SYSTEM MCQs COLLECTION

NSCT Prep

Free MCQ Practice for NSCT Test Preparation

Databases

1530 Multiple Choice Questions

nsctprep.dev

This dataset is created and compiled by Muhammad Abdullah Awais

Easy Questions

510 questions

Q1. What is a database?

A. A specialized type of programming language
B. An organized collection of structured data
C. A specific type of network protocol used
D. A physical hardware component in a system

Answer: B

Q2. Which of the following is a popular Database Management System (DBMS)?

A. CSS
B. HTML
C. MySQL
D. FTP

Answer: C

Q3. What does DBMS stand for?

A. Data Backup Management System
B. Database Management System
C. Database Modification Software
D. Digital Base Management System

Answer: B

Q4. Which of these is an advantage of using a database system?

A. Data cannot be shared between users
B. Data redundancy is significantly reduced
C. Overall storage cost increases greatly
D. Programs become much harder to write

Answer: B

Q5. What is data redundancy?

A. Algorithmic compression of data
B. Unnecessary duplication of data
C. Permanent loss of stored data
D. Cryptographic encryption of data

Answer: B

Q6. Who manages and controls access to a database?

A. Business Systems Analyst
B. Database Administrator (DBA)
C. Network Systems Engineer
D. Web Interface Designer

Answer: B

Q7. Which of the following is NOT a type of database user?

A. DBA administrator
B. Database end user
C. SQL programmer
D. Compiler program

Answer: D

Q8. What is the primary goal of a DBMS?

A. To manage and configure physical hardware components in a server rack
B. To provide a convenient and efficient way to store and retrieve data
C. To design and develop interactive websites and web applications online
D. To replace the functionality of the underlying operating system entirely

Answer: B

Q9. In a file-based system, data is stored in:

A. Random access memory storage only
B. Separate files for each application
C. Cloud-based storage systems only
D. A single centralized data repository

Answer: B

Q10. Data independence means:

A. Changes to data storage do not affect application programs
B. Applications and data definitions are treated as the same
C. Stored data values cannot be modified by any user at all
D. Data items are not stored permanently on the storage disk

Answer: A

Q11. The two-tier architecture in databases consists of:

A. Three layer tiers
B. Client and Server
C. Only Client tier
D. Only Server tier

Answer: B

Q12. In a client-server architecture, the database resides on the:

A. Client side
B. Neither side
C. Both equally
D. Server side

Answer: D

Q13. What is a centralized database system?

A. A system designed to operate without any DBMS software layer
B. A system that runs exclusively on cloud computing platforms
C. A system with data distributed across multiple remote locations
D. A system where data is stored and managed at a single location

Answer: D

Q14. Which tier in three-tier architecture handles the user interface?

A. Presentation tier
B. Database data tier
C. None of these tiers
D. Business logic tier

Answer: A

Q15. What does a query processor do?

A. Backs up data to external storage
B. Interprets and executes database queries
C. Stores data on physical disk drives
D. Manages user accounts and profiles

Answer: B

Q16. The storage manager in a DBMS is responsible for:

A. Creating graphical user interfaces
B. Managing how data is stored on disk
C. Writing and executing SQL queries
D. Configuring network communication

Answer: B

Q17. Which component of DBMS translates high-level queries into low-level operations?

A. Transaction manager
B. Query compiler module
C. Buffer manager module
D. File manager module

Answer: B

Q18. In a three-tier architecture, the middle tier is called:

A. Data warehouse store
B. Application server
C. Database server
D. Web browser client

Answer: B

Q19. What is the role of the buffer manager?

A. Creates and maintains database schemas and table structures
B. Manages the transfer of data between disk and main memory
C. Manages user authentication and role-based access control
D. Processes and optimizes SQL queries for efficient execution

Answer: B

Q20. A database engine is also known as:

A. Query language
B. Storage engine
C. Front-end application
D. Report generator

Answer: B

Q21. Which data model organizes data in tables of rows and columns?

A. Relational model
B. Object model
C. Hierarchical model
D. Network model

Answer: A

Q22. In the hierarchical data model, data is organized as a:

A. Graph structure
B. Star structure
C. Tree structure
D. Table structure

Answer: C

Q23. An Entity-Relationship (ER) diagram is used to:

A. Design network topologies for servers
B. Create application interfaces for users
C. Model the conceptual design of a database
D. Write SQL queries against stored tables

Answer: C

Q24. In an ER diagram, a rectangle represents:

A. An entity type
B. A relationship
C. An attribute
D. A primary key

Answer: A

Q25. An attribute in a data model represents:

A. A constraint on records
B. A property of an entity
C. A table in the database
D. A relationship between entities

Answer: B

Q26. Which symbol represents a relationship in an ER diagram?

A. Ellipse
B. Rectangle
C. Triangle
D. Diamond

Answer: D

Q27. A primary key uniquely identifies:

A. A relationship instance
B. A single column value
C. Each record in a table
D. An entire database schema

Answer: C

Q28. In the network data model, the relationship between records is:

A. Many-to-many using sets
B. No relationships allowed
C. Hierarchical only structure
D. One-to-one relationships

Answer: A

Q29. The relational data model was proposed by:

A. Edgar Dijkstra
B. Charles Bachman
C. E.F. Codd
D. Peter Chen

Answer: C

Q30. Which of these is NOT a data model?

A. Relational model
B. Hierarchical model
C. Compiler model
D. Network model

Answer: C

Q31. In a relational database, a table is also called a:

A. Tree
B. Graph
C. Document
D. Relation

Answer: D

Q32. A row in a relational table is called a:

A. Schema
B. Attribute
C. Domain
D. Tuple

Answer: D

Q33. A column in a relational table is called:

A. An attribute
B. A relation
C. A key value
D. A tuple

Answer: A

Q34. A primary key must be:

A. Unique and not null
B. Duplicate in values
C. Foreign key based
D. Nullable and optional

Answer: A

Q35. A foreign key is:

A. A key used only for sorting records in a particular order
B. The first column of any table regardless of its data content
C. An attribute that is always unique across all related tables
D. An attribute that references the primary key of another table

Answer: D

Q36. The domain of an attribute specifies:

A. The number of rows containing that attribute
B. The name assigned to identify that attribute
C. The set of allowed values for that attribute
D. The table name where the attribute exists

Answer: C

Q37. The degree of a relation refers to:

A. The number of tables in database
B. The number of tuples (rows) stored
C. The number of keys in the schema
D. The number of attributes (columns)

Answer: D

Q38. The cardinality of a relation refers to:

A. The number of keys in the table
B. The number of foreign key links
C. The number of tuples (rows)
D. The number of attributes (columns)

Answer: C

Q39. Which of the following is a property of a relational table?

A. Rows must be ordered
B. Each row is unique
C. Columns can be duplicated
D. Attributes are ordered

Answer: B

Q40. NULL in a database represents:

A. An empty string character
B. A zero numeric value only
C. A space character in data
D. An unknown or missing value

Answer: D

Q41. The selection operation in relational algebra:

A. Selects tuples (rows) that satisfy a condition
B. Selects specific columns from a given relation
C. Renames a relation or its attribute identifiers
D. Combines two relations into a single relation

Answer: A

Q42. The projection operation in relational algebra:

A. Deletes tuples from a stored relation
B. Selects specific columns from a relation
C. Joins two relations on common attributes
D. Selects specific rows from a given relation

Answer: B

Q43. The union of two relations R and S contains:

A. Only tuples that are common to both R and S
B. All tuples from both R and S without duplicates
C. The Cartesian product of relations R and S
D. Tuples in R but not present in relation S

Answer: B

Q44. Which relational algebra operation renames a relation or its attributes?

A. Rename (< •
B. Projection (< •
C. Division (÷)
D. Selection (<2•

Answer: A

Q45. The Cartesian product of relations R and S is denoted as:

A. R "* S
B. R × S
C. R ") S
D. R " S

Answer: B

Q46. Set difference (R " S) returns:

A. Common tuples found in R and S
B. Tuples in S that are not in R
C. Tuples in R that are not in S
D. All tuples from both R and S

Answer: C

Q47. For union compatibility, two relations must have:

A. The same number of attributes with compatible domains
B. The same name for both relations being combined
C. The same number of tuples stored in both relations
D. Identical primary keys defined on both data relations

Answer: A

Q48. The intersection of R and S (R ") S) returns:

A. Tuples that are only in relation R
B. Tuples that are only in relation S
C. All tuples from relations R and S
D. Tuples that are in both R and S

Answer: D

Q49. Which operation is fundamental in relational algebra (cannot be derived from others)?

A. Intersection
B. Division
C. Selection
D. Natural join

Answer: C

Q50. The natural join operation:

A. Renames attributes within a relation for query convenience
B. Returns all tuples from both relations regardless of matching
C. Combines tuples from two relations based on common attributes
D. Removes duplicates from a single relation without conditions

Answer: C

Q51. SQL stands for:

A. Standard Question Language
B. Structured Query Language
C. System Query Logic
D. Simple Query Language

Answer: B

Q52. Which SQL statement is used to retrieve data from a database?

A. DELETE
B. INSERT
C. UPDATE
D. SELECT

Answer: D

Q53. Which SQL clause is used to filter rows?

A. HAVING
B. ORDER BY
C. WHERE
D. GROUP BY

Answer: C

Q54. Which SQL statement adds new rows to a table?

A. UPDATE SET
B. INSERT INTO
C. ALTER TABLE
D. CREATE INDEX

Answer: B

Q55. Which SQL statement modifies existing data?

A. DROP
B. INSERT
C. UPDATE
D. SELECT

Answer: C

Q56. Which SQL statement removes rows from a table?

A. ALTER
B. DELETE
C. TRUNCATE
D. DROP

Answer: B

Q57. Which SQL command creates a new table?

A. CREATE TABLE
B. NEW TABLE
C. INSERT TABLE
D. ADD TABLE

Answer: A

Q58. Which SQL command permanently removes a table?

A. DESTROY TABLE
B. REMOVE TABLE
C. DELETE TABLE
D. DROP TABLE

Answer: D

Q59. The ORDER BY clause is used to:

A. Filter rows by value
B. Join tables on keys
C. Sort the result set
D. Group rows together

Answer: C

Q60. Which keyword is used to remove duplicate rows from a result?

A. FILTER
B. UNIQUE
C. REMOVE
D. DISTINCT

Answer: D

Q61. A view in SQL is:

A. A stored procedure in the database server
B. A virtual table based on a SELECT query
C. A physical copy of a table stored on disk
D. A type of index on one or more columns

Answer: B

Q62. Which SQL command creates a view?

A. INSERT VIEW
B. MAKE VIEW
C. BUILD VIEW
D. CREATE VIEW

Answer: D

Q63. A stored procedure is:

A. A precompiled set of SQL statements stored in the database
B. An index structure built on one or more columns of a table
C. A type of view that shows data from multiple source tables
D. A constraint that enforces data integrity rules on a table

Answer: A

Q64. A trigger in SQL is:

A. A constraint that restricts the values allowed in a specific column
B. A type of view that displays filtered data from the underlying table
C. A manual query written by the user and executed on demand each time
D. A procedure that automatically executes when a specific event occurs

Answer: D

Q65. What is a cursor in SQL?

A. A pointer to a table that references its location in data storage
B. A database object used to retrieve and process rows one at a time
C. A type of index structure created on one or more table columns
D. A constraint on a column that restricts the allowed data values

Answer: B

Q66. Which SQL statement is used to grant privileges?

A. PERMIT
B. ALLOW
C. ENABLE
D. GRANT

Answer: D

Q67. Which SQL statement removes privileges?

A. DENY
B. DISABLE
C. REVOKE
D. REMOVE

Answer: C

Q68. An assertion in SQL is:

A. A view that provides a virtual representation of the underlying data
B. A stored procedure that runs on a scheduled basis within the server
C. A constraint that specifies a condition the database must always satisfy
D. A type of trigger that fires automatically when data changes occur

Answer: C

Q69. What is a transaction in SQL?

A. A type of view that provides a virtual representation of table data
B. A single SELECT query that retrieves data from the database tables
C. An index structure that speeds up data retrieval from large tables
D. A sequence of operations treated as a single logical unit of work

Answer: D

Q70. The COMMIT statement in SQL:

A. Creates a savepoint within the current active transaction
B. Starts a new transaction and initializes the session state
C. Undoes all changes made in the current transaction entirely
D. Saves all changes made in the current transaction permanently

Answer: D

Q71. Normalization in databases is the process of:

A. Adding redundant data to speed up query processing
B. Deleting all data from every table in the database
C. Creating as many tables as possible in the schema
D. Organizing data to reduce redundancy and dependency

Answer: D

Q72. First Normal Form (1NF) requires:

A. All attributes are restricted to numeric
B. No foreign keys exist within the tables
C. All attributes contain atomic values only
D. Tables have only one column of any data

Answer: C

Q73. A functional dependency X !' Y means:

A. The value of Y uniquely determines the value of X
B. X and Y always contain the same identical values
C. X and Y are completely unrelated to each other
D. The value of X uniquely determines the value of Y

Answer: D

Q74. Redundancy in a database leads to:

A. Insertion, deletion, and update anomalies
B. Better data integrity and consistency
C. Faster queries and improved performance
D. Simplified database design structure

Answer: A

Q75. An update anomaly occurs when:

A. A table is created for the first time in the schema
B. A new row is inserted into an existing data table
C. The same data must be changed in multiple places
D. An index is dropped from a column in the database

Answer: C

Q76. Second Normal Form (2NF) requires:

A. Only atomic values without any other constraints
B. 1NF and no partial dependency on the primary key
C. No functional dependencies of any kind at all
D. All attributes must serve as keys in the table

Answer: B

Q77. Third Normal Form (3NF) requires:

A. No attributes defined at all in it
B. Only 1NF without other constraints
C. 2NF and no transitive dependencies
D. All attributes are primary key parts

Answer: C

Q78. A transitive dependency exists when:

A. Two keys depend on each other in a circular dependency
B. A key attribute depends on a non-key attribute directly
C. A non-key attribute depends on another non-key attribute
D. No dependencies exist between any attributes at all

Answer: C

Q79. A partial dependency exists when:

A. A non-key attribute depends on the entire composite primary key
B. A non-key attribute depends on part of a composite primary key
C. No dependencies exist between any attributes in the data table
D. All attributes depend on all keys defined in the relation schema

Answer: B

Q80. Denormalization is:

A. Removing all tables and their data from the entire database
B. Intentionally introducing redundancy to improve read performance
C. Deleting all indexes to reduce overall storage requirements
D. Normalizing the schema to a higher normal form than before

Answer: B

Q81. ACID properties of a transaction stand for:

A. Atomicity, Concurrency, Isolation, Durability
B. Addition, Consistency, Integrity, Durability
C. Atomicity, Consistency, Isolation, Durability
D. Atomicity, Consistency, Isolation, Dependency

Answer: C

Q82. Atomicity means that a transaction:

A. Runs in isolation from all other transactions
B. Is either fully completed or fully rolled back
C. Is stored permanently on the disk after start
D. Can be partially completed and then paused

Answer: B

Q83. Consistency ensures that a transaction:

A. Never fails regardless of the operations within
B. Is stored permanently on disk after the commit
C. Runs concurrently with other active transactions
D. Brings the database from one valid state to another

Answer: D

Q84. Isolation ensures that:

A. Transactions are never allowed to execute concurrently
B. All transactions see each other's changes immediately
C. Only one transaction can exist in the system at a time
D. Concurrent transactions do not interfere with each other

Answer: D

Q85. Durability ensures that:

A. Data is stored only in volatile memory and not on stable disk
B. Changes can be undone at any time even after transaction commit
C. Committed transaction changes persist even after system failure
D. Transactions are temporary and their changes are never saved

Answer: C

Q86. The COMMIT operation in a transaction:

A. Pauses the active transaction
B. Makes all changes permanent
C. Undoes all current changes
D. Starts a new transaction

Answer: B

Q87. The ROLLBACK operation in a transaction:

A. Undoes all changes made by the transaction
B. Starts a new transaction in the session
C. Deletes the database and all stored data
D. Saves all changes and makes them permanent

Answer: A

Q88. A transaction that has not yet been committed or rolled back is called:

A. Committed
B. Active
C. Aborted
D. Terminated

Answer: B

Q89. Which of the following is a transaction state?

A. All of these: Active, Partially committed, Committed, Failed, Aborted
B. Only Aborted is considered a valid state for a transaction process
C. Only Active is considered a valid state for a transaction in a DBMS
D. Only Committed is considered a valid state for a transaction to be

Answer: A

Q90. A schedule in transaction management is:

A. A single SQL query executed against the database tables
B. A backup plan for recovering from a system level failure
C. A type of index structure for speeding up data lookups
D. A sequence of operations from one or more transactions

Answer: D

Q91. Concurrency control ensures:

A. Larger storage capacity on the disk drives
B. Better network speed between client nodes
C. Faster hardware performance for the system
D. Correct execution of concurrent transactions

Answer: D

Q92. A lock in database concurrency control is:

A. A mechanism to control access to a data item
B. A backup method for protecting against loss
C. A type of query for retrieving stored data
D. A type of index for speeding up data search

Answer: A

Q93. A shared lock allows:

A. Writing by multiple transactions to the same data item
B. No access at all to the locked data item by any party
C. Only one transaction to read the data item at any time
D. Multiple transactions to read the data item simultaneously

Answer: D

Q94. An exclusive lock allows:

A. Multiple transactions to write to the same data item
B. Multiple transactions to read from the same data item
C. Only one transaction to read and write the data item
D. No operations at all on the locked data item by anyone

Answer: C

Q95. A deadlock in database systems occurs when:

A. All transactions commit at the same time without any conflicts found
B. A single transaction is very slow in completing its data operations
C. Two or more transactions are waiting for each other to release locks
D. No transactions are running in the database system at the moment

Answer: C

Q96. The Two-Phase Locking (2PL) protocol has two phases:

A. Lock phase and Unlock phase only
B. Growing phase and Shrinking phase
C. Read phase and Write phase only
D. Start phase and End phase only

Answer: B

Q97. Which of the following can break a deadlock?

A. Adding more locks to the waiting transactions
B. Creating more indexes on the locked data items
C. Aborting one of the deadlocked transactions
D. Increasing the buffer size for the data manager

Answer: C

Q98. Starvation in concurrency control occurs when:

A. All transactions complete quickly without any delays or blocking
B. A transaction waits indefinitely because others keep getting priority
C. The database runs out of storage space on the physical disk
D. No locks are used by any of the transactions in the database

Answer: B

Q99. A lock manager is responsible for:

A. Designing schemas for new data models
B. Creating tables in the database schema
C. Writing SQL queries for data retrieval
D. Granting and releasing locks on data items

Answer: D

Q100. The lock compatibility matrix shows:

A. How to create new tables in the database schema now
B. Which lock types can coexist on the same data item
C. The execution plans for all queries in the database
D. Index structures for optimizing data retrieval speed

Answer: B

Q101. Database recovery is the process of:

A. Deleting all data from the tables in the entire database
B. Restoring the database to a correct state after a failure
C. Creating new tables and defining their column structures
D. Optimizing queries by choosing better execution plans

Answer: B

Q102. A transaction log (write-ahead log) records:

A. Table structures and their column definitions in the database schema
B. Only SELECT queries that have been executed against the database tables
C. User login information and authentication credentials stored securely
D. All changes made by transactions before they are applied to the database

Answer: D

Q103. The redo operation:

A. Creates a new database from scratch without using any data from the previous state
B. Re-applies changes of committed transactions that may not have been written to disk
C. Undoes all changes of committed transactions and reverts to the original data state
D. Deletes the entire log file and removes all records of previous transaction actions

Answer: B

Q104. The undo operation:

A. Reverses changes made by uncommitted transactions
B. Creates a backup of the current database on disk
C. Optimizes queries for faster execution performance
D. Applies committed changes to the database on disk

Answer: A

Q105. A checkpoint in database recovery is:

A. A type of lock used to prevent concurrent access to shared data items
B. An SQL statement used to query data from one or more database tables
C. A point at which all modified data is written to disk and recorded in the log
D. A type of index structure built on columns for faster data retrieval

Answer: C

Q106. The Write-Ahead Log (WAL) protocol requires:

A. Logs are only written to disk after the transaction has committed fully
B. No log is needed for any database operations under this recovery protocol
C. Data must be written to disk before the corresponding log records are made
D. Log records must be written to disk before the corresponding data changes

Answer: D

Q107. A system crash typically causes:

A. Loss of data in main memory but not on disk
B. No data loss of any kind after the crash
C. Loss of all data everywhere in the system
D. Loss of data on disk but not in main memory

Answer: A

Q108. A media failure refers to:

A. Application crash caused by a software defect
B. Network slowness affecting query response times
C. Disk or storage device failure causing data loss
D. Memory overflow due to excessive data loading

Answer: C

Q109. Database backups are important for:

A. Reducing table size by removing unnecessary data rows
B. Recovering from media failures and catastrophic events
C. Creating new databases from the backed up data copies
D. Improving query speed and reducing execution latency

Answer: B

Q110. An incremental backup stores:

A. The entire database every single time
B. Only the database schema definitions
C. Only the changes since the last backup
D. Only the transaction log file entries

Answer: C

Q111. An index in a database is used to:

A. Replace primary key usage
B. Slow down query results
C. Speed up data retrieval
D. Increase storage redundancy

Answer: C

Q112. A primary index is built on:

A. An unordered file without sorting
B. A foreign key only in the relation
C. Any non-key attribute in the table
D. The primary key of an ordered file

Answer: D

Q113. A dense index has:

A. An entry for every block in the data file
B. Only one entry for the entire data file
C. An entry for every record in the data file
D. No entries at all in the index structure

Answer: C

Q114. A sparse index has:

A. An entry for every individual record stored
B. An entry for each block or group of records
C. Entries only for deleted records in the file
D. No entries at all in the entire index file

Answer: B

Q115. A B-tree is a:

A. Linear list with sequential access
B. Balanced search tree used for indexing
C. Binary search tree with two children
D. Unbalanced tree without constraints

Answer: B

Q116. In a B+ tree, all actual data pointers are stored in:

A. All nodes equally
B. Leaf nodes only
C. Root node only
D. Internal nodes only

Answer: B

Q117. Hashing is used for:

A. Direct access to records using a hash function
B. Sorting data into a specific order on columns
C. Sequential access to all records in the file
D. Creating views on top of the base table data

Answer: A

Q118. A secondary index is built on:

A. The ordering field of the file
B. The primary key field only
C. A non-ordering field of a file
D. A foreign key field only

Answer: C

Q119. A clustered index means:

A. Data rows are stored randomly without any ordering
B. Data rows are physically stored in the index order
C. No physical ordering is implied by the index at all
D. Multiple indexes exist on the same column of data

Answer: B

Q120. Which data structure is commonly used for implementing indexes?

A. Linked list nodes
B. Stack structure
C. Queue structure
D. B+ tree structure

Answer: D

Q121. Query processing involves:

A. Only creating tables without any query functionality
B. Only writing queries without any execution or output
C. Only storing data without any retrieval functionality
D. Parsing, optimizing, and executing database queries

Answer: D

Q122. A query execution plan describes:

A. The user who wrote the query and their access privilege details
B. The table creation date and the schema modification history
C. The sequence of operations the DBMS will perform to execute a query
D. The query syntax and grammatical structure in the SQL language

Answer: C

Q123. Query optimization aims to:

A. Find the most efficient way to execute a query
B. Make queries longer and more complex to write
C. Increase the number of disk accesses required
D. Slow down query processing for all operations

Answer: A

Q124. The EXPLAIN command in SQL shows:

A. The database schema layout
B. The query execution plan
C. User access permissions
D. The table data content

Answer: B

Q125. A full table scan reads:

A. Only indexed rows stored
B. Only the last row stored
C. Every row in the table
D. Only the first row found

Answer: C

Q126. Using an index for a query is generally:

A. Always slower than performing a full table scan
B. Faster than a full table scan for selective queries
C. The same speed as a full table scan on all data
D. Not recommended for any database query at all

Answer: B

Q127. Which join algorithm compares every tuple in one relation with every tuple in another?

A. Nested Loop Join
B. Sort-Merge Join
C. Index Nested Join
D. Hash Join method

Answer: A

Q128. The cost of a query is typically measured in:

A. Number of SQL statements used
B. Lines of code in the procedure
C. Number of disk I/O operations
D. Number of tables in the schema

Answer: C

Q129. Selectivity of a condition refers to:

A. The size of the database measured in total bytes
B. The number of columns selected in the projection
C. The fraction of tuples that satisfy the condition
D. The number of tables referenced in the SQL query

Answer: C

Q130. A query optimizer can be:

A. Only cost-based approach
B. Rule-based or cost-based
C. Neither rule nor cost based
D. Only rule-based approach

Answer: B

Q131. Database security aims to protect the database from:

A. Faster queries and improved execution performance
B. Unauthorized access, modification, and destruction
C. Larger storage capacity on the available disk space
D. Better indexing and optimized data retrieval speed

Answer: B

Q132. Authentication in database security verifies:

A. The identity of a user
B. The number of indexes
C. The speed of queries
D. The size of tables

Answer: A

Q133. Authorization in database security determines:

A. The operating system and hardware details
B. What actions a user is permitted to perform
C. The database version and release number
D. The user password and login credentials

Answer: B

Q134. The SQL GRANT statement is used to:

A. Create tables in database
B. Remove privileges from users
C. Delete data from tables
D. Give privileges to users

Answer: D

Q135. The SQL REVOKE statement is used to:

A. Remove previously granted privileges
B. Insert data into existing table records
C. Grant new privileges to database users
D. Create views on top of the base tables

Answer: A

Q136. A database role is:

A. An index structure built on columns for faster data retrieval
B. A type of table used to store data records within the database
C. A named collection of privileges that can be assigned to users
D. A query optimization technique to improve execution planning

Answer: C

Q137. Encryption in database security is used to:

A. Compress data to reduce the amount of storage space it requires
B. Create indexes on columns for faster retrieval of requested records
C. Convert data into an unreadable format to prevent unauthorized access
D. Speed up queries by creating cached copies of frequently used results

Answer: C

Q138. An audit trail in a database records:

A. All operations performed on the database for security monitoring
B. Only DDL operations like CREATE, ALTER, and DROP table commands
C. Only user logins and their session authentication event details
D. Only SELECT queries that have been executed against stored tables

Answer: A

Q139. SQL injection is a type of:

A. Indexing strategy used for speeding up data retrieval from the tables
B. Security attack that exploits vulnerabilities in SQL query construction
C. Database optimization technique that improves overall query performance
D. Backup method used for creating copies of the database for recovery

Answer: B

Q140. The principle of least privilege means:

A. All users should have full access to every table in the database
B. Users should have only the minimum privileges needed for their tasks
C. No privileges should be granted to any user in the database
D. Only administrators should be allowed to use the database system

Answer: B

Q141. A distributed database is:

A. A database stored only in the cloud without any local presence
B. A database stored on a single machine in one physical location
C. A database spread across multiple sites connected by a network
D. A backup copy of a database kept for disaster recovery purposes

Answer: C

Q142. Data fragmentation in distributed databases means:

A. Duplicating the entire database at every site simultaneously
B. Compressing data to reduce storage requirements at each site
C. Losing data during transfer between sites on the network link
D. Dividing a relation into smaller parts stored at different sites

Answer: D

Q143. Horizontal fragmentation divides a relation by:

A. Rows and columns both
B. Rows (tuples) only
C. Neither rows nor cols
D. Columns (attributes)

Answer: B

Q144. Vertical fragmentation divides a relation by:

A. Rows (tuples) only
B. Columns (attributes)
C. Neither of the above
D. Columns and rows both

Answer: B

Q145. Data replication in distributed databases means:

A. Compressing data to reduce transfer size
B. Deleting data from all distributed sites
C. Encrypting data before sending it across
D. Storing copies of data at multiple sites

Answer: D

Q146. The transparency goal in distributed databases means:

A. Data cannot be distributed across multiple sites
B. Users must know all data locations before querying
C. Only administrators can query distributed data
D. Users should not be aware that data is distributed

Answer: D

Q147. Location transparency means:

A. Data is always stored locally on the user's own machine
B. Users must specify the exact data location in every query
C. Users do not need to know where data is physically stored
D. No remote access is possible from any site in the system

Answer: C

Q148. A distributed DBMS (DDBMS) manages:

A. Only local databases without any distribution
B. A database distributed across multiple sites
C. Only a centralized database on one machine
D. Only cloud databases hosted by third parties

Answer: B

Q149. Which of the following is an advantage of distributed databases?

A. Slower queries being preferred by users
B. Higher costs being beneficial overall
C. Improved reliability and availability
D. Increased complexity as an advantage

Answer: C

Q150. Fragmentation transparency means:

A. Users are unaware that data is fragmented
B. Data cannot be fragmented at any site
C. Only fragments can be queried directly
D. Users must reassemble fragments manually

Answer: A

Q151. NoSQL stands for:

A. Not Only SQL format
B. Newer SQL approach
C. Non-Standard Queries
D. No Structured Queries

Answer: A

Q152. Which of the following is a NoSQL database?

A. MySQL
B. PostgreSQL
C. Oracle DB
D. MongoDB

Answer: D

Q153. NoSQL databases are designed primarily for:

A. Replacing all SQL databases entirely in every use case now
B. Only relational data organized in normalized table schemas
C. Small structured datasets only stored in relational tables
D. Scalability and handling large volumes of unstructured data

Answer: D

Q154. A document-oriented database stores data as:

A. Graphs with nodes and edge links
B. Fixed-size blocks on the disk drive
C. Tables with rows and columns only
D. Documents (typically JSON or BSON)

Answer: D

Q155. A key-value store maps:

A. Tables to their schemas
B. Rows to table columns
C. Queries to result sets
D. A unique key to a value

Answer: D

Q156. Redis is an example of a:

A. Key-value store
B. Column-family store
C. Graph database
D. Relational database

Answer: A

Q157. A column-family store organizes data into:

A. Documents in JSON format
B. Traditional rows and tables
C. Columns and column families
D. Graphs with node and edges

Answer: C

Q158. Apache Cassandra is an example of a:

A. Relational database
B. Document database
C. Graph database
D. Column-family store

Answer: D

Q159. A graph database stores data as:

A. Documents in JSON formats
B. Key-value pairs only stored
C. Nodes, edges, and properties
D. Tables and rows of records

Answer: C

Q160. Neo4j is an example of a:

A. Document database
B. Key-value store
C. Relational database
D. Graph database

Answer: D

Q161. A data warehouse is:

A. A transactional database designed for processing day-to-day operations in real-time mode
B. A subject-oriented, integrated, time-variant, non-volatile collection of data for decisions
C. A backup storage mechanism that periodically copies data for disaster recovery plans
D. A type of NoSQL database that supports flexible schemas and horizontal scaling systems

Answer: B

Q162. OLAP stands for:

A. Offline Analytical Processing
B. Online Application Programming
C. Online Automated Processing
D. Online Analytical Processing

Answer: D

Q163. OLTP stands for:

A. Online Text Processing
B. Online Transaction Processing
C. Online Table Processing
D. Offline Transaction Processing

Answer: B

Q164. The ETL process stands for:

A. Export, Translate, Link
B. Enter, Transfer, List
C. Extract, Transfer, Log
D. Extract, Transform, Load

Answer: D

Q165. A fact table in a data warehouse contains:

A. Quantitative measures and foreign keys to dimension tables
B. Only dimension attributes used for filtering and grouping
C. Only text descriptions without any numeric measure values
D. Only primary keys without any quantitative measure data

Answer: A

Q166. A dimension table in a data warehouse contains:

A. Descriptive attributes used for filtering and grouping
B. Log entries recording all data warehouse load operations
C. Transaction records from the operational source systems
D. Quantitative measures like sales amounts and quantities

Answer: A

Q167. The star schema consists of:

A. Multiple fact tables connected to each other in it
B. A central fact table connected to dimension tables
C. Only dimension tables without any fact table data
D. Only fact tables without any dimension table data

Answer: B

Q168. Data mining is:

A. Writing SQL queries for retrieving records from relational tables
B. Creating database backups for disaster recovery and data safety
C. The process of discovering patterns and insights from large datasets
D. Storing data in a warehouse for long-term historical analysis use

Answer: C

Q169. A data mart is:

A. A type of OLTP system designed for transactional data processing
B. A subset of a data warehouse focused on a specific business area
C. A complete data warehouse containing all enterprise data sources
D. A NoSQL database designed for handling unstructured data at scale

Answer: B

Q170. Which of the following is an OLAP operation?

A. INSERT
B. Drill-down
C. UPDATE
D. DELETE

Answer: B

Q171. What is a schema in a database context?

A. A query used to retrieve specific data
B. A backup copy of the entire database
C. The overall design or structure of data
D. A security policy for user permissions

Answer: C

Q172. Which component stores metadata about a database?

A. Query processor
B. Storage engine
C. Data dictionary
D. Buffer manager

Answer: C

Q173. What does data integrity refer to in databases?

A. Amount of storage space data occupies
B. Number of users accessing data at once
C. Accuracy and consistency of stored data
D. Speed of data retrieval operations

Answer: C

Q174. Which language is used to define database structure?

A. DCL
B. DML
C. TCL
D. DDL

Answer: D

Q175. What is a database instance?

A. The design blueprint of the database schema
B. The software used to manage the database
C. The snapshot of data at a given moment
D. The physical disk where data is stored

Answer: C

Q176. Which of these is a function of a DBMS?

A. Rendering graphical user interface views
B. Compiling high-level source code files
C. Routing packets across network devices
D. Managing concurrent data access safely

Answer: D

Q177. What is a query in database terminology?

A. A method to compress stored records
B. A protocol for network communication
C. A type of database backup procedure
D. A request to retrieve or modify data

Answer: D

Q178. Which of these describes a flat-file database?

A. Data distributed across multiple server nodes
B. Data stored in a single two-dimensional table
C. Data organized in a hierarchical tree format
D. Data linked through complex network pointers

Answer: B

Q179. What does ACID stand for in databases?

A. Allocation, Computation, Indexing, Deletion
B. Access, Control, Integrity, Design
C. Atomicity, Consistency, Isolation, Durability
D. Authentication, Concurrency, Integration, Data

Answer: C

Q180. What is the role of a query processor?

A. It encrypts data before storage use
B. It parses and executes user queries
C. It backs up data to external drives
D. It manages physical disk allocation

Answer: B

Q181. What is a client-server database architecture?

A. A system where all processing happens on one machine
B. A system that does not use any network connections
C. A system where data is stored only on client machines
D. A system where clients request services from a server

Answer: D

Q182. What is the function of the storage manager in a DBMS?

A. It handles network routing between nodes
B. It compiles application source code files
C. It creates user interface forms and reports
D. It manages data storage on physical media

Answer: D

Q183. Which tier handles the user interface in a three-tier architecture?

A. Data tier
B. Logic tier
C. Storage tier
D. Presentation tier

Answer: D

Q184. What is a database buffer in the system architecture?

A. A backup copy of the entire database files
B. A network connection between two servers
C. A temporary memory area for data transfer
D. A permanent storage location on the disk

Answer: C

Q185. What does the query compiler do in a DBMS?

A. It handles backup and recovery processes
B. It stores data permanently on disk drives
C. It translates queries into executable plans
D. It manages user authentication credentials

Answer: C

Q186. In a two-tier architecture, which components communicate directly?

A. Two separate database servers communicate
B. The middleware and storage tier communicate
C. The client and database server communicate
D. Two client applications communicate together

Answer: C

Q187. What is the role of the transaction manager?

A. Ensuring transactions execute correctly and safely
B. Designing the user interface for applications
C. Compressing data files for efficient disk storage
D. Routing network requests between server clusters

Answer: A

Q188. Which memory component holds recently accessed disk pages?

A. Permanent storage
B. Cache buffer pool
C. Network registers
D. Virtual memory

Answer: B

Q189. What is a centralized database system?

A. A system spread across many geographic locations
B. A system with no persistent data storage at all
C. A system using only cloud-based storage services
D. A system running on a single computer at one site

Answer: D

Q190. What is the purpose of the data communication component?

A. To enable communication between DBMS and users
B. To perform mathematical calculations on data
C. To define the structure of database tables
D. To compress data before writing to disk drives

Answer: A

Q191. What is a data model in database design?

A. A physical storage device for database files
B. A conceptual framework for organizing data
C. A network protocol for data transmission
D. A programming language for writing queries

Answer: B

Q192. Which data model organizes data into tables of rows and columns?

A. Network model
B. Hierarchical model
C. Relational model
D. Object model

Answer: C

Q193. In the hierarchical data model, data is organized as:

A. A graph with many-to-many connections
B. An unstructured collection of documents
C. A tree with parent-child relationships
D. A flat two-dimensional table structure

Answer: C

Q194. What is an entity in the ER model?

A. A constraint applied to a column in a table
B. A real-world object represented in a database
C. A relationship between two database tables
D. A query used to retrieve records from data

Answer: B

Q195. What does an attribute represent in a data model?

A. A connection between two separate entities
B. A property or characteristic of an entity
C. A method for sorting records in a table
D. A rule that constrains data entry values

Answer: B

Q196. What is a relationship in the ER model?

A. A method for indexing table records fast
B. An association between two or more entities
C. A column in a database table structure
D. A constraint on a single attribute value

Answer: B

Q197. Which symbol represents an entity in an ER diagram?

A. Triangle shape
B. Rectangle shape
C. Oval shape
D. Diamond shape

Answer: B

Q198. What does cardinality specify in a relationship?

A. The number of entity instances in a link
B. The data type of attributes in an entity
C. The access permission level for database users
D. The storage format for records on a disk

Answer: A

Q199. What is a composite attribute in the ER model?

A. An attribute that can be divided into subparts
B. An attribute that has multiple values at once
C. An attribute derived from other stored values
D. An attribute that cannot be subdivided further

Answer: A

Q200. Which data model uses nodes and edges to represent data?

A. Graph data model
B. Columnar model
C. Relational model
D. Flat file model

Answer: A

Q201. What is a tuple in a relational database?

A. A column in a relational database table
B. A key in a relational database design
C. A table in a relational database schema
D. A row in a relational database table

Answer: D

Q202. What is a primary key in a relational table?

A. A column that always contains null values
B. A column that uniquely identifies each row
C. A column that links to an external file
D. Any column that contains text data values

Answer: B

Q203. What is a foreign key used for?

A. Calculating aggregate functions on data sets
B. Encrypting data in a database column field
C. Linking one table to another related table
D. Sorting records in ascending order by value

Answer: C

Q204. What does the term relation refer to in relational databases?

A. A connection between two different databases
B. A table with rows and columns of data
C. A constraint applied to a single column
D. A query that joins multiple data sources

Answer: B

Q205. What is a domain in a relational database?

A. A type of relationship between two tables
B. A network address for the database server
C. A set of allowable values for an attribute
D. A backup strategy for database recovery use

Answer: C

Q206. What does a candidate key represent?

A. A key that is always the foreign key value
B. A key used only for sorting table records
C. A key that allows duplicate values in rows
D. A key that might be selected as primary key

Answer: D

Q207. What is the degree of a relation?

A. The number of constraints on it
B. The number of rows in the table
C. The number of keys in the table
D. The number of columns in the table

Answer: D

Q208. What are null values used to represent?

A. The empty string in a text data column
B. Missing or unknown information in a field
C. The number zero in a numeric column field
D. A default value assigned to every record

Answer: B

Q209. What is a superkey in relational databases?

A. A set of attributes that uniquely identifies rows
B. A key used only in object-oriented databases
C. A key with exactly one attribute column only
D. A key that cannot contain any null value at all

Answer: A

Q210. What is an attribute in a relational table?

A. A column within the relational table structure
B. A row within the relational table structure
C. A constraint applied to the entire table
D. A relationship between two different tables

Answer: A

Q211. What is relational algebra?

A. A data storage format for flat files
B. A programming language for web development
C. A procedural query language for relations
D. A network protocol for data transfer

Answer: C

Q212. What does the selection operation do?

A. It combines two relations into one table
B. It picks specific rows matching a condition
C. It picks specific columns from a relation
D. It removes duplicate rows from a relation

Answer: B

Q213. What does the projection operation do?

A. It selects specific rows based on conditions
B. It calculates aggregate values for a group
C. It joins two relations on a common column
D. It selects specific columns from a relation

Answer: D

Q214. What is the Cartesian product of two relations?

A. The difference of rows between both tables
B. All combinations of rows from both tables
C. Only matching rows from both of the tables
D. The intersection of rows from both tables

Answer: B

Q215. What does the union operation require?

A. Relations must be union-compatible to combine
B. Relations must share a common primary key
C. Relations must be stored in the same file
D. Relations must have different attribute counts

Answer: A

Q216. What does the set difference operation produce?

A. All rows appearing in either of two relations
B. Only rows common to both of the two relations
C. A sorted version of the first input relation
D. Rows in the first relation but not the second

Answer: D

Q217. What is a natural join operation?

A. A join that only works on single-column tables
B. A join that matches rows on common attributes
C. A join that removes all duplicate column values
D. A join that produces the Cartesian product only

Answer: B

Q218. What does the rename operation do in relational algebra?

A. It sorts the tuples in a relation by key order
B. It deletes a relation from the database schema
C. It creates a backup copy of the existing table
D. It changes the name of a relation or attribute

Answer: D

Q219. What is the intersection of two relations?

A. Tuples present in both relations simultaneously
B. Tuples in the first but not in the second one
C. All tuples from both relations combined together
D. Tuples in the second but not in the first one

Answer: A

Q220. Which operation combines selection and join in one step?

A. Rename operation
B. Projection operation
C. Division operation
D. Theta join operation

Answer: D

Q221. What does SQL stand for?

A. Structured Query Language
B. Sequential Query Logic
C. Standard Query Library
D. System Query Lookup

Answer: A

Q222. Which SQL command retrieves data from a table?

A. UPDATE
B. DELETE
C. INSERT
D. SELECT

Answer: D

Q223. Which clause filters rows in a SELECT statement?

A. ORDER BY
B. WHERE
C. GROUP BY
D. HAVING

Answer: B

Q224. What does the INSERT command do in SQL?

A. Modifies existing rows in a table
B. Adds new rows into a table
C. Deletes rows from a table
D. Creates a new database table

Answer: B

Q225. Which SQL keyword removes duplicate rows from results?

A. FILTER
B. PRIMARY
C. DISTINCT
D. UNIQUE

Answer: C

Q226. What does the UPDATE command do in SQL?

A. Removes rows from the table
B. Retrieves rows from a table
C. Modifies existing rows in table
D. Creates a new table in database

Answer: C

Q227. Which SQL command deletes rows from a table?

A. CLEAR
B. DELETE
C. DROP
D. REMOVE

Answer: B

Q228. What does the ORDER BY clause do?

A. Sorts the result set by columns
B. Groups rows by a specified column
C. Limits the number of output rows
D. Filters rows based on a condition

Answer: A

Q229. Which SQL command creates a new table?

A. CREATE TABLE
B. BUILD TABLE
C. MAKE TABLE
D. NEW TABLE

Answer: A

Q230. What is the wildcard character in SQL SELECT?

A. Hash symbol
B. Asterisk symbol
C. Ampersand symbol
D. Dollar symbol

Answer: B

Q231. What is a stored procedure in SQL?

A. A single SQL query saved as a bookmark
B. A temporary table created during a query
C. A precompiled set of SQL statements stored
D. A backup copy of the database schema file

Answer: C

Q232. What is a view in SQL?

A. A physical table stored on the disk drive
B. A log file recording all query activity
C. A virtual table based on a query result
D. A backup of the original database table

Answer: C

Q233. What is a trigger in SQL databases?

A. A type of index for faster data retrieval
B. Code that executes automatically on events
C. A manual command executed by the user only
D. A constraint that limits column data types

Answer: B

Q234. What is an index used for in SQL?

A. Creating backup copies of table records
B. Defining relationships between two tables
C. Encrypting data stored in database tables
D. Speeding up data retrieval from tables

Answer: D

Q235. What does the GRANT command do in SQL?

A. Gives users specific access permissions
B. Creates new columns in existing tables
C. Backs up data to an external location
D. Removes existing tables from a database

Answer: A

Q236. What is a cursor in SQL?

A. A method for creating temporary table data
B. A type of database constraint on columns
C. A graphical pointer on the user interface
D. A mechanism to traverse query result rows

Answer: D

Q237. What does the REVOKE command do?

A. It deletes data from a specified table
B. It creates new user accounts in database
C. It removes previously granted permissions
D. It modifies the structure of a table

Answer: C

Q238. What is a transaction in SQL?

A. A unit of work with multiple operations
B. A single SELECT query against one table
C. A method for defining table constraints
D. A type of database index on a column

Answer: A

Q239. What does COMMIT do in SQL?

A. Deletes all data from the current table
B. Creates a new savepoint within transaction
C. Saves all changes made in a transaction
D. Undoes all changes made in a transaction

Answer: C

Q240. What does ROLLBACK do in SQL?

A. Saves all changes made in the transaction
B. Grants permissions to a specified database user
C. Undoes all changes made in the transaction
D. Creates a new table in the active database

Answer: C

Q241. What is database normalization?

A. Organizing data to reduce redundancy and dependency
B. Adding redundant data to improve query speed
C. Encrypting data for security and privacy protection
D. Compressing data to save physical disk storage space

Answer: A

Q242. What is the first normal form (1NF)?

A. All candidate keys are properly identified and used
B. Each column contains only atomic indivisible values
C. All non-key attributes depend on the primary key
D. No transitive dependencies exist in the relation

Answer: B

Q243. What is a functional dependency?

A. A relationship between two separate database tables
B. A method for sorting records in ascending order
C. A constraint where one attribute determines another
D. A technique for encrypting stored data values

Answer: C

Q244. What is data redundancy in databases?

A. Data that is encrypted for security purposes
B. Data that is properly normalized in tables
C. Unnecessary repetition of data across tables
D. Data stored in compressed file formats only

Answer: C

Q245. What problem does normalization primarily solve?

A. Data anomalies from redundancy
B. Slow network connection speeds
C. Lack of user interface design
D. Insufficient hardware capacity

Answer: A

Q246. What is an insertion anomaly?

A. Inability to add data without other unrelated data
B. Difficulty deleting unwanted records from tables
C. Errors when updating records across multiple tables
D. Problems with reading data from large database files

Answer: A

Q247. What is a deletion anomaly?

A. Creating circular references between related tables now
B. Failing to update all copies of redundant data values
C. Losing unrelated data when deleting a specific record
D. Adding duplicate records accidentally to a table field

Answer: C

Q248. What is an update anomaly?

A. Inconsistency from failing to update all data copies
B. Inability to insert new records into a database table
C. Failure to create proper indexes on table columns here
D. Accidentally deleting records from a database table now

Answer: A

Q249. What is denormalization?

A. Intentionally adding redundancy for performance gain
B. Converting a relational database to a flat file type
C. The process of achieving first normal form status
D. Removing all indexes from database table columns

Answer: A

Q250. What does 2NF require beyond 1NF?

A. No transitive dependencies on the primary key
B. Every column must have a unique non-null value
C. No partial dependencies on the primary key set
D. All attributes must be of the same data type

Answer: C

Q251. What is a database transaction?

A. A report generated from database queries
B. A backup copy of the database schema
C. A connection between two database servers
D. A logical unit of work on a database

Answer: D

Q252. What does the Atomicity property ensure?

A. All or none of the operations are completed
B. Changes persist even after system failures
C. Database remains consistent after transaction
D. Transactions execute in complete isolation always

Answer: A

Q253. What does the Consistency property ensure?

A. All operations in a transaction execute together
B. Database moves from one valid state to another
C. Transactions run without any interference at all
D. Transaction results are permanently stored safely

Answer: B

Q254. What does the Isolation property ensure?

A. All operations complete or none complete at all
B. Changes are saved permanently after a commit
C. Concurrent transactions do not interfere with others
D. The database schema remains unchanged after commits

Answer: C

Q255. What does the Durability property ensure?

A. Transactions can be rolled back at any later time
B. Multiple transactions can run at the same time now
C. Data types remain consistent across all table rows
D. Committed changes persist even after system failure

Answer: D

Q256. What does COMMIT do in transaction processing?

A. It creates a new transaction from the current state
B. It saves all changes made during the transaction
C. It undoes all changes made during the transaction
D. It temporarily pauses the current active transaction

Answer: B

Q257. What does ROLLBACK do in transaction processing?

A. Pauses the transaction until the user resumes it later
B. Saves all pending changes to the database permanently
C. Undoes all changes since the last commit or savepoint
D. Creates a backup of the database before rolling back

Answer: C

Q258. What is a schedule in transaction processing?

A. An ordered sequence of operations from transactions
B. A configuration file for database server parameters
C. A backup plan for the database recovery process
D. A list of all users currently connected to database

Answer: A

Q259. What is a serial schedule?

A. Transactions execute one after another without overlap
B. Transactions execute simultaneously on different processors
C. Transactions execute their operations in interleaved order
D. Transactions execute in reverse order of their submission

Answer: A

Q260. What is a transaction log used for?

A. Storing user passwords and authentication credentials
B. Managing network connections between client and server
C. Recording all changes for recovery and audit purposes
D. Displaying query results to the end user interface

Answer: C

Q261. What is concurrency control in databases?

A. Backing up data to prevent data loss scenarios
B. Managing simultaneous access by multiple users
C. Encrypting data for secure transmission purposes
D. Compressing data to save disk storage space used

Answer: B

Q262. What is a lock in database concurrency control?

A. A method for compressing database table records
B. A technique for encrypting sensitive data fields
C. A mechanism to restrict access to data items
D. A backup mechanism for database recovery tasks

Answer: C

Q263. What is a shared lock used for?

A. Preventing all access to the data item entirely
B. Allowing only one transaction to access the data
C. Allowing multiple transactions to read the data
D. Allowing only write access to the data item

Answer: C

Q264. What is an exclusive lock used for?

A. Giving one transaction sole access for writing
B. Sharing write access among multiple transactions
C. Allowing multiple users to read data together
D. Preventing all transactions from accessing data

Answer: A

Q265. What is a deadlock in database systems?

A. A backup process that runs during peak usage hours
B. A transaction that completes faster than expected
C. Two transactions waiting for each other indefinitely
D. A query that returns more results than anticipated

Answer: C

Q266. What is starvation in concurrency control?

A. A transaction completes too quickly for proper logging
B. A transaction produces incorrect results due to bad data
C. A transaction waits indefinitely for a resource access
D. A transaction uses too much memory during its execution

Answer: C

Q267. What does the lock manager do in a DBMS?

A. It compiles SQL queries into executable code plans
B. It creates user accounts and manages permissions
C. It manages physical storage allocation on disk drives
D. It grants and releases locks on data items safely

Answer: D

Q268. What is a lock table used for?

A. Maintaining backup copies of all locked data items
B. Storing the actual data values of locked records
C. Tracking which locks are held by which transactions
D. Recording all SQL queries executed by each user

Answer: C

Q269. What is lock compatibility?

A. Whether two lock types can coexist on same item
B. Whether a lock persists after system restart event
C. Whether a lock can be transferred between users
D. Whether a lock applies to one or multiple tables

Answer: A

Q270. What is the purpose of a timeout in deadlock handling?

A. To extend the duration of held locks indefinitely
B. To increase the priority of waiting transactions
C. To abort a transaction after waiting too long
D. To speed up slow-running queries automatically

Answer: C

Q271. What is database recovery?

A. Adding new data records to database tables
B. Optimizing queries for faster execution speed
C. Restoring a database after a failure event
D. Creating user accounts for database access

Answer: C

Q272. What is a system crash in database context?

A. A routine update of the database software used
B. A planned maintenance shutdown of the server
C. An unexpected failure that disrupts operations
D. A scheduled backup of the database to disk

Answer: C

Q273. What is a transaction log?

A. A record of all changes made by transactions
B. A backup copy of the database schema only
C. A list of all database users and permissions
D. A configuration file for server parameters

Answer: A

Q274. What is the purpose of a database backup?

A. To create a copy of data for recovery purposes
B. To speed up query processing performance times
C. To encrypt sensitive data for security compliance
D. To normalize database tables for better design

Answer: A

Q275. What does the REDO operation do in recovery?

A. Deletes all log records from the transaction log file
B. Creates a new backup of the database on the disk
C. Reapplies changes from committed transactions in log
D. Undoes changes from committed transactions in log

Answer: C

Q276. What does the UNDO operation do in recovery?

A. Applies changes from committed transactions to data
B. Compresses the transaction log for storage savings
C. Reverses changes from uncommitted transactions safely
D. Creates new checkpoints in the transaction log file

Answer: C

Q277. What is a checkpoint in database recovery?

A. A lock that prevents concurrent access to data
B. A point where the database state is synchronized
C. A user account with limited database permissions
D. A query that checks for data integrity violations

Answer: B

Q278. What is the difference between volatile and non-volatile storage?

A. Non-volatile storage is faster but loses data on power loss
B. Neither type of storage can lose data under any conditions
C. Volatile storage is faster but loses data on power loss
D. Both types of storage retain data after power is lost

Answer: C

Q279. What is a disk failure in database context?

A. A routine defragmentation of the disk drive unit
B. A malfunction that causes loss of stored disk data
C. A scheduled cleaning of the disk read-write heads
D. A planned upgrade of the disk storage hardware

Answer: B

Q280. What is the purpose of a log buffer?

A. Temporarily holding log records before writing to disk
B. Encrypting data before it is stored on the disk drive
C. Backing up the entire database to an external device
D. Storing user queries before they are executed later

Answer: A

Q281. What is an index in a database?

A. A backup method for database recovery tasks
B. A security mechanism for user authentication
C. A data structure for fast data retrieval access
D. A complete copy of the database table data

Answer: C

Q282. What is a primary index in a database?

A. An index that only supports range-based queries
B. An index built on the primary key of the table
C. An index built on a non-key column of a table
D. An index that stores all table data within itself

Answer: B

Q283. What is a secondary index?

A. An index that is only used for backup purposes
B. An index that replaces the primary key entirely
C. An index on the ordering key of the data file
D. An index on a non-ordering field of a data file

Answer: D

Q284. What is a dense index?

A. An index that compresses data for storage efficiency
B. An index that stores data without any key structure
C. An index with an entry for every data record stored
D. An index with an entry for some data records only

Answer: C

Q285. What is a sparse index?

A. An index with entries for only some data records
B. An index with an entry for every single data record
C. An index stored entirely in main memory at all times
D. An index that does not use any key values at all

Answer: A

Q286. What is a B-tree in database indexing?

A. A linear list of pointers to database records
B. A balanced tree structure for disk-based indexing
C. A binary search tree for in-memory data only
D. A hash-based structure for exact match queries

Answer: B

Q287. What is a hash index used for?

A. Efficient pattern matching on text data values
B. Efficient exact match queries on key values
C. Efficient range queries on sorted data values
D. Efficient sequential scanning of all records

Answer: B

Q288. What is a clustered index?

A. An index where data order differs from index order
B. An index where data order matches the index order
C. An index created automatically on every new column
D. An index that works only on temporary table data

Answer: B

Q289. What is sequential file organization?

A. Records stored in a tree structure on disk blocks
B. Records stored in a hash table on the disk blocks
C. Records stored in random order on the disk blocks
D. Records stored in sorted order on the disk blocks

Answer: D

Q290. What is a heap file organization?

A. Records are stored in the order they are inserted
B. Records are stored in sorted order by primary key
C. Records are stored using a B-tree index structure
D. Records are stored in a hash table data structure

Answer: A

Q291. What is query processing in a database?

A. Managing user accounts and access permissions sets
B. Designing the physical layout of storage on disk
C. Translating and executing queries against the data
D. Creating backup copies of database tables regularly

Answer: C

Q292. What is a query execution plan?

A. A list of all users who can run the query
B. A log of all previously executed database queries
C. A backup schedule for the database operations
D. The strategy used to execute a database query

Answer: D

Q293. What does a query optimizer do?

A. It creates indexes on all columns automatically now
B. It translates SQL into a different programming language
C. It selects the most efficient query execution plan
D. It encrypts queries before sending to the server

Answer: C

Q294. What is a full table scan?

A. Reading only indexed rows from the data table
B. Reading only the first row of the database table
C. Reading rows from multiple tables at the same time
D. Reading every row in the table sequentially stored

Answer: D

Q295. What is the purpose of query parsing?

A. Sorting the query results in alphabetical order
B. Checking the syntax and validity of SQL queries
C. Executing the query against the stored database
D. Backing up the query text for future reference

Answer: B

Q296. What is a cost-based query optimizer?

A. An optimizer that selects the plan with lowest cost
B. An optimizer that ignores indexes during planning
C. An optimizer that always chooses the first valid plan
D. An optimizer that only works with simple queries

Answer: A

Q297. What are statistics used for in query optimization?

A. Estimating the cost of different execution plans
B. Encrypting data before storage on disk drives
C. Displaying query results to end users on screen
D. Creating user accounts and password management

Answer: A

Q298. What is an index scan in query processing?

A. Scanning all rows in a table without using any index
B. Using an index to locate specific rows efficiently
C. Scanning multiple tables simultaneously for matches
D. Scanning only the first page of the data table

Answer: B

Q299. What is query decomposition?

A. Encrypting query text to prevent unauthorized access
B. Combining multiple queries into a single large query
C. Breaking a complex query into simpler sub-operations
D. Removing queries from the execution queue permanently

Answer: C

Q300. What is the role of the query evaluation engine?

A. It manages physical storage allocation on the disk
B. It executes the plan generated by the optimizer
C. It creates user interfaces for database applications
D. It designs the database schema and table layouts

Answer: B

Q301. What is database security?

A. Protecting data from unauthorized access or damage
B. Compressing data to save storage space on disk
C. Optimizing queries for faster execution speeds
D. Sorting data in alphabetical order in all tables

Answer: A

Q302. What is authentication in database security?

A. Compressing database files for efficient storage
B. Verifying the identity of a user accessing data
C. Encrypting data stored in the database tables
D. Backing up data to prevent loss from failures

Answer: B

Q303. What is authorization in database security?

A. Determining what actions a user is allowed to do
B. Verifying user identity before granting access
C. Encrypting data before storing it in the database
D. Creating backup copies of the database regularly

Answer: A

Q304. What does the GRANT command do?

A. Creates a new database on the server system
B. Removes user access to database objects entirely
C. Gives specific permissions to users or roles
D. Deletes data from a specified database table

Answer: C

Q305. What does the REVOKE command do?

A. Adds new permissions to a user or role account
B. Encrypts all data in the specified database table
C. Creates new user accounts in the database system
D. Removes previously granted permissions from users

Answer: D

Q306. What is data encryption in database security?

A. Duplicating data across multiple storage devices
B. Converting data into an unreadable coded format
C. Sorting data in a specific predetermined order
D. Compressing data to reduce its storage size need

Answer: B

Q307. What is a database role?

A. A named collection of related database privileges
B. A physical server that hosts the database system
C. A backup strategy for disaster recovery planning
D. A type of query used for data retrieval purposes

Answer: A

Q308. What is an audit trail in database security?

A. A method for optimizing database query performance
B. A record of all access and changes to the database
C. A technique for compressing data stored on disk drives
D. A strategy for distributing data across multiple servers

Answer: B

Q309. What is SQL injection?

A. A method for optimizing SQL query performance time
B. An attack that inserts malicious SQL code in input
C. A process for encrypting SQL queries before running
D. A technique for backing up SQL database regularly

Answer: B

Q310. What is access control in database security?

A. Compressing database tables to save storage space
B. Optimizing the speed of data retrieval operations
C. Regulating who can access what data and operations
D. Creating indexes for faster query execution speeds

Answer: C

Q311. What is a distributed database?

A. A database accessible only through a local area network
B. A database spread across multiple connected computers
C. A database that only stores metadata and no actual data
D. A database stored entirely on a single server machine

Answer: B

Q312. What is data fragmentation in distributed databases?

A. Compressing data to reduce network transfer overhead
B. Encrypting data before distributing to remote servers
C. Breaking a relation into smaller parts across sites
D. Combining data from multiple sources into one table

Answer: C

Q313. What is horizontal fragmentation?

A. Replicating the entire table across all network sites
B. Combining multiple tables into a single unified table
C. Dividing a table by selecting specific rows from it
D. Dividing a table by selecting specific columns from it

Answer: C

Q314. What is vertical fragmentation?

A. Dividing a table by selecting specific rows based on value
B. Dividing a table by selecting specific columns from table
C. Combining rows from different tables into a unified view
D. Storing complete copies of a table at every network site

Answer: B

Q315. What is data replication in distributed databases?

A. Encrypting data before transmission between all sites
B. Dividing data into fragments across different sites
C. Compressing data to reduce the overall storage needed
D. Storing copies of data at multiple sites for access

Answer: D

Q316. What is a distributed DBMS?

A. Software that manages data across multiple network sites
B. Software that only handles backup and recovery operations
C. Software that manages files on a single computer only
D. Software that designs user interfaces for database access

Answer: A

Q317. What is location transparency?

A. Users must establish connections to each site manually
B. Users can only access data stored on their local site
C. Users access data without knowing its physical location
D. Users must specify which site stores the requested data

Answer: C

Q318. What is a site in a distributed database?

A. A computer or node that hosts part of the database
B. A column in a distributed database table structure
C. A type of query used in distributed data processing
D. A backup copy of the distributed database for safety

Answer: A

Q319. What is fragmentation transparency?

A. Users access data without knowing how it is fragmented
B. Users must specify fragment names in all their queries
C. Users can only query one fragment at a time per query
D. Users must manually join fragments when querying data

Answer: A

Q320. What is a distributed query?

A. A query that is scheduled to run at a future time date
B. A query that runs on a single database server only
C. A query that only retrieves metadata from the catalog
D. A query that accesses data from multiple network sites

Answer: D

Q321. What does NoSQL stand for?

A. Not Only SQL databases
B. Networked Open SQL system
C. No Structured Query Language
D. Non-Sequential Query Logic

Answer: A

Q322. Which is an example of a document-based NoSQL database?

A. MongoDB
B. MySQL
C. Redis
D. Neo4j

Answer: A

Q323. What is a key-value store?

A. A database that maps unique keys to their values
B. A database that uses SQL for all data operations
C. A database that stores data in relational tables
D. A database that only stores numeric data values

Answer: A

Q324. Which is an example of a graph database?

A. MongoDB
B. Neo4j
C. Cassandra
D. Redis

Answer: B

Q325. What type of NoSQL database is Redis?

A. Document store
B. Column-family store
C. Graph database store
D. Key-value store

Answer: D

Q326. What type of NoSQL database is Cassandra?

A. Key-value store type
B. Column-family store
C. Graph database type
D. Document store type

Answer: B

Q327. What is a document database?

A. A database storing data in graph node structures
B. A database storing data in fixed relational tables
C. A database storing only text document file contents
D. A database storing data in flexible document formats

Answer: D

Q328. What is horizontal scaling in NoSQL databases?

A. Adding more resources to a single existing server
B. Adding more servers to distribute the workload
C. Upgrading the software on the existing servers
D. Reducing the number of servers in the cluster

Answer: B

Q329. What is schema flexibility in NoSQL databases?

A. Schema changes require rebuilding the entire database
B. Schema must be defined before any data can be inserted
C. Records in the same collection can have different fields
D. All records must follow the exact same fixed schema

Answer: C

Q330. What is the main advantage of NoSQL over relational databases?

A. More mature and established query language support
B. Better support for complex multi-table join queries
C. Better scalability for large unstructured data sets
D. Stronger ACID transaction support across tables

Answer: C

Q331. What is a data warehouse?

A. A central repository for integrated historical data
B. A physical building where servers are maintained
C. A system for real-time transaction processing only
D. A temporary storage for data during transmission

Answer: A

Q332. What does ETL stand for in data warehousing?

A. Evaluate, Test, Launch processing
B. Encrypt, Transfer, Load processing
C. Extract, Transform, Load processing
D. Edit, Track, Log processing steps

Answer: C

Q333. What is a fact table in a data warehouse?

A. A table containing descriptive text about entities
B. A table containing measurable business event data
C. A table containing database configuration settings
D. A table storing user credentials and permissions

Answer: B

Q334. What is a dimension table?

A. A table storing system configuration parameters
B. A table providing descriptive context for facts
C. A table storing numeric measurement values only
D. A table containing temporary processing results

Answer: B

Q335. What is OLAP in data warehousing?

A. Online Application Logging and Processing
B. Online Analytical Processing of stored data
C. Optimized Linear Analysis of Procedures
D. Offline Logical Access and Programming

Answer: B

Q336. What is a star schema?

A. A schema without any fact or dimension tables used
B. A schema with many fact tables and one dimension
C. A schema where all tables are fully normalized form
D. A schema with one fact table and many dimensions

Answer: D

Q337. What is data mining?

A. Deleting old data from the database storage files
B. Backing up data to external storage media drives
C. Manually entering data into database tables one row
D. Discovering patterns and knowledge from large data

Answer: D

Q338. What is the purpose of data integration?

A. Splitting data into smaller separate fragments only
B. Encrypting data before storing in the data warehouse
C. Combining data from multiple sources into unified view
D. Deleting duplicate records from all source databases

Answer: C

Q339. What is a data mart?

A. A complete enterprise-wide data warehouse system
B. A tool for mining patterns from large data volumes
C. A method for encrypting warehouse data at rest now
D. A subset of a warehouse for a specific department

Answer: D

Q340. What is the difference between OLTP and OLAP?

A. OLTP and OLAP serve the exact same purpose in databases
B. OLTP is newer technology while OLAP is the older version
C. OLTP is for analytics while OLAP is for transactions only
D. OLTP handles transactions while OLAP handles analytics

Answer: D

Q341. Which type of database stores data in key-value pairs?

A. Relational database
B. Key-value store
C. Hierarchical database
D. Network database

Answer: B

Q342. What is the main purpose of a database management system?

A. To compile programming code
B. To manage and organize data efficiently
C. To design user interfaces
D. To configure network settings

Answer: B

Q343. Which of the following is an example of a relational DBMS?

A. MongoDB
B. Redis
C. MySQL
D. Cassandra

Answer: C

Q344. What does DDL stand for in database systems?

A. Data Description Logic
B. Data Definition Language
C. Database Design Layer
D. Dynamic Data Linking

Answer: B

Q345. Which language is most commonly used to interact with relational databases?

A. Python
B. Java
C. SQL
D. C++

Answer: C

Q346. What is a record in a database table?

A. A column header
B. A single row of data
C. The table name
D. A database file

Answer: B

Q347. Which of the following best describes data consistency?

A. Data is stored in multiple formats
B. Data follows defined rules and remains accurate
C. Data is always encrypted
D. Data is compressed for storage

Answer: B

Q348. What is a field in a database?

A. An entire table
B. A single piece of data within a record
C. A database backup
D. A type of query

Answer: B

Q349. Which of the following is NOT a benefit of using a DBMS?

A. Reduced data redundancy
B. Improved data security
C. Increased hardware speed
D. Data backup and recovery

Answer: C

Q350. What does DML stand for in database terminology?

A. Data Manipulation Language
B. Data Modeling Logic
C. Database Management Layer
D. Dynamic Memory Loading

Answer: A

Q351. What is the simplest form of database architecture?

A. Three-tier architecture
B. Distributed architecture
C. Single-tier architecture
D. Cloud architecture

Answer: C

Q352. In database architecture, what does the term 'tier' refer to?

A. A type of database table
B. A logical or physical layer in the system
C. A security protocol
D. A backup mechanism

Answer: B

Q353. Which component interprets and executes database queries?

A. Storage engine
B. Query processor
C. Network adapter
D. Display driver

Answer: B

Q354. What is the purpose of the disk manager in a DBMS?

A. To display query results
B. To manage the allocation of space on disk
C. To handle user authentication
D. To optimize network traffic

Answer: B

Q355. In a two-tier architecture, which component sends requests to the server?

A. Database engine
B. Client application
C. Storage manager
D. Buffer pool

Answer: B

Q356. What does the term 'back-end' refer to in database architecture?

A. The user interface
B. The server side that manages the database
C. The network cable
D. The display monitor

Answer: B

Q357. Which architecture separates presentation, logic, and data layers?

A. Single-tier architecture
B. Two-tier architecture
C. Three-tier architecture
D. Peer-to-peer architecture

Answer: C

Q358. What is the role of the authorization manager in a DBMS?

A. To optimize queries
B. To verify user permissions before allowing operations
C. To manage disk space
D. To handle data backup

Answer: B

Q359. What does a database server primarily do?

A. Renders web pages
B. Processes and responds to database requests
C. Manages email services
D. Compiles application code

Answer: B

Q360. Which DBMS component temporarily holds data pages in memory?

A. Query parser
B. Buffer pool
C. Log manager
D. Schema manager

Answer: B

Q361. What is a simple attribute in the ER model?

A. An attribute that can be divided into sub-parts
B. An attribute that cannot be divided further
C. An attribute with multiple values
D. An attribute derived from other attributes

Answer: B

Q362. What does a diamond shape represent in an ER diagram?

A. An entity
B. An attribute
C. A relationship
D. A primary key

Answer: C

Q363. What is a key attribute in an ER diagram?

A. Any attribute of an entity
B. An attribute that uniquely identifies each entity instance
C. An attribute that stores encrypted data
D. An attribute that links to another table

Answer: B

Q364. In the ER model, what is a one-to-many relationship?

A. One entity instance relates to at most one instance of another entity
B. One entity instance can relate to many instances of another entity
C. Many entity instances relate to many instances of another entity
D. No entity instances are related

Answer: B

Q365. What is the purpose of an ER diagram?

A. To write SQL queries
B. To visually represent the database structure and relationships
C. To configure network settings
D. To define programming functions

Answer: B

Q366. Which data model represents data as a collection of objects?

A. Relational model
B. Hierarchical model
C. Object-oriented data model
D. Flat file model

Answer: C

Q367. In the ER model, what is a many-to-many relationship?

A. Each entity instance relates to exactly one other
B. Multiple instances of both entities can relate to each other
C. Only one entity participates in the relationship
D. The relationship has no cardinality

Answer: B

Q368. What does an oval shape represent in an ER diagram?

A. An entity
B. A relationship
C. An attribute
D. A constraint

Answer: C

Q369. What is a strong entity in the ER model?

A. An entity that has a primary key and can exist independently
B. An entity that depends on another entity for identification
C. An entity with the most attributes
D. An entity connected to all other entities

Answer: A

Q370. In the network data model, what type of relationship structure is used?

A. Tree structure
B. Graph structure allowing many-to-many relationships
C. Simple flat tables
D. Star schema

Answer: B

Q371. What is an alternate key in a relational database?

A. A key that is always numeric
B. A candidate key that is not selected as the primary key
C. A key used only for backup purposes
D. A key that allows duplicate values

Answer: B

Q372. What is referential integrity in a relational database?

A. All tables must have the same number of rows
B. Foreign key values must match existing primary key values or be null
C. All attributes must have default values
D. Tables must be stored on the same disk

Answer: B

Q373. What does the term 'relation' mean in the relational model?

A. A connection between two databases
B. A two-dimensional table of data
C. A physical storage device
D. A type of programming function

Answer: B

Q374. What is entity integrity in a relational database?

A. Every table must have at least one row
B. The primary key of a relation cannot contain null values
C. All attributes must be of the same data type
D. Each table must have exactly one foreign key

Answer: B

Q375. How many candidate keys can a relation have?

A. Exactly one
B. At most two
C. One or more
D. None

Answer: C

Q376. What is the purpose of a composite key?

A. To encrypt data in a column
B. To use two or more attributes together to uniquely identify a tuple
C. To link two databases
D. To store multiple values in one field

Answer: B

Q377. What is a relation schema?

A. The data stored in a table at a given time
B. The definition of a relation including its name and attributes
C. A backup copy of a relation
D. A query that retrieves data from a relation

Answer: B

Q378. In a relational table, can two rows be identical?

A. Yes, any number of duplicate rows are allowed
B. No, each row must be unique
C. Only if the table has no primary key defined
D. Only in temporary tables

Answer: B

Q379. What is the role of a foreign key?

A. To uniquely identify rows in its own table
B. To establish a link between data in two tables
C. To encrypt sensitive data
D. To automatically generate unique IDs

Answer: B

Q380. What does the term 'arity' of a relation refer to?

A. The number of rows in the relation
B. The number of columns (attributes) in the relation
C. The size of the relation in bytes
D. The number of constraints on the relation

Answer: B

Q381. Which relational algebra operation removes duplicate rows from the result?

A. Projection
B. Selection
C. Cartesian product
D. Rename

Answer: A

Q382. What does the selection operation filter?

A. Columns from a relation
B. Rows from a relation based on a condition
C. Duplicate tables
D. Database users

Answer: B

Q383. What does the projection operation do in relational algebra?

A. Selects specific rows from a table
B. Selects specific columns from a relation
C. Joins two tables together
D. Renames a relation

Answer: B

Q384. What is the result of a Cartesian product of two relations?

A. Only matching rows from both relations
B. All possible combinations of rows from both relations
C. The intersection of both relations
D. An empty relation

Answer: B

Q385. For the union operation, what must be true about the two relations?

A. They must have the same name
B. They must be union compatible with the same number and types of attributes
C. They must be stored in the same database
D. They must have the same number of rows

Answer: B

Q386. What does the set difference operation R - S return?

A. All tuples in both R and S
B. Tuples in R that are not in S
C. Tuples in S that are not in R
D. The union of R and S

Answer: B

Q387. Which operation combines two union-compatible relations and removes duplicates?

A. Cartesian product
B. Natural join
C. Union
D. Set difference

Answer: C

Q388. What is the rename operation used for in relational algebra?

A. To delete a relation
B. To change the name of a relation or its attributes
C. To create a copy of a relation
D. To merge two relations

Answer: B

Q389. How many fundamental operations exist in relational algebra?

A. Three
B. Five
C. Six
D. Eight

Answer: C

Q390. What is the intersection of two relations R and S?

A. All tuples in R or S
B. Tuples that appear in both R and S
C. Tuples that appear in R but not S
D. An empty set

Answer: B

Q391. Which SQL clause is used to filter rows based on a condition?

A. ORDER BY
B. GROUP BY
C. WHERE
D. HAVING

Answer: C

Q392. Which SQL statement is used to add new rows to a table?

A. UPDATE
B. INSERT INTO
C. ALTER TABLE
D. CREATE TABLE

Answer: B

Q393. What does the SELECT DISTINCT keyword do?

A. Selects all rows including duplicates
B. Selects only unique rows, removing duplicates
C. Selects the first row only
D. Selects rows in a random order

Answer: B

Q394. Which SQL keyword is used to sort the result set?

A. SORT BY
B. ARRANGE BY
C. ORDER BY
D. FILTER BY

Answer: C

Q395. What does the SQL COUNT() function do?

A. Counts the number of tables in a database
B. Counts the number of rows that match a criteria
C. Counts the number of databases on the server
D. Counts the number of characters in a column

Answer: B

Q396. Which SQL command is used to remove a table from the database?

A. DELETE TABLE
B. REMOVE TABLE
C. DROP TABLE
D. DESTROY TABLE

Answer: C

Q397. What does the SQL UPDATE statement do?

A. Adds new rows to a table
B. Modifies existing data in a table
C. Deletes rows from a table
D. Creates a new table

Answer: B

Q398. Which SQL operator is used to search for a pattern in a column?

A. MATCH
B. LIKE
C. FIND
D. SEARCH

Answer: B

Q399. What does the SQL DELETE statement do?

A. Removes a table structure
B. Removes specific rows from a table
C. Removes a database
D. Removes a column from a table

Answer: B

Q400. Which keyword is used to combine results from two SELECT statements?

A. MERGE
B. COMBINE
C. UNION
D. APPEND

Answer: C

Q401. What is a trigger in SQL?

A. A button in the database interface
B. A special stored procedure that automatically executes when a specified event occurs
C. A type of database index
D. A command to start the database server

Answer: B

Q402. What is an SQL view?

A. A graphical display of the database
B. A virtual table based on the result of a SELECT query
C. A physical copy of a table
D. A type of report

Answer: B

Q403. What is the purpose of the GRANT command in SQL?

A. To create a new table
B. To give users specific access privileges on database objects
C. To delete user accounts
D. To optimize queries

Answer: B

Q404. What does the REVOKE command do in SQL?

A. Reverses a transaction
B. Removes access privileges previously granted to users
C. Restores deleted data
D. Reverts schema changes

Answer: B

Q405. What is an assertion in SQL?

A. A statement that prints debug information
B. A constraint that specifies a condition the database must always satisfy
C. A type of stored procedure
D. A SQL comment

Answer: B

Q406. What is the purpose of the COMMIT command?

A. To start a new transaction
B. To save all changes made during the current transaction permanently
C. To undo changes
D. To lock the database

Answer: B

Q407. What does the ROLLBACK command do?

A. Moves the database to a new server
B. Undoes all changes made during the current transaction
C. Deletes the entire database
D. Creates a backup of the database

Answer: B

Q408. What is a CHECK constraint in SQL?

A. A constraint that checks disk space
B. A constraint that ensures values in a column satisfy a specified condition
C. A constraint that checks user passwords
D. A constraint that validates SQL syntax

Answer: B

Q409. What is the main goal of database normalization?

A. To increase data redundancy
B. To organize data to reduce redundancy and improve data integrity
C. To make queries run faster
D. To add more tables to the database

Answer: B

Q410. What is First Normal Form (1NF)?

A. A table with no foreign keys
B. A table where each column contains only atomic (indivisible) values
C. A table with at least three columns
D. A table stored in the first file

Answer: B

Q411. What is a repeating group in database design?

A. A group of users who access the database repeatedly
B. Multiple values stored in a single column violating 1NF
C. A frequently executed query
D. A backup that runs on a schedule

Answer: B

Q412. What does Second Normal Form (2NF) require?

A. The table must be in 1NF and have no partial dependencies on a composite key
B. The table must have at least two columns
C. All columns must be numeric
D. The table must have exactly two candidate keys

Answer: A

Q413. What is a partial dependency?

A. A dependency on the entire primary key
B. A non-key attribute that depends on only part of a composite primary key
C. A dependency between two foreign keys
D. A dependency that only sometimes holds

Answer: B

Q414. What does Third Normal Form (3NF) eliminate?

A. All foreign keys
B. Transitive dependencies of non-key attributes on the primary key
C. All null values
D. All composite keys

Answer: B

Q415. What type of anomaly occurs when deleting a row causes loss of unrelated data?

A. Insertion anomaly
B. Update anomaly
C. Deletion anomaly
D. Selection anomaly

Answer: C

Q416. Which normal form addresses the issue of repeating groups?

A. Second Normal Form
B. Third Normal Form
C. First Normal Form
D. Boyce-Codd Normal Form

Answer: C

Q417. What does the Atomicity property of a transaction ensure?

A. The transaction is very small
B. Either all operations in the transaction are completed or none are
C. The transaction uses atomic data types
D. The transaction runs on a single processor

Answer: B

Q418. What does the Consistency property of a transaction ensure?

A. All transactions run at the same speed
B. A transaction brings the database from one valid state to another valid state
C. All columns have consistent data types
D. The database is consistent in size

Answer: B

Q419. What does the Isolation property of a transaction ensure?

A. Transactions are stored in isolated tables
B. Concurrent transactions do not interfere with each other
C. The database is isolated from the network
D. Each transaction runs on a separate server

Answer: B

Q420. What does the Durability property of a transaction ensure?

A. The transaction runs for a long duration
B. Once a transaction is committed, its changes persist even after system failures
C. The database hardware is durable
D. The transaction cannot be undone

Answer: B

Q421. What are the possible states of a transaction?

A. Start, Run, Stop
B. Active, Partially Committed, Committed, Failed, Aborted
C. Open, Closed, Locked
D. Read, Write, Delete

Answer: B

Q422. What happens when a transaction is aborted?

A. Its changes are permanently saved
B. All changes made by the transaction are undone
C. The database is shut down
D. Other transactions are paused

Answer: B

Q423. What is a schedule in the context of transactions?

A. A time table for database backups
B. The chronological order in which operations from concurrent transactions are executed
C. A list of database users
D. A plan for database design

Answer: B

Q424. What is a read operation in a transaction?

A. Reading the database schema
B. Retrieving a data item value from the database
C. Reading the transaction log
D. Reading user input

Answer: B

Q425. What is concurrency control in database systems?

A. A method to control the number of database tables
B. A mechanism to manage simultaneous access to the database by multiple transactions
C. A technique to control query speed
D. A tool to manage database backups

Answer: B

Q426. What is a lock in the context of concurrency control?

A. A physical security device for the server
B. A mechanism that restricts access to a data item by other transactions
C. A type of database table
D. A method for encrypting data

Answer: B

Q427. What is a shared lock (S-lock)?

A. A lock shared across multiple databases
B. A lock that allows multiple transactions to read a data item simultaneously
C. A lock that allows both reading and writing
D. A lock that prevents all access

Answer: B

Q428. What is an exclusive lock (X-lock)?

A. A lock available only to the DBA
B. A lock that allows only one transaction to both read and write a data item
C. A lock that blocks all database operations
D. A lock used during database installation

Answer: B

Q429. What is the purpose of the two-phase locking (2PL) protocol?

A. To lock data in two different databases
B. To ensure serializability by dividing a transaction into a growing phase and a shrinking phase
C. To perform two types of locks on each data item
D. To execute transactions in two steps

Answer: B

Q430. In the two-phase locking protocol, what is the growing phase?

A. The phase where the database grows in size
B. The phase where a transaction acquires locks but cannot release any
C. The phase where new tables are created
D. The phase where transactions are added to the schedule

Answer: B

Q431. In the two-phase locking protocol, what is the shrinking phase?

A. The phase where the database shrinks in size
B. The phase where a transaction releases locks but cannot acquire new ones
C. The phase where tables are deleted
D. The phase where the transaction log is trimmed

Answer: B

Q432. What is starvation in the context of concurrency control?

A. When a database runs out of storage
B. When a transaction waits indefinitely because other transactions keep getting priority
C. When the server runs out of memory
D. When the network connection is lost

Answer: B

Q433. What is a lock manager in a DBMS?

A. A person who manages physical locks on the server room
B. A DBMS component that handles lock requests, grants, and releases
C. A software for managing encryption keys
D. A tool for managing database user accounts

Answer: B

Q434. What is a transaction log (journal)?

A. A list of all database users
B. A file that records all changes made to the database for recovery purposes
C. A log of network traffic
D. A list of all SQL queries ever written

Answer: B

Q435. What is a database backup?

A. A secondary database server
B. A copy of the database stored separately for recovery purposes
C. A redundant index on a table
D. An extra user account

Answer: B

Q436. What does the redo operation do during recovery?

A. Re-executes all queries from the beginning
B. Reapplies committed transactions whose changes may not have been written to disk
C. Restarts the database server
D. Recreates all database tables

Answer: B

Q437. What does the undo operation do during recovery?

A. Removes the database entirely
B. Reverses the changes of uncommitted transactions to restore consistency
C. Undoes the last backup
D. Removes the recovery log

Answer: B

Q438. What is the purpose of the write-ahead log rule?

A. To ensure writes happen before reads
B. To ensure log records are written before the corresponding data changes are flushed to disk
C. To write data ahead of the schedule
D. To log all user activities before database access

Answer: B

Q439. What is a system failure in database terminology?

A. A failure in the database design process
B. A failure that loses the contents of main memory but not disk storage
C. A failure that destroys all data permanently
D. A failure in the user interface

Answer: B

Q440. What is a media failure in database terminology?

A. A failure in multimedia content
B. A failure that damages the physical storage medium, potentially losing disk data
C. A failure in the display monitor
D. A failure in the network media

Answer: B

Q441. What is an incremental backup?

A. A backup of the entire database
B. A backup that captures only the changes made since the last backup
C. A backup performed incrementally faster each time
D. A backup that increments the database version

Answer: B

Q442. What is a primary index?

A. The first index created on a table
B. An index built on the primary key of an ordered file
C. An index that can only store integers
D. The most important index in the database

Answer: B

Q443. What is a B-tree?

A. A binary tree used for sorting
B. A balanced search tree where all leaf nodes are at the same level, used for indexing
C. A tree that only stores binary data
D. A tree with exactly two branches

Answer: B

Q444. What is the advantage of using indexes?

A. They reduce storage space
B. They speed up data retrieval operations
C. They make all queries faster without any trade-off
D. They eliminate the need for primary keys

Answer: B

Q445. What is the disadvantage of having too many indexes?

A. Queries become slower to read
B. Insert, update, and delete operations become slower as indexes must be maintained
C. The database becomes read-only
D. Indexes use no additional storage

Answer: B

Q446. How many clustered indexes can a table have?

A. As many as needed
B. Exactly one
C. Exactly two
D. None

Answer: B

Q447. What is query optimization?

A. Writing queries in the shortest way possible
B. The process of selecting the most efficient execution plan for a query
C. Removing unnecessary queries from the database
D. Optimizing the database for more queries

Answer: B

Q448. What is the role of the query parser?

A. To execute the query
B. To check the syntax and validate the SQL query before optimization
C. To store query results
D. To create database tables

Answer: B

Q449. What is a table scan (full table scan)?

A. Scanning the table structure for errors
B. Reading every row in a table to find matching records
C. Scanning only the index
D. A security scan of the table

Answer: B

Q450. Why is using an index generally faster than a full table scan?

A. Indexes are stored on faster disks
B. Indexes narrow down the search to specific records without scanning every row
C. Indexes compress the data for faster reading
D. Indexes always return fewer results

Answer: B

Q451. What does the EXPLAIN command show in SQL?

A. An explanation of SQL syntax
B. The execution plan the DBMS will use for a query
C. The database schema
D. A list of all indexes

Answer: B

Q452. What is a join operation in the context of query optimization?

A. Joining two databases together
B. Combining rows from two or more tables based on related columns
C. Joining user sessions
D. Adding columns to a table

Answer: B

Q453. What is cost-based optimization?

A. Optimizing the monetary cost of database licenses
B. Selecting the execution plan with the lowest estimated resource cost
C. Minimizing the cost of database hardware
D. Optimizing query writing costs

Answer: B

Q454. What is a query tree (query plan tree)?

A. A tree-shaped database
B. A tree structure representing the relational algebra operations of a query
C. A decision tree for choosing queries
D. A directory structure for storing queries

Answer: B

Q455. What are database statistics used for in query optimization?

A. Generating reports about database usage
B. Providing the optimizer with data distribution information to estimate operation costs
C. Tracking user login statistics
D. Measuring network performance

Answer: B

Q456. What is data encryption?

A. Compressing data to save space
B. Converting data into an unreadable format that can only be decoded with a key
C. Backing up data to a remote location
D. Indexing data for faster access

Answer: B

Q457. What is an access control list (ACL) in database security?

A. A checklist for database design
B. A list specifying which users have access to which database objects and what operations they can perform
C. A list of all database connections
D. A list of recently accessed tables

Answer: B

Q458. What is a database audit trail?

A. A hiking trail near the data center
B. A record of all activities performed on the database for monitoring and accountability
C. A trail of deleted data
D. A path for data migration

Answer: B

Q459. What is a database firewall?

A. A physical wall around the server room
B. A security layer that monitors and filters database traffic to prevent unauthorized access
C. A backup power supply
D. A type of database index

Answer: B

Q460. What is data masking?

A. Hiding the database from the network
B. Replacing sensitive data with fictitious but realistic data to protect privacy
C. Compressing data to reduce size
D. Encrypting data at rest

Answer: B

Q461. What is the advantage of data replication?

A. It saves storage space
B. It improves data availability and local read performance
C. It eliminates the need for backups entirely
D. It speeds up write operations

Answer: B

Q462. What is a homogeneous distributed database?

A. A database with the same data at all sites
B. A distributed database where all sites use the same DBMS software
C. A database that stores only one type of data
D. A database with uniform table sizes

Answer: B

Q463. What is a heterogeneous distributed database?

A. A database with mixed data types
B. A distributed database where different sites may use different DBMS software
C. A database with unequal table sizes
D. A database that stores data in different formats

Answer: B

Q464. What is location transparency in distributed databases?

A. The database hides where data is physically stored from users
B. The database reveals the location of all data to users
C. Transparency about the data center location
D. Showing users the physical address of each server

Answer: A

Q465. What is replication transparency?

A. Making replication visible to all users
B. Users are unaware that data is replicated; the system manages copies automatically
C. Transparent replication logs
D. Showing users all copies of data

Answer: B

Q466. Which of the following is a document-based NoSQL database?

A. MySQL
B. PostgreSQL
C. MongoDB
D. Oracle

Answer: C

Q467. Which of the following is an example of a key-value store?

A. MongoDB
B. Redis
C. Neo4j
D. Cassandra

Answer: B

Q468. What is a column-family database?

A. A database organized by families of users
B. A NoSQL database that stores data in columns grouped into column families
C. A database with family-related data only
D. A relational database feature

Answer: B

Q469. Which of the following is a graph database?

A. Redis
B. Cassandra
C. Neo4j
D. HBase

Answer: C

Q470. What is the main advantage of NoSQL databases over relational databases?

A. They always provide ACID guarantees
B. They offer better scalability and flexibility for unstructured or semi-structured data
C. They always have better query capabilities
D. They require less storage

Answer: B

Q471. What is a document in a document-based NoSQL database?

A. A text file on disk
B. A self-contained data record often in JSON or BSON format with flexible schema
C. A PDF stored in the database
D. A database manual

Answer: B

Q472. What does horizontal scaling mean in the context of NoSQL?

A. Adding more columns to tables
B. Adding more servers to handle increased load
C. Making the server hardware more powerful
D. Increasing disk space on one server

Answer: B

Q473. What is vertical scaling?

A. Adding more servers to the cluster
B. Increasing the capacity of a single server by adding more CPU, RAM, or storage
C. Adding more rows to a table
D. Stacking databases vertically

Answer: B

Q474. What is ETL in data warehousing?

A. Electronic Transfer Link
B. Extract, Transform, Load - the process of moving data from sources into the data warehouse
C. Extended Table Logic
D. Enterprise Technology Layer

Answer: B

Q475. What is a dimension table in a data warehouse?

A. A table that stores database dimensions (size)
B. A table containing descriptive attributes used to filter, group, and label fact data
C. A table with three-dimensional data
D. A temporary table

Answer: B

Q476. What is a star schema in data warehousing?

A. A schema shaped like a star on the database diagram
B. A schema with a central fact table connected to multiple dimension tables, forming a star pattern
C. A schema used in astronomy databases
D. A schema with star-rated data

Answer: B

Q477. What is OLAP?

A. Online Automatic Processing
B. Online Analytical Processing - technology for complex analytical queries on multidimensional data
C. Open Link Access Protocol
D. Offline Loading and Processing

Answer: B

Q478. What is a measure in a data warehouse fact table?

A. The physical size of the table
B. A numeric value that can be aggregated, such as sales amount or quantity
C. A unit of measurement for storage
D. A count of dimension tables

Answer: B

Q479. What is a data cube in OLAP?

A. A cubic-shaped storage device
B. A multidimensional representation of data allowing analysis across multiple dimensions
C. A backup cube for data protection
D. A rubik's cube made of data drives

Answer: B

Q480. What is a DEFAULT constraint in SQL?

A. The default database setting
B. A constraint that assigns a default value to a column when no value is specified during insertion
C. A constraint that defaults to NULL
D. The default table in a database

Answer: B

Q481. What is the purpose of the UNIQUE constraint in SQL?

A. To make all rows unique across all columns
B. To ensure that all values in a column or combination of columns are distinct
C. To create a unique database
D. To uniquely identify the database

Answer: B

Q482. What is an unnormalized table?

A. A table without a name
B. A table that contains repeating groups or non-atomic values violating 1NF
C. A table with too many columns
D. A table that is not indexed

Answer: B

Q483. What is the relationship between normalization and data redundancy?

A. Normalization increases data redundancy
B. Normalization reduces data redundancy by organizing data into separate related tables
C. Normalization has no effect on redundancy
D. Normalization eliminates all data from the database

Answer: B

Q484. What is a write operation in a transaction?

A. Writing documentation about the database
B. Modifying the value of a data item in the database
C. Writing a query to the console
D. Creating a new table

Answer: B

Q485. What does it mean when a transaction is in the 'committed' state?

A. The transaction is waiting to start
B. The transaction has completed successfully and its changes are permanently recorded
C. The transaction has been cancelled
D. The transaction is still running

Answer: B

Q486. Why is concurrency control necessary in database systems?

A. To make the database run faster
B. To ensure data consistency and isolation when multiple transactions access data simultaneously
C. To control the number of database users
D. To manage disk space

Answer: B

Q487. What is a full database backup?

A. A backup that only stores schema information
B. A complete copy of the entire database at a specific point in time
C. A backup of the backup system
D. A partial backup of frequently accessed data

Answer: B

Q488. What is the purpose of a log file in database recovery?

A. To log user login attempts
B. To record all database modifications so transactions can be redone or undone during recovery
C. To store query execution plans
D. To track disk space usage

Answer: B

Q489. What is the search key of an index?

A. A physical key to unlock the database
B. The attribute or set of attributes used to look up records in the index
C. A password for accessing indexed data
D. The primary key of the indexed table

Answer: B

Q490. Can a table have multiple non-clustered indexes?

A. No, only one index per table
B. Yes, a table can have multiple non-clustered indexes on different columns
C. Only if the table has no primary key
D. Only two non-clustered indexes are allowed

Answer: B

Q491. What is an index entry composed of?

A. Only the table name
B. A search key value and a pointer to the corresponding data record or block
C. Only the row number
D. A copy of the entire row

Answer: B

Q492. What happens to indexes when a row is inserted into a table?

A. Indexes are unaffected by insertions
B. All relevant indexes must be updated to include entries for the new row
C. Indexes are automatically deleted
D. Only the primary index is updated

Answer: B

Q493. Why might a database optimizer choose not to use an available index?

A. Indexes can never be ignored
B. When the optimizer estimates a full table scan would be faster, such as when most rows match the query
C. The optimizer always uses indexes
D. Indexes are only used for SELECT queries

Answer: B

Q494. What is a query rewrite in the optimization process?

A. Rewriting the query in a different programming language
B. Transforming a query into an equivalent but more efficient form before generating execution plans
C. Manually rewriting queries for better readability
D. Rewriting queries to fix syntax errors

Answer: B

Q495. What is the purpose of a password policy in database security?

A. To make passwords decorative
B. To enforce rules about password strength, expiration, and reuse to prevent unauthorized access
C. To store passwords in plain text
D. To share passwords among users

Answer: B

Q496. What is a privilege in database security?

A. A type of database table
B. A right or permission granted to a user to perform specific operations on database objects
C. A special type of data
D. A database backup method

Answer: B

Q497. What is the difference between system privileges and object privileges?

A. There is no difference
B. System privileges allow operations on the system level while object privileges allow operations on specific database objects
C. System privileges are for hardware while object privileges are for software
D. Object privileges are more powerful

Answer: B

Q498. Why is it important to regularly update database software?

A. To change the database color theme
B. To patch security vulnerabilities and protect against newly discovered threats
C. To increase storage space
D. To delete old data

Answer: B

Q499. What is a distributed database management system (DDBMS)?

A. A DBMS installed on multiple monitors
B. Software that manages a distributed database, making distribution transparent to users
C. A mobile database application
D. A database stored on a USB drive

Answer: B

Q500. What is a site (node) in a distributed database?

A. A website that hosts the database
B. A physical or logical location that has its own local database and processing capability
C. A web page with database information
D. A user workstation

Answer: B

Q501. What is the benefit of distributing a database across multiple sites?

A. It makes the database harder to use
B. It improves reliability, availability, and performance by placing data closer to where it is used
C. It reduces the total amount of data stored
D. It eliminates the need for security

Answer: B

Q502. What is a communication network in a distributed database system?

A. A social network for database users
B. The network infrastructure that connects different database sites and enables data transfer between them
C. A chat system within the database
D. An internal database messaging system

Answer: B

Q503. What is a collection in MongoDB?

A. A group of database servers
B. A grouping of documents, similar to a table in relational databases
C. A collection of SQL queries
D. A backup set

Answer: B

Q504. What is a node in a graph database?

A. A network server
B. An entity or data point represented as a vertex in the graph, such as a person or product
C. A connection between two databases
D. A cluster of documents

Answer: B

Q505. What are the four key characteristics of a data warehouse according to Bill Inmon?

A. Fast, Reliable, Secure, Scalable
B. Subject-oriented, Integrated, Time-variant, and Non-volatile
C. Structured, Indexed, Normalized, and Backed-up
D. Large, Distributed, Real-time, and Encrypted

Answer: B

Q506. What is a data source in the context of data warehousing?

A. The power source for the data center
B. An operational system, external file, or other origin from which data is extracted for the warehouse
C. The database administrator
D. A query result

Answer: B

Q507. What is a dashboard in business intelligence?

A. The front panel of a database server
B. A visual display of key performance indicators and metrics from the data warehouse for decision making
C. A database configuration page
D. A network monitoring tool

Answer: B

Q508. What is the staging area in a data warehouse architecture?

A. A theater stage for database presentations
B. An intermediate storage area where extracted data is cleaned and transformed before loading into the warehouse
C. The main warehouse storage
D. A backup location

Answer: B

Q509. What does it mean to revoke a privilege from a database user?

A. To grant additional privileges
B. To remove a previously granted permission from the user
C. To delete the user account
D. To reset the user password

Answer: B

Q510. What is network latency in the context of distributed databases?

A. The speed of the database processor
B. The time delay for data to travel between sites across the network
C. The size of the database
D. The number of concurrent users

Answer: B

Medium Questions

510 questions

Q511. Which level of data abstraction describes how data is actually stored?

A. View level
B. Logical level
C. External level
D. Physical level

Answer: D

Q512. The three-schema architecture was proposed by:

A. IEEE standards body
B. ISO organization
C. W3C consortium
D. ANSI/SPARC committee

Answer: D

Q513. Which of the following is a disadvantage of DBMS?

A. Increased data redundancy across tables
B. Program-data dependence across modules
C. High initial cost of hardware and software
D. Data inconsistency between applications

Answer: C

Q514. What is metadata in a database context?

A. Temporary data
B. Data about data
C. Deleted data
D. Encrypted data

Answer: B

Q515. Which component of DBMS is responsible for ensuring data consistency?

A. Integrity manager
B. Authorization manager
C. Buffer pool manager
D. File system manager

Answer: A

Q516. A data dictionary stores:

A. Application programs and their code
B. Actual user data stored in tables
C. Metadata about the database structure
D. Backup files created periodically

Answer: C

Q517. Which of the following describes logical data independence?

A. Ability to change applications without affecting the underlying data
B. Ability to change the logical schema without affecting external schemas
C. None of these options correctly describes logical data independence
D. Ability to change physical storage without affecting the logical schema

Answer: B

Q518. What is a database instance?

A. The database management software application
B. The overall architectural design of the database
C. A backup copy of the database stored on disk
D. The collection of data stored at a particular moment

Answer: D

Q519. A database schema is:

A. A data structure used for indexing records
B. A language used for querying the database
C. The overall design or structure of the database
D. The actual data values stored in the tables

Answer: C

Q520. Which of the following is NOT a function of a DBMS?

A. Generating source code for applications
B. Manipulating and querying stored data
C. Managing concurrent transaction access
D. Defining and managing data structures

Answer: A

Q521. In the three-schema architecture, which schema defines the user's view?

A. External schema
B. Internal schema
C. Conceptual schema
D. Physical schema

Answer: A

Q522. Which DBMS component ensures that the database remains in a consistent state after a failure?

A. Authorization manager
B. Buffer manager
C. Recovery manager
D. Query optimizer

Answer: C

Q523. The data definition language (DDL) compiler processes:

A. Application programs written in the host language
B. Schema definitions and stores metadata in the catalog
C. User-submitted queries for data retrieval operations
D. Data insertion commands for populating new tables

Answer: B

Q524. In a two-tier architecture, the application logic resides:

A. Exclusively on the client tier
B. Exclusively on the server tier
C. On a separate dedicated middle tier
D. On the client side or server side

Answer: D

Q525. What is a database catalog?

A. A system database containing metadata about the database structure
B. A comprehensive list of all users who have database access rights
C. An index structure built on primary keys of all related tables
D. A sequential log file that records all committed transactions

Answer: A

Q526. Which architecture is most suitable for web-based database applications?

A. Peer-to-peer architecture
B. Two-tier architecture
C. Single-tier architecture
D. Three-tier architecture

Answer: D

Q527. The runtime database processor executes:

A. Backup and recovery operations only
B. Data definition language statements only
C. Schema compilation processes only
D. Privileged commands and query plans at runtime

Answer: D

Q528. What is the purpose of the DML precompiler?

A. To compile DDL commands and update the system catalog
B. To extract DML commands embedded in application programs
C. To manage concurrency control among active transactions
D. To optimize queries by selecting the best execution plan

Answer: B

Q529. In a parallel database system, the main goal is:

A. Reduced overall storage cost on disk
B. Simplified query writing for users
C. Eliminating the need for any backups
D. Improved performance through parallelism

Answer: D

Q530. Shared-nothing architecture means:

A. All processors share the memory and disk
B. Only disk is shared among all processors
C. Each processor has its own memory and disk
D. Only memory is shared among processors

Answer: C

Q531. In an Enhanced ER (EER) model, specialization is:

A. Adding new relationships between existing entity types
B. Combining multiple entity types into a single entity
C. The process of defining subclasses of an entity type
D. Removing attributes from entities during refinement

Answer: C

Q532. A weak entity in an ER diagram:

A. Cannot be uniquely identified by its own attributes alone
B. Is always fully independent of any other entity type
C. Cannot have any attributes or properties defined on it
D. Has a complete primary key defined on its own attributes

Answer: A

Q533. Total participation in an ER diagram means:

A. The relationship is optional for all entities
B. No entity participates in the defined relationship
C. Only some entities participate in the relationship
D. Every entity must participate in the relationship

Answer: D

Q534. A composite attribute is:

A. An attribute that holds multiple values for one entity
B. An attribute that serves as the key for identification
C. An attribute whose value is derived from other attributes
D. An attribute that can be divided into smaller sub-attributes

Answer: D

Q535. A multivalued attribute is represented in an ER diagram by:

A. A single ellipse
B. A rectangle shape
C. A dashed ellipse
D. A double ellipse

Answer: D

Q536. In an ER diagram, a double rectangle represents:

A. A weak entity
B. An attribute
C. A relationship
D. A strong entity

Answer: A

Q537. Generalization in EER modeling is:

A. Adding new attributes to an existing entity class
B. Removing a class entirely from the ER model
C. Splitting a single class into multiple subclasses
D. Combining subclasses into a generalized superclass

Answer: D

Q538. A derived attribute is one whose value:

A. Must be entered manually by users
B. Cannot change once it is set initially
C. Is always NULL in every record stored
D. Can be computed from other attributes

Answer: D

Q539. The cardinality ratio in an ER model specifies:

A. The overall size of the database measured in stored records
B. The number of attributes defined on a single entity type
C. The total number of entities present in the entire ER model
D. The number of relationship instances an entity can participate in

Answer: D

Q540. An object-oriented data model supports:

A. Only tree structures and parent-child links
B. Only tables, rows, and column structures
C. Only flat files and sequential access paths
D. Encapsulation, inheritance, and polymorphism

Answer: D

Q541. A candidate key is:

A. Any arbitrary attribute present in the table
B. The same concept as a foreign key reference
C. A minimal superkey that uniquely identifies tuples
D. Always the first column defined in the schema

Answer: C

Q542. A superkey is:

A. Any set of attributes that uniquely identifies tuples
B. A key that is permitted to contain NULL values
C. Always limited to a single attribute in the table
D. The same concept as a foreign key in a relation

Answer: A

Q543. Referential integrity constraint ensures that:

A. A foreign key value must match a primary key value in the referenced table or be NULL
B. All columns across every table in the database must have the same compatible data type
C. All tables in the database schema must have exactly the same total number of data rows
D. Primary key values can be duplicated across different rows within the same data table

Answer: A

Q544. Entity integrity constraint states that:

A. Foreign keys must always be unique in tables
B. Tables must have at least one row of data
C. All attributes in every row must have values
D. No primary key attribute can have a NULL value

Answer: D

Q545. An alternate key is:

A. A composite key made up of two or more attributes
B. A candidate key that is not chosen as the primary key
C. A key that is permitted to contain NULL value entries
D. The same concept as a foreign key in another table

Answer: B

Q546. A composite key is:

A. A key with only one single attribute value
B. A foreign key referencing another table
C. A key that allows NULL values in rows
D. A key consisting of two or more attributes

Answer: D

Q547. Which of the following is NOT a relational integrity constraint?

A. Entity integrity
B. Sequential constraint
C. Referential integrity
D. Domain constraint

Answer: B

Q548. A relation schema defines:

A. Only the primary key definition of the relation in question
B. The structure of a relation including attribute names and domains
C. The actual data values currently stored within the relation
D. Only the foreign keys that reference other related tables

Answer: B

Q549. In relational algebra, the Cartesian product of two relations R and S produces:

A. All possible combinations of tuples from R and S
B. The intersection of tuples in both R and S
C. Only tuples that match on common attributes
D. The union of all tuples from both R and S

Answer: A

Q550. A relation in 1NF requires that:

A. The table has only one column of values
B. There are no foreign keys in the relation
C. All values are strictly numeric in type
D. All attribute values are atomic (indivisible)

Answer: D

Q551. The theta join of relations R and S is:

A. A rename operation applied to change the relation name
B. A Cartesian product followed by selection with a condition
C. A set union combining all tuples from two given relations
D. A simple projection on selected attributes from the relation

Answer: B

Q552. An equi-join is a special case of theta join where:

A. No condition is applied to the join result
B. The result of the join is always an empty set
C. The condition uses only inequality comparisons
D. The condition uses only equality comparisons

Answer: D

Q553. The division operation R ÷ S returns:

A. The Cartesian product of both relations R and S fully
B. Tuples in R that are associated with every tuple in S
C. All tuples from relation S regardless of the condition
D. All tuples from relation R regardless of the condition

Answer: B

Q554. Left outer join includes:

A. All tuples from the right relation and matching from the left side
B. Only tuples that have matching values in both of the joined tables
C. No tuples at all if there is no matching condition between tables
D. All tuples from the left relation and matching tuples from the right

Answer: D

Q555. In relational algebra, which operation can be derived from the fundamental operations?

A. Projection
B. Intersection
C. Division
D. Selection

Answer: B

Q556. The semi-join of R and S returns:

A. All tuples from both relations R and S combined
B. The complete Cartesian product of R and S result
C. Only the tuples from relation S not in relation R
D. Tuples from R that have matching tuples in S

Answer: D

Q557. Anti-join of R and S returns:

A. All tuples from R regardless of matching in S
B. The complete Cartesian product of R and S both
C. Tuples from R that have no matching tuples in S
D. All tuples that match between relations R and S

Answer: C

Q558. Relational calculus is:

A. A procedural query language that specifies how to retrieve data
B. A data definition language for creating tables and constraints
C. A physical storage language for defining data file structures
D. A non-procedural query language that describes what to retrieve

Answer: D

Q559. Tuple relational calculus uses:

A. Tuple variables that range over relations
B. Domain variables that range over values
C. Only table names without any variables
D. No variables of any kind are required

Answer: A

Q560. Domain relational calculus uses:

A. SQL queries written in standard syntax format
B. No variables of any kind within the expression
C. Domain variables that range over attribute domains
D. Tuple variables that range over entire relations

Answer: C

Q561. Which SQL clause is used to filter groups in a GROUP BY query?

A. WHERE
B. HAVING
C. ORDER BY
D. LIMIT

Answer: B

Q562. An INNER JOIN returns:

A. All rows from the left table with NULLs for right
B. Only rows that have matching values in both tables
C. All rows from the right table with NULLs for left
D. All rows from both the left and the right tables

Answer: B

Q563. Which aggregate function returns the number of rows?

A. COUNT()
B. SUM()
C. AVG()
D. MAX()

Answer: A

Q564. A LEFT JOIN returns:

A. The Cartesian product of both tables without any condition
B. All rows from the right table and matching from the left
C. Only rows that have matching values in both joined tables
D. All rows from the left table and matching rows from the right

Answer: D

Q565. Which SQL constraint ensures a column cannot have NULL values?

A. CHECK
B. UNIQUE
C. NOT NULL
D. DEFAULT

Answer: C

Q566. The LIKE operator is used with which wildcard characters?

A. # and @
B. % and _
C. * and ?
D. & and ^

Answer: B

Q567. What does the IN operator do in SQL?

A. Checks if a value matches any value in a list
B. Performs an inner join between two given tables
C. Inserts data into a specified target table row
D. Indexes a column for faster search performance

Answer: A

Q568. Which SQL clause limits the number of rows returned?

A. LIMIT (or TOP in SQL Server)
B. GROUP BY (aggregation clause)
C. HAVING (group filter clause)
D. WHERE (filter condition clause)

Answer: A

Q569. A subquery is:

A. A query that modifies existing data
B. A query that creates new table objects
C. A stored procedure in the database
D. A query nested inside another query

Answer: D

Q570. The BETWEEN operator selects values:

A. Within a given inclusive range
B. Outside a specified range only
C. That are NULL valued entries
D. Equal to a single specific value

Answer: A

Q571. A materialized view differs from a regular view in that it:

A. Cannot be refreshed or updated at all
B. Physically stores the result set on disk
C. Is always virtual and never materialized
D. Does not use SELECT queries for data

Answer: B

Q572. Window functions in SQL are used to:

A. Delete rows from tables based on specified filtering conditions
B. Create new tables by defining their schema and column structures
C. Perform calculations across a set of rows related to the current row
D. Modify table structure by adding or removing column definitions

Answer: C

Q573. The ROW_NUMBER() function in SQL:

A. Deletes duplicate rows from the result set of the current query
B. Returns the primary key value from the underlying data table
C. Assigns a unique sequential number to each row within a partition
D. Counts the total number of rows stored in the entire data table

Answer: C

Q574. A Common Table Expression (CTE) is:

A. A trigger that fires automatically in response to a data change
B. A stored procedure that encapsulates reusable database SQL logic
C. A permanent table that persists after the query has been executed
D. A temporary named result set used within a single SQL statement

Answer: D

Q575. The RANK() function differs from ROW_NUMBER() in that:

A. RANK() never produces ties in the ranking of any result set rows
B. RANK() assigns the same rank to ties and skips subsequent ranks
C. ROW_NUMBER() handles ties by assigning them the same rank value
D. They are identical and produce exactly the same results always

Answer: B

Q576. A recursive CTE is used for:

A. Querying hierarchical or recursive data structures
B. Creating indexes on one or more table columns
C. Optimizing joins between tables in the query
D. Dropping tables from the database permanently

Answer: A

Q577. The SAVEPOINT statement in SQL:

A. Commits the transaction and makes all changes fully permanent
B. Starts a new transaction and initializes all session variables
C. Creates a point within a transaction to which you can roll back
D. Drops a table and removes it permanently from the data schema

Answer: C

Q578. An INSTEAD OF trigger:

A. Executes after the triggering action ends
B. Never executes regardless of the trigger
C. Executes before the triggering action runs
D. Executes in place of the triggering action

Answer: D

Q579. The MERGE statement in SQL combines:

A. Only DDL operations such as CREATE, ALTER, and DROP tables
B. Only INSERT and DELETE operations without any UPDATE logic
C. Only SELECT operations for querying data from source tables
D. INSERT, UPDATE, and DELETE operations based on a condition

Answer: D

Q580. What is a sequence in SQL?

A. A type of trigger that fires on data modification
B. An object that generates sequential numeric values
C. An ordered view of data filtered by a set condition
D. A sorted table arranged by a specific column value

Answer: B

Q581. Boyce-Codd Normal Form (BCNF) is stricter than 3NF because:

A. Every determinant must be a candidate key
B. It allows partial dependencies on some keys
C. It requires no functional dependencies at all
D. It requires only atomic values in each field

Answer: A

Q582. A relation in BCNF is always in:

A. 4NF
B. 1NF
C. 5NF
D. 3NF

Answer: D

Q583. A multivalued dependency X !'!' Y exists when:

A. Y is a primary key attribute that uniquely identifies every tuple
B. A set of values of Y is determined by X, independent of other attributes
C. X functionally determines Y through a standard functional dependency
D. No dependency exists between the attributes X and Y in the relation

Answer: B

Q584. Fourth Normal Form (4NF) requires:

A. No functional dependencies of any kind in the schema
B. BCNF and no non-trivial multivalued dependencies
C. All attributes serve as keys in the relation structure
D. Only 1NF without any additional constraints applied

Answer: B

Q585. A lossless join decomposition ensures that:

A. Some data is always permanently lost during the decomposition process
B. Joining the decomposed relations reproduces the original relation exactly
C. Extra spurious tuples are generated when joining decomposed relations
D. Only key attributes are preserved in the decomposed relation fragments

Answer: B

Q586. Which normal form deals with join dependencies?

A. Third Normal Form (3NF)
B. First Normal Form (1NF)
C. Second Normal Form (2NF)
D. Fifth Normal Form (5NF)

Answer: D

Q587. Armstrong's axiom of reflexivity states:

A. If X determines Y, then XZ determines YZ for any Z
B. If Y is a subset of X, then X determines Y
C. If X determines Y and Y determines Z, then X determines Z
D. If X determines Y, then Y also determines X always

Answer: B

Q588. Armstrong's axiom of augmentation states:

A. If X determines Y and Y determines Z, then X determines Z
B. If X determines Y, then Y also determines X in reverse
C. If X determines Y, then XZ determines YZ for any attribute Z
D. If Y is a subset of X, then X functionally determines Y

Answer: C

Q589. The canonical (minimal) cover of a set of FDs:

A. Is always empty regardless of the original dependency set
B. Contains all possible FDs derivable from the original set
C. Has maximum redundancy among all its functional dependencies
D. Has no redundant dependencies and no extraneous attributes

Answer: D

Q590. A dependency-preserving decomposition ensures that:

A. All original functional dependencies can be enforced without joining decomposed tables
B. All tables in the decomposition must have exactly the same schema and attributes
C. Joins are never needed between any of the resulting decomposed relation fragments
D. Some functional dependencies are always lost during the process of decomposition

Answer: A

Q591. A serial schedule is one where:

A. Only one transaction is allowed in the database system total
B. Operations are randomly ordered across multiple transactions
C. Transactions execute one after another without interleaving
D. Transactions execute concurrently with operations interleaved

Answer: C

Q592. A serializable schedule is:

A. A schedule whose result is equivalent to some serial schedule
B. A schedule where all transactions run simultaneously at once
C. A schedule that is always a serial schedule by definition
D. A schedule with no transactions present within its timeline

Answer: A

Q593. Conflict serializability is tested using:

A. Hash table (bucket-based structure)
B. B-tree index (balanced tree structure)
C. ER diagram (entity relationship model)
D. Precedence graph (serialization graph)

Answer: D

Q594. Two operations conflict if:

A. They belong to different transactions, access the same data item, and at least one is a write
B. They are both read operations accessing the same data item from within different transactions
C. They belong to the same transaction and access different data items within the transaction
D. They access different data items from different transactions without any write operations

Answer: A

Q595. A dirty read occurs when:

A. No read operation occurs between any concurrent transactions
B. A transaction reads data that it has written by itself before
C. A transaction reads data written by an uncommitted transaction
D. A transaction reads data that has already been fully committed

Answer: C

Q596. A non-repeatable read occurs when:

A. No read operation is performed during the entire transaction
B. All reads from a transaction return the same consistent value
C. A transaction reads a value once and never accesses it again
D. A transaction reads the same item twice and gets different values

Answer: D

Q597. A phantom read occurs when:

A. A transaction re-executes a query and finds new rows added by another transaction
B. A transaction reads no rows at all from the table during its entire execution
C. A transaction creates a table and populates it with data from another data source
D. A transaction deletes rows from the table and those changes are then committed

Answer: A

Q598. The READ UNCOMMITTED isolation level allows:

A. No anomalies of any kind during transaction processing
B. Dirty reads, non-repeatable reads, and phantom reads
C. Only phantom reads but prevents all other read anomalies
D. Only non-repeatable reads but prevents all dirty reads

Answer: B

Q599. The SERIALIZABLE isolation level provides:

A. Complete isolation from other transactions
B. Only dirty read prevention mechanism
C. Faster performance than other levels
D. No isolation between any transactions

Answer: A

Q600. A cascading rollback occurs when:

A. No rollback is needed because every transaction completes without any failures
B. Only one transaction rolls back while all other transactions continue normally
C. One transaction's rollback forces other dependent transactions to also roll back
D. All transactions commit successfully without any errors during their execution

Answer: C

Q601. Strict Two-Phase Locking requires that:

A. No locks are needed for any transaction under this locking protocol
B. Only shared locks are used and exclusive locks are never acquired
C. All exclusive locks are held until the transaction commits or aborts
D. Locks can be released at any time during the transaction processing

Answer: C

Q602. Rigorous Two-Phase Locking requires:

A. Locks are released immediately after each individual use
B. All locks (shared and exclusive) are held until commit or abort
C. Only exclusive locks are held until the transaction commits
D. No growing phase exists in this locking protocol at all

Answer: B

Q603. A wait-die deadlock prevention scheme:

A. Allows younger transactions to wait; older ones are rolled back
B. Allows older transactions to wait; younger ones are rolled back
C. Prevents all waiting by immediately aborting every transaction
D. Kills all transactions to resolve and prevent deadlock situations

Answer: B

Q604. A wound-wait deadlock prevention scheme:

A. All transactions wait without any preemption at all now
B. No preemption occurs under any circumstances in the system
C. Older transactions preempt younger ones; younger ones wait
D. Younger transactions preempt older ones; older ones wait

Answer: C

Q605. A deadlock detection algorithm uses:

A. A hash table index
B. A wait-for graph
C. An ER diagram model
D. A B-tree structure

Answer: B

Q606. Timestamp-based concurrency control assigns:

A. Random access to data items without any order
B. Locks to all data items accessed by transactions
C. A unique timestamp to each transaction for ordering
D. Priority based on the size of each transaction

Answer: C

Q607. In timestamp ordering, if a transaction tries to write a data item with a later read timestamp:

A. The data item gets deleted
B. The transaction is rolled back
C. The read timestamp is updated
D. The write proceeds normally

Answer: B

Q608. Optimistic concurrency control assumes that:

A. Conflicts are frequent and locks must always be acquired
B. All transactions conflict and none can proceed at once
C. No validation is needed at any point during processing
D. Conflicts are rare, so validation is done at commit time

Answer: D

Q609. The three phases of optimistic concurrency control are:

A. Start, Process, End
B. Read, Validation, Write
C. Read, Write, Commit
D. Lock, Execute, Unlock

Answer: B

Q610. Lock granularity refers to:

A. The number of transactions in the system
B. The type of query being currently executed
C. The size of the data item that is locked
D. The speed of acquiring locks on the item

Answer: C

Q611. Deferred update (no-undo/redo) recovery technique:

A. Requires neither undo nor redo operations during the recovery process
B. Writes changes immediately and may need undo for uncommitted changes
C. Does not write changes to disk until commit, so undo is never needed
D. Does not use a log file for recording any of the transaction changes

Answer: C

Q612. Immediate update recovery technique:

A. Does not use a log file for any of the transaction record entries
B. Writes changes only after commit, requiring only redo operations
C. Requires only redo operations during the crash recovery process
D. Writes changes to disk before commit, requiring both undo and redo

Answer: D

Q613. Shadow paging recovery technique:

A. Maintains two page tables: current and shadow
B. Requires checkpoints at all times to function
C. Uses only a log file for the recovery process
D. Cannot handle any system crashes that may occur

Answer: A

Q614. After a crash, which transactions need to be redone?

A. All transactions regardless of their commit state now
B. Committed transactions whose changes may not be on disk
C. Uncommitted transactions that were active during crash
D. Only aborted transactions that failed before the crash

Answer: B

Q615. After a crash, which transactions need to be undone?

A. All transactions regardless of their state at the time of the crash
B. No transactions need to be undone during the recovery after a crash
C. Transactions that were active (not committed) at the time of crash
D. Committed transactions that completed successfully before the crash

Answer: C

Q616. The ARIES recovery algorithm stands for:

A. Automated Recovery and Integrity Enforcement System
B. Advanced Recovery and Index Enhancement System
C. Algorithm for Redo and Integrity Evaluation System
D. Algorithm for Recovery and Isolation Exploiting Semantics

Answer: D

Q617. ARIES uses which three phases during recovery?

A. Scan, Parse, Execute
B. Analysis, Redo, Undo
C. Lock, Execute, Unlock
D. Read, Write, Commit

Answer: B

Q618. A log sequence number (LSN) is:

A. The total number of active transactions
B. A type of lock used in concurrency control
C. The size of the log file on stable storage
D. A unique identifier for each log record

Answer: D

Q619. Fuzzy checkpointing differs from standard checkpointing in that:

A. No log entry is written to record the occurrence of the checkpoint
B. The database does not need to stop processing during the checkpoint
C. All processing must stop completely during the checkpoint operation
D. It occurs only once during the entire lifetime of the database

Answer: B

Q620. Physiological logging means:

A. Log records describe changes at the page level with logical operations within the page
B. Only logical operation names are logged without any reference to the physical pages
C. Only physical byte-level changes are logged for each modification to the data pages
D. No logging occurs at all for any modification made to the database during processing

Answer: A

Q621. In a B+ tree of order n, an internal node can have at most:

A. n+1 keys in the node
B. n keys and n children
C. n children and n-1 keys
D. 2n children in total

Answer: C

Q622. The height of a B+ tree with n records and order p is approximately:

A. p*log(n)
B. n * p
C. log_p(n)
D. n / p

Answer: C

Q623. Static hashing has which disadvantage?

A. Fixed number of buckets may lead to overflow as data grows
B. It requires B+ trees to function as the base structure
C. It is too slow for any practical database use case at all
D. It cannot handle any data stored in the database tables

Answer: A

Q624. Extendible hashing handles growth by:

A. Deleting old records to make room for new data entries
B. Using a B+ tree instead of hashing for the index data
C. Increasing the number of hash functions used for keys
D. Doubling the directory and splitting buckets as needed

Answer: D

Q625. Linear hashing handles growth by:

A. Using a fixed number of buckets without any changes
B. Splitting all buckets simultaneously at the same time
C. Splitting buckets in a linear order, one at a time
D. Never splitting any buckets regardless of the load

Answer: C

Q626. A composite index is an index on:

A. A stored procedure reference
B. Two or more columns combined
C. A view of the base table data
D. A single column of the table

Answer: B

Q627. A covering index is one that:

A. Indexes all tables in the database regardless of whether they need indexing
B. Has no columns included in it and serves no purpose for query optimization
C. Covers only the primary key column and does not include any other attributes
D. Contains all columns needed to satisfy a query without accessing the base table

Answer: D

Q628. A bitmap index is most efficient for:

A. Primary key columns that uniquely identify each row
B. Columns with unique values that never repeat at all
C. Columns with high cardinality (many distinct values)
D. Columns with low cardinality (few distinct values)

Answer: D

Q629. An index-organized table (IOT) stores:

A. Data separately from all index structures created
B. Data only in memory without disk-based persistence
C. Data directly in the index structure (B+ tree)
D. Data in flat files without any indexing structures

Answer: C

Q630. The term fan-out in a B+ tree refers to:

A. The number of records stored in the entire data file
B. The maximum number of pointers (children) in a node
C. The number of leaf nodes at the bottom of the tree
D. The total height of the tree from root to leaf level

Answer: B

Q631. Sort-Merge Join works by:

A. Using a full table scan on both relations without any index use
B. Sorting both relations on the join attribute and then merging them
C. Creating a hash table on one relation and probing with the other
D. Using nested loops to compare every tuple in both the relations

Answer: B

Q632. Hash Join works by:

A. Nested iteration comparing every pair of tuples from relations
B. Sorting both relations on the join attribute and merging results
C. Using indexes on both tables to find matching join attributes
D. Building a hash table on one relation and probing with the other

Answer: D

Q633. An index nested loop join uses:

A. No index at all on either of the two relations being joined here
B. An index on the inner relation to find matching tuples efficiently
C. An index on the outer relation only for scanning its tuple values
D. A full scan of both relations without using any index structures

Answer: B

Q634. Pushing selections down in a query tree:

A. Increases the size of intermediate results
B. Removes selections entirely from the plan
C. Has no effect on the intermediate results
D. Reduces the size of intermediate results

Answer: D

Q635. Materialization in query processing means:

A. Not storing any intermediate results and processing everything in one single pass
B. Using only indexes to process the query without accessing the base table data
C. Computing and storing the result of each operation before passing it to the next
D. Only computing the final result without evaluating any intermediate operations

Answer: C

Q636. Pipelining in query processing means:

A. Processing operations one at a time in a strict sequential fashion
B. Storing all intermediate results on disk before the next operation
C. Passing tuples from one operation to the next as they are produced
D. Using only a single CPU core for all database query computations

Answer: C

Q637. The catalog statistics used by a cost-based optimizer include:

A. Only the table name without any other statistical information
B. Only the primary key definition of each table in the catalog
C. Number of tuples, number of distinct values, number of blocks
D. Only column names without cardinality or distribution details

Answer: C

Q638. Heuristic optimization rules include:

A. Perform joins before selections and projections always
B. Always use Cartesian products first before other joins
C. Perform selections and projections as early as possible
D. Avoid using indexes for any data retrieval operations

Answer: C

Q639. The join ordering problem in query optimization refers to:

A. Ordering rows within a single table by a specific column
B. Ordering columns in a table based on their data type size
C. Determining the most efficient order to join multiple tables
D. Sorting query results in ascending or descending sequence

Answer: C

Q640. Dynamic programming is used in query optimization for:

A. Creating tables and defining column data types
B. Parsing SQL queries into abstract syntax trees
C. Managing locks in the concurrency control layer
D. Finding the optimal join order for multiple tables

Answer: D

Q641. Mandatory Access Control (MAC) is based on:

A. Security labels and clearance levels assigned to subjects and objects
B. User discretion in deciding who can access which database objects
C. No access control at all with every user having full open access
D. Role assignments only without any additional security label checks

Answer: A

Q642. Discretionary Access Control (DAC) allows:

A. Data owners to grant and revoke access to other users
B. No access control with all data open to every user
C. Access based on security clearance only in the system
D. Only the DBA to control access to all database objects

Answer: A

Q643. Role-Based Access Control (RBAC) assigns privileges to:

A. Individual users only without grouping
B. Queries only without role management
C. Tables only without any user control
D. Roles, which are then assigned to users

Answer: D

Q644. The Bell-LaPadula model enforces:

A. Only read access controls without write restrictions
B. No read-up and no write-down rules for confidentiality
C. No rules at all for controlling information access
D. Only write access controls without read restrictions

Answer: B

Q645. Data masking in database security:

A. Encrypts entire databases for secure data storage
B. Deletes all data from the tables in the database
C. Replaces sensitive data with realistic but fake data
D. Creates copies of data for backup and recovery use

Answer: C

Q646. A view can provide security by:

A. Deleting sensitive records from the database tables
B. Encrypting all data stored in the underlying tables
C. Creating new tables with restricted column schemas
D. Restricting which columns and rows a user can see

Answer: D

Q647. Virtual Private Database (VPD) provides:

A. Only audit logging without any enforcement of access policies
B. Column-level encryption only without any row-level restrictions
C. No security features of any kind for the database application
D. Row-level security by automatically adding conditions to queries

Answer: D

Q648. Transparent Data Encryption (TDE) encrypts:

A. Query results only sent to clients
B. Log files only on the disk storage
C. Data in transit only on the network
D. Data at rest on the storage media

Answer: D

Q649. Database activity monitoring (DAM) is used to:

A. Optimize storage allocation across available disk partitions
B. Speed up queries by caching frequently accessed data in memory
C. Create backups of the database on a scheduled periodic basis
D. Monitor and record database activity in real-time for security

Answer: D

Q650. The GRANT OPTION in SQL allows:

A. Creating tables and views in the database schema at will
B. Revoking all privileges from every user in the database
C. Deleting databases and all their data from the server
D. A grantee to further grant the same privilege to other users

Answer: D

Q651. The two-phase commit (2PC) protocol ensures:

A. Only read operations complete
B. No atomicity guarantees at all
C. Only local transaction atomicity
D. Atomicity of distributed transactions

Answer: D

Q652. In the 2PC protocol, the coordinator first sends:

A. A prepare (vote) message to all participants
B. A commit message directly to all the sites
C. An abort message to cancel the transaction
D. A query to all participants for data first

Answer: A

Q653. If any participant votes No in 2PC, the coordinator:

A. Ignores the vote and proceeds with commit
B. Sends an abort message to all participants
C. Retries the transaction automatically again
D. Sends a commit message to all participants

Answer: B

Q654. A distributed query involves:

A. No data access from any site at all
B. Only local data access on one site
C. Accessing data from a single table
D. Accessing data from multiple sites

Answer: D

Q655. The CAP theorem states that a distributed system can guarantee at most:

A. Two of three: Consistency, Availability, Partition Tolerance
B. None of these properties can be guaranteed in any system
C. Only one property out of the three possible in the theorem
D. All three properties at the same time without any trade-offs

Answer: A

Q656. Semi-join optimization in distributed queries:

A. Increases data transfer by sending all records across the link
B. Uses full Cartesian products between all fragments in the join
C. Reduces data transfer by sending only relevant join attributes
D. Eliminates all joins from the distributed query execution plan

Answer: C

Q657. Mixed fragmentation combines:

A. No fragmentation is used at all
B. Only vertical fragmentation types
C. Horizontal and vertical fragmentation
D. Only horizontal fragmentation types

Answer: C

Q658. Full replication means:

A. Every site has a complete copy of the entire database
B. No replication exists at any site in the system at all
C. Only one site has the data and others have no copies
D. Data is partially replicated at only a few of sites

Answer: A

Q659. The blocking problem in 2PC occurs when:

A. No failures occur during the execution of the two-phase commit protocol
B. The coordinator fails after sending prepare but before sending the decision
C. All participants commit successfully without any failures in the protocol
D. The network is fast and reliable with no message delays or lost packets

Answer: B

Q660. Distributed deadlock detection is more complex because:

A. All transactions are local and never span multiple sites
B. Deadlocks can involve transactions at multiple sites
C. Deadlocks never occur in distributed database systems
D. Only one site has transactions running at any time

Answer: B

Q661. The BASE properties in NoSQL stand for:

A. Batch Analysis, Stable Environment for data processing
B. Binary Access, Sequential Evaluation of stored records
C. Basic Access, Secure Entry with authentication support
D. Basically Available, Soft state, Eventually consistent

Answer: D

Q662. NoSQL databases typically favor:

A. Strong consistency over availability in all distributed environments
B. Availability and partition tolerance over strong consistency (AP in CAP)
C. Neither availability nor consistency for distributed data systems
D. Only consistency without any consideration for availability at all

Answer: B

Q663. Sharding in NoSQL databases refers to:

A. Compressing data to reduce storage on each server node
B. Horizontally partitioning data across multiple servers
C. Vertically splitting a table into separate column groups
D. Encrypting data before storing it on distributed servers

Answer: B

Q664. MongoDB stores data in which format?

A. CSV (Comma Separated)
B. Plain text (Flat file)
C. XML (Markup Language)
D. BSON (Binary JSON)

Answer: D

Q665. In MongoDB, a collection is analogous to:

A. A column in a relational database
B. A database in a relational system
C. A table in a relational database
D. A row in a relational database

Answer: C

Q666. MapReduce in NoSQL is used for:

A. Managing transactions and ensuring ACID properties on the data
B. Creating indexes on columns for improving data retrieval speed
C. Single-record queries for retrieving individual data items fast
D. Processing large datasets in parallel across distributed systems

Answer: D

Q667. Which NoSQL database type is best for social network relationship queries?

A. Key-value store
B. Document database
C. Column-family store
D. Graph database

Answer: D

Q668. Denormalization in NoSQL is common because:

A. It reduces storage requirements on each server in cluster
B. It enforces referential integrity between related records
C. It avoids expensive joins and improves read performance
D. It simplifies updates by reducing the number of writes

Answer: C

Q669. Apache HBase is built on top of:

A. MongoDB
B. Redis
C. Hadoop HDFS
D. MySQL

Answer: C

Q670. Eventual consistency in NoSQL means:

A. No consistency is guaranteed between replicas at any point in time whatsoever
B. Data is never consistent across replicas regardless of the time that passes
C. All reads always return the latest written value from any replica in the cluster
D. All replicas will converge to the same state given enough time without new updates

Answer: D

Q671. The snowflake schema extends the star schema by:

A. Removing dimension tables from the schema completely
B. Adding more fact tables to handle additional measures
C. Normalizing dimension tables into sub-dimension tables
D. Denormalizing all tables into one single flat structure

Answer: C

Q672. Roll-up in OLAP means:

A. Drilling down to see more detailed level of data
B. Selecting a specific dimension value for filtering
C. Aggregating data to a higher level of summarization
D. Rotating the data cube axes for a different view

Answer: C

Q673. Drill-down in OLAP means:

A. Rotating the data cube to view a different perspective
B. Navigating from summarized data to more detailed data
C. Aggregating data to higher levels of summarization
D. Removing dimensions from the cube for simplification

Answer: B

Q674. Slice operation in OLAP:

A. Selects a single value for one dimension, creating a sub-cube
B. Aggregates all dimensions into one single summary result
C. Selects multiple values for multiple dimensions simultaneously
D. Rotates the cube to provide an alternative data perspective

Answer: A

Q675. Dice operation in OLAP:

A. Selects a single dimension value for data filtering
B. Rotates the cube axes for an alternative data view
C. Selects specific values for two or more dimensions
D. Aggregates data across all available dimension types

Answer: C

Q676. Pivot (rotate) in OLAP:

A. Deletes a dimension from the cube structure entirely
B. Rotates the data axes to provide an alternative view
C. Merges cubes from different source data warehouses
D. Adds new data records into the existing data cube

Answer: B

Q677. MOLAP (Multidimensional OLAP) stores data in:

A. A multidimensional array (cube) structure
B. Key-value stores for fast data retrieval
C. Flat files stored on the file system disk
D. Relational tables with rows and columns

Answer: A

Q678. ROLAP (Relational OLAP) stores data in:

A. Relational database tables
B. Flat files on disk storage
C. NoSQL database documents
D. Multidimensional arrays

Answer: A

Q679. Association rule mining discovers:

A. Database schemas and their normalization structure in the system
B. Index structures and their configuration for faster data access
C. Relationships between items (e.g., items frequently bought together)
D. SQL queries and their execution plans for optimized performance

Answer: C

Q680. The Apriori algorithm is used in data mining for:

A. Finding frequent itemsets and association rules
B. Classification of records into known classes
C. Regression analysis on continuous value data
D. Clustering data into groups of similar items

Answer: A

Q681. What distinguishes logical data independence from physical data independence?

A. Logical is used in NoSQL while physical is used in relational databases only
B. Logical shields from schema changes while physical shields from storage changes
C. Logical handles storage changes while physical handles schema modifications
D. Logical applies to queries while physical applies to transactions exclusively

Answer: B

Q682. Which mapping exists between the external and conceptual schema levels?

A. External-internal mapping
B. External-conceptual mapping
C. Schema-instance mapping
D. Conceptual-internal mapping

Answer: B

Q683. In the three-schema architecture, the conceptual level describes:

A. Individual user views of specific data subsets
B. How data is physically stored on disk devices
C. The overall logical structure of the database
D. The network protocols used for data transfer

Answer: C

Q684. Which type of database language handles data access permissions?

A. Data Control Language handles permissions
B. Data Definition Language handles permissions
C. Transaction Control Language handles them
D. Data Manipulation Language handles them

Answer: A

Q685. What is the difference between a database schema and a database state?

A. Schema is the data values while state is the structure definition
B. Schema applies to queries while state applies to transactions only
C. Schema is the structure definition while state is the current data
D. Schema is temporary storage while state is permanent data storage

Answer: C

Q686. Which of the following best describes program-data independence?

A. Program logic is separated from the data structure definitions
B. Programs and data are stored on the same physical disk drives
C. Data files are independent of the operating system platform used
D. Programs automatically generate data without user intervention

Answer: A

Q687. What is the self-describing nature of a database system?

A. The system generates documentation without any user involvement
B. The system repairs itself when errors are automatically detected
C. The system contains both data and its own metadata description
D. The system creates backups at regular predetermined time intervals

Answer: C

Q688. Which database model was the first to be widely commercially adopted?

A. Hierarchical data model
B. Object-oriented model
C. Network data model
D. Relational data model

Answer: A

Q689. What is the purpose of a database catalog?

A. To compress data for efficient disk storage
B. To maintain metadata about database objects
C. To store actual user data records in tables
D. To provide backup copies of all user data

Answer: B

Q690. Which characteristic differentiates a DBMS from a traditional file system?

A. DBMS cannot handle concurrent access but file systems can do it
B. DBMS supports data abstraction while file systems do not offer it
C. DBMS uses flat files while file systems use structured tables
D. DBMS requires less memory than file systems for similar data sets

Answer: B

Q691. In a three-tier architecture, what role does the application server play?

A. It serves as intermediary between client and database
B. It only handles user interface rendering on screen
C. It manages the operating system kernel processes
D. It directly stores data in physical database files

Answer: A

Q692. What advantage does a parallel database architecture provide?

A. Simplified system administration with fewer components
B. Improved query performance through concurrent processing
C. Reduced hardware costs compared to single server systems
D. Elimination of the need for any backup or recovery plan

Answer: B

Q693. Which component of DBMS architecture handles authorization checks?

A. File manager component
B. Buffer manager component
C. Disk space manager component
D. Authorization manager component

Answer: D

Q694. What is the shared-nothing architecture in parallel databases?

A. All processors share a common memory and disk pool
B. Each processor has its own private memory and disk
C. All processors share memory but have private disk units
D. Each processor has private memory but shares all disks

Answer: B

Q695. How does connection pooling improve database performance?

A. By encrypting all data connections between nodes
B. By compressing data packets sent over the network link
C. By duplicating queries across multiple server instances
D. By reusing existing connections instead of creating new

Answer: D

Q696. What distinguishes embedded database systems from server-based ones?

A. Embedded systems only work with object-oriented languages
B. Embedded systems require dedicated server hardware always
C. Embedded systems support more concurrent users overall
D. Embedded systems run within the application process itself

Answer: D

Q697. In DBMS architecture, what does the disk manager handle?

A. Allocating and managing space on storage devices
B. Rendering graphical interfaces for database users
C. Parsing and optimizing SQL queries for execution
D. Managing user sessions and authentication tokens

Answer: A

Q698. What is the benefit of a shared-disk architecture in clusters?

A. Data is automatically replicated to every single node
B. Any node can access any data on the shared storage
C. Each node operates with completely independent data
D. No network communication is needed between the nodes

Answer: B

Q699. Which architectural approach best supports database scalability?

A. Single monolithic server with maximum hardware resources
B. Standalone desktop database with local file storage only
C. Distributed architecture with horizontal partitioning used
D. Centralized file-based system with sequential data access

Answer: C

Q700. What is the purpose of the catalog manager in DBMS?

A. To maintain and provide access to system metadata
B. To manage physical disk read and write operations
C. To encrypt data before transmission over networks
D. To handle user interface rendering for the client

Answer: A

Q701. What is a weak entity in the ER model?

A. An entity that cannot participate in any relationships
B. An entity that depends on another entity for identity
C. An entity that contains only a single attribute value
D. An entity that has its own unique primary key defined

Answer: B

Q702. How does the network data model differ from the hierarchical model?

A. Network model has no relationships while hierarchical has them
B. Network model only supports one-to-one cardinality mappings
C. Network model allows many-to-many while hierarchical does not
D. Network model uses tables while hierarchical uses trees only

Answer: C

Q703. What is generalization in the Enhanced ER model?

A. Removing attributes from entities to simplify the schema
B. Combining entities into a higher-level general entity type
C. Breaking a general entity into specific subtype entities
D. Creating relationships between two unrelated entity types

Answer: B

Q704. What is the difference between total and partial participation?

A. Total is for strong entities while partial is for weak ones
B. Total applies to attributes while partial applies to entities
C. Total means some entities participate while partial means all
D. Total means all entities participate while partial means some

Answer: D

Q705. In the ER model, what is a multivalued attribute?

A. An attribute that is shared between multiple entities
B. An attribute that cannot store any null values at all
C. An attribute with exactly one value for each entity
D. An attribute with multiple values for a single entity

Answer: D

Q706. What is specialization in the Enhanced ER model?

A. Removing relationships between entities to reduce complexity
B. Merging multiple entities into one general supertype entity
C. Converting attributes into separate independent entity types
D. Defining subtypes of an entity based on distinguishing traits

Answer: D

Q707. What does a ternary relationship involve in the ER model?

A. A relationship connecting exactly two entity types together
B. A relationship with three attributes and no entity types
C. A relationship connecting an entity type to itself only
D. A relationship connecting exactly three entity types at once

Answer: D

Q708. What is the object-relational data model?

A. A model designed exclusively for unstructured text data
B. A model using only hierarchical tree data structures
C. A model combining relational and object-oriented features
D. A model that eliminates the concept of tables entirely

Answer: C

Q709. What is a derived attribute in the ER model?

A. An attribute that uniquely identifies each entity instance
B. An attribute whose value is entered manually by users
C. An attribute that connects two entities in a relationship
D. An attribute whose value is computed from other attributes

Answer: D

Q710. What does the concept of aggregation address in ER modeling?

A. Combining multiple attributes into a single composite one
B. Treating a relationship set as a higher-level entity set
C. Splitting an entity into multiple independent sub-entities
D. Removing redundant relationships from the ER diagram view

Answer: B

Q711. What is the difference between a superkey and a candidate key?

A. A candidate key is a minimal superkey without extra attributes
B. A superkey cannot contain null values but candidate keys can
C. A candidate key allows duplicate values but superkeys do not
D. A superkey is always a single attribute unlike candidate key

Answer: A

Q712. What does entity integrity constraint specifically require?

A. Primary key attributes cannot contain any null values
B. Every table must have at least one foreign key defined
C. All attributes in a table must have default values set
D. Foreign keys must reference existing primary key values

Answer: A

Q713. What is referential integrity in relational databases?

A. Each table must have at least two candidate key options
B. Foreign key values must match existing primary key values
C. All attributes must have the same domain type definition
D. Each row must have a unique value in every single column

Answer: B

Q714. What is the cardinality of a relation?

A. The number of candidate keys defined in the table
B. The number of tuples or rows currently in the table
C. The number of attributes or columns in the table
D. The number of foreign key constraints on the table

Answer: B

Q715. What is a composite key in relational databases?

A. A key consisting of two or more attributes
B. A key consisting of a single attribute only
C. A key that allows null values in all parts
D. A key that is automatically generated by DBMS

Answer: A

Q716. What is the difference between natural and surrogate keys?

A. Natural keys are system-generated while surrogate keys are not
B. Natural keys come from data while surrogate keys are generated
C. Natural keys are always composite while surrogate ones are not
D. Natural keys must be numeric while surrogate keys are textual

Answer: B

Q717. What does the closure of a set of functional dependencies represent?

A. The complete set of all functional dependencies that hold
B. The total count of candidate keys in a given relation set
C. The minimal number of attributes needed for a primary key
D. The maximum number of tuples a relation can store safely

Answer: A

Q718. What is an alternate key in relational databases?

A. Any attribute that is part of the primary key itself
B. A foreign key that references a composite primary key
C. A candidate key that is not chosen as primary key
D. A key that only exists in temporary database tables

Answer: C

Q719. What does the relational model information principle state?

A. All information is represented as values in relations
B. Some data must be stored outside of relation tables
C. Data can be stored in hierarchical tree structures
D. Relations can contain duplicate tuples when necessary

Answer: A

Q720. What is a relation schema compared to a relation instance?

A. Schema defines queries while instance defines storage format
B. Schema is the data while instance is the structure definition
C. Schema and instance are always identical in all situations
D. Schema is the structure while instance is the actual data set

Answer: D

Q721. How does a left outer join differ from a natural join?

A. Left outer join requires explicit join conditions unlike natural
B. Left outer join only works on tables with identical schemas
C. Left outer join excludes all unmatched tuples entirely
D. Left outer join preserves all tuples from the left relation

Answer: D

Q722. What is the division operation used for in relational algebra?

A. Splitting a relation into two halves by row count
B. Calculating arithmetic division on numeric column values
C. Removing duplicate tuples from a single relation quickly
D. Finding tuples associated with all values in another set

Answer: D

Q723. What is the equivalence rule for pushing selection through a join?

A. Selections on attributes of one relation can precede the join
B. Selections can never be applied before join operations
C. Selections must always follow projection operations in order
D. Selections and projections are interchangeable in all cases

Answer: A

Q724. How does semijoin differ from natural join?

A. Semijoin returns only attributes from one input relation
B. Semijoin returns attributes from both input relations used
C. Semijoin does not require any common attributes to operate
D. Semijoin always produces more tuples than natural join does

Answer: A

Q725. What does the extended projection operation allow?

A. Arithmetic expressions and renaming in the output columns
B. Filtering rows based on aggregate function result values
C. Joining multiple relations without specifying join conditions
D. Only selecting existing columns from the base relation

Answer: A

Q726. In relational algebra, which operation is NOT a fundamental operation?

A. Union operation
B. Selection operation
C. Intersection operation
D. Cartesian product

Answer: C

Q727. What is the antijoin operation in relational algebra?

A. Tuples from left relation that have no match in right
B. Tuples from right relation that have no match in left
C. Tuples from both relations that match on join condition
D. All possible combinations of tuples from both relations

Answer: A

Q728. How is the generalized projection different from standard projection?

A. It can only select a single column from the relation
B. It removes the requirement for union compatibility checks
C. It allows arithmetic operations on attributes in output
D. It automatically eliminates null values from the results

Answer: C

Q729. What does the grouping operation (gamma) do in extended relational algebra?

A. It renames the relation and all of its attribute names used
B. It sorts tuples by a specified attribute in ascending order
C. It groups tuples and applies aggregate functions to groups
D. It removes all duplicate tuples from the relation entirely

Answer: C

Q730. What is a full outer join in relational algebra?

A. A join that only returns tuples from the right side table
B. A join preserving unmatched tuples from both left and right
C. A join that produces the Cartesian product without any filter
D. A join preserving only matching tuples from both relations

Answer: B

Q731. What is the difference between WHERE and HAVING clauses?

A. WHERE works only with joins while HAVING works with all queries
B. WHERE is used in DDL while HAVING is used in DML statements
C. WHERE filters individual rows while HAVING filters groups
D. WHERE filters after grouping while HAVING filters before it

Answer: C

Q732. Which type of join returns rows only when there is a match in both tables?

A. Inner join
B. Right outer join
C. Left outer join
D. Full outer join

Answer: A

Q733. What does the GROUP BY clause do in SQL?

A. Sorts results in descending order by default
B. Removes duplicate values from a single column only
C. Limits the number of rows in the result output
D. Groups rows sharing common values for aggregation

Answer: D

Q734. What is a subquery in SQL?

A. A query nested inside another SQL query statement
B. A query that creates a temporary table permanently
C. A query that runs automatically on a fixed schedule
D. A query that only works with a single table at once

Answer: A

Q735. What is the purpose of the LIKE operator in SQL?

A. To perform exact matching on numeric columns only
B. To perform pattern matching on string column values
C. To compare two tables for identical structure setup
D. To join tables based on a common column attribute

Answer: B

Q736. What does the BETWEEN operator do in SQL?

A. Joins two tables on a specified column attribute
B. Checks if a value exists in a list of values
C. Filters values within a specified inclusive range
D. Sorts results between ascending and descending order

Answer: C

Q737. What is the difference between DELETE and TRUNCATE?

A. DELETE can use WHERE clause while TRUNCATE removes all rows
B. DELETE is DDL while TRUNCATE is DML in SQL classification
C. DELETE is faster than TRUNCATE for large table operations
D. DELETE removes the table while TRUNCATE only removes rows

Answer: A

Q738. Which aggregate function counts non-null values in a column?

A. MAX function
B. AVG function
C. COUNT function
D. SUM function

Answer: C

Q739. What does the IN operator do in SQL?

A. Checks if a value falls within a numeric range
B. Verifies if a column has null values present
C. Performs pattern matching on string data values
D. Checks if a value matches any value in a list

Answer: D

Q740. What is a self-join in SQL?

A. A join between two completely different database tables
B. A join that requires no ON clause or condition specified
C. A join where a table is joined with itself using aliases
D. A join that automatically matches all columns in tables

Answer: C

Q741. What is the difference between a stored procedure and a function in SQL?

A. Functions can modify data while procedures cannot modify it
B. Procedures cannot take parameters while functions always can
C. Procedures can return values while functions cannot do so
D. Functions must return a value while procedures need not do

Answer: D

Q742. What is a materialized view and how does it differ from a regular view?

A. Materialized views store data physically while regular do not
B. Materialized views cannot be refreshed while regular ones can
C. Materialized views use less storage than regular views always
D. Materialized views are always up to date while regular are not

Answer: A

Q743. What are the different types of triggers based on timing?

A. FAST, MEDIUM, and SLOW trigger performance speed types
B. BEFORE, AFTER, and INSTEAD OF trigger execution types
C. START, MIDDLE, and END trigger position location types
D. FIRST, SECOND, and THIRD trigger priority level types

Answer: B

Q744. What is dynamic SQL and when is it used?

A. SQL that only works with temporary tables created at runtime
B. SQL statements constructed and executed at runtime dynamically
C. SQL that automatically optimizes queries without any user input
D. SQL embedded directly in application source code at compile

Answer: B

Q745. What is the purpose of a savepoint in SQL transactions?

A. To lock all tables preventing concurrent access by others
B. To create a backup of the entire database at that moment
C. To permanently commit all changes up to that specific point
D. To mark a point within a transaction for partial rollback

Answer: D

Q746. What is the difference between clustered and non-clustered indexes?

A. Clustered indexes use more disk space than non-clustered type
B. Clustered indexes sort table data while non-clustered do not
C. Clustered indexes are slower than non-clustered for all uses
D. Clustered indexes cannot be created on primary key columns

Answer: B

Q747. What is a Common Table Expression (CTE) used for?

A. Creating database users with specific permission levels set
B. Encrypting sensitive data before storing it in the database
C. Permanently storing query results in a physical table file
D. Defining temporary named result sets within a single query

Answer: D

Q748. What does the EXPLAIN command do in SQL?

A. It describes the structure of a database table in detail
B. It shows the execution plan for a given SQL query statement
C. It documents the purpose of stored procedures and functions
D. It provides help text for all available SQL commands listed

Answer: B

Q749. What is a window function in SQL?

A. A function that manages database connection windows open
B. A function that creates graphical windows in applications
C. A function that performs calculations across row sets
D. A function that displays query results in separate panels

Answer: C

Q750. What is the purpose of the CASCADE option in DROP TABLE?

A. It drops the table and all dependent objects together
B. It only drops the table if it contains no data rows
C. It converts the table to a view before deletion occurs
D. It creates a backup before dropping the specified table

Answer: A

Q751. What is a partial dependency in database normalization?

A. A key attribute depends on a non-key attribute in table
B. A non-key attribute depends on another non-key attribute
C. A non-key attribute depends on the entire composite key
D. A non-key attribute depends on part of a composite key

Answer: D

Q752. What additional requirement does 3NF impose over 2NF?

A. No transitive dependencies on the primary key column
B. All tables must have at least two candidate key options
C. Every attribute must be part of some candidate key set
D. All attributes must be of the same domain type value

Answer: A

Q753. What is a transitive dependency?

A. Attribute A depends on a foreign key in another table
B. Attribute A depends on B which depends on primary key
C. Attribute A and B both depend on the same primary key
D. Attribute A depends directly on the primary key only

Answer: B

Q754. How does BCNF differ from 3NF?

A. BCNF allows transitive dependencies that 3NF does not
B. BCNF is less restrictive than 3NF for all relations
C. BCNF requires every determinant to be a candidate key
D. BCNF only applies to tables with a single candidate key

Answer: C

Q755. What is a multivalued dependency?

A. An attribute depends on the concatenation of all other ones
B. One attribute value determines a set of independent values
C. Two attributes are always equal in every tuple of table
D. One attribute value determines exactly one other value

Answer: B

Q756. What does the fourth normal form (4NF) address?

A. Transitive dependencies among non-key attributes used
B. Join dependencies that cannot be reduced any further
C. Partial dependencies on the composite primary key
D. Multivalued dependencies that cause data redundancy

Answer: D

Q757. What is lossless join decomposition?

A. A decomposition that always results in fewer tables total
B. A decomposition where some data may be lost permanently
C. A decomposition where original data can be fully recovered
D. A decomposition that eliminates all foreign keys from tables

Answer: C

Q758. What is dependency preservation in normalization?

A. All functional dependencies can be checked within tables
B. Dependencies are only preserved in the original table form
C. Dependencies are stored in a separate metadata table file
D. Some functional dependencies may be lost during the split

Answer: A

Q759. When is denormalization justified in database design?

A. When read performance is prioritized over write efficiency
B. When all tables are already in first normal form status
C. When the database has too few tables for organization
D. When the database has no functional dependencies at all

Answer: A

Q760. What is the difference between lossless and lossy decomposition?

A. Lossless recovers original data while lossy creates spurious tuples
B. Lossless is always faster while lossy requires more processing
C. Lossless uses more storage while lossy uses compression methods
D. Lossless adds extra data while lossy preserves original data

Answer: A

Q761. What is a serializable schedule?

A. A schedule that does not require any concurrency control
B. A schedule equivalent in result to some serial schedule
C. A schedule where transactions run in strict serial order
D. A schedule that allows all types of data conflicts always

Answer: B

Q762. What is the difference between conflict serializability and view serializability?

A. View serializability allows more schedules than conflict type
B. Conflict serializability is less restrictive than view type
C. View serializability requires all transactions to be serial
D. Conflict serializability only applies to read operations here

Answer: A

Q763. What is a dirty read in transaction processing?

A. Reading data from a backup file instead of the live database
B. Reading data written by an uncommitted transaction directly
C. Reading data from a table that has no index defined on it
D. Reading data that has been committed by another transaction

Answer: B

Q764. What is a non-repeatable read anomaly?

A. A transaction reads the same row twice and gets same results
B. A transaction cannot read any rows because they are all locked
C. A transaction reads the same row twice and gets different results
D. A transaction reads rows from a table that does not exist yet

Answer: C

Q765. What is a cascading rollback in transaction management?

A. Automatically committing all pending transactions at the same time
B. Rolling back a single transaction without affecting any others
C. Restarting the database server to clear all active transactions
D. One transaction rollback causing other transactions to rollback

Answer: D

Q766. What is the write-ahead logging (WAL) protocol?

A. Neither log records nor data pages need to be written to disk
B. Log records are written to disk before data pages are flushed
C. Both log records and data pages are written simultaneously
D. Data pages are written to disk before log records are made

Answer: B

Q767. What is a recoverable schedule?

A. A schedule that automatically recovers from all hardware failures
B. A schedule where no transaction ever needs to be rolled back
C. A schedule where transactions commit only after dependencies commit
D. A schedule where all transactions execute in strict serial order

Answer: C

Q768. What are the different transaction isolation levels in SQL?

A. Read Uncommitted, Read Committed, Repeatable Read, Serializable
B. Private, Shared, Protected, and Public access levels for data
C. Fast, Normal, Slow, and Batch processing modes for transactions
D. Low, Medium, High, and Maximum security levels for transactions

Answer: A

Q769. What is a checkpoint in database recovery?

A. A point where all current data is deleted from the database
B. A point where the database schema is backed up to a file
C. A point where database state is synchronized with the log
D. A point where new transactions are prevented from starting

Answer: C

Q770. What is a cascadeless schedule?

A. A schedule allowing transactions to read uncommitted data freely
B. A schedule where transactions only read data from committed ones
C. A schedule where all transactions execute without any commits
D. A schedule that permits cascading rollbacks to occur frequently

Answer: B

Q771. What is the two-phase locking protocol?

A. Transactions use only shared locks in the first phase entirely
B. Transactions acquire all locks first then release all at end
C. Transactions can acquire and release locks in any order freely
D. Transactions have a growing phase and a shrinking phase

Answer: D

Q772. How does the wait-die scheme handle deadlock prevention?

A. Younger transactions wait while older ones are rolled back
B. Older transactions wait while younger ones are rolled back
C. Transactions are randomly selected for rollback on each conflict
D. All transactions are rolled back when a conflict is detected

Answer: B

Q773. How does the wound-wait scheme handle deadlock prevention?

A. Older transactions are rolled back while younger ones wait
B. Older transactions preempt younger ones while younger wait
C. Both transactions are always rolled back on any conflict
D. Transactions are processed strictly in first-come first-served

Answer: B

Q774. What is timestamp-based concurrency control?

A. Scheduling transactions alphabetically by their name values
B. Using random numbers to determine transaction execution order
C. Assigning locks based on the size of each transaction query
D. Ordering transactions by timestamps to ensure serializability

Answer: D

Q775. What is optimistic concurrency control?

A. Transactions acquire locks before accessing any data items
B. Transactions are executed one at a time in serial order only
C. Transactions always use exclusive locks on all data items used
D. Transactions proceed without locks and validate before commit

Answer: D

Q776. What is a lock escalation?

A. Converting an exclusive lock to a shared lock type
B. Releasing all locks held by a transaction at once
C. Converting fine-grained locks to coarser-grained locks
D. Transferring locks from one transaction to another one

Answer: C

Q777. What is the difference between pessimistic and optimistic concurrency control?

A. Pessimistic uses timestamps while optimistic uses lock-based
B. Pessimistic locks data preemptively while optimistic validates at commit
C. Pessimistic allows dirty reads while optimistic prevents all of them
D. Pessimistic is always faster than optimistic concurrency approaches

Answer: B

Q778. What is intention locking in a multigranularity system?

A. Locking individual rows before accessing a full table record
B. Preventing any transaction from acquiring exclusive lock types
C. Locking the entire database before any transaction can start
D. Signaling intent to lock at a finer granularity level below

Answer: D

Q779. What is the phantom problem in concurrency control?

A. A transaction reads data that was never written to database
B. New rows appear in repeated queries due to concurrent inserts
C. Data becomes corrupted due to simultaneous write operations
D. A transaction loses its locks after a network disconnection

Answer: B

Q780. What is multiversion concurrency control (MVCC)?

A. Maintaining a single version of data with exclusive locking
B. Maintaining multiple versions of data for concurrent access
C. Deleting old versions of data immediately after each update
D. Compressing multiple data versions into a single stored copy

Answer: B

Q781. What is the difference between immediate and deferred database modification?

A. Immediate writes to log first while deferred writes to disk first
B. Immediate uses checkpoints while deferred uses log records for data
C. Immediate only works for read operations while deferred handles writes
D. Immediate writes changes to disk before commit while deferred waits

Answer: D

Q782. What is shadow paging in database recovery?

A. Creating shadows of all user queries for later replay at anytime
B. Maintaining two page tables: current and shadow for safe recovery
C. Encrypting page data to prevent unauthorized access to records
D. Duplicating the entire database to a secondary server for backup

Answer: B

Q783. What is the steal/no-steal policy in buffer management?

A. Steal and no-steal both prevent uncommitted pages from being flushed
B. No-steal allows uncommitted pages while steal prevents their flushing
C. Steal allows uncommitted pages to be flushed while no-steal prevents
D. Steal refers to memory allocation while no-steal refers to disk usage

Answer: C

Q784. What is the force/no-force policy in buffer management?

A. Force uses redo logging while no-force uses undo logging exclusively
B. Force writes all pages at checkpoint while no-force writes at commit
C. Force prevents dirty pages from being flushed while no-force allows
D. Force requires all pages written at commit while no-force does not

Answer: D

Q785. What is media recovery versus crash recovery?

A. Media recovery restores from disk failure while crash handles system crash
B. Media recovery handles software bugs while crash recovery handles disk
C. Media recovery is faster than crash recovery in all failure situations
D. Media recovery uses logs while crash recovery uses backups exclusively

Answer: A

Q786. What information does a log record typically contain?

A. Only the transaction ID and the timestamp of each operation
B. Only the names of tables that were accessed during the transaction
C. Only the SQL query text that was executed by the user application
D. Transaction ID, data item, old value, new value, and operation type

Answer: D

Q787. What is fuzzy checkpointing?

A. A checkpoint that runs automatically without any administrator control
B. A checkpoint that allows transactions to continue during the process
C. A checkpoint that only saves metadata without any actual data pages
D. A checkpoint that stops all transactions until it completes fully

Answer: B

Q788. What is the purpose of the log sequence number (LSN)?

A. To assign priority levels to different transactions for scheduling
B. To uniquely identify and order log records in the transaction log
C. To measure the physical size of each log record stored on disk
D. To count the total number of transactions in the database system

Answer: B

Q789. What is a warm backup versus a cold backup?

A. Cold backup is taken while database is online while warm is offline
B. Warm backup is taken while database is online while cold is offline
C. Warm backup only copies logs while cold backup copies entire database
D. Cold backup is faster than warm backup in all operational scenarios

Answer: B

Q790. What is point-in-time recovery?

A. Recovering only the tables that were modified since last backup
B. Recovering the database schema without any of the actual data
C. Recovering the database to a specific moment before the failure
D. Recovering the database to its most recent committed state only

Answer: C

Q791. What is the difference between a B-tree and a B+ tree?

A. B+ tree requires more levels than B-tree for the same data volume
B. B-tree supports range queries while B+ tree only supports exact match
C. B+ tree stores data only in leaves while B-tree stores in all nodes
D. B-tree stores data in leaves only while B+ tree stores in all nodes

Answer: C

Q792. What is a composite index?

A. An index on a single column of a database table
B. An index built on two or more columns together
C. An index that is shared between multiple tables
D. An index that uses both hash and tree structures

Answer: B

Q793. What is index selectivity and why does it matter?

A. The time required to create the index on a large data table
B. The size of the index file compared to the original data file
C. The ratio of distinct values to total records affecting efficiency
D. The number of tables that reference a particular index entry

Answer: C

Q794. What is extendible hashing?

A. A static hash function that never changes bucket count
B. A hashing technique that eliminates all collision handling
C. A hash function that only works with string data values
D. A dynamic hashing scheme that grows and shrinks as needed

Answer: D

Q795. What is a bitmap index best suited for?

A. Columns with very high cardinality like unique identifiers
B. Tables that are frequently updated with new data records
C. Columns with low cardinality like gender or status fields
D. Queries that require exact match lookups on primary keys

Answer: C

Q796. What is the advantage of a covering index?

A. It works only with aggregate queries on numeric data types
B. It automatically updates when the table schema is changed
C. It satisfies a query entirely without accessing table data
D. It requires more storage space than a regular index type

Answer: C

Q797. What is linear hashing?

A. A dynamic hashing scheme that splits buckets incrementally
B. A hashing scheme that uses a fixed number of buckets always
C. A hash function that only produces sequential hash values
D. A hashing technique that stores data in a linear linked list

Answer: A

Q798. What is the difference between static and dynamic hashing?

A. Static is always faster than dynamic for all query types used
B. Static uses trees while dynamic uses linear search methods
C. Static supports range queries while dynamic only supports exact
D. Static has fixed buckets while dynamic adjusts bucket count

Answer: D

Q799. What is a multi-level index?

A. An index structure with multiple levels like a tree hierarchy
B. An index on multiple columns of the same table structure
C. An index shared across multiple databases on the network
D. An index that indexes other indexes for circular referencing

Answer: A

Q800. What is the difference between ordered and unordered file organization?

A. Ordered files use hashing while unordered files use tree indexes
B. Ordered files cannot be indexed while unordered files always can
C. Ordered files are stored in RAM while unordered are stored on disk
D. Ordered files maintain sort order while unordered files do not

Answer: D

Q801. What is the difference between a nested loop join and a hash join?

A. Hash join only works with equality conditions while nested loop not
B. Hash join builds a hash table while nested loop iterates through rows
C. Nested loop is always faster than hash join for all data sizes
D. Nested loop requires sorted input while hash join does not need it

Answer: B

Q802. What is a merge join and when is it most efficient?

A. Merge join sorts and merges inputs and is efficient when pre-sorted
B. Merge join uses nested loops and is efficient for indexed columns
C. Merge join hashes both inputs and is efficient for small tables
D. Merge join uses random access and is efficient for all query types

Answer: A

Q803. What is query rewriting in optimization?

A. Changing the query results to match expected output format
B. Adding additional columns to the query output automatically
C. Transforming a query into an equivalent but more efficient form
D. Replacing all queries with stored procedure calls instead

Answer: C

Q804. What is selectivity estimation in query optimization?

A. The process of selecting the best join algorithm for queries
B. The process of selecting which indexes to create on tables
C. Estimating the total storage required for query result output
D. Estimating the fraction of rows that satisfy a given condition

Answer: D

Q805. What is the purpose of a histogram in query optimization?

A. Tracking the number of queries executed over time periods
B. Displaying query performance metrics to database operators
C. Providing data distribution information for cost estimation
D. Graphically showing the structure of the execution plan tree

Answer: C

Q806. What is pipelining in query execution?

A. Passing results from one operation to the next immediately
B. Executing all operations in parallel on separate processors
C. Storing intermediate results to disk before continuing
D. Repeating operations until the correct result is obtained

Answer: A

Q807. What is materialization in query execution?

A. Executing all operations in a single pass without stopping
B. Storing intermediate results temporarily before next step
C. Deleting temporary tables after query execution completes
D. Converting query results into a permanent database table

Answer: B

Q808. What is predicate pushdown in query optimization?

A. Removing all filter conditions to speed up processing time
B. Duplicating filter conditions across multiple plan operators
C. Moving filter conditions closer to the data source in plan
D. Moving filter conditions to the end of the execution plan

Answer: C

Q809. What does the catalog store that helps query optimization?

A. The source code of all stored procedures and functions used
B. The network configuration for distributed query processing
C. Table sizes, column statistics, and index information data
D. The actual data rows from all database tables and views

Answer: C

Q810. What is a left-deep join tree in query optimization?

A. A tree where both inputs of each join are intermediate results
B. A tree where the left input of each join is a base table
C. A tree where the right input of each join is a base table
D. A tree where all joins are performed simultaneously in parallel

Answer: C

Q811. What is the difference between discretionary and mandatory access control?

A. Discretionary allows owner control while mandatory uses labels
B. Discretionary is managed by system while mandatory by data owner
C. Discretionary applies to tables while mandatory applies to rows
D. Discretionary uses encryption while mandatory uses authentication

Answer: A

Q812. What is role-based access control (RBAC)?

A. Encrypting data based on the role of the user accessing data
B. Assigning permissions to roles that are then granted to users
C. Auditing user actions based on their assigned role in system
D. Assigning permissions directly to each individual user account

Answer: B

Q813. How does parameterized querying prevent SQL injection attacks?

A. It blocks all user input from being used in SQL queries
B. It validates user input using regular expressions only here
C. It encrypts the entire SQL query before sending to server
D. It separates SQL code from user input preventing injection

Answer: D

Q814. What is transparent data encryption (TDE)?

A. Encrypting data using user-provided keys for each session
B. Encrypting data at rest without requiring application changes
C. Encrypting data only when it is transmitted over a network
D. Encrypting only the database schema and not the actual data

Answer: B

Q815. What is the principle of least privilege in database security?

A. Users should share a single account to simplify access control
B. Users should have only minimum access needed for their tasks
C. Users should have maximum access to complete tasks efficiently
D. Users should have temporary access that expires after each day

Answer: B

Q816. What is database masking?

A. Replacing sensitive data with realistic but fake values shown
B. Removing all data from a database for security purposes only
C. Encrypting all database files using symmetric key algorithms
D. Compressing database tables to obscure the stored data values

Answer: A

Q817. What is a security label in mandatory access control?

A. A classification assigned to data and users for access rules
B. A type of encryption key used to protect data at rest only
C. A log entry recording when data was last accessed by users
D. A physical tag attached to the database server hardware

Answer: A

Q818. What is the purpose of database activity monitoring?

A. Managing storage allocation across multiple database disk drives
B. Optimizing query performance by tracking slow-running queries
C. Tracking and analyzing database activities for security threats
D. Scheduling automated backups during off-peak operational hours

Answer: C

Q819. What is column-level security in databases?

A. Restricting access to specific columns within a table view
B. Encrypting individual columns using different key algorithms
C. Restricting access to entire database tables for all users
D. Compressing specific columns to reduce storage size needed

Answer: A

Q820. What is a virtual private database (VPD)?

A. A database that is only accessible through a VPN connection
B. A feature that adds predicates to queries based on user policy
C. A temporary database created for each user session separately
D. A database running on a private physical server exclusively

Answer: B

Q821. What is the CAP theorem in distributed databases?

A. Compression, Authentication, Privacy are the three key goals
B. Concurrency, Atomicity, Persistence must all be maintained
C. Consistency, Availability, Performance can all be guaranteed
D. Consistency, Availability, Partition tolerance: pick two of three

Answer: D

Q822. What is the two-phase commit protocol used for?

A. Replicating data between primary and secondary server sites
B. Fragmenting tables into horizontal and vertical data parts
C. Ensuring atomicity of transactions across distributed sites
D. Optimizing queries across multiple distributed database sites

Answer: C

Q823. What is the difference between homogeneous and heterogeneous distributed databases?

A. Homogeneous uses same DBMS at all sites while heterogeneous not
B. Heterogeneous uses same DBMS at all sites while homogeneous not
C. Heterogeneous has more data than homogeneous distributed database
D. Homogeneous stores data at one site while heterogeneous multiple

Answer: A

Q824. What is eventual consistency in distributed systems?

A. All nodes always have identical data at every single moment
B. All nodes will converge to the same data state over time
C. Only the primary node maintains the consistent data state
D. Data is never consistent across any nodes in the system

Answer: B

Q825. What is a federated database system?

A. A database system that only supports read-only query operations
B. An integration of multiple autonomous database systems together
C. A single database distributed across multiple data centers
D. A backup system that replicates data to geographically distant

Answer: B

Q826. What is distributed deadlock and how is it detected?

A. Deadlock involving transactions across multiple sites via graph
B. Deadlock that only affects read-only transactions across sites
C. Deadlock that automatically resolves without any intervention
D. Deadlock that occurs only on a single site in the system

Answer: A

Q827. What is the difference between synchronous and asynchronous replication?

A. Synchronous only works with two nodes while asynchronous many
B. Synchronous replicates to all nodes before commit while async not
C. Asynchronous replicates to all nodes before commit while sync not
D. Both methods replicate at exactly the same speed and timing

Answer: B

Q828. What is a global schema in a distributed database?

A. A schema that only exists at the primary coordinator node
B. A temporary schema created for each distributed query session
C. A physical schema that defines storage at each individual site
D. A unified logical view of the entire distributed database

Answer: D

Q829. What is the semi-join operation used for in distributed databases?

A. Joining all columns from two tables at different remote sites
B. Splitting a table into horizontal fragments for distribution
C. Replicating the result of a join across all available sites
D. Reducing data transfer by sending only relevant join columns

Answer: D

Q830. What is sharding in distributed databases?

A. Replicating the entire database to every node in the cluster
B. Encrypting data during transmission between distributed nodes
C. Compressing data before storing it on distributed server nodes
D. Partitioning data across nodes where each holds a data subset

Answer: D

Q831. What is the BASE model in NoSQL databases?

A. Batch Access, Sequential Execution, Cached data approach
B. Binary Allocation, Sorted Entries, Encrypted data approach
C. Basically Available, Soft state, Eventually consistent approach
D. Buffered Access, Static Entries, Compressed data approach

Answer: C

Q832. How does a column-family store differ from a relational table?

A. Column-family requires a fixed schema like relational tables do
B. Column-family does not support any form of data querying at all
C. Column-family groups columns into families with dynamic columns
D. Column-family stores data in rows like relational tables exactly

Answer: C

Q833. What is the advantage of document databases for hierarchical data?

A. They require multiple collections for each hierarchy level
B. They require joins to represent nested data structures
C. They store nested data naturally within a single document
D. They flatten all hierarchical data into a single flat table

Answer: C

Q834. What is a graph database best suited for?

A. Storing large volumes of unstructured text documents only
B. Storing time-series data from Internet of Things sensors
C. Processing simple key-value lookups at maximum speed
D. Modeling and querying highly connected relationship data

Answer: D

Q835. What is eventual consistency in NoSQL systems?

A. All replicas are always perfectly synchronized at every moment
B. Consistency is only guaranteed during maintenance windows only
C. Replicas may temporarily diverge but will converge over time
D. Data is never consistent across any of the database replicas

Answer: C

Q836. What is data denormalization in NoSQL and why is it used?

A. Compressing data to reduce storage requirements on disk drives
B. Removing all redundancy from the data model for efficiency
C. Encrypting data for security before storing in the database
D. Duplicating data to avoid expensive join operations at query

Answer: D

Q837. What is the difference between embedded and referenced documents in MongoDB?

A. Referenced stores related data together while embedded links by ID
B. Embedded is for read operations while referenced is for write only
C. Both approaches store data identically with no practical difference
D. Embedded stores related data together while referenced links by ID

Answer: D

Q838. What is MapReduce in the context of NoSQL databases?

A. A strategy for backing up NoSQL databases to external storage
B. A programming model for processing large data sets in parallel
C. A method for encrypting data stored in NoSQL database systems
D. A technique for replicating data across multiple server nodes

Answer: B

Q839. What is a wide-column store?

A. A database that only supports columns with wide text data types
B. A database that stores each row with exactly the same columns
C. A database where each row can have different sets of columns
D. A database where column width is fixed at creation time always

Answer: C

Q840. What is the role of a partition key in NoSQL databases?

A. It manages user authentication and access control rights
B. It defines the schema structure for each document record
C. It determines how data is distributed across cluster nodes
D. It encrypts data before storing it in the database system

Answer: C

Q841. What is a snowflake schema and how does it differ from a star schema?

A. Snowflake has no dimension tables while star schema has many
B. Snowflake normalizes dimension tables while star keeps denormalized
C. Snowflake has multiple fact tables while star schema has one only
D. Snowflake is faster for queries while star schema is slower ones

Answer: B

Q842. What is the difference between Extract and Transform phases in ETL?

A. Extract cleans data while Transform retrieves data from sources
B. Extract and Transform perform the exact same operations on data
C. Extract loads data into warehouse while Transform retrieves it
D. Extract retrieves source data while Transform cleans and converts

Answer: D

Q843. What is a slowly changing dimension (SCD)?

A. A dimension whose attributes change infrequently over time
B. A dimension that is updated in real-time with every change
C. A dimension that never changes after initial data loading
D. A dimension that is deleted and recreated each data load

Answer: A

Q844. What is the drill-down operation in OLAP?

A. Moving from detailed data to summarized aggregate level
B. Sorting data by a specified column in descending order only
C. Moving from summary data to more detailed granular level
D. Filtering data to show only records from a specific time

Answer: C

Q845. What is the roll-up operation in OLAP?

A. Moving from aggregated summary to detailed granular data
B. Adding new dimension tables to the existing schema design
C. Deleting old data from the data warehouse storage tables
D. Moving from detailed data to higher-level aggregated summary

Answer: D

Q846. What is a surrogate key in a data warehouse dimension table?

A. A natural business key from the source operational system
B. A composite key made from multiple business attribute columns
C. A foreign key referencing the primary fact table records
D. An artificial system-generated key independent of source data

Answer: D

Q847. What is data granularity in a data warehouse?

A. The frequency at which the data warehouse backup is performed
B. The level of detail at which data is stored in the warehouse
C. The total number of dimension tables in the warehouse schema
D. The level of encryption applied to stored warehouse data

Answer: B

Q848. What is the slice operation in OLAP?

A. Selecting a single value on one dimension to view data
B. Aggregating data across all available dimensions at once
C. Filtering data based on a range of values in a dimension
D. Rotating the data cube axes to change the perspective view

Answer: A

Q849. What is the dice operation in OLAP?

A. Rotating the axes of the data cube for different views
B. Selecting data by filtering on a single dimension value
C. Aggregating data from the lowest to highest detail level
D. Selecting a sub-cube by filtering on multiple dimensions

Answer: D

Q850. What is the pivot operation in OLAP?

A. Filtering data to show only specific dimension values only
B. Merging two dimension tables into a single combined table
C. Rotating the data cube axes to get a different perspective
D. Adding new measures to the existing fact table in warehouse

Answer: C

Q851. Which type of data independence allows changes to the storage structure without affecting the conceptual schema?

A. Logical data independence
B. Physical data independence
C. Schema independence
D. View independence

Answer: B

Q852. In the three-schema architecture, which schema represents individual user views?

A. Internal schema
B. Conceptual schema
C. External schema
D. Physical schema

Answer: C

Q853. What is the primary difference between a database and a data warehouse?

A. A database is for operational data while a data warehouse is for analytical data
B. A database is larger than a data warehouse
C. A data warehouse only stores text data
D. There is no difference between them

Answer: A

Q854. Which DBMS component is responsible for enforcing integrity constraints?

A. Query optimizer
B. Integrity constraint manager
C. Buffer manager
D. File manager

Answer: B

Q855. What is the purpose of a view in a database system?

A. To store physical data on disk
B. To provide a virtual table derived from base tables
C. To manage network connections
D. To compile SQL statements

Answer: B

Q856. Which of the following is a characteristic of the relational model?

A. Data is stored in tree structures
B. Data is represented as tables with rows and columns
C. Data must be stored sequentially
D. Each record must have a parent record

Answer: B

Q857. What role does the data dictionary play in a DBMS?

A. It stores actual user data
B. It stores metadata about the database structure
C. It performs data encryption
D. It manages user sessions

Answer: B

Q858. Which of the following describes a multi-user DBMS?

A. A DBMS that runs on a single processor only
B. A DBMS that allows multiple users to access the database concurrently
C. A DBMS that stores data in multiple formats
D. A DBMS limited to read-only access

Answer: B

Q859. What is the difference between intension and extension of a database?

A. Intension is the schema while extension is the current data
B. Intension is the data while extension is the schema
C. Both refer to the same concept
D. Intension refers to queries while extension refers to indexes

Answer: A

Q860. Which of the following is an example of a DCL command?

A. CREATE TABLE
B. SELECT
C. GRANT
D. UPDATE

Answer: C

Q861. What is the primary benefit of connection pooling in database architecture?

A. It encrypts all database connections
B. It reuses existing connections to reduce overhead of creating new ones
C. It automatically backs up data
D. It compresses query results

Answer: B

Q862. In a shared-memory parallel architecture, what is the main bottleneck?

A. Network bandwidth between nodes
B. Memory bus contention when many processors access shared memory
C. Disk storage capacity
D. Application code complexity

Answer: B

Q863. What is the function of the log manager in DBMS architecture?

A. It manages user login sessions
B. It records all database modifications for recovery purposes
C. It formats query output
D. It manages network connections

Answer: B

Q864. Which architecture is most suitable for large-scale internet applications?

A. Single-tier architecture
B. Two-tier architecture
C. N-tier or multi-tier architecture
D. Standalone architecture

Answer: C

Q865. What distinguishes a fat client from a thin client in database architecture?

A. Fat clients have larger monitors
B. Fat clients process more application logic locally while thin clients rely on the server
C. Fat clients use more network bandwidth
D. There is no real difference

Answer: B

Q866. In DBMS architecture, what does the term 'middleware' refer to?

A. Hardware between the client and server
B. Software that connects different applications or system components
C. The middle table in a database
D. A type of storage device

Answer: B

Q867. What is the role of the query optimizer in DBMS architecture?

A. To check SQL syntax
B. To determine the most efficient execution plan for a query
C. To store query results
D. To manage user connections

Answer: B

Q868. What is an embedded database system?

A. A database that runs within an application process without a separate server
B. A database stored on embedded hardware only
C. A database that cannot be queried
D. A database used only for IoT devices

Answer: A

Q869. In database architecture, what is a stored procedure?

A. A procedure for physically storing disks
B. A precompiled collection of SQL statements stored in the database
C. A method for backing up data
D. A procedure for installing the DBMS

Answer: B

Q870. What advantage does the three-tier architecture offer over two-tier for security?

A. It uses stronger encryption algorithms
B. It adds an intermediate layer that prevents direct client access to the database
C. It requires fewer passwords
D. It eliminates the need for firewalls

Answer: B

Q871. What is partial participation in an ER diagram?

A. All entity instances must participate in the relationship
B. Some entity instances may not participate in the relationship
C. The relationship only involves one entity
D. The entity has no attributes

Answer: B

Q872. How is a weak entity identified in an ER diagram?

A. By a single rectangle
B. By a double-bordered rectangle
C. By a dashed rectangle
D. By a bold rectangle

Answer: B

Q873. What is a partial key (discriminator) of a weak entity?

A. A key that uniquely identifies the weak entity globally
B. An attribute that partially identifies weak entity instances within their owner
C. A foreign key reference
D. A candidate key of a strong entity

Answer: B

Q874. In the EER model, what does disjoint specialization mean?

A. An entity can belong to multiple subclasses simultaneously
B. An entity can belong to only one subclass at a time
C. All entities must belong to a subclass
D. No entity belongs to any subclass

Answer: B

Q875. What is the difference between specialization and generalization in EER?

A. They are the same process
B. Specialization is top-down while generalization is bottom-up
C. Specialization is bottom-up while generalization is top-down
D. Neither relates to entity hierarchies

Answer: B

Q876. How is a derived attribute represented in an ER diagram?

A. With a solid oval
B. With a dashed oval
C. With a double oval
D. With an underlined oval

Answer: B

Q877. What is an identifying relationship for a weak entity?

A. Any relationship the weak entity participates in
B. The relationship that provides the weak entity with part of its identification
C. A self-referencing relationship
D. A many-to-many relationship

Answer: B

Q878. In the ER model, what is a binary relationship?

A. A relationship involving exactly two entity types
B. A relationship stored in binary format
C. A relationship with two attributes
D. A relationship between an entity and itself

Answer: A

Q879. What does the participation constraint specify in an ER relationship?

A. The data types of related entities
B. Whether existence of an entity depends on being related to another entity
C. The maximum number of entities in a table
D. The storage format of relationships

Answer: B

Q880. What is an associative entity in ER modeling?

A. A strong entity with multiple keys
B. An entity created to resolve a many-to-many relationship
C. An entity with only derived attributes
D. An entity without any relationships

Answer: B

Q881. What is a functional dependency in relational databases?

A. A dependency between two database servers
B. A constraint where one attribute set determines another
C. A physical storage dependency
D. A relationship between database users

Answer: B

Q882. What is the update anomaly in relational databases?

A. An error that occurs during database installation
B. Inconsistency caused when the same data must be updated in multiple rows
C. A problem with database backup procedures
D. An error in the query optimizer

Answer: B

Q883. What is the insertion anomaly in a relational table?

A. Inability to insert a new row because required data for another entity is missing
B. An error in the INSERT SQL statement syntax
C. A performance issue during batch inserts
D. A conflict between two simultaneous insertions

Answer: A

Q884. What is the deletion anomaly in relational databases?

A. A failure to delete data from disk
B. Unintended loss of data when deleting a row that contains other useful information
C. A conflict during concurrent deletions
D. An error in the DELETE SQL syntax

Answer: B

Q885. What is a natural join in the relational model?

A. A join that only uses primary keys
B. A join that combines tuples from two relations based on common attributes with equal values
C. A join that preserves all rows from both tables
D. A join that creates a Cartesian product

Answer: B

Q886. What constraint does the ON DELETE CASCADE option enforce?

A. Prevents deletion of any row with a primary key
B. Automatically deletes referencing rows when the referenced row is deleted
C. Logs all deletion operations
D. Requires manual confirmation before deletion

Answer: B

Q887. What is a self-referencing foreign key?

A. A foreign key that references a column in the same table
B. A foreign key that references itself recursively with no limit
C. A key that auto-increments
D. A key used for database replication

Answer: A

Q888. What is the closed world assumption in the relational model?

A. The database is closed for modifications after deployment
B. Only the data present in the database is considered true; absent data is considered false
C. All tables must be fully populated before queries
D. The database cannot communicate with external systems

Answer: B

Q889. What is the difference between ON DELETE SET NULL and ON DELETE RESTRICT?

A. SET NULL deletes the child row while RESTRICT sets it to null
B. SET NULL sets the foreign key to null while RESTRICT prevents the deletion
C. Both options behave identically
D. SET NULL removes the constraint while RESTRICT adds one

Answer: B

Q890. What is a theta join in relational algebra?

A. A join using only the equality condition
B. A join that combines relations based on any comparison condition
C. A join that produces a Cartesian product
D. A join exclusive to temporal databases

Answer: B

Q891. How does a left outer join differ from an inner join?

A. Left outer join only returns rows from the left table
B. Left outer join includes all rows from the left table, padding with nulls where there is no match
C. Left outer join is faster than inner join
D. There is no difference

Answer: B

Q892. In relational algebra, can the intersection operation be expressed using other fundamental operations?

A. No, intersection is a fundamental operation
B. Yes, R ") S = R - (R - S)
C. Yes, R ") S = R × S
D. Yes, R ") S = R "* S

Answer: B

Q893. What is the result of a full outer join?

A. Only matching rows from both tables
B. All rows from both tables with nulls where no match exists
C. Only the Cartesian product
D. An empty relation

Answer: B

Q894. What does the grouping (gamma) operation do in extended relational algebra?

A. Groups tuples by specified attributes and applies aggregate functions
B. Groups tables into schemas
C. Groups users by access level
D. Groups queries for batch processing

Answer: A

Q895. What is a semijoin in relational algebra?

A. Half of a natural join result
B. A join that returns tuples from the first relation that have matching tuples in the second
C. A join that uses only half the attributes
D. A partial Cartesian product

Answer: B

Q896. Why is relational algebra considered procedural?

A. Because it uses procedures stored in the database
B. Because it specifies the sequence of operations to obtain the result
C. Because it requires procedural programming languages
D. Because it processes data in procedure calls

Answer: B

Q897. What is relational completeness?

A. A database that stores all possible data
B. A query language that can express any query expressible in relational algebra
C. A relation that has no null values
D. A database with all constraints defined

Answer: B

Q898. What is the difference between WHERE and HAVING clauses in SQL?

A. There is no difference
B. WHERE filters rows before grouping while HAVING filters groups after grouping
C. HAVING is used before WHERE in query execution
D. WHERE only works with numeric columns

Answer: B

Q899. What is a correlated subquery in SQL?

A. A subquery that runs independently of the outer query
B. A subquery that references columns from the outer query
C. A query that uses the CORRELATE keyword
D. A query that joins two subqueries

Answer: B

Q900. What does the COALESCE function do in SQL?

A. Combines two tables
B. Returns the first non-null value from a list of expressions
C. Counts all null values
D. Converts null to zero

Answer: B

Q901. What is the purpose of the GROUP BY clause in SQL?

A. To sort the result set
B. To group rows with the same values in specified columns for aggregate functions
C. To filter individual rows
D. To join tables together

Answer: B

Q902. What does the EXISTS keyword do in SQL?

A. Checks if a table exists in the database
B. Returns TRUE if the subquery returns one or more rows
C. Checks if a column exists in a table
D. Verifies database connection

Answer: B

Q903. What is the purpose of the CASE expression in SQL?

A. To define a database case study
B. To provide conditional logic within a SQL query
C. To change the case of text values
D. To handle exception cases

Answer: B

Q904. What is an alias in SQL?

A. A permanent rename of a table
B. A temporary name given to a table or column for the duration of a query
C. A backup name for a database
D. A type of database index

Answer: B

Q905. What is the difference between BEFORE and AFTER triggers?

A. BEFORE triggers execute before the triggering event while AFTER triggers execute after
B. BEFORE triggers are faster than AFTER triggers
C. AFTER triggers cannot modify data
D. There is no practical difference

Answer: A

Q906. What is a Common Table Expression (CTE) in SQL?

A. A table shared across multiple databases
B. A temporary named result set defined within a query using the WITH clause
C. A permanent table created with CREATE TABLE
D. A type of materialized view

Answer: B

Q907. What is the difference between a function and a stored procedure in SQL?

A. Functions must return a value and can be used in expressions while procedures may not return a value
B. Functions cannot take parameters
C. Procedures always return a value
D. There is no difference

Answer: A

Q908. What is dynamic SQL?

A. SQL that adjusts to different time zones
B. SQL statements constructed and executed at runtime as strings
C. SQL that automatically optimizes itself
D. SQL used only in dynamic websites

Answer: B

Q909. What is the WITH GRANT OPTION used for in SQL privileges?

A. To grant access for a limited time
B. To allow the grantee to further grant the same privileges to other users
C. To grant read-only access
D. To grant access to system tables

Answer: B

Q910. What is an updatable view in SQL?

A. Any view that has been recently modified
B. A view through which INSERT, UPDATE, and DELETE operations can be performed on the base tables
C. A view that refreshes automatically
D. A view with an auto-increment column

Answer: B

Q911. What is the purpose of the INSTEAD OF trigger?

A. To execute instead of a SELECT query
B. To replace the default DML action on a view with custom logic
C. To run instead of other triggers
D. To execute when an error occurs instead of raising the error

Answer: B

Q912. What is a row-level trigger versus a statement-level trigger?

A. Row-level triggers fire once per affected row while statement-level triggers fire once per SQL statement
B. Row-level triggers are faster
C. Statement-level triggers process each row individually
D. They execute at the same frequency

Answer: A

Q913. What is embedded SQL?

A. SQL stored inside HTML pages
B. SQL statements written within a host programming language like C or Java
C. SQL embedded in PDF documents
D. SQL that runs only on embedded systems

Answer: B

Q914. What is Boyce-Codd Normal Form (BCNF) and how does it differ from 3NF?

A. BCNF is the same as 3NF
B. BCNF is stricter than 3NF, requiring every determinant to be a candidate key
C. BCNF is weaker than 3NF
D. BCNF only applies to tables with composite keys

Answer: B

Q915. Can a relation be in 3NF but not in BCNF?

A. No, 3NF always implies BCNF
B. Yes, when a non-candidate key attribute determines part of a candidate key
C. No, BCNF is a weaker form than 3NF
D. Yes, but only in tables with no candidate keys

Answer: B

Q916. What is Fourth Normal Form (4NF)?

A. A table in BCNF with no trivial multivalued dependencies
B. A table with four or fewer columns
C. A table with exactly four candidate keys
D. A table that has been normalized four times

Answer: A

Q917. What is the purpose of decomposition in normalization?

A. To delete unnecessary data
B. To split a relation into smaller relations to eliminate anomalies while preserving data
C. To combine multiple tables into one
D. To encrypt data across multiple tables

Answer: B

Q918. What is a canonical cover (minimal cover) of functional dependencies?

A. The largest possible set of functional dependencies
B. A minimal set of functional dependencies equivalent to the original set with no redundancy
C. A set of dependencies that covers all tables
D. The first set of dependencies defined during design

Answer: B

Q919. When is denormalization appropriate in database design?

A. Always, as it improves all aspects of performance
B. When read performance is critical and the overhead of maintaining redundancy is acceptable
C. Never, as it violates database theory
D. Only when the database has fewer than 10 tables

Answer: B

Q920. What is the synthesis approach to normalization?

A. Combining all tables into one
B. Building relations from a given set of functional dependencies using a bottom-up approach
C. Creating tables randomly and then checking normal forms
D. Synthesizing data from external sources

Answer: B

Q921. What is a prime attribute?

A. The most important attribute in a table
B. An attribute that is part of any candidate key
C. An attribute that stores primary data
D. The first attribute defined in a table

Answer: B

Q922. What does the decomposition algorithm for BCNF guarantee?

A. Both lossless join and dependency preservation
B. Lossless join but not necessarily dependency preservation
C. Dependency preservation but not lossless join
D. Neither lossless join nor dependency preservation

Answer: B

Q923. What is a conflict in transaction schedules?

A. A disagreement between database administrators
B. When two operations from different transactions access the same data item and at least one is a write
C. When two transactions have the same name
D. When the database runs out of storage

Answer: B

Q924. What is conflict serializability?

A. A schedule where all conflicts are resolved
B. A schedule that is equivalent to a serial schedule by swapping non-conflicting operations
C. A schedule with no conflicts
D. A schedule where conflicts are ignored

Answer: B

Q925. How is conflict serializability tested using a precedence graph?

A. By counting the number of nodes
B. A schedule is conflict serializable if and only if its precedence graph has no cycles
C. By checking if the graph is fully connected
D. By verifying all nodes have equal edges

Answer: B

Q926. What is a cascading rollback?

A. Rolling back only the last operation of a transaction
B. When the rollback of one transaction forces rollbacks of other dependent transactions
C. Rolling back changes in alphabetical order
D. A rollback that happens automatically at midnight

Answer: B

Q927. What is the difference between a strict and a cascadeless schedule?

A. They are the same thing
B. Cascadeless prevents cascading rollbacks; strict additionally prevents reading uncommitted data items until the writing transaction commits or aborts
C. Strict schedules allow more concurrency
D. Cascadeless is stricter than strict schedules

Answer: B

Q928. What is view serializability?

A. A schedule based on database views
B. A schedule where each transaction sees the same data values as in some serial schedule
C. A schedule that only involves view operations
D. A schedule verified by viewing the database

Answer: B

Q929. What is a blind write in transactions?

A. A write performed by an unauthorized user
B. A write operation performed without first reading the data item
C. A write that produces no output
D. A write to a hidden column

Answer: B

Q930. What is strict two-phase locking (Strict 2PL)?

A. A 2PL variant where locks are acquired in strict order
B. A 2PL variant where all exclusive locks are held until the transaction commits or aborts
C. A 2PL variant that uses only shared locks
D. A 2PL variant with no growing phase

Answer: B

Q931. What is rigorous two-phase locking?

A. A 2PL variant that is more flexible
B. A 2PL variant where all locks (both shared and exclusive) are held until commit or abort
C. A 2PL variant that uses no locks
D. A 2PL variant for read-only transactions

Answer: B

Q932. What is the wait-die deadlock prevention scheme?

A. Older transactions wait; younger ones are rolled back if they would need to wait
B. Younger transactions always wait for older ones
C. Both transactions are rolled back
D. The database picks a random transaction to abort

Answer: A

Q933. What is the wound-wait deadlock prevention scheme?

A. Older transactions roll back younger ones; younger transactions wait for older ones
B. Younger transactions always get priority
C. Both transactions always wait
D. Deadlocks are detected after they occur

Answer: A

Q934. How is deadlock detection performed using a wait-for graph?

A. By counting the number of edges
B. A deadlock exists if and only if the wait-for graph contains a cycle
C. By checking if the graph has more than 10 nodes
D. By measuring the graph's depth

Answer: B

Q935. What is lock granularity in concurrency control?

A. The size of the lock data structure
B. The level at which locking is applied: database, table, page, or row level
C. The number of transactions that can hold a lock
D. The duration a lock is held

Answer: B

Q936. What is the trade-off between fine-grained and coarse-grained locking?

A. Fine-grained locking is always better
B. Fine-grained locking allows more concurrency but has higher lock management overhead
C. Coarse-grained locking allows more concurrency
D. There is no trade-off

Answer: B

Q937. What is an intention lock in multi-granularity locking?

A. A lock that indicates the transaction intends to lock the database
B. A lock placed on a higher-level node indicating that a finer-grained lock is held on a descendant
C. A lock that is planned but not yet acquired
D. A lock used during database design

Answer: B

Q938. What is a livelock in concurrency control?

A. A lock that keeps the database alive
B. A situation where a transaction repeatedly retries an operation but never succeeds
C. A lock on live streaming data
D. A lock that never expires

Answer: B

Q939. What is the ARIES recovery algorithm?

A. A simple backup tool
B. An advanced recovery algorithm using WAL with three phases: analysis, redo, and undo
C. An encryption algorithm for log files
D. A database compression algorithm

Answer: B

Q940. In ARIES, what does the analysis phase do?

A. Analyzes query performance
B. Scans the log from the last checkpoint to determine dirty pages and active transactions at crash time
C. Analyzes the database schema
D. Performs statistical analysis on data

Answer: B

Q941. What is the difference between steal and no-steal buffer management policies?

A. Steal allows modified pages to be flushed before commit while no-steal does not
B. Steal refers to taking data from other transactions
C. No-steal prevents any buffer modifications
D. They refer to how memory is allocated

Answer: A

Q942. What is the difference between force and no-force buffer management policies?

A. Force requires all modified pages to be written to disk at commit time while no-force does not
B. Force means using forced writes for better performance
C. No-force prevents any writes to disk
D. They refer to the physical force required to insert disks

Answer: A

Q943. What combination of buffer management policies does ARIES use?

A. Force, No-steal
B. No-force, Steal
C. Force, Steal
D. No-force, No-steal

Answer: B

Q944. What is a log sequence number (LSN)?

A. A serial number for database licenses
B. A unique identifier assigned to each log record to maintain ordering
C. A number identifying the database sequence
D. A network sequence number

Answer: B

Q945. What is physiological logging?

A. Logging based on physical health metrics
B. A logging approach that identifies pages physically but logs changes logically within each page
C. A form of logging used only for physical databases
D. Logging the physical location of data on disk

Answer: B

Q946. What is a compensation log record (CLR) in ARIES?

A. A log record for financial compensation
B. A log record written during undo to describe the reversal of an action
C. A log record that compensates for lost data
D. A log record for replication

Answer: B

Q947. What is the purpose of a fuzzy checkpoint?

A. A checkpoint with imprecise data
B. A checkpoint that does not halt transaction processing while recording the checkpoint information
C. A checkpoint performed at random intervals
D. A checkpoint that only saves partial data

Answer: B

Q948. What is a B+ tree and how does it differ from a B-tree?

A. They are identical structures
B. In a B+ tree, all data pointers are at the leaf level and leaves are linked, unlike B-trees where data pointers exist at all levels
C. A B+ tree is always smaller than a B-tree
D. B+ trees do not support range queries

Answer: B

Q949. What is a hash index?

A. An index that uses a hash function to map keys directly to their storage location
B. An index that stores data in alphabetical order
C. An index used only for text fields
D. An index that requires no storage space

Answer: A

Q950. What is a covering index?

A. An index that covers the entire disk
B. An index that contains all columns needed by a query, eliminating the need to access the base table
C. An index that covers all tables in the database
D. An index with a protective encryption layer

Answer: B

Q951. When would a hash index be preferred over a B+ tree index?

A. For range queries
B. For exact match (equality) lookups where range queries are not needed
C. For pattern matching queries
D. Hash indexes are always preferred

Answer: B

Q952. What is a bitmap index?

A. An index stored as an image file
B. An index that uses bit arrays to represent the presence or absence of values for each indexed column
C. An index on binary data columns
D. An index that maps bits to bytes

Answer: B

Q953. What is the order (or degree) of a B-tree?

A. The number of levels in the tree
B. The maximum number of children a node can have
C. The number of records in the tree
D. The total number of nodes

Answer: B

Q954. What is index selectivity?

A. The ability to select which index to use
B. The ratio of distinct indexed values to total number of records, indicating how well an index discriminates
C. How selective the DBA is when creating indexes
D. The speed of index selection

Answer: B

Q955. What is a non-clustered index?

A. An unorganized index
B. An index whose logical order differs from the physical order of the data rows
C. An index without a B-tree structure
D. An index that cannot be used for searching

Answer: B

Q956. What is the nested loop join algorithm?

A. A join that creates nested tables
B. For each row in the outer table, it scans the entire inner table to find matches
C. A join that only works with nested queries
D. A loop that nests multiple joins together

Answer: B

Q957. What is the sort-merge join algorithm?

A. A join that sorts results after merging
B. Both tables are sorted on the join attribute and then merged together in a single pass
C. A join that only works on sorted tables
D. A combination of sorting and hashing

Answer: B

Q958. What is the hash join algorithm?

A. A join that uses hash codes for security
B. One table is hashed into buckets on the join key, then the other table is probed against these buckets
C. A join that produces hashed output
D. A join between hash tables

Answer: B

Q959. What is the difference between a rule-based and cost-based query optimizer?

A. They produce identical plans
B. Rule-based uses predefined heuristic rules while cost-based uses statistical cost estimates to choose plans
C. Rule-based is always better
D. Cost-based does not use any rules

Answer: B

Q960. Why is pushing selections down in a query tree an important optimization?

A. It makes the query tree shorter
B. It reduces the size of intermediate results early, decreasing processing costs for subsequent operations
C. It removes the need for joins
D. It changes the query semantics for better results

Answer: B

Q961. What is the difference between materialization and pipelining in query execution?

A. They are the same technique
B. Materialization stores intermediate results on disk while pipelining passes results directly between operations
C. Pipelining is always slower
D. Materialization uses less memory

Answer: B

Q962. What are histograms used for in query optimization?

A. Displaying query results graphically
B. Providing detailed data distribution information to improve cardinality estimation accuracy
C. Showing database usage patterns
D. Graphing query execution times

Answer: B

Q963. What is cardinality estimation in query optimization?

A. Estimating the number of tables needed
B. Predicting the number of rows produced by each operation in the query plan
C. Estimating the number of users
D. Calculating the number of indexes required

Answer: B

Q964. What is an index-only scan?

A. A scan that creates an index
B. A query execution method that retrieves all needed data from the index without accessing the base table
C. A scan that only checks if an index exists
D. A scan of the index metadata

Answer: B

Q965. What is SQL injection and how can it be prevented?

A. A method to inject data faster; prevented by faster hardware
B. An attack inserting malicious SQL through user input; prevented by parameterized queries and input validation
C. An SQL optimization technique; prevented by using indexes
D. A database backup method; prevented by RAID

Answer: B

Q966. What is mandatory access control (MAC)?

A. Access control that all users must follow voluntarily
B. Access control enforced by the system based on security labels and clearance levels, not user discretion
C. Optional access restrictions
D. Access control managed by users themselves

Answer: B

Q967. What is discretionary access control (DAC)?

A. Access control based on time of day
B. Access control where data owners grant or revoke permissions to other users at their discretion
C. Access control enforced by the operating system only
D. Access control that cannot be changed

Answer: B

Q968. What is the Bell-LaPadula model in database security?

A. A model for database performance tuning
B. A multilevel security model that enforces no read up and no write down policies
C. A model for database backup
D. A relational data model

Answer: B

Q969. What is the difference between encryption at rest and encryption in transit?

A. They are the same type of encryption
B. At rest protects stored data on disk while in transit protects data moving across networks
C. At rest encrypts network data while in transit encrypts stored data
D. At rest is stronger than in transit

Answer: B

Q970. What is a privilege cascade in database security?

A. A waterfall display of privileges
B. When revoking a privilege from a user also revokes it from users who received it through that user
C. Cascading permissions to all database objects
D. A list of privileges sorted by priority

Answer: B

Q971. What is column-level security?

A. Security that applies to all columns equally
B. Restricting user access to specific columns within a table rather than the entire table
C. Encryption of column headers
D. A security column added to every table

Answer: B

Q972. What is row-level security (RLS)?

A. Security applied to each row of the audit log
B. Filtering rows based on user attributes so users can only see rows they are authorized to access
C. Encrypting individual rows
D. Adding a security row to each table

Answer: B

Q973. What is the two-phase commit (2PC) protocol?

A. A commitment to finish in two phases of development
B. A protocol that coordinates distributed transactions by having a prepare phase and a commit/abort phase
C. A two-step backup process
D. A protocol for connecting two database servers

Answer: B

Q974. What is the role of the coordinator in the two-phase commit protocol?

A. To optimize queries across sites
B. To initiate the commit protocol, collect votes from participants, and make the final commit or abort decision
C. To balance load across database sites
D. To manage user authentication

Answer: B

Q975. What is the difference between full replication and partial replication?

A. Full replication stores more data per site
B. Full replication copies the entire database to every site while partial replication copies only selected fragments
C. Full replication is always faster
D. Partial replication is not possible in distributed databases

Answer: B

Q976. What is a distributed catalog in a distributed database?

A. A catalog distributed to all users
B. Metadata about the location, fragmentation, and replication of data across sites
C. A printed catalog of database features
D. A central list of all distributed databases

Answer: B

Q977. What is the global deadlock problem in distributed databases?

A. A deadlock affecting the entire internet
B. A deadlock involving transactions at multiple sites that cannot be detected by any single site's local deadlock detector
C. A deadlock in the global catalog
D. A worldwide database outage

Answer: B

Q978. What is mixed fragmentation in distributed databases?

A. A mix of correct and corrupted data
B. A combination of horizontal and vertical fragmentation applied to a relation
C. Mixing data from different databases
D. Fragmenting data in a random order

Answer: B

Q979. What is the principle of local autonomy in distributed databases?

A. Each site must be autonomous from the network
B. Each site maintains control over its local data and can operate independently
C. Users can autonomously create databases
D. Local sites do not need backup

Answer: B

Q980. How does MongoDB store data internally?

A. In relational tables
B. In BSON (Binary JSON) documents organized into collections
C. In key-value pairs only
D. In column families

Answer: B

Q981. What is sharding in NoSQL databases?

A. Breaking data into pieces and discarding some
B. Distributing data across multiple servers by partitioning it based on a shard key
C. Creating backup copies of data
D. Encrypting data into shards

Answer: B

Q982. What is the difference between embedded documents and references in MongoDB?

A. They are the same concept
B. Embedded documents nest related data within a document while references store relationships using document IDs
C. Embedded documents are faster to delete
D. References are used only for primary keys

Answer: B

Q983. What is the Cassandra data model?

A. A relational model with SQL support
B. A partitioned row store using partition keys for distribution and clustering keys for sorting within partitions
C. A pure key-value model
D. A document-based model

Answer: B

Q984. What is the CAP theorem's significance for NoSQL database design?

A. It has no relevance to NoSQL databases
B. It forces NoSQL designers to choose between consistency, availability, and partition tolerance trade-offs
C. It guarantees all three properties for NoSQL databases
D. It only applies to SQL databases

Answer: B

Q985. What is a time-series database?

A. A database that tracks the time of each query
B. A specialized database optimized for storing and querying time-stamped data points
C. A database that only works during certain times
D. A temporary database

Answer: B

Q986. What is the aggregation framework in MongoDB?

A. A framework for adding data
B. A pipeline-based data processing framework that transforms and analyzes documents through stages
C. A framework for managing database clusters
D. A tool for database administration

Answer: B

Q987. What is the difference between SCD Type 1 and SCD Type 2?

A. They handle dimension changes identically
B. Type 1 overwrites old values while Type 2 creates new rows to preserve history
C. Type 1 preserves history while Type 2 does not
D. Type 1 is for fact tables while Type 2 is for dimension tables

Answer: B

Q988. What is a surrogate key in a data warehouse?

A. A natural business key
B. A system-generated artificial key used as the primary key in dimension tables instead of natural keys
C. A temporary key used during ETL
D. A backup of the primary key

Answer: B

Q989. What is the difference between MOLAP, ROLAP, and HOLAP?

A. They are different programming languages
B. MOLAP uses multidimensional storage, ROLAP uses relational storage, and HOLAP combines both approaches
C. They are different database vendors
D. MOLAP is the fastest and ROLAP is the slowest

Answer: B

Q990. What is a conformed dimension in data warehousing?

A. A dimension that conforms to naming conventions
B. A dimension that has the same meaning and content when referenced across multiple fact tables or data marts
C. A dimension that has been validated
D. A dimension with confirmed data

Answer: B

Q991. What is a surrogate key and when is it used in relational design?

A. A key borrowed from another table
B. A system-generated artificial key used when no natural candidate key is suitable
C. A temporary key used during testing
D. A key that replaces foreign keys

Answer: B

Q992. What is the extended projection operation in relational algebra?

A. A projection that includes all columns
B. A projection that allows arithmetic expressions and string operations on attributes in the result
C. A projection that extends the table with new rows
D. A projection applied to multiple tables

Answer: B

Q993. What is the assignment operation used for in relational algebra?

A. Assigning users to database roles
B. Storing the result of a relational algebra expression in a temporary relation variable
C. Assigning data types to attributes
D. Assigning primary keys to tables

Answer: B

Q994. What is the purpose of the IN operator in SQL?

A. To insert new rows
B. To check if a value matches any value in a specified list or subquery
C. To index a column
D. To include all columns in the result

Answer: B

Q995. What is the difference between INNER JOIN and CROSS JOIN in SQL?

A. They are identical
B. INNER JOIN matches rows based on a condition while CROSS JOIN produces all combinations of rows from both tables
C. CROSS JOIN requires a join condition
D. INNER JOIN produces more rows

Answer: B

Q996. What is the NULL-safe equality operator in SQL?

A. The = operator handles NULL automatically
B. The IS NOT DISTINCT FROM operator that treats two NULLs as equal
C. There is no such operator
D. The <=> operator that ignores NULL values

Answer: B

Q997. What is a materialized view refresh strategy?

A. Refreshing the browser to view database changes
B. The method used to update a materialized view's stored data, either completely (full) or with only changes (incremental)
C. Clearing the view cache
D. Restarting the database to refresh views

Answer: B

Q998. What is an extraneous attribute in a functional dependency?

A. An attribute that is extra in the table
B. An attribute that can be removed from the left or right side of a dependency without changing the closure
C. An attribute added by mistake
D. An attribute from an external database

Answer: B

Q999. What is the non-repeatable read problem?

A. A read operation that takes too long
B. When a transaction reads the same data item twice and gets different values because another transaction modified it
C. A read that cannot be repeated due to errors
D. A read of non-existent data

Answer: B

Q1000. What is the READ UNCOMMITTED isolation level?

A. A level where no reads are allowed
B. The lowest isolation level that allows dirty reads, non-repeatable reads, and phantom reads
C. A level that prevents all anomalies
D. A level that only allows uncommitted writes

Answer: B

Q1001. What is the difference between an aborted transaction and a failed transaction?

A. They are the same state
B. A failed transaction has encountered an error and cannot proceed, while aborted means it has been rolled back
C. Aborted transactions can resume while failed ones cannot
D. Failed transactions are committed while aborted ones are not

Answer: B

Q1002. What is lock compatibility in concurrency control?

A. Whether locks are compatible with the hardware
B. The rules determining which lock types can coexist on the same data item, such as shared locks being compatible with each other
C. Whether locks can be used across databases
D. Compatibility of lock software versions

Answer: B

Q1003. What is a differential backup and how does it differ from an incremental backup?

A. They are identical
B. A differential backup captures all changes since the last full backup while an incremental captures changes since the last backup of any type
C. Differential is smaller than incremental
D. Incremental captures more data

Answer: B

Q1004. What is a unique index?

A. An index on a unique table
B. An index that enforces uniqueness on the indexed columns, preventing duplicate values
C. The only index in the database
D. An index with a unique name

Answer: B

Q1005. What is index fragmentation and how does it affect performance?

A. Physical breaking of the index file
B. When index pages become disordered or have excessive free space due to insertions and deletions, degrading scan performance
C. Splitting an index across databases
D. Creating too many index levels

Answer: B

Q1006. What is the selectivity of a predicate in query optimization?

A. How selective the DBA is about indexing
B. The fraction of rows in a table that satisfy the predicate condition, used to estimate result sizes
C. The order in which predicates are evaluated
D. The number of predicates in a query

Answer: B

Q1007. What is a database honeypot?

A. A storage container for database files
B. A decoy database system designed to attract and detect unauthorized access attempts
C. A type of database optimization
D. A backup database server

Answer: B

Q1008. What is the principle of separation of duties in database security?

A. Separating the database into different files
B. Dividing critical security functions among different individuals so no single person has complete control
C. Separating development and production databases
D. Keeping security logs separate from data

Answer: B

Q1009. What is a distributed join operation?

A. Joining multiple distributed systems
B. A join operation that requires data from multiple sites, involving data transfer across the network
C. Joining two local tables
D. A social gathering of database administrators

Answer: B

Q1010. What is the difference between tight coupling and loose coupling in distributed databases?

A. They refer to how tightly cables connect servers
B. Tight coupling shares resources like memory and is more integrated, while loose coupling connects independent systems via messages
C. Tight coupling is always faster
D. There is no practical difference

Answer: B

Q1011. What is the catalog management challenge in distributed databases?

A. Managing physical catalogs of database products
B. Deciding where to store the distributed catalog (centralized, replicated, or partitioned) while balancing lookup speed and update cost
C. Organizing database documentation
D. Managing vendor catalogs

Answer: B

Q1012. What is an edge in a graph database?

A. The boundary of the database
B. A connection between two nodes representing a relationship, which can have its own properties
C. The last record in a collection
D. A database boundary condition

Answer: B

Q1013. What is the replica set in MongoDB?

A. A set of replicated queries
B. A group of MongoDB instances that maintain the same data set, providing redundancy and high availability
C. A collection of duplicate documents
D. A set of backup scripts

Answer: B

Q1014. What is the MapReduce programming model used in NoSQL?

A. A model for mapping database locations
B. A data processing paradigm where Map transforms data into key-value pairs and Reduce aggregates values by key
C. A model for reducing database size
D. A navigation model for data maps

Answer: B

Q1015. What is a pivot operation in OLAP?

A. Rotating the OLAP server hardware
B. Rotating the data cube to view data from a different perspective by swapping dimensions on axes
C. Pivoting from one database to another
D. Changing the database server

Answer: B

Q1016. What is the difference between additive, semi-additive, and non-additive measures?

A. They refer to the number of measures in a fact table
B. Additive measures can be summed across all dimensions, semi-additive across some, and non-additive cannot be meaningfully summed
C. Additive measures add rows while non-additive do not
D. They are different data types for measures

Answer: B

Q1017. What is a role-playing dimension?

A. A dimension used in role-based access control
B. A single physical dimension table referenced multiple times in a fact table with different meanings, like date used as order date and ship date
C. A dimension for storing user roles
D. A dimension that changes behavior

Answer: B

Q1018. What is a galaxy schema (fact constellation)?

A. A schema based on astronomical data
B. A data warehouse schema containing multiple fact tables that share common dimension tables
C. A schema with star-shaped dimensions
D. A schema used only in space databases

Answer: B

Q1019. What is ELT and how does it differ from ETL?

A. They are the same process
B. ELT loads raw data into the target first then transforms it there, while ETL transforms data before loading
C. ELT is faster because it skips extraction
D. ETL loads data while ELT extracts it

Answer: B

Q1020. What is a dimension hierarchy in a data warehouse?

A. The importance ranking of dimensions
B. A logical structure within a dimension that defines levels of aggregation, such as day to month to quarter to year
C. A hierarchy of dimension tables
D. The order in which dimensions are created

Answer: B

Hard Questions

510 questions

Q1021. In the ANSI/SPARC three-schema architecture, which mapping exists between the conceptual and internal schemas?

A. Application/Internal mapping
B. External/Conceptual mapping
C. Conceptual/Internal mapping
D. External/Internal mapping

Answer: C

Q1022. Which of the following best describes the impedance mismatch problem?

A. Differences between database model and programming language data types
B. Mismatch between primary key constraints and foreign key references
C. Slow network connectivity between the client and the database server
D. Hardware incompatibility between different types of storage devices

Answer: A

Q1023. Self-describing nature of a database system means:

A. The system contains metadata along with the data itself
B. Users do not need to interact with it in any way
C. The database automatically generates its own queries
D. Data stored in the system is automatically encrypted

Answer: A

Q1024. Program-data independence is primarily achieved through:

A. Avoiding the use of all index structures
B. Using only one single programming language
C. Data abstraction and catalog separation
D. Upgrading to faster processing hardware

Answer: C

Q1025. Which of the following is a consequence of violating data independence?

A. Concurrent transaction handling is improved across the system
B. Data security and access controls are substantially improved
C. The system achieves significantly faster query processing speed
D. Application programs must be modified when data storage changes

Answer: D

Q1026. The database approach provides controlled redundancy by:

A. Encrypting all data stored in the system
B. Avoiding the use of keys or constraints
C. Integrating data and applying normalization
D. Storing data in multiple separate files

Answer: C

Q1027. Which type of DBMS supports complex data types like spatial and multimedia data?

A. Flat file based DBMS
B. Network model DBMS
C. Hierarchical tree DBMS
D. Object-Relational DBMS

Answer: D

Q1028. In a multi-user DBMS environment, the concurrency control subsystem is essential to:

A. Perform scheduled database backup operations to external storage
B. Ensure correct results when multiple users access data simultaneously
C. Speed up the processing of single-user queries on a local database
D. Manage the allocation of available disk space across the system

Answer: B

Q1029. Which of the following database utilities is used to reorganize data for better performance?

A. Statistical analysis utility
B. Backup restore utility
C. Reorganization utility
D. Data loading utility

Answer: C

Q1030. Data abstraction in DBMS hides complexity at different levels. Which sequence is correct from lowest to highest?

A. External !’ Conceptual !’ Internal
B. Conceptual !’ Internal !’ External
C. External !’ Internal !’ Conceptual
D. Internal !’ Conceptual !’ External

Answer: D

Q1031. In a shared-disk parallel architecture, the potential bottleneck is:

A. The interconnection network to the shared disk
B. The CPU processing speed of each processor
C. Query parsing and syntax validation steps
D. Memory allocation among running transactions

Answer: A

Q1032. The query optimizer in a DBMS uses which approach to find the best execution plan?

A. Cost-based optimization
B. Random selection method
C. User preference settings
D. Alphabetical table ordering

Answer: A

Q1033. Which type of parallelism involves executing different operations of a query simultaneously?

A. Inter-query parallelism
B. None of these types
C. Intra-query parallelism
D. Pipeline parallelism

Answer: A

Q1034. In a DBMS architecture, the catalog manager is responsible for:

A. Managing user connections and authentication sessions
B. Handling network protocols and data transmission tasks
C. Processing application logic and business rule validation
D. Storing and retrieving schema information and statistics

Answer: D

Q1035. The concept of data independence is directly supported by:

A. Using only a single schema without any additional views
B. Storing all data exclusively in flat file format on disk
C. Avoiding the use of views or external schema definitions
D. The three-level architecture with mappings between levels

Answer: D

Q1036. In a DBMS, the lock manager is part of which subsystem?

A. Storage management subsystem
B. Communication layer subsystem
C. Concurrency control subsystem
D. Query processing subsystem

Answer: C

Q1037. What is the main advantage of shared-memory parallel architecture?

A. Lower hardware costs compared to other parallel architectures
B. Fast communication between processors through shared memory
C. No synchronization needed between the concurrent processors
D. No need for an operating system to coordinate the processes

Answer: B

Q1038. The DBMS buffer replacement strategy LRU stands for:

A. Least Recently Used
B. Last Record Updated
C. Latest Read Utility
D. Least Required Update

Answer: A

Q1039. Which DBMS component is responsible for enforcing authorization and integrity constraints?

A. Network manager and router module
B. Query parser and syntax analyzer
C. Storage manager and file handler
D. Authorization and integrity manager

Answer: D

Q1040. In a multi-tier architecture, which tier contains the business rules?

A. Data tier (bottom tier)
B. Application tier (middle tier)
C. Client tier (front tier)
D. Presentation tier (top tier)

Answer: B

Q1041. In the EER model, an overlapping specialization means:

A. An entity can belong to multiple subclasses
B. No entity can be specialized at all
C. An entity belongs to exactly one subclass
D. Subclasses have no common superclass type

Answer: A

Q1042. A category (union type) in the EER model represents:

A. A superclass entity that has no subclasses
B. An entity defined without any attributes
C. A relationship with no participants at all
D. A subclass with multiple possible superclasses

Answer: D

Q1043. The concept of aggregation in the ER model is used to:

A. Treat a relationship set as a higher-level entity
B. Create weak entities that depend on other types
C. Delete entities from the conceptual data model
D. Merge two databases into a single schema model

Answer: A

Q1044. In a ternary relationship among three entities A, B, and C, the cardinality constraint is specified:

A. Only for the strongest entity participating here
B. Only for one entity in the ternary relationship
C. Cardinality does not apply to ternary relationships
D. For each pair of entities with respect to the third

Answer: D

Q1045. Disjoint constraint (d) on specialization means:

A. An entity can be a member of at most one subclass
B. Subclasses can overlap with shared memberships
C. There is no constraint on subclass membership
D. An entity must belong to all subclasses at once

Answer: A

Q1046. Attribute inheritance in EER means:

A. Attributes are shared across all entities uniformly
B. Subclass entities inherit all attributes of the superclass
C. Attributes cannot be inherited between entity types
D. Superclass inherits all attributes from the subclass

Answer: B

Q1047. A lattice in the EER model occurs when:

A. An entity has no attributes or properties defined on it
B. A superclass has no subclasses defined in the model at all
C. A subclass has more than one superclass (multiple inheritance)
D. All entities in the model are weak entity types entirely

Answer: C

Q1048. The Object-Relational model extends the relational model by adding:

A. Hierarchical parent-child relationships between records
B. Network-style pointers linking records across the model
C. Only more tables and relations without other extensions
D. Support for complex data types, methods, and inheritance

Answer: D

Q1049. In the Entity-Relationship model, the participation constraint and cardinality ratio together define:

A. Query optimization parameters used
B. Physical storage structure of tables
C. Only the key constraints of entities
D. Structural constraints of a relationship

Answer: D

Q1050. A recursive relationship is one where:

A. No entity type participates in the relationship that is defined
B. The same entity type participates more than once in different roles
C. Only weak entity types participate in the relationship instance
D. Two different entity types participate together in one relationship

Answer: B

Q1051. Codd's Rule 0 (Foundation Rule) states:

A. SQL is the only query language allowed for use within the database system
B. Indexes are mandatory for every attribute present in all database tables
C. All data must be stored exclusively in flat files without any table structure
D. A relational DBMS must manage data entirely through its relational capabilities

Answer: D

Q1052. The information rule (Codd's Rule 1) states that:

A. Pointers can be used to represent data relationships
B. All information must be represented as values in tables
C. Metadata need not be stored in tables or relations
D. Data can be stored in external files outside the system

Answer: B

Q1053. A relation is in BCNF if:

A. For every non-trivial functional dependency X to Y, X is a superkey
B. All attributes in the relation are prime (part of some key)
C. It has no multivalued dependencies among any of its attributes
D. It satisfies only the requirements of first normal form (1NF)

Answer: A

Q1054. The closure of a set of attributes X (X+) under a set of functional dependencies is:

A. The set of all foreign keys that reference attribute X
B. The set of all attributes functionally determined by X
C. The total number of individual attributes within X
D. The set of all tables in the database containing X

Answer: B

Q1055. Armstrong's axioms include all EXCEPT:

A. Augmentation
B. Reflexivity
C. Distribution
D. Transitivity

Answer: C

Q1056. A lossless-join decomposition guarantees that:

A. The original relation can be reconstructed by joining the decomposed relations
B. No foreign keys are ever needed between the resulting decomposed relations
C. Join operations are always faster when applied to the decomposed relations
D. Some data will always be permanently lost during the decomposition process

Answer: A

Q1057. The dependency preservation property ensures that:

A. Only primary key dependencies are preserved during the decomposition of tables
B. All functional dependencies can be checked within individual decomposed relations
C. No dependencies are maintained after decomposition of the original relation
D. Dependencies are stored separately from the decomposed relation structures

Answer: B

Q1058. A surrogate key is:

A. A system-generated unique identifier with no business meaning
B. A natural key derived directly from the business domain
C. A composite key made up of multiple business attributes
D. A foreign key referencing a primary key in another table

Answer: A

Q1059. Codd defined how many rules for a fully relational database?

A. 12 rules (numbered 0-12, so 13 total)
B. 10 rules (numbered 1 through 10 only)
C. 5 rules (covering the basic properties)
D. 20 rules (for comprehensive coverage)

Answer: A

Q1060. The Closed World Assumption in relational databases means:

A. Facts not present in the database are considered false
B. Nothing stored in the database is considered to be true
C. All possible facts about the domain are stored in it
D. The database cannot be modified once it has been created

Answer: A

Q1061. A query language is relationally complete if it can express:

A. All queries expressible in relational calculus
B. Only join operations between related tables
C. Only aggregate functions on grouped data
D. Only selection and projection operations used

Answer: A

Q1062. The aggregate operation in extended relational algebra:

A. Performs Cartesian products between two unrelated relations
B. Renames relations and their associated attributes in queries
C. Selects individual tuples based on a simple filter condition
D. Computes functions like SUM, AVG, COUNT on groups of tuples

Answer: D

Q1063. Which of the following cannot be expressed in basic relational algebra?

A. Projection of attributes from relations
B. Selection with various filter conditions
C. Recursive queries like transitive closure
D. Natural join of two compatible relations

Answer: C

Q1064. The concept of safe expressions in relational calculus ensures:

A. Queries always run faster than other approaches
B. No NULL values appear in the query results
C. All tuples in the relation are always selected
D. The result of a query is a finite set of tuples

Answer: D

Q1065. In tuple relational calculus, the expression {t | P(t)} means:

A. Update tuples in the relation where P is a false value
B. Delete all tuples from the relation where P is true value
C. The set of all tuples t such that predicate P(t) is true
D. Insert tuples into the relation that satisfy predicate P

Answer: C

Q1066. The outer union operation:

A. Only works with relations that have identical schemas and matching attributes
B. Is the same as regular union and produces exactly identical results every time
C. Takes the union of relations that are not union-compatible by padding with NULLs
D. Removes all NULL values from the result set after combining both relations

Answer: C

Q1067. Full outer join preserves:

A. Only tuples from the left relation with NULL for non-matching rows
B. Only tuples from the right relation with NULL for non-matching rows
C. All tuples from both relations, padding with NULLs where necessary
D. Only tuples that have matching values in both of the joined tables

Answer: C

Q1068. The multiset (bag) version of relational algebra differs from set-based algebra in that:

A. It removes all duplicates
B. It allows duplicate tuples
C. It does not support joins
D. It only supports selection

Answer: B

Q1069. Which algebraic equivalence rule is used for query optimization?

A. Removing all joins from the query plan
B. Pushing selections down the query tree
C. Eliminating projections from the query
D. Reversing the order of query operations

Answer: B

Q1070. The assignment operation in relational algebra is used to:

A. Delete relations permanently from the database schema and storage system
B. Assign the result of a relational algebra expression to a temporary variable
C. Modify existing tuples by updating their attribute values in the table
D. Create new databases and define their schema structure from scratch

Answer: B

Q1071. A correlated subquery is different from a regular subquery because it:

A. References columns from the outer query and executes once per outer row
B. Does not reference the outer query and runs completely independently
C. Is always faster than any other type of subquery in all situations
D. Executes only once regardless of how many rows the outer query returns

Answer: A

Q1072. The COALESCE function in SQL:

A. Calculates the average of numeric values
B. Returns the last value in a specified list
C. Concatenates strings from multiple columns
D. Returns the first non-NULL value in a list

Answer: D

Q1073. What is the difference between DELETE and TRUNCATE?

A. TRUNCATE can have a WHERE clause to filter which specific rows should be removed from tables
B. DELETE is faster than TRUNCATE because it removes all rows without any logging of the changes
C. They are identical in function and there is no difference between them in any database system
D. DELETE can have a WHERE clause and is logged; TRUNCATE removes all rows and is minimally logged

Answer: D

Q1074. The CASE expression in SQL is used for:

A. Managing concurrent transactions
B. Dropping indexes from a table
C. Creating tables in the database
D. Conditional logic within a query

Answer: D

Q1075. A CROSS JOIN produces:

A. An empty result set with no rows
B. The intersection of two data tables
C. Only rows matching between tables
D. The Cartesian product of two tables

Answer: D

Q1076. Which SQL statement is used to create an index?

A. CREATE INDEX
B. ADD INDEX
C. INSERT INDEX
D. MAKE INDEX

Answer: A

Q1077. The EXISTS operator in SQL:

A. Returns TRUE if the subquery returns at least one row
B. Checks for NULL values in a specified column only
C. Returns the count of rows matching in the subquery
D. Creates a new table based on the subquery results

Answer: A

Q1078. A self-join is:

A. A cross join of two tables
B. An outer join between tables
C. A join of a table with itself
D. A join without any conditions

Answer: C

Q1079. UNION vs UNION ALL in SQL:

A. They are identical in behavior and produce exactly the same result set
B. UNION ALL removes duplicates while UNION keeps all including duplicates
C. UNION removes duplicates; UNION ALL keeps all rows including duplicates
D. UNION is faster because it skips the deduplication processing entirely

Answer: C

Q1080. What does the NVL/IFNULL/ISNULL function do?

A. Replaces NULL with a specified value
B. Creates NULL constraints on columns
C. Deletes all rows containing NULLs
D. Counts NULL values in the column

Answer: A

Q1081. The DENSE_RANK() function differs from RANK() in that:

A. They are identical in their ranking behavior
B. DENSE_RANK() only works with string values
C. DENSE_RANK() skips ranks after ties occur
D. DENSE_RANK() does not skip ranks after ties

Answer: D

Q1082. Lateral joins (LATERAL keyword) allow:

A. The right side of the join to reference columns from the left side
B. Only cross joins to be performed without any join filter condition
C. Joining tables without specifying any condition between the rows
D. Only inner joins to be used between the two tables in the query

Answer: A

Q1083. The PARTITION BY clause in window functions:

A. Creates new tables from partitioned subsets of the original data
B. Divides the result set into partitions for function calculation
C. Deletes rows from specific partitions based on filter conditions
D. Physically partitions a table into separate storage file segments

Answer: B

Q1084. Dynamic SQL refers to:

A. SQL without variables or parameters of any kind
B. SQL statements constructed and executed at runtime
C. SQL that runs only during the compilation phase
D. Static predefined queries that never change at all

Answer: B

Q1085. The PIVOT operation in SQL:

A. Joins two tables together
B. Transforms rows into columns
C. Transforms columns into rows
D. Creates an index on data

Answer: B

Q1086. SQL injection can be prevented by:

A. Using parameterized queries and prepared statements
B. Storing user passwords in plain text in database
C. Disabling authentication on the database server
D. Using dynamic SQL without any input validation

Answer: A

Q1087. A user-defined function (UDF) in SQL differs from a stored procedure in that:

A. A UDF must return a value and can be used in SELECT statements
B. A stored procedure can be used directly within SELECT queries
C. A UDF cannot return any values to the calling SQL statement
D. They are identical in every way with no functional difference

Answer: A

Q1088. The GROUPING SETS clause in SQL:

A. Groups all rows into one single group regardless of the criteria
B. Removes grouping and returns individual rows from the result set
C. Allows specifying multiple grouping combinations in a single query
D. Sorts the result set based on the specified column or expression

Answer: C

Q1089. ROLLUP and CUBE in SQL GROUP BY are used for:

A. Generating subtotals and grand totals across multiple dimensions
B. Sorting data within each group by a specified column expression
C. Creating indexes on columns that are used in GROUP BY clauses
D. Deleting grouped data from the tables based on aggregate values

Answer: A

Q1090. Temporal tables in SQL track:

A. Network latency and response metrics
B. Only the current state of stored data
C. Deleted tables and their old schemas
D. Historical changes to data over time

Answer: D

Q1091. Fifth Normal Form (5NF) eliminates:

A. Join dependencies that are not implied by candidate keys
B. Multivalued dependencies among the stored data attributes
C. Functional dependencies between attributes in the table
D. All types of dependencies regardless of their categories

Answer: A

Q1092. A relation is in Domain-Key Normal Form (DKNF) if:

A. Every constraint is a logical consequence of domain constraints and key constraints
B. It satisfies only the requirements of first normal form and nothing beyond that
C. All attributes in the relation must serve as keys for unique identification
D. It has no constraints of any type defined on its attributes or relationships

Answer: A

Q1093. The synthesis algorithm for 3NF decomposition:

A. Always produces a BCNF decomposition that preserves all functional dependencies
B. Uses the canonical cover to create a dependency-preserving, lossless decomposition
C. Only works for binary relations with exactly two attributes in the schema
D. Removes all dependencies from the relation during the decomposition procedure

Answer: B

Q1094. It is possible to have a decomposition that is lossless but NOT dependency-preserving in:

A. BCNF decomposition
B. 1NF decomposition
C. Any decomposition
D. No decomposition

Answer: A

Q1095. The chase algorithm is used to:

A. Create indexes on frequently queried columns
B. Optimize storage allocation across disk pages
C. Test whether a decomposition is lossless
D. Test whether a query is syntactically correct

Answer: C

Q1096. A trivial functional dependency is one where:

A. The right side is a subset of the left side
B. The left side is empty with no attributes
C. No attributes are involved in the FD
D. Both sides are identical key attributes

Answer: A

Q1097. An extraneous attribute in an FD X !' Y is an attribute that:

A. Can be removed from X or Y without changing the closure
B. Cannot be removed under any circumstance from the FD
C. Must always be present in the functional dependency rule
D. Is a primary key attribute that cannot ever be removed

Answer: A

Q1098. Normalization theory is based on which mathematical concept?

A. Functional dependencies and their properties
B. Probability theory and statistical methods
C. Set theory only without other foundations
D. Graph theory only without other foundations

Answer: A

Q1099. When BCNF decomposition is not dependency-preserving, the alternative is:

A. Go to 4NF which eliminates all multivalued dependencies from the relation
B. Accept 3NF which guarantees both lossless join and dependency preservation
C. Denormalize completely to avoid any decomposition of the original table
D. Use 1NF only which requires just atomic values in every attribute field

Answer: B

Q1100. A fully functionally dependent attribute means:

A. Removing any attribute from the left side of the FD destroys the dependency
B. The attribute has no dependencies of any kind on other attributes at all
C. The attribute is part of the key and contributes to unique identification
D. The attribute depends on a single column only in the current relation schema

Answer: A

Q1101. View serializability is:

A. Not related to serializability in any way
B. Less restrictive than conflict serializability
C. The same as conflict serializability criteria
D. More restrictive than conflict serializability

Answer: B

Q1102. A recoverable schedule ensures that:

A. A transaction commits only after all transactions whose data it read have committed
B. All transactions must run serially with no concurrent execution allowed at all
C. Transactions can commit in any order regardless of their data reading dependencies
D. Rollbacks are never needed because all transactions always complete successfully

Answer: A

Q1103. A cascadeless schedule is one where:

A. All transactions must abort immediately
B. Transactions only read committed data
C. Transactions can read uncommitted data
D. No reads are allowed in any transaction

Answer: B

Q1104. A strict schedule requires that:

A. No transaction reads or writes data written by an uncommitted transaction
B. Transactions can freely read uncommitted data from other transactions
C. Commits are not required for transactions to be considered complete
D. Only serial schedules are allowed for all transactions in the system

Answer: A

Q1105. The precedence graph for testing conflict serializability has:

A. Nodes for SQL statements and edges for their execution ordering
B. Nodes for transactions and edges for conflicting operation pairs
C. Nodes for data items and edges for transactions accessing them
D. No nodes or edges are used in this graph representation at all

Answer: B

Q1106. Testing view serializability is:

A. NP-complete
B. Linear time
C. Logarithmic time
D. Constant time

Answer: A

Q1107. The READ COMMITTED isolation level prevents:

A. No anomalies of any kind
B. All anomalies including phantoms
C. Only dirty reads alone
D. Dirty and non-repeatable reads

Answer: C

Q1108. The REPEATABLE READ isolation level prevents:

A. All anomalies including phantom reads
B. Dirty reads and non-repeatable reads
C. Only dirty reads from transactions
D. No anomalies of any kind at all

Answer: B

Q1109. In a multi-version concurrency control (MVCC) system:

A. Transactions must be serial with no concurrent execution allowed at all
B. Multiple versions of data items are maintained for concurrent access
C. Only one version of each data item exists at any given point in time
D. Locks are always required for every read and write operation performed

Answer: B

Q1110. A blind write is:

A. A committed transaction that has been successfully applied
B. A query optimization technique that reduces disk access
C. A read operation that is not followed by any write at all
D. A write operation without reading the current value first

Answer: D

Q1111. Multiple Granularity Locking uses intention locks. An Intention Shared (IS) lock means:

A. No lower-level locks exist on any of the descendant nodes in the tree
B. The item cannot be locked by any transaction under any circumstances
C. A transaction intends to acquire shared locks on finer-granularity items
D. A transaction has exclusively locked the entire item and all subitems

Answer: C

Q1112. An Intention Exclusive (IX) lock means:

A. A transaction currently has a shared lock on the item and its subitems
B. No locks are needed for any data items under this locking protocol
C. A transaction intends to acquire exclusive locks on finer-granularity items
D. The transaction is aborted because it cannot acquire the needed lock

Answer: C

Q1113. A Shared Intention Exclusive (SIX) lock means:

A. No intention locks exist on any node in the hierarchy and all nodes are freely open
B. The node is shared-locked and the transaction intends to exclusively lock finer items
C. The node is exclusively locked by the transaction and no other can access it at all
D. The transaction is complete and all locks have been released from every node here

Answer: B

Q1114. In the Thomas Write Rule (modification to timestamp ordering):

A. Obsolete writes are ignored instead of causing rollback
B. All writes cause rollback regardless of timestamp order
C. Reads are ignored and only writes are processed by it
D. Timestamps are not used in this modified protocol

Answer: A

Q1115. The phantom problem in locking can be solved by:

A. Allowing dirty reads between queries
B. Using only row-level locks in tables
C. Index locking or predicate locking
D. Removing all indexes from the tables

Answer: C

Q1116. In multi-version timestamp ordering (MVTO):

A. Only one version of each data item is maintained
B. Reads always block writes until they are completed
C. No timestamps are used for ordering transactions
D. Each write creates a new version of the data item

Answer: D

Q1117. The wound-wait scheme is a preemptive protocol because:

A. Younger transactions always preempt older ones holding locks on data items
B. No rollback ever occurs because all transactions complete without conflict
C. An older transaction can force a younger one holding a needed lock to roll back
D. Only reads are preempted while writes always proceed without interruption

Answer: C

Q1118. In tree locking protocol (for B+ tree indexes), a transaction must:

A. Use no locks at all when accessing data through the index structure
B. Lock children before parents when traversing the index tree structure
C. Lock a node before locking its children, and unlock parent before child
D. Lock all nodes simultaneously before accessing any data in the tree

Answer: C

Q1119. Lock escalation is the process of:

A. Converting coarse-grained locks into many more fine-grained locks
B. Converting many fine-grained locks into fewer coarse-grained locks
C. Adding locks continuously without ever releasing any of them
D. Removing all locks from the data items in the locked collection

Answer: B

Q1120. Snapshot isolation provides each transaction with:

A. Only read access to the data without any ability to write modifications
B. No isolation at all between concurrently executing database transactions
C. Access to the latest uncommitted data written by other active transactions
D. A consistent snapshot of the database as of the transaction's start time

Answer: D

Q1121. In ARIES, the analysis phase determines:

A. The set of dirty pages and active transactions at crash time
B. Only committed transactions that completed before the crash
C. The hardware failure cause and the affected disk components
D. Only the last checkpoint that was recorded in the log file

Answer: A

Q1122. In ARIES, the redo phase:

A. Replays all logged changes from the earliest dirty page LSN forward
B. Starts from the end of the log and processes records in reverse
C. Only replays committed transactions and ignores uncommitted ones
D. Undoes all transactions that were active at the time of the crash

Answer: A

Q1123. In ARIES, Compensation Log Records (CLRs) are:

A. Checkpoint records that mark points of database consistency
B. Log records for committed transactions only in the system
C. Log records written during undo to record the undo actions
D. Index maintenance records for B-tree structure updates

Answer: C

Q1124. The steal/no-force buffer management policy means:

A. Uncommitted pages can be written to disk (steal); committed need not be forced immediately (no-force)
B. Committed pages are always forced to disk immediately; uncommitted pages are never written to disk
C. Only committed pages are present on disk; uncommitted data remains exclusively in the memory buffer
D. All pages are always kept in memory; nothing is written to disk until the entire database shuts down

Answer: A

Q1125. No-steal/force policy would mean:

A. Uncommitted pages are never written to disk, and committed pages are always forced to disk
B. Neither undo nor redo is needed with this policy since both cases are eliminated
C. Uncommitted pages can be written to disk at any time before the transaction commits
D. Committed pages may not be on disk immediately and can remain only in memory buffer

Answer: A

Q1126. The redoLSN in ARIES is:

A. The checkpoint LSN that marks the most recent consistency checkpoint
B. The last LSN that was recorded in the transaction log before the crash
C. The LSN from which redo must start (earliest recLSN among dirty pages)
D. The LSN of the last commit record written to the stable storage log

Answer: C

Q1127. Write-ahead logging requires that log records be written to stable storage before:

A. The transaction starts its initial read operation
B. Any query is executed against the database tables
C. The corresponding data page is flushed to disk
D. The user submits the query to the database system

Answer: C

Q1128. In ARIES, when a transaction is undone, the undo follows:

A. The commit order of other transactions in the same schedule
B. The table creation order defined in the database schema file
C. The lastLSN chain backward through that transaction's log records
D. A random order without any specific sequence or priority rules

Answer: C

Q1129. Remote backup systems for disaster recovery provide:

A. A copy of the database at a geographically separate site
B. Better index structures for improving search performance
C. Improved normalization of tables to reduce data redundancy
D. Faster query processing through parallel execution plans

Answer: A

Q1130. The concept of idempotent redo means:

A. Applying redo multiple times produces the same result as applying it once
B. Redo can only be applied once and cannot be repeated after that point
C. Redo changes the result each time it is applied to the database state
D. Redo is never repeated because recovery always succeeds on first try

Answer: A

Q1131. In a B+ tree, the minimum occupancy of internal nodes (non-root) is:

A. All pointers must be full
B. One pointer (child) as minimum
C. Exactly n pointers at all times
D. Ceiling of n/2 pointers (children)

Answer: D

Q1132. The global depth in extendible hashing indicates:

A. The total number of records stored across all the buckets
B. The number of buckets currently allocated in the structure
C. The number of bits used to determine the directory entry
D. The number of hash functions used for the key computation

Answer: C

Q1133. Local depth in extendible hashing indicates:

A. The depth of the B+ tree from the root node to the leaf node level
B. The total number of collisions that have occurred in the hash table
C. The number of bits used to determine which bucket a record belongs to
D. The size of the directory that maps hash values to bucket addresses

Answer: C

Q1134. A hash join in query processing works by:

A. Partitioning both relations using the same hash function then joining matching partitions
B. Using an index on the join attribute to look up matching tuples from the inner relation
C. Sorting both relations on the join attribute first and then merging the sorted result sets
D. Using nested loops where each tuple of one relation is compared with every tuple of other

Answer: A

Q1135. R-trees are used to index:

A. Text data and documents only
B. Temporal data and timestamps
C. One-dimensional key values only
D. Spatial (multidimensional) data

Answer: D

Q1136. A function-based index is created on:

A. A user role that defines access permissions and grants
B. A database name used to identify the schema information
C. A table name stored in the database system catalog data
D. The result of a function applied to one or more columns

Answer: D

Q1137. An inverted index is commonly used in:

A. Foreign key enforcement
B. Numerical computations only
C. Primary key lookups only
D. Full-text search engines

Answer: D

Q1138. The order of a B+ tree affects performance because:

A. A higher order means more keys per node, reducing tree height and disk I/O
B. A higher order always increases the tree height and requires more disk access
C. Lower order is always better because smaller nodes fit in memory more easily
D. Order has no effect on performance and does not change disk I/O requirements

Answer: A

Q1139. Bulk loading a B+ tree is more efficient than individual insertions because:

A. It creates an unbalanced tree that must be rebalanced later
B. It sorts data first and builds the tree bottom-up, minimizing splits
C. It skips creating the tree entirely and uses a flat file instead
D. It uses no sorting and inserts records in their original order

Answer: B

Q1140. A skip list can be used as an alternative to B+ trees for indexing because:

A. It provides O(log n) average search, insert, and delete with simpler implementation
B. It uses less memory than any other data structure used for indexing purposes
C. It is always faster than B+ trees for every type of database query and operation
D. It requires no randomization and is fully deterministic in its performance

Answer: A

Q1141. The Selinger optimizer (System R) uses:

A. Only heuristic rules without any cost estimation being performed
B. Random plan selection from all possible execution plan choices
C. Dynamic programming with interesting orders to find optimal plans
D. No optimization at all with queries running as they are written

Answer: C

Q1142. An interesting sort order in query optimization is:

A. Any random sort order applied to tuples without considering the query operations
B. A sort order used only for displaying the final result to the end user query
C. A sort order that increases cost by requiring additional sorting in later steps
D. A sort order produced by one operation that is useful for a subsequent operation

Answer: D

Q1143. The number of possible join orders for n relations is:

A. Grows exponentially (Catalan number)
B. Always n squared for n relations
C. Always n for n given relations total
D. Exactly n! for n relations total

Answer: A

Q1144. In cost-based optimization, the estimated cost of a nested loop join R "È S is approximately:

A. Always 1 block access regardless of the sizes of relations R and S used
B. b_R plus b_S (sum of blocks from both R and S relations) in simple analysis
C. n_R plus n_S (sum of tuples from both R and S relations) in simple analysis
D. n_R times b_S plus b_R (b is blocks, n_R is tuples in R) in simple analysis

Answer: D

Q1145. Histogram statistics improve query optimization by:

A. Increasing query complexity by adding more steps to the optimization process
B. Providing more accurate selectivity estimates for non-uniform data distributions
C. Removing the need for indexes by providing direct access to all stored records
D. Making all estimates uniform regardless of the actual data value distribution

Answer: B

Q1146. A left-deep join tree is preferred by many optimizers because:

A. It is the only valid tree shape for all joins
B. It avoids all disk I/O during the join process
C. It minimizes the number of tables in joins
D. It allows pipelining of intermediate results

Answer: D

Q1147. Adaptive query processing differs from traditional optimization by:

A. Only optimizing before execution begins and then running without changes
B. Adjusting the execution plan during query execution based on runtime statistics
C. Never changing the plan once it has been selected by the query optimizer
D. Ignoring all statistics and using a fixed plan for every query submitted

Answer: B

Q1148. The concept of equivalent query expressions means:

A. Identical query text that is written exactly the same way in SQL
B. Different relational algebra expressions that produce the same result
C. Queries that produce different results from the same input relations
D. Queries executed on different databases with different table schemas

Answer: B

Q1149. Predicate pushdown in distributed queries:

A. Ignores remote data sources completely and only queries local tables
B. Sends all data to a central location first before any filtering occurs
C. Sends filter conditions to remote data sources to reduce data transfer
D. Removes all predicates from the query and returns all unfiltered data

Answer: C

Q1150. Columnar storage benefits analytical queries because:

A. No indexes are needed for any query on columnar data store
B. Row storage is used instead of column-oriented data format
C. All columns are always read regardless of the query needs
D. Only the needed columns are read from disk, reducing I/O

Answer: D

Q1151. Statistical database security addresses the problem of:

A. Managing user roles and their assigned privileges set
B. Creating backup copies of the database for recovery
C. Inferring individual data from aggregate query results
D. Encrypting data at rest on the physical storage media

Answer: C

Q1152. The inference problem in database security occurs when:

A. No queries are allowed to be executed against the tables
B. Users deduce restricted information from permitted queries
C. Users have direct database access without any restrictions
D. All data is encrypted and cannot be read without the key

Answer: B

Q1153. Label-based security in databases assigns:

A. Sensitivity labels to data rows and clearance levels to users
B. Labels only to tables without any row-level classification
C. No labels to any data regardless of its sensitivity or nature
D. Random access permissions to all users without any restrictions

Answer: A

Q1154. Homomorphic encryption allows:

A. Faster decryption of data using optimized algorithms
B. Only reading encrypted data without any modifications
C. Computations on encrypted data without decrypting it
D. Only writing encrypted data without reading it first

Answer: C

Q1155. Database firewall works by:

A. Encrypting all data stored in the database tables and log files
B. Only blocking network traffic without inspecting SQL statements
C. Creating physical firewalls around the database server hardware
D. Monitoring SQL traffic and blocking unauthorized or suspicious queries

Answer: D

Q1156. The k-anonymity privacy model ensures:

A. Only k users can access the database at any given time during concurrent operations
B. Data is stored in k copies across different sites for redundancy and availability
C. All data is encrypted using a key of length k bits for secure storage and retrieval
D. Each record is indistinguishable from at least k-1 other records on quasi-identifiers

Answer: D

Q1157. Differential privacy provides:

A. Full access to all data without any restrictions or privacy controls
B. Only physical security measures without any mathematical foundations
C. No privacy guarantees of any kind for the data stored in the database
D. Mathematical guarantees that query results do not reveal individual data

Answer: D

Q1158. A privilege graph is used to:

A. Track the flow of granted privileges among users
B. Design ER diagrams for conceptual data modeling
C. Optimize query execution plans for better speed
D. Create indexes on columns for fast data retrieval

Answer: A

Q1159. Cascading revocation in SQL means:

A. Revoking a privilege also revokes it from all users who received it through the original grant
B. No revocation occurs and the privilege remains active for all users who currently have it
C. Only the specified user loses the privilege and no other users are affected by the revocation
D. All database privileges for every user are revoked regardless of the specific grant involved

Answer: A

Q1160. Column-level encryption differs from TDE in that:

A. It encrypts specific columns rather than entire database files
B. It encrypts only at the network level rather than at storage level
C. It encrypts only table names rather than the actual column data
D. It is less secure than TDE because it covers fewer data elements

Answer: A

Q1161. The three-phase commit (3PC) protocol improves on 2PC by:

A. Using only one phase for the commit decision
B. Removing the prepare phase from the protocol
C. Adding a pre-commit phase to avoid blocking
D. Ignoring all failures that occur at any site

Answer: C

Q1162. In distributed databases, the global query optimization must consider:

A. Only CPU time without considering disk or network access costs
B. Only disk I/O costs without considering network or CPU overhead
C. Only local processing costs without considering network overhead
D. Network communication costs in addition to local processing costs

Answer: D

Q1163. Paxos is a protocol used for:

A. Achieving consensus in distributed systems
B. Indexing data for faster retrieval queries
C. Query optimization in database management
D. Schema design for database normalization

Answer: A

Q1164. The RAFT consensus algorithm is designed to be:

A. Faster than all other protocols in every distributed environment
B. A replacement for SQL in distributed relational database systems
C. More understandable than Paxos while providing the same guarantees
D. A type of index for organizing data in distributed hash tables

Answer: C

Q1165. In a federated database system:

A. No integration is possible between any databases at all
B. Autonomous databases cooperate to provide integrated access
C. Only one database exists in the entire distributed system
D. All databases are identical copies of the same data schema

Answer: B

Q1166. The eventual consistency model guarantees that:

A. No consistency is provided between any of the replicas in the system
B. Immediate consistency at all times for every read at every site node
C. All reads always return the latest written value at every site instantly
D. All replicas will converge to the same value if no new updates are made

Answer: D

Q1167. The quorum-based protocol requires:

A. Qr = Qw = 1 for consistency meaning only one replica needs to respond
B. Read quorum (Qr) + Write quorum (Qw) > total replicas (N) for consistency
C. No quorum is needed and any single replica can respond to operations
D. Qr + Qw less than N for consistency between read and write quorums

Answer: B

Q1168. Vector clocks in distributed systems are used to:

A. Capture causal relationships between events at different sites
B. Measure CPU speed and processing capacity at each network node
C. Optimize queries by selecting the best execution plan for joins
D. Tell wall-clock time accurately across all distributed node sites

Answer: A

Q1169. Distributed hash tables (DHTs) are used in:

A. Peer-to-peer distributed storage systems
B. Centralized databases on a single machine
C. Single-machine databases without networks
D. Only for indexing within one database node

Answer: A

Q1170. The coordinator selection problem in distributed databases can be solved by:

A. Manual intervention always requiring a human administrator
B. No algorithm exists for solving the coordinator selection
C. Random selection only without any formal algorithmic approach
D. Election algorithms like the Bully algorithm or Ring algorithm

Answer: D

Q1171. The concept of tunable consistency in Cassandra means:

A. Consistency cannot be changed once the Cassandra cluster has been configured
B. Only strong consistency is available for all queries in the entire cluster
C. The consistency level can be configured per query (ONE, QUORUM, ALL, etc.)
D. Only eventual consistency is possible and no other level can be selected

Answer: C

Q1172. A wide-column store differs from a traditional columnar database in that:

A. It is identical to a relational database model
B. All rows must have the same columns defined
C. Each row can have a different set of columns
D. It uses only fixed schemas without flexibility

Answer: C

Q1173. CRDTs (Conflict-free Replicated Data Types) are used in distributed NoSQL to:

A. Automatically resolve conflicts in replicated data without coordination
B. Compress data to reduce the network transfer size between replica sites
C. Encrypt data before replicating it across the distributed cluster nodes
D. Create indexes on replicated data for faster query retrieval operations

Answer: A

Q1174. The gossip protocol in distributed NoSQL systems is used for:

A. Transaction management for multi-site commit control
B. Index creation on distributed data across all nodes
C. Query optimization for distributed join operations
D. Spreading information about cluster state across nodes

Answer: D

Q1175. NewSQL databases aim to provide:

A. Only NoSQL features without any ACID guarantee support
B. No guarantees at all for either consistency or scaling
C. SQL-like guarantees (ACID) with NoSQL-like scalability
D. Only traditional SQL features without scalability gains

Answer: C

Q1176. Examples of NewSQL databases include:

A. Google Spanner, CockroachDB, VoltDB
B. All of the above are NewSQL types
C. MongoDB, Redis, Apache Cassandra
D. MySQL, PostgreSQL, Oracle Database

Answer: A

Q1177. The LSM (Log-Structured Merge) tree used in many NoSQL databases provides:

A. Only sequential reads without support for random access query patterns
B. High read throughput only without any optimization for write operations
C. High write throughput by buffering writes in memory and merging to disk
D. No performance benefits over traditional B-tree index structures at all

Answer: C

Q1178. Google's Bigtable is:

A. A traditional relational database with SQL support
B. A key-value store only without column family data
C. A distributed, sparse, multi-dimensional sorted map
D. A graph database using nodes and edge connections

Answer: C

Q1179. Amazon DynamoDB uses:

A. B+ trees for all data storage and access
B. Consistent hashing for data partitioning
C. Only single-node storage without sharding
D. Relational tables internally for all data

Answer: B

Q1180. Multi-model databases support:

A. Only relational data organized in standard normalized table format on disk
B. Only flat files stored sequentially without any structured data model in use
C. Multiple data models (document, graph, key-value) in a single database engine
D. Only one data model such as relational without any support for other models

Answer: C

Q1181. A slowly changing dimension (SCD) Type 2 handles changes by:

A. Overwriting the old value directly in place without keeping any history of changes
B. Adding a new row with the new value and preserving the old row with history tracking
C. Deleting the old row entirely and inserting a brand new row with the changed value
D. Adding a new column to store the updated value alongside the previous column value

Answer: B

Q1182. A degenerate dimension is:

A. A dimension key in the fact table that has no corresponding dimension table
B. A dimension that is always empty and contains no rows of data whatsoever
C. A dimension table with many attributes and a large number of data records
D. A normalized dimension table that has been split into multiple sub-tables

Answer: A

Q1183. The galaxy schema (fact constellation) contains:

A. No dimension tables in the schema at all
B. Only dimension tables without fact tables
C. Only one fact table with no shared dimensions
D. Multiple fact tables sharing dimension tables

Answer: D

Q1184. Clustering in data mining groups:

A. Data by primary key values in ascending sort order
B. Data based on known categories and predefined classes
C. Similar data points together without predefined labels
D. Data randomly without any similarity considerations

Answer: C

Q1185. Decision tree classification in data mining:

A. Builds a tree structure to classify data based on attribute values
B. Creates random groupings of data without any logical partitioning
C. Performs regression only for predicting continuous numeric values
D. Finds association rules between items in transactional data sets

Answer: A

Q1186. The K-means algorithm is a:

A. Classification algorithm using labeled training data sets
B. Association rule algorithm for finding item relationships
C. Regression algorithm for predicting continuous value data
D. Clustering algorithm that partitions data into K groups

Answer: D

Q1187. A materialized view in a data warehouse is refreshed:

A. Only when the database restarts after a scheduled shutdown
B. Periodically or on demand to reflect source data changes
C. In real-time always with zero latency from source updates
D. Never after creation regardless of any source data changes

Answer: B

Q1188. Data lake differs from a data warehouse in that it:

A. Is always smaller in size than a data warehouse
B. Requires data transformation before storage occurs
C. Stores raw, unstructured data in its native format
D. Only stores structured data in normalized table form

Answer: C

Q1189. The concept of data lineage in data warehousing refers to:

A. The size of tables measured in rows and storage bytes
B. Tracking the origin and transformation history of data
C. The number of queries executed against the data store
D. The age of the database since its initial creation date

Answer: B

Q1190. Real-time data warehousing aims to:

A. Minimize the latency between source data changes and warehouse availability
B. Store data offline in cold storage without any access for active queries
C. Avoid using ETL processes entirely and load data without transformation
D. Only process batch data on a periodic schedule without real-time updates

Answer: A

Q1191. In a scenario where multiple applications need different views of employee data, which architecture feature best addresses this?

A. External schema views tailored to each application need
B. Conceptual schema duplication across each separate application
C. Internal schema customization for each application program
D. Physical storage partitioning for each individual application

Answer: A

Q1192. A university database must track students, courses, and enrollments. Which data independence level is impacted when changing the index structure on the student table?

A. Physical data independence is impacted directly
B. Neither level of independence is affected here
C. Logical data independence is impacted directly
D. Both logical and physical independence equally

Answer: A

Q1193. When a DBMS enforces referential integrity, what specific problem is being prevented?

A. Excessive storage consumption from duplicate entries
B. Unauthorized users gaining access to sensitive records
C. Slow query performance due to missing index structures
D. Dangling references to non-existent related records

Answer: D

Q1194. An enterprise migrates from file-based to DBMS. Which challenge is most critical during this transition?

A. Selecting a new operating system for all server machines
B. Reducing the total number of end user workstations
C. Converting existing data and ensuring schema consistency
D. Upgrading monitor displays for the database administrators

Answer: C

Q1195. A company needs to change the conceptual schema by adding a new attribute to a table without affecting existing applications. This exemplifies:

A. View-level independence from external schema changes
B. Instance-level independence from data value updates
C. Logical data independence from schema modifications
D. Physical data independence from storage changes

Answer: C

Q1196. In the ANSI/SPARC architecture, which component is responsible for transforming requests between schema levels?

A. The transaction coordinator
B. The schema mapping engine
C. The query optimizer module
D. The concurrency controller

Answer: B

Q1197. A hospital database requires that patient records cannot be deleted if they have associated billing entries. Which DBMS feature enforces this?

A. Triggers that log all deletion attempts made
B. Check constraints on the billing amount field
C. Referential integrity with restrict delete
D. Indexing on the patient identifier column

Answer: C

Q1198. When evaluating DBMS overhead versus file systems for a small single-user application, which factor most favors using a file system?

A. The application must enforce complex integrity constraints
B. The application needs minimal setup and lower overhead
C. The application demands sophisticated query optimization
D. The application requires complex multi-user concurrency

Answer: B

Q1199. A database system uses deferred modification during transaction execution. What does this mean regarding write operations?

A. Writes are duplicated to two separate disk locations each
B. Writes go directly to disk before the transaction commits
C. Writes are encrypted before being stored on the disk drive
D. Writes are held in buffer until transaction commits first

Answer: D

Q1200. An organization has separate databases for HR, Finance, and Inventory with overlapping employee data. Which integration approach best reduces redundancy?

A. Creating a centralized database with shared employee records
B. Allowing each department to independently define employee data
C. Using flat files to export and import data between departments
D. Maintaining separate databases with manual data synchronization

Answer: A

Q1201. A company experiences slow queries during peak hours with a two-tier architecture. Which architectural change would best address this?

A. Increasing the network bandwidth between client and server
B. Migrating to a three-tier architecture with application server
C. Replacing the client hardware with more powerful machines
D. Adding more storage disks to the existing database server

Answer: B

Q1202. In a shared-nothing parallel database, how is data redistribution handled during query joins across partitions?

A. All data is consolidated to a single node before joining
B. Joins across partitions are not supported in this setup
C. Relevant data is shipped between nodes over the network
D. Data is copied to a shared memory area for join processing

Answer: C

Q1203. A web application handles thousands of concurrent users. Which DBMS architectural feature is most critical for maintaining response times?

A. Connection pooling and efficient buffer management
B. Manual memory allocation by the database operator
C. Verbose logging of every individual query executed
D. Schema versioning for backward compatibility support

Answer: A

Q1204. When designing a fault-tolerant database architecture, which strategy best ensures continuous availability?

A. Disabling transaction logging to improve performance speed
B. Using active-passive replication with automatic failover
C. Storing all data on a single high-performance server
D. Relying solely on periodic backup tapes for recovery

Answer: B

Q1205. In a multi-tier architecture, which tier should implement data validation business rules for optimal security?

A. In the application logic tier on the middle server
B. Distributed equally across all three architectural tiers
C. Exclusively in the presentation tier on the client side
D. Only in the data tier within stored database procedures

Answer: A

Q1206. A distributed database system uses a two-phase commit protocol. What happens if the coordinator fails after sending prepare messages?

A. The transaction is automatically rolled back at all sites
B. Each participant independently decides to commit or abort
C. All participants automatically commit their transactions
D. Participants remain blocked until the coordinator recovers

Answer: D

Q1207. An organization needs to process analytical queries without affecting transactional performance. Which architectural pattern best achieves this?

A. Using a single-tier architecture with query prioritization
B. Implementing separate OLTP and OLAP systems with data feeds
C. Disabling indexing on the transactional database for speed
D. Running both workloads on the same database server instance

Answer: B

Q1208. In a cloud-based database architecture, what does the concept of elastic scaling specifically refer to?

A. Permanently allocating maximum resources for peak loads
B. Reducing database features to lower computational overhead
C. Manually adding hardware when usage increases over time
D. Automatically adjusting resources based on current demand

Answer: D

Q1209. When implementing a microservices architecture, what database pattern avoids tight coupling between services?

A. Services communicate by directly reading each others data
B. Each microservice owns its own private database instance
C. A single transaction spans multiple microservice databases
D. All microservices share a single centralized database

Answer: B

Q1210. A DBMS uses write-ahead logging in its architecture. What constraint does this impose on buffer management?

A. Buffer management operates independently of log records
B. Data pages must be written before any log records at all
C. Buffer pages can be flushed to disk at any time freely
D. Log records must be written to disk before data pages

Answer: D

Q1211. A university needs to model a scenario where a student can major in multiple departments, and each department offers multiple majors. Which relationship type best represents this?

A. Many-to-many between student and department entities
B. Unary relationship on the student entity type alone
C. One-to-many from department entity to student entity
D. One-to-one between student and department entities

Answer: A

Q1212. When converting an ER diagram with a ternary relationship to a relational schema, how many foreign keys does the resulting relation typically contain?

A. No foreign keys since ternary cannot be directly mapped
B. Two foreign keys referencing the two main entity types
C. Three foreign keys referencing each participating entity
D. One foreign key referencing the primary entity only

Answer: C

Q1213. In the Enhanced ER model, an overlapping specialization constraint means:

A. Subtypes must collectively include all supertype entities
B. An entity can belong to at most one subtype at a time
C. No entity can belong to any of the defined subtype sets
D. An entity can belong to multiple subtypes simultaneously

Answer: D

Q1214. A hospital system models doctors, patients, and treatments. A doctor treats a patient with a specific treatment. Which ER construct best models this scenario?

A. An aggregation of doctor-patient with treatment entity
B. Binary relationship between doctor and patient only
C. Two separate binary relationships for each pair type
D. A ternary relationship among all three entity types

Answer: D

Q1215. When modeling a recursive relationship where an employee manages other employees, what special consideration is needed?

A. The relationship requires two separate entity types always
B. A weak entity must be created for managed employees here
C. Role names must distinguish the two participations clearly
D. Recursive relationships cannot have any cardinality ratios

Answer: C

Q1216. In an EER diagram, the category (union type) construct is used when:

A. An entity belongs to exactly one of multiple supertype sets
B. An entity is always independent of all other entity types
C. An entity has no relationship with any supertype entity set
D. An entity must belong to all listed supertype entity sets

Answer: A

Q1217. A library system has Books, Authors, and Publishers. A book can have multiple authors, and each author may work with multiple publishers. If we need to track which publisher published which author-specific book, the best modeling approach is:

A. Two binary relationships with an aggregation construct used
B. Three binary relationships connecting each pair of entities
C. A single ternary relationship among Book, Author, Publisher
D. A single binary relationship between Book and Author only

Answer: C

Q1218. When mapping a total specialization with disjoint constraint to a relational schema, which approach uses the fewest tables?

A. Create a single table with a type attribute discriminator
B. Create tables for only the subtypes with inherited attributes
C. Create tables for supertype and all subtypes separately
D. Create one table per attribute of the supertype entity type

Answer: B

Q1219. In object-oriented data models, what does the concept of encapsulation specifically provide?

A. Hiding internal state and exposing only defined methods
B. Direct access to all internal data attributes of objects
C. Conversion of objects into flat relational table records
D. Automatic replication of objects across multiple servers

Answer: A

Q1220. A social network needs to model users who can be friends with each other and also belong to groups. Which combination of ER constructs most accurately captures this?

A. A ternary relationship among user, friendship, and group
B. Only binary relationships between user and group entities
C. A recursive relationship on user plus binary with groups
D. An aggregation of user entity with a group attribute added

Answer: C

Q1221. Given a relation R(A, B, C, D) with functional dependencies A->B, B->C, and C->D, what is the candidate key?

A. The attribute set consisting of C only
B. The attribute set consisting of D only
C. The attribute set consisting of A only
D. The attribute set consisting of B only

Answer: C

Q1222. In a relation with attributes {A, B, C, D, E} and FDs {AB->C, C->D, D->E}, what is the closure of {A, B}?

A. The closure is the set {A, B, D, E} only
B. The closure is the set {A, B, C} only
C. The closure is the set {C, D, E} without A B
D. The closure is the set {A, B, C, D, E}

Answer: D

Q1223. A table has a composite primary key (StudentID, CourseID). Which statement about this design is correct?

A. The combination must be unique but individual parts may repeat
B. Each attribute alone must uniquely identify every record
C. Only StudentID needs to be unique across all table records
D. Neither StudentID nor CourseID can be null independently

Answer: A

Q1224. If a relation has five candidate keys each consisting of two attributes from a set of six, how many superkeys exist at minimum?

A. Only one superkey exists in this relation total
B. Exactly five superkeys exist in this relation
C. At least five but potentially many more superkeys
D. Exactly ten superkeys exist in this relation

Answer: C

Q1225. A database designer must choose between a natural key (SSN) and a surrogate key (auto-increment ID) for a Person table. Which consideration most favors the surrogate key?

A. Surrogate keys carry inherent meaning about the entity
B. Surrogate keys are stable and unaffected by real-world changes
C. Natural keys provide better query performance in all cases
D. Natural keys are always shorter than surrogate key values

Answer: B

Q1226. In a relation R(A, B, C) with the functional dependency AB->C and no other non-trivial FDs, which normal form is R in?

A. Second normal form but not third due to transitivity
B. Third normal form since no transitive dependency exists
C. Boyce-Codd normal form since AB is a superkey of R
D. First normal form only since it has partial dependency

Answer: C

Q1227. When designing a relational schema for a many-to-many relationship between Students and Courses with an enrollment date attribute, which approach is correct?

A. Add enrollment date to the Students table as a new column
B. Store enrollment date in a separate unrelated lookup table
C. Create a junction table with enrollment date as attribute
D. Add enrollment date to the Courses table as a new column

Answer: C

Q1228. If Armstrong axioms are applied to derive A->C from A->B and B->C, which specific axiom is used?

A. Transitivity axiom of functional dependencies
B. Decomposition axiom of functional dependencies
C. Reflexivity axiom of functional dependencies
D. Augmentation axiom of functional dependencies

Answer: A

Q1229. A relation has attributes {W, X, Y, Z} with FDs {W->X, X->Y, WY->Z}. What are all the candidate keys?

A. The candidate keys are both {W} and {X} individually
B. The only candidate key is the single attribute {W}
C. The only candidate key is the attribute set {W, Y}
D. The only candidate key is the attribute set {W, X}

Answer: B

Q1230. In a relational database, what specific problem does the phantom read anomaly cause in concurrent transactions?

A. A transaction reads data that was already committed earlier
B. A transaction sees new rows that match its previous query range
C. A transaction cannot read any rows due to an exclusive lock
D. A transaction overwrites changes made by another transaction

Answer: B

Q1231. Given relations R(A,B) and S(B,C), which expression finds A values associated with all C values in S?

A. Divide the projection of AB from R by projection of B from S
B. Divide the projection of AC from R join S by projection of C from S
C. Divide the natural join of R and S by relation R on A only
D. Project A from the natural join of R and S directly

Answer: B

Q1232. When optimizing a relational algebra expression, which transformation typically yields the greatest performance improvement?

A. Replacing all natural joins with Cartesian products instead
B. Pushing selections down to be applied as early as possible
C. Performing projections after all joins are completed fully
D. Delaying all selections until the final result is computed

Answer: B

Q1233. In relational algebra, how can the intersection operation be expressed using fundamental operations?

A. A minus the result of A minus B gives the intersection
B. A union B minus the Cartesian product of A and B together
C. A union B minus A minus B gives the intersection result
D. The Cartesian product of A and B filtered by equality check

Answer: A

Q1234. A query requires finding employees who work in every department. Which relational algebra operation is most appropriate?

A. Division of employee-department pairs by all departments
B. Natural join of employee and department relations directly
C. Projection of employee names from the department relation
D. Selection on the Cartesian product with a count condition

Answer: A

Q1235. When converting SELECT DISTINCT A FROM R WHERE B > 5 to relational algebra, the correct expression is:

A. R join sigma_{B>5}(R) using natural join with selection first
B. pi_A(sigma_{B>5}(R)) applying selection before projection
C. sigma_{B>5}(pi_A(R)) applying projection before selection
D. pi_A(R) / sigma_{B>5}(R) using division after projection

Answer: B

Q1236. In query tree optimization, why is it beneficial to perform projections early in the expression tree?

A. Early projections increase the number of attributes processed
B. Early projections automatically sort the intermediate results
C. Early projections eliminate the need for any join operations
D. Early projections reduce tuple width reducing memory overhead

Answer: D

Q1237. Given R(A,B,C) and S(C,D,E), which expression correctly computes R left outer join S on R.C = S.C?

A. (R join S) union ((R - pi_{A,B,C}(R join S)) times {(null,null)})
B. (R join S) union (pi_{A,B,C}(R join S) - R) with null padding
C. R times S filtered by R.C equals S.C with duplicates removed
D. pi_{A,B}(R) join sigma_{D>0}(S) with null values for unmatched

Answer: A

Q1238. When two selection conditions are conjunctive (AND), which algebraic equivalence holds?

A. sigma_{p AND q}(R) = sigma_p(sigma_q(R)) using cascaded selections
B. sigma_{p AND q}(R) = sigma_p(R) times sigma_q(R) using product
C. sigma_{p AND q}(R) = sigma_p(R) union sigma_q(R) using union
D. sigma_{p AND q}(R) = sigma_p(R) - sigma_q(R) using set difference

Answer: A

Q1239. In the context of query optimization, what is the cost of a Cartesian product of relations with m and n tuples respectively?

A. The cost is proportional to the sum of m plus n tuples
B. The cost is proportional to the minimum of m or n tuples
C. The cost is proportional to the maximum of m or n tuples
D. The cost is proportional to the product of m times n tuples

Answer: D

Q1240. A relational algebra expression uses multiple joins. Which join ordering heuristic typically minimizes intermediate result sizes?

A. Always join the two smallest intermediate results first
B. Always join the two largest relations first in the query
C. Join relations in the order they appear in the SQL query
D. Join relations in alphabetical order of their table names

Answer: A

Q1241. A query needs to find the second highest salary from an Employee table. Which approach is correct?

A. SELECT MIN(salary) FROM Employee WHERE salary > (SELECT MIN(salary) FROM Employee)
B. SELECT DISTINCT salary FROM Employee ORDER BY salary LIMIT 1
C. SELECT MAX(salary) FROM Employee WHERE salary < (SELECT MAX(salary) FROM Employee)
D. SELECT MAX(salary) FROM Employee without any additional filter

Answer: C

Q1242. What is the difference between UNION and UNION ALL in SQL?

A. UNION works with different columns while UNION ALL needs same
B. UNION removes duplicates while UNION ALL keeps all duplicates
C. UNION keeps duplicates while UNION ALL removes all duplicates
D. UNION is for joins while UNION ALL is for subquery operations

Answer: B

Q1243. In a correlated subquery, when is the inner query executed?

A. Once for each row processed by the outer query statement
B. Only when the outer query returns an empty result set
C. Once before the outer query begins its execution cycle
D. Once after the outer query completes all its processing

Answer: A

Q1244. Which SQL construct is best for assigning a rank with gaps to rows based on a column value?

A. NTILE window function
B. RANK window function
C. ROW_NUMBER window function
D. DENSE_RANK window function

Answer: B

Q1245. A query uses multiple joins. In which order does the SQL engine typically process the clauses?

A. SELECT, FROM, WHERE, GROUP BY, HAVING, ORDER BY sequence
B. ORDER BY, SELECT, FROM, WHERE, GROUP BY, HAVING sequence
C. WHERE, FROM, SELECT, GROUP BY, ORDER BY, HAVING sequence
D. FROM, WHERE, GROUP BY, HAVING, SELECT, ORDER BY sequence

Answer: D

Q1246. What is the purpose of the COALESCE function in SQL?

A. It returns the first non-null value from its arguments
B. It concatenates multiple string values into one result
C. It converts a value from one data type to another type
D. It calculates the average of all provided numeric values

Answer: A

Q1247. When using a Common Table Expression (CTE), what advantage does it provide over a regular subquery?

A. CTEs allow recursive queries that subqueries cannot handle
B. CTEs bypass all security permissions on underlying base tables
C. CTEs are always faster than equivalent subquery expressions
D. CTEs automatically create permanent tables in the database

Answer: A

Q1248. A query needs to find departments where ALL employees earn more than 50000. Which approach is correct?

A. SELECT dept FROM Employee GROUP BY dept HAVING MAX(salary) > 50000
B. SELECT dept FROM Employee GROUP BY dept HAVING MIN(salary) > 50000
C. SELECT dept FROM Employee GROUP BY dept HAVING AVG(salary) > 50000
D. SELECT dept FROM Employee WHERE salary > 50000 without grouping

Answer: B

Q1249. What is the difference between EXISTS and IN when used with subqueries?

A. EXISTS only works with correlated subqueries and IN never does
B. EXISTS and IN always produce different results for same data
C. EXISTS checks for value matches while IN checks row existence
D. EXISTS returns true if subquery returns any rows at all

Answer: D

Q1250. In SQL, what does the PIVOT operation accomplish?

A. It transforms row values into separate column headers
B. It merges multiple tables into a single combined view
C. It sorts data by multiple columns in alternating order
D. It splits a single column into multiple separate tables

Answer: A

Q1251. A recursive CTE is used to traverse an employee hierarchy. What component prevents infinite recursion?

A. A termination condition in the recursive member WHERE clause
B. The LIMIT clause applied to the final outer SELECT query
C. An index on the self-referencing foreign key column used
D. The DISTINCT keyword in the recursive member query

Answer: A

Q1252. When should a materialized view be refreshed using ON COMMIT versus ON DEMAND?

A. ON COMMIT for infrequently changed data and ON DEMAND for frequent
B. ON COMMIT for read-heavy workloads and ON DEMAND for write-heavy
C. ON COMMIT when data freshness is critical and ON DEMAND when not
D. ON COMMIT for large tables and ON DEMAND for small table sizes

Answer: C

Q1253. A trigger needs to maintain an audit trail of all salary changes. Which trigger type and timing combination is most appropriate?

A. BEFORE SELECT trigger on the employee table records
B. INSTEAD OF DELETE trigger on the employee table only
C. AFTER UPDATE trigger on the employee salary column
D. BEFORE INSERT trigger on the employee table only

Answer: C

Q1254. In SQL, what is the difference between RANK(), DENSE_RANK(), and ROW_NUMBER() for tied values?

A. All three produce identical results for tied values always
B. ROW_NUMBER skips ranks while RANK and DENSE_RANK assign same to ties
C. RANK and DENSE_RANK handle ties while ROW_NUMBER assigns unique numbers
D. RANK skips numbers for ties while DENSE_RANK and ROW_NUMBER do not

Answer: C

Q1255. A stored procedure needs to handle errors gracefully. Which SQL mechanism is most appropriate?

A. Implementing TRY-CATCH or EXCEPTION handling blocks within
B. Using nested IF-ELSE statements for every possible error
C. Wrapping every statement in a separate transaction context
D. Checking return values of every SQL statement manually each

Answer: A

Q1256. When creating a covering index, what specific advantage does it provide for query performance?

A. It compresses the data to reduce overall storage requirements
B. It allows the query to be answered from the index alone
C. It eliminates the need for primary keys on the covered table
D. It automatically partitions the table across multiple disks

Answer: B

Q1257. A database has a view defined on a join of three tables. Under which condition is this view updatable?

A. When all three tables have identical primary key definitions
B. When the update affects columns from only one base table
C. When the view includes aggregate functions on joined columns
D. When the view includes a DISTINCT clause in its definition

Answer: B

Q1258. What is the difference between statement-level and row-level triggers?

A. Statement-level only works with SELECT while row-level with all DML
B. Statement-level is faster than row-level in every possible case
C. Statement-level fires for each row while row-level fires once
D. Statement-level fires once per statement while row-level per row

Answer: D

Q1259. When using the MERGE statement, what advantage does it provide over separate INSERT, UPDATE, DELETE statements?

A. MERGE automatically creates indexes on the target table columns
B. MERGE combines insert, update, and delete in a single atomic step
C. MERGE is always faster than individual DML statements executed
D. MERGE bypasses all constraint checking for improved performance

Answer: B

Q1260. A query uses LAG and LEAD window functions. What do these functions specifically provide?

A. They access data from preceding and following rows respectively
B. They count the number of rows before and after current row
C. They calculate running totals across all rows in partition
D. They determine the first and last values within each partition

Answer: A

Q1261. Given R(A, B, C, D) with FDs {A->B, B->C, C->D}, decompose into BCNF relations.

A. R1(A,D), R2(B,C) splitting based on attribute positions
B. R1(A,B,C,D) remains as a single table in BCNF form
C. R1(A,C), R2(B,D) with cross-referencing foreign key links
D. R1(A,B), R2(B,C), R3(C,D) each with proper determinants

Answer: D

Q1262. A relation R(Student, Course, Instructor) has FDs {Student,Course->Instructor, Instructor->Course}. What normal form is R in?

A. R is only in 2NF because of a partial dependency present
B. R is only in 1NF because of a transitive dependency issue
C. R is in 3NF but not BCNF due to the second dependency
D. R is in BCNF since all determinants are candidate keys

Answer: C

Q1263. When decomposing a relation into 3NF using the synthesis algorithm, what is the first step?

A. Identify all candidate keys of the original relation first
B. List all attributes and their domains in the relation schema
C. Compute a minimal cover of the functional dependencies set
D. Create a separate table for each attribute in the relation

Answer: C

Q1264. A relation has attributes {A, B, C, D, E} with FDs {AB->C, C->D, D->E, E->A}. How many candidate keys does this relation have?

A. Exactly two candidate keys exist in this relation
B. Exactly one candidate key exists in this relation
C. Exactly three candidate keys exist in this relation
D. Exactly four candidate keys exist in this relation

Answer: D

Q1265. What specific problem can arise when decomposing a relation into BCNF that does not occur with 3NF decomposition?

A. BCNF decomposition may lose some functional dependencies
B. BCNF decomposition cannot handle composite primary key sets
C. BCNF decomposition always results in lossy join operations
D. BCNF decomposition always creates more tables than needed

Answer: A

Q1266. A table stores Employee(EmpID, Name, DeptID, DeptName, DeptLocation). With FD DeptID->DeptName,DeptLocation, what anomalies exist?

A. No anomalies exist because the table is fully normalized
B. Update, insertion, and deletion anomalies all exist here
C. Only deletion anomalies exist due to DeptID dependency
D. Only insertion anomalies exist due to the composite key

Answer: B

Q1267. Fifth normal form (5NF) addresses which specific type of dependency?

A. Join dependencies that cannot be expressed as simpler ones
B. Inclusion dependencies between attributes of related tables
C. Functional dependencies between key and non-key attributes
D. Multivalued dependencies between independent attribute sets

Answer: A

Q1268. In the chase algorithm for testing lossless join, what does the algorithm verify?

A. Whether the original relation has any candidate key defined
B. Whether the resulting tables are all in BCNF normal form
C. Whether the decomposition allows lossless reconstruction
D. Whether all functional dependencies are preserved in split

Answer: C

Q1269. A relation R(A,B,C) has multivalued dependencies A->>B and A->>C. What is the correct 4NF decomposition?

A. Keep R(A,B,C) as a single table since it is already in 4NF
B. Decompose into R1(A,B) and R2(A,C) removing the dependency
C. Decompose into R1(B,C) and R2(A) with a mapping table added
D. Decompose into R1(A,B,C) and R2(A) with a foreign key link

Answer: B

Q1270. When is a relation in Domain-Key Normal Form (DKNF)?

A. When all functional dependencies have superkeys as determinants
B. When all constraints are consequences of domain and key constraints
C. When no multivalued dependencies exist in the relation schema
D. When all attributes are of the same domain type and key type

Answer: B

Q1271. A banking system processes concurrent transfers between accounts. If T1 transfers from A to B and T2 transfers from B to A simultaneously, what specific ACID property prevents incorrect total balances?

A. Atomicity ensures each transfer fully completes or not at all
B. All three properties above work together to prevent the issue
C. Isolation ensures the transfers do not interfere with each
D. Durability ensures the balances survive server failures

Answer: C

Q1272. In a schedule where T1 reads X, T2 writes X, T1 reads X again, which anomaly occurs and at which isolation level is it first prevented?

A. Non-repeatable read prevented at Repeatable Read level first
B. Dirty read anomaly prevented at Read Uncommitted level first
C. Phantom read anomaly prevented at Serializable level first
D. No anomaly occurs because both transactions read the data

Answer: A

Q1273. When testing conflict serializability, what does a cycle in the precedence graph indicate?

A. The schedule needs more transactions to determine its result
B. The schedule is not conflict serializable and has conflicts
C. The schedule is conflict serializable and safe to execute
D. The schedule has no conflicts between any transactions at all

Answer: B

Q1274. A long-running analytical transaction keeps getting aborted due to conflicts with short OLTP transactions. Which strategy best addresses this?

A. Using snapshot isolation for the analytical transaction query
B. Increasing the timeout period for the analytical transaction
C. Converting the analytical query into multiple short transactions
D. Disabling concurrency control for the analytical transaction

Answer: A

Q1275. In a distributed transaction using the two-phase commit protocol, what happens during the prepare phase?

A. All participants send their data to the coordinator node
B. The coordinator unilaterally decides to commit or abort all
C. All participants commit their local transactions immediately
D. The coordinator asks participants if they can commit safely

Answer: D

Q1276. A system uses strict two-phase locking. What additional guarantee does this provide over basic two-phase locking?

A. It prevents deadlocks from occurring between transactions
B. It holds all exclusive locks until the transaction commits
C. It allows locks to be released during the growing phase
D. It eliminates the need for any lock management overhead

Answer: B

Q1277. When using write-ahead logging with UNDO/REDO recovery, what must be true before a dirty page can be flushed to disk?

A. The page must have been read by at least one other transaction
B. The buffer pool must be completely full of dirty data pages
C. The log record for the update must be on stable storage
D. The transaction must have already committed successfully

Answer: C

Q1278. A system uses multiversion concurrency control (MVCC). What is the primary advantage over lock-based concurrency?

A. Readers do not block writers and writers do not block readers
B. MVCC eliminates the need for any transaction isolation level
C. MVCC uses less memory than lock-based systems in all cases
D. MVCC guarantees serializability without any additional effort

Answer: A

Q1279. In the ARIES recovery algorithm, what are the three phases of recovery?

A. Prepare, Commit, and Acknowledge phases for transaction recovery
B. Scan, Parse, and Execute phases for query recovery management
C. Analysis, Redo, and Undo phases for crash recovery processing
D. Backup, Restore, and Verify phases for data recovery operations

Answer: C

Q1280. A transaction T1 reads a set of rows matching a condition, then T2 inserts a new row matching that condition, then T1 re-reads. What anomaly is this and how is it prevented?

A. Phantom read anomaly prevented by Serializable isolation level
B. Dirty read anomaly prevented by Read Committed isolation level
C. Non-repeatable read prevented by Repeatable Read isolation level
D. Lost update anomaly prevented by Read Uncommitted isolation level

Answer: A

Q1281. In a system using strict two-phase locking, a transaction T1 holds an exclusive lock on item X and T2 requests a shared lock on X. What happens?

A. T1 is forced to release its exclusive lock for T2 to proceed
B. T2 is permanently aborted and must restart from the beginning
C. T2 immediately gets the shared lock alongside the exclusive
D. T2 must wait until T1 releases the exclusive lock on item X

Answer: D

Q1282. A database uses multigranularity locking with IS, IX, S, SIX, and X modes. Which lock mode allows reading an entire table while intending to write some rows?

A. SIX (Shared with Intention Exclusive) lock on table
B. IS (Intention Shared) lock mode on the entire table
C. S (Shared) lock mode on the entire table resource
D. IX (Intention Exclusive) lock mode on the table

Answer: A

Q1283. When using timestamp ordering, what happens when a transaction tries to write an item with a read timestamp greater than its own?

A. The write proceeds normally without any restriction applied
B. The transaction is rolled back and restarted with new timestamp
C. The read timestamp is updated to match the write transaction
D. The item is locked until both transactions complete their work

Answer: B

Q1284. In optimistic concurrency control, what are the three phases of transaction execution?

A. Prepare, Commit, and Cleanup phases for transaction flow
B. Read, Validation, and Write phases for each transaction
C. Parse, Optimize, and Execute phases for query processing
D. Lock, Execute, and Release phases for each transaction

Answer: B

Q1285. A system detects deadlocks using a wait-for graph. What does a cycle in this graph indicate?

A. No conflicts exist between the transactions in the cycle
B. A deadlock exists among the transactions in the cycle now
C. The transactions in the cycle have all committed already
D. All transactions in the cycle can proceed without issues

Answer: B

Q1286. When using predicate locking to prevent phantoms, what is the main practical limitation?

A. Predicate locking requires all transactions to run serially
B. Predicate locking is computationally expensive to implement
C. Predicate locks cannot prevent any type of data anomaly
D. Predicate locks only work with exclusive lock modes applied

Answer: B

Q1287. In snapshot isolation, what specific anomaly can still occur that true serializability prevents?

A. Phantom reads from newly inserted rows matching conditions
B. Dirty reads from uncommitted transaction data values
C. Write skew where two transactions write based on stale reads
D. Non-repeatable reads from concurrent update operations

Answer: C

Q1288. A database uses tree protocol locking. What constraint does this protocol impose on lock acquisition?

A. All locks must be acquired simultaneously before any operation
B. Locks must be acquired in alphabetical order of data items
C. Locks can only be acquired on leaf nodes of the tree structure
D. First lock can be on any item but subsequent locks follow tree

Answer: D

Q1289. When comparing wound-wait and wait-die for deadlock prevention, which scheme causes more unnecessary rollbacks and why?

A. Wait-die causes more because younger transactions die on conflict
B. Neither scheme causes any unnecessary rollbacks by their design
C. Wound-wait causes more because it preempts younger transactions
D. Both schemes cause exactly the same number of rollbacks always

Answer: A

Q1290. In a real-time database system, how should the concurrency control protocol be modified to meet timing constraints?

A. Use standard two-phase locking with no modifications at all
B. Disable concurrency control entirely to maximize throughput
C. Use priority-based protocols that consider transaction deadlines
D. Lock all data items at the start to prevent any waiting time

Answer: C

Q1291. In the ARIES recovery algorithm, what is the purpose of the analysis phase?

A. To determine dirty pages and active transactions at crash time
B. To redo all committed transactions found in the log records
C. To create a new checkpoint after the recovery process completes
D. To undo all uncommitted transactions found in the log records

Answer: A

Q1292. During ARIES redo phase, why does it repeat history by redoing all logged actions?

A. To optimize the recovery process by skipping unnecessary operations
B. To restore the database to its exact state at the moment of crash
C. To ensure only committed transaction changes are applied to data
D. To verify that all log records are correctly formatted and valid

Answer: B

Q1293. In a system using steal/no-force policy, which recovery operations are needed?

A. Neither UNDO nor REDO operations are needed for the system
B. Only REDO operations are needed for committed transactions
C. Both UNDO and REDO operations are needed for full recovery
D. Only UNDO operations are needed for uncommitted transactions

Answer: C

Q1294. A system crash occurs during a checkpoint. How does the recovery process handle this incomplete checkpoint?

A. The incomplete checkpoint is treated as if it completed normally
B. The recovery process ignores the incomplete checkpoint entirely
C. The database is restored from the most recent full backup only
D. Recovery uses the previous complete checkpoint as starting point

Answer: D

Q1295. In physiological logging used by ARIES, what does a log record describe?

A. The logical operation applied to a specific physical page only
B. The exact byte-level changes made to the physical disk blocks
C. The complete SQL statement that was executed by the transaction
D. The schema changes made to the table definition during update

Answer: A

Q1296. A database uses asynchronous replication for disaster recovery. What is the primary risk of this approach?

A. Network bandwidth usage is higher than synchronous replication
B. The primary server may experience increased write latency
C. The replica server cannot handle any read query operations
D. Some committed transactions may be lost if the primary fails

Answer: D

Q1297. How does the compensation log record (CLR) work in the ARIES algorithm?

A. CLR creates a backup of the data before any changes are applied
B. CLR records the original changes made by a committed transaction
C. CLR logs the undo action and points to the next record to undo
D. CLR prevents any further modifications to the recovered data items

Answer: C

Q1298. A system needs sub-second recovery time. Which technique best achieves this requirement?

A. Shadow paging with deferred modification and periodic checkpoints
B. In-memory database with synchronous replication to a standby
C. Full database backup every hour with log replay for recovery
D. Traditional checkpoint-based recovery with large log intervals

Answer: B

Q1299. In a distributed database, what additional complexity does recovery face compared to centralized systems?

A. Distributed databases do not require any transaction log records
B. All sites must use identical hardware and software configurations
C. Recovery in distributed systems is always faster than centralized
D. Coordinating recovery across multiple sites with network issues

Answer: D

Q1300. When using logical undo during recovery, why might a physical undo be insufficient?

A. Physical undo is always sufficient for all recovery scenarios
B. Concurrent transactions may have modified the same page since then
C. Physical undo cannot read the log records written to stable storage
D. Logical undo is faster than physical undo in every possible case

Answer: B

Q1301. A B+ tree with order 4 and 1000 records needs at minimum how many levels including the root?

A. Two levels are needed for this tree configuration setup
B. Four levels are needed for this tree configuration setup
C. Five levels are needed for this tree configuration setup
D. Three levels are needed for this tree configuration setup

Answer: D

Q1302. When should a database administrator choose a hash index over a B+ tree index?

A. When the indexed column has very low cardinality values only
B. When equality queries dominate and data size is predictable
C. When the table is frequently updated with many insert operations
D. When range queries are the predominant access pattern needed

Answer: B

Q1303. In a B+ tree, what happens during a leaf node split when inserting into a full node?

A. The leaf splits into two and the middle key propagates upward
B. The entire tree is rebuilt from scratch with new node sizes
C. All records in the full node are redistributed to sibling nodes
D. The insertion is rejected until space becomes available in node

Answer: A

Q1304. A query accesses 5% of records in a table with one million rows. Is using an index beneficial compared to a full table scan?

A. Index is always beneficial regardless of the percentage accessed
B. Index is never beneficial when accessing more than 1% of records
C. Full table scan is better because 5% means too many random reads
D. It depends on whether the access is sequential or random pattern

Answer: D

Q1305. What is the advantage of using a B+ tree over extendible hashing for concurrent access?

A. B+ tree never requires any locking for concurrent operations
B. B+ tree allows finer granularity locking on individual nodes
C. Extendible hashing cannot support more than one concurrent user
D. Extendible hashing always requires locking the entire directory

Answer: B

Q1306. In a database with write-heavy workload, what is the impact of having too many indexes on a table?

A. Write performance degrades because each index must be updated too
B. There is no impact because indexes only affect read query operations
C. Write performance improves because indexes speed up all operations
D. Read performance degrades because indexes conflict with each other

Answer: A

Q1307. What is a partial index and when is it most useful?

A. An index that is partially stored in memory and partially on disk
B. An index created during off-peak hours for performance reasons
C. An index on all rows but only some columns of the data table
D. An index on a subset of rows matching a predicate condition

Answer: D

Q1308. How does a log-structured merge tree (LSM tree) differ from a B+ tree in write performance?

A. B+ trees are faster for writes due to their balanced tree structure
B. Both structures have identical write performance characteristics
C. LSM trees are faster for writes by using sequential append-only IO
D. LSM trees are slower for writes due to frequent merge operations

Answer: C

Q1309. A table has a composite index on columns (A, B, C). Which queries can efficiently use this index?

A. Queries filtering on A, or A and B, or A and B and C together
B. Only queries filtering on all three columns A, B, and C at once
C. Queries filtering on any single column A or B or C independently
D. Only queries filtering on column C alone can use this index

Answer: A

Q1310. What is the write amplification problem in B+ tree indexes and how do LSM trees address it?

A. LSM trees have worse write amplification than B+ trees in all cases
B. B+ trees have no write amplification issue at all in practice
C. Write amplification only occurs in hash-based indexes not in trees
D. B+ trees cause multiple page writes per update while LSM trees batch

Answer: D

Q1311. When joining three tables R, S, and T, how does the optimizer determine the best join order?

A. It always joins in the order specified in the SQL FROM clause
B. It enumerates possible orders and estimates the cost of each
C. It randomly selects a join order and uses that for execution
D. It always joins the two smallest tables first without analysis

Answer: B

Q1312. A query has a selection condition with OR predicates on different indexed columns. How does the optimizer typically handle this?

A. It creates a temporary composite index on both columns at once
B. It uses index union (bitmap OR) to combine results from indexes
C. It ignores all indexes and performs a full table scan always
D. It only uses the index on the first column mentioned in query

Answer: B

Q1313. In a distributed database, what additional factor must the query optimizer consider compared to centralized systems?

A. The color scheme of the user interface displaying query results
B. Only the CPU cost of processing each query operation step
C. Only the memory usage of temporary tables during query execution
D. The network communication cost of transferring data between sites

Answer: D

Q1314. What is the difference between rule-based and cost-based optimization?

A. Cost-based does not consider indexes while rule-based always does so
B. Rule-based uses predefined rules while cost-based uses data statistics
C. Rule-based is always more accurate than cost-based for complex plans
D. Rule-based considers data statistics while cost-based uses fixed rules

Answer: B

Q1315. How does the optimizer handle correlated subqueries for performance?

A. It ignores correlated subqueries and returns approximate results
B. It always executes correlated subqueries exactly as they are written
C. It may decorrelate them by transforming into joins when possible
D. It converts all correlated subqueries into temporary table scans

Answer: C

Q1316. What is adaptive query processing?

A. Modifying the execution plan during runtime based on actual data
B. Processing queries without any optimization or planning phase
C. Using a fixed plan regardless of the data distribution changes
D. Processing queries only during off-peak hours for efficiency

Answer: A

Q1317. When estimating the cost of a join operation, what is the most critical factor?

A. The estimated size of the intermediate join result produced
B. The number of columns selected in the final output projection
C. The order of columns listed in the SELECT clause of the query
D. The names of the tables being joined in the SQL statement

Answer: A

Q1318. A query execution plan shows a sort-merge join between two large unsorted tables. Under what condition would a hash join be more efficient?

A. When both tables have clustered indexes on the join attribute
B. When the join condition involves a range-based comparison only
C. When one table fits in memory and an equality join is needed
D. When both tables are already sorted on the join attribute key

Answer: C

Q1319. What is the purpose of maintaining multiple access paths in the optimizer search space?

A. To simplify the optimization process by reducing total choices
B. To find the plan with the lowest cost among various alternatives
C. To guarantee that indexes are never used for any query plans
D. To ensure every query produces exactly the same result always

Answer: B

Q1320. How do parameterized queries affect query optimization?

A. They force the optimizer to always perform a full table scan only
B. They prevent the optimizer from generating any execution plan
C. They allow plan caching and reuse across different parameter values
D. They eliminate the need for any statistics collection on columns

Answer: C

Q1321. An application stores passwords in a database. Which approach provides the strongest security?

A. Hashing passwords with salt using a slow algorithm like bcrypt
B. Encrypting passwords with a reversible symmetric algorithm
C. Encoding passwords with Base64 encoding for basic obfuscation
D. Storing passwords in plaintext for easy comparison checks

Answer: A

Q1322. A database administrator needs to implement row-level security for a multi-tenant application. Which approach is most effective?

A. Using views with security predicates based on tenant context
B. Using application-level filtering without any database controls
C. Creating separate tables for each tenant in the same database
D. Giving each tenant a completely separate database instance here

Answer: A

Q1323. What is the risk of a second-order SQL injection compared to a first-order injection?

A. Second-order only affects SELECT queries while first-order all
B. Stored malicious input executes later when used in another query
C. Second-order is less dangerous because the input is validated
D. There is no practical difference between first and second order

Answer: B

Q1324. When implementing database encryption, what is the key management challenge?

A. Encryption keys do not need any management after initial creation
B. Securely storing, rotating, and distributing keys without exposure
C. Key management only matters for symmetric encryption not asymmetric
D. Encryption keys are automatically managed by the operating system

Answer: B

Q1325. A healthcare database must comply with HIPAA regulations. Which database security measures are most critical?

A. Only password protection and regular backups are sufficient here
B. Access control, encryption, audit trails, and data masking combined
C. Only data encryption at rest without any other security measures
D. Only network firewalls without any database-level security controls

Answer: B

Q1326. What is the difference between data-at-rest encryption and data-in-transit encryption?

A. At-rest is optional while in-transit is always mandatory requirement
B. At-rest protects network data while in-transit protects stored data
C. Both protect data at the same stage of the data lifecycle process
D. At-rest protects stored data while in-transit protects network data

Answer: D

Q1327. In a privilege escalation attack on a database, what specific vulnerability is exploited?

A. The attacker gains higher privileges than originally authorized
B. The attacker corrupts data by writing invalid values to tables
C. The attacker creates excessive connections to overload the server
D. The attacker exploits weak network protocols to intercept data

Answer: A

Q1328. A database uses fine-grained auditing. How does this differ from standard auditing?

A. Fine-grained auditing captures specific conditions and column access
B. Standard auditing provides more detail than fine-grained auditing
C. Fine-grained auditing is performed only at the application level
D. Fine-grained auditing only records failed access attempts made

Answer: A

Q1329. What security risk does dynamic SQL pose that static SQL does not?

A. Dynamic SQL is more vulnerable to injection if input not sanitized
B. Dynamic SQL is slower and uses more memory than static SQL code
C. Dynamic SQL bypasses all authentication checks automatically always
D. Static SQL cannot access any tables while dynamic SQL can access all

Answer: A

Q1330. When implementing database security for a cloud-hosted database, what additional concern must be addressed compared to on-premises?

A. Cloud databases are inherently more secure than on-premises ones
B. Shared responsibility model and data sovereignty across regions
C. Only network security matters for cloud-hosted database instances
D. Cloud databases do not require any encryption for stored data

Answer: B

Q1331. In a distributed database using two-phase commit, what is the blocking problem and how does three-phase commit address it?

A. Participants block if coordinator fails; 3PC adds precommit phase
B. Three-phase commit eliminates all network communication needed
C. Blocking only occurs during the abort phase of commit protocol
D. Blocking occurs when a participant crashes before prepare phase

Answer: A

Q1332. A distributed database must choose between strong consistency and high availability during a network partition. According to CAP, which trade-off must be made?

A. The system must choose between consistency or availability only
B. The system should shut down until the partition is resolved
C. The system can maintain both consistency and availability fully
D. The system must sacrifice both consistency and availability

Answer: A

Q1333. When optimizing a distributed query that joins tables at different sites, what is the primary cost to minimize?

A. The CPU processing time at each individual site in network
B. The memory consumption at the coordinator site during join
C. The network data transfer cost between the distributed sites
D. The disk storage used for temporary tables at each site node

Answer: C

Q1334. In a multi-master replication setup, how are write conflicts typically resolved?

A. Using conflict resolution policies like last-writer-wins or merge
B. By routing all writes to a single designated primary master node
C. By preventing all writes until conflict is manually resolved
D. By rejecting all conflicting writes and requiring resubmission

Answer: A

Q1335. What is the Paxos consensus algorithm used for in distributed databases?

A. Replicating data between nodes using synchronous write operations
B. Optimizing query execution plans across distributed network nodes
C. Partitioning tables into fragments for distributed storage placement
D. Achieving agreement on a value among distributed nodes despite failures

Answer: D

Q1336. In a geo-distributed database, what is the primary challenge of maintaining ACID transactions across data centers?

A. Network latency between data centers increases transaction time
B. Geo-distribution eliminates the need for any concurrency control
C. Data centers always have identical hardware and network configs
D. Each data center uses different database management software

Answer: A

Q1337. What is the difference between eager and lazy replication in terms of consistency guarantees?

A. Eager provides eventual consistency while lazy provides strong
B. Both provide identical consistency guarantees in all situations
C. Eager has no consistency guarantee while lazy guarantees it all
D. Eager provides strong consistency while lazy provides eventual

Answer: D

Q1338. A company uses a distributed database with read replicas. How does read-your-writes consistency differ from strong consistency?

A. Read-your-writes ensures all users see the same data at all
B. Read-your-writes and strong consistency are the same guarantee
C. Read-your-writes ensures a user sees their own writes quickly
D. Read-your-writes provides weaker guarantees than eventual one

Answer: C

Q1339. In a distributed database using vector clocks, what do they provide that physical timestamps cannot?

A. Automatic conflict resolution without any additional logic now
B. Higher precision timing than physical clock implementations can
C. Exact wall-clock time ordering of all events across all nodes
D. Causal ordering of events without requiring synchronized clocks

Answer: D

Q1340. When designing a distributed database schema, what is the key consideration for choosing between horizontal and vertical fragmentation?

A. Choose based on query access patterns and data locality needs
B. Vertical fragmentation always results in less network transfer
C. The choice has no impact on query performance or utilization
D. Horizontal fragmentation is always preferred over vertical form

Answer: A

Q1341. When should a system architect choose a document database over a relational database for a new application?

A. When the application has a fixed well-defined relational schema
B. When the application requires complex multi-table join operations
C. When the data model is hierarchical and schema evolves frequently
D. When strong ACID transactions across entities are the top priority

Answer: C

Q1342. In a distributed NoSQL database using consistent hashing, what happens when a new node is added to the cluster?

A. All data is completely redistributed across all existing nodes
B. All existing nodes must be taken offline during the process
C. The new node starts empty and receives no data automatically
D. Only a portion of data from adjacent nodes is redistributed

Answer: D

Q1343. What is the gossip protocol used for in distributed NoSQL databases?

A. Propagating cluster state information between nodes efficiently
B. Encrypting data during transmission between cluster nodes only
C. Executing distributed queries across all nodes in the cluster
D. Rebalancing data partitions when nodes join or leave cluster

Answer: A

Q1344. In a NoSQL database handling a social media feed, why might a fan-out-on-write approach be chosen over fan-out-on-read?

A. Fan-out-on-write reduces write complexity and storage overhead
B. Fan-out-on-write requires less storage than fan-out-on-read
C. Fan-out-on-read eliminates the need for any caching mechanisms
D. Fan-out-on-write provides faster read times at cost of writes

Answer: D

Q1345. What is the trade-off between strong consistency and partition tolerance in a distributed NoSQL system?

A. Strong consistency improves availability during network partition
B. Maintaining strong consistency during partitions requires blocking
C. Partition tolerance is never needed in modern cloud environments
D. Strong consistency and partition tolerance have no trade-off

Answer: B

Q1346. In Cassandra data model, what is the significance of the clustering key?

A. It specifies the replication factor for the data across nodes
B. It determines the consistency level for read-write operations
C. It determines which node stores the data in the cluster ring
D. It defines the sort order of data within a partition on disk

Answer: D

Q1347. When designing a NoSQL schema for a real-time analytics application, what is the primary modeling principle?

A. Design schema based on entity relationships like relational models
B. Design schema to maximize normalization for data integrity only
C. Design schema to minimize data redundancy above all other goals
D. Design schema based on the query patterns the application needs

Answer: D

Q1348. What is a vector clock conflict resolution limitation in eventually consistent NoSQL systems?

A. Vector clocks grow in size with the number of participating nodes
B. Vector clocks cannot detect any concurrent write conflicts at all
C. Vector clocks only work with key-value stores and no other types
D. Vector clocks require perfectly synchronized physical time clocks

Answer: A

Q1349. In a multi-model NoSQL database, what advantage does combining document and graph capabilities provide?

A. It allows modeling both structured data and complex relationships
B. It reduces storage requirements compared to single-model databases
C. It guarantees ACID transactions across all data model types used
D. It eliminates the need for any indexing on the stored data sets

Answer: A

Q1350. A NoSQL database uses quorum-based replication with N=5, W=3, R=3. What consistency guarantee does this provide?

A. Strong consistency since W+R exceeds N ensuring overlap exists
B. Eventual consistency with no read-your-writes guarantee at all
C. No consistency guarantee regardless of the quorum configuration
D. Consistency only for write operations not for read operations

Answer: A

Q1351. When designing an enterprise data warehouse, what is the key difference between Inmon top-down and Kimball bottom-up approach?

A. Inmon builds data marts first while Kimball builds warehouse
B. Both approaches build the data warehouse in exactly same order
C. Inmon uses star schemas while Kimball uses snowflake schemas
D. Inmon builds normalized warehouse first while Kimball marts first

Answer: D

Q1352. In a Type 2 slowly changing dimension, how are historical changes tracked?

A. Changes are not tracked and old values are permanently deleted
B. The existing record is overwritten with the new attribute values
C. A new record is added with effective dates preserving old ones
D. A separate history table stores all previous versions of records

Answer: C

Q1353. What is a degenerate dimension in data warehouse design?

A. A dimension containing only null values for all its attributes
B. A dimension key in the fact table without a dimension table
C. A dimension stored as a separate fully normalized table entity
D. A dimension that has been deleted from the warehouse schema

Answer: B

Q1354. When designing a fact table, what is the difference between additive, semi-additive, and non-additive measures?

A. Non-additive measures can be summed across time dimension only
B. Semi-additive can be summed across all dimensions but additive cannot
C. Additive measures can be summed across all dimensions while others not
D. All three types of measures can be summed across any dimension

Answer: C

Q1355. A company implements a real-time data warehouse. What architectural change is needed compared to traditional batch ETL?

A. Use only flat files instead of database tables for data storage
B. Implement change data capture and micro-batch or stream processing
C. Replace ETL with ELT and remove all transformation logic entirely
D. Remove all indexes from the warehouse for faster write performance

Answer: B

Q1356. In data mining, what is the difference between classification and clustering?

A. Both methods use identical approaches with labeled training data sets
B. Clustering uses labeled data while classification uses unlabeled data
C. Classification uses labeled data while clustering uses unlabeled data
D. Classification is unsupervised while clustering is supervised learning

Answer: C

Q1357. What is a factless fact table and when is it used in data warehouse design?

A. A fact table that stores future predictions instead of facts
B. A fact table with no numeric measures recording only events
C. A fact table with no dimension keys linking to other tables
D. A fact table that is temporarily empty during initial loading

Answer: B

Q1358. When implementing ELT instead of ETL, what is the key architectural difference?

A. ELT does not require any extraction from source systems at all
B. Data is loaded first then transformed inside the target warehouse
C. Data is transformed before loading into the target warehouse
D. ELT skips the loading step entirely and queries sources directly

Answer: B

Q1359. What is a conformed dimension in Kimball dimensional modeling and why is it important?

A. A dimension that has been fully normalized into sub-dimensions
B. A dimension shared across multiple data marts for consistency
C. A dimension containing only surrogate keys without attributes
D. A dimension used in only one specific data mart exclusively here

Answer: B

Q1360. In association rule mining, what do the metrics support and confidence measure?

A. Support measures rule accuracy while confidence measures frequency
B. Support measures query performance while confidence measures quality
C. Support measures item frequency while confidence measures strength
D. Support measures data volume while confidence measures processing

Answer: C

Q1361. In a DBMS, what does the concept of impedance mismatch refer to?

A. Mismatch between different hardware configurations
B. Mismatch between the data model used by the application and the database
C. Mismatch between primary and foreign key types
D. Mismatch between storage media speeds

Answer: B

Q1362. Which of the following correctly describes the closed world assumption in databases?

A. All data in the database is encrypted by default
B. Facts not present in the database are considered false
C. The database cannot accept new data after initial loading
D. All queries must return results

Answer: B

Q1363. What is the significance of Codd's 12 rules in database systems?

A. They define network protocol standards
B. They establish criteria for a system to qualify as a relational DBMS
C. They specify hardware requirements for databases
D. They outline data encryption algorithms

Answer: B

Q1364. Which Codd rule states that all information in a relational database must be represented as values in tables?

A. Rule 0: Foundation Rule
B. Rule 1: Information Rule
C. Rule 5: Comprehensive Data Sublanguage Rule
D. Rule 8: Physical Data Independence

Answer: B

Q1365. In database systems, what does the principle of minimal redundancy aim to achieve?

A. Maximum storage utilization by duplicating data
B. Storing each fact in exactly one place to avoid update anomalies
C. Removing all indexes to save space
D. Eliminating all foreign key relationships

Answer: B

Q1366. What problem does the lost update anomaly cause in concurrent database access?

A. Data is permanently deleted from the database
B. One transaction overwrites changes made by another concurrent transaction
C. Indexes become corrupted during updates
D. The database schema is modified unintentionally

Answer: B

Q1367. Which of the following scenarios best illustrates the need for logical data independence?

A. Changing disk storage from HDD to SSD
B. Adding a new column to a table without modifying existing applications
C. Moving the database to a different server
D. Increasing the buffer pool size

Answer: B

Q1368. What is the phantom read problem in database concurrency?

A. A transaction reads data that has been permanently deleted
B. A transaction reads rows that were inserted by another uncommitted transaction
C. A transaction cannot find previously existing records
D. The database returns empty result sets randomly

Answer: B

Q1369. In the context of database evolution, what advantage did the relational model provide over the hierarchical model?

A. Faster physical storage access
B. Data independence and flexible querying through declarative languages
C. Built-in support for multimedia data
D. Automatic data replication

Answer: B

Q1370. What is the dirty read problem in database transactions?

A. Reading data from a corrupted disk sector
B. Reading uncommitted data written by another transaction that may roll back
C. Reading outdated cached data from the buffer
D. Reading data without proper authentication

Answer: B

Q1371. In a shared-nothing parallel database architecture, how is fault tolerance typically achieved?

A. By sharing memory across all nodes
B. Through data replication across multiple independent nodes
C. By using a single centralized backup server
D. Through shared disk access

Answer: B

Q1372. What is the primary challenge of implementing a shared-disk architecture?

A. Limited network bandwidth
B. Managing concurrent disk access and cache coherence across nodes
C. Insufficient disk storage
D. Slow query parsing

Answer: B

Q1373. In the context of DBMS architecture, what is the write-ahead log (WAL) protocol?

A. All write operations are performed after user confirmation
B. Log records must be written to stable storage before the corresponding data changes are written to disk
C. Data is written ahead of any read operation
D. Logs are written only after transaction completion

Answer: B

Q1374. What is the purpose of the ARIES recovery algorithm in DBMS architecture?

A. To optimize query execution plans
B. To provide efficient and correct crash recovery using WAL
C. To manage user authentication
D. To allocate memory buffers

Answer: B

Q1375. In a multi-tier architecture, what is the benefit of using an application server as a connection broker?

A. It eliminates the need for a database entirely
B. It multiplexes many client connections into fewer database connections, reducing server load
C. It automatically generates SQL queries
D. It compresses all data before storage

Answer: B

Q1376. What is the CAP theorem's relevance to distributed database architecture?

A. It states that a database can achieve consistency, availability, and partition tolerance simultaneously
B. It proves that a distributed system can guarantee at most two of consistency, availability, and partition tolerance
C. It defines the maximum number of concurrent users
D. It specifies the minimum hardware requirements

Answer: B

Q1377. What is the difference between inter-query and intra-query parallelism?

A. Inter-query parallelism is faster than intra-query parallelism
B. Inter-query runs different queries in parallel while intra-query parallelizes a single query
C. Inter-query uses more memory while intra-query uses more disk
D. They are the same concept with different names

Answer: B

Q1378. In DBMS architecture, what is the purpose of the checkpoint mechanism?

A. To mark the completion of a user session
B. To periodically flush modified buffers to disk and record a checkpoint in the log to limit recovery time
C. To check database integrity constraints
D. To verify user permissions

Answer: B

Q1379. What architectural pattern does database sharding implement?

A. Vertical partitioning across different schemas
B. Horizontal partitioning of data across multiple independent database instances
C. Replication of entire databases to standby servers
D. Compression of data into fewer storage nodes

Answer: B

Q1380. In a replicated database architecture, what is the split-brain problem?

A. A query that is too complex for the optimizer
B. When network partition causes two nodes to independently accept writes, leading to conflicting states
C. When a database table is split across two different schemas
D. When CPU usage is divided between query processing and logging

Answer: B

Q1381. When converting an ER diagram to a relational schema, how is a multivalued attribute handled?

A. It is stored as a comma-separated list in the parent table
B. A separate table is created with a foreign key referencing the parent entity
C. It is ignored during conversion
D. It becomes a primary key in the parent table

Answer: B

Q1382. In the EER model, what is a shared subclass (multiple inheritance)?

A. A subclass that inherits attributes from only one superclass
B. A subclass that is a subset of two or more superclasses
C. A subclass with no attributes of its own
D. A superclass with no subclasses

Answer: B

Q1383. When mapping a ternary relationship to a relational schema, how is the resulting table structured?

A. It contains only the primary key of one entity
B. It contains the primary keys of all three participating entities as foreign keys
C. It creates three separate binary relationship tables
D. It merges all three entities into a single table

Answer: B

Q1384. What is the temporal data model's primary extension to the relational model?

A. Adding spatial coordinates to every record
B. Incorporating time dimensions to track data changes over time
C. Replacing tables with graphs
D. Removing the need for primary keys

Answer: B

Q1385. In the context of ER modeling, what is the minimum cardinality constraint?

A. The maximum number of relationships an entity can have
B. The minimum number of relationship instances each entity must participate in
C. The minimum number of attributes an entity must have
D. The minimum number of tables in a database

Answer: B

Q1386. How does the attribute-defined specialization differ from user-defined specialization in EER?

A. Attribute-defined uses a discriminating attribute while user-defined is determined manually
B. Attribute-defined is faster to process
C. User-defined cannot have disjoint constraints
D. There is no difference between them

Answer: A

Q1387. What happens to a total specialization constraint when converting an EER diagram to a relational schema?

A. It is automatically enforced by the schema
B. It cannot be directly enforced in the relational model and requires application-level or trigger enforcement
C. It is converted to a CHECK constraint
D. It becomes a foreign key constraint

Answer: B

Q1388. In the object-relational data model, what are user-defined types (UDTs)?

A. Types that only the database administrator can use
B. Custom data types created by users that can encapsulate attributes and methods
C. Temporary types used only during queries
D. Types that replace all standard SQL types

Answer: B

Q1389. What is the concept of closure in the context of data model operations?

A. The database cannot accept new connections
B. The result of any operation on relations is also a relation
C. All transactions must be closed before new ones begin
D. The database is shut down properly

Answer: B

Q1390. When mapping overlapping specialization to a relational schema, which approach is most appropriate?

A. Create a single table for all subclasses with a type column
B. Create separate tables for each subclass with redundant superclass attributes
C. Create separate tables for each subclass with foreign keys to the superclass table
D. Merge all entities into the superclass table only

Answer: C

Q1391. What is Armstrong's axiom of reflexivity in functional dependencies?

A. If Y is a subset of X, then X functionally determines Y
B. If X determines Y, then Y determines X
C. If X determines Y and Y determines Z, then X determines Z
D. If X determines Y, then XZ determines YZ

Answer: A

Q1392. What is the transitive rule in Armstrong's axioms?

A. If X -> Y, then Y -> X
B. If X -> Y and Y -> Z, then X -> Z
C. If X -> Y, then XZ -> YZ
D. If Y is a subset of X, then X -> Y

Answer: B

Q1393. What is the closure of a set of attributes under a set of functional dependencies?

A. The set of all attributes that can be functionally determined from the given attributes
B. The removal of all redundant attributes
C. The set of all tables referencing the attributes
D. The physical storage allocation for the attributes

Answer: A

Q1394. What is a lossless join decomposition?

A. A decomposition where no data is physically deleted
B. A decomposition where the original relation can be perfectly reconstructed by joining the decomposed relations
C. A decomposition that reduces storage by 50%
D. A decomposition that creates no foreign keys

Answer: B

Q1395. What is a dependency-preserving decomposition?

A. A decomposition where all data dependencies are documented
B. A decomposition where all original functional dependencies can be enforced using the decomposed relations
C. A decomposition that preserves physical storage format
D. A decomposition that maintains the same number of rows

Answer: B

Q1396. What is a multivalued dependency in relational theory?

A. A dependency where an attribute has multiple data types
B. A dependency where one attribute set determines a set of values independently of other attributes
C. A dependency between multiple databases
D. A dependency that only exists in NoSQL databases

Answer: B

Q1397. What is a join dependency in relational database theory?

A. A constraint that two tables must always be joined in queries
B. A constraint where a relation can be losslessly decomposed into multiple relations and reconstructed via natural join
C. A dependency between join operations
D. A physical storage constraint for joined tables

Answer: B

Q1398. What is a trivial functional dependency?

A. A dependency that is obvious to the database designer
B. A dependency X -> Y where Y is a subset of X
C. A dependency that involves only primary keys
D. A dependency that can be derived from other dependencies

Answer: B

Q1399. In relational algebra, what is the difference between division and other operations?

A. Division is faster than other operations
B. Division finds tuples in one relation that are associated with all tuples in another relation
C. Division removes duplicate tuples
D. Division is used for aggregation

Answer: B

Q1400. What does the augmentation rule in Armstrong's axioms state?

A. If X -> Y, then X -> YZ for any Z
B. If X -> Y, then XZ -> YZ for any attribute set Z
C. If X -> Y and Y -> Z, then X -> Z
D. If XY -> Z, then X -> Z and Y -> Z

Answer: B

Q1401. How can the division operation R ÷ S be expressed using fundamental relational algebra operations?

A. R ÷ S = R - S
B. R ÷ S = < ,…'"' Ò ÀX((< ,…'"' r 2’ Ò "' v†W&R , Ò GG ibutes of R not in S
C. R ÷ S = R ") S
D. R ÷ S = <2…" r 2•

Answer: B

Q1402. What is the algebraic equivalence rule for pushing selection through a join?

A. Selections cannot be pushed through joins
B. <62…" ,, 2’ Ò Ãc(R) "È S if condition c involves only attributes of R
C. <62…" ,, 2’ Ò " ,, Ãc(S) regardless of which relation c references
D. Selections must always be applied after joins

Answer: B

Q1403. In query optimization, why is pushing projections down the query tree beneficial?

A. It increases the number of columns processed
B. It reduces the width of intermediate results, decreasing memory and processing costs
C. It makes queries more readable
D. It eliminates the need for selection operations

Answer: B

Q1404. What is the difference between tuple relational calculus and domain relational calculus?

A. They are identical in expressive power and syntax
B. Tuple calculus uses tuple variables ranging over relations while domain calculus uses variables ranging over attribute domains
C. Tuple calculus is more powerful than domain calculus
D. Domain calculus cannot express joins

Answer: B

Q1405. What does the concept of safe expressions mean in relational calculus?

A. Expressions that do not cause runtime errors
B. Expressions guaranteed to produce a finite result set
C. Expressions that do not modify the database
D. Expressions that use only primary keys

Answer: B

Q1406. How is the outer union operation different from the standard union?

A. Outer union requires union compatibility
B. Outer union can combine relations that are not union compatible by using nulls for missing attributes
C. Outer union removes all duplicates while union does not
D. There is no difference

Answer: B

Q1407. What is the commutative property of natural join in relational algebra?

A. R "È S is not equal to S "È R
B. R "È S = S "È R, meaning the order of operands does not affect the result
C. Natural join can only be applied in one direction
D. Commutativity only holds for equi-joins

Answer: B

Q1408. In relational algebra, what is the associative property of join?

A. Joins cannot be reordered
B. (R "È S) "È T = R "È (S "È T), allowing join reordering for optimization
C. Only two relations can be joined at a time
D. Associativity only applies to Cartesian products

Answer: B

Q1409. What is the expressive power relationship between relational algebra and safe relational calculus?

A. Relational algebra is more powerful than relational calculus
B. Relational calculus is more powerful than relational algebra
C. They are equivalent in expressive power for safe expressions
D. Neither can express joins

Answer: C

Q1410. What is the deferred operation approach in evaluating relational algebra expressions?

A. All operations are executed immediately
B. Operations are composed into a query tree and optimized before execution
C. Operations are deferred until the database is idle
D. Only selection operations are deferred

Answer: B

Q1411. What is the execution order of clauses in a SQL SELECT statement?

A. SELECT, FROM, WHERE, GROUP BY, HAVING, ORDER BY
B. FROM, WHERE, GROUP BY, HAVING, SELECT, ORDER BY
C. FROM, SELECT, WHERE, GROUP BY, ORDER BY, HAVING
D. SELECT, WHERE, FROM, HAVING, GROUP BY, ORDER BY

Answer: B

Q1412. What is a recursive Common Table Expression (CTE) used for in SQL?

A. Creating temporary indexes
B. Querying hierarchical or recursive data structures like organizational charts
C. Improving query performance automatically
D. Creating permanent views

Answer: B

Q1413. What is the difference between RANK() and DENSE_RANK() window functions?

A. They always produce the same result
B. RANK() leaves gaps after ties while DENSE_RANK() does not
C. DENSE_RANK() leaves gaps after ties while RANK() does not
D. RANK() is faster than DENSE_RANK()

Answer: B

Q1414. What is a lateral join in SQL?

A. A join that connects tables horizontally
B. A join where the right-hand subquery can reference columns from the left-hand table
C. A standard cross join with an alias
D. A join used exclusively in NoSQL databases

Answer: B

Q1415. What does the SQL MERGE statement do?

A. Merges two databases into one
B. Performs insert, update, or delete operations in a single statement based on match conditions
C. Combines query results like UNION
D. Merges table schemas

Answer: B

Q1416. What is the difference between a scalar subquery and a table subquery?

A. They are the same thing
B. A scalar subquery returns exactly one value while a table subquery returns a result set
C. A scalar subquery is faster
D. A table subquery cannot be used in WHERE clauses

Answer: B

Q1417. What is a window function's PARTITION BY clause used for?

A. To partition a table into separate physical storage
B. To divide the result set into partitions to which the window function is independently applied
C. To create table partitions for archiving
D. To split a query into parallel operations

Answer: B

Q1418. What is the purpose of the ROWS BETWEEN clause in window functions?

A. To limit the total number of rows returned
B. To define the window frame specifying which rows relative to the current row are included in the calculation
C. To filter rows before the window function is applied
D. To specify the number of rows per page

Answer: B

Q1419. What is the NTH_VALUE() window function used for?

A. To find the nth database in a server
B. To retrieve the value of an expression at the nth row within a window frame
C. To calculate the nth percentile
D. To insert values at the nth position

Answer: B

Q1420. What is SQL injection and how is it prevented?

A. A method to inject performance optimizations; prevented by disabling indexes
B. A security vulnerability where malicious SQL is inserted into queries; prevented using parameterized queries
C. A technique to insert bulk data; prevented by limiting batch sizes
D. A method to inject stored procedures; prevented by disabling triggers

Answer: B

Q1421. What is the purpose of EXPLAIN or EXPLAIN PLAN in SQL?

A. To explain SQL syntax to beginners
B. To display the execution plan showing how the DBMS will process a query
C. To explain error messages in detail
D. To document stored procedures

Answer: B

Q1422. What is a pivot operation in SQL?

A. Rotating a table physically on screen
B. Transforming row data into column data for cross-tabulation reporting
C. Reversing the order of rows
D. Swapping primary and foreign keys

Answer: B

Q1423. What is the difference between OLTP and OLAP in terms of SQL usage?

A. OLTP uses SQL while OLAP does not
B. OLTP uses simple short transactions while OLAP uses complex analytical queries
C. OLTP is for analytics while OLAP is for transactions
D. There is no difference

Answer: B

Q1424. What is a temporal table in SQL?

A. A table that is deleted after each session
B. A system-versioned table that automatically tracks data changes over time
C. A table used only for temporary calculations
D. A table with timestamp columns

Answer: B

Q1425. What are the ACID properties' implications for SQL transaction design?

A. They require all queries to be written in uppercase
B. They ensure that transactions are atomic, consistent, isolated, and durable for reliable processing
C. They define the SQL syntax rules
D. They specify hardware requirements for databases

Answer: B

Q1426. What is the purpose of the GROUPING SETS clause in SQL?

A. To create multiple user groups
B. To specify multiple grouping combinations in a single query instead of multiple UNION queries
C. To set group permissions
D. To group database connections

Answer: B

Q1427. What is a deferred constraint in SQL?

A. A constraint that is never checked
B. A constraint whose enforcement is postponed until the end of the transaction
C. A constraint applied only to archived data
D. A constraint that is always disabled

Answer: B

Q1428. What is the FETCH FIRST or LIMIT clause used for with OFFSET?

A. To limit the database size
B. To implement pagination by skipping a number of rows and returning a limited set
C. To limit the number of database connections
D. To restrict the number of columns returned

Answer: B

Q1429. What is the difference between CUBE and ROLLUP in SQL?

A. They produce identical results
B. ROLLUP generates hierarchical subtotals while CUBE generates subtotals for all combinations of grouping columns
C. CUBE is faster than ROLLUP
D. ROLLUP generates more rows than CUBE

Answer: B

Q1430. What is Fifth Normal Form (5NF) or Project-Join Normal Form?

A. A relation with five or fewer attributes
B. A relation in 4NF with no non-trivial join dependencies that are not implied by candidate keys
C. A relation that requires five decompositions
D. A theoretical form that cannot be achieved in practice

Answer: B

Q1431. What is Domain-Key Normal Form (DKNF)?

A. A form where all constraints are based on domains and keys only
B. A form that normalizes domain names
C. A form that requires key encryption
D. A form applicable only to key-value databases

Answer: A

Q1432. Why might a BCNF decomposition not preserve all functional dependencies?

A. BCNF does not use functional dependencies
B. Some dependencies may span across decomposed relations and cannot be checked without joining them
C. BCNF only considers primary keys
D. Dependencies are automatically removed during decomposition

Answer: B

Q1433. What is the minimal cover algorithm used for in normalization?

A. To find the minimum number of tables
B. To reduce a set of functional dependencies to an equivalent minimal set
C. To cover all tables with indexes
D. To minimize the number of rows in a table

Answer: B

Q1434. In the context of normalization, what is a spurious tuple?

A. A tuple with all null values
B. An invalid tuple generated by incorrectly joining decomposed relations
C. A tuple that violates a domain constraint
D. A duplicate tuple in a relation

Answer: B

Q1435. How does the 3NF synthesis algorithm ensure dependency preservation?

A. By keeping all attributes in one table
B. By creating a separate relation for each functional dependency in the minimal cover
C. By ignoring dependencies during decomposition
D. By enforcing all dependencies through triggers

Answer: B

Q1436. What is the relationship between multivalued dependencies and Fourth Normal Form?

A. 4NF ignores multivalued dependencies
B. 4NF requires that all non-trivial multivalued dependencies have a superkey on their left side
C. Multivalued dependencies only exist in 1NF
D. 4NF eliminates all functional dependencies

Answer: B

Q1437. What is over-normalization and what are its consequences?

A. Having too many columns in a table
B. Excessive decomposition that leads to too many joins, degrading query performance
C. Having too many indexes
D. Normalizing data that should be kept as text

Answer: B

Q1438. What is the Heath's theorem in normalization?

A. Every relation can be decomposed into 1NF
B. If R(A,B,C) has A->B, then R can be losslessly decomposed into R1(A,B) and R2(A,C)
C. All relations must have at least three attributes
D. Normalization always preserves dependencies

Answer: B

Q1439. What is the difference between the decomposition and synthesis approaches to achieving 3NF?

A. They always produce identical results
B. Decomposition starts from a universal relation and decomposes, while synthesis builds relations from functional dependencies
C. Synthesis is a top-down approach while decomposition is bottom-up
D. Decomposition is always preferred over synthesis

Answer: B

Q1440. What is the Thomas Write Rule in transaction processing?

A. All writes must be timestamped
B. An outdated write can be ignored if a more recent transaction has already written to the same item
C. Writes must always be performed before reads
D. Thomas must approve all writes

Answer: B

Q1441. What is the write-ahead logging (WAL) protocol for transaction recovery?

A. All writes are performed ahead of other operations
B. Log records must be written to stable storage before the actual database modifications
C. All log files must be written to tape
D. Write operations must precede read operations

Answer: B

Q1442. What is the difference between the READ COMMITTED and REPEATABLE READ isolation levels?

A. They provide the same guarantees
B. READ COMMITTED prevents dirty reads while REPEATABLE READ also prevents non-repeatable reads
C. REPEATABLE READ is less restrictive than READ COMMITTED
D. READ COMMITTED prevents phantom reads while REPEATABLE READ does not

Answer: B

Q1443. What is the SERIALIZABLE isolation level and what anomalies does it prevent?

A. It only prevents dirty reads
B. It prevents all anomalies including dirty reads, non-repeatable reads, and phantom reads
C. It only prevents phantom reads
D. It provides no isolation guarantees

Answer: B

Q1444. What is a distributed transaction and what protocol manages it?

A. A transaction that runs on multiple distributed disks; managed by RAID
B. A transaction spanning multiple database nodes; managed by two-phase commit protocol
C. A transaction distributed across time zones; managed by clock sync
D. A transaction split into smaller parts; managed by thread pools

Answer: B

Q1445. What is a compensation transaction?

A. A transaction that pays for database usage
B. A transaction that semantically undoes the effects of a previously committed transaction
C. A transaction that compensates for hardware failures
D. A backup transaction that runs automatically

Answer: B

Q1446. What is snapshot isolation and how does it differ from true serializability?

A. They are identical
B. Snapshot isolation provides each transaction a consistent view at its start but may allow write skew anomalies that serializability prevents
C. Snapshot isolation is stronger than serializability
D. Snapshot isolation does not prevent dirty reads

Answer: B

Q1447. What is the write skew anomaly in transaction processing?

A. When a write operation skews data distribution
B. When two concurrent transactions read overlapping data and make disjoint updates that together violate a constraint
C. When disk writes are not aligned properly
D. When write operations are slower than reads

Answer: B

Q1448. What is the phantom problem in concurrency control and how does it differ from a non-repeatable read?

A. They are the same problem
B. Phantoms involve new rows appearing in a re-executed query while non-repeatable reads involve changes to existing rows
C. Phantoms are caused by hardware issues
D. Non-repeatable reads involve new rows

Answer: B

Q1449. What is predicate locking and why is it used?

A. Locking based on column names
B. Locking based on the condition (predicate) of a query to prevent phantom reads
C. Locking only primary key columns
D. Locking rows in a predetermined order

Answer: B

Q1450. What is index locking (next-key locking) as a practical alternative to predicate locking?

A. Locking the entire index structure
B. Locking index entries and gaps between them to prevent phantoms without the overhead of predicate locking
C. Locking only the first index entry
D. Creating a lock on every index in the database

Answer: B

Q1451. In the multiple-granularity locking protocol, what does an IS (Intention Shared) lock on a table indicate?

A. The transaction intends to insert new rows
B. The transaction intends to acquire shared locks on some rows within the table
C. The table is being shared across databases
D. The index structure is being modified

Answer: B

Q1452. What is the SIX (Shared with Intention Exclusive) lock mode used for?

A. Locking six data items simultaneously
B. A transaction that reads all rows but intends to update only some rows in the locked item
C. A special lock for six-column tables
D. A combination of six different lock types

Answer: B

Q1453. What is the validation phase in optimistic concurrency control?

A. Validating SQL syntax before execution
B. Checking at commit time whether the transaction has conflicted with other transactions during its execution
C. Validating user credentials
D. Checking database schema validity

Answer: B

Q1454. What is the difference between deadlock prevention and deadlock detection?

A. They are the same approach
B. Prevention avoids deadlocks proactively using protocols while detection allows deadlocks and resolves them when found
C. Prevention is always more efficient
D. Detection prevents deadlocks from occurring

Answer: B

Q1455. In timestamp ordering, what happens when a transaction tries to read a data item written by a younger transaction?

A. The read proceeds normally
B. The transaction is rolled back and restarted with a new timestamp
C. The younger transaction is rolled back
D. Both transactions are aborted

Answer: B

Q1456. What is the concept of lock escalation in database systems?

A. Increasing the lock timeout duration
B. Automatically converting many fine-grained locks to fewer coarse-grained locks to reduce overhead
C. Escalating lock priorities for important transactions
D. Moving locks from memory to disk

Answer: B

Q1457. What is the relationship between two-phase locking and serializability?

A. 2PL does not guarantee serializability
B. 2PL is a sufficient but not necessary condition for conflict serializability
C. 2PL is both necessary and sufficient for serializability
D. 2PL only guarantees view serializability

Answer: B

Q1458. In ARIES recovery, why does the redo phase redo all logged actions including those of uncommitted transactions?

A. To simplify the algorithm
B. To restore the exact state of the database at crash time (repeating history) before selectively undoing uncommitted changes
C. To commit all active transactions
D. To verify log integrity

Answer: B

Q1459. What is the role of the dirty page table in ARIES recovery?

A. A list of pages with errors
B. It tracks modified pages in the buffer pool with their recovery LSN to determine the starting point for redo
C. A list of pages that need cleaning
D. A table storing deleted pages

Answer: B

Q1460. What is the concept of idempotent redo in recovery algorithms?

A. Redo operations that can only be performed once
B. Redo operations that produce the same result regardless of how many times they are applied
C. Redo operations that are twice as fast
D. Redo operations that skip duplicate records

Answer: B

Q1461. What is the difference between physical and logical logging?

A. Physical logging is done on paper while logical is digital
B. Physical logging records the actual before/after data values while logical logging records the operation performed
C. Physical logging is faster
D. They produce identical log files

Answer: B

Q1462. In a no-steal, force buffer policy, why is recovery simpler but performance worse?

A. Because no recovery is ever needed
B. No undo is needed (uncommitted data never on disk) and no redo is needed (committed data always on disk), but forcing writes at commit is expensive
C. Because the log is not needed
D. Because all transactions are serialized

Answer: B

Q1463. What is the shadow paging recovery technique?

A. Creating backup pages in shadow memory
B. Maintaining a shadow page table pointing to the original pages while current modifications go to new pages
C. Using dark mode for the database interface
D. Logging shadows of transactions

Answer: B

Q1464. What is the nested top actions concept in ARIES?

A. Transactions nested inside other transactions
B. A mechanism to protect certain actions from being undone during rollback while still being part of the transaction
C. Top-level actions that contain sub-actions
D. Actions performed at the top of the log file

Answer: B

Q1465. How does log-based recovery handle distributed transactions after a crash?

A. Each node recovers independently without coordination
B. The coordinator consults its log to determine the commit decision and informs participants to redo or undo accordingly
C. All distributed transactions are automatically aborted
D. Recovery is not possible for distributed transactions

Answer: B

Q1466. What is the purpose of the transaction table in ARIES?

A. A table listing all tables in the database
B. It tracks the state and last LSN of each active transaction for determining which transactions to undo
C. A table of transaction prices
D. A lookup table for transaction types

Answer: B

Q1467. What is the difference between hot backup and cold backup?

A. Hot backup uses heated storage media
B. Hot backup is taken while the database is running while cold backup requires shutting down the database
C. Cold backup is faster than hot backup
D. They produce identical backups

Answer: B

Q1468. What is a function-based index?

A. An index on a column that stores function names
B. An index built on the result of a function or expression applied to one or more columns
C. A function that returns index values
D. An index that uses functional programming concepts

Answer: B

Q1469. What is the difference between a B+ tree index and a log-structured merge tree (LSM tree)?

A. They are the same structure with different names
B. B+ trees optimize for reads with in-place updates while LSM trees optimize for writes using sequential I/O and periodic compaction
C. LSM trees cannot perform range queries
D. B+ trees are always superior to LSM trees

Answer: B

Q1470. What is a partial index?

A. An incomplete index that does not work correctly
B. An index built on a subset of rows that satisfy a specified condition
C. An index on half the columns of a table
D. An index that partially sorts data

Answer: B

Q1471. What is the relationship between index height and query performance in a B+ tree?

A. Index height has no effect on performance
B. Each additional level of height adds one more disk I/O to the search path
C. Taller indexes are always faster
D. Height only affects insertion performance

Answer: B

Q1472. What is a space-partitioning index like an R-tree used for?

A. Indexing relational tables only
B. Indexing multidimensional spatial data such as geographic coordinates and geometric objects
C. Partitioning disk space
D. Indexing data across multiple tablespaces

Answer: B

Q1473. What is the write amplification problem in index maintenance?

A. Writes become louder on disk
B. A single logical write operation causes multiple physical writes due to index updates and page splits
C. Writing is amplified to be faster
D. The write buffer size increases

Answer: B

Q1474. What is an inverted index and where is it commonly used?

A. An index that sorts data in reverse order
B. A mapping from content (words or terms) to their locations in documents, commonly used in full-text search engines
C. An index that inverts primary and foreign keys
D. A reversed B-tree

Answer: B

Q1475. What is the query optimization search space problem?

A. Finding the optimal physical location for the database
B. The number of possible execution plans grows exponentially with the number of relations, making exhaustive search impractical
C. Finding enough disk space for queries
D. The search for optimal query syntax

Answer: B

Q1476. What is dynamic programming used for in query optimization?

A. Dynamically creating database tables
B. Efficiently finding optimal join orderings by building solutions from smaller subproblems
C. Programming dynamic SQL statements
D. Dynamically allocating memory for queries

Answer: B

Q1477. What is the concept of interesting orders in query optimization?

A. Ordering results in a visually appealing way
B. Sort orders produced by operations that are useful for later operations like merge joins or ORDER BY
C. The most interesting queries to optimize
D. Ordering tables by their creation date

Answer: B

Q1478. What is the difference between a left-deep join tree and a bushy join tree?

A. Left-deep trees are deeper than bushy trees
B. Left-deep trees have only base relations as the right input, enabling pipelining, while bushy trees allow any join result as input
C. Bushy trees are always more efficient
D. Left-deep trees cannot be optimized

Answer: B

Q1479. What is the impact of correlated subqueries on query performance?

A. They always improve performance
B. They can be very expensive because the subquery is re-evaluated for each row of the outer query
C. They have no impact on performance
D. They speed up the outer query

Answer: B

Q1480. What is the Semi-Join Reduction technique in distributed query optimization?

A. Reducing the number of joins by half
B. Reducing data transfer by sending only the join column values to a remote site to filter before transferring matching rows
C. Joining only half the table
D. Reducing the join condition to one column

Answer: B

Q1481. What is the role of a plan cache (query cache) in query optimization?

A. To cache database backups
B. To store previously computed execution plans for reuse, avoiding repeated optimization overhead
C. To cache query results permanently
D. To store database statistics

Answer: B

Q1482. What is the problem with parameter sniffing in query optimization?

A. Parameters are stolen by unauthorized users
B. The optimizer creates a plan based on the first parameter value which may be suboptimal for other values
C. Parameters are not passed correctly to the query
D. Sniffing causes network latency

Answer: B

Q1483. What is the Volcano/Iterator model of query execution?

A. A model based on volcanic data processing
B. Each operator implements a next() method, pulling one tuple at a time from child operators in a demand-driven fashion
C. A model for processing data in eruptions
D. A model that stores data in volcanic rock formations

Answer: B

Q1484. What is an inference attack in database security?

A. An attack that infers database passwords from usernames
B. An attack where users deduce confidential data from permitted queries by combining and analyzing non-sensitive results
C. An attack on the database inference engine
D. An attack that targets the CPU inference pipeline

Answer: B

Q1485. What is differential privacy in the context of databases?

A. Using different passwords for different databases
B. A mathematical framework that adds noise to query results to prevent identification of individuals while maintaining statistical utility
C. Privacy between different database systems
D. Different levels of privacy for different users

Answer: B

Q1486. What is the polyinstantiation technique in multilevel secure databases?

A. Creating multiple instances of the database server
B. Allowing different versions of the same data at different security levels to prevent inference
C. Installing the database multiple times
D. Creating multiple copies of tables for performance

Answer: B

Q1487. What is homomorphic encryption and its relevance to database security?

A. Encryption that uses similar keys for all data
B. Encryption that allows computations on encrypted data without decrypting it first
C. Encryption that only works on homogeneous data
D. Encryption applied uniformly across all tables

Answer: B

Q1488. What is the Biba model and how does it complement Bell-LaPadula?

A. The Biba model is identical to Bell-LaPadula
B. Biba focuses on integrity (no write up, no read down) while Bell-LaPadula focuses on confidentiality
C. Biba focuses on availability while Bell-LaPadula focuses on integrity
D. They cannot be used together

Answer: B

Q1489. What is k-anonymity and its limitations in database privacy?

A. Anonymizing k users from the database
B. Ensuring each record is indistinguishable from at least k-1 other records on quasi-identifiers, but vulnerable to homogeneity and background knowledge attacks
C. Making data available to k users only
D. A key management protocol for k users

Answer: B

Q1490. What is database activity monitoring (DAM)?

A. Monitoring disk activity for defragmentation
B. Independent monitoring of database activities in real-time to detect and prevent unauthorized access and policy violations
C. Monitoring the dam that stores database water cooling
D. Monitoring CPU activity during queries

Answer: B

Q1491. What is the covert channel problem in multilevel secure databases?

A. A hidden database administrator account
B. An unintended communication path that allows information to leak between security levels through indirect means
C. A covert backup channel
D. A channel used for database replication

Answer: B

Q1492. What is attribute-based access control (ABAC) and how does it extend RBAC?

A. Controlling access based on data attributes only
B. Making access decisions based on attributes of users, resources, actions, and environment, providing finer-grained control than role-based approaches
C. Adding attributes to roles
D. A type of encryption based on attributes

Answer: B

Q1493. What is the challenge of key management in database encryption?

A. Keys are too small to manage
B. Securely storing, rotating, and distributing encryption keys without exposing them while maintaining access to encrypted data
C. Keys take too much storage space
D. There are no challenges with key management

Answer: B

Q1494. What is the three-phase commit (3PC) protocol and what problem does it solve?

A. A three-step backup process
B. An extension of 2PC that adds a pre-commit phase to avoid blocking when the coordinator fails
C. A three-step data migration process
D. A protocol requiring three coordinators

Answer: B

Q1495. What is the Raft consensus algorithm and how does it compare to Paxos?

A. Raft is unrelated to distributed systems
B. Raft achieves similar consensus guarantees as Paxos but is designed to be more understandable with a clear leader-election mechanism
C. Raft is less reliable than Paxos
D. Raft works only in non-distributed settings

Answer: B

Q1496. What is eventual consistency in distributed databases?

A. Consistency that is never achieved
B. A model where all replicas will converge to the same state eventually if no new updates are made, trading immediate consistency for availability
C. Consistent performance across all queries
D. A guarantee that data is always consistent

Answer: B

Q1497. What is vector clock and how is it used in distributed databases?

A. A synchronized physical clock across all sites
B. A logical clock mechanism that tracks causal ordering of events across distributed nodes using a vector of counters
C. A clock that measures query execution time
D. A hardware clock in the database server

Answer: B

Q1498. What is the CALM theorem and its implication for distributed databases?

A. It states that all distributed systems must be calm and stable
B. It states that logically monotonic programs can be made consistent without coordination, guiding coordination-free distributed designs
C. It defines the maximum load for distributed queries
D. It specifies memory allocation for distributed systems

Answer: B

Q1499. What is a conflict-free replicated data type (CRDT)?

A. A data type that prevents all database conflicts
B. A data structure designed so concurrent updates from different replicas can be merged automatically without conflicts
C. A data type that only allows read operations
D. A replicated column type

Answer: B

Q1500. What is the quorum-based protocol in distributed databases?

A. A vote by database administrators
B. A protocol requiring read and write operations to access a minimum number of replicas (quorum) to ensure consistency
C. A protocol for electing a new database coordinator
D. A minimum number of users needed to access the database

Answer: B

Q1501. What is the network partition handling challenge in distributed databases related to the CAP theorem?

A. Managing network cable partitions
B. During a network partition, the system must choose between maintaining consistency (rejecting operations) or availability (allowing operations with potential inconsistency)
C. Partitioning network traffic for different queries
D. Managing database partitions across different networks

Answer: B

Q1502. What is Google Spanner's approach to distributed consistency?

A. It ignores consistency for performance
B. It uses TrueTime API with synchronized atomic clocks to achieve external consistency (linearizability) across globally distributed data
C. It uses eventual consistency only
D. It restricts all data to a single region

Answer: B

Q1503. What is the consistent hashing technique used in NoSQL databases?

A. Using the same hash function across all operations
B. A technique that maps both data and nodes to a ring, minimizing data redistribution when nodes are added or removed
C. Hashing data to ensure consistency
D. A type of encryption for NoSQL data

Answer: B

Q1504. What is the gossip protocol used for in distributed NoSQL systems?

A. Spreading rumors about database performance
B. A protocol where nodes periodically exchange state information with random peers to disseminate cluster membership and health data
C. A chatting protocol between database users
D. A protocol for sharing query results

Answer: B

Q1505. What is the LSM (Log-Structured Merge) tree and why do many NoSQL databases use it?

A. A tree for logging system messages
B. A write-optimized data structure that batches writes in memory and periodically merges sorted runs to disk
C. A tree structure for merge operations only
D. A logging mechanism for SQL databases

Answer: B

Q1506. What is tunable consistency in Cassandra?

A. Consistency that tunes database performance
B. The ability to configure the consistency level per query, choosing between ONE, QUORUM, ALL, and other levels
C. Automatic consistency management
D. A fixed consistency setting

Answer: B

Q1507. What is the concept of anti-entropy in distributed NoSQL databases?

A. A measure of data disorder
B. A mechanism that compares and synchronizes data across replicas to repair inconsistencies
C. The opposite of data entropy
D. A technique to compress data

Answer: B

Q1508. What are tombstones in NoSQL databases like Cassandra?

A. Markers for inactive user accounts
B. Special markers that indicate deleted data, retained temporarily to ensure deletion propagates across all replicas
C. Physical markers on disk sectors
D. A type of index entry

Answer: B

Q1509. What is the difference between CP and AP systems in the context of the CAP theorem?

A. CP systems are faster than AP systems
B. CP systems sacrifice availability during partitions to maintain consistency, while AP systems sacrifice consistency to remain available
C. AP systems always have better performance
D. CP and AP are identical in practice

Answer: B

Q1510. What is a Bloom filter and how is it used in NoSQL databases?

A. A filter for cleaning database data
B. A probabilistic data structure that efficiently tests whether an element might be in a set, used to avoid unnecessary disk reads
C. A filter for bloom-type queries
D. A data quality filter

Answer: B

Q1511. What is the Dynamo architecture model that influenced many NoSQL databases?

A. A model for dynamic query processing
B. A decentralized, highly available key-value architecture using consistent hashing, vector clocks, and sloppy quorum with hinted handoff
C. A centralized database management model
D. A model for database security

Answer: B

Q1512. What is a multi-model database and what advantage does it offer?

A. A database that models multiple data centers
B. A database supporting multiple data models (document, graph, key-value) in a single engine, reducing operational complexity
C. A database with multiple backup models
D. A database that runs multiple instances

Answer: B

Q1513. What is the Kimball methodology for data warehouse design?

A. A methodology focused on building a centralized enterprise data warehouse first
B. A bottom-up approach using dimensional modeling, building data marts first that can be integrated into an enterprise warehouse
C. A methodology that avoids using star schemas
D. A methodology that does not use ETL processes

Answer: B

Q1514. What is the Inmon methodology for data warehouse design?

A. A bottom-up approach starting with data marts
B. A top-down approach building a normalized enterprise data warehouse first, from which data marts are derived
C. A methodology that uses only star schemas
D. A methodology that avoids normalization

Answer: B

Q1515. What is a factless fact table?

A. A fact table with errors
B. A fact table that records events or conditions without any numeric measures, capturing only the relationship between dimensions
C. A table that is not a fact table
D. A fact table with null values

Answer: B

Q1516. What is a junk dimension in data warehouse design?

A. A dimension containing useless data
B. A dimension that combines miscellaneous low-cardinality flags and indicators into a single table
C. A dimension that should be deleted
D. A dimension with corrupted data

Answer: B

Q1517. What is a degenerate dimension?

A. A dimension that has degraded over time
B. A dimension key in the fact table that has no corresponding dimension table, such as an order number
C. A dimension with no attributes
D. A dimension that is no longer used

Answer: B

Q1518. What is change data capture (CDC) and how is it used in data warehousing?

A. Capturing screenshots of data changes
B. A technique that identifies and captures only the data that has changed in source systems since the last extraction for efficient ETL
C. Capturing data from change management systems
D. A method for tracking schema changes

Answer: B

Q1519. What is a data lake and how does it differ from a data warehouse?

A. A data lake is a smaller version of a data warehouse
B. A data lake stores raw, unprocessed data in its native format while a data warehouse stores processed, structured data
C. They are the same concept
D. A data lake only stores water-related data

Answer: B

Q1520. What is the concept of grain in dimensional modeling?

A. The texture of stored data
B. The level of detail represented by each row in a fact table, defining what one row means
C. The granularity of storage hardware
D. The size of each data block

Answer: B

Q1521. What is a periodic snapshot fact table versus a transaction fact table?

A. They store the same data differently
B. A transaction fact table records individual events while a periodic snapshot captures cumulative state at regular intervals
C. Periodic snapshots are always larger
D. Transaction fact tables contain no measures

Answer: B

Q1522. What is an accumulating snapshot fact table?

A. A fact table that accumulates errors
B. A fact table that tracks the lifecycle of a process with multiple date columns representing milestones, updating rows as the process progresses
C. A fact table that only grows and never changes
D. A temporary fact table

Answer: B

Q1523. What is the LEAD() window function used for in SQL?

A. To create a leading index
B. To access a row at a specified physical offset after the current row within the partition
C. To lead data migration
D. To find the first row in a partition

Answer: B

Q1524. What is the saga pattern for long-running transactions?

A. A storytelling approach to database design
B. A pattern that splits a long transaction into a sequence of local transactions, each with a compensating transaction for rollback
C. A type of database narrative log
D. A sequential access pattern for data

Answer: B

Q1525. What is the strict timestamp ordering protocol?

A. Ordering transactions by their creation date
B. A protocol where a transaction that attempts to read or write a data item written by a younger transaction is rolled back and restarted
C. Ordering log entries by time
D. Strictly formatting timestamps in the log

Answer: B

Q1526. What is the fill factor of a B+ tree index and why does it matter?

A. The percentage of the database that is indexed
B. The percentage of space in each index page that is filled with data, affecting split frequency and space utilization
C. The total size of the index file
D. The ratio of index pages to data pages

Answer: B

Q1527. What is a skip list and how is it used as an alternative index structure?

A. A list that skips every other record
B. A probabilistic data structure using multiple linked list levels where higher levels skip over elements for O(log n) average search time
C. A list of skipped transactions
D. An index that skips null values

Answer: B

Q1528. What is the cost model used to evaluate whether an index scan or table scan is more efficient?

A. The cost of purchasing the index software
B. Comparing estimated I/O costs: index scan cost = index traversal I/Os + data page I/Os versus table scan cost = total data pages
C. The monetary cost of running the query
D. The time it takes to create the index

Answer: B

Q1529. What is the Selinger optimization algorithm?

A. An algorithm for selecting database hardware
B. A dynamic programming-based algorithm from IBM System R that finds optimal join orderings by considering access paths and interesting orders
C. A security algorithm for databases
D. An algorithm for selecting storage structures

Answer: B

Q1530. What is the blocking problem in the two-phase commit protocol?

A. Network traffic blocking data transfer
B. If the coordinator fails after sending prepare but before sending the final decision, participants holding locks are blocked indefinitely waiting
C. Users being blocked from accessing the database
D. Queries being blocked by long-running transactions

Answer: B