UNIT-I

Data, Database, Database Management System:

Data: means known facts that can be recorded and that have some meaning. eg. names, telephone numbers, address.

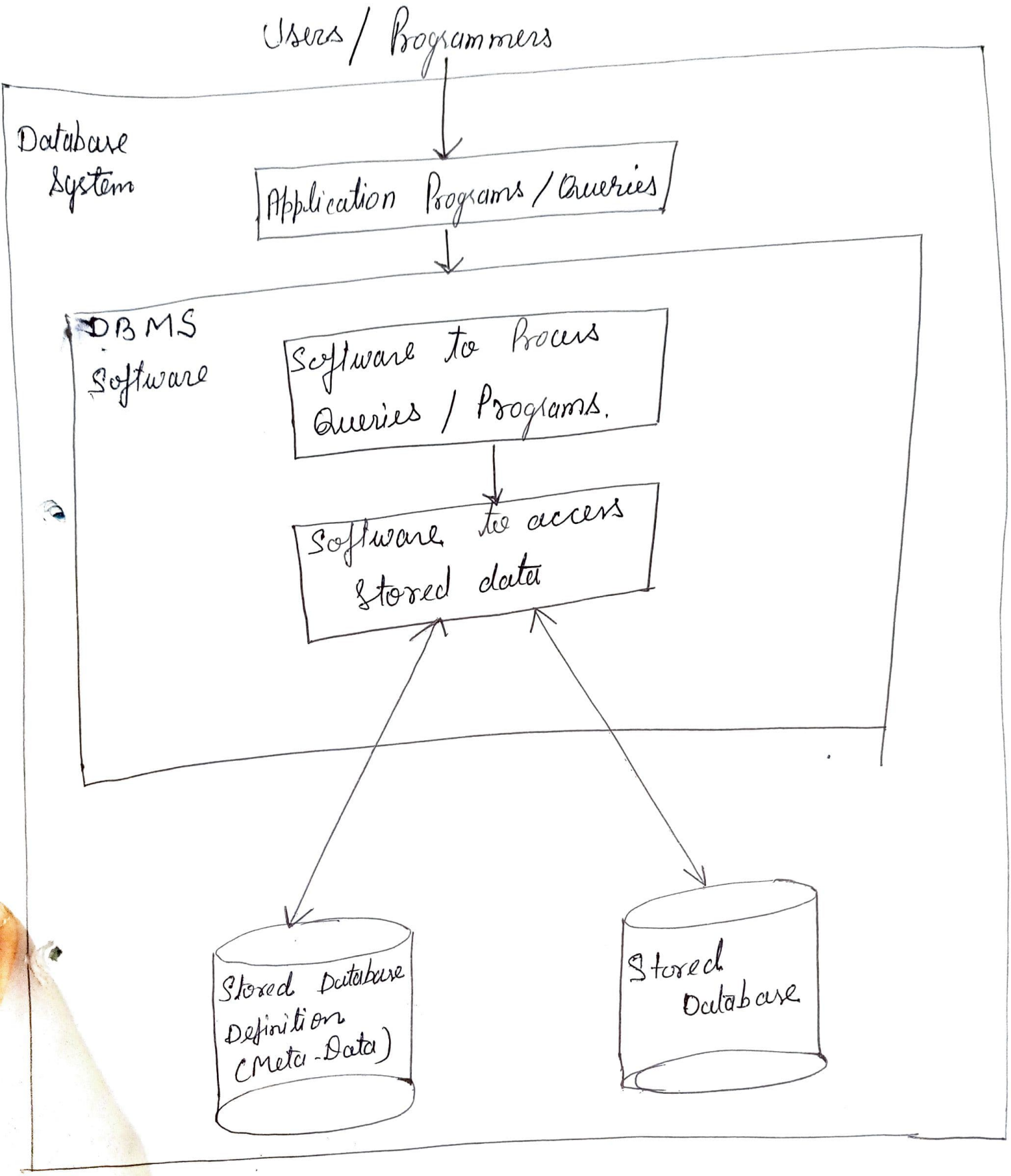
Database: It is a collection of logically related data.
eg. student records.
Example of commercial database is Amazon.com

Database Management System (DBMS): It is a collection of programs that enables users to create and maintain a database.
eg. Oracle, DB2, Sybase, SQL Server, MySQL.
Microsoft Access.

/* File Vs Database */

File: A collection of records or documents dealing with one organisation, person area or subject.
- manual (Paper) files.
- computer files.

Database: A collection of similar records with relationships between records.
- Bibliographic, statiscal, busines data etc.



A simplified database system environment

Applications of DBMS :-

1. Banking. — customer information, accounts, loans, transactions
2. Airlines — reservation and schedule information.
3. Universities — student information, course registration
4. Credit card Transaction — for purchase on credit card & generation of monthly statements.
5. Telecommunication — records of call made, monthly bill, etc.
6. Finance :- for storing information about holdings, sales, and purchase of financial instrument such as stocks and bonds.
7. Sales :- for customer, product and purchase information.

Components of DBMS.

(1) Hardware - The H/w is the actual computer system used for keeping and accessing the database conventional DBMS Hardware consists of secondary storage devices such as hard disk.

(2) Software - S/w is the actual DBMS between the physical database and the users of the systems. All the request from the user for accessing the database are handled by DBMS.

(3) Data :- Raw facts & figures..

(4) User :- There are no. of users who can access or retrieve the data on demand using the application and the interfaces provided by the DBMS.

The users of database can be classified into the following groups -

(1) Naive users
(2) Online Users
(3) Sophisticated users
(4) Specialised users
(5) Application programmers
(6) DBA - Database Administrator.

① Naive Users:- Those who need not be aware of the presence of the database system. They are the end users of the database who work through a menu driven application programs.

② Online Users:- Those users who may communicate with database directly through an online terminal

③ Sophisticated User:- They are those user who interact with the system without writing the program Instead they form their request in database query language.

④ Specialized User:- These user who write specialized database application.

⑤. Application programmer:- These users who are responsible for developing the application programs or user-interface. The application program could be written in high level language.

⑥. DBA - Database administrator - It is a person or the group incharge for implementing the database system within the organisation. The DBA has all the privledge allowed by the DBMS and can assign or remove the privilages from the Users.

DBMS Benefits

- Minimal data redundancy.
- consistency of data.
- Integration of data.
- Sharing of data.
- Ease of application development
- Security, privacy, and integrity controls.
- Data accessibility and responsiveness
- Data independence.
- Reduced program maintenance.

Disadvantages of DBMS

- Problem associated with centralisation
- cost of software / hardware
- complexity of backup and recovery.

operation perform on Database -

1- Insertion
2 - Updation
3 - Deletion
4 - Retrive
5 - Sorting.

Problem with file processing system:

- Inconsistent data
- Inflexibility
- Limited data sharing
- Poor enforcement of standards
- Excessive program maintenance.
- Duplication of data.
- Difficulty in representing the data from the user point of view.
- Data security - the security of data is low in the file based system because the data is maintain in a flat file is easy accessible.
- Transactional problem:- This system does not satisfy transactional properties called ACID properties, A → Atomicity, C → Consistency, I → Isolation, D → Durability.
- It can lead to loss of data integrity.

Problem with file processing system:

- Inconsistent data
- Inflexibility
- Limited data sharing
- Poor enforcement of standards
- Excessive program maintenance.
- Duplication of data.
- Difficulty in representing the data from the user point of view.
- Data security - the security of data is low in the file based system because the data is maintain in a flat file is easy accessible.
- Transactional problem:- This system does not satisfy transactional properties called ACID properties, A → Atomicity, C → Consistency, I → Isolation, D → Durability.
- It can lead to loss of data integrity.

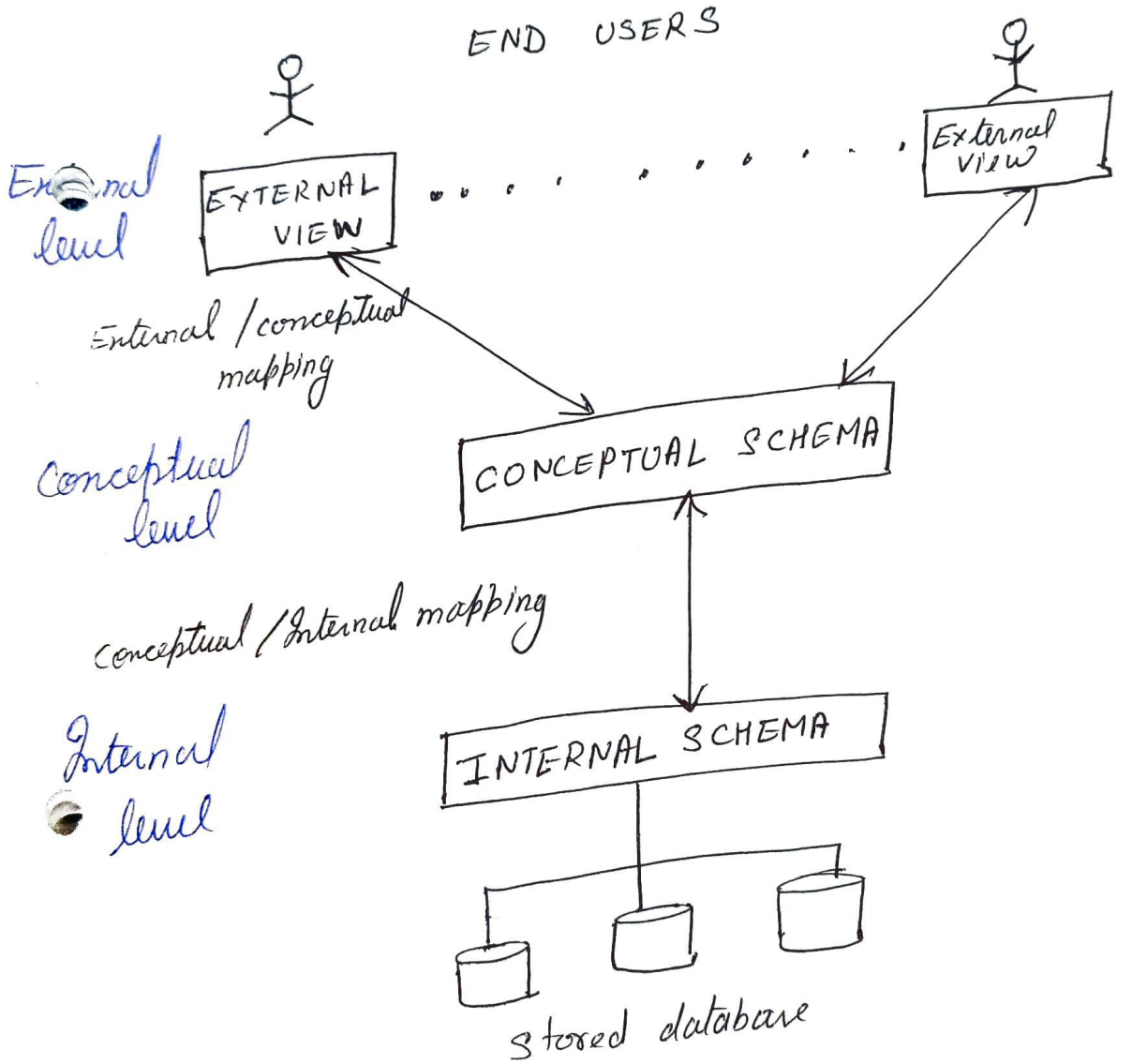
File Management eg c++, or COBOL Program	Database Menergement eg. Oracle, Sybase - .
(1). Small System	(1) Large system
(2). Relatively cheap	(2) relatively expensive.
(3). few files	(3) many files
(4) Simple structure	(4) complex structure
(5) little preliminary design	(5) Vast preliminary design.
(6) integrity left to programmer	(6) inbuilt integrity checking.
(7) no security	(7) Rigorous security.
(8) Simple back up/recovery	(8) complex backup/recovery.
(9) often single user	(9) multiple user

Responsibility of DBA :-

1. Makes the decision concerning the content of the database.
2. Plans the storage structure and access strategy
3. Provides the support to the users
4. Defines the security and integrity checks
5. Backup and recovery strategies
6. Monitoring the performance and responding to the changes in the requirements.

Architecture of Database Systems

In 1975 the American National Standards Institute Standards Planning and Requirements Committee (ANSI - SPARC) proposed three - level architecture identified three levels of abstractions.



These levels are sometimes referred to as schemas or views

1. External or User level:- This level describes the user's or application program's view of the database.

2. Conceptual level:- This level describes the organisation's view of all the data in the database, the relationship between the data and the constraints applicable to the database. This level describes a logical view of the database ie. a view lacking implementation detail.

3. The Internal or Physical Level: This level describes the way in which data is stored and the way in which data may be accessed. This level describes a Physical view of the database.

Each level is described in terms of schema — a map of the database.

Physical schema: The physical schema describes the physical storage structure of the database.

conceptual or logical schema:— The logical schema describes the structure of the whole database for a community of users. The conceptual schema describes the entities, data types, relationships, user operations and constraints.

Lecture - 3

Data Independence:

① **Logical data independency:-** It indicates that the conceptual schema can be changed without effecting the existing external schema.

② **Physical data independency:** It indicates that the physical storage structure or devices can be changed without effecting the conceptual schema.

Instances: The collection of information stored in the database at a particular moment is called an instance of a database

Schemas: The overall design of the database is called the database schema, ie, the description of a database is called database schema which is specified during database design and is not expected to change frequently.

Data Definition Language (DDL)

- Specification notation for defining the database schema

eg. create table account (
account_no char(10),
branch_name char(10),
balance integer);

The commands come under DDL are —

1. create
2. Alter
3. Remove
4. Drop.

It is a set of SQL commands used to create, modify and delete database structure but not the data. These commands are normally not used by a general user, who should be accessing the database via an application.

ER Model (Entity Relationship Model)

1. The entity-relationship model is based on a perception of the world as consisting of a collection of basic objects (entities) and relationships among these objects.

Notation used for ER Diagrams are –


Entity


Weak Entity


Relationship


attribute

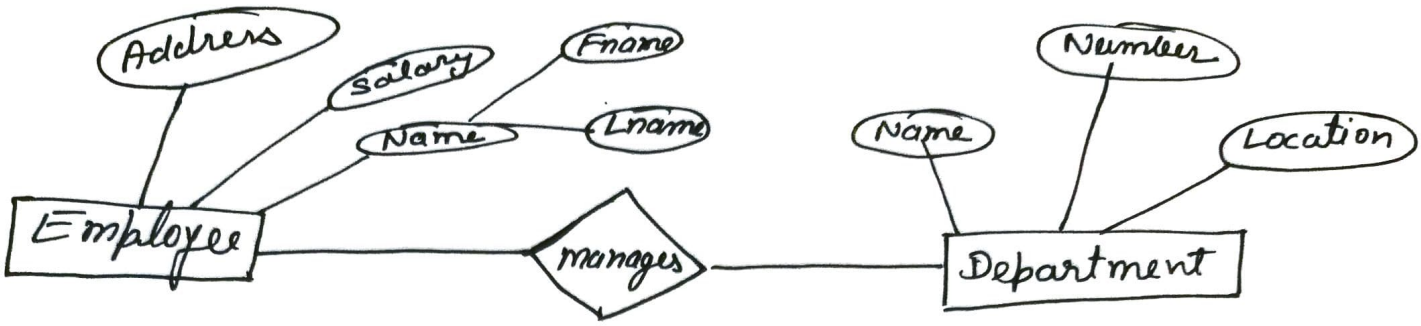

key attribute


Multivalued attribute


composite attribute

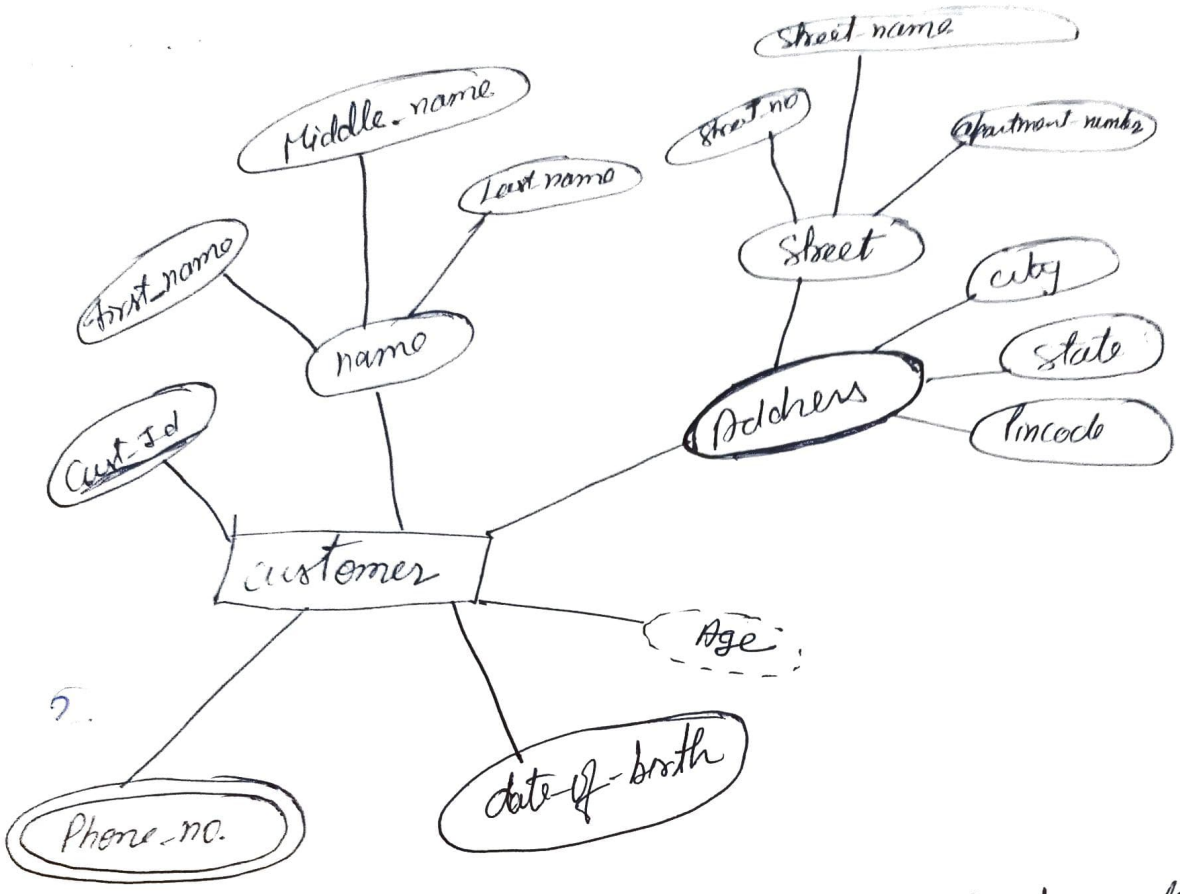

Derived attribute

eg.



Cardinality in ER-Diagram

- one to one (1:1)
- one to many (1:N) or Many to one (N:1)
- Many to many (M:N).



ER diagram with composite, multivalued & derived attributes.

key attribute →

composite " →

derived " →

multivalued " →

Database languages & Interfaces.

- Data Definition Language (DDL)

→ SQL commands used to create, modify, & delete database structure but not data

1. Create ~~table~~ - used to create database or its objects table, (Index, views etc.)
2. Drop - is used to delete objects from the database,
3. Alter - used to Alter the structure of the database
4. Truncate - used to remove all records from a table, including all spaces allocated for the records are removed.
5. Comment - used To comments to the data dictionary.
6. Rename - used to rename an object existing in the database.

Data Manipulation Language (DML)

It is used for accessing and manipulating the data organized by the appropriate data model.
DML also known as query language.

commands used in DML are —

1. Insert
2. Delete
3. Select
4. Update

Data control language (DCL): control the access permission of database &

commands are —

① Grant
② Revoke

TCL (Transaction control Language)

→ commit – commit a transaction
→ Rollback – Rollback a transaction in case of any error.
→

KEY

Entities and relationships are distinguishable using various keys —

Key: A key is a combination of one or more attribute, eg. social-security number, combination of name and social-security number.

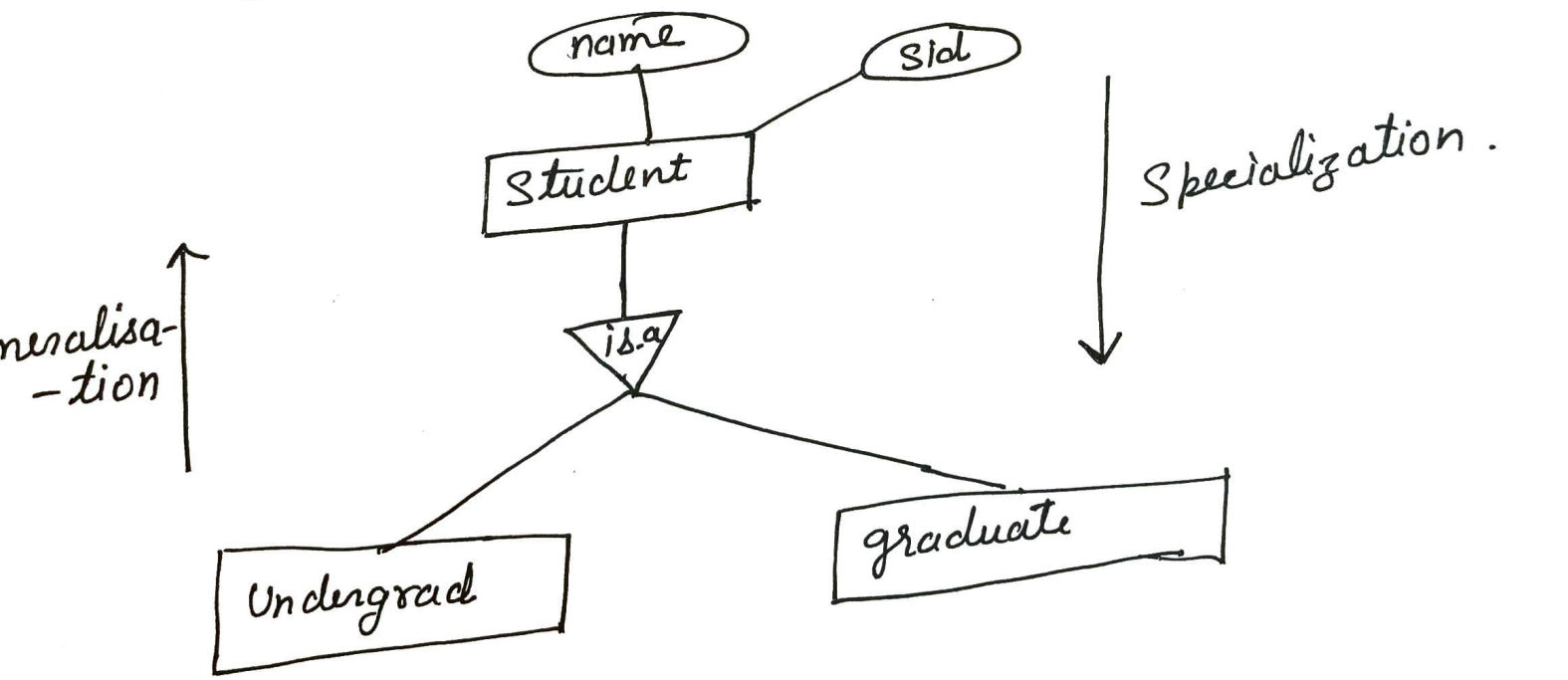
Super key: A super key is a key defined either for an entity set or relationship set that uniquely identifies an entity eg. social-security number, Phone number, combination of name and social-security number

Candidate key: A candidate key is a minimal super key that uniquely identifies either an entity or a relationship eg. social-security number, candidate key is a minimal superkey.

Primary key: A primary key is a candidate key that is chosen by the database designer to identify the entities of an entity set, Primary key always be unique & not null.

Generalization: is the result of computing the union of two or more entity sets to produce a higher-level entity set. It represents the containment relationship that exists between the higher-level entity set and one or more lower level entity sets.

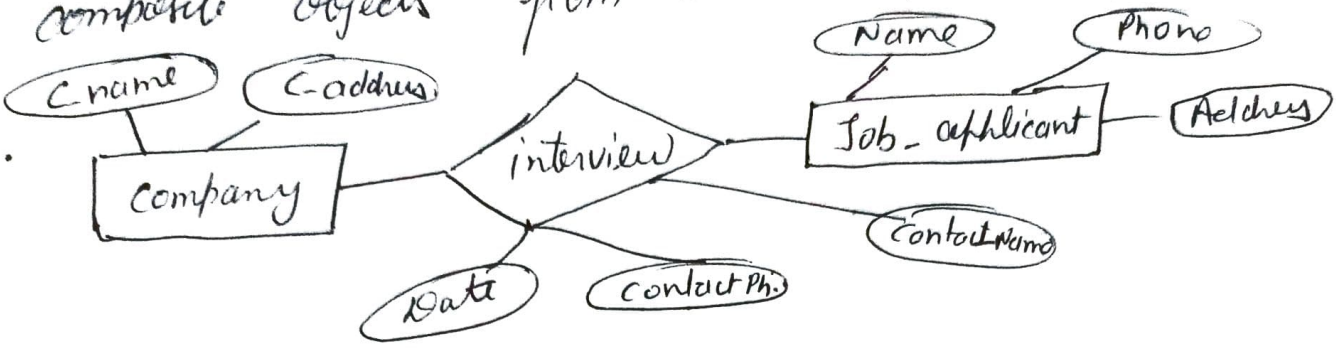
Specialisation: constructs the lower level entity sets that are a subset of a higher level entity set.



Aggregation

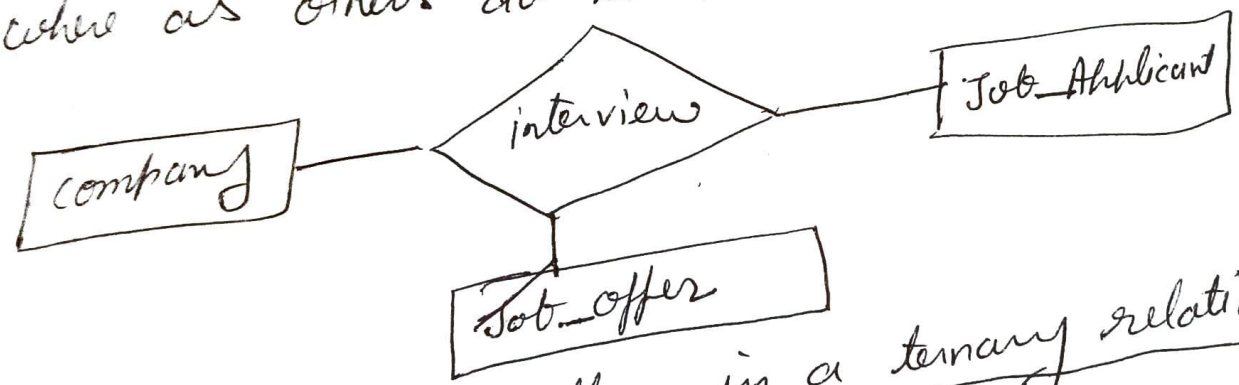
Aggregation is an abstraction concept for building composite objects from their component objects.

eg.



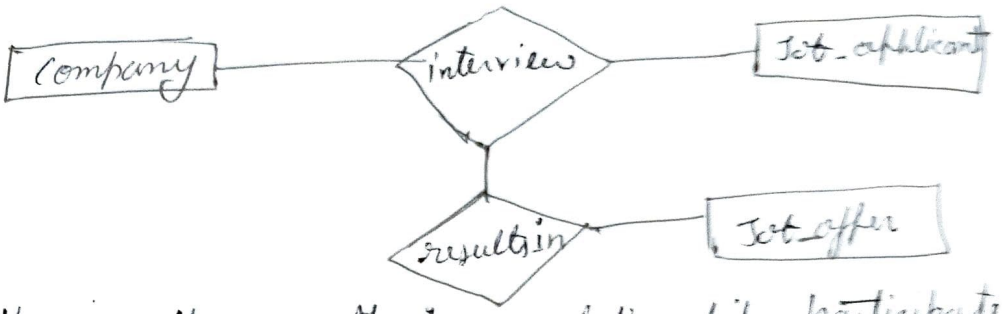
(a) The relationship type Interview.

Suppose that some interviews results in job offers, while as others do not.



(b) including Job_offer is a ternary relation type (incorrect)

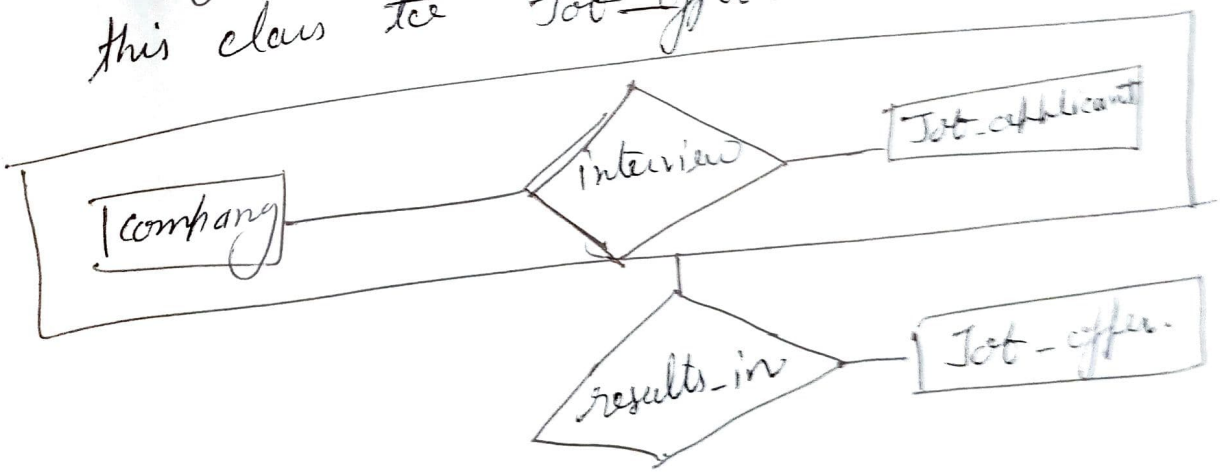
is incorrect because it requires each interview relationship instance to have a job offer.



(c) Having the results in relationship participate in other relationships (not allowed in ER).

incorrect because ER Model does not allow relationship among relationship.

One way to represent this situation is to create a higher-level aggregate class composed of Company, Job_applicant, and interview & relate this class to Job_offer.

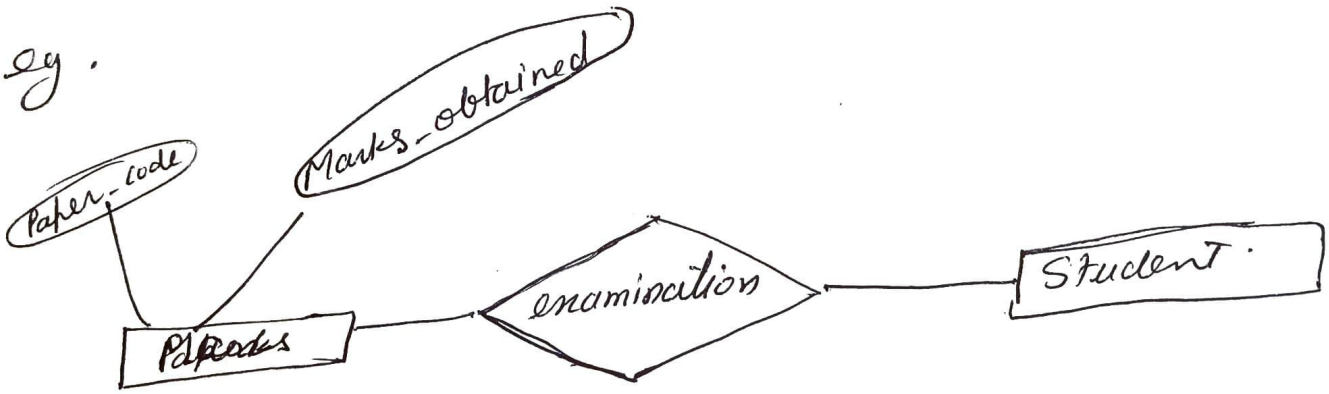


Reduction of ER diagram to tables.

converting a database representation from an E-R diagram to a table format is the basis for deriving a relational database design from ER Diagram.

Let A be an entity set with descriptive attributes $x_1, x_2 \ldots x_k$. We represent this entity by a table called A with k different columns.

eg.



The Marks table

Paper_code	Marks_obtained
MAS-401	70
NEC-402	75
NCS-401	80
NCS-402	85

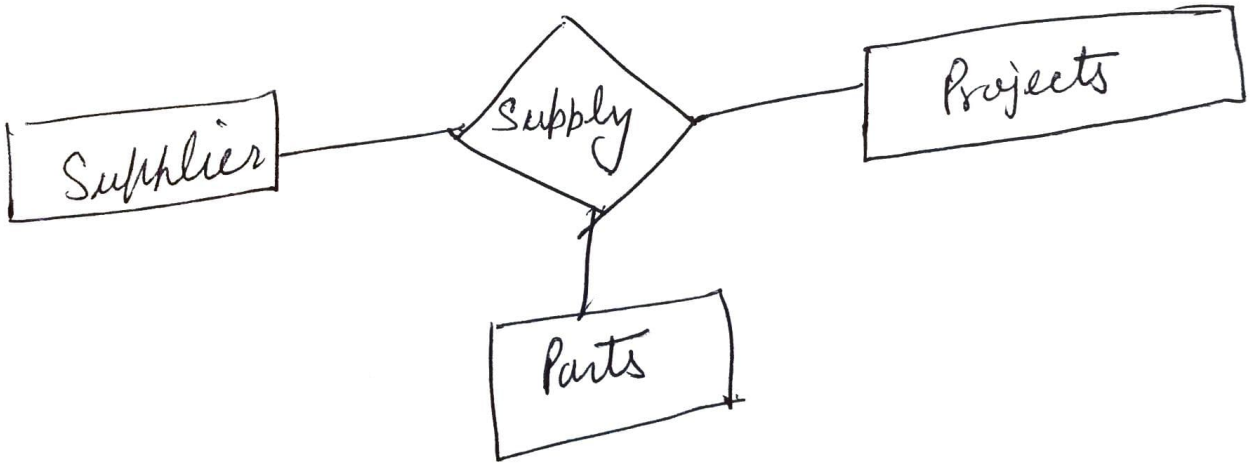
Relationship of higher degree

The degree of a relation type is the number of a relationship type participating entity sets.

A relationship type of degree two is called binary.

A relationship type of degree three is called ternary.

Binary relationship



Interesting Facts about Database Management Systems

- The term 'Database Management System' was coined in the late 1960s.
- Relational Databases - Edgar F. Codd, created the relational database model, that maintain data records in tables.
  oracle - was lauched in 1980
  MySQL,
  DB2,
  MS SQL Server

  MySQL
  Ms Acces.
- Post Relational databases —
  It comprised of all the NoSQL DBMS such as
  MongoDB, HBase, Cassandra, Redis and couchDB
- Traditional database management system —
  DB2, DBASE III, DBASE III+ FOXPRO etc.
- You can communicate with RDBMS using Structured Query Language (SQL).
- VDL (view definition Language)
  Specify user's view and their mapping to the conceptual schema.
- SDL (Storage Definition Language) - This language specify the internal schema. It may specify the mapping between two schemas.