# *Database Fundamentals*

Database Fundamentals

- Introduction

- Types of Databases

- Relational Database

- Entity Relationships

- Normalization

- First Normalization

- Second Normalization

- Third Normalization

- Boyce Codd Normalization

Introduction

- Data collected from different activities have to be stored electronically

- This data should be readily available for updation and analysis

- Records should be arranged in a systematic order and should identified based on certain attributes

- Such a structure becomes a repository of information

Why we need a database
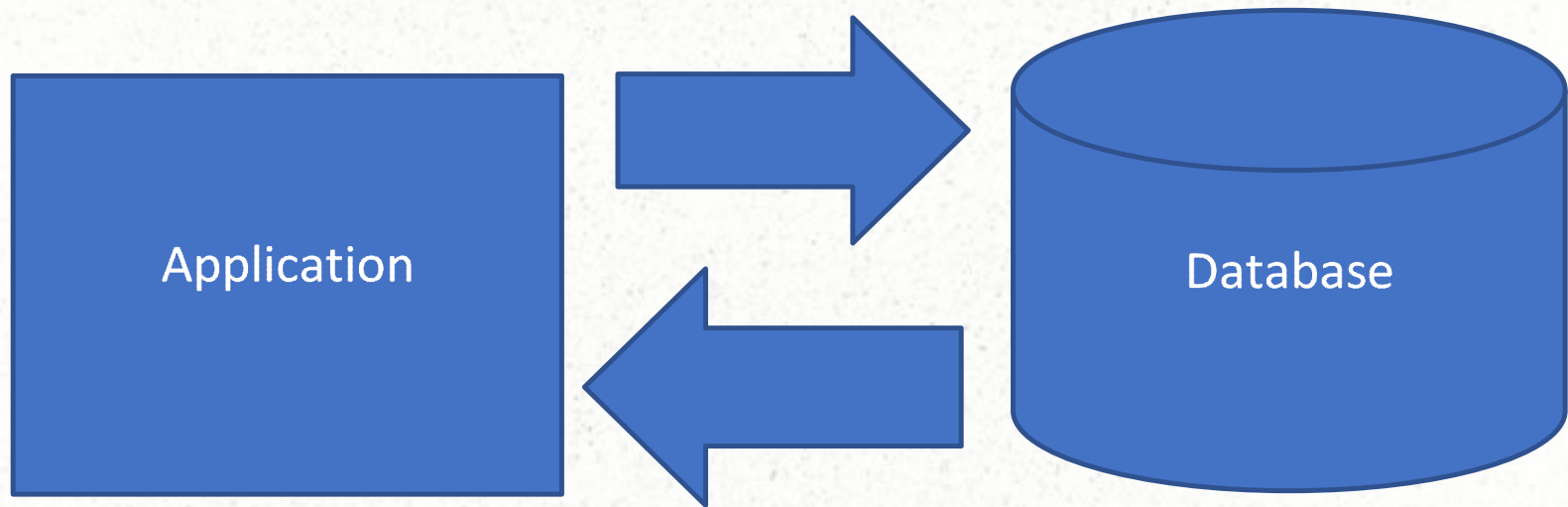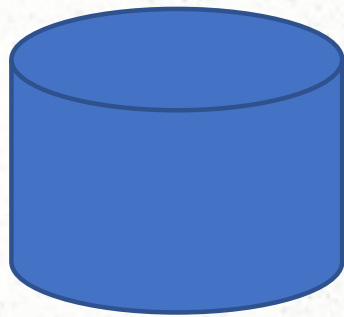
Application → Database

Database → Application
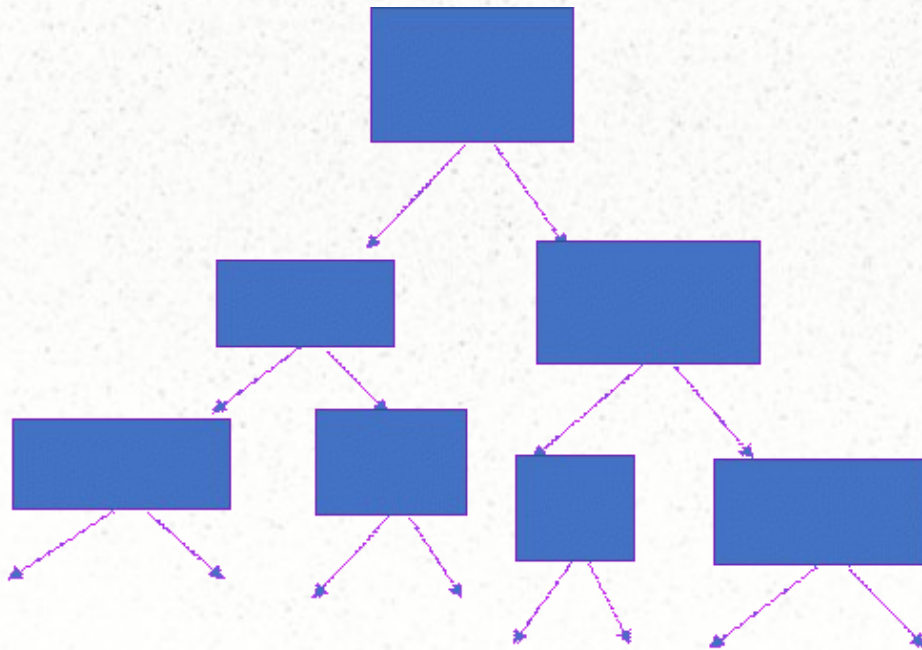
Types of Models for Database Management System

- Hierarchical database

- Network database

- Relational database

- NoSQL database

- Object-oriented database

Hierarchical Model

- In this type of model data is organized in a tree like structure ,where records have a parent child relationship

-

Hierarchical Database

- In a hierarchical database,one parent can have many child nodes,but a child can have only one parent.

- The first record in the database is root record

- The relationship formed between Parent and child nodes is one to many

Advantage of Hierarchical Database

- Traversal through records is easy

- Nodes can be mapped and identified easily

Disadvantage of Hierarchical Database

- Rigid structure

- Multiple nodes with same data will lead to data redundancy

- Movement of record from child to parent or parent to child can be a difficult task

Network Database Model

- A network database model is a progression of hierarchical database model .

- In a network model every child can have multiple parents

- The nodes have a many to one relationship

-

Network Database Model
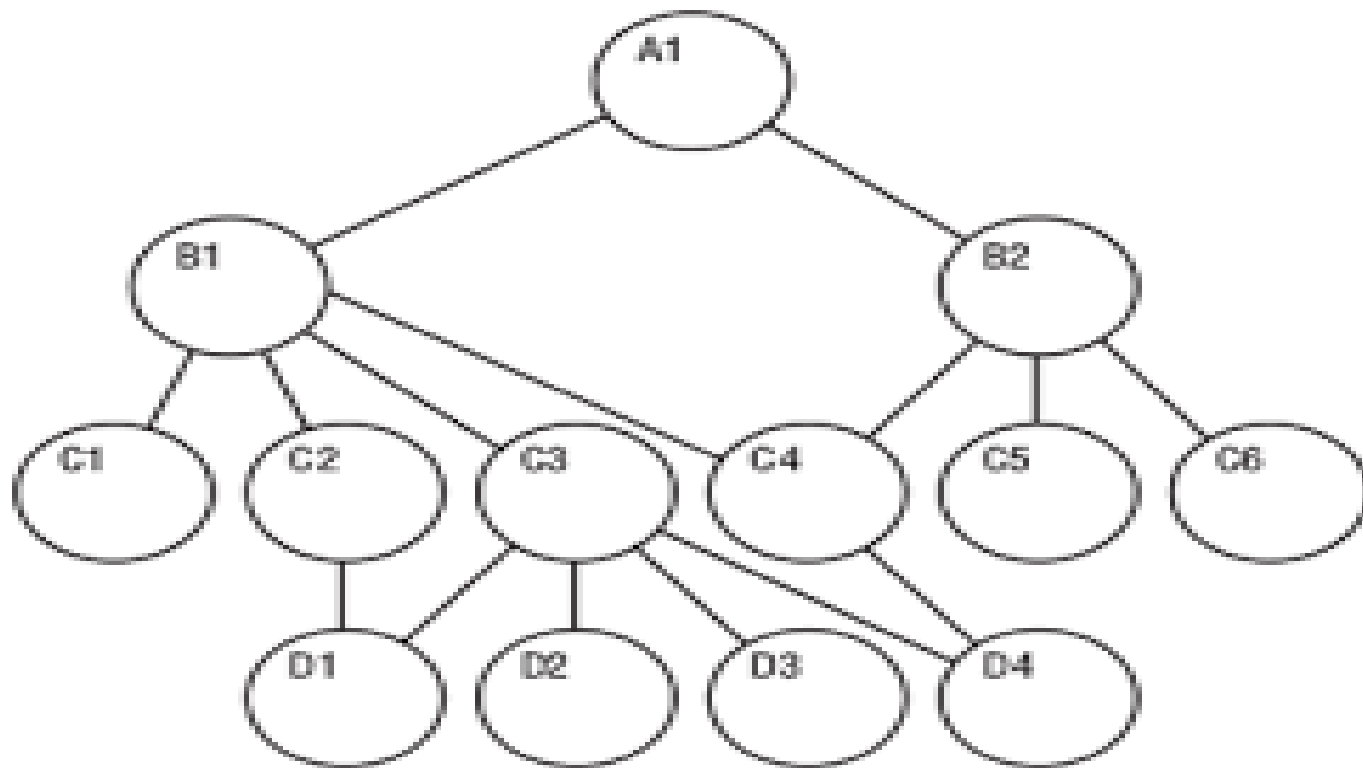
Advantage of Network Model

- Nodes in network model can have Many to Many relationship

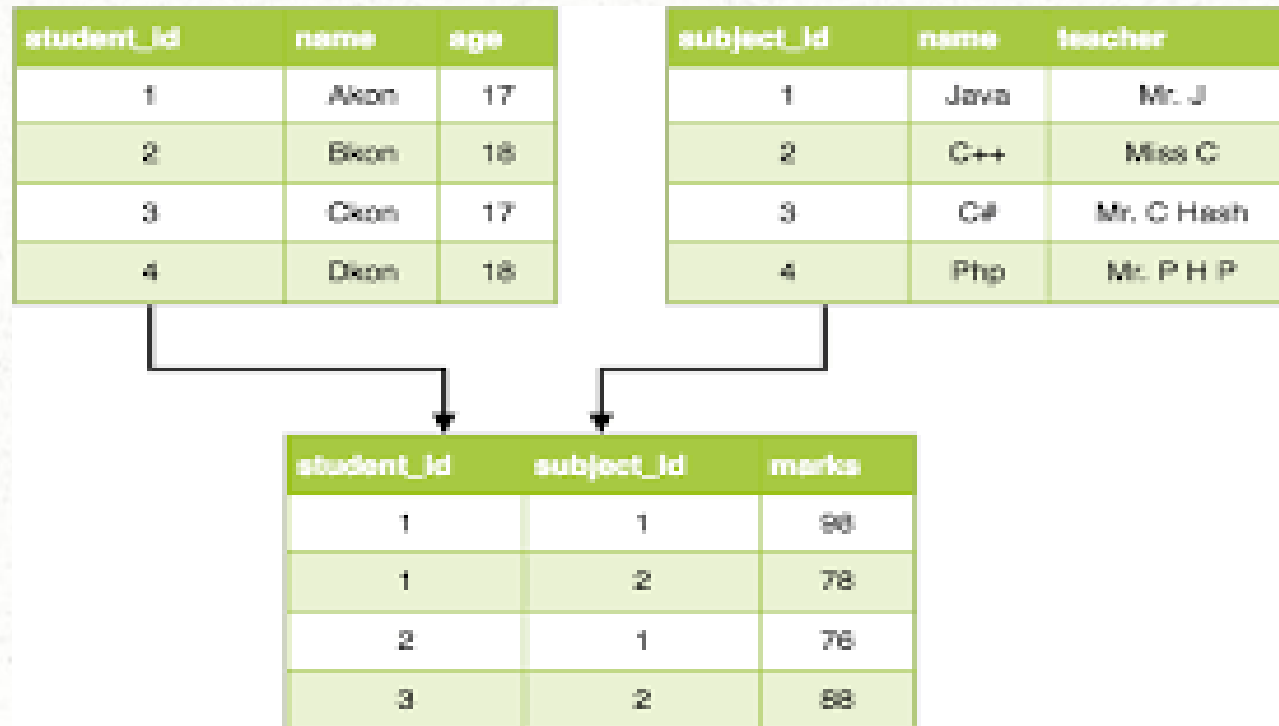# *Disadvantage of Network model*

- Complex structure makes it tough to implement ,understand and traverse

-

Relational Model

- In a relational model records are stored as table known as relation

- In a relation each row is known as tuple and each column is an attribute

- The tables are related to each other ,this helps in storing data efficiently and reducing redundancy through normalization

Relational Database

student_Id	name	age
1	Akon	17
2	Bkon	18
3	Ckon	17
4	Dkon	18

subject_Id	name	teacher
1	Java	Mr. J
2	C++	Miss C
3	C#	Mr. C Hash
4	Php	Mr. P H P

student_Id	subject_Id	marks
1	1	98
1	2	78
2	1	76
3	2	88

Advantages of Relational Database

- Data stored in tabular format is easy to understand

- Data redundancy can be removed

- Tables can be related to one another

- Flexible in comparison to other models

- Has a separate set of Querying language called SQL

NoSQL database

- NoSQL stands for not only SQL

- NoSQL databases are non tabular structures

- Data is stored differently then Relational database in document ,key value pair or graph format

- Some of the popular NoSQL database are MongoDb ,Apache Ignite,Redis

Object Oriented Databases

- Object Oriented Database is based on Object Oriented Programming

- Data is represented and stored as objects in Database

- OODbs have many features like extensibility,custom data models and provision for modeling complex data

Entity Relation Ship Diagram

- ER Diagram stands for Entity Relationship Diagram, also known as ERD is a diagram that displays the relationship of entity sets stored in a database. In other words, ER diagrams help to explain the logical structure of databases. ER diagrams are created based on three basic concepts: entities, attributes and relationships.

Symbol	Description
▭	Represents Entity
⬭	Represents Attribute
◇	Represents Relationship
—	Links Attribute(s) to entity set(s) or Entity set(s) to Relationship set(s)
⬭⬭	Represents Multivalued Attributes
⬭ (dotted)	Represents Derived Attributes
═	Represents Total Participation of Entity
▭▭	Represents Weak Entity
◇◇	Represents Weak Relationships
⬭⬭⬭	Represents Composite Attributes
⬭ (underlined)	Represents Key Attributes / Single Valued Attributes
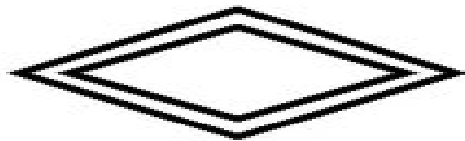
Crow's foot notation symbols:

- One
- Many
- One (and only one)
- Zero or one
- One or many
- Zero or many

ER Diagram

Student

Employee

Vehicle

Teacher

ER Diagram

Phone no

Enmployee Name

Employee

Salary

Department

Job Role

Employee ID

Designation
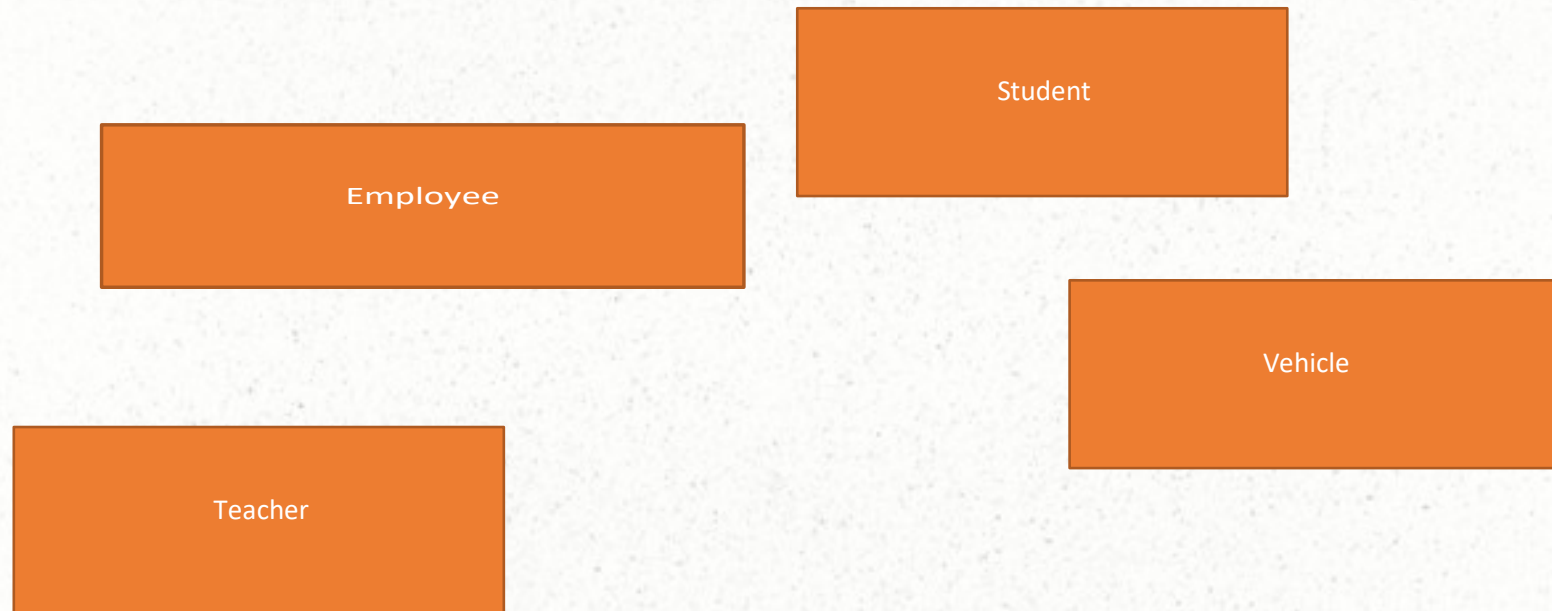
Exercise-1

- Draw an ER Diagram for below given entities

- Student

- Teacher

- Vehicle

- Exam

- Book

- Author

ER Diagram

Teacher — Teaches — Student

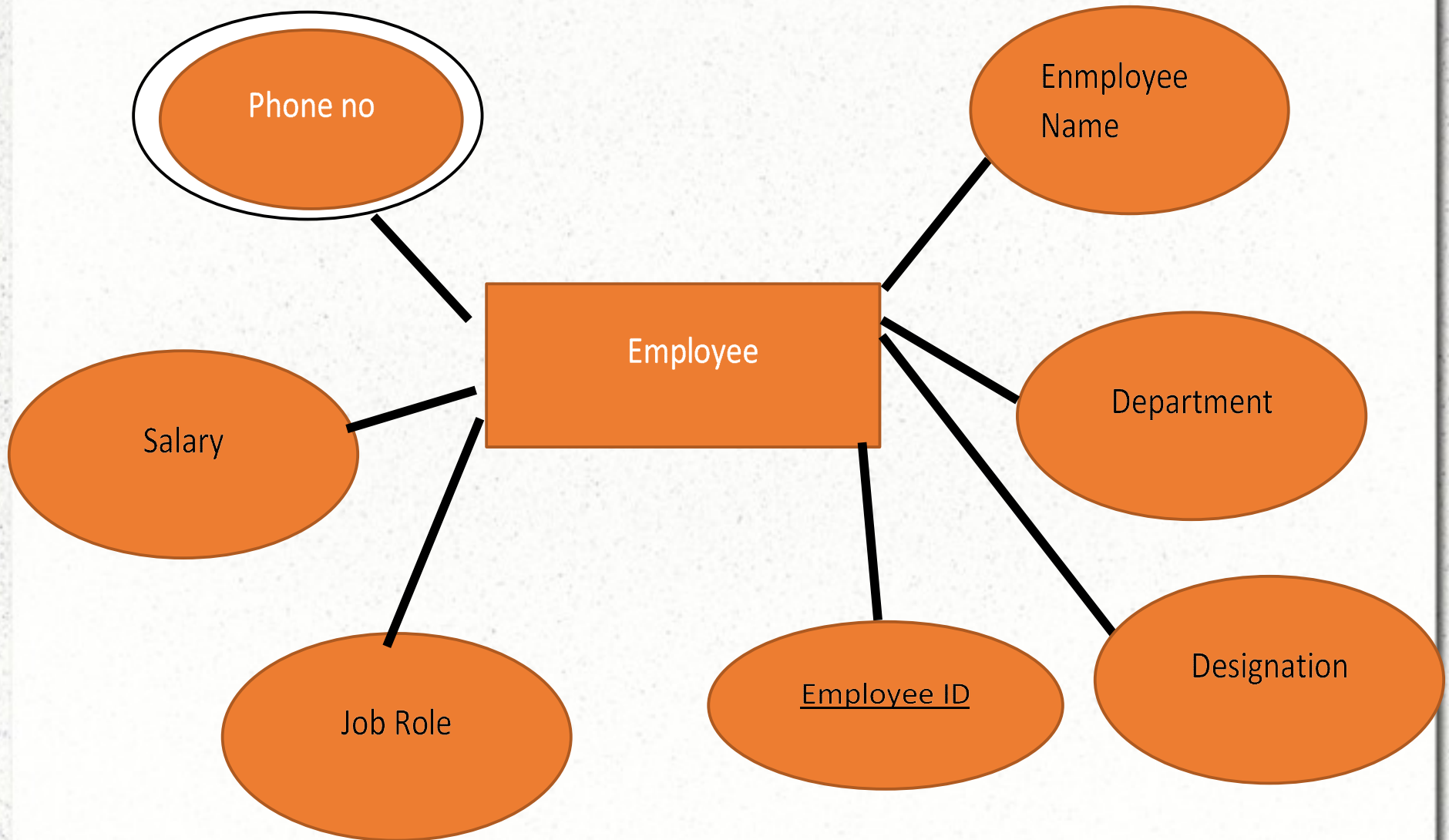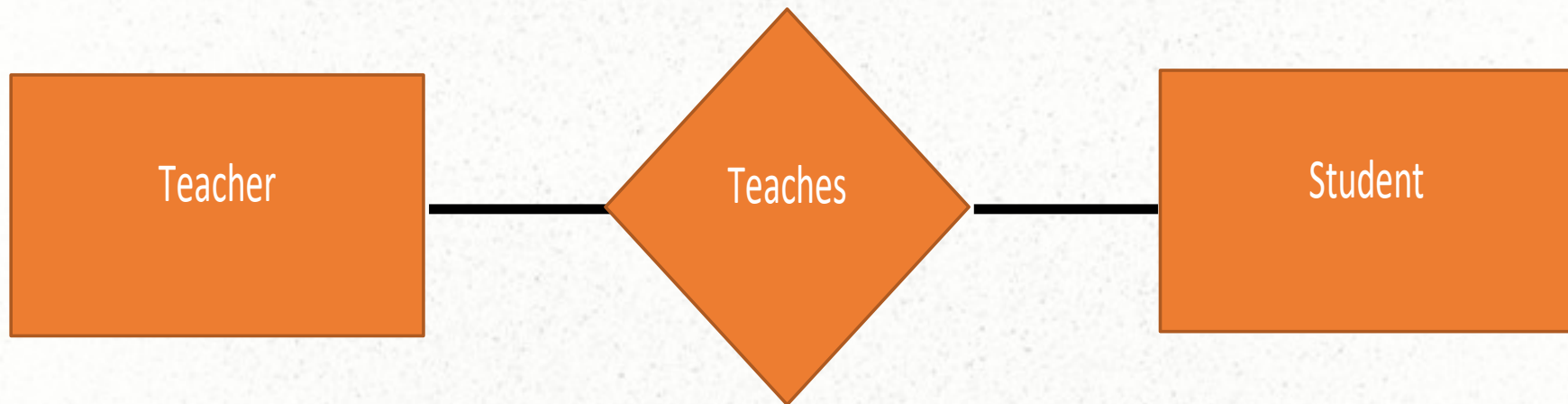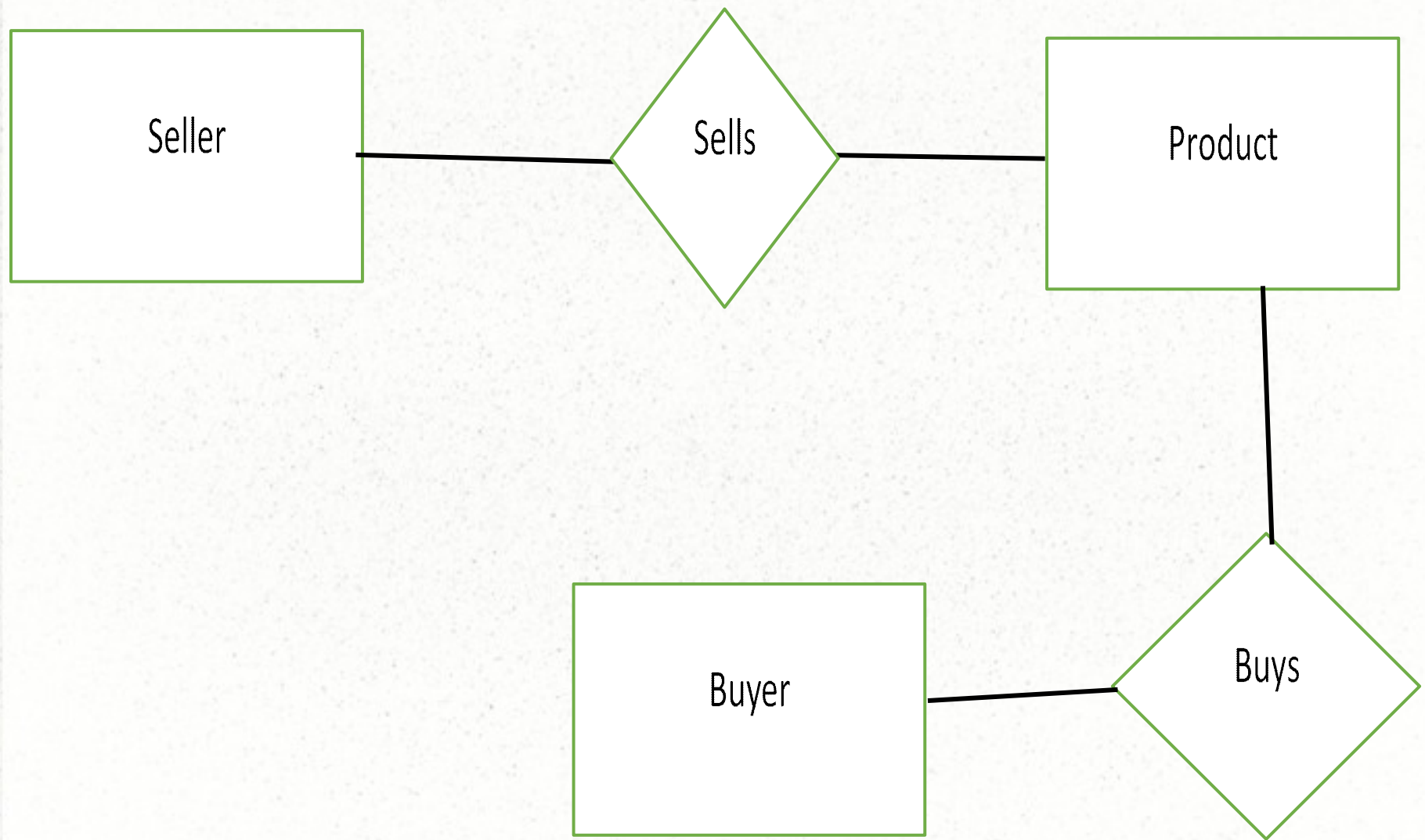# *Exercise 2*

- Identify and draw relationship diagram between these group of entities

- Student ,Subject ,Teacher

- Student ,Exam

- Employee ,Organization,Salary

- Doctor,Patient,Medicine

Data Redundancy

- Data Redundancy is duplication of data in database .

- Data Redundancy complicates database ,and increases data retrieval time .

- Also Data Redundancy can cause a wastage of space ,and maintenance of database can be difficult

How to remove Data Redundancy

- One way to remove data Redundancy is through Normalization of tables in database

- Normalization is the process through which duplicate values are removed from database by way of decomposing table and reducing redundancy

- There are four types of Normalization ,

- First ,Second ,Third and Boyce Codd Normalization

First Normalization

- If we look at the table below

Customer Name	Address	Telephone Number
Peter	23 old street	9983344343
Peter	23 old street	9986633774
Sam	45 baker street	8834300340
Mathew	59 Garden House	8899393432
Mathew	59 Garden House	8877577342

First Normalization

What are the issues with this table

- The values are getting repeated because customer has multiple contact numbers

- If we try to update address of a customer ,it will have to be updated at multiple locations

- If we try to delete a record ,It has to be deleted from multiple location

- This can create inconsistencies

First Normalization

- Solution

  The table has to be converted in First Normalization form

- A table is said to be First Normalization form when

  attributes in the table are single or atomic in nature

Solution

Customer Name	Address	Telephone Number1	Telephone Number2
Peter	23 old street	9983344343	9986633774
Sam	45 baker street	8834300340	
Mathew	59 Garden House	8899393432	8877577342

Second Normalization

- Let's look at the table below

Studentid	Student Name	Subject Name	Teacher
1	Peter	Java	Mr James
2	Sam	CSharp	Mr Erric
3	Edgar	Java	Mr James
4	Mathew	Visual Basic	Mr Harry
5	William	Java	Mr James
6			

Second Normalization

- The table has a primary key student id

- Student Name is dependent on student id

- But Teacher is dependent on subject(decided by subject taken by student)

- If we update Teacher for a subject ,it has to updated at multiple places.

- If the student wishes to change subject ,It has to be changed at multiple places

Second Normalization

- A table is said to be in Second Normalization form if

  It is in first normalization form

  - ◆ There should be no Partial Dependency of key.That is any non-prime attribute of re-lation should not be dependent on any subset of candidate Key

  - ◆

Studentid	Student Name	Subject ID
1	Peter	
2	Sam	
3	Edgar	
4	Mathew	
5	William	
6		

Subject id	Subect Name	Teacher

A Case of Transitive Dependency

- A--->B------->C

  A transitive dependency is generated when an attribute is dependent on primary key transitively

-

An example of Transitive dependency

BookId	GenerId	GenerType	Price
1	2	Horror	$120
2	1	Comic	$340
3	2	Horror	$150
4	3	Action	$130
5	1	Comic	$220

Third Normalization

- A Table is said to be in 3NF when

  It is in 2NF

  There is no transitive functional dependency

Boyce Codd Normalization

- A table complies with BCNF if it is in 3NF and for every functional dependency X->Y, X should be the super key of the table.

eid	e_city	dept	dept_type	dept_no_emp
1	Delhi	Sales	Rev	20
1	Delhi	Marketing	Rev	5
2	Mumbai	Training	IT	7
2	Mumbai	Development	IT	12
3	Chennai	Accounts	Acc	5

- Functional dependencies in the table above:

- emp_id -> emp_city

- dept -> {dept_type, dept_no_emp}

- Candidate key: {emp_id, dept}

eid	City
1	Delhi
2	Mumbai
3	Chennai

Dept	Dept_type	dept_no_emp
Sales	REV	20
Marketing	REV	5
Training	IT	7
Development	IT	12
Accounts	ACC	5

eid	dept
1	sales
1	Marketing
2	Training
2	DEV