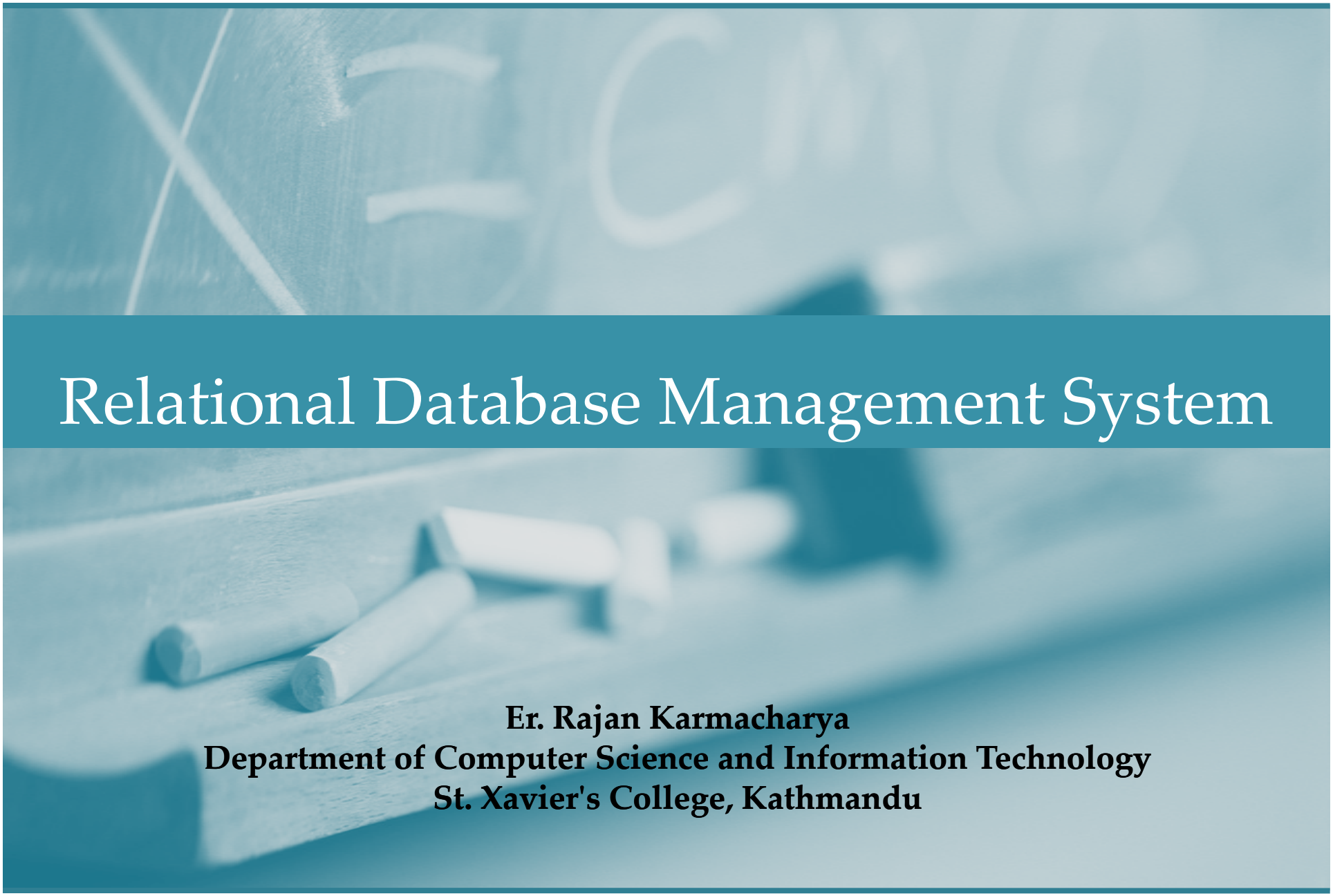
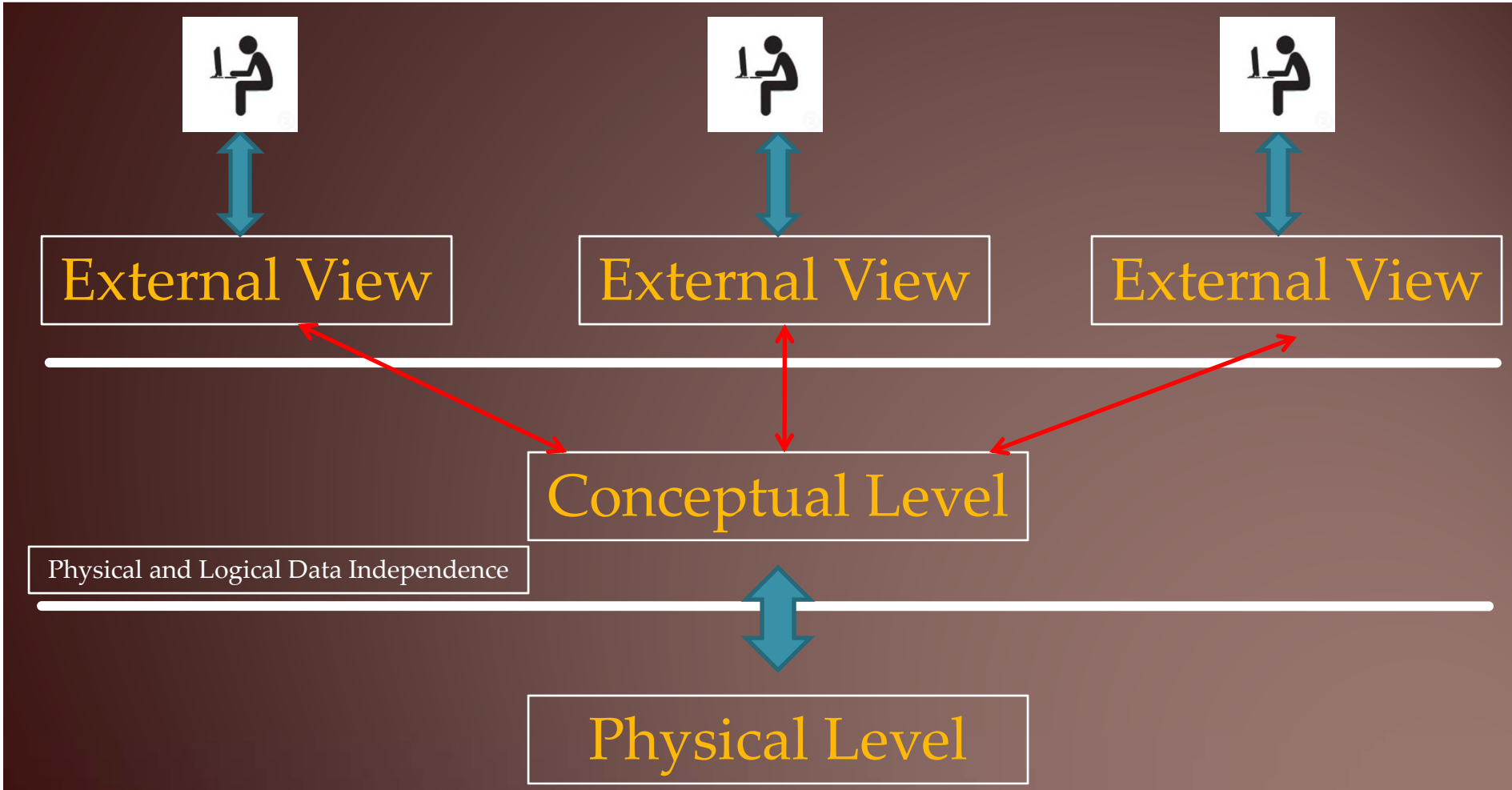


Relational Database Management System

Er. Rajan Karmacharya
Department of Computer Science and Information Technology
St. Xavier's College, Kathmandu

Database Design



Database Design

- Physical Level
  - Lowest level where certain physical components organize and store raw data
  - They also include control structures that track the location and format of the stored data elements.
- Conceptual Level
  - Isolates the data storage details to the physical level
  - Tables, views, procedures , triggers exist at this level
  - If the underlying hardware and OS changes, the consequences are limited to the interface between the physical and conceptual level.
- External Level
  - Highest and outermost layer which presents varying external tailored views of the application to the different users.

Entity

- An entity is a class of persons, places, objects, events or concepts in the real world that is distinguishable from other objects
- Person : contractor, teacher, employee, student
- Place: zone, country, branch
- Object : tool, machine, building, product
- Event : sale, award, registration, renewal
- Concept : qualification, account, course

Attribute

- Attributes are descriptive properties possessed by each member of an entity.
- Attributes are also called element, property or field.
- The values for each attribute are defined in terms of three properties viz data type, domain and default.
- Data type defines what type of data can be stored in that attribute
- Domain defines what values an attribute can legitimately take on
- Default value is the value that will be recorded if not specified by the user.

Emp_ID	Emp_Name	Emp_Designation	Emp_Contact
E001	Andrew Mathews	Computer Operator	9841444444
E015	Pemba Lama	System Analyst	9841555555
E016	Siris Bashyal	Programmer	9843333333
E099	Bishal Dahal	Data Analyst	9849999999

Emp_ID, Emp_Name, Emp_Designation,Emp_Contact are attributes

Relationships

- Conceptually, entities and attributes do not exist in isolation.
- The things they represent interact with and impact one another to support the business mission.
- Relationship is a natural business association that exists between one or more entities.
- Types of relationships
  - One to One ---------- driver and car
  - One to many---------- teacher and student
  - Many to many---------- books and readers
  - Many to one ---------- students and college

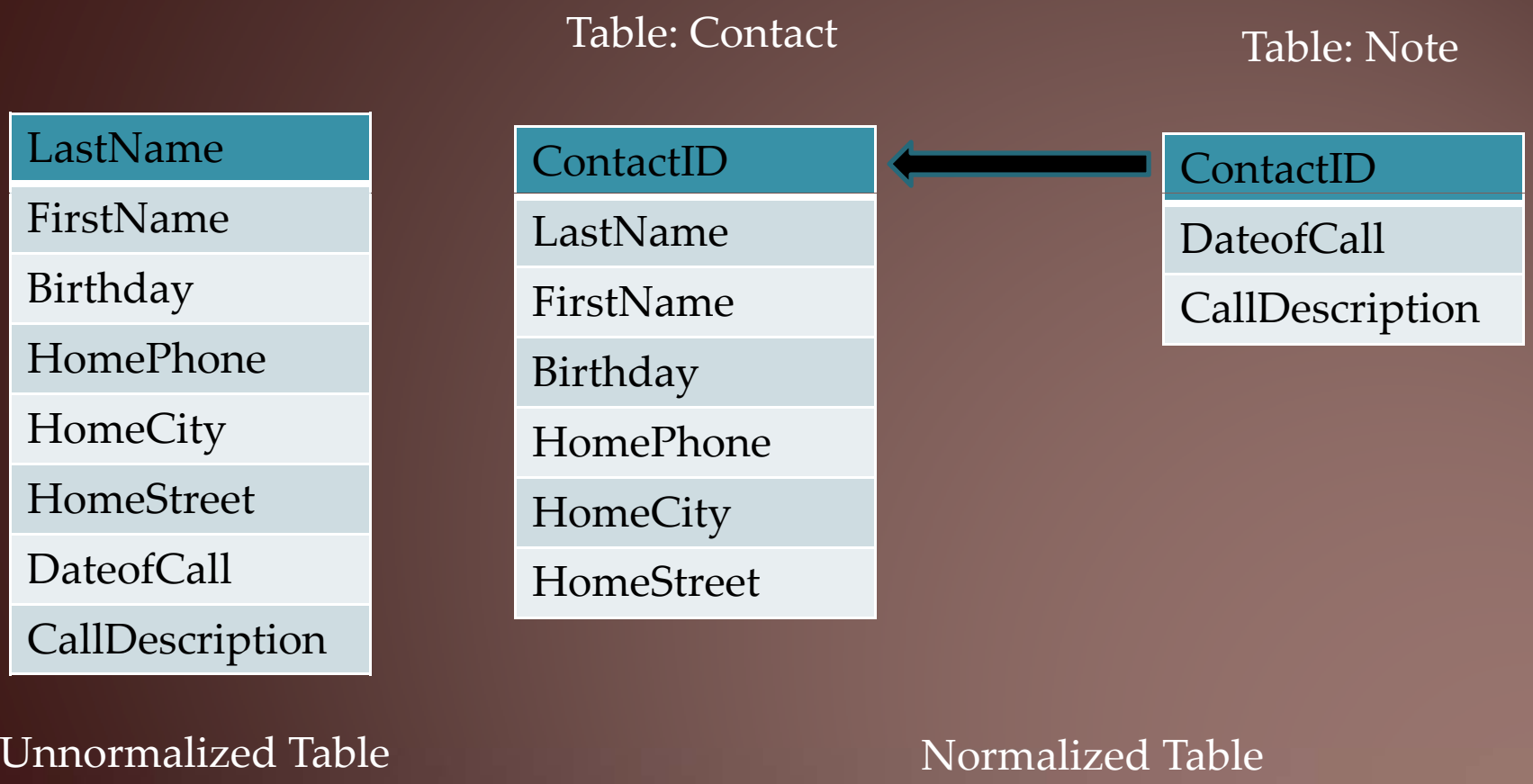
Normalization

- Normalization is the process of efficiently organizing data in a database with two goals in mind
- First goal: eliminate redundant data
  - for example, storing the same data in more than one table
- Second Goal: ensure data dependencies
  - for example, only storing related data in a table
- Is the process of splitting tables to minimize data redundancy and establishing relations between tables.
- Provides flexibility, data consistency and avoids anomalies while inserting, deleting and updating data

Benefits of Normalization

- Less storage space
- Quicker updates
- Less data inconsistency
- Clearer data relationships
- Easier to add data
- Flexible Structure

Normalization



Forms of Normalization

- The breaking down of a table may undergo series of stages called NORMAL FORMS.
- A higher level of normalization cannot be achieved unless previous levels have been satisfied.
  - First Normal Form (1NF)
  - Second Normal Form (2NF)
  - Third Normal Form (3NF)

First Normal Form (1NF)

- A table is said to be in first normal form if it has no repeating groups.
- For each cell in a table, there can be only one value.
- If a group of items repeats, it should be split into a new table.

SalesOrderNo
Date
CustomerName
CustomerAddr
ItemNo
Description
Quantity
Unit Price

Unnormalized Table

Table: Customer

SalesOrderNo
Date
CustomerNo
CustomerName
CustomerAddr

Table: Sales

SalesOrderNo
ItemNo
Description
Quantity
UnitPrice

Table in 1NF

Second Normal Form (2NF)

- A table is said to be in Second Normal Form if it is already in first normal form and every non key columns depends on the entire key.

Table: Sales

SalesOrderNo
ItemNo
Description
Quantity
UnitPrice

- Here, the key field is *SalesOrderNo*.
- The field description depends on *ItemNo* rather than *SalesOrderNo* whereas *Quantity* and *UnitPrice* are not dependent on *ItemNo* as they may be different for different sale order.
- In this case, the table is split again.
- The columns that depend on the key are kept in one table and rest on another table

Table: Sales

SalesOrderNo
ItemNo
Quantity
UnitPrice

Table: Items

ItemNo
Description

Table in 2 NF

Third Normal Form (3 NF)

- A table is said to be in Third Normal Form if is already in second normal form and if non key columns are not dependent on each other.
- There should not be any hidden dependencies among non key columns

Table: Customer

SalesOrderNo
Date
CustomerNo
CustomerName
CustomerAddr

- Here non key fields *CustomerName* and *CustomerAddr* depend on *CustomerNo* but not on *SalesOrderNo.*
- This hidden dependency is removed by splitting the table.

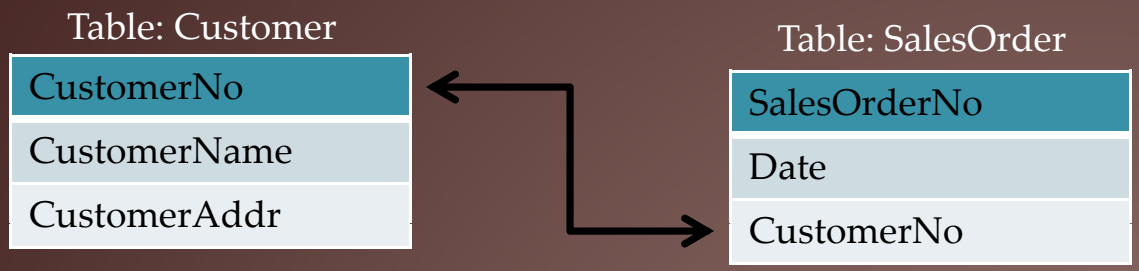
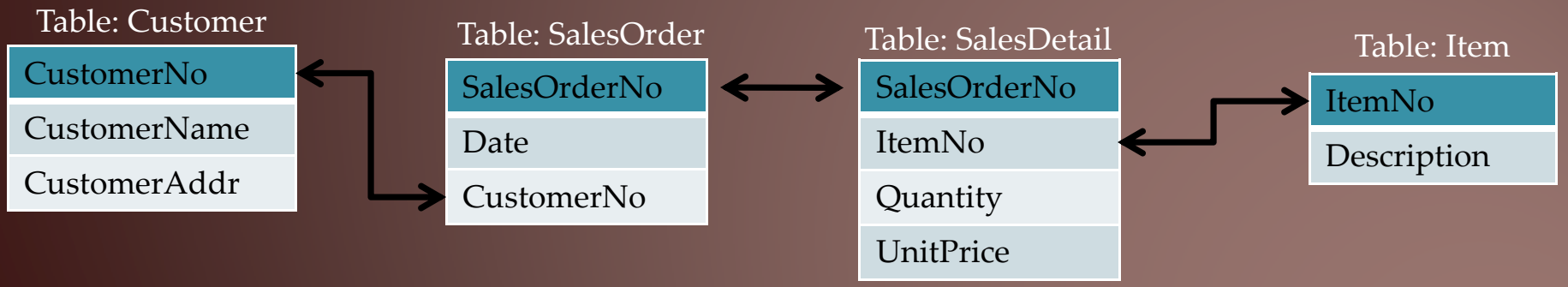


Table in 3NF



Final Normalized Table

Any Queries??