Third Normal Form

Normalisation

Normalisation is the process used to ensure that the logical design process has produced high-quality structures for the tables within a database. By quality, we mean structures that will perform optimally, and do not contain any duplicate information.

If a logical structure is normalised, then each attribute in the structure will have the same importance. This is useful during the physicalisation process, when the database is actually created using SQL, because indexes can be assigned to each attribute to make searching and retrieval on that attribute faster.

Normalisation is a useful verification tool, which indicates amendments to be made to the logical design process, but does not substitute for the conceptual and logical design. All design stages should be undertaken to ensure that the design covers all aspects of the required system.

The Normalisation Process

The process of normalising a logical model can be summarised as follows:

1. Convert all table structures to First Normal Form
2. Identify Functional Dependencies
3. Convert all table structures from First Normal Form to Second Normal Form
4. Convert all table structures from Second Normal form to Third Normal Form and (very rarely) into Boyce Codd Normal Form

Problems with Second Normal Form

A relation in second normal form can still show data insert and update anomalies, as shown by the ORDERS table below which is in Second Normal Form (key fields in this case are Customer Name, Product ID and Date).

ORDERS

Customer Name	**Product ID**	**Date**	Quantity	%Discount
James Smith	23	12/12/2003	50	10
James Smith	24	15/12/2003	100	15
Martin Jones	23	2/11/2002	50	10
Alex Hayley	23	15/1/2003	150	20

Considering the operations we wish to do to the table:

1. Update – what happens if the quantity for the first sale to James Smith changes? This means that the value in the %Discount field must also be changed.
2. Insert - A value for %Discount cannot be inserted until a value for quantity is known.

For database tables **Functional Dependencies** describe how the value of one attribute depends on the value of another – i.e. when you change one attribute, you must also change another. They need to be eliminated between NON-KEY fields.

Listing the functional dependencies for the above table:

- Quantity → %Discount
- Customer Name, Product ID, Date → Quantity
- Customer Name, Product ID, Date → %Discount

In other words:

- the value of %Discount is determined by the Quantity purchased, as well as by the values of Customer Name, Product ID and Date.
- the value for Quantity is in turn dependant on the values of Customer Name, Product ID and Date.

This situation can be described in general as follows:

- The value of a non-key attribute A must be changed when the value of another non-key attribute B changes
- The value of B must be changed when the primary key changes

The situation is known as a **transitive dependence**.

Third Normal Form

Formally: A relation is in third normal form (3NF) if and only if it is in Second Normal Form (2NF) and every non-key attribute is non-transitively dependent on the primary key.

Therefore a table is in third normal form only if the NON-KEY attributes are:

- Mutually independent (one non-key field is not dependent on any other non-key fields)
- Fully dependant on the primary key

A NON-KEY attribute is any attribute that does not participate in the primary key of the table concerned. Attributes are mutually independent if none of them is functionally dependant on any combination of the others. Independence implies that any attribute can be updated separately to any other, without affecting the value of any other.

The table above should therefore be decomposed (split into two or more tables) to overcome the problems caused by Second Normal Form. This decomposition will bring the resulting tables into Third Normal Form.

Customer Name	**Product ID**	**Date**	Quantity
James Smith	23	12/12/2003	50
James Smith	24	15/12/2003	100

Martin Jones	23	2/11/2002	50
Alex Hayley	23	15/1/2003	150

Quantity	%Discount
50	10
100	15
150	20