

## Normalization

- Formal process of decomposing relations with anomalies to produce smaller, **wellstructured** and **stable** relations
- *Primarily a tool to validate and improve a logical design so that it satisfies certain constraints that **avoid unnecessary***

## Well-Structured Relations

- A relation that contains minimal data redundancy and allows users to insert, delete, and update rows without causing data inconsistencies
- Goal is to avoid (*minimize*) anomalies
  - **Insertion Anomaly** – adding new rows forces user to create duplicate data
  - **Deletion Anomaly** – deleting a row may cause loss of other data representing completely different facts

## Functional Dependencies

- **Functional Dependency**: The value of one attribute (the *determinant*) determines the value of another attribute.
  - $A \rightarrow B$  reads “Attribute B is functionally dependent on A”
  - $A \rightarrow B$  means if two rows have same value of A they necessarily have same value of B
  - FDs are determined by **semantics**: **You can’t** say that a FD exists just by looking at data. But can say whether it does

# Functional Dependencies and Keys

- **Functional Dependency**: The value of one attribute (the *determinant*) determines the value of another attribute.
- **Candidate Key**
  - Attribute that uniquely identifies a row in a relation
  - Could be a combination of (*non-redundant*) attributes
  - Each non-key field is functionally dependent on every candidate key

