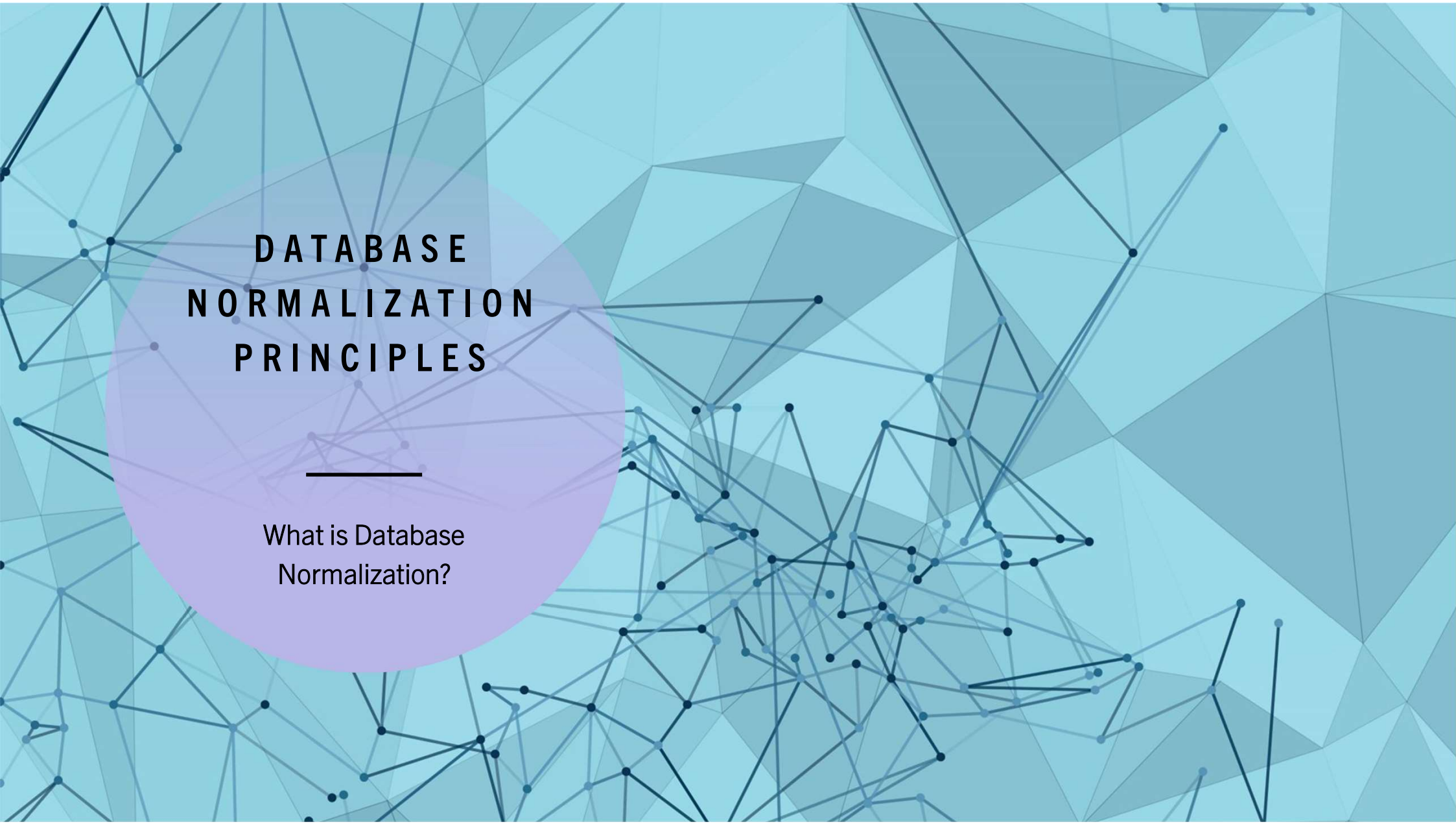

DATABASE
NORMALIZATION
PRINCIPLES
What is Database
Normalization?

WHAT IS DATABASE NORMALIZATION?

- Database normalization is a process used to organize a database into a structured format to reduce redundancy and improve data integrity. This process involves decomposing tables into smaller tables and defining relationships between them.
- Normalization typically follows these forms:
- 1NF (First Normal Form)
- Objective: Ensure atomicity by eliminating repeating groups and ensuring each column contains atomic values.
- Rules:

1. Each column should contain only one value.
2. Each row must be unique.
3. Each column must have a unique name.

EXAMPLE:

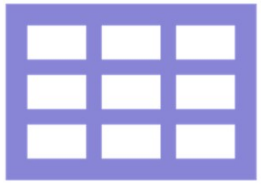
UNNORMALIZED TABLE:

OrderID	ProductIDs	CustomerID
101	1, 2, 3	501
102	4, 5	502

NORMALIZED TO 1NF:

OrderID	ProductID	CustomerID
101	1	501
101	2	501
101	3	501
102	4	502
102	5	502

2NF (SECOND NORMAL FORM)



- Objective: Eliminate partial dependency by ensuring all non-key attributes are fully dependent on the primary key.
- Rules:

1. The table must meet all 1NF rules.
2. Remove partial dependencies (i.e., non-prime attributes should depend on the whole primary key, not just a part).

EXAMPLE:
1NF TABLE:

OrderID	ProductID	CustomerName	ProductName
101	1	Alice	Laptop
101	2	Alice	Mouse
102	3	Bob	Keyboard

NORMALIZED TO 2NF: ORDERS TABLE:

OrderID	CustomerName
101	Alice
102	Bob

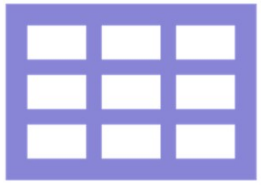
PRODUCTS TABLE:

ProductID	ProductName
1	Laptop
2	Mouse
3	Keyboard

ORDERDETAILS TABLE:

OrderID	ProductID
101	1
101	2
102	3

3NF (THIRD NORMAL FORM)



- **Objective**: Eliminate transitive dependency by ensuring non-key attributes depend only on the primary key.
- **Rules**:

1. The table must meet all 2NF rules.
2. Remove transitive dependencies (i.e., non-key attributes should not depend on other non-key attributes).

EXAMPLE:

2NF TABLE:

CustomerID	CustomerName	City	State
501	Alice	New York	NY
502	Bob	Los Angeles	CA

NORMALIZED TO 3NF: CUSTOMERS TABLE:

CustomerID	CustomerName	CityID
501	Alice	1
502	Bob	2

CITIES TABLE:

CityID	City	State
1	New York	NY
2	Los Angeles	CA

ER diagrams

- ER-modeling is a data modeling technique used in software engineering to produce a conceptual data model of a information system.
- Diagrams created using this ER-modeling technique are called Entity-Relationship Diagrams, or ER diagrams or ERDs.

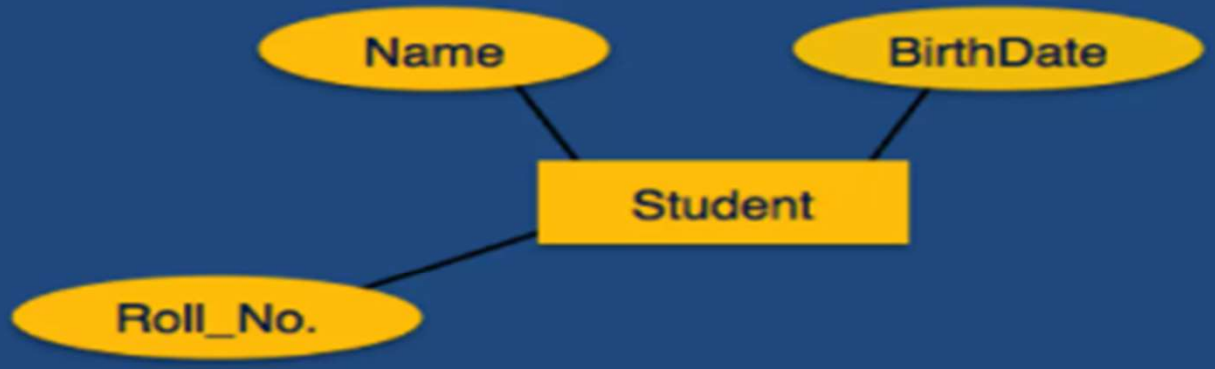
Entity

- Entities are represented by means of rectangles. Rectangles are named with the entity set they represent.

Entities in a school database

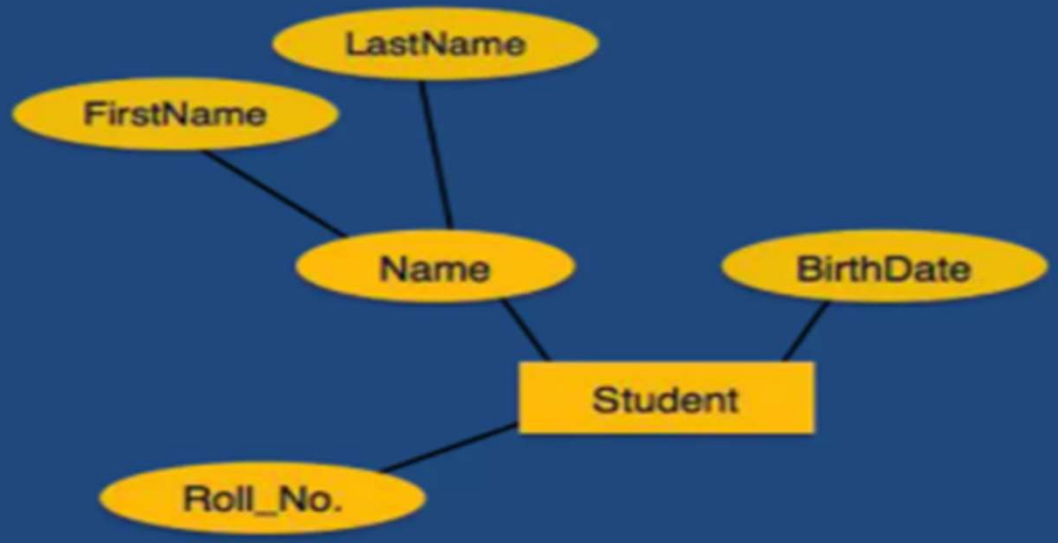
Attributes

- Attributes are properties of entities. Attributes are represented by means of eclipses. Every eclipse represents one attribute and is directly connected to its entity (rectangle).

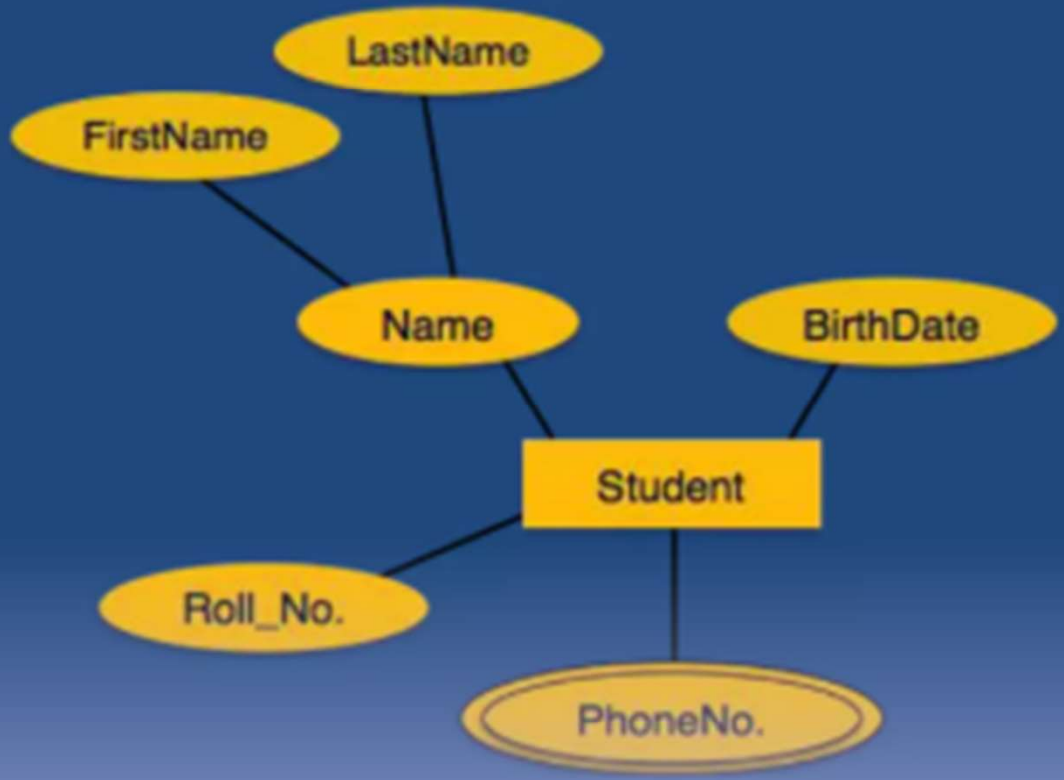


composite attributes

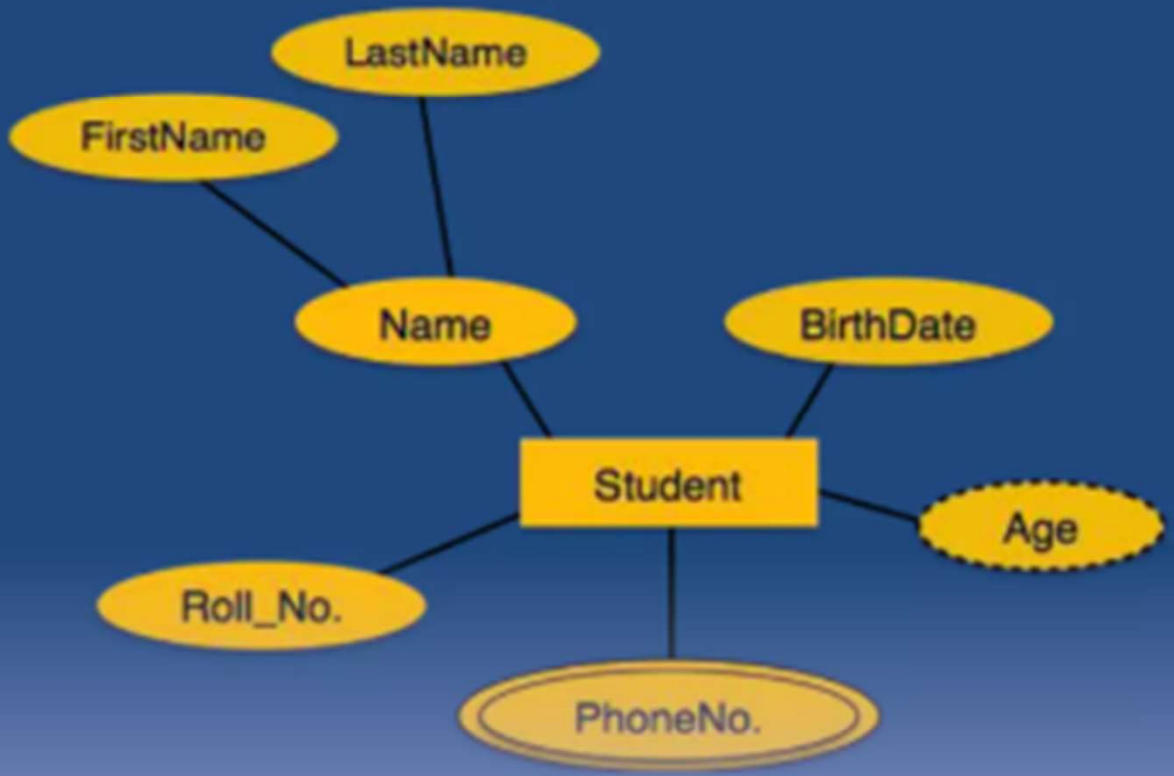
- If the attributes are ***composite***, they are further divided in a tree like structure. Every node is then connected to its attribute. That is composite attributes are represented by eclipses that are connected with an eclipse.



- ***Multivalued*** attributes are depicted by double eclipse.



- ***Derived*** attributes are depicted by dashed eclipse.



Relationship

- Relationships are represented by diamond shaped box. Name of the relationship is written in the diamond-box. All entities (rectangles), participating in relationship, are connected to it by a line.

One-to-one

- When only one instance of entity is associated with the relationship, it is marked as '1'. This image below reflects that only 1 instance of each entity should be associated with the relationship. It depicts one-to-one relationship
- The 1:1 (read as “one to one”) relationship means that one instance of the parent entity is associated with one instance of the child entity.

One-to-many

- When more than one instance of entity is associated with the relationship, it is marked as 'N'. This image below reflects that only 1 instance of entity on the left and more than one instance of entity on the right can be associated with the relationship. It depicts one-to-many relationship
- In this kind of relationship, a single instance of a parent entity is associated with many instances of a child entity; however, the child entity instance is related to only one instance of the parent.

Many-to-one

- When more than one instance of entity is associated with the relationship, it is marked as 'N'. This image below reflects that more than one instance of entity on the left and only one instance of entity on the right can be associated with the relationship. It depicts many-to-one relationship.
- In this case, many instances of a parent entity can relate to many instances of a child entity.

Many-to-many

- This image below reflects that more than one instance of entity on the left and more than one instance of entity on the right can be associated with the relationship. It depicts many-to-many relationship