DECOMPOSITION

Decomposition means divide a large and complex table into multiple small and easy tables. This is done to remove redundancy, anomalies and inconsistency of a database. This is a first stage towards normalization.

Suppose if we have a relational schema R, in which we have attributes as per given below

A_1, A_2, A_3…………A_n

So R = {A_1, A_2, A_3…………A_n}

If we decompose it in small parts then R will be divided in following parts:

R_1, R_2……..Rx

These all relational schemas belongs to the original one R

R_1, R_2……..Rx =R

Also we can write that union of all these subset belongs to original set R

R_1 U R_2 U R_3 ……..U Rx = R

Here R_1, R_2……..Rx <= R

Also 1<= i <= x (i= number of relation like 1,2,3…..x)

Decomposition is further divided into two parts:

1. Lossless Decomposition
2. Lossy Decomposition

LOSSLESS DECOMPOSITION:

Loss means data loss while decomposing a relational table. A lossless decomposition is somewhat in which data is not lose because JOIN is used.

First, we decompose a large table into small appropriate tables, then applied natural join to reconstruct the original table.

This is a student database relational table:

Sid	Name (Not Null)	Subject (Not Null)	Mobile	Address
1	Raj	English	65468154	51, Vaishalinagar
2	Jyoti	Home Science	87668545	4a, Sukhsagar
3	Vikash	Maths	26865948	H7, Civil Lines
1	Harsh	Maths	Null	R32, Gokul Villa
3	Ajay	Science	86516529	26, Karoli

We can decompose it in two simple tables as per given below:

Student Subject Details

Sid	Name (Not Null)	Subject (Not Null)
1	Raj	English
2	Jyoti	Home Science
3	Vikash	Maths
1	Harsh	Maths
3	Ajay	Science

Student Personal Details

Sid	Mobile	Address
1	65468154	51, Vaishalinagar
2	87668545	4a, Sukhsagar
3	26865948	H7, Civil Lines
1	Null	R32, Gokul Villa
3	86516529	26, Karoli

If we want to see a common table then we can apply Natural JOIN between both tables like this:

Student Subject Details ⋈ Student Personal Details

Sid	Name (Not Null)	Subject (Not Null)	Mobile	Address
1	Raj	English	65468154	51, Vaishalinagar
2	Jyoti	Home Science	87668545	4a, Sukhsagar
3	Vikash	Maths	26865948	H7, Civil Lines
1	Harsh	Maths	Null	R32, Gokul Villa
3	Ajay	Science	86516529	26, Karoli

In this operation no data loss occurs, so this is good option to consider for decomposition.

LOSSY DECOMPOSITION:

In this, the decomposition is performed in such a manner that the data will be loose. Let's take an example:

Sid	Name (Not Null)	Subject (Not Null)	Mobile	Address
1	Raj	English	65468154	51, Vaishalinagar
2	Jyoti	Home Science	87668545	4a, Sukhsagar
3	Vikash	Maths	26865948	H7, Civil Lines
1	Harsh	Maths	Null	R32, Gokul Villa
3	Ajay	Science	86516529	26, Karoli

If we divide this student details table in two sections as per given below:

Student Subject Details

Sid	Name (Not Null)	Subject (Not Null)
1	Raj	English
2	Jyoti	Home Science
3	Vikash	Maths
1	Harsh	Maths
3	Ajay	Science

Student Personal Details

Mobile	Address
65468154	51, Vaishalinagar
87668545	4a, Sukhsagar
26865948	H7, Civil Lines
Null	R32, Gokul Villa
86516529	26, Karoli

In this Student Personal Details table, the SID column is not included, so now we don't know that these mobiles numbers and address belongs to whom.

So always decompose a table in such a manner that the data may be easily reconstructed and retrieved.

UNIVERSAL RELATION:

As we know: $R = \{A_1, A_2, A_3 \ldots\ldots\ldots\ldots A_n\}$

Also we know that $R_1, R_2 \ldots\ldots\ldots Rx = R$

And $R_1 \cup R_2 \cup R_3 \ldots\ldots\ldots \cup Rx = R$

So here $R = \{A_1, A_2, A_3 \ldots\ldots\ldots\ldots A_n\}$is also known as universal Relation because R holds all the attributes.

We can write Universal Relation as **R(U)**.