UNIT - 2

Data Modelling Using ER & Models :-

→ Enterprise :-

Entity type :- It is name, thing etc.; These are the data Object about which information is to be collected

Attributes :- characteristic of entity OR Data elements & Data fields.

Type of Attributes :-

1) Single Value Attributes :- 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 Attributes :- 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 Age Date of Birth / Date of Birth

6) Null value :- Empty.

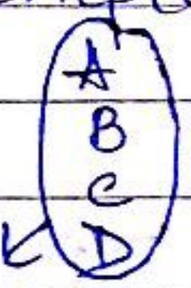
Entity Types :- Collection of Entity that share the same attribute.

eg :- Employee → Name, Age, Salary.

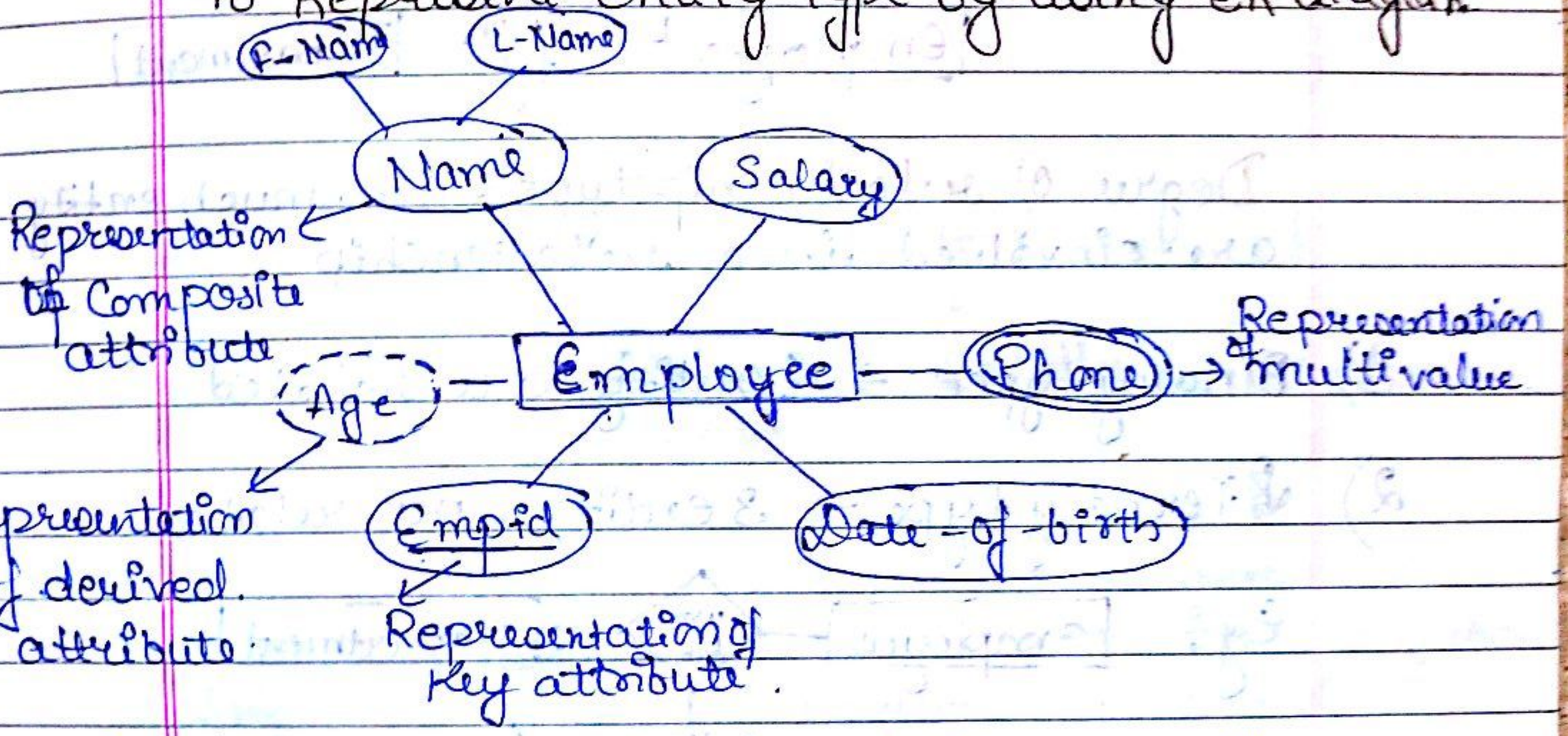
Entity type

Passion
Page No.
Date

Entity Set :- Employee — Name, Age, Salary

EntitySet 

To Represent Entity Type by using ER diagram

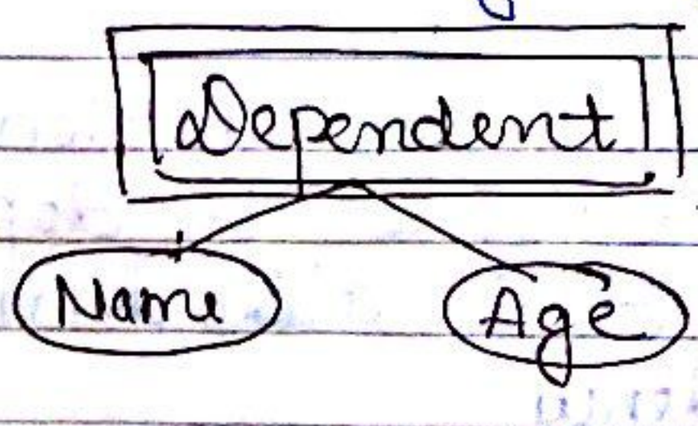


→ Key attribute :- Key attribute are those attribute which is uniquely identify the record.

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

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

Eg of weak Entity type



Domain of the attributes :- Set of possible values.

To join two entities → Relationship :- Representation of Relationship

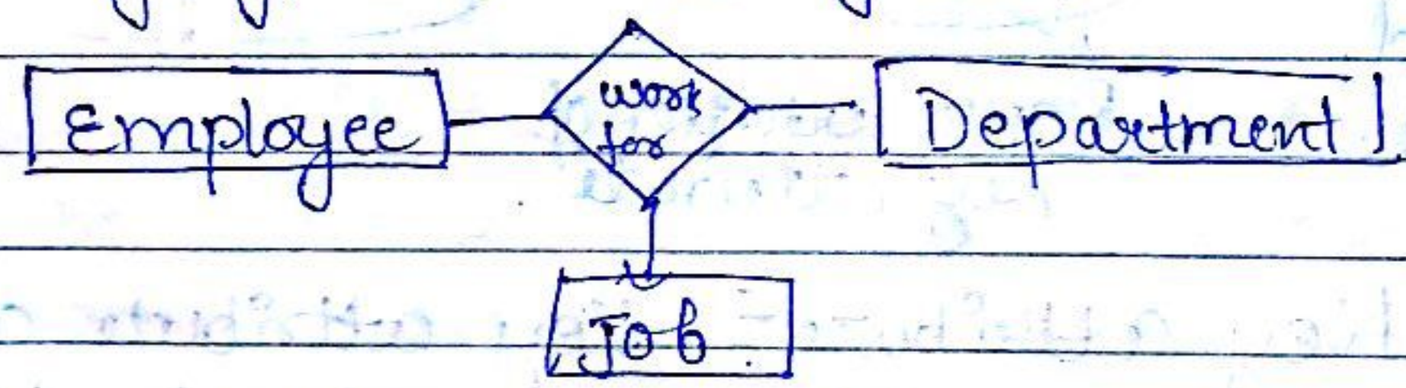
Relationship type :-

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

Eg :-



Relationship Constraints :-
↳ Participation constraints
↳ Cardinality ratio

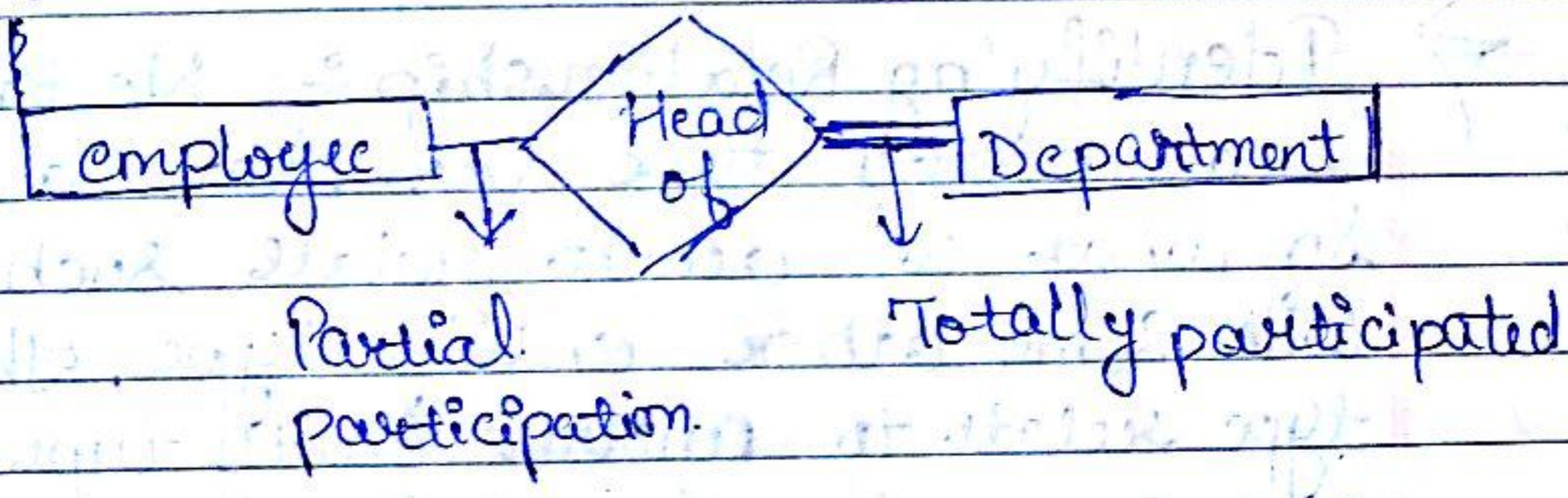
① Participation Constraints

Total participation Partial Participation

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

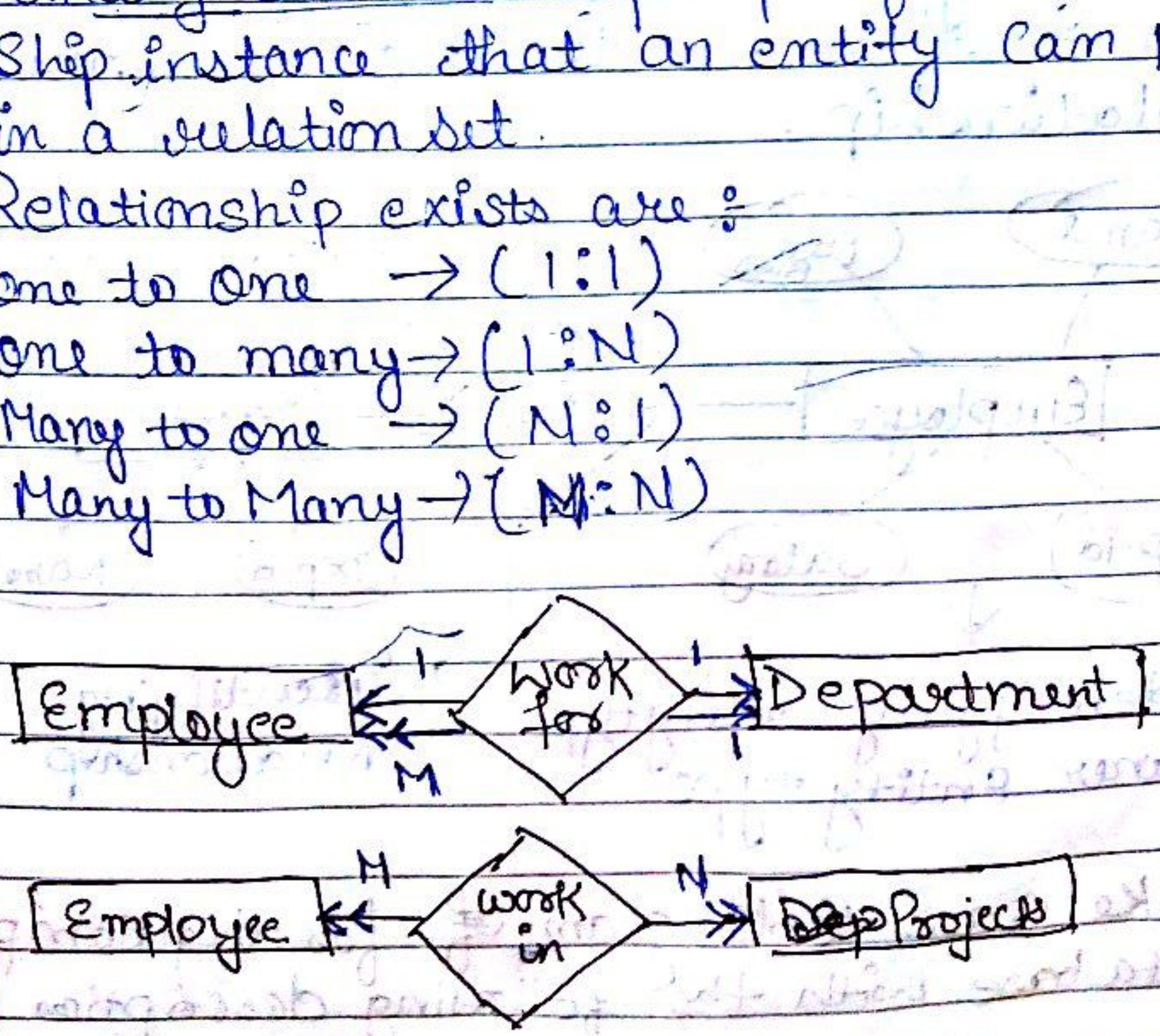
In E-R diagram it is represented as a double line connecting the participating Entity type two the relationship

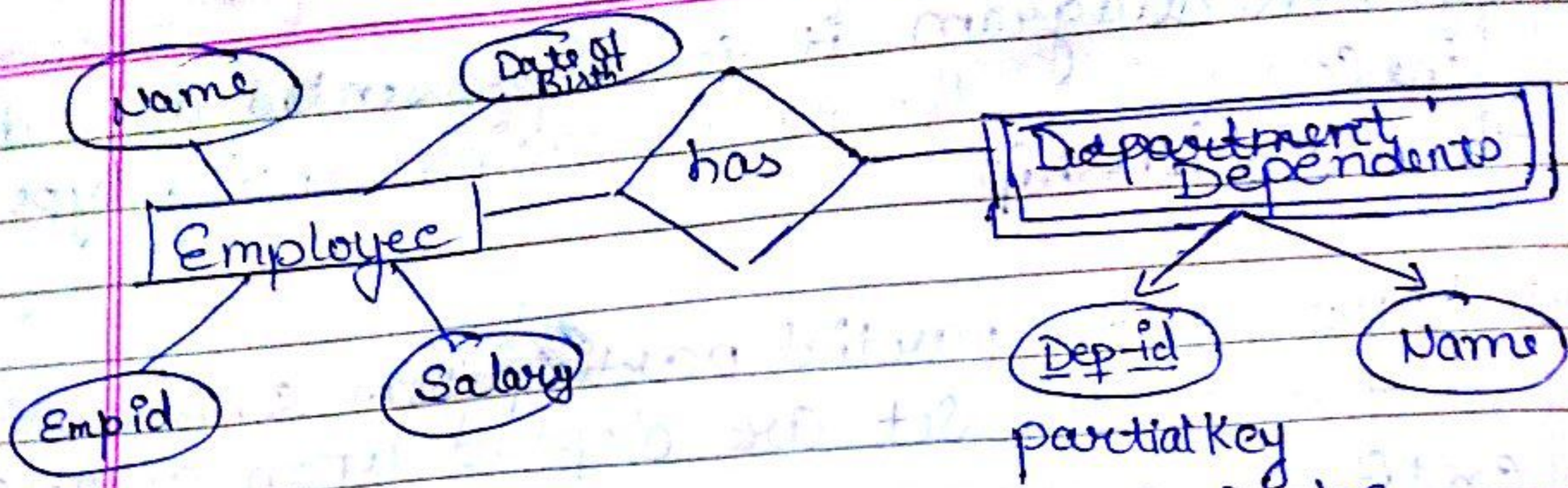
① Partial participation: In partial participation some entities in the entity set are depend upon another entity



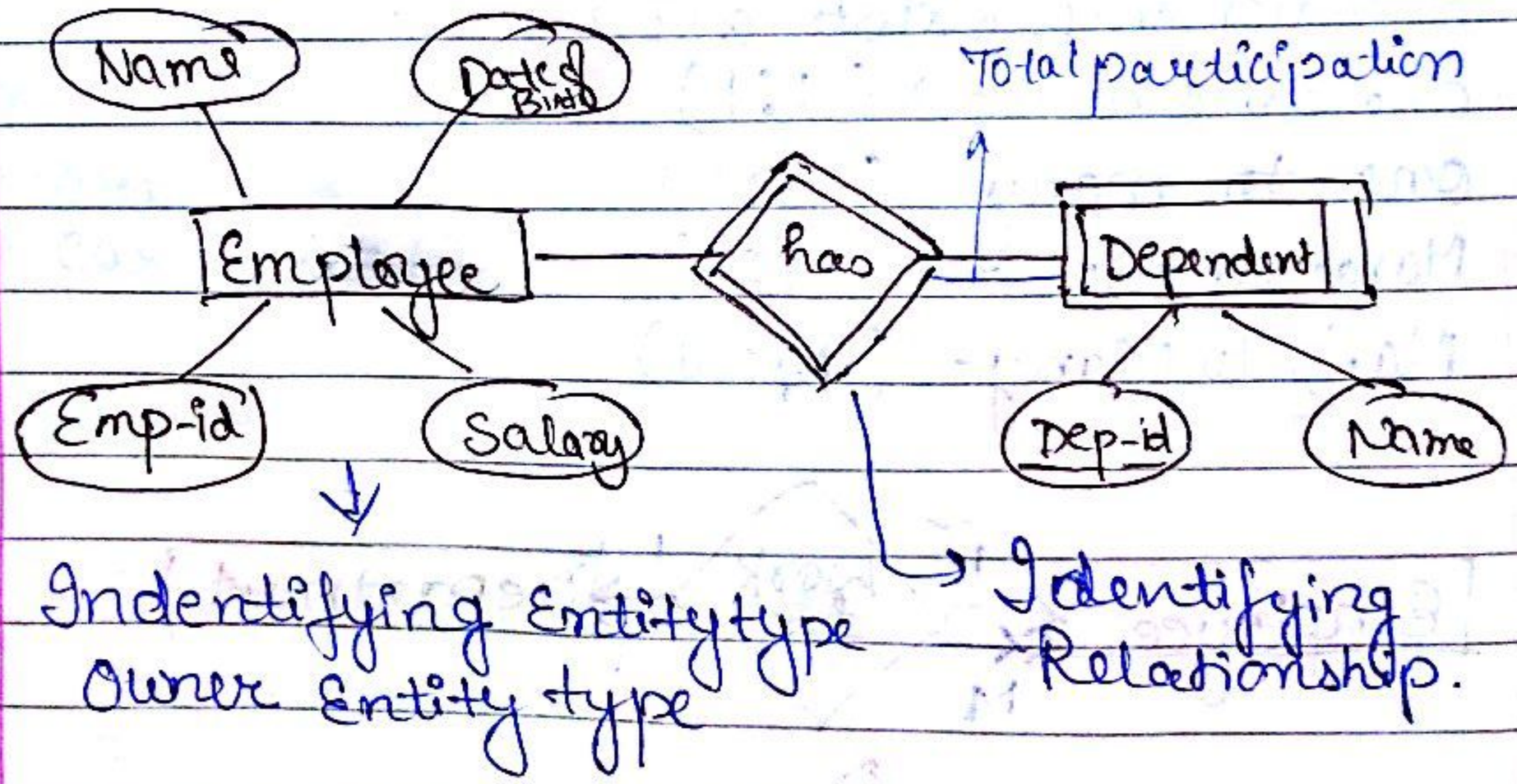
② Cardinality ration :- Cardinality ratio for binary relationship specifies the no. of relation ship instance that an entity can participate in a relation set.
Relationship exists are :-
one to one → (1:1)
one to many → (1:N)
Many to one → (N:1)
Many to Many → (M:N)

eg :-





→ Identifying Relationship :- We know that a weak entity type does have a key attribute. So what we do to relate such entity type with some other entity type. The weak entity type relate to 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.



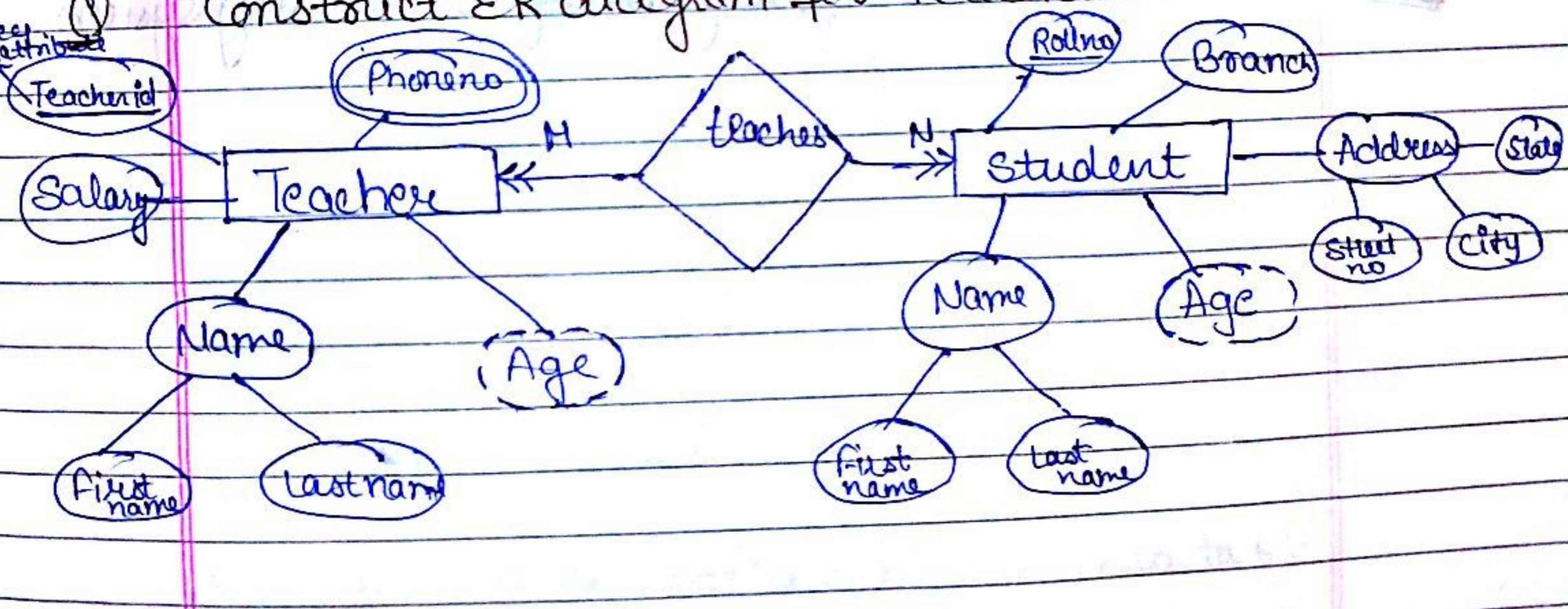
9 Make an E-R diagram of for the company database with the following description

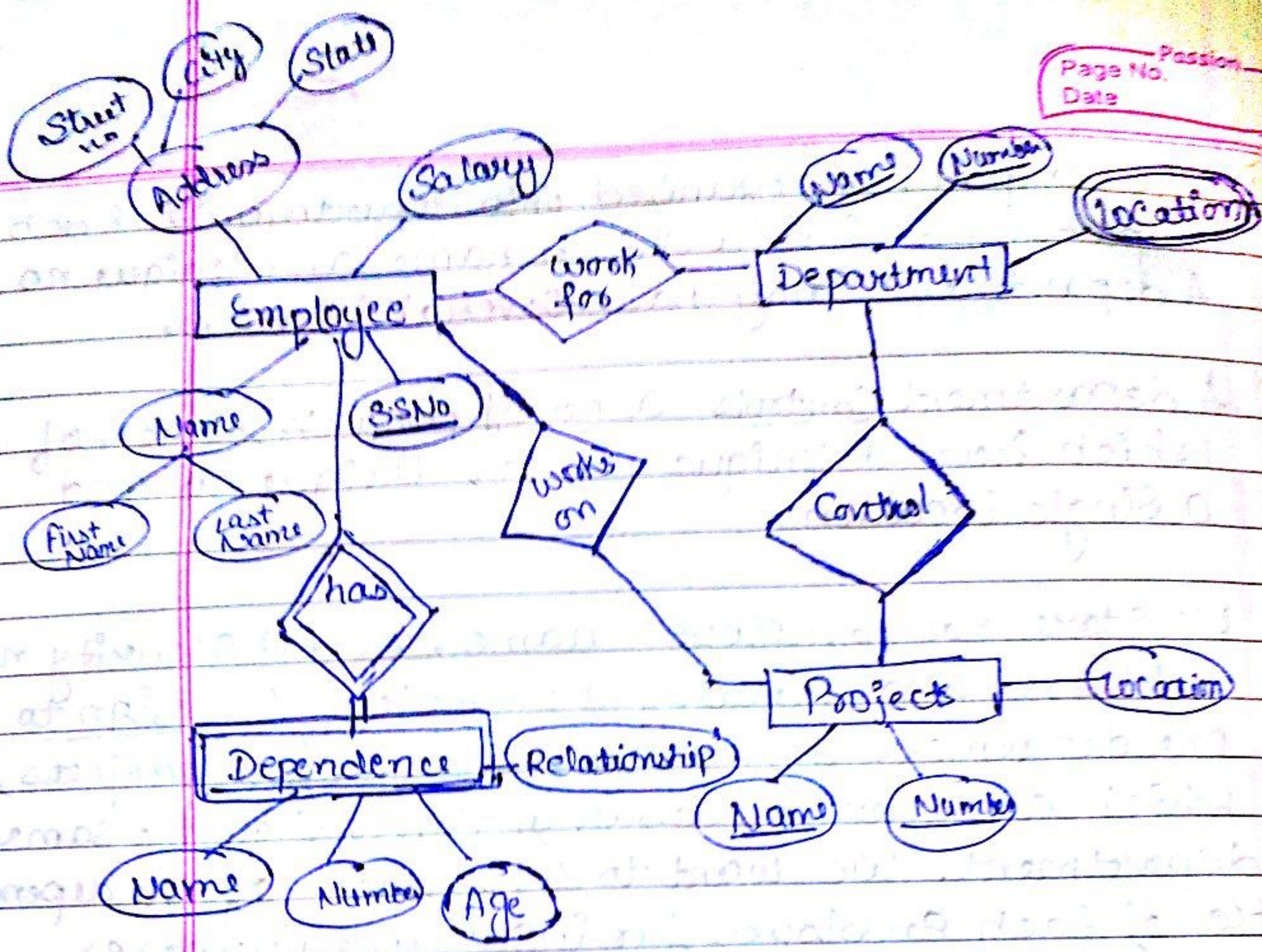
(1) The company is organised into departments. Each department has a unique name and unique no. A department may have several locations

(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 assign to One deparment but may work on several projects, which are not necessarily Contool by the same department. We want to Keep track of the dependen -ce of Each Employee for insurance purposes.

Q Construct ER diagram for Teacher Student database



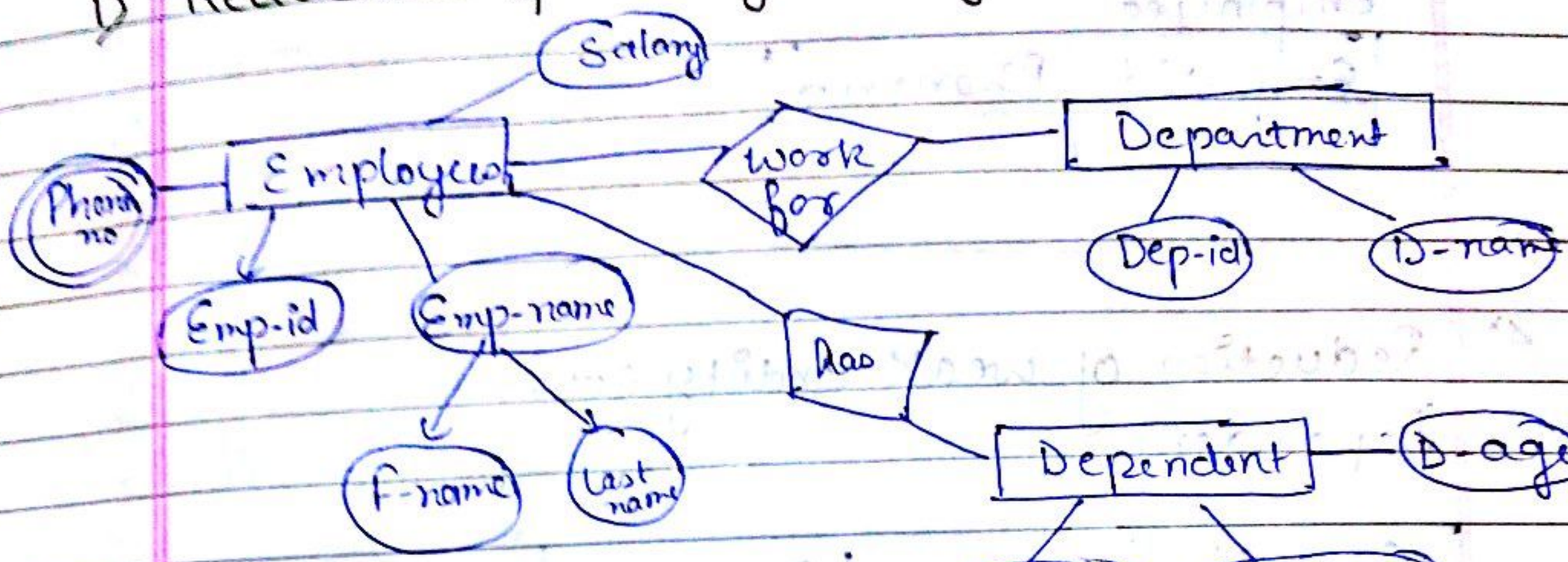


Page No. Passion
Date

Que Construct an E-R diagram of University System

Reduction of E-R diagram into table

1) Reduction of strong entity set into table



Key attributes eg.

Employee

Emp id	E-name	Salary

Department

Dep-id	Dept name

Reduction of composite attribute into table

Employee

Emp-id	f-name	l-name	Salary

3) Reduction of multivalued attributes:
Phoneno.

Employee

Emp-id	Phoneno

4) Reduction of weak entity set:
Dependent

Emp-id	Depno	Depname	Depage

5) Reduction of relationship sets.
Work for

Emp-id	Department id

RDBMS:-
3 Components:-
- Data Structure
- Data integrity
- Data Manipulation

Candidate key is a subset of superkey.

Key:- It is a set of one or more columns whose combined values are unique among all the

Page No.
Date

(1) ~~Candidate Key~~ :
occurances in a given table

Types of keys :-
1) Candidate Key
2) Super key
3) Primary key
4) foreign key
5) Alternate key
6) Artificial key
7) Composite key

1) Candidate Key :- They are those attributes of the relational, which have the properties of uniqueness and irreducibility.

a) Uniqueness :- No legal value of R ever contain two distinct tuples with same values for R.

b) Irreducibility :- No proper subset of K has the uniqueness property.

For example if the Combination of (Name, Class) is unique, then it can be identified as the candidate key if and only if Name and class individually are not unique.

2) Super Key :- Super key follow the properdy of Uniqueness, but not irreducibility
A Super key has a uniqueness property but not necessarily the irreducibility property
A candidate key is a special case of a Super key

eg:- If Roll-no is unique in relation STUDENT then the set of attribute (Roll-no, Name, Class) is a super key for a relation STUDENT, these set of attributes are also unique, but this combination of Key is not having the property of irreducibility

→ Relation of patient in which Patient-number is unique. The patient-numbers is a candidate key and (Patient-number, Patient name) is a Superkey.
'A superset of a candidate key is a Superkey.

3) Primary Key:- The primary key is an attribute or a set of attributes that uniquely identify a specific instance of an entity. Primary key cannot contain any null value because we cannot uniquely identify multiple null values.

4) Alternate Key:- Exactly one of those candidate keys is chosen as the primary key and the remainder, if any, are then called Alternate Keys. An alternate key is a function of all candidate key minus primary key.

5) Composite keys:- A primary key that is made up of more than one attribute is known as Composite key.

Ex:- WORK

Employee ID	Project ID	Hours-Worked
01	01	200
01	02	120
02	01	50
02	03	120
03	03	160
03	04	200

It is Subset of superkey

6) Artifical keys:- An artificial key is one that has no meaning to the business or organization. Artificial key are permitted when
(1) no attribute has all the primary key properties
(2) the primary key is large and complex

eg:-

Enrollment		
Student	Class	row-id
AK		PK

7) Foreign Key:- Foreign key are the attributes of the table which refers to primary key of some another table. Foreign key are used to link together two or more different tables which have some form of relationship with each other. These foreign keys is a reference to the tuple of a table from which it was taken, this tuple being called as Referenced or target tuple.

Target Attribute

Foreign key

Department → Target table

eg:- Employeee

Emp-id	Name	Salary	Dep-id
1	Abc	30000	10
2	Xyz	40000	20
3	Pqr	50000	30

Dep-id	Dep-Name
10	Sales
20	Market
30	Product

eg:- Student

Rno	Name	Class-code
1	A	2
2	B	1
3	C	-

Class.

Classcode	Name
1	B.TECH
2	B.TECH
3	BBA

(2) Data integrity :- Basically it consist of two rules

1) Entity integrity Rule
2) Referential Integrity Rule

(1) Entity Integrity Rule :- This rules states that in a relation, the value of the attribute of a primary key cannot be null.

(2) Referential Integrity :- It states that if a foreign key exists in a relation, either the foreign key value must match the primary key value of some tuple in its home relation or the foreign key value must be completely null.

Passion
Page No.
Date

Employee

Emp-id	Name	Salary	Dep-id
1	A	3000	10
2	B	4000	20
3	C	5000	30
4	D	6000	40

Department

Dep-id	Dept name
1	Sales
2	Market
3	Product

either Should match or null.

Enterprise Constraints :-

Codd's Rule :- Founder of RDBMS is Dr. E.F Codd
There are 12 rule are as following :-

(1) Information rule :- All information represented in form of tables

(2) Guaranteed Access rule :- To Access any info we use key

(3) Comprehensive data Sublanguage rule :- There should be a particular language to support RDBMS rule.

(4) View updating rules :- If we change any data in Database then it should be automatically updated in all records or tables

(5) High level Insert, update & Delete :- The language we used it should contain these rule insert, update & Delete and perform all function.

(6) Physical data independency :- Change in lower level not effect higher lvel and this rule support

(7) Logical Data idependency :- Change in conceptual level and doesnot effect external level.

(8) Integrity Independency :- RDBMS support all Integrity rule.

(9) Non Subversion rule :- Any language we use to access the database and, that language will support our Integrity independency.

(10) Systematic treatment of null Value :- There should be special treatment of null value.

(11) Data Description rule :- Data we describe it shall be in form of table

(12) Distribution Independence :- It should be platform independent

Data Manipulation :-
(1) Relational Algebra
(2) Relational calculus.

Q Difference b/w Relational Algebra & Relational Calculus.

	Relational Algebra	Relational Calculus.
(1)	It is procedural language In this we follow the procedure & step	(1) It is a non procedural language In this we automatically get output
(2)	We can combine 2 or more table to get an another table	

Both are non user friendly.

Relation Algebra :- A (+) B

Relational operator :- Type of Relational operator
(1) Traditional Set Operator
(2) Special operator

(1) Traditional Set operators :- → Union
→ Intersection
→ Difference
→ Cartesian Product

Union :- Combine two table

Employee

Emp-id	Name	Dept-id

Dep-id

Dep-id	Name

→ Intersection:- Common element

→ Difference:- The Difference between two set S1 & S2 produces a set, which contain all the members of one set, which are not in the other.

R

Cust.name	Cust-Status
Ram	Good
Shyam	Excellent

S

Cust-name	Cust-Status
Karan	Average
Ram	Good

(1) R U S

Cust.name	Cust.Status
Ram	Good
Shyam	Excellent
Karan	Average

(2) R ∩ S

Cust.name	Cust.Status
Shyam	
Ram	Good

(3) R - S ⇒ Shyam.

(4) R x S = { (Ram, Karan) (R R), (S K), (S R) }

ii) Special operators:- → Selection
→ Projection
→ Joins
→ Division.

(1) Selection operation:- It yields a horizontal subset of a given relation that is that subset of row should be selected with in the given relation for which a predicate

Page No. Passion
Date

Condition is satisfied. Sign of selection is (σ)

Employee :-

Emp-id	Name	Salary

For row: σ salary > 10,000 (Employee)

② Projection:- The Projection operation on a table, simply from another table by copying specified columns from the original table Symbol of projection is (Π)

To select name of employee :-

Πname (Employee)

for salary :-

Πsalary (Employee)

We want the name of all the employee having salary less than 70,000

σsal. Πsalary σ salary > 70,000 [Πname (Employee)]

Page No.
Date

Codd's Rules :-

(1) Information Rule :- All Information in a relational database including tables names, Column names is represented in the form of tables. This Simple View of data Speedsup design and learning process. User productivity is improved since knowledge of only one language is necessary to access all data such as description of the table and attribute definitions, integrity constraints

(2) Guranteed Access Rule :- Every piece of data in relational database, can be accessed by using a primary Key value that identifies the row and Column name. User productivity is improved since there is no need to resort to using physical pointers or address.

(3) Comprehensive Data Sub-language Rule :- The RDBMS may support several languages. But atleast one of them should allow the user to do all following define table, query & update data, Set integrity Constraints, Set authorizations & define transaction. User productivity is improved since there is just one approach that can be used for all database operations.

(4) View updating Rule :- Any view that can be updated theoretically can be updated using the RDBMS. Data Consistency is ensured since the changes made in the View are

transmitted to base table and vice-versa

(5) High level Insert, update & Delete: The RDBMS support insertion, updating and deletion at a table level The performance is improved since the commands acts on a set of records rather than one record at a time.

(6) Physical Data independence: The e Database administrators can make change to the physical access and storage method, which improve performance and domot require changes in the application programs or request

(7) Logical Data independence: Logical changes in tables and views such as adding/delleting columns or changing field lengths need not necessitate modifications in the programs ~~or in~~

(8) Integrity Independence: Integrity constraints are stored in the one line Catalog or data dictionary and can therefore be changed without necessitating changes in the application program

(9) Non Subversion Rule: If the RDBMS has a language that accesses the information of records at a time, this language should not be used bypass the integrity constraints. This is necessary for data integrity

(10) Systematic Treatment of Null values :- In RDBMS null value should be supported for the representation of missing ~~information~~ and in applicable information. The database management must have a consistent method for representing null values.

(11) Database Description Rule :- The description of database is stored and maintained in the form of tables

(12) Distribution Independence :- The RDBMS package must have distribution independence.

Comparision :-

	Relation Algebra	Relational Calculus.
(1)	It is a procedural method for solving Queries	It is non-procedural method of solving the Queries
(2)	The Solution to the data base access problem using a relational algebra is obtained by stating what is required and what are the steps to obtain that information	The solution to the database access problem using a relational Calculus is obtained simply by stating what is required and letting the system find the answer
	It is used as vehicle for implementation of Relational Calculus	Relation Calculus Queries are Converted into equivalent relational algebra format by using Codd's reduction Algorithm and then it is implemented

	with the help of relational algebra operators
(4) Relational algebra operators are used as a yardistick for measuring the expressive power of any given language.	A lang. is said to be Complete if it is at least powerful as the Calculus that is, if any relation definable by some expression of the calculus is also definable by some expression of the language in question

Difference :-

(1). Relational Algebra is a procedural language that can be used to tell the DBMS how to build a new relation from one or more relation in the database.

- Relational Calculus is non-procedural language that can be used to formulate the definition of relation in terms of one or more database relation.

(2) Relation Algebra :- user has to specify what is required and what are the procedure or steps to obtain output.

Relational Calculus :- user just specifies what is required and need not to specify how to obtain it