Database Design

Normalization

Normalization & Normal Forms

- A set of conditions on table structure that improves maintenance.
- Normalization removes processing anomalies:
  - Update
  - Inconsistent Data
  - Addition
  - Deletion

1st Normal Form
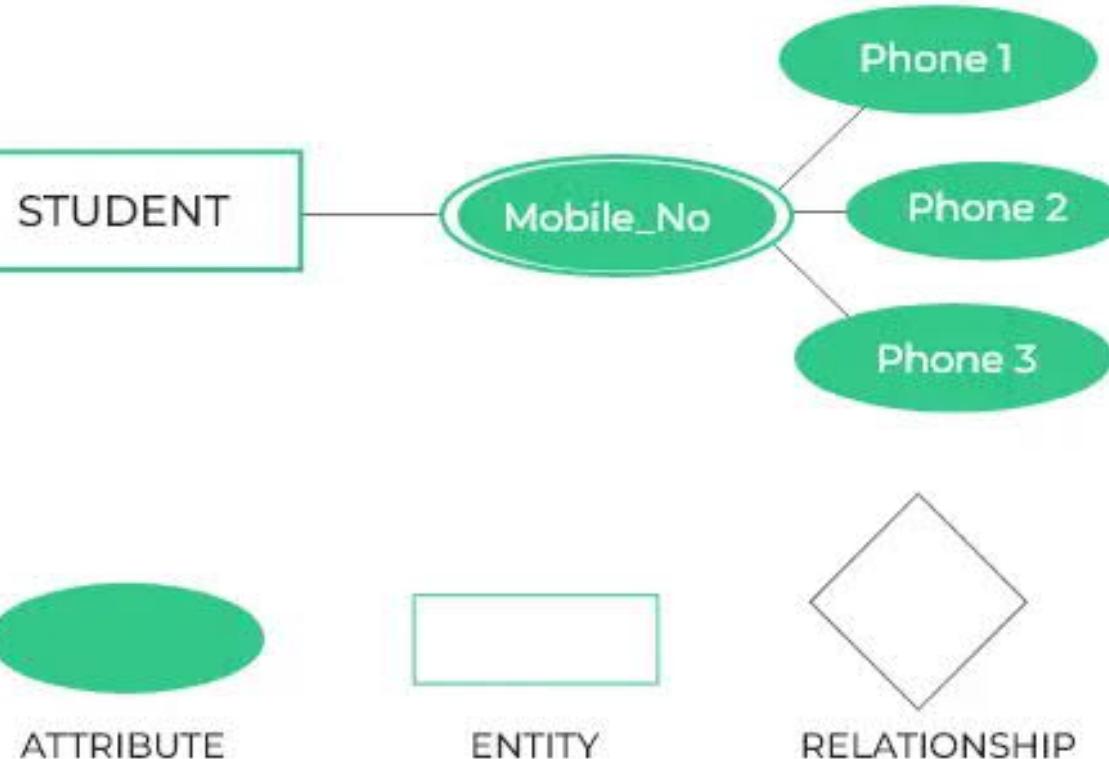
- Table has a primary key
- Table has no:
  - Multivalued attribute, or
  - Repeating groups

A <u>multivalued attribute</u> is an attribute that may have several values for one record

A <u>repeating group </u>is a set of one or more multivalued attributes that are related

Multivalued Attribute

MultiValued Attribute

STUDENT — Mobile_No — Phone 1 / Phone 2 / Phone 3

ATTRIBUTE   ENTITY   RELATIONSHIP

Repeating Group

Repeating group

Employee

ENO	Name	Province	PayDate1	Amount1	PayDate2	Amount2
E001	Somchai	Khon Kaen	15/04/2004	5,000.00	30/04/2004	5,000.00
E002	Sompong	Sarakham	15/04/2004	4,500.00	30/04/2004	4,500.00
E003	Somchay	Ubon	15/04/2004	5,200.00	30/04/2004	5,200.00

Employee

ENO	Name	Province
E001	Somchai	Khon Kaen
E002	Sompong	Sarakham
E003	Somchay	Ubon

PayCheck

ENO	PayDate	Amount
E001	15/04/2004	5,000.00
E001	30/04/2004	5,000.00
E002	15/04/2004	4,500.00
E002	30/04/2004	4,500.00
E003	15/04/2004	5,200.00
E003	30/04/2004	5,200.00

Repeating Group (Example 2)

Example: Repeating Groups

Figure 3.8a Hotel Invoice with Two Repeating Groups — Sample Invoice

GRANDVIEW HOTEL
Sea Bluffs, California

Invoice Number: 1234   Arrival Date: 10/12/2003
Customer Name: Mary Jones
                           Fred Jones
                           Sally Jones

10/12/2003	Room	$ 99.00
10/12/2003	Food	$ 37.55
10/12/2003	Phone	$ 2.50
10/12/2003	Tax	$ 15.00
10/13/2003	Room	$ 99.00
10/13/2003	Food	$ 47.90
10/13/2003	Tax	$ 15.00
	Total	$ 315.95

(a)

3753 X2 - 2004

14

Example (Order Table)

☐ Multivalued attribute:

OrderNumber, OrderDate, {PartNumber})

[12491 | 9/02/2001 | BT04, BZ66   ]

☐ Repeating group:

OrderNumber, OrderDate, {PartNumber, NumberOrdered}

[12491 | 9/02/2001 | (BT04, 1), (BZ66, 1)]

Normalization: 1NF
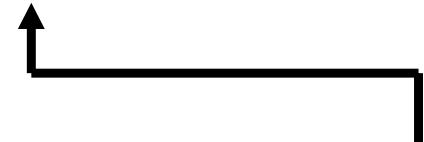
- Every multi-valued attribute and repeating group becomes a new table with the appropriate foreign key relationships preserved.
- Remove nested repeating groups from the outside in

Order(OrderNumber, OrderDate, {PartNumber, {Supplier}})
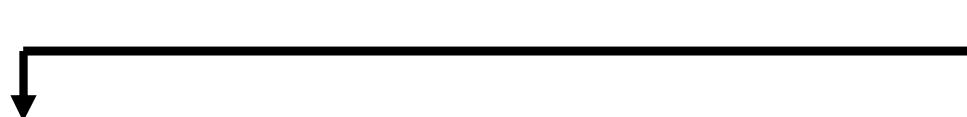
Example: 1NF

Order(<u>OrderNumber</u>, OrderDate,

           {PartNumber, {Supplier}})

Order(<u>OrderNumber</u>, OrderDate)

        Order-Part(<u>OrderNumber</u>, <u>PartNumber</u>)

Part(<u>PartNumber</u>, {Supplier})

Example: 1NF (cont.)

Part(<u>PartNumber</u>, {Supplier})

Part(<u>PartNumber</u>)

Part-Supplier(PartNumber, SupplierNum)

Supplier(<u>SupplierNum</u>)

2nd Normal Form (2NF)

Partial Dependencies

- CourseRegistration(StudentID, CourseID, StudentName, CourseName)
  - Primary Key: (StudentID, CourseID)
  - Non-key attributes: StudentName, CourseName
- Dependencies
  - StudentName → StudentID (depends only on StudentID)
  - CourseName → CourseID (depends only on CourseID)

Partial Dependencies (cont...)

- Continuing previous example:
  - CourseRegistration(StudentID, CourseID, StudentName, CourseName)
    - Primary Key: (StudentID, CourseID)
    - Non-key attributes: StudentName, CourseName
- StudentName and CourseName are partially dependent on the composite key because they don't need both StudentID and CourseID to be identified.

2nd Normal Form

- No partial dependencies

No attribute depends on only some of the attributes of a concatenated key.

Order-Part

[OrderNumber | PartNumber | PartDescription]

Create a new table with PartNumber key.

3rd Normal Form (3NF)

Transitive Dependency

- Employee(EmpID, EmpName, DeptID, DeptName)
- Primary Key: EmpID
- Dependencies:
    - EmpName → EmpID (✔)
    - DeptID → EmpID (✔)
    - DeptName → DeptID
        - Problem: DeptName depends on DeptID, which depends on EmpID

3rd Normal Form

- 3rd Normal Form: no transitive dependencies

  Transitive dependency means that a <u>non-key</u> attribute depends on another <u>non-key</u> attribute(s) which in turn, depends on the primary key

- If A → B and B → C, then A → C is a transitive dependency.

3NF (Before & After)

Student_ID	Student_Name	Dept_ID	Dept_Name	Dept_Loc
101	Alice Smith	D01	Science	Building A
102	Bob Johnson	D02	Arts	Building B
103	Carol White	D03	Commerce	Building C

Student_ID	Student_Name	Dept_ID
101	Alice Smith	D01
102	Bob Johnson	D02
103	Carol White	D03

Dept_ID	Dept_Name	Dept_Loc
D01	Science	Building A
D02	Arts	Building B
D03	Commerce	Building C

Normalization

- Eliminates Data Redundancy
- Improves Data Integrity
- Avoids Insertion, Update, and Deletion Anomalies
- Supports Scalability and Maintainability
- Logical Data Organization for better understanding

Normalization

- 1NF
  - Keys & no multivalued attributes or repeating groups
- 2NF
  - 1NF & all attributes depend on all key components
- 3NF
  - 2NF & No transitive dependency exists between non-prime attributes and the primary key.