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 those data.

The collection of data, usually refered to as the database, contains information relevant to an enterprize.

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 v/s Traditional file accessing approach -

Major disadvantages of traditional file accessing approach are -

(1) Data redundancy and inconsistency.
(2) Difficulty in accessing data
(3) Data isolation
(3) Integrity problems
(4) Atomicity problems
(5) Concurrent-access anomalies
(6) Security problems

③ Advantages of database systems -

(1) Compactness
(2) Speed
(3) less drudgery ~~(mechanical)~~ (machine are 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) Subschemas → 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 next 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 -

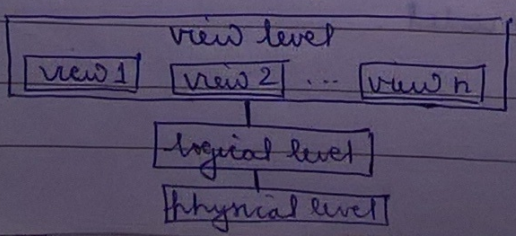
To provide abstract view of the data hiding the details of how data are stored and ~~makain~~ maintained, is known as data abstraction.

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 those 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 those data

⑨ DBMS Interfaces -

Types of interfaces provided by the DBMS include -

Menu-Based Interfaces for Web Clients or Browsing -

→ Present user with list of options (menus).

→ Lead user through formulation of request.

→ Query is composed of selection of options from menu displayed by system

Form-Based Interfaces -

→ Display a form to each user

→ User can fill out form to insert new data or fill out only certain entries

→ Designed & Programmed for naive users as interfaces to canned transactions

Graphical User Interfaces -

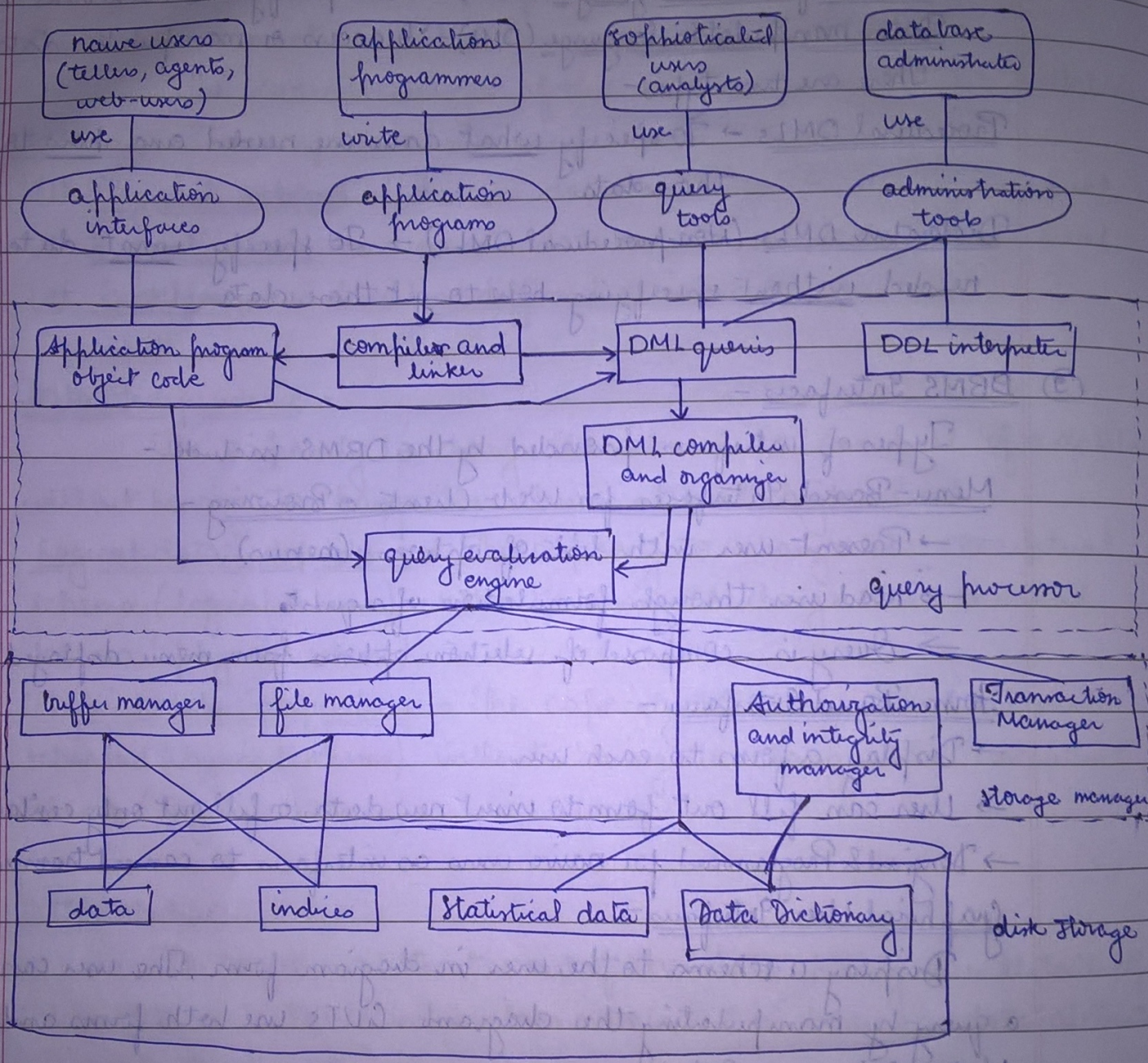
Display a schema to the user in diagram form. The user can specify a query by manipulating the diagram. GUI's use both forms and menus.

Natural language Interfaces -

→ 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.

(10) Database Structure (DBMS Architecture) -



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 ~~box~~ engine

(11) Database Administrator (DBA) -

A person who has central control of both the data and the programs that access those data over 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 perceives 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 set of entities of the same type that share the same properties, or attributes.

(2) Attributes - An entity is represented by set of attributes. Attributes are descriptive 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 attributes - Having a single value.

Multi-valued attributes - Having set of values.

Derived attributes - Having value derived from other attributes

(3) Entity types -

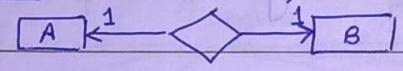
Strong entity set - A set having a primary key.

Weak entity set - A set does not have a primary key.

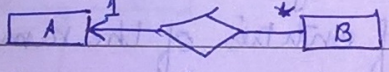
④ Entity-Relationship Diagram -
Overall logical structure of a database graphically. Major components are -

- Rectangles - entity sets
- Ellipses - attributes
- Diamonds - Relationship sets
- lines - link attributes to entity sets & entity sets to relationship sets
- Double-ellipses - Multivalued attributes
- Dashed ellipses - Derived attributes
- Double lines - Total participation of an entity in a relationship set
- Double Rectangles - Represent weak entity sets.

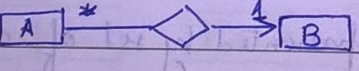
⑤ Mapping Cardinalities (Association) -

→ One to One - 

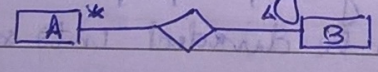
→ ~~One to Many~~ - An entity in A is associated with at most one entity in B, and an entity in B is associated with at most one entity in A

→ 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

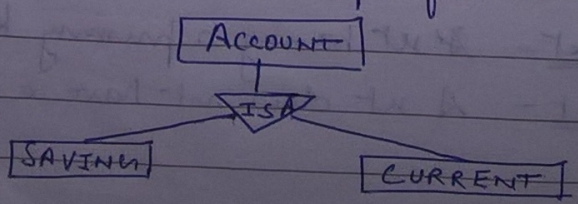
→ Many to One - 

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 subgrouping within an entity set. It is a top down approach based on some specific attributes of entity types within a given entity.



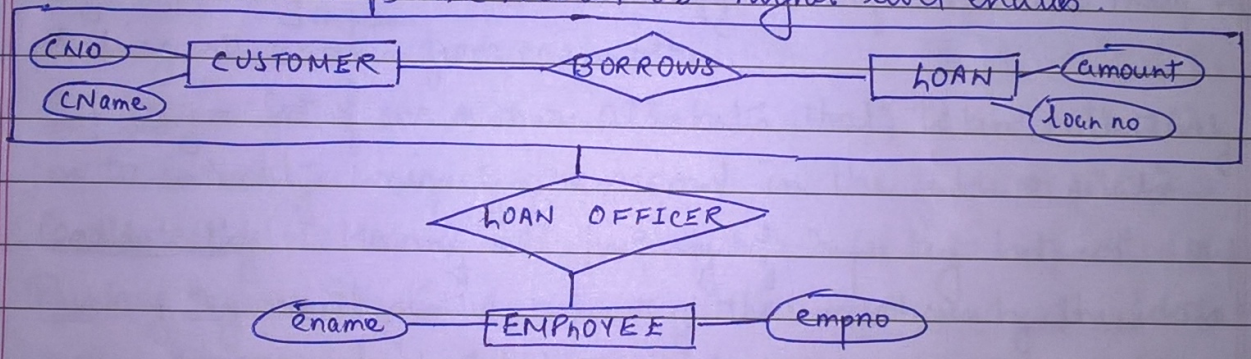
Generalization - The refinement from an initial entity set into successive levels of entity subgroupings represents a top-down design process

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 -

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



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

- ER model → entities & relationships
- Object-oriented model → collection of objects, use of class and objects
- Semantic model → inter dependencies among entities
- Functional model → use of functions (objects and its properties & relationships)

Data models

Record-based data model - Also describe data at logical & view levels, fixed format records

- Relational model → in the form of Table
- Network model → Using pointers
- Hierarchical model → collection of trees