What is decomposition?

Decomposition is the process of breaking down in parts or elements.

It replaces a relation with a collection of smaller relations.

It breaks the table into multiple tables in a database.

It should always be lossless, because it confirms that the information in the original relation can be accurately reconstructed based on the decomposed relations.

If there is no proper decomposition of the relation, then it may lead to problems like loss of information.

Properties of Decomposition

Following are the properties of Decomposition,

1. Lossless Decomposition

2. Dependency Preservation

3. Lack of Data Redundancy

1. Lossless Decomposition

Decomposition must be lossless. It means that the information should not get lost from the relation that is decomposed.

It gives a guarantee that the join will result in the same relation as it was decomposed.

Example:

Let's take 'E' is the Relational Schema, With instance 'e'; is decomposed into: E1, E2, E3, ... En; With instance: e1, e2, e3, ... en, If e1 ⋈ e2 ⋈ e3 . . . . ⋈ en, then it is called as 'Lossless Join Decomposition'.

In the above example, it means that, if natural joins of all the decomposition give the original relation, then it is said to be lossless join decomposition.

Example: <Employee_Department> Table

Eid	Ename	Age	City	Salary	Deptid	DeptName
E001	ABC	29	Pune	20000	D001	Finance
E002	PQR	30	Pune	30000	D002	Production

E003	LMN	25	Mumbai	5000	D003	Sales
E004	XYZ	24	Mumbai	4000	D004	Marketing
E005	STU	32	Bangalore	25000	D005	Human Resource

Decompose the above relation into two relations to check whether a decomposition is lossless or lossy.

Now, we have decomposed the relation that is Employee and Department.

Relation 1 : <Employee> Table

Eid	Ename	Age	City	Salary
E001	ABC	29	Pune	20000
E002	PQR	30	Pune	30000
E003	LMN	25	Mumbai	5000
E004	XYZ	24	Mumbai	4000
E005	STU	32	Bangalore	25000

Employee Schema contains (Eid, Ename, Age, City, Salary).

Relation 2 : <Department> Table

Deptid	Eid	DeptName
D001	E001	Finance
D002	E002	Production
D003	E003	Sales
D004	E004	Marketing
D005	E005	Human Resource

Department Schema contains (Deptid, Eid, DeptName).

So, the above decomposition is a Lossless Join Decomposition, because the two relations contains one common field that is 'Eid' and therefore join is possible.

Now apply natural join on the decomposed relations.

Employee ⋈ Department

Eid	Ename	Age	City	Salary	Deptid	DeptName
E001	ABC	29	Pune	20000	D001	Finance
E002	PQR	30	Pune	30000	D002	Production
E003	LMN	25	Mumbai	5000	D003	Sales
E004	XYZ	24	Mumbai	4000	D004	Marketing
E005	STU	32	Bangalore	25000	D005	Human Resource

Hence, the decomposition is Lossless Join Decomposition.

If the <Employee> table contains (Eid, Ename, Age, City, Salary) and <Department> table contains (Deptid and DeptName), then it is not possible to join the two tables or relations, because there is no common column between them. And it becomes Lossy Join Decomposition.

2. Dependency Preservation

Dependency is an important constraint on the database.

Every dependency must be satisfied by at least one decomposed table.

If {A → B} holds, then two sets are functional dependent. And, it becomes more useful for checking the dependency easily if both sets in a same relation.

This decomposition property can only be done by maintaining the functional dependency.

In this property, it allows to check the updates without computing the natural join of the database structure.

3. Lack of Data Redundancy

Lack of Data Redundancy is also known as a Repetition of Information.

The proper decomposition should not suffer from any data redundancy.

The careless decomposition may cause a problem with the data.

The lack of data redundancy property may be achieved by Normalization process.

Decomposition is lossless if:

X intersection Y -> X, that is: all attributes common to both X and Y functionally determine ALL the attributes in X.

X intersection Y -> Y, that is: all attributes common to both X and Y functionally determine ALL the attributes in Y

If X intersection Y forms a superkey of either X or Y, the decomposition of R is a lossless decomposition.