

Normalization is a process in database design that helps create "good" relations within a relational database schema. A relational database schema consists of a set of relations, and each relation represents a set of attributes. The way attributes are grouped into relations is crucial, and normalization or schema refinement plays a significant role in determining the quality of these relations.

Decomposition in database design can sometimes lead to problems that affect the quality and integrity of the relational schema. Two important properties of decompositions are the loss-less join property and the dependency preserving property. The loss-less join property ensures that no information is lost when combining the decomposed relations back together, while the dependency preserving property ensures that all functional dependencies are preserved during the decomposition process.

The closure of a set  $F$  of functional dependencies (FDs) represents the set  $F^+$  that includes all FDs that can be inferred from  $F$ . In other words,  $F^+$  contains all the additional dependencies that can be derived from the given set of FDs. The closure of a set of attributes  $X$  with respect to  $F$ , denoted as  $X^+$ , represents the set of all attributes that are functionally determined by  $X$ . It can be calculated by iteratively applying certain inference rules, such as IR1, IR2, and IR3, using the FDs in  $F$ .

The relational model is a conceptual framework for representing a database as a collection of relations. A relation in the relational model corresponds to a table in a database and consists of two main components: the relation schema and the relation instance. The relation schema describes the structure of the relation and includes information such as the name of the relation (table), the name of each field (attribute), and the domain of each field. The domain defines the set of permissible values for a particular attribute and is characterized by a domain name and its associated values.

In a database, a constraint can be established between two relations, referred to as the referencing relation and the referenced relation. The referencing relation contains tuples (rows) with foreign key attributes (FK) that reference the primary key attributes (PK) of the referenced relation. To visually represent this relationship, an arrow is drawn from the foreign key attribute in the referencing relation to the primary key attribute in the referenced relation.

To ensure data integrity and prevent inconsistencies, constraints or conditions are specified on a relational schema. These constraints, often referred to as integrity constraints (ICs), define rules that the data in the database must adhere to. A database is considered a legal instance when it satisfies all the constraints specified on its schema. The role of the database management system (DBMS) is to enforce these constraints, allowing only legal instances to be stored in the database. When the database schema is defined by the database administrator (DBA) or end user, they specify the constraints that must be met by the data.