paper setters are requested to keep their appointments strictly confidential. including your passwords
- The question paper should be such that a candidate of decided ability and good preparation in that course can reasonably be expected to answer completely within the allotted time.
- Ensure that the questions are not used in previous examinations or anywhere else, like a model question paper.

- Ensure that the questions are balanced in terms of syllabus coverage, marks, and time taken to answer each question.
- The detailed answers for all the questions are to be given in the key (scheme of valuation) in softcopy, either as Word, pdf, or scanned images (TIF, JPG, or PNG).
- Follow the time schedule since many follow-up activities like moderation, formatting, and printing out the question paper rely on your punctuality.



Signature with date