INSTITUTE OF AERONAUTICAL ENGINEERING
(Autonomous)
Dundigal, Hyderabad - 500 043

COMPUTER SCIENCE AND ENGINEERING
DEFINITIONS AND TERMINOLOGY

Course Title	**DATABASE MANAGEMENT SYSTEMS**				
Course Code	AITC05				
Program	B.Tech				
Semester	IV	CSE			
Course Type	Core				
Regulation	UG20				
Course Structure	Theory			Practical	
	Lecture	Tutorials	Credits	Laboratory	Credits
	3	-	3	-	-
Course Coordinator	Mrs. V Divyavani, Assistant Professor				

COURSE OBJECTIVES:
The students will try to learn:

I	Efficient ways of designing database by encapsulating data requirements for business and organizational scenarios
II	Analysing and developing sophisticated queries in database language SQL for extracting information from large datasets
III	Enhancing skills in developing and managing data efficiently in related engineering problems.

COURSE OUTCOMES:
After successful completion of the course, students should be able to:

CO 1	**Outline** the importance of database system, RDBMS and its functionalities for voluminous data storage and management..	Understand
CO 2	**Model** the real world database systems using Entity Relationship Diagrams from the requirement specification.	Apply
CO 3	**Construct** queries in Relational Algebra, Relational Calculus and SQL to retrieve desired information.	Apply
CO 4	**Identify** appropriate normalization technique using dependencies for controlling the redundancy of data in database.	Apply

CO 5	**Demonstrate** ACID properties of Transaction processing, currency control protocols and recovery to preserve the database in a consistent state.	Understand
CO 6	**Organize** data storage and file organization techniques using tree and hash indices for effective query processing..	Apply

DEFINITION AND TERMINOLOGY:

S.No	DEFINITION	CO's
	MODULE I	
	CONCEPTUAL MODELING INTRODUCTION	
1	**What is Data?**	CO 1
	Data is a raw and unorganized fact that required to be processed to make it meaningful.	
2	**What is Information?**	CO 3
	Information is a set of data which is processed in a meaningful way according to the given requirement.	
3	**Define Database?**	CO 2
	A database is an organized collection of data, generally stored and accessed electronically from a computer system. Where databases are more complex they are often developed using formal design and modeling techniques.	
4	**Define Database Management System?**	CO 1
	It is a collection of programs that enables user to create and maintain a database. In other words it is general-purpose software that provides the users with the processes of defining, constructing and manipulating the database for various applications.	
5	**Define Database System?**	CO 1
	The database and DBMS software together is called as Database system.	
6	**What is File system?**	CO 1
	A file system is a process that manages how and where data on storage disk, typically a hard disk drive (HDD), is stored, accessed and managed.	
7	**Define Data abstraction?**	CO 1
	Database systems are made-up of complex data structures. To ease the user interaction with database, the developers hide internal irrelevant details from users. This process of hiding irrelevant details from user is called data abstraction.	
8	**What is Instance?**	CO 1
	The data stored in database at a particular moment of time is called instance of database. .	
9	**What is Schema?**	CO 1
	Design of a database is called the schema.	

10	**Define Data Independence?**	CO 1
	Data independence means that "The application is independent of the storage structure and access strategy of data". In other words, the ability to modify the schema definition in one level should not affect the schema definition in the next higher level.	
11	**What is Conceptual design?**	CO 1
	Conceptual design is the first stage in the database design process. The goal at this stage is to design a database that is independent of database software and physical details. The output of this process is a conceptual data model that describes the main data entities, attributes, relationships, and constraints of a given problem domain. This design is descriptive and narrative in form.	
12	**What is Relational Model?**	CO 1
	Relational Model represents how data is stored in Relational Databases. A relational database stores data in the form of relations (tables).	
13	**What is redundancy?**	CO 1
	Redundancy is defined as repetition of data in database.	
14	**What is DDL?**	CO 1
	DDL is defined as Data Definition Language which is used to define the database schema.	
15	**What is DML?**	CO 1
	DML is defined as Data Manipulation Language which is used for managing data within schema objects.	
16	**What is TCL?**	CO 1
	TCL is defined as Transaction control language which is used to manage the changes made by DML-statements.	
17	**What is DCL?**	CO 1
	DCL is defined as Data control Language which is used to control access to data stored in a database.	
18	**What is Cardinality?**	CO 1
	Number of rows in a relation.	
19	**State DBA?**	CO 1
	A database administrator (DBA) directs or performs all activities related to maintaining a successful database environment.	
20	**State Integrity Constraints?**	CO 1
	Integrity constraints provide a way of ensuring that changes made to the database by authorized users do not result in a loss of data consistency.	
21	**Define Super key?**	CO 1
	Super Key is the superset of primary key. The super key contains a set of attributes, including the primary key, which can uniquely identify any data row in the table.	

22	**Define Candidate key?** The candidate keys in a table are defined as the set of keys that is minimal and can uniquely identify any data row in the table.	CO 1
23	**Define Primary Key?** The primary key is selected from one of the candidate keys and becomes the identifying key of a table. It can uniquely identify any data row of the table.	CO 1
24	**Define Foreign Key?** A foreign key is an attribute value in a table that acts as the primary key in another another. Hence, the foreign key is useful in linking together two tables. Data should be entered in the foreign key column with great care, as wrongly entered data can invalidate the relationship between the two tables.	CO 1
25	**What is Hierarchical model?** Hierarchical model, data is organized into a tree like structure with each record ismhaving one parent record and many children. The main drawback of this model is that, it can have only one to many relationships between nodes.	CO 1
	MODULE II	
	RELATIONAL APPROACH	
1	**Define entity set?** An entity set is a set of entities of the same type . Entity sets need not be disjoint. For example, the entity set employee (all employees of a bank) and the entity set customer (all customers of the bank) may have members in common.	CO 3
2	**Define entity relationship model?** A conceptual model is a specific representation of the structure of data as entities and relationships, and is generally defined in a domain-specific language (DSL) that implements the concepts of the EDM.	CO 3
3	**Define attribute set?** An attribute set can be defined by a single attribute or by a set of attributes to apply to specific products.	CO 5
4	**Define key attribute?** Keys are very important part of Relational database model. They are used to establish and identify relationships between tables and also to uniquely identify any record or row of data inside a table. A Keycan be a single attribute or a group of attributes, where the combination may act as a key.	CO 5
5	**Define Derived attribute?** A derived attribute is an attribute whose value is calculated (derived) from other attributes. The derived attribute need not be physically stored within the database; instead, it can be derived by using an algorithm.As the name implies, multi-valued attributes may have many values.	CO 3

6	**Define Multivalued attribute?**	CO 2
	A multivalued attribute can have more than one value at a time for an attribute. For ex., skills of a surgeon is a multivalued attribute since a surgeon can have more than one skill.	
7	**Define Composite attribute**	CO 3
	A simple attribute is one component that is atomic. Acomposite attribute has multiple components, each of which is atomic or composite. Figure 4. ER diagram notation for composite attribute domain, name. Another way to classify attributes is either as single-valued or multi-valued.	
8	**Define Simple attribute?**	CO 3
	Simple attribute - Simple attributes are atomic values, which cannot be divided further. For example, a student's phone number is an atomic value of 10 digits. Composite attribute - Composite attributesare made of more than one simple attribute. For example, a student s complete name may have first name and last name.	
9	**Define Cardinality?**	CO 3
	Cardinality in DBMS defines the maximum number of relationship instances in which an entity can participate.	
10	**Define participation constraints?**	CO 3
	Participation constraint defines the least number of relationship instances in which an entity must participate.	
11	**Define Generalization?**	CO 3
	Generalization is like a bottom-up approach in which two or more entities of lower level combine to form a higher level entity if they have some attributes in common. ... In generalization, entities are combined to form a more generalized entity, i.e., subclasses are combined to make a superclass.	
12	**Define aggregation?**	CO 3
	Relationship with its corresponding entities is aggregated into a higher level entity. For example: Center entity offers the Course entity act as a single entity in the relationship which is in a relationship with another entity visitor.	
13	**Define specialization?**	CO 3
	Specialization is a top-down approach, and it is opposite to Generalization. In specialization, one higher level entity can be broken down into two lower level entities.Specialization is used to identify the subset of an entity set that shares some distinguishing characteristics. .	
14	**Define Total participation Constraint?**	CO 3
	Every entity instance must be connected through the relationship to another instance of the other participating entity types is termed as total participation constraint.	

15	**What is partial participation Constraint?** If All instances need not participate in relationship then the constraint is partial participation.	CO 3
16	**What is Role of Entity set?** An entity set participates in several relationship sets and also participate more than once in a single relationship set which is known as Roles of entity set.	CO 3
17	**What is weak entity set?** An entity set that does not have a primary key is referred to as a weak entity set.	CO 3
18	**List types of joins?** Various forms of join operation are: Inner Joins: • Theta join • EQUI join • Natural join Outer join: • Left Outer Join • Right Outer Join • Full Outer Join	CO 3
19	**What is Strong entity set?** An entity set that have a primary key is referred to as a Strong entity set.	CO 3
20	**How to convert a weak entity set into strong?** Convert a weak entity set into strong entity set by attaching a primary key to it. The primary key of a weak entity set is formed by the primary key of the strong entity set on which the weak entity set is existence dependent, plus the weak entity set's discriminator	CO 3
21	**Define Attribute Inheritance?** Attribute inheritance is a lower-level entity set inherits all the attributes and relationship participation of the higher-level entity set to which it is linked.	CO 3
22	**What is a ISA relationship?** The ISA relationship is referred to as superclass - subclass relationship to represent generalization – specialization.	CO 3
MODULE III		
SQL QUERY BASICS RDBMS NORMALIZATION		
1	**What is Relational algebra?** Relational Algebra is procedural query language, which takes relation as input and generates relation as output. Relational algebra mainly provides theoretical foundation for relational databases and SQL.	CO 3
2	**List Relational algebra operations?** Relational algebra operations are selection, projection, set operations, renaming, joins and division.	CO 1
3	**What is a selection operation?** The select operation selects tuples that satisfy a given predicate. It is denoted by sigma (σ). Notation: σ p(r) .	CO 2

4	**What is a Projection operation?**	CO 3
	This operation shows the list of those attributes that we wish to appear in the result. Rest of the attributes is eliminated from the table. It is denoted by. Notation: A1, A2, An (r)	
5	**Define BETWEEN operator?**	CO 6
	The BETWEEN operator selects values within a given range. The values can be numbers, text, or dates.	
6	**What are Union operations?**	CO 6
	Union is a binary relation. Suppose there are two tuples R and S. The union operation contains all the tuples that are either in R or S or both in R and S.It eliminates the duplicate tuples. It is denoted by U Notation: R U S	
7	**What are Intersection operations?**	CO 6
	Intersection is a binary relation. Suppose there are two tuples R and S. The set intersection operation contains all tuples that are in both R and S. It is denoted by intersection ∩. Notation: R ∩ S	
8	**What is Set difference operations?**	CO 6
	Set difference is a binary relation. Suppose there are two tuples R and S. The set intersection operation contains all tuples that are in R but not in S.It is denoted by intersection minus (-). Notation: R - S	
9	**What is renaming operations?**	CO 5
	The rename operation is used to rename the output relation. It is denoted by rho (ρ). ρ(STUDENT1, STUDENT)	
10	**Define Fading effects?**	CO 2
	In wireless communications, fading is variation of the attenuation of a signal with various variables. These variables include time, geographical position.	
11	**Define frequency selective fading?**	CO 6
	Frequency selective fading is a radio propagation anomaly caused by partial cancellation of a radio signal .	
12	**Define like operator?**	CO 6
	Like operator is used for pattern matching in sql and denoted by wildcard symbol %.	
13	**What are natural join operations?**	CO 6
	A natural join is the set of tuples of all combinations in R and S that are equal on their common attribute names. It is denoted by inf.	
14	**What is a sub query?**	CO 3
	a subquery is a query within a query. These subqueries can reside in the WHERE clause, the FROM clause, or the SELECT clause.	
15	**Define Aggregate functions?**	CO 6
	Aggregate functions compute a single result from a set of input values. Few functions are count, avg, max, min, sum .	

16	**What is domain relational calculus?** In domain relational calculus, filtering variable uses the domain of attributes. Domain relational calculus uses the same operators as tuple calculus. It uses logical connectives ∧ (and), ∨ (or) and × (not).It uses Existential ∃ and Universal Quantifiers (∀) to bind the variable. Notation: a1, a2, a3, ..., an — P (a1, a2, a3, ... ,an) Where a1, a2 are attributes P stands for formula built by inner attributes.	CO 6
17	**What is tuple relational calculus?** The tuple relational calculus is specified to select the tuples in a relation. The result of the relation can have one or more tuples. In TRC, filtering variable uses the tuples of a relation. Notation: {T - P (T)} or {T - Condition (T)} WhereT is the resulting tuples P(T) is the condition used to fetch T.	CO 6
18	**Define Theta join?** The general case of JOIN operation is called a Theta join. It is denoted by symbol A infinity B Theta join can use any conditions in the selection criteria.	CO 6
19	**What is fading in communication?** In an inner join, only those tuples that satisfy the matching criteria are included, while the rest are excluded.	CO 6
20	**What is natural Join?** Natural join can only be performed if there is a common attribute (column) between the relations. The name and type of the attribute must be same.	CO 6
MODULE IV		
TRANSACTION MANAGEMENT		
1	**What is Normalization?** Normalization is a process of organizing the data in database to avoid data redundancy, insertion anomaly, update anomaly deletion anomaly.	CO 1
2	**Define Decomposition?** Decomposition is a process of dividing a complex relation into simple sub relations.	CO 6
3	**Define Functional Dependency?** A functional dependency is a constraint between two sets of attributes in a relation from a database. It is a relationship that exists when one attribute uniquely determines another attribute.	CO 4
4	**What is redundancy?** Redundancy is defined as repetition of data in database.	CO 3
5	**What is Trigger?** A trigger is a stored procedure in database which automatically invokes whenever a special event in the database occurs.	CO 6
6	**What is an Assertion?** An assertion ensures a certain condition will always exist in the database.	CO 4

7	**Define Normal forms?**	CO 5
	Normal form is used to reduce redundancy from the database table. Types of normal forms are: 1NF,2NF,3NF,4NF and 5NF	
8	**List out anomalies in dbms?**	CO 4
	Insertion, deletion and update anomalies.	
9	**Define 1st normal form?**	CO 2
	A relation is in first normal form if and only if the domain of each attribute contains only atomic values, and the value of each attribute contains only a single value from that domain.	
10	**Define 2nd normal form?**	CO 3
	The Second Normal Form states that it should meet all the rules of 1NF and there must be no partial dependences of any of the columns on the primary key.	
11	**Define 3rd normal form?**	CO 2
	The Third Normal Form states that it should meet all the rules of second normal form and no primary fields are dependent on the other primary key.	
12	**Define 4th normal form?**	CO 4
	The Fourth Normal Form states that it should meet all the rules of 3NF and there are no non-trivial multivalued dependencies other than a candidate key.	
13	**Define BCNF normal form?**	CO 6
	It is a slightly stronger version of the third normal form. A table is in BCNF if every functional dependency X $\rightarrow$ Y, X is the super key of the table.	
14	**Define 5th normal form?**	CO 4
	The Fifth Normal Form states that it should meet all the rules of 4NF and not contains any join dependency and joining should be lossless.	
15	**Define prime attributes?**	CO 1
	Prime attributes are attributes of the relation which exist in at least one of the possible candidate keys.	
16	**Define non-prime attributes?**	CO 3
	A non-prime attribute of R is an attribute that does not belong to any candidate key.	
17	**Define a join dependency?**	CO 2
	A join dependency is a constraint on the set of legal relations over a database scheme.	
18	**Define Multi valued dependency?**	CO 4
	Multivalued dependency occurs when there are more than one independent multivalued attributes in a table.	

19	**Define Transitive dependency?**	CO 3
	When an indirect relationship causes functional dependency it is called Transitive Dependency. If P -> Q and Q -> R is true, then P-> R is a transitive dependency..	
20	**What is Insertion Anomaly ?**	CO 2
	An insertion anomaly is the inability to add data to the database due to absence of other data.	
21	**When Deletion Anomaly occurs?**	CO 5
	Deletion Anomalies happen when the deletion of unwanted information causes desired information to be deleted as well.	
22	**What is Update anomaly?**	CO 6
	Update Anomaly is to update a value, it need update multiple rows. Update anomalies are due to redundancy.	
MODULE V		
DATA STORAGE AND QUERY PROCESSING		
1	**Define transaction in DBMS?**	CO 5
	A transaction is a unit of program execution that accesses and/or updates various data items..	
2	**State the property Atomicity of a Transaction?**	CO 6
	Atomicity of a transaction is either all operations of the transaction are properly reflected in the database or none of them.	
3	**What is effect Durability of Transaction?**	CO 3
	Durability ensures that after a transaction completes successfully, the changes it has made to the database persist, even if there are system failures.	
4	**Indicate the importance of Isolation property of a Transaction?**	CO 6
	Although multiple transactions may execute concurrently, each transaction must be unaware of other concurrently executing transactions and result as it is executing alone.	
5	**How Consistency of a transaction preserved?**	CO 4
	Execution of a transaction in isolation preserves the consistency of the database.	
6	**Define the role of Recovery Management component?**	CO 1
	The recovery-management component of a database system implements the support for atomicity and durability.	
7	**Define Schedule of transactions?**	CO 5
	Schedule is a sequences of instructions that specify the chronological order in which all instructions of concurrent transactions are executed.	
8	**Define a serializable Schedule?**	CO 2
	A (possibly concurrent) schedule is serializable if it is equivalent to a serial schedule .	

9	**When a schedule is View serializable?**	CO 3
	A schedule S is view serializable if it is view equivalent to a serial schedule.	
10	**When two instructions are conflict to each other?**	CO 6
	Instructions li and lj of transactions Ti and Tj respectively, conflict if and only if there exists some item Q accessed by both li and lj, and at least one of these instructions wrote Q.	
11	**Define Recoverable schedule?**	CO 5
	A schedule is recoverable if a transaction Tj reads a data item previously written by a transaction Ti , then the commit operation of Ti appears before the commit operation of Tj.	
12	**Define a Cascade less schedules?**	CO 4
	A Schedule is Cascadeless schedules in which cascading rollbacks cannot occur; for each pair of transactions Ti and Tj such that Tj reads a data item previously written by Ti, the commit operation of Ti appears before the read operation of Tj.	
13	**What are various possible failures may occur during concurrent transactions?**	CO 3
	Two main issues to deal with: Failures of various kinds, such as hardware failures and system crashes Concurrent execution of multiple transactions.	
14	**How Concurrency control schemes ensure Isolation of transactions?**	CO 3
	Concurrency control schemes are mechanisms to achieve isolation, that is, to control the interaction among the concurrent transactions in order to prevent them from destroying the consistency of the database.	
15	**Define precedence graph with reference to graph based protocols?**	CO 2
	Precedence graph is a direct graph where the vertices are the transactions (names) and arcs are drawn from Ti to Tj if the two transaction conflict, and Ti accessed the data item on which the conflict arose earlier.	
16	**What is purpose of Lock in Lock based protocols?**	CO 4
	A lock is a mechanism to control concurrent access to a data item.	
17	**State the functions of Growing Phase in two phase locking protocol?**	CO 1
	In Growing Phase of locking protocol, a transaction may obtain locks and transaction may not release locks.	
18	**State the functions of Shrinking Phase in two phase locking protocol?**	CO 4
	In Shrinking Phase of two phase locking protocol a transaction may release locks and transaction may not obtain locks.	
19	**Distinguish ordered indices from hash indices?**	CO 3
	Ordered indices: search keys are stored in sorted order where as Hash indices: search keys are distributed uniformly across buckets using a hash function.	

20	**List various index evaluation metrics?**	CO 4
	Index evaluation techniques are : Access types, Access time, Insertion time, Deletion time, Space overhead.	
21	**Differentiate Dense index from Sparse Index?**	CO 5
	In Dense index index record appears for every search-key value in the file But Sparse Index contains index records for only some Applicable search-key values .	
22	**How B-Tree and B+-tree are different?**	CO 6
	B-Tree is Similar to B+-tree, but B-tree allows search-key values to appear only once; but in B+-tree some search keys appear repeatedly in various levels of nodes. In B-tree; an additional pointer field for each search key in a non leaf node must be included.	
23	**Write characteristics of Extendable hashing function?**	CO 1
	Extendable hashing one form of dynamic hashing Hash function generates values over a large range typically b bit integers, with b = 32. At any time use only a prefix of the hash function to index into a table of bucket addresses.	
24	**Distinguish static hashing from dynamic hashing?**	CO 4
	In static hashing, function h maps search-key values to a fixed set of B of bucket addresses and dynamic hashing Allows the hash function to be modified dynamically and map to variable set of buckets.	
25	**Write Advantages and disadvantages of extendable hashing?**	CO 5
	Benefits of extendable hashing are : Hash performance does not degrade with growth of file, Minimal space overhead Disadvantages of extendable hashing are: Extra level of indirection to find desired record, Bucket address table may itself become very big.	

Course Coordinator: **HOD CSE**
Mrs. V Divyavani, Assistant Professor