

DBMS

Data : Data is the raw material that can be processed for any computing machine eg; employee name, product name, name of the student, marks of the student, any number, image

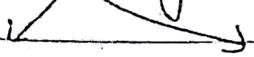
Information : It is the data that has been converted into more useful or intelligible form eg report card sheet.

Why we need information

- 1) To gain knowledge about the surroundings
- 2) To keep the system upto date
- 3) To know about the rules and regulation of the society :

Knowledge : Human mind purposefully organise the information and evaluate it to produce knowledge eg 238 is a data and Marks of student is information and the hard work required to get mark is knowledge

Knowledge



Fact Based Heuristic Based

- 1) Fact Based : The knowledge gain from fundamental through experiment
- 2) Heuristic Based : It is the knowledge of good practice

and good judgement like hypothesis

Difference between data and information

Data	Information
① Data is the raw fact	It is the processed form of data
② It is not significant to a business	It is significant to a business
③ Data are atomic level piece of information	It is a collection of data
④ Data does not help in decision making	It helps in decision making
⑤ eg: product name, name of student	eg: report card sheet

Database: The related information when placed in an organised form makes a database or an organised collection of related information is known as database eg Dictionary, Telephone Directory, Mobile contact

- Operations performed on Database
- 1) Insertion
 - 2) Updation
 - 3) Deletion
 - 4) Retrieval
 - 5) Sorting

Difference between Computerised databases and manual databases

Traditional File System:

File System: A file system is the method of storing and organising the computer files and the data they contain to make it easy to find and access them.

Characteristics of file system:

- 1) It is a group of files for storing the data of an organisation.
- 2) Each file is independent from one another.
- 3) Each file is called a flat file.
- 4) Files are designed by using the program written in programming languages such as C, C++.

Limitations / Disadvantages of File processing system:

- 1) Separated and isolated data.
- 2) Duplication of data: 1) It costs time and money.
2) It takes up additional storage space.
3) It can lead to loss of data integrity.
- 3) Data dependencies: files and records were described by specific physical format that were coded in the application program by the programs.
- 4) Difficulty in representing the data from the user point of view.
- 5) Data security: The security of data is low in the file based system because the data maintained in a flat file is easily accessible.

- Transactional problems: This system called
 1) Transactional properties A → Atomicity
 2) Durability
 3) Isolation
 4) Concurrency problems: When multiple user
 access a same piece of data at a
 same interval of time then it is called
 concurrency of system. When two or more
 user read the data simultaneously like to
 update the file simultaneously
 then it is called in a problem

Building Block of Database :-

- 1) Column / Fields
- 2) Rows / Tuple / Record
- 3) Table

DBMS (Data Base Management System): It
is the software system that allows the user to
define, to create and maintain the database
and provide controlled access to the data.

Application of Database :-

- 1) Library System
- 2) Banking System
- 3) ATM

Database: Mysql, Oracle, Sql Server, DB2, Microsoft Access

Components of DBMS :-

Hardware :- The hardware is the actual computer system used for keeping and accessing the database. Conventional DBMS hardware consists of secondary storage devices such as hard disk. Database runs on this range of machine from micro computers to main frames.

Software :- Software is the actual DBMS between the physical database and the users of the system. All the request from the user for accessing the database are handled by DBMS.

Data :-

Users :- 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 the database can be classified into the following groups.

Naive Users

Online Users

Sophisticated Users

Specialized Users

Application programmers

DBA - Database Administrator

- 1) General Application User: A user who uses a range of applications and interacts with the system through a menu and the keyboard. They type in the indicated keys or commands.

2) Online User: Those users who may communicate with the system directly through an online terminal or indirectly through their application program.

3) Sophisticated User: They are those users who interact with the system without writing the program. Instead, they use their own language in database query language.

4) Specialized User: Those users who write specialized database application that do not fit into the functional database processing framework.

5) Application Programmer: Those users who are responsible for developing the application programs or user interface. The application programs could be written in high-level languages.

6) DBA - Database Administrator: It is a person or the group in charge for implementing the database system within the organization. The DBA has all the privilege (all power) given by the DBMS and can assign or remove the privileges from the users.

Procedure :-

Disadvantages of DBMS :-

- 1) Complexity
- 2) Size
- 3) Performance
- 4) Higher impact of failure
- 5) Cost of DBMS

Differentiate between File Management System & DBMS.

- * **Master file :-** Master file are those file which remain static. These do not change.
- o **Transaction file :-** Transaction file are those file which is dynamic in nature. We can make changes.
- o **Instances :-** The situation data in the database at a particular moment of time is called an instance.
- o **Schema :-** The overall design of the database is called Schema or Description of database.
- o **Subschema :-** It is the subset of the schema and inherit the same property that a schema has. It gives the user a window through which he/she can view only that part of database which is of interest to him.

Architecture of DBMS :-

There is 3 level

- 1) External level
- 2) Conceptual Level
- 3) Internal Level

Objective of three Level Architecture or Space

3 level Architecture :-

The objective is to separate each user's view of the data from the way the database is physically represented.

There are several reasons

- 1) The internal structure of the database should be unaffected while changes to the physical aspects of storage.
- 2) The DBA should be able to change the conceptual structure of the database without affecting all other users.

1) External level | view level :- This level describes that part of the database that is relevant to each user. This level insulates the users from the details of conceptual and the internal level.

2) Conceptual level | logic level :- This level describes what data is stored into the database and the relationship among the data.

It represents :-

- (a) All the entities, attributes and their relationships
- (b) The constraints on the data

- .) security and integrity information /storage level
- ;) Internal level: It is the physical representation of the database on the computer. This level describes how the data is stored in the database. It covers the data structure and file organisation used to store the data on storage devices.

Schemas:

- .) External Schema
- ;) Conceptual Schema
- ;) Internal Schema

i) External Schema: The external view is described by means of a schema called External Schema that corresponds to different views of the data.

ii) Conceptual Schema: The conceptual view is defined by Conceptual Schema, which describes all the entities, attributes and their relationship with the integrity constraints.

;) Internal Schema: Internal level is defined by internal Schema, which is a complete description of the internal model.

There is only 1 Conceptual Schema and 1 internal Schema per database and more than 1 external Schema.

Schema is also known as Intension.

NOTE :-

Instance / Extension of database

Mapping between the levels :-

- 1) External | Conceptual Mapping
- 2) Conceptual | Internal Mapping

1) External | Conceptual Mapping :- Each external Schema is related to the Conceptual Schema by external Conceptual Mapping. This Mapping gives the Correspondence among the records and the relationships of the external & Conceptual Views. There is a mapping from a particular logical record in the external View to one or more Conceptual record in the Conceptual View.

2) Conceptual | Internal Mapping :- Conceptual Schema is related to Internal Schema by Conceptual Internal Mapping. Mapping between the Conceptual and Internal level Specify the method of deriving the Conceptual Record from physical database

Data Independence :-

- 1) Logical data independency
- 2) Physical data independency

1) Logical data independency :- It indicates that the Conceptual Schema can be changed without affecting the existing external Schema. The changes would be absorbed by the mapping

between external and conceptual level

- 2) Physical data independence: It indicates that the physical storage structure or devices can be changed without affecting the conceptual schema. The change would be absorbed by the conceptual internal mapping.

→ Logical data independence is much more difficult than physical data independence as it requires the flexibility in the design of the database and programmer has to see the future requirement & modification in the design

o Limitation of file processing system:

- 1) Separated and Isolated Data: To make a decision, a user might need data from two separate files. First the files were evaluated by analysts and programmers to determine the specific data required from each file and the relationship between the data and then application could be written in a programming language to process and extract the needed data.

- 2) Difficulty in representing data from the user's view: To create useful application for the user, often data from various files must be combined. In file processing, it was difficult to determine relationships between isolated data in order to meet user application.

Components of DBMS :-
Data :- It is the most important component of DBMS environment from the end users point of view.

One of the major features of database is that actual data are separated from the programs that use the data. A database should always be designed, built and populated for a particular audience and for a specific purpose.

Procedures :- Procedures refer to the instructions and rules that govern the design and use of the database. The users of the system and the staff that manage the database require documented procedures on how to use or run the system.

Disadvantages of DBMS :-

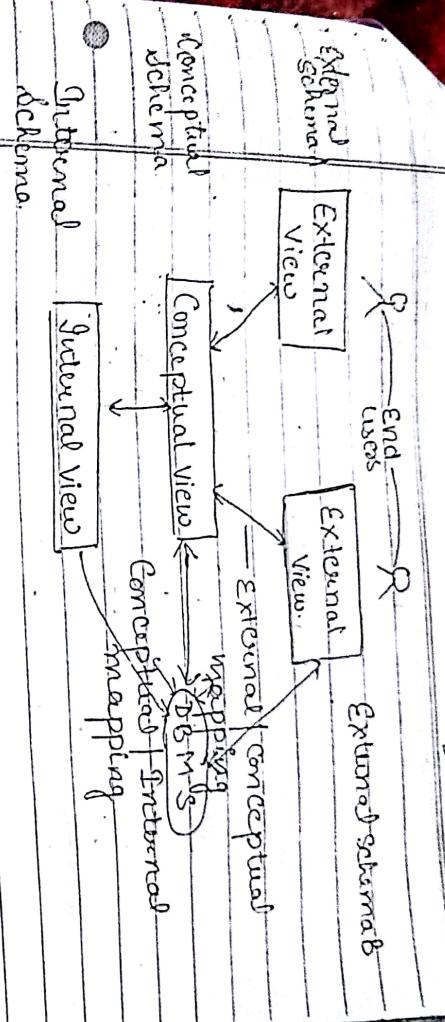
- 1) Complexity :- The provision of the functionality that is expected of a good DBMS makes the DBMS an extremely complex piece of software. Database designers, developers, database administrators and end-users must understand this functionality to take full advantage of it. Failure to understand the system can lead to bad design decisions, which can have serious consequences for an organization.

- 3) Performance: A file Based System is written for a specific application such as invoicing. As a result performance is generally very good. However the DBMS is written to be more general, to cater for many applications rather than just one.
- 4) Higher impact of a failure: The centralization of resources increases the vulnerability of the system. Since all user and applications rely on the availability of the DBMS, the failure of any component can bring operation to a halt.

File Management eg C++

or C or C# Program

- | <u>File Management eg C++</u> | <u>Database Management</u> |
|---------------------------------------------|-----------------------------------------|
| 1) Small System | 1) Large System |
| 2) Relatively cheap | 2) relatively expensive |
| 3) few files | 3) many files |
| 4) files are files | 4) files not necessarily files |
| 5) Simple structure | 5) complex structure |
| 6) little preliminary design | 6) vast, preliminary design. |
| 7) integrity left to application programmer | 7) no option inbuilt integrity checking |
| 8) no security | 8) numerous security |
| 9) simple, primitive | 9) complex & sophisticated |
| 10) backup/recovery | 10) backup/recovery |
| 11) often single user | 11) multiple users |



Internal schema.

Role of DBA (Database Administrator): It is a person or group in charge for implementing DBMS in an organisation. The DBA job requires high degree of technical expertise. DBA consist of a team of people rather than just one person.

Responsibilities 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) Interprets backup and recovery statements.
- 6) Monitoring the performance and responding to the changes in the requirements.

Data Dictionary or Meta Data :-

- It is of two type
1) Active Data Dictionary.
2) Passive Data Dictionary.

A Meta Data is the data about the data. It is the self describing nature of database. It also holds the following information about each data element in the database such as names, types, range of values, access authorization, indicate which application uses the data.

Meta Data is used by the developers to develop the program, queries to manage and manipulate the data.

Dictionary:-

- 1) Active Data : It is managed automatically by the data management Software. It is always consistent with the current structure of the Database.
- 2) Passive Data Dictionary : It is the one use for documentation purposes. It is managed by the user of the system and is modified manually by the users.

Database Languages :-

- ① Data Definition Language (DDL) :- It is a language that allows the user to define the data and their relationship to

Other type of data. Command are
Create
Alter
Rename
Drop

Data Manipulation language (DML) : It is a language that provides a set of operations to support the basic data manipulation operation on the data held in the database.

Command used are :

Insert

Delete .

Select ..

Update

Data Control language (DCL) : Command are
GRANT
REVOKE

Data Models : There are three type of data Models.

Object based

Record based logical models

Physical Based

It can be define as an integrated collection of concepts for describing and manipulating the data, relationship between the data and constraints on the data in an organization.

It Comprises of three component

Structural part : There are rule which help

- 2) Manipulative part : which type of operation is apply on model.
- 3) Integrity Rules :

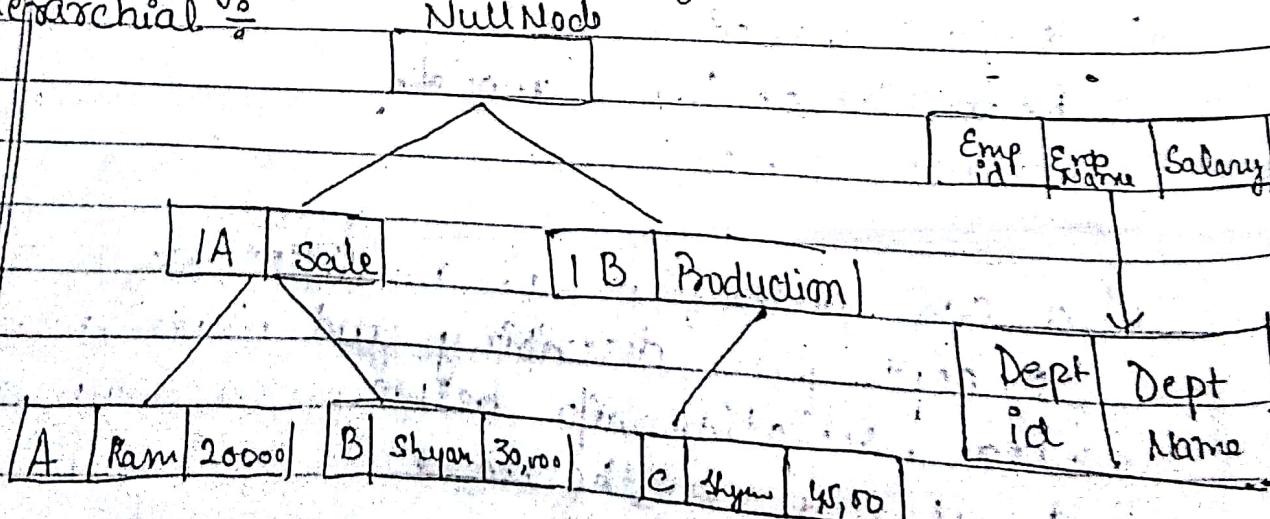
Data model are divided into three category
object based : It uses a concept such as entities, attribute and their relationship. This model can be used to describe the data at the conceptual and external level.
Eg E-R models

2) Physical based : These model describe how the data is stored in the computer. This model is used to describe the data at the internal level.

3) Record based logical models : These models are used in describing the data at the logical and the view level.

These models are used to specify the overall logical structure of the database

Eg of Hierarchical :



1) Hierarchical model:

- i) It is based on tree structure. It consists of collection of records that are connected to each other by links. The tree structure like in a hierarchical model is known as rooted tree. The root node of that tree is an empty node. So Hierarchical model is a collection of rooted trees and the relationship exists in the Hierarchical model is one-to many and many-to-one.

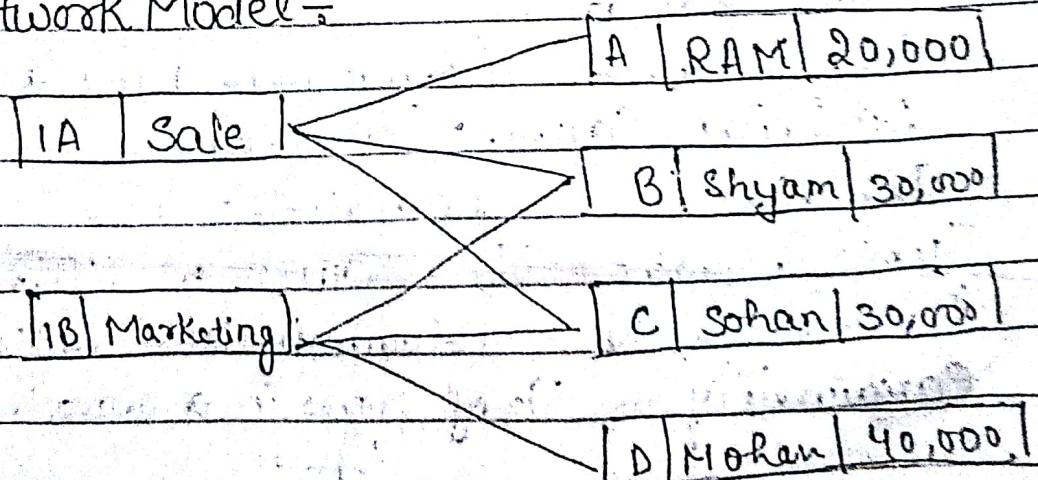
Advantages:

- 1) It is easy to understand.
- 2) More efficient than ER model

Disadvantage:

- 1) Data inconsistency occurs when the parent node is deleted that result in the deletion of the child node.
- 2) Wastage of storage space.
- 3) Complex to design.
- 4) Absence of structural independency.

2) Network Model:



For one-to-many →

One-to-one →

Date

It is based on graph structure. It consists of collection of records, which are connected to each other by links.

Advantages:

- It is easy to design than the Hierarchical model.
- Data access is easy in the network model.

Disadvantages:

- It is complex to design than a relational model.
- Efficiency are less than the relational model.
- Absence of structural independency.

3) Relational Model: Relational model stores data in form of tables.

Employee

Empid	Name	Salary
1	A	20,000
2	B	30,000
3	C	40,000
4	D	50,000

Department

Depid	Name
A	Sales
B	Marketing

↓
Tuple Variable

Table → Relation

Tuple → Rows (Each Row of data)

Attribute → Column. (Each column in the tuple)

Domain → Set of permitted values

Tuples Variable → Any value of a Tuple

Degree → No of Column in a relation

Cardinality → No of rows in a relation. → 3 in emp. (Tuples)

In Relational model many to many relationship can be easily implemented.

It is useful for representing most of the real world object and relationship among them.

Relational model does not maintain physical connection among records. Data is organized logically in the form of rows and columns and stored in tables.

UNIT - 2

Data Modelling Using ER Models :-

→ Enterprise :-

Entity type :- It is name, thing etc. These are the data object about which information is collected.

Attribute :- Prop characteristic of entity or Data elements & Data fields.

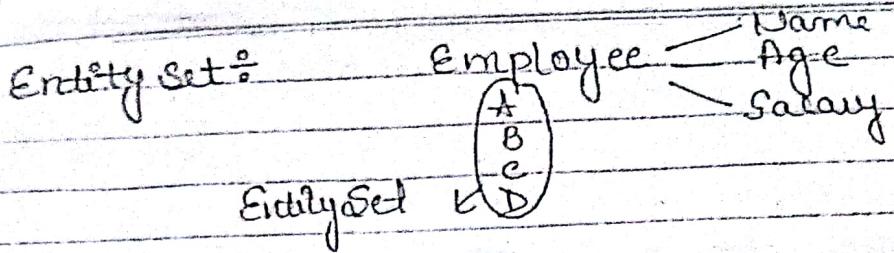
Type of Attributes :-

- 1) Single Value Attribute :- Those attribute which contain a single value. For eg. Age, Salary etc.
- 2) Multivalued Attribute :- That contain more than one value ^{of a single entity} for eg. Phone no.
- 3) Composite Attribute :- Those attribute which can be further divided. for eg. name → First name, Last name, Date of Birth etc.
- 4) Simple or Atomic attributes :- Those attributes which can not be further divided. for eg. Age.
- 5) Stored Attributes :- Attributes which can ^{not} be derived from another attribute. Date of Birth.
- 6) Null values :- Empty.

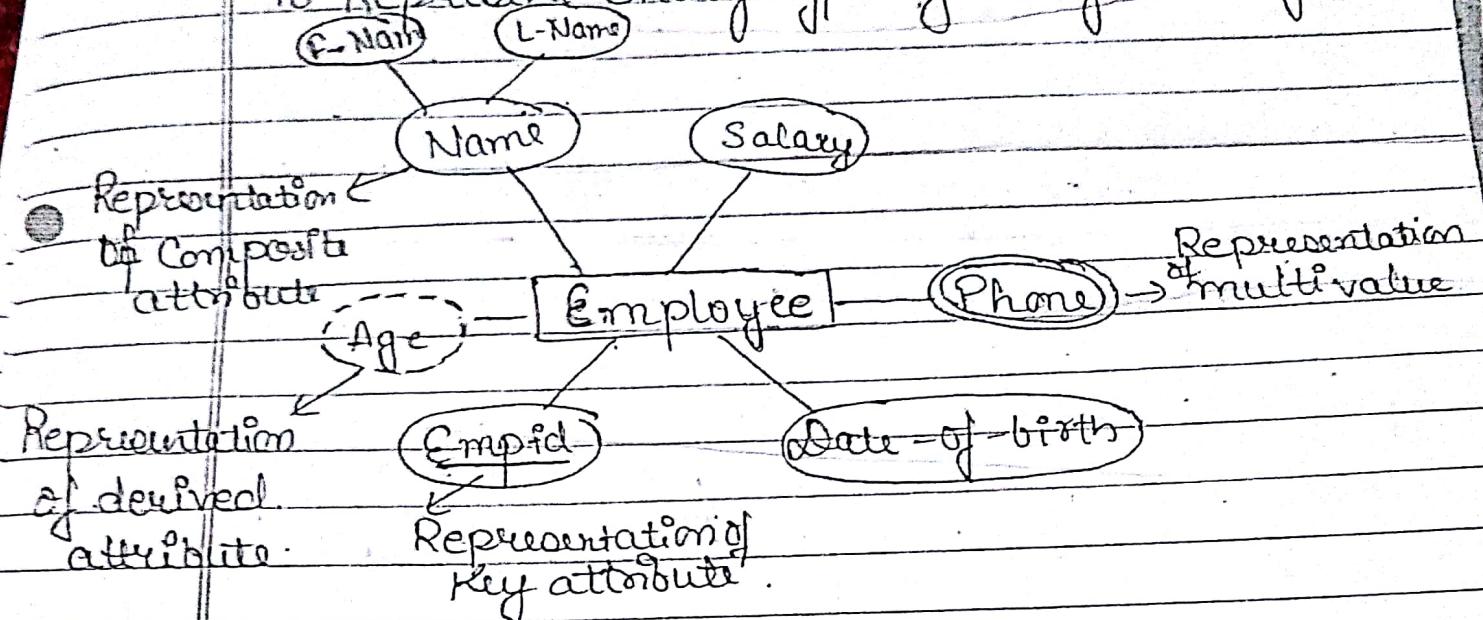
Entity Types :- Collection of Entity that Share the Same attribute.

e.g. Employee → Name
→ Age

Entity type. Salary



To Represent Entity Type by Using ER diagram

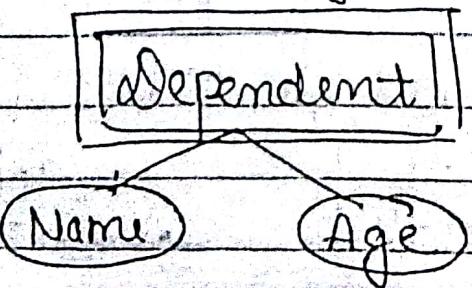


Key attribute: Key attribute are those attribute which is uniquely identify the record.

Weak Entity type: Those entity type in which no key attributes is present

Strong Entity type: Those strong Entity type which consist of key attributes.

Eg of weak Entity type



Set of possible

Domain of the attributes?

Values

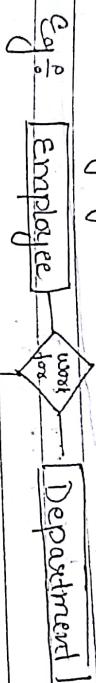
Representation of Relationship

Job has two Relationship types
Employee → Relationship type



Degree of relationship type? How much entity are involved in a relationship

- 1) Binary type → 2 entity are related
- 2) Ternary type → 3 entity are related



Relationship Constraints:

↳ Participation Constraints

↳ cardinality ratio

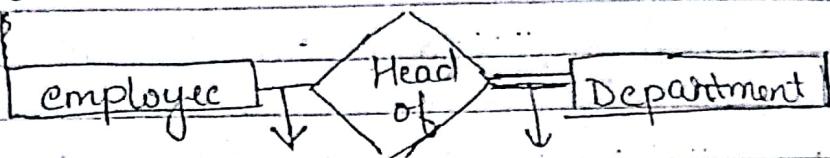
① Participation Constraints

Total participation Partial Participation

- ① Total Participation: In total participation every entity in the empty set must be dependent on another entity. It is also known as existence dependency.

In E-R diagram it is represented as a double line connecting the participating entity type two - the relationship.

- ② Partial participation In partial participation some entities in the entity set are dependent upon another entity



Partial participation.

Totally participated

- ③ Cardinality relation : Cardinality ratio for binary relationship specifies the no. of relationship instances that an entity can participate in a relation set.

Relationship exists are :-

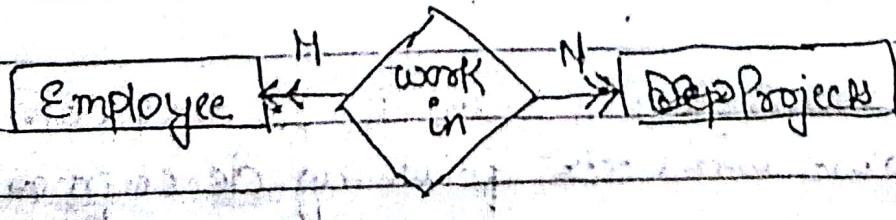
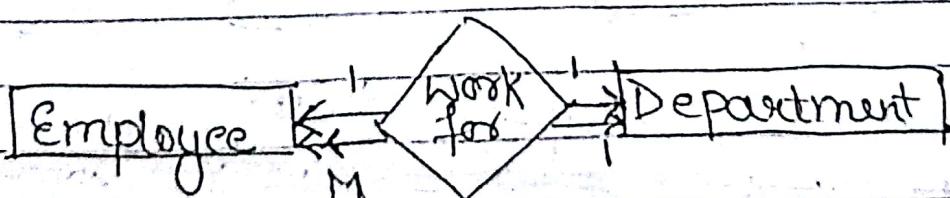
One to One $\rightarrow (1:1)$

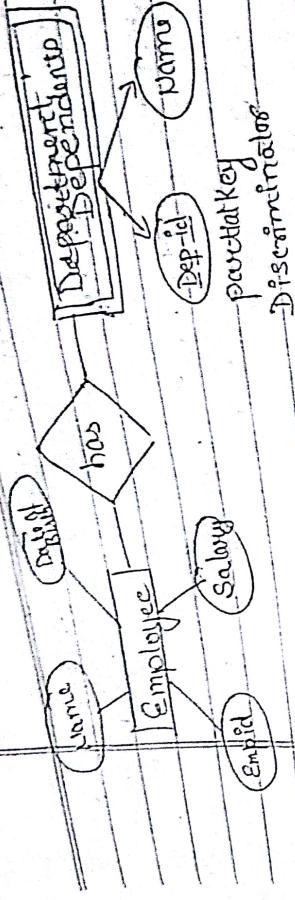
One to many $\rightarrow (1:N)$

Many to one $\rightarrow (N:1)$

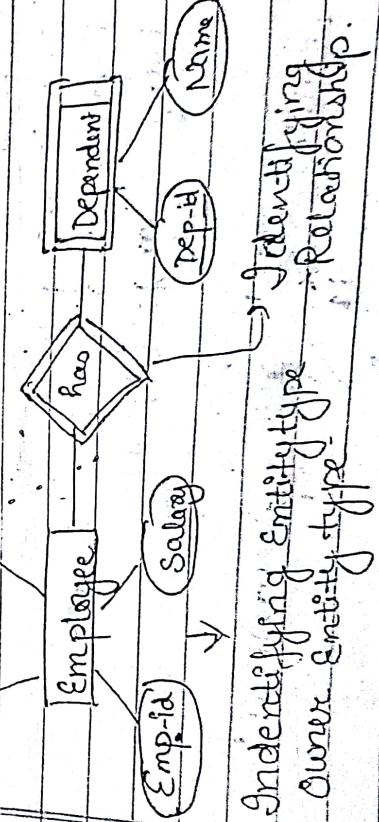
Many to Many $\rightarrow (N:N)$

e.g:-





→ Identifying Relationship :- He knows that a relationship type due to weak entity type do what we do to handle such entity type with some other entity type. If we break entity type in combination with another entity type in combination with some of their attributes value. we call this other entity type the identifying or the owner entity type and we call the relationship type that relates to its owner the identifying relationship.



Identifying Entity-type
Owner Entity-type
Relationship

→ Identifying Entity-type
Relationship

Q Make an E-R diagram of data base with the following description

(1) The Company is organised into departments. Each department has a unique name and unique no. A department may have several locations.

(2) A department controls a no of projects, each of which has a unique name, Unique no, and a single location.

(3) We store each employee name, Social security no, address and salary. An employee is assigned to one department but may work on several projects which are not necessarily controlled by the same department. We want to keep track of the dependence of each employee for insurance purposes.

