Birla Institute of Technology & Science, Pilani
Second Semester 2022-2023
Database Systems (CS F212)
Comprehensive Examination -- Part A [Closed Book]

Maximum marks 20 Maximum Time Allowed : 30 Minutes

Name	
BITS ID	
Invigilator Sign	

Note: Write the answers in the space provided ONLY. You can do the rough work on the last page of the main answer sheet. Overwriting and writing anywhere except space provided will not be evaluated.

Ques 1) Compare the time-stamping protocol and the two-phase locking protocol in terms of their approach to Conflict serializatility, deadlock freedom, View serializatility, Recoverability and Prevention from Cascading rollback. Write **Ensured** or **NOT Ensured** only in the respective cells of table below. *[No explanation is required.]* ***[6 Marks]***

Assuming If the schedules allowed from time stamp protocol are all conflict serializable write Ensured otherwise write NOT Ensured. If the scheduled allowed from two-phase locking protocol, ensure deadlock freedom write Ensured otherwise write NOT Ensured. And so on for every cell in table.

	Conflict serializatility	Deadlock handling	View serializatility	Recoverability	Prevention from Cascading rollback
Time-stamping protocol					
Rigorous two-phase locking protocol					
Simple two- phase locking protocol					

Ques 2) Consider a database table named "Projects" with the following attributes: *Employee_ID (EID), Project_ID (PID), Role, and Skill*. The following functional dependencies are given:

1. EID, PID → Role
2. EID, Skill → Role

Additionally, the following multi-valued dependencies are given:

1) EID →→ PID
2) EID →→ Skill

a) Calculate the closure of each attribute (EID, PID, Role, and Skill) based on the given functional dependencies. *No explanation is required.* ***[2 marks]***

Attribute	Closure
EID	
PID	
Role	
Skill	

b) Determine all the candidate key(s) for the table Projects. *[Use comma separator, with No explanation.]* **[2 Marks]**

Candidate Keys	

c) Decompose the given table into 4NF. Write decomposed table and their attributes, as well as any primary and foreign key(s). *[It's not necessary that there are 5 different Relations in final answer, Use the number of the number of correct ones and leave the Rest Rows blank.]* **[4 Marks]**

	Decomposed 4NF Relation	Its Candidate Key(s)	Its Foreign Key(s)
R1			
R2			
R3			
R4			
R5			

Ques 3) Given the following SQL query, rewrite it using Relational Algebra notation:
*Assume the **Customers** table has the attributes **customer_id** and **customer_name** and the **Orders** table has the attributes **order_id**, **customer_id** and **order_date**.* **[2 Marks]**

```
SELECT c.customer_name, o.order_id
FROM customers c
JOIN orders o ON c.customer_id = o.customer_id
WHERE o.order_date >= '2022-01-01';
```

Ques 4) Which of the following is/are true? Tick against that. *[All correct answers marked and no wrong options marked will fetch marks]* **[4 Marks]**

More and more latest developed applications prefer NoSQL databases (Column databases/ Graph databases/ Key-value stores/ Document databases) and not relational database schema because.

☐ A. NoSQL databases allows Horizontal scaling and Sharding.
☐ B. Schema-less storage allows faster storage, flexibility and easier retrieval.
☐ C. NoSQL allows more specific API designed for the data models used
☐ D. NoSQL databases often trade-off consistency for availability and partition tolerance, adhering to the CAP theorem, hence more transactions are allowed that are eventually consistent.

Birla Institute of Technology & Science, Pilani
Second Semester 2022-2023
Database Systems (CS F212)
Comprehensive Examination -- Part B [Open Book] Maximum marks: 85

Note: After completing each question, place the final answer in a separate box located at the end of that particular question.
When answering a question with multiple parts, make sure to address all parts of the question together in one cohesive response. Do not attempt individual parts separately at different times, as any parts submitted independently will not be considered for evaluation.

Ques 1) Answer the following questions. *[All three parts are independent of each other]* [5+5 = 10 marks]

a) Consider the following schedule S of transactions T1, T2, T3, and T4:

S: R1(A), R2(A), W1(A), R3(B), R4(B), W2(A), W3(B), R1(B), W1(B), R3(C), W3(C), R4(C), W4(C), R2(B), W2(B).

Create a precedence graph for the given schedule S and Determine if the given schedule S is **conflict serializable**. If it is, find an equivalent serial schedule. If not, explain why it is not conflict serializable.

b) On a *different* schedule S of transactions T1, T2, and T3:

S: R1(A), W1(A), R2(A), W2(A), R1(B), W1(B), R3(B), W3(B), R2(C), W2(C), R3(C), W3(C)

Determine if the given schedule S is **view serializable**. If it is, provide an equivalent view serial schedule and explain your reasoning. Otherwise, explain why not.

Ques 2) Answer the following questions. *[All three parts are independent of each other]* [5+5+5 = 15 marks]

a) Consider the following schedule S of transactions T1, T2, and T3:

S: R1(A), W1(A), R2(A), W2(A), R3(B), W3(B), R1(B), W1(B), Commit T1, R2(B), W2(B), R3(C), W3(C), Commit T3, Commit T2.

Determine if the given schedule S is **recoverable**. If it is not, explain why.

b) Consider the following schedule S of transactions T1, T2, and T3:

S: R1(A), W1(A), R2(B), W2(B), R2(A), W2(A), R1(B), W1(B), R3(C), W3(C), R3(B), W3(B), Commit T3, Commit T1, Commit T2.

Determine if the given schedule S is **recoverable.** Also determine if the given schedule S is **cascadeless**. Give proper justification.

c) Consider two schedules SA and SB of transactions T1, T2, and T3:

SA: R1(A), W1(A), R2(A), R1(B), W1(B), Commit T1, W2(A), R2(B), W2(B), commit T2, R3(A), R3(B), W3(A), W3(B), Commit T3.

SB: R1(A), W1(A), R2(B), W2(B), R1(B), W1(B), R2(A), W2(A), R3(B), Commit T1, W3(B), R3(A), W3(A), commit T2, Commit T3.

Determine if the schedules SA and SB are **recoverable and cascadeless**. Give proper justification.

Ques 3) Answer the following questions. *[Part b) is related to part a)]*

a) Consider the following schedule S of transactions T1, T2, and T3 operating on data items A, B, and C:

S: R1(A), R2(A), W1(A), R2(B), W2(A), R3(C), W2(B), W1(B), R1(C), W1(C), R3(A), W3(C)

Apply the **simple 2PL** protocol to the given schedule S and list all the lock/unlock operations required for each transaction to maintain concurrency control. Identify if any deadlocks occur while applying the simple 2PL protocol. **Show diagrammatically**.

b) Apply the **strict and conservative 2PL** variants on the schedule given in part a). Show diagrammatically. [5+5 = 10 marks]

Ques 4) Assume the following schedule S of transactions T1, T2, and T3 operating on data items A, B, and C:

S: R1(A), R2(A), W1(A), R2(B), W2(A), R3(C), W2(B), W1(B), R1(C), W1(C), R3(A), W3(C)

Assume the transactions follow a time-stamping protocol.

Apply the basic **time-stamping protocol** to the given schedule S, providing the read and write time-stamps for each transaction and data item. Identify any conflicts and specify how they are resolved.

Assume that the schedule started at time 10 and every read/write operation takes one-time unit to execute. E.g. In the above schedule R1(A) is executed at time 10, R2(A) is executed at time 11, W1(A) is executed at time 12, and so on. [5 marks]

Ques 5) A database system uses a log-based crash recovery protocol. Transactions T1, T2, T3, and T4 operate on data items A, B, and C, and their operations are logged as follows:

L: [START T1], [START T2], [T1, A, 10, 20], [T2, B, 15, 25], [T1, C, 30, 40], [COMMIT T1], [START T3], [T2, A, 20, 30], [T3, C, 40, 50], [COMMIT T2], [T4, A, 30, 35], [START CKPT (T3, T4)], [END CKPT], [COMMIT T3], POINT Z

Assume a crash occurs at POINT Z. *[CKPT means checkpoint]*

a) Perform the crash recovery process for this log, and indicate the final state of data items A, B, and C after recovery. Explain the output of "undo" and "redo" buckets/operations during the crash recovery process.

b) A database system uses a Write-Ahead Logging (WAL) protocol for crash recovery. Transactions T1, T2, and T3 operate on data items A, B, and C, and their operations are logged as follows:

L: [START T1], [START T2], [T1, A, 10, 20], [T2, B, 15, 25], [T1, C, 30, 40], [COMMIT T1], [T2, A, 20, 30], [START T3], [T3, C, 40, 50], [COMMIT T2], [COMMIT T3] – CRASH--

Assume the system crashes before any of the changes are written to the disk or we are not sure about all the changes written to the disk. Perform the necessary crash recovery steps using the log L, and indicate the final state of data items A, B, and C after recovery. [5+5 = 10 marks]

Ques 6) Answer the following questions. [All three parts are independent of each other]

a) Consider a database table named "Students" with the following attributes: Student_ID (SID), Name, and Age. The Student_ID (SID) is unique for each student record. Assume that there

are 10,000 records in the table and that each file block can hold 100 records. A primary index is created on Student_ID.

Calculate the total number of file blocks required to store the entire Students table. Determine the maximum and minimum number of disk block accesses required for a binary search on the primary index when searching for a specific Student_ID. If the primary index is stored using a B+-tree index with a maximum node degree of 3, calculate the maximum height of the B+-tree index.

b) Consider a database table named "Products" with the following attributes: Product_ID (PID), Name, Price, and Description. The Product_ID (PID) is unique for each product record. A B+-tree index with a maximum node degree of 3 is created for the Product_ID attribute based on the following instructions. In order to create B+-tree index, the following Product_IDs are added in sequence: 1, 3, 5, 8, 12, 15, and 18. Show the final state of the B+-tree index and all intermediate states. Calculate the number of disk block accesses required to search for a specific Product_ID using the B+-tree index. Assume that a new product with Product_ID 9 is added to the table. Show how the B+-tree index is updated, and indicate the final state of the index. [5+5 = 10 marks]

Ques 7) Consider a database table named "Courses" with the following attributes: Course_ID (CID), Instructor, Semester, and Room. The following functional dependencies are given:

1) **CID → Instructor**
2) **CID, Semester → Room**
3) **Semester, Room → CID**

a) Calculate the closure of each attribute (CID, Instructor, Semester, and Room) based on the given functional dependencies.

b) Determine all the candidate key(s) for the table. Explain your reasoning.

c) Identify the highest normal form (1NF, 2NF, 3NF, or BCNF) that the given table satisfies. Justify your answer. [5+5+5 = 15 marks]

Ques 8) Consider a database table named "Students" with the following attributes: Student_ID (SID), Name, Course_ID (CID), Grade, and Hobby. The following functional dependencies are given:

1) **SID → Name**
2) **SID, CID → Grade**

Additionally, the following multi-valued dependencies are given:

1) **SID →→ CID** *[Multivalued dependency]*
2) **SID →→Hobby** *[Multivalued dependency]*

a) Decompose the given table into 4NF. Show each decomposed table and their attributes.

b) For each decomposed table, identify the candidate key(s) and any foreign key(s). [5+5 = 10 marks]