UNIT I RELATIONAL DATABASES

Purpose of Database System – Views of data – Data Models – Database System Architecture – Introduction to relational databases – Relational Model – Keys – Relational Algebra – SQL fundamentals – Advanced SQL features – Embedded SQL– Dynamic SQL

UNIT-I / PART-A

1.	**Define database management system?** Database management system (DBMS) is a collection of interrelated data and a set of programs to access those data.
2.	**List any five applications of DBMS.** Banking, Airlines, Universities, Credit card transactions, Tele communication, Finance, Sales, Manufacturing, Human resources.
3.	**What is the purpose of Database Management System? *(Nov/Dec 14)*** Data redundancy and inconsistency , Difficulty in accessing data, Data isolation, Integrity problems, Atomicity problems and Concurrent access anomalies
4.	**Define instance and schema?** **Instance:** Collection of data stored in the data base at a particular moment is called an Instance of the database **Schema:** The overall design of the data base is called the data base schema.
5.	**Define the terms 1) physical schema 2) logical schema.** **Physical schema:** The physical schema describes the database design at the physical level, which is the lowest level of abstraction describing how the data are actually stored. **Logical schema:** The logical schema describes the database design at the logical level, which describes what data are stored in the database and what relationship exists among the data.
6.	**What is a data model? List the types of data models used?** A data model is a collection of conceptual tools for describing data, data relationships, data semantics and consistency constraints.
7.	**Define- relational algebra.** The relational algebra is a procedural query language. It consists of a set of operations that take one or two relation as input and produce a new relation as output.
8.	**What is a data dictionary?** A data dictionary is a data structure which stores meta data about the structure of the database i.e. the schema of the database.
9.	**List out the operations of the relational algebra** The Six basic operators Select, project, union, set difference, Cartesian product and Rename.
10.	**Define relational data model** Relational model use a collection of tables to represent both data and the relationships among those data. Each table has a multiple columns and each column has unique name.
11.	**Explain Semi structured data model** ✓ Specification of data where individual data item of same type may have different sets of attributes ✓ Sometimes called schema less or self-describing ✓ XML is widely used to represent this data model
12.	**Define Object based data model** Object based data model can be seen as extension of the E-R model with notion of encapsulation, methods and object identify.
13.	**Explain Hierarchical data model** ✓ The Hierarchical data model organizes data in a tree structure. There is hierarchy of parent and child data segments. ✓ This model uses parent child relationship. ✓ 1:M Mapping between record type

Arunai Engineering College

14. **Define Network Model**
- ✓ Some data were more naturally modeled with more than one parent per child.
- ✓ This model permitted the modeling of M:N relationship

15. **Write the characteristics that distinguish the Database approach with the File-based approach.** ***(Apr/May 15)(Nov/Dec 16)***

File-based System.

1. Separation and isolation of data
2. Duplication of data
3. Incompatible file formats
4. Data dependence

1. Control of data redundancy
2. Data consistency
3. Sharing of data
4. Improved data integrity
5. Improved security

16. **What are the disadvantages of file processing system?*(May/June 16)***

The file processing system has the following major disadvantages:
- ✓ Data redundancy and inconsistency ✓
- Integrity Problems
- ✓ Security Problems
- ✓ Difficulty in accessing data ✓
- Data isolation.

17. **Define query language?**

A query is a statement requesting the retrieval of information. The portion of DML that involves information retrieval is called a query language.

18. **List the string operations supported by SQL?**

1) Pattern matching Operation
2) Concatenation
3) Extracting character strings
4) Converting between uppercase and lower case letters.

19. **List out some date functions.**
- ✓ To_date
- ✓ To_char(sysdate,'fmt')
- ✓ d,dd,ddd,mon,dy,day,y,yy,yyy,yyyy,year,month,mm

20. **What is the use of sub queries?**

A sub query is a select-from-where expression that is nested with in another query. A common use of sub queries is to perform tests for set membership, make set comparisons, and determine set cardinality.

21. **Name the categories of SQL command? *(May/June 16)***

SQL commands are divided in to the following categories:
1. Data - definition language
2. Data manipulation language
3. Data Query language
4. Data control language
5. Data administration statements
6. Transaction control statements

22. **List the SQL domain Types?**

SQL supports the following domain types.
Char (n) , varchar (n), int , numeric (p,d) , float(n) , date.

Anna University Engineering College

23. **What are aggregate functions? And list the aggregate functions supported by SQL?**
Aggregate functions are functions that take a collection of values as input and return a single value. Aggregate functions supported by SQL are
- ✓ Average: avg
- ✓ Minimum: min
- ✓ Maximum: max
- ✓ Total: sum Count: count

24. **What is the difference between char and varchar2 data type?**
- ✓ Char and varchar2 are data types which are used to store character values.
- ✓ Char is static memory allocation; varchar2 is dynamic memory allocation.

25. **How to add primary key to a table with suitable query?**
Alter table <table name> add primary key(column);

26. **Differentiate static and dynamic SQL.** ***(Nov/Dec 14,15,16) (Apr/May 15)***

Static SQL	Static SQL
The SQL statements do not change each time the program is run is called Static SQL.	The SQL statements do not change each time the program is run is called Static SQL.
Static SQL is compiled and optimized prior to its execution	Static SQL is compiled and optimized prior to its execution
The statement is prepared before the program is executed and the operational form of the statement persists beyond the execution of the program.	The statement is prepared before the program is executed and the operational form of the statement persists beyond the execution of the program.

27. **Why does SQL allow duplicate tuples in a table or in a query result?** ***(Nov/Dec 15)***
If key constraint is not set on a relation every result in a relation will be considered as a tuple and hence SQL allows duplicate tuples in a table. Distinct keyword is used to avoid duplicate tuples in the result.

28. **Define: DDL, DML, DCL and TCL.** ***(Nov/Dec 14,16)(Apr/May 15)***

- Create
- Alter
  - ✓ Add
  - ✓ Modify
  - ✓ Drop
- Rename
- Drop

DML Commands:
- Insert
- Select
- Update
- Delete

DCL commands
- Grant - Provide access privilege to user
- Revoke - Get back access privilege from user

TCL commands
- Commit
- Rollback
- Save point

29. **What is the use of Union and intersection operation?**
Union: The result of this operation includes all tuples that are either in r1 or in r2 or in both r1 and r2.Duplicate tuples are automatically eliminated.
Intersection: The result of this relation includes all tuples that are in both r1 and r2.

Arunai Engineering College

30. **What is embedded SQL? What are its advantages?**
The SQL standard defines embedded of SQL in a variety of programming languages such as C, Java, and Cobol. A language to which SQL queries are embedded is referred to as a host language, and the SQL structures permitted in the host language comprise embedded SQL.
The basic form of these languages follows that of the System R embedding of SQL into PL/I. EXEC SQL statement is used to identify embedded SQL request to the preprocessor EXEC SQL <embedded SQL statement > END_EXEC

UNIT-I / PART-B

1. Explain the purpose and components of DBMS in detail.
2. List out the disadvantages of File system over DB & explain it in detail.
3. List out the operations of the relational algebra and explain with suitable examples.***(Nov/Dec 16)***
4. i) With the help of a neat block diagram explain the basic architecture of a database management system***.(Nov/Dec 15)(May/June 16)***
ii) What are the advantages of having a centralized control of data? Illustrate your answer with suitable example. ***(Nov/Dec 15)***
5. Briefly explain about views of data.***(May/June 16)***
6. Discuss about (i) Data Models (ii) Mapping cardinalities. ***(Nov/Dec 14)***
7. Explain about data definition language and data manipulation language in SQL with examples.***(Nov/Dec 14)(May/June 16)***
9. Explain about data control language and TCL in SQL with examples.
10. Design an employee detail relation and explain referential integrity using SQL queries.
11. Consider a student registration database comprising of the below given table schema. Student File

Student Number	Student Name	Address	Telephone

Course File

Course Number	Description	Hours	Professor Number

Professor File

Professor Number	Name	Office

Registration file

Student Number	Course Number	Date

Consider a suitable sample of tuples/records for the above mentioned tables and write DML statements (SQL) to answer for the queries listed below.
1. Which courses does a specific professor teach?
2. What courses does specific professors?
3. Who teaches a specific course and where is his/her office?
4. For a specific student number, in which courses is the student registered and what is his/her name?
5. Who are the professors for a specific student?
6. Who is the student registered in a specific course? ***(Apr/May 15)***

12. Explain about SQL Fundamentals.***(May/June 16)***
13. Describe the six clauses in the syntax of an SQL query, and show what type of constructs can be specified in each of the six clauses. Which of the six clauses are required and which are optional? ***(Nov/Dec 15)***

Arunai Engineering College

14.	Assume the following table. Degree(degcode, name, subject) Candidate(seatno, degcode, semester, month, year, result) Marks(seatno, degcode, semester, month, year, papcode, marks) Degcode-degree code, Name-name of the degree (MSc, MCOM) Subject-subject of the course. E.g. Phy, Papcode- Paper code E.g. A1 Solve the following queries using SQL: (i) Write a SELECT statement to display all the degree codes which are there in the candidate table but not present in degree table in the order of degcode. (ii) Write a SELECT statement to display the name of all the candidates who have got less than 40 marks in exactly 2 subjects. (iii)Write a SELECT statement to display the name, subject and number of candidates for all degrees in which there are less than 5 candidates. (iv) Write a SELECT statement to display the names of all the candidates who have got highest total marks in MSc.,(Maths) ***(Nov/Dec 15)***

UNIT II DATABASE DESIGN

Entity-Relationship model – E-R Diagrams – Enhanced-ER Model – ER-to-Relational Mapping – Functional Dependencies – Non-loss Decomposition – First, Second, Third Normal Forms, Dependency Preservation – Boyce/Codd Normal Form – Multi-valued Dependencies and Fourth Normal Form – Join Dependencies and Fifth Normal Form

UNIT-II / PART-A

1.	**Explain entity relationship model?*(May/June 16)*** The entity relationship model is a collection of basic objects called entities and relationship among those objects. An entity is a thing or object in the real world that is distinguishable from other objects.
2.	**What is relationship? Give examples** A relationship is an association among several entities. Example: A depositor relationship associates a customer with each account that he/she has.
3.	**What are stored and derived attributes?** Stored attributes: The attributes stored in a data base are called stored attributes. Derived attributes: The attributes that are derived from the stored attributes are called derived attributes
4.	**What are composite attributes?** Composite attributes can be divided in to sub parts. The degree of relationship type is the number of participating entity types.
5.	**What is a weak entity? Give example. *(Nov/Dec 16)*** It is an entity that cannot be identified uniquely without considering some primary key attributes of another identifying owner entity. An example is including Dependent information for employees for insurance purposes.
6.	**What are attributes? Give examples.** An entity is represented by a set of attributes. Attributes are descriptive properties possessed by each member of an entity set. Example: possible attributes of customer entity are customer name, customer id, Customer Street, customer city.
7.	**Mention the 2 forms of integrity constraints in ER model?** ✓ Key declarations ✓ Form of a relationship

Arunai Engineering College

8.	**What is the use of integrity constraints?** Integrity constraints ensure that changes made to the database by authorized users do not result in a loss of data consistency. Thus integrity constraints guard against accidental damage to the database
9.	**List some security violations (or) name any forms of malicious access.** 1) Unauthorized reading of data 2) Unauthorized modification of data 3) Unauthorized destruction of data.
10.	**What is a primary key?** Primary key is a set of one or more attributes that can uniquely identify record from the relation; it will not accept null values and redundant values. A relation can have only one primary key.
11.	**What is called query processing?** Query processing refers to the range of activities involved in extracting data from a database.
12.	**What is called a query evaluation plan?** A sequence of primitive operations that can be used to evaluate be query is a query evaluation plan or a query execution plan.
13.	**What is called as an N-way merge?** The merge operation is a generalization of the two-way merge used by the standard in-memory sort-merge algorithm. It merges N runs, so it is called an N-way merge.
14.	**What is a super key?** A super key is a set of one or more attributes that collectively allows us to identify uniquely an entity in the entity set.
15.	**What is foreign key?** A relation schema r1 derived from an ER schema may include among its attributes the primary key of another relation schema r2.this attribute is called a foreign key from r1 referencing r2.
16.	**What is the difference between unique and primary key?** Unique and primary key are keys which are used to uniquely identify record from the relation. But unique key accepts null values; primary key does not accept null values.
17.	**What does the cardinality ratio specify?** Mapping cardinalities or cardinality ratios express the number of entities to which another entity can be associated. Mapping cardinalities must be one of the following: One to one, One to many, Many to one and Many to many.
18.	**Explain the two types of participation constraint.** ✓ Total: The participation of an entity set E in a relationship set R is said to be total if every entity in E participates in at least one relationship in R. ✓ Partial: if only some entities in E participate in relationships in R, the participation of entity set E in relationship R is said to be partial.
19.	**Define Tuple variable?** Tuple variables are used for comparing two tuples in the same relation. The tuple variables are defined in the from clause by way of the as clause.
20.	**What is first normal form?** The domain of attribute must include only atomic (simple, indivisible) values.
21.	**What is 2NF?** Relation schema R is in 2NF if it is in 1NF and every non-prime attribute An in R is fully functionally dependent on primary key.
22.	**What is meant by domain key normal form?** Domain/key normal form (DKNF) is a normal form used in database normalization which requires that the database contains no constraints other than domain constraints and key constraints.

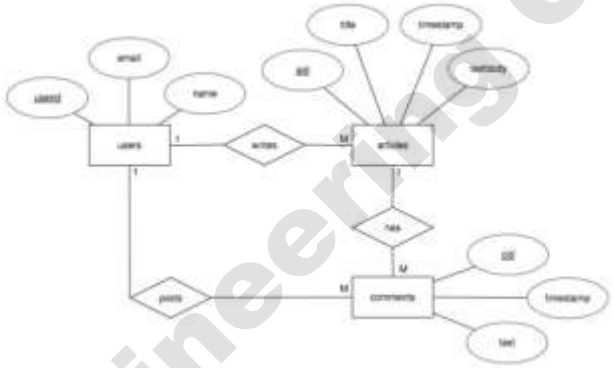
Arunai Engineering College

23.	**Define Functional dependency. *(Apr/May 15)*** In a given relation R, X and Y are attributes. **Attribute Y is functionally dependent on attribute X if each value of X determines EXACTLY ONE value of Y**, which is represented as X → Y (X can be composite in nature). We say here "x determines y" *or "y is functionally dependent on x" Empid → Ename*
24.	**Define full functional dependency.** The removal of any attribute A from X means that the dependency does not hold any more.
25.	**Explain about partial functional dependency?** X and Y are attributes. Attribute Y is partially dependent on the attribute X only if it is dependent on a sub-set of attribute X.
26.	**What you meant by transitive functional dependency?** **Transitive dependency** is a functional dependency which holds by virtue of transitivity. A transitive dependency can occur only in a relation that has three or more attributes. Let A, B, and C designates three distinct attributes (or distinct collections of attributes) in the relation. Suppose all three of the following conditions hold: 1.A $\rightarrow$ B 2. It is not the case that B $\rightarrow$ A 3. B $\rightarrow$ C Then the functional dependency A $\rightarrow$ C (Which follows from 1 and 3 by the axiom of transitivity) is a transitive dependency.

UNIT-II / PART-B

1.	ER model by taking Hospital management/Banking System/University Database as case study***(Nov/Dec 14)***
2.	Explain the various components of ER diagram with examples.
3.	Discuss about (i) Data Models (ii) Mapping cardinalities. ***(Nov/Dec 14)***
4.	Explain functional dependency in database design with its properties.
5.	Design an E-R diagram for keeping track of the exploits of your favourite sports team. You should store the matches played, the scores in each match, the players in each match and individual player statistics for each match. Summary statistics should be modelled as derived attributes.
6.	Construct an E-R diagram for a car insurance company whose customers own one or more cars each. Each car has associated with it zero to any number of recorded accidents. Each insurance policy covers one or more cars, and has one or more premium payments associated with it. Each payment is for a particular period of time set of customers, and the date when the payment was received. ***(Nov/Dec 16)***
7.	A car rental company maintains a database for all vehicles in its current fleet. For all vehicles, it includes the vehicle identification number, license number, manufacturer, model, date of purchase, and color. Special data are included for certain types of vehicles. • Trucks: cargo capacity. • Sports cars: horsepower, renter age requirement. • Vans: number of passengers. • Off-road vehicles: ground clearance, drivetrain (four- or two-wheel drive). Construct an ER model for the car rental company database. ***(Nov/Dec 15)***

Arunai Engineering College

8.	State the need for Normalization of a Database and Explain the various Normal Forms (1st, 2nd, 3rd, BCNF, 4th, 5th and Domain- Key) with suitable examples. ***(Apr/May 15)(Nov/Dec 14,16)***
9.	Draw E – R Diagram for the "Restaurant Menu Ordering System", which will facilitate the food items ordering and services within a restaurant. The entire restaurant scenario is detailed as follows. The Customer is able to view the food items menu, call the waiter, place orders and obtain the final bill through the computer kept in their table. The waiters through their wireless tablet PC are able to initialize a table for customers, control the table functions to assist customers, orders, send orders to food preparation staff (chef) and finalize the customer's bill. The food preparation staffs (Chefs), with their touch-display interface to the system, are able to view orders sent to the kitchen by waiters. During preparation, they are able to let the waiter know the status of each item, and can send notification when items are completed. The system should have full accountability and logging facilities, and should support supervisor actions to account for exceptional circumstances, such as a meal being refunded or walked out on. ***(Apr/May 15)***
10.	For the ER diagram given below explain ER to Relational mapping procedures. 

UNIT III TRANSACTIONS

Transaction Concepts – ACID Properties – Schedules – Serializability – Concurrency Control – Need for Concurrency – Locking Protocols – Two Phase Locking – Deadlock – Transaction Recovery – Save Points – Isolation Levels – SQL Facilities for Concurrency and Recovery.

UNIT-III/ PART-A

1.	**Give the reasons for allowing concurrency?** The reasons for allowing concurrency is if the transactions run serially, a short transaction may have to wait for a preceding long transaction to complete, which can lead to unpredictable delays in running a transaction. So concurrent execution reduces the unpredictable delays in running transactions.
2.	**What is average response time?** The average response time is that the average time for a transaction to be completed after it has been submitted.
3.	**What are the two types of serializability?** The two types of serializability is Conflict serializability, View serializability.
4.	**Differentiate strict two phase locking protocol and rigorous two phase locking protocol.*(May/June 16)*** ✓ In strict two phases locking protocol all exclusive mode locks taken by a transaction is held until that transaction commits. ✓ Rigorous two phase locking protocol requires that all locks be held until the Transaction commits.

Arunai Engineering College

5. **How the time stamps are implemented**
   - ✓ Use the value of the system clock as the time stamp. That is a transaction's time stamp is equal to the value of the clock when the transaction enters the system.
   - ✓ Use a logical counter that is incremented after a new timestamp has been assigned; that is the time stamp is equal to the value of the counter.
6. **What are the different modes of lock?** The modes of lock are:
   - ✓ Shared
   - ✓ Exclusive
7. **What are the time stamps associated with each data item?**
   - ✓ W-timestamp (Q) denotes the largest time stamp if any transaction that executed WRITE (Q) successfully.
   - ✓ R-timestamp (Q) denotes the largest time stamp if any transaction that executed READ (Q) successfully.
8. **Define blocks?**
   The database system resides permanently on nonvolatile storage, and is partitioned into fixed-length storage units called blocks.
9. **Define deadlock?**
   Neither of the transaction can ever proceed with its normal execution. This situation is called deadlock
10. **Define the phases of two phase locking protocol**
    Growing phase: a transaction may obtain locks but not release any lock.
    Shrinking phase: a transaction may release locks but may not obtain any new locks.
11. **Define upgrade and downgrade?**
    It provides a mechanism for conversion from shared lock to exclusive lock is known as upgrade.
    It provides a mechanism for conversion from exclusive lock to shared lock is known as downgrade.
12. **What is a database graph?**
    The partial ordering implies that the set D may now be viewed as a directed acyclic graph, called a database graph.
13. **What are uncommitted modifications?**
    The immediate-modification technique allows database modifications to be output to the database while the transaction is still in the active state. Data modifications written by active transactions are called uncommitted modifications.
14. **What is meant by buffer blocks?**
    The blocks residing temporarily in main memory are referred to as buffer blocks.
15. **Define shadow paging.**
    An alternative to log-based crash recovery technique is shadow paging. This technique needs fewer disk accesses than do the log-based methods.
16. **Define page.**
    The database is partitioned into some number of fixed-length blocks, which are referred to as pages.
17. **Explain current page table and shadow page table.**
    The key idea behind the shadow paging technique is to maintain two page tables during the life of the transaction: the current page table and the shadow page table. Both the page tables are identical when the transaction starts. The current page table may be changed when a transaction performs a write operation.

Anand Engineering College

18. **What is transaction?**
Collections of operations that form a single logical unit of work are called transactions.

19. **What are the drawbacks of shadow-paging technique?**
- ✓ Commit Overhead
- ✓ Data fragmentation
- ✓ Garbage collection

20. **What is meant by garbage collection.*(May/June 16)***
Garbage may be created also as a side effect of crashes. Periodically, it is necessary to find all the garbage pages and to add them to the list of free pages. This process is called garbage collection.

21. **What are the properties of transaction? Or Write the ACID properties of Transaction.** ***(Nov/Dec 14) (Apr/May 15)(May/June 16)***
Atomicity , Consistency, Isolation and Durability

22. **What is recovery management component?**
Ensuring durability is the responsibility of a software component of the base system called the recovery management component.

23. **When is a transaction rolled back?**
Any changes that the aborted transaction made to the database must be undone. Once the changes caused by an aborted transaction have been undone, then the transaction has been rolled back.

24. **Give an example of two phase commit protocol. *(Nov/Dec 15)***
Client want all or nothing transactions and Transfer either happens or nothing at all.

25. **What are the states of transaction?**
The states of transaction are
- ✓ Active
- ✓ Partially committed
- ✓ Failed
- ✓ Aborted
- ✓ Committed
- ✓ Terminated

26. **What is a shadow copy scheme?**
It is simple, but efficient, scheme called the shadow copy schemes. It is based on making copies of the database called shadow copies that one transaction is active at a time. The scheme also assumes that the database is simply a file on disk.

27. **What is serializability? How it is tested? *(Nov/Dec 14,16)***
A (possibly concurrent) schedule is serializable if it is equivalent to a serial schedule.
Precedence graph is used to test the serializability

28. **Mention the approaches of deadlock recovery**
The common solution is to roll back one or more transactions to break the deadlock
- ✓ Selection of victim
- ✓ Rollback
- ✓ Partial
- ✓ Total and Starvation.

Arunai Engineering College

29. **What is meant by concurrency control? *(Nov/Dec 15)***

A transaction is a particular execution of the program. When multiple transactions are trying to access the same shareable resource, many problems arise if the access control is not done properly. Mechanisms to which access control can be maintained is called Concurrency control.

30. **List the four conditions for deadlock. *(Nov/Dec 16)***

- ✓ mutual exclusion
- ✓ hold and wait or partial allocation
- ✓ no pre-emption
- ✓ resource waiting or circular wait

31. **What are different isolation levels in database?**

- ✓ Serializable.
- ✓ Repeatable reads.
- ✓ Read committed.
- ✓ Read uncommitted.
- ✓ Dirty reads.
- ✓ Non-repeatable reads.
- ✓ Phantom reads.
- ✓ Isolation levels vs read phenomena.

32. **What is isolation in database?**

Transaction isolation is an important part of any transactional system. It deals with consistency and completeness of data retrieved by queries unaffecting a user data by other user actions. A database acquires locks on data to maintain a high level of isolation.

UNIT-III / PART-B

1.	Discuss view serializability and conflict serializability ***(Nov/Dec 15)***
2.	Write short notes on Transaction State and discuss the properties of transaction.
3.	Briefly describe two phase locking in concurrency control techniques. ***(Nov/Dec 14,16)***
4.	Explain the concepts of concurrent execution in Transaction processing system.***(Nov/Dec 14)***
5.	Explain Transaction concept with an example. ***(Nov/Dec 14)***
6.	Explain about dead lock recovery algorithm with an example.
7.	Illustrate Granularity locking method in concurrency control.
8.	Describe Database Recovery concepts.
9.	What is concurrency control? How is it implemented in DBMS? Illustrate with a suitable example. ***(Nov/Dec 14)***
10.	Briefly explain about Two phase commit and three phase commit protocols. ***(Apr/May 15) (May/June 16) (Nov/Dec 14)***
11.	What is deadlock? How does it occur? How transactions be written to (i) Avoid deadlock (ii) guarantee correct execution. Illustrate with suitable example. ***(Nov/Dec 14,15,16)***

Arunai Engineering College

12.	Explain about Locking Protocols. ***(May/June 16)***
13.	Consider the following extension to the tree-locking protocol, which allows both shared and exclusive locks: • A transaction can be either a read-only transaction, in which case it can request only shared locks, or an update transaction, in which case it can request only exclusive locks. • Each transaction must follow the rules of the tree protocol. Read-only transactions may lock any data item first, whereas update transactions must lock the root first. Show that the protocol ensures serializability and deadlock freedom. ***(Nov/Dec 16)***
14.	Consider the following schedules. The actions are listed in the order they are schedule, and prefixed with transaction name. S1: T1: R(X), T2: R(x), T1: W(Y), T2: W(Y), T1: R(Y), T2: R(Y) S2:T3: R(X), T1: R(X), T1: W(Y), T2: R (Z), T2: W (Z), T3: R (Z) For each of the schedules, answer the following questions: i. What is the precedence graph for the schedule? ii. Is the schedule conflict-serializable? If so, what are all the conflict equivalent serial schedules? iii. Is the schedule view-serializable? If so, what are all the view equivalent serial schedules? ***(Apr/May 15)***

15.	Consider the following two transactions:
	T_1: read(A);
	read(B);
	if $A = 0$ **then** $B := B +$
	1; write(B).
	T_2: read(B);
	read(A);
	if $B = 0$ **then** $A := A + 1$;
	write(A).
	Add lock and unlock instructions to transactions $T31$ and $T32$, so that they observe the two-phase locking protocol. Can the execution of these transactions result in a deadlock? ***(Nov/Dec 16)***

UNIT IV IMPLEMENTATION TECHNIQUES

RAID – File Organization – Organization of Records in Files – Indexing and Hashing –Ordered Indices – B+ tree Index Files – B tree Index Files – Static Hashing – Dynamic Hashing – Query Processing Overview – Algorithms for SELECT and JOIN operations – Query optimization using Heuristics and Cost Estimation.

UNIT-IV / PART-A

1.	**What is B-Tree?** • A B-tree eliminates the redundant storage of search-key values. • It allows search key values to appear only once.
2.	**What is a B+-Tree index?** A B+-Tree index takes the form of a balanced tree in which every path from the root of the root of the root of the tree to a leaf of the tree is of the same length.

Arunai Engineering College

3.	**What is a hash index?** A hash index organizes the search keys, with their associated pointers, into a hash file structure
4.	**Define seek time.** The time for repositioning the arm is called the seek time and it increases with the distance that the arm is called the seek time.
5.	**Define rotational latency time.** The time spent waiting for the sector to be accessed to appear under the head is called the rotational latency time.
6.	**What is called mirroring?** The simplest approach to introducing redundancy is to duplicate every disk. This technique is called mirroring or shadowing.
7.	**What are the two main goals of parallelism?** • Load –balance multiple small accesses, so that the throughput of such accesses increases. • Parallelize large accesses so that the response time of large accesses is reduced.
8.	**What is an index?** An index is a structure that helps to locate desired records of a relation quickly, without examining all records
9.	**What are the factors to be taken into account when choosing a RAID level?** • Monetary cost of extra disk storage requirements. • Performance requirements in terms of number of I/O operations • Performance when a disk has failed and Performances during rebuild.
10.	**What are the types of storage devices?** Primary storage, Secondary storage, Tertiary storage, Volatile storage, Nonvolatile storage
11.	**What is called remapping of bad sectors?** If the controller detects that a sector is damaged when the disk is initially formatted, or when an attempt is made to write the sector, it can logically map the sector to a different physical location.
12.	**Define software and hardware RAID systems?*(May/June 16)*** RAID can be implemented with no change at the hardware level, using only software modification. Such RAID implementations are called software RAID systems and the systems with special hardware support are called hardware RAID systems.
13.	**Define hot swapping?** Hot swapping permits the removal of faulty disks and replaces it by new ones without turning power off. Hot swapping reduces the mean time to repair.
14.	**What are the ways in which the variable-length records arise in database systems?** Storage of multiple record types in a file, Record types that allow variable lengths for one or more fields, Record types that allow repeating fields.
15.	**What are the two types of blocks in the fixed –length representation? Define them.** ✓ Anchor block: Contains the first record of a chain. ✓ Overflow block: Contains the records other than those that are the first record of a chain.

Arunai Engineering College

16. **What is hashing file organization?**
In the hashing file organization, a hash function is computed on some attribute of each record. The result of the hash function specifies in which block of the file the record should be placed.

17. **What are called index-sequential files?**
The files that are ordered sequentially with a primary index on the search key are called index-sequential files.

18. **Define Primary index and Secondary Index**
It is in a sequentially ordered file, the index whose search key specifies the sequential order of the file. Also called clustering index. The search key of a primary index is usually but not necessarily the primary key. It is an index whose search key specifies an order different from the sequential order of the file. Also called non clustering index.

19. **Give an example of a join that is not a simple equi-join for which partitioned parallelism can be used.** ***(Nov/Dec 15)***
$r \; join \; (r.A = s.B) \wedge (r.A < s.c)$

20. **Differentiate static and dynamic hashing.** ***(Apr/May 15) (Nov/Dec 14,15)***

Static Hashing	Dynamic Hashing
In static hashing, when a search-key value is provided, the hash function always computes the same address.	Hash function, in dynamic hashing, is made to produce a large number of values and only a few are used initially.
The number of buckets provided remains unchanged at all times i.e. fixed	Dynamic hashing provides a mechanism in which data buckets are added and removed dynamically and on-demand .i.e. no. of buckets not fixed.
Space and overhead is more	Minimum space and less overhead
As file grows performance decreases	Performance do not degrade as file grows

21. **List out the mechanisms to avoid collision during hashing.*(Nov/Dec 16)***
- ✓ In overflow chaining, the overflow buckets of a given bucket are chained together in a linked list.
- ✓ Above scheme is called closed hashing. An alternative, called open hashing, which does not use overflow buckets, is not suitable for database applications.

22. **What are the disadvantages of B-Tree over B+ Tree?** ***(Nov/Dec 16)***
- ✓ Only small fraction of all search-key values are found early
- ✓ Non-leaf nodes are larger. Thus, B-Trees typically have greater depth than corresponding B+-Tree
- ✓ Insertion and deletion more complicated than in B+-Trees
- ✓ Implementation is harder than B+-Trees.
- ✓ Not possible to sequentially scan a table by just looking at leafs.

23. **What is called query processing?**
Query processing refers to the range of activities involved in extracting data from a database.

24. **What is called a query evaluation plan?**
A sequence of primitive operations that can be used to evaluate be query is a query evaluation plan or a query execution plan.

Arunai Engineering College

25.	**Explain "Query optimization"?*(May/June 16)*** Query optimization refers to the process of finding the lowest cost method of evaluating a given query.
26.	**State the need for Query Optimization. *(Apr/May 15)*** The query optimizer attempts to determine the most efficient way to execute a given query by considering the possible query plans.
	UNIT-IV / PART-B
1.	Describe File Organization.
2.	Define RAID and Briefly Explain RAID techniques.***(Nov/Dec 14, 15, 16) (Apr/May 15,16)***
3.	Explain Secondary storage devices.
4.	Explain about static and dynamic hashing with an example
5.	Explain about Multidimensional and parallel with an example
6.	Explain about ordered indices with an example
7.	Explain about B+ trees indexing concepts with an example ***(Nov/Dec 14)(May/June 16)***
8.	Explain about B trees indexing concepts with an example ***(Nov/Dec 14)***
9.	Illustrate indexing and hashing techniques with suitable examples. ***(Nov/Dec 15)***
10.	Explain about Query optimization with neat Diagram. ***(Nov/Dec 14,16)***
11.	Give a detailed description about Query processing and Optimization.Explain the cost estimation of Query Optimization ***(Nov/Dec 14).***
12.	Discuss about join order optimization and heuristic optimization algorithm. ***(Apr/May 15)***
13.	Briefly explain about Query Processing***(May/June 16)***
	UNIT V ADVANCED TOPICS Distributed Databases: Architecture, Data Storage, Transaction Processing – Object-based Databases: Object Database Concepts, Object-Relational features, ODMG Object Model, ODL, OQL – XML Databases: XML Hierarchical Model, DTD, XML Schema, XQuery – Information Retrieval: IR Concepts, Retrieval Models, Queries in IR systems.
	UNIT-V / PART-A
1.	**What is homogeneous distributed database and heterogeneous distributed database** A homogeneous distributed database has identical software and hardware running all databases instances, and may appear through a single interface as if it were a single database. A heterogeneous distributed database may have different hardware, operating systems, database management systems, and even data models for different databases.
2.	**Define Distributed Database Systems. *(Nov/Dec 16)*** Database spread over multiple machines (also referred to as sites or nodes).Network interconnects the machines. Database shared by users on multiple machines is called Distributed Database Systems
3.	**What are the types of Distributed Database** ✓ Homogenous distributed DB ✓ Heterogeneous distributed DB
4.	**Define fragmentation in Distributed Database** The system partitions the relation into several fragment and stores each fragment at different sites Two approaches : ✓ Horizontal fragmentation ✓ Vertical fragmentation

Arunai Engineering College

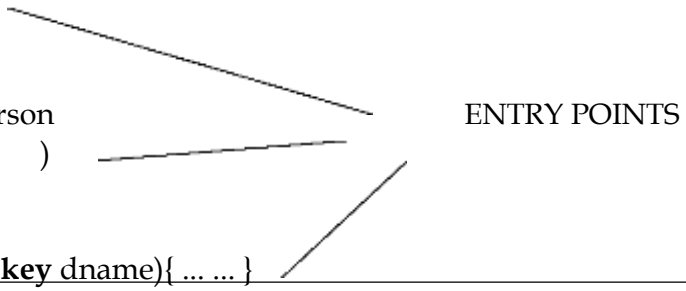
5.	**Define Database replication.** Database replication can be used on many database management systems, usually with a master/slave relationship between the original and the copies. The master logs the updates, which then ripple through to the slaves. The slave outputs a message stating that it has received the update successfully, thus allowing the sending of subsequent updates.
6.	**What is the advantage of OODB?** An integrated repository of information that is shared by multiple users, multiple products, multiple applications on multiple platforms.
7.	**What is Object database System?** An object database is a database management system in which information is represented in the form of objects as used in object-oriented programming. Object-relational databases are a hybrid of both approaches.
8.	**What are the advantages of OODB?** An integrated repository of information that is shared by multiple users, multiple products, multiple applications on multiple platforms. It also solves the following problems: 1. The semantic gap: The real world and the Conceptual model is very similar. 2. Impedance mismatch: Programming languages and database systems must be interfaced to solve application problems. But the language style, data structures, of a programming language (such as C) and the DBMS (such as Oracle) are different. The OODB supports general purpose programming in the OODB framework. 3. New application requirements: Especially in OA, CAD, CAM, CASE, object-orientation is the most natural and most convenient.
9.	**How do you define types in object relational feature in oracle?** Oracle allows us to define types similar to the types of SQL. The syntax is CREATE TYPE t AS OBJECT (list of attributes and methods);
10.	**Define ODMG Object model?** The ODMG object model is the data model upon which the object definition langauge (ODL) and object query language (OQL) are based.
11.	**Define ODL.** ODL langauge is used to create object specifications: • classes and interfaces - Using the specific langauge bindings to specify how ODL • constructs can be mapped to constructs in specific programming • language, such as C++, SMALLTALK, and JAVA.
12.	**Define Information Retrieval.** It is an activity of obtaining information resources relevant to an information need from a collection of information resources.
13.	**Define Relevance Ranking.** ***(Nov/Dec 14)*** A system in which the search engine tries to determine the theme of a site that a link is coming from.
14.	**Can we have more than one constructor in a class? If yes, explain the need for such a situation.** ***(Nov/Dec 15)*** Yes, default constructor and constructor with parameter
15.	**Define XML Database.** An XML database is a data persistence software system that allows data to be stored in XML format. These data can then be queried, exported and serialized into the desired format. XML databases are usually associated with document-oriented databases.

Arunai Engineering College

16. **Define OQL with syntax.**

- Entry point to the database: needed for each query which can
- be any named *persistent object*:
  - the name of the extent of a class

class Person
(**extent** persons
key ssn)
{ ... ... }
class Faculty extends Person
(**extent** faculy)
{ ... ... } **class**
Department
(**extent** departmet **key** dname){ ... ... }



17. **Define Crawling and indexing the web.** ***(Nov/Dec 14)***

Web Crawling is the process of search engines combing through web pages in order to properly index them. These "web crawlers" systematically crawl pages and look at the keywords contained on the page, the kind of content, all the links on the page, and then returns that information to the search engine's server for indexing. Then they follow all the hyperlinks on the website to get to other websites. When a search engine user enters a query, the search engine will go to its index and return the most relevant search results based on the keywords in the search term. Web crawling is an automated process and provides quick, up to date data.

18. **How does the concept of an object in the object-oriented model differ from the concept of an entity in the entity-relationship model?*(Nov/Dec 16)***

An entity is simply a collection of variables or data items. An object is an encapsulation of data as well as the methods (code) to operate on the data. The data members of an object are directly visible only to its methods. The outside world can gain access to the object's data only by passing pre-defined messages to it and these messages are implemented by the methods.

19. **Is XML Hierarchical?**

XML documents have a hierarchical structure and can conceptually be interpreted as a tree structure, called an XML tree. XML documents must contain a root element (one that is the parent of all other elements). All elements in an XML document can contain sub elements, text and attributes.

20. **What is DTD?**

A document type definition (DTD) contains a set of rules that can be used to validate an XML file. After you have created a DTD, you can edit it manually, adding declarations that define elements, attributes, entities, and notations, and how they can be used for any XML files that reference the DTD file.

21. **What is the use of XML Schema?**

XML Schema is commonly known as XML Schema Definition (XSD). It is used to describe and validate the structure and the content of XML data. XML schema defines the elements, attributes and data types. Schema element supports Namespaces.

22. **What is Xpath and Xquery?**

XPath can be used to navigate through elements and attributes in an XML document. XPath is a syntax for defining parts of an XML document. XPath uses path expressions to navigate in XML documents. XPath contains a library of standard functions. XPath is a major element in XSLT and in XQuery.

23. **Define Keyword Queries.**

Keyword-based queries are the simplest and most commonly used forms of IR queries: the user just enters keyword combinations to retrieve documents.

Anand Engineering College

24.	**What are the Types of Queries in IR Systems** • Keyword Queries. Boolean Queries • Phrase Queries • Proximity Queries • Wildcard Queries • Natural Language Queries.
25.	**State the steps to create DTD.** Create a new DTD, complete the following steps: 1. Create a project to contain the DTD if needed. 2. In the workbench, click File > New > Other and select XML > DTD. Click Next. 3. Select the project or folder that will contain the DTD. 4. In the File name field, type the name of the DTD, for example MyDTD.dtd. The name of your DTD file must end with the extension .dtd 5. Click Next. 6. Optional: You can use a DTD template as the basis for your new DTD file. To do so, click the Use DTD Template check box, and select the template you want to use. 7. Click Finish.
	UNIT-V / PART-B
1.	Explain about Object Oriented Databases and XML Databases.
2.	Explain in detail (i) Information Retrieval (iii) Transaction processing ***(Nov/Dec 14)***
3.	Write short notes on Distributed Transactions. ***(Nov/Dec 14)***
4.	Explain in detail the Client - Server Architecture for DDBMS
5.	Suppose an Object Oriented database had an object A, which references object B, which in turn references object C. Assume all objects are on disk initially? Suppose a program first dereferences A, then dereferences B by following the reference from A, and then finally dereferences C. Show the objects that are represented in memory after each dereference, along with their state. ***(Nov/Dec 15)***
6.	Suppose that you have been hired as a consultant to choose a database system for your client's application. For each of the following applications, state what type of database system (relational, persistent programming language–based OODB, object relational; do not specify a commercial product) you would recommend. Justify your recommendation. (i)A computer-aided design system for a manufacturer of airplanes. (ii)A system to track contributions made to candidates for public office. (iii)An information system to support the making of movies. ***(Nov/Dec 16)***
7.	Give the DTD for an XML representation of the following nested-relational schema *Emp = (ename, ChildrenSet* setof*(Children), SkillsSet* setof*(Skills))* *Children = (name, Birthday)* *Birthday = (day, month, year)* *Skills = (type, ExamsSet* setof*(Exams)).* *Exams = (year, city)* ***(Nov/Dec 16)***
8.	Explain XML Schema with an example.
9.	Explain various queries in IR Systems with an example.
10.	Explain ODL and OQL with an example.
11.	Explain ODMG – Object Model in detail.

Arunai Engineering College