

# **Exam Preparation**

## **Unit 1: Introduction**

[ 4 hrs ]

1. Database Management System purpose and applications,
2. Database Systems vs File Systems
3. View of Data- Data Abstraction (Physical, logical and view level, Data Independence)
4. Instances and Schemas,
5. Database Languages (DDL, DML and DCL)
6. Database and Application Architecture- Database System Architecture and Database Application Architecture (two-tier and three-tier)

**[1 question 7 or 8 marks]**

## **Important Topics/Questions**

1. Database definition, objectives, application areas, pros, cons, purpose.
2. Explain the importance of DBMS. How do DBMs help in the field of IT? Also explain how DBMS is accessed using various DDL, DML, and DCL languages.
3. Differences between DDL, DML, and DCL languages
4. Differentiate between the File system and DBMS. (Modern database vs traditional)
5. Schema vs instances.
6. Explain different levels of schema.
7. Why do we need multilayer architecture in DBMS? Describe the 3 schema architectures.
8. Describe the 3 levels of DBMS (physical, logical, and external)
9. Difference between Physical and Logical Data Independence. Any 5 points. Also, list the major steps you would take to set up a database for a particular enterprise.
10. Explain Database System Architecture and Database Application Architecture with examples and differences.
11. What are the major responsibilities of DBMS? For each responsibility explain the problem that would arise if the responsibility were not discharged.
12. What is data abstraction? Explain different levels of abstraction and also mention how it is different from data independence.

**Unit 2: ER and Relational Model [8 hours]**

1. Introduction to ER Model:
2. Entity sets, attributes and values, Relationship sets-participation, entity's role, descriptive attributes, degree of relationship set, Mapping Cardinalities, Attributes- simple, composite, single-valued, multi-valued, derived, Entity-Relationship (ER) Diagram, Specialization, Generalization, and Aggregation
3. Key- Super key, Candidate key and Primary key, Strong and Weak Entity Sets
4. Introduction to Relational Model
5. Reducing ER diagrams to Relational Schema
6. Structure of Relational Databases, Database Schema, Keys, Schema Diagrams
7. Relational Algebra

**[2 or 3 questions 15 to 22 marks]**

**Important Topics/Questions**

1. Draw an ER diagram or Schema diagram (it can be any diagram). Uses of Er diagram. Explain ER design issues. What are the Uses of ERD?
2. What are the steps to reduce the ER diagram into a schema diagram? Explain
3. Explain 4 different kinds of data models and also list out the Merits and demerits of the Data model. Also, mention differences between Hierarchical, ER model, and relational models.
4. One RA Question (it can be also any question)
5. Differentiate between Strong and weak entities.
6. Differentiate between Generalization and specialization.
7. How Aggregation and Inheritance helps in ERD.
8. What is Key? Why do we need a key in DBMS? Explain different types of Keys. Also mention differences between Candidate, Primary, and Foreign Keys.
9. What are the extended features of the ER diagram? How Aggregation and Inheritance Help in ERD.

10. Explain different set operations Used in RA. What are the rules for using Union and Intersection?
11. What is union compatibility (type compatibility)? Explain with an example.
12. Explain different types of Joins. Differentiate between Inner, Outer, and Natural join.

**Unit 3: Structured Query Language [5 hours]**

1. Structured Query Language (SQL)- SQL DDL and DML,
2. Basic Structure of SQL Queries, DDL queries, Basic Operations (Rename, String, Attribute Specification in the select clause, order by, where-clause), Set Operations, Null values, Aggregate Functions,
3. Nested Queries,
4. Join Expressions (Natural Join, Join Conditions, Outer Joins),
5. Modification of Database (delete, insert into, update),
6. Views, Stored Procedures

**[1 or 2 questions 7 or 15 marks]**

**Important Topics/Questions**

1. SQL-related Question (It can be any Question)
2. What is view? How to create a view and update it.
3. Stored procedure
4. Different types of joins in SQL.

**Unit 4: Relational Database Design [ 6 hrs ]**

1. Integrity constraints- Domain Constraints, Entity Integrity Constraints, Referential Integrity Constraints, Assertions and Triggers
2. Features of Good Relational Designs
3. Functional dependencies and Armstrong's Axioms
4. Closure of a Set of Functional Dependencies and Closure of Attribute Sets
5. Database Normalization and Normal Forms- 1NF, 2NF, 3NF and BCNF
6. Denormalization for Performance

**[2 or 3 questions 15 to 22 marks]**  
**Important Topics/Questions**

1. What are constraints? Explain different types of Constraints. Also, mention the difference between entity, referential, domain, and key constraints.
2. What is functional dependency? Explain different types of FD.
3. What are Armstrong Axioms? Explain the different rules of it.
4. What are the features of good relational design?
5. What is closure? How to find closure of any FD? Explain with examples.
6. What is normalization and denormalization? Explain different types of Normalization.
7. What are anomalies in the Database Explain their types and how to resolve them
8. Differences between 1 NF, 2NF, 3NF and BCNF.
9. Short notes on Assertions and Triggers.
10. Any Normalization (Numerical) Related Question (From class and Assignment)

**Unit 5: Security** [ 3

hrs ]

1. Security and integrity violations
2. Access control
3. Authorization
4. Security and Views

**[1 question 7 or 8 marks]**  
**Important Topics/Questions**

1. What is security? How it can be violated? Explain different security levels.
2. Needs for database security.
3. What is Authentication and Authorization? Mention the difference between them.
4. What are the different Authentication techniques explain.
5. What is the use of Grant and Revoke command in Authorization?

6. What are access control and access protection? Explain its Types(MAC and DAC) with their difference.
7. What is encryption and decryption?
8. Describe GRANT and REVOKE functions and explain how it relate to security. What types of privileges may be granted? How rights could be revoked?
9. What is cryptography? Explain its types.
10. Why is database security important?

---

## **Unit 6 : Query Processing and Optimization**

[ 5

hrs ]

1. Introduction to Query Processing
2. Equivalence of Expressions
3. Query Cost Estimation
4. Query Optimization
  
5. Query evaluation and execution plan

**[1 or 2 questions 7 to 12 marks]**

### **Important Topics/Questions**

1. Explain the basic steps in query processing with a diagram in RDBMS. What is pipelining in query evaluation? Explain with an example.
2. What is the task of the evaluation engine in the query processing? Explain cost-based query optimization.
3. What is Equivalent expressions?
4. How are equivalence rules for relational algebra helpful for query optimization? Explain with an example.
5. Differentiate between the pipelining approach and materialization approach for query optimization in detail.
6. How do you evaluate the performance of a magnetic disk? What are the optimization techniques to reduce the disk block access?
7. What is a Cost estimation, Query optimization, and query tree? Give 3 different examples of query tree which consists of the join of 3 different tables.
8. Prove the given equivalence rules with examples.

**equivalents**

- 1 Conjunctive selection operations can be deconstructed into a sequence of individual selections.  

$$\sigma_{\theta_1 \wedge \theta_2}(E) = \sigma_{\theta_1}(\sigma_{\theta_2}(E))$$
- 2 Selection operations are commutative.  

$$\sigma_{\theta_1}(\sigma_{\theta_2}(E)) = \sigma_{\theta_2}(\sigma_{\theta_1}(E))$$
- 3 Only the last in a sequence of projection operations is needed, the others can be omitted.  

$$\Pi_{i_k}(\Pi_{i_{k-1}}(\dots(\Pi_{i_1}(E))\dots)) = \Pi_{i_k}(E)$$
- 4 Selections can be combined with Cartesian products and theta joins.
  - $\sigma_\theta(E_1 \times E_2) = E_1 \bowtie_0 E_2$
  - $\sigma_{\theta_1}(E_1 \bowtie_{\theta_2} E_2) = E_1 \bowtie_{\theta_1 \wedge \theta_2} E_2$
- 5 Theta-join operations (and natural joins) are commutative.  

$$E_1 \bowtie_0 E_2 = E_2 \bowtie_0 E_1$$
- 6 a) Natural join operations are associative:  

$$(E_1 \bowtie E_2) \bowtie E_3 = E_1 \bowtie (E_2 \bowtie E_3)$$
b) Theta joins are associative in the following manner:  

$$(E_1 \bowtie_{\theta_1} E_2) \bowtie_{\theta_2 \wedge \theta_3} E_3 = E_1 \bowtie_{\theta_2 \wedge \theta_3} (E_2 \bowtie_{\theta_2} E_3)$$
where  $\theta_2$  involve attributes from only  $E_2$  and  $E_3$ .
- 7 The selection operation distributes over the theta join operation under the following two conditions:
  - (a) When all the attributes in  $\theta_0$  involve only the attributes of one of the expressions ( $E_i$ ) being joined.  

$$\sigma_{\theta_0}(E_1 \bowtie_0 E_2) = (\sigma_{\theta_0}(E_1)) \bowtie_0 E_2$$
  - (b) When  $\theta_1$  involves only the attributes of  $E_1$  and  $\theta_2$  involves only the attributes of  $E_2$ .  

$$\sigma_{\theta_1 \wedge \theta_2}(E_1 \bowtie_0 E_2) = (\sigma_{\theta_1}(E_1)) \bowtie_0 (\sigma_{\theta_2}(E_2))$$

**Unit 7 : Storage Management and Indexing [ 3**

hrs ]

1. File Organization- fixed length records, variable length records
2. Organization of Records in File- Heap, Sequential and Indexed sequential file organizations.
3. B+ Tree Index Files
4. Hash Indices

**[1 or 2 questions 7 to 12 marks]****Important Topics/Questions**

1. What is file organization Explain how you organize files using a hash index.
2. Differences between Heap, Sequential, and indexed sequential file organizations.
3. What are the different Organization of Records in the file?
4. Differences between fixed and variable length records. Explain the slotted page structure in detail.
5. Explain static and dynamic hashing. Also, mention 5 differences between them.
6. Explain different collision-solving techniques of hashing.
7. Explain different storage media used for file organization.

8. Differences between Sparse and dense index. Any 5
9. Differences between Primary and secondary index. Any 5.
- 10.1 B+ tree related numerical. (follow class examples)
11. Hash index.

---

## **Unit 8 : Transactions and Concurrency Control [ 4 hrs ]**

1. Transaction Concepts
2. Transaction Model and State Diagram
3. ACID properties of transaction
4. Serializability- conflict and view serializability
5. SQL Standard Isolation Levels
6. Concurrency Control- Lock-Based Protocols and Graph-Based Protocols

**[1 or 2 questions 7 to 12 marks]**

### **Important Topics/Questions**

1. What do you mean by transaction Processing System? Explain with an example. Explain transaction management in the database.
2. “Concurrent execution of transactions is more important when data are in memory and transaction are very short”. Justify this statement.
3. What are the states of the Transaction?
4. Explain the properties of the Transaction. (ACID)
5. Differences between serial and serializable schedules.
6. Differences between Conflict and view serializable.
7. Since every conflict-serializable schedule is view serializable, why do we emphasize conflict serializability rather than view serializable?
8. How schedules can be tested for serializability? Explain with examples.  
What are the different techniques used for testing serializability?
9. Explain different types of concurrency Problems. Why we need concurrency control schemes explain different types of methods.
10. What is a lost update problem how to prevent it?
11. Explain Graph-based protocol in detail with an example. Also, mention lock-based vs Graph-based protocols.
12. What are the differences between 2PL and strict 2PL techniques

13. Any numerical related to serializability (From class and assignment examples)
14. How schedules can be tested for serializability? Explain with examples.  
What are the different techniques used for testing serializability

### **Unit 9 : Crash Recovery**

[ 3 ]

hrs ]

1. Failure classification
2. Recovery and Atomicity- log records, database modification, concurrency control and recovery, transaction commit, Redo and Undo Transactions using Log, Check Points,
3. Recovery Algorithm Using Log Records- Transaction Rollback, Recovery After a System Crash, Optimizing Commit Processing
4. High Availability Using Remote Backup System

**[1 question 7 or 8 marks]**

### **Important Topics/Questions**

1. What is a crash in dbms? Explain different kinds of failures that occur in transactions
2. Explain log-based recovery and its advantages and disadvantages with an example also mention the differences between Deferred and Immediate database modification.
3. Differences between shadow paging and log-based recovery.
4. Explain shadow paging with its advantages and disadvantages.
5. How do you recover when a failure results in loss of data from the nonvolatile storage?
6. Explain the purpose of the checkpoint mechanism. What is done during checkpoint?
7. Remote Backup System, Undo redo process,
8. For the given example write a log record for both Immediate and Deferred database modification.

$\langle T_0 \text{ start} \rangle$	$\langle T_0 \text{ start} \rangle$	$\langle T_0 \text{ start} \rangle$
$\langle T_0, A, 1000, 950 \rangle$	$\langle T_0, A, 1000, 950 \rangle$	$\langle T_0, A, 1000, 950 \rangle$
$\langle T_0, B, 2000, 2050 \rangle$	$\langle T_0, B, 2000, 2050 \rangle$	$\langle T_0, B, 2000, 2050 \rangle$

$\langle T_0 \text{ commit} \rangle$	$\langle T_0 \text{ commit} \rangle$	$\langle T_0 \text{ commit} \rangle$
$\langle T_1 \text{ start} \rangle$	$\langle T_1 \text{ start} \rangle$	$\langle T_1 \text{ start} \rangle$
$\langle T_1, C, 700, 600 \rangle$	$\langle T_1, C, 700, 600 \rangle$	$\langle T_1, C, 700, 600 \rangle$

(a)

(b)

(c)

$\langle T_0 \text{ start} \rangle$	$\langle T_0 \text{ start} \rangle$	$\langle T_0 \text{ start} \rangle$
$\langle T_0, A, 950 \rangle$	$\langle T_0, A, 950 \rangle$	$\langle T_0, A, 950 \rangle$
$\langle T_0, B, 2050 \rangle$	$\langle T_0, B, 2050 \rangle$	$\langle T_0, B, 2050 \rangle$
	$\langle T_0 \text{ commit} \rangle$	$\langle T_0 \text{ commit} \rangle$
	$\langle T_1 \text{ start} \rangle$	$\langle T_1 \text{ start} \rangle$
	$\langle T_1, C, 600 \rangle$	$\langle T_1, C, 600 \rangle$
		$\langle T_1 \text{ commit} \rangle$

(a)

(b)

(c)

**Unit 10 : Emerging Trend in Databases**

[ 4

hrs ]

1. NoSQL Databases- Characteristics, Categories, Advantages
2. Object Oriented Database and ORM
3. Distributed Databases-
4. Distributed Ledged Technology: Blockchain, Cryptocurrency, Blockchain Properties

**[1 or 2 questions 7 to 12 marks]****Important Topics/Questions**

1. A healthcare organization manages patient data stored across different locations. Analyze how distributed databases can ensure data integrity and security while providing fast access to critical information.
2. Explain the concept of Distributed Databases. Mention the differences between Distributed and Centralized databases with their pros, cons, and applications.
3. What are NoSQL Databases explain its characteristics. How it differs from SQL. Explain its- Characteristics, Categories, and Advantages. Also, explain any 2 NoSQL database systems.
4. What are object-oriented database models? Explain the advantages and disadvantages of object-oriented databases over relational databases.
5. Short notes on Blockchain, Cryptocurrency, ORM
6. Compare Blockchain with a relational database. Is it possible to modify the data once it is written in a block?
7. Can you explain what ORM Associations are and outline their significance in database management?