UNIT-1

DBMS Concepts and Architecture:

1. A database-management system (DBMS) is a collection of interrelated data and a set of programs to access these data. The collection of data, usually referred to as the database, contains information relevant to an enterprise. The primary goal of a DBMS is to provide a way to store and retrieve database information that is both convenient and efficient.

2. Database approach vs Traditional file accession approach -
   Major disadvantage of traditional file accession approach -
   (1) Data redundancy and inconsistency.
   (2) Difficulty in accessing data.
   (3) Data isolation.
   (4) Integrity problems.
   (5) Atomicity problems.
   (6) Concurrent access anomalies.
   (7) Security problems.

3. Advantages of database systems -
   (1) Compactness
   (2) Speed
   (3) Less drudgery & mechanized (machine can be better than hands)
   (4) Accurate, up-to-date information is available on demand at any time
   (5) Protection.

4. Data Models - Structure of a database or a collection of conceptual tools for describing data, data relationships, data semantics, and consistency constraints. There are two data models -
   (1) Entity-Relationship (E-R) model
   (2) Relational model.
5. Schemas and Instances -

- The collection of information stored in the database at a particular moment is called an instance of the database.
- The overall design of the database is called the database schema. There are three schemas -
  1. Physical schema - describes database design at the physical level
  2. Logical schema - describes database design at the logical level
  3. Subschema - describes different views of the database

6. Data Independence -

- The capacity to change the schema at one level of a database without having to change the schema at another higher level.
  - Logical data independence - The capacity to change the conceptual schema (logical schema) without having to change the internal schema (physical schema).
  - Physical data independence - The capacity to change the internal schema (physical schema) without having to change the conceptual schema (logical schema).

7. Data Abstraction -

- To provide an abstract view of the data hiding the details of how data are stored and maintained, it is known as Data Abstraction.
- Three levels of abstraction are -
  1. Physical level - lowest level of abstraction describes how the data are actually stored
  2. Logical level - Next higher level of abstraction describes what data are stored in the database, and what relationships exist among these data
  3. View level - highest level of abstraction describes only part of the entire database
metadata - data about data

8. Database languages -
   Data definition language (DDL) - To specify the database schema
   Data manipulation language (DML) - To access or manipulate data
   There are two types -
   Procedural DMLs - To specify what data are needed and how to get those data
   Declarative DMLs (Non-procedural DMLs) - To specify what data are needed without specifying how to get them data

9. DBMS Interfaces -
   Types of interfaces provided by the DBMS include -
   Menu-Based Interface for Web Clients or Browsing -
     → Present user with list of options (menus)
     → lead user through formulation of request
     → Query is composed of selection of items from menu displayed by system
   Form-Based Interface -
     → Display a form to each user
     → Users can fill out form to insert new data or fill out only certain entries
     → Designed for novice users as interface to canned transactions
   Graphical User Interface -
     → Display a schema to the user in diagram form. The user can specify a query by manipulating the diagram. GUIs use both forms and menus
   Natural Language Interface -
     → Accept requests in written English or other languages and attempt to understand them
     → Interface has its own schema, and a dictionary of important words. Uses the schema and dictionary to interpret a natural language request
Database Structure (DBMS Architecture) -

- Have users (tellie, agents, web-users)
  - Use
  - Application Interfaces
  - Application Programs
    - Application Programs
      - Compiler and Linker
      - DML compiler and optimizer
      - Query evaluation engine
      - Cache manager
      - File manager
      - Buffer manager
      - Authorization and integrity manager
    - DDL interpreter
    - Transaction manager
      - Storage manager
      - Disk storage
      - Data
      - Index
      - Statistical data
      - Referential Dictionary

A storage manager 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. Query processor components include DDL interpreter, DML compiler, and query evaluation engine.
Database Administration (DBA) -

A person who has central control of both the data and the programs that access that data on the system is called a DBA.

Functions of DBA are:
1. Schema definition
2. Storage structure and access method definition
3. Schema and physical organization modification
4. Granting of authorization for data access
5. Routine Maintenance

ER data model:
The Entity-relationship (E-R) data models perceive the real world as consisting of basic objects, called entities, and relationships among these objects.

1. Entity - It is a "thing" or "object" in the real world.
   An entity set is a set of entities of the same type that share the same properties, or attributes.

2. Attributes - An entity is represented by a set of attributes. Attributes are characteristics, properties possessed by each member of an entity set.
   - Simple attributes - Having simple properties
   - Composite attributes - Can be divided into subparts (other attributes)
   - Single-valued attribute - Having a single value
   - Multi-valued attribute - Having set of values
   - Derived attribute - Value derived from other attributes

3. Entity types -
   - Strong entity set - It is having a primary key.
   - Weak entity set - A set class not have a primary key.
4. **Entity-Relationship Diagram**
   - **Overall logical structure of a database graphically.**
   - **Major components are:**
     - **Rectangle** - entity set
     - **Ellipses** - attribute
     - **Diamonds** - relationship set
     - **Link attributes** to entity set & entity set to relationship set.
     - **Double-ellipses** - multivalued attribute
     - **Boxed ellipses** - derived attribute
     - **Double lines** - total participation of an entity in a relationship set
     - **Double Rectangle** - represent weak entity set.

5. **Mapping Cardinalities (Association):**
   - **One to One**
   - **One to Many**
     - An entity in A is associated with at most one entity in B, and an entity in B is associated with at most one entity in A.
   - **One to Many**
     - An entity in A is associated with any number of entities in B, however, an entity in B is associated with at most one entity in A.
   - **Many to One**
     - An entity in A is associated with at most one entity in B, however, an entity in B is associated with any number of entities in A.
   - **Many to Many**
     - An entity in A is associated with any number of entities in B, and an entity in B is associated with any number of entities in A.

6. **Specialization:**
   - The process of subtyping within an entity set. It is a top-down approach based on some specific attributes of entity types within a given entity.
   - **Diagram:**
     - **Account**
     - **SAVING**
     - **CURRENT**
Generalization - The refinement from an initial entity set into
decreased levels of entity subconcepting represents a top down
design process

Generalization - Group of entities are combined together into a
single group based on some common features, some common attributes.
It is a bottom up approach.

Generalization is a simple inversion of specialization.

7. Aggregation -

The limitation of E-IR model is that it cannot express
relationships among relationships. Aggregation is an abstraction through
which relationships are treated as higher level entities.

[Diagram of relationships between entities]

8. Object-based data model - describe data at the logical and view levels.

- ER model
- Object-oriented model
- Semantic model
- Functional model

Record-based data model - Also describe data at logical & view level, stored
in the form of tables

- Relational model
- Network model
- Hierarchical model

- Entity-Relationship
- Object-oriented model
- Semantic model
- Functional model