



### CS6302- Database Management Systems

#### Two Marks

#### UNIT-I

#### INTRODUCTION TO DBMS

##### 1. What is database?

A database is logically coherent collection of data with some inherent meaning, representing some aspect of real world and which is designed, built and populated with data for a specific purpose.

##### 2. Define Database Management System.

A Database management System consists of a collection interrelated data and set of programs to access those data. The collection of data, usually referred to as the database, contains information about one particular enterprises

##### 3. Advantages of DBMS?

- Redundancy is controlled
- Unauthorized access is restricted.
- Providing multiple user interface
- Enforcensing integrity constraints.
- Providing backup and recovery.

##### 4. Disadvantages in File Processing System.

- Data redundancy & inconsistency
- Difficult in accessing data
- Data isolation
- Data integrity
- Concurrent access is not possible.
- Security Problems.

##### 5. Define Data independence.

The ability to modify a schema definition in only level without affecting a Schema definition in the next higher level is called data independence.

##### 6. Define Data Models and list the types of Data Model.

Underlying the structure of database is the data mode: a collection of conceptual tools for describing data, data relationships, data semantics and consistency constraints. The various data model that have been proposed fall in to three different groups: Object-based logical model, record-based logical model and physical model.

##### 7. What is E-R model?

The entity-relationship data model is based on perception of a real world that consists of a collection of basic objects, called entities and of relationships among these object

##### 8. Define entity and entity set.

A entity is a “thing” or “object” in the real world that is distinguisgble from other objects. For



example, each person in an entity. The set of all entities of the same type are termed and entity set.

### 9. What is Weak entity set?

An entity set may not have sufficient attributes to form a primary key, and its primary key compromises of its partial key and primary key of its parent entity, the its is said to be Weak Entity set.

### 10. Define relationship and relationship set.

A relationship is an association among several entities. For example, a depositor relationship associated a customer with each account. The set of all relationships of the same type are termed a relationship set.

### 11. What is object Oriented Model?

The model is based on collection of object. An object contains values stored in instance variables within the object. An object also contains bodies of code that operate on the object. These bodies of code are called methods. Objects that contain same type of values and the same methods are grouped together into classes.

### 12. Define Record-Based Logical Models.

Record-based logical models are used in describing data at the logical and levels. They are used both to specify the overall structure of the database and provide a high-level description of the implementation.

### 13. Define Relational model.

The relational model uses a collection of tables to represent both data and the relationships among those data. Each table has multiple columns, and each column has a unique name.

### 14. Define Network model.

Data in the network model are represented by collection of records and relationships among data are represented by links, which can be viewed as pointers. The records in the database are organized as collections of arbitrary graphs.

### 15. Define Hierarchical Model

The hierarchical model is similar to the network models in the sense that data and relationship among data are represented by records and links respectively. It differs from the network model in that the records are organized as collection of trees rather than arbitrary graphs.

### 16. List the role of DBA.

The person who has central control over the system is called database administrator. The function of the DBA includes the following:

Schema definition

Storage structure and access-method definition

Schema and physical-organization modification

Granting of authorization for data access

Integrity-constraint specification

### 17. List the different type of database system user.

There are four different types of database-system users, differentiated by the way they expect to interact with the system.



# SYED AMMAL ENGINEERING COLLEGE

(An ISO 9001: 2008 certified Institution)

Dr. E.M.Abdullah Campus, Ramanathapuram – 623 502.

## Department of Computer Science and Engineering



- Application programmers
- Sophisticated Users
- Specialized users
- Naïve users.

### 18. Write about the role of Transaction manager.

Tm is responsible for ensuring that the database remains in a consistent state despite system failures. The TM also ensures that concurrent transaction executions proceed without conflicting.

### 19. Write about role of storage manager.

A SM is a program module that provides the interface between the low-level data stored in the database and the application programs and queries submitted to the system. The SM is responsible for interaction with the data stored on disk.

### 20. Define attributes.

Entities are described in a database by a set of attributes. For example, the attributes account-number and balance describes one particular account in a bank.

### 21. Define mapping constraints.

An E-R enterprise schema may define certain constraints to which the contents of a database must conform. Two of the most important types of constraints are mapping cardinalities: express the number of entities to which another entity can be associated via relationship set.

### 22. Define Relational algebra.

A general expression in the relational algebra is constructed out of smaller sub expressions.

### 23. Define Relational calculus.

A tuple relational calculus expression is of the form  $\{t/P(t)\}$  where P is a formula. Several tuple variables may appear in a formula.

### 24. List possible operations in Relation algebra

Select, project, theta join, equijoin, union, intersection, difference, Cartesian product, division.

### 25. Is it possible for several attributes to have the same domain? Illustrate your answer with suitable example.

Yes, several attributes can have the same domain. Attributes: name, address, belong to the same domain contains all text strings of certain length.

### 26. What are the problems in data redundancy?

Multiple updates, storage space is wasted

## UNIT II

### SQL & QUERY OPTIMIZATION

#### 1. Define Aggregate Functions in SQL?

Aggregate functions are functions that take a collection of values as input and return a single value. SQL offers five built-in aggregate functions:

- Average: avg
- Minimums: min



- Maximum : max
- Total: sum
- Count: count

### 2. Define Nested Subqueries.

SQL provides a mechanism for the nesting of sub queries. A sub query is a select-from-where expression that is nested within another query. A common use of sub queries is to perform tests for set membership, set comparison, and set cardinality.

### 3. Define Integrity constraints.

Integrity constraints provide a means of ensuring that changes made to the database by authorized user do not result in a loss of data consistency. Thus Integrity constraints guard against accidental damage to the database.

### 4. Define referential Integrity.

Often, we wish to ensure that a value that appears in one relation for a given set of attributes also appears for a certain set of attributes in another relation. This condition is called referential integrity.

### 5. Define Assertions.

An assertion is a predicate expressing a condition that we wish the database always satisfied. E.g. create assertion <assertion-name> check <predicate>

### 6. Define Triggers.

A trigger is a statement that is executed automatically by the system as a side effect of a modification to the database. To design a trigger mechanism, we must meet two requirements:

1. Specify the conditions under which the trigger is to be executed.
2. Specify the actions to be taken when the trigger executed.

### 7. Define Functional Dependency

- Constraints on the set of legal relations.
- Require that the value for a certain set of attributes determines uniquely the value for another set of attributes.
- A functional dependency is a generalization of the notion of a *key*.

### 8. List the pitfalls in Relational Database Design.

1. Repletion of information
2. Inability to represent certain information.

### 9. Define normalization?

It is a process of analyzing the given relation schemas based on their functional Dependencies(FDs) and primary key to achieve the properties.

- Minimizing redundancy
- Minimizing insertion, deletion and update anomalies.

### 10. List the properties of decomposition.

1. Lossless join
2. Dependency preservation
3. No repletion of information



### 11. Define 1NF?

A relation schema R is in 1NF if the domains of all attributes of R are atomic

### 12. Define 2NF?

A relation schema R is in 2NF if every non prime attribute A in R is fully functionally dependent on the primary key of R.

### 13. Define 3NF?

A relation schema R is in third normal form if, whenever a nontrivial functional dependency  $X \rightarrow A$  holds in R, either (a) X is super key of R or (b) A is a prime attribute of R.

### 14. What is BCNF?

A relation schema R is in BCNF if it is in 3NF and satisfies additional constraints that for every FE  $X \rightarrow A$ , X must be a candidate key.

### 15. What is query?

A query with respect to DBMS relates to user commands that are used to interact with a database. The query language can be classified into data definition language and data manipulation language.

### 16. What do you mean by Correlated subquery?

Subqueries, or nested queries, are used to bring back a set of rows to be used by the parent query. Depending on how the subquery is written, it can be executed once for the parent query or it can be executed once for each row returned by the parent query. If the subquery is executed for each row of the parent, this is called correlated subquery.

### 17. Define SQL and state the difference between SQL and other conventional programming Language.

SQL is a nonprocedural language that is designed specifically for data access operations on normalized relational database structures. The primary difference between SQL and other conventional programming languages is that SQL statements Specify what data operations should be performed rather than how to perform them.

### 18. What is database trigger is a PL/SQL block that can defined to automatically

execute for insert, update and deleted statements against a table .The trigger can be defied to execute once for the entire statement or once for every row that is inserted, updated, or deleted. For any one table, there are twelve events for which you can define database triggers.

### 19. What are Armstrong rules?

Reflexive rule: If Y is subset or equal to X the  $X \rightarrow Y$

Augmentation rule: If  $X \rightarrow Y$  then  $XZ \rightarrow YZ$ .

Transitive Rule: If  $\{x \rightarrow Y, Y \rightarrow z\}$  then  $X \rightarrow Z$

The above three are known as Armstrong Rule

### 20. What are the privileges that can be granted on a table by a user to others?

Insert, Update, delete, select, references, index, execute, alter, all

### 21. What is the difference between TRUNCATE and DELETE command?

TRUNCATE is a DDL command whereas DELECTE is a DML command. Hence DELETE operation can be rolled back, but TRUNCATE operation cannot be rolled back. WHERE clause can be used with DELETE and not with TRUNCATE.



### 22. What is the use of CASCADE CONSTRAINTS?

When this clause is used with the DROP command, a parent table can be dropped even when a child table exists.

### 23. Is it possible for several attributes to have the same domain? Illustrate your answer with suitable examples.

Possibility for several attributes to have same domain. Yes, several attributes to have same domain. For example Attributes: name, address, belong to same domains contains all text string of certain length.

### 24. List out the field level constraints that can be associated with relational table.

Unique, Check, Not null, primary key, foreign key

### 25. Justify the need for normalization.

Need for normalization-To avoid insertion, deletion and updating anomalies. To ensure that there is no redundant data. To ascertain that the database design is perfect.

### 26. Name the different type of joins supported in SQL.

Equi join, Natural join, self join, and outer join.

## UNIT – III

### TRANSACTION AND CONCURRENCY CONTROL

#### 1. What is transaction?

A transaction is collection of operations that form logical unit of work .A database system must ensure proper execution of transaction despite failures either the entire transaction executes or none of it does.

#### 2. Define atomicity?

Either all operations of the transaction are reflected properly in the database or none are.

#### 3. Define Durability?

A transaction completes successfully then changes I t has made to the database persist even if there are system failure.

#### 4. What is transaction-management component?

Ensuring atomicity is the responsibility of the database system itself specifically, it is handled by a component called the transaction-management component.

#### 5. List the properties of traction.

1. Atomicity 2. consistency 3. Isolation 4. Durability.

#### 6. Define concurrency control?

Serializability of schedules generated by concurrently executing transactions can be ensure through one of a variety of mechanisms called concurrency control.

#### 7. What are the tow operation access data in transaction?

Read(x)- transfer data item x from database.

Write(x)- transfer data item x from the local buffer.



8. What do you mean by read only transaction?

The data base operation in a transaction do not update the database but only retrieve data, the transaction is called a read-only transaction.

9. What are the steps followed in Executing read(x)command in transaction?

1. Find the address of the disk block that contains item x.
2. Copy that disk block in to a buffer in main memory.
3. Copy item x from the buffer to the program variable named x.

10. What are the steps followed in executing write(x) command in transaction?

1. find the address of disk block that contain item x.
2. Copy that disk block into buffer in main memory.
3. Copy item x from the program variable named x into its correct location in the buffer
4. Store the update block from the buffer back to disk.

11. List out the transaction states?

- i) active ii) Partially Committed iii) Failed iv) Abort v) committed.

12. What are the need for concurrency?

- i) Improved throughput and resource utilization
- ii) Reduced waiting time.

13. Define Schedule.

The chromosomal order in which instructions are executed in the system.

14. When the schedule is called serial?

Each serial schedule consists of a sequence of instruction from various transaction where the instruction belonging to one single transaction appear together in that schedule.

15. When two operation in schedule are said to be conflict?

- i) Two operation belong to different transaction
- ii) Two operation access the same item x
- iii) At least one of the operation is write-item(x)

16. Define cascading rollback?

An uncommitted transaction has to be roll back because it read an item from a transaction that failure.

17. When the schedule is said to cascade less?

A schedule is said to be cascadeless or avoid cascading roll back if every transaction in the schedule reads only items that were written by committed transactions.

18. Define serializability.

Any schedule produced by concurrent processing of set of transaction will have an effect equivalent to a schedule produce when these transaction are rule serially is some order of guarantees this called serializability.

19. When the serializability is equivalent?



# SYED AMMAL ENGINEERING COLLEGE

(An ISO 9001: 2008 certified Institution)

Dr. E.M.Abdullah Campus, Ramanathapuram – 623 502.

## Department of Computer Science and Engineering



Transaction is serializable if it is equivalent to some serial schedule of the same  $n$  transaction.

20. When two schedule are called result equivalent?

Two schedule are called result equivalent if they produce the same final state of the database.

21. How to find the schedule is conflict serialization or non using procedure graph?

The graph has a cycle then schedules not conflict serialization if the graph contain no cycle the schedule is conflict serialization.

22. Define Concurrency control?

The system to control the interacting among the concurrency transactions and this control is a achieved through one of the varying of mechanism called concurrency control.

23. Define lock?

Lock is variable associated with a data item. Lock are used as a means of synchronizing the access by concurrent transaction to the database item.

24. Define lock table?

System maintain record for the items that are currently locked in lock table that which could be organized as a hash file.

25. What you mean by lock conversion?

A transaction that already holds a lock on item  $x$  is allowed under certain condition to apply the Lock from one locked state to another.

26. List out the two-phase locking?

- 1) Growing phase: A transaction may obtain locks but may not replace any lock.
- 2) Shrinking phase: A transition may release lock but may not obtain any new locks.

27. When the system is said to deadlock?

A system is in a deadlock state if there exists a set of transactions such that every transition in the set is waiting for another transition in the set.

28. What are the methods in deadlock?

- (i) Deadlock prevention (ii) Deadlock detection and deadlock recovery

29. Define timestamps?

Timestamp are typically based on the order in which transition are started.

30. What are the different methods for deadlock prevention?

- i) Wait-die
- ii) Would-wait

31. How detect the deadlock?

A deadlock exists in the system if and only if the wait-for graph contains a cycle.

32. What are the methods to follow the recovery from deadlock?

1. Selection of a victim
2. Roll back.
3. Starvation



33. List out type of failure?

1) Transaction failure 2) System crash 3) Disk failure

34. List out the draw backs of shadow-page?

1) Commit overhead 2) Data fragmentation 3) Garbage collection

### UNIT – IV

#### TRENDS IN DATABASE TECHNOLOGY

##### 1. Explain Optical Storage Device?

The most popular form of optical storage is the compact disk read-only memory, can be read by a laser. Optical storage is the write-once, read-many disk, which allows data to be written once, but does not allow them to be erased and rewritten.

##### 2. Define cache.

The cache is the fastest and most costly form of storage. Cache memory is small; its use is managed by the operating system.

##### 3. Define disk controller?

It is an interface between the computer system and the actual hardware of the disk drive. Accept high-level command to read or write a sector. It attaches checksums to each sector that is written. It also performs remapping of bad sectors.

##### 4. Define RAID.

It is collectively called redundant arrays of inexpensive disk, have been proposed to address the performance and reliability issues. RAID are used for their higher reliability and higher data transfer rate. RAID stands of independent, instead of inexpensive.

##### 5. Define file organization.

A file is organized logically as a sequence of records. These records are mapped on to disk blocks. File are the basic construct in operating system.

##### 6. Define Hash indices?

Indices are based on the values being distributed uniformly across a range of buckets. The bucket to which a value is assigned is determined by a function called hash function.

##### 7. Define dense index?

An index record appears for every search-key value in the file. The index record contains the search-key value and pointer to the first data record with that search key value.

##### 8. Define sparse index?

An index record is created for only some of the clause. Each index record contains a search-key value and a pointer to the first data record with that search-key value. To locate a record we find the index entry with the largest search-key value that is less than or equal to the search-key value.

##### 9. Differentiate between dense index and sparse index.

Dense index-An index record appears for every search-key value in the file. Sparse index-An index record appears for only some of the search-key values.

##### 10. Explain B+ tree index structure?



## Department of Computer Science and Engineering

The b+ tree index structure is the most widely used of several index Structures that maintain their efficiency despite insertion and deletion of data. A B+ tree index takes the form of a balanced tree in which ever path from the root of the tree to a leaf of the tree is the same length.

### 11. Define Static Hashing?

File organization based on the technique of hashing allow us to avoid accessing an index structure. Hashing also provides a way of constructing indices.

### 12. Define Query processing?

Query processing refers to the range of activities involved in extracting data form a database. These activities include translation of queries expressed in high-level database language into expression that cab be implemented at the physical level of the file system.

### 13. What are the steps involved in query processing?

1. Parsing and translation
2. Optimization
3. Evaluation

### 14. What is indexing and What are the different kinds of indexing?

Indexing is a technique for determining how quickly specific data can be found. Types:

- Binary search style indexing
- B-Tree indexing
- Inverted list indexing
- Memory resident table
- Table indexing

### 15. What is meant by query optimization?

The phase that identifies an efficient execution plan for evaluating a query that has the least estimated cost is referred to as query optimization.

### 16. What is Buffer Manager?

It is program module, which is responsible for fetching data from disk storage into main memory and deciding what data to be cache in memory.

### 17. What is hashing?

The conversion of a column's primary key value to a database page number on which the row will be stored. Retrieval operations that specify the key column value use the same hashing algorithm and can locate the row directly. Hashing provide fat retrieval for data that contains a unique key value.

### 18. List out disadvantage of Hashing?

Lacks locality and sequential Retrieval by key Inability to use Duplicate keys Pre-allocation of Space is not possible. Complexity

### 19. Define data striping?

Data striping is used to utilized parallelism to improve disk reliability and performance. Data striping distributes data transparently over multiple disks to make them appear as a single large, fast disk.



# SYED AMMAL ENGINEERING COLLEGE

(An ISO 9001: 2008 certified Institution)

Dr. E.M.Abdullah Campus, Ramanathapuram – 623 502.



## Department of Computer Science and Engineering

### 20. Define double buffering?

The CPU can start processing a block once a transfer to main memory is completed, at the same time the disk I/O processor can be reading and transferring the next block into a different buffer. This technique is called double buffering.

### 21. What is file descriptor?

A file descriptor (or file header) includes information that describes the file, such as the field names and their data types and the addresses of the file blocks on disk

### 22. List out the operation of files?

OPEN, FIND, FINDNEXT, READ, INSEART, DELETE, MODIFY, CLOSE, REORGANIZE

### 23. What is the purpose of buffer manager?

To control data movement from main memory and secondary memory.

### 24. What are the cost components for query execution?

Access cost to secondary storage, storage cost, memory usage cost, computation cost, communication cost.

### 25. What are the properties of B trees?

Balance tree. Inner nodes having pointer to its children leaf node having pointer to actual data order of the tree is  $n$  then each node we can have  $n-1$  elements.

### 26. How does pipelining improve query-evaluation efficiency? Explain.

Pipeline eliminates the cost of reading and writing temporary relation.

## UNIT V ADVANCED TOPICS

### 1. What is Data warehousing?

**Data warehousing:** It is the process that is used to integrate and combine data from multiple sources and format into a single unified schema. So it provides the enterprise with a storage mechanism for its huge amount of data.

### 2. Define Integrity.

Integrity refers to the process of ensuring that a database remains an accurate reflection of the universe of discourse it is modeling or representing. In other words there is a close correspondence between the facts stored in the database and the real world it models.

### 3. What are Deductive Databases?

A **deductive database** system specifies rules through a **declarative language** - a language in which we specify what to achieve rather than how to achieve it. An **inference engine** within the system can deduce new facts from the database by interpreting these rules. The model used for deductive databases is related to the relational data model and also related to the field of **logic programming** and the **Prolog** language.

### 4. Mention two features of parallel Databases.

It is used to provide speedup, where queries are executed faster because more resources, such as processors and disks, are provided. It is also used to provide scale up, where increasing workloads are handled without increased response time, via an increase in the degree of parallelism.



# SYED AMMAL ENGINEERING COLLEGE

(An ISO 9001: 2008 certified Institution)

Dr. E.M.Abdullah Campus, Ramanathapuram – 623 502.



## Department of Computer Science and Engineering

### 5. Mention two features of Multimedia databases.

- The multimedia database systems are to be used when it is required to administrate a huge amount of multimedia data objects of different types of data media (optical storage, video, tapes, audio records, etc.) so that they can be used (that is, efficiently accessed and searched) for as many applications as needed.
- The Objects of Multimedia Data are: text, images, graphics, sound recordings, video recordings, signals, etc., that are digitalized and stored.

### 6. Define concurrency control.

Concurrency control is the activity of coordinating concurrent accesses to a database in a multiuser system. Concurrency control allows user to access a database in a multi-programmed fashion while preserving the consistency of the data.

### 7. What is Persistence?

Persistence is the property of an object through which its existence transcends time i.e. (the object continues to exist after its creator ceases to exist), and/or space (i.e. the object's location moves from the address space in which it was created).