Database Management System (DBMS)

Lecture-33

Dharmendra Kumar

November 8, 2020

Normalization

To overcome these anomalies we need to normalize the data. In the next section we will discuss different types of normal forms. These normal forms are:-

1. First normal form(1NF)
2. Second normal form(2NF)
3. Third normal form(3NF)
4. Boyce-Codd normal form(BCNF)
5. Fourth normal form(4NF)
6. Fifth normal form(5NF)

First normal form(1NF):

As per the rule of first normal form, an attribute (column) of a table cannot hold multiple values. It should hold only atomic values.

Consider the following table:-

emp_id	emp_name	emp_address	emp_mobile
101	Herschel	New Delhi	8912312390
102	Jon	Kanpur	8812121212 , 9900012222
103	Ron	Chennai	7778881212
104	Lester	Bangalore	9990000123 , 8123450987

Table 1: Employee

This table is not in 1NF as the rule says "each attribute of a table must have atomic (single) values", the emp_mobile values for employees Jon & Lester violates that rule.

To make the table complies with 1NF we should have the data like this:

emp_id	emp_name	emp_address	emp_mobile
101	Herschel	New Delhi	8912312390
102	Jon	Kanpur	8812121212
102	Jon	Kanpur	9900012222
103	Ron	Chennai	7778881212
104	Lester	Bangalore	9990000123
104	Lester	Bangalore	8123450987

Table 2: Revised Employee table

Now this table is in 1NF.

Second normal form(2NF)

Consider a relation schema R with set of functional dependencies F.
A relation schema R is said to be in second normal form(2NF) if
(1) It is in 1NF.
(2) Every non-prime attribute is fully functionally dependent on candidate key.

Prime attribute

An attribute which is a part of any candidate key is said to be prime attribute. And which is not part of any candidate key is said to be non-prime attribute.

Partial functional dependency

A partial dependency means if the non-key attributes depend on the part of candidate key then it is said to be partial dependency.
An attribute is fully functional dependent on a set of attributes α, if it is functionally dependent on only α and not on any of its proper subset.

Example: Consider a relation schema R = (A,B,C,D) with only one candidate key as {A,B}. In this case, A and B are prime attributes. C and D are non-prime attributes.

Note: A relation with a single-attribute primary key is automatically in at least 2NF.

Example: Consider following functional dependencies in relation R(A, B, C, D)
AB $\rightarrow$ C
BC $\rightarrow$ D
Is this relation in 2NF?
Solution:
First find candidate keys. Here candidate key is A,B. Clearly non-prime attributes are C and D. From functional dependencies, C and D are fully functional dependent, therefore R is in 2NF.

Example: Consider a relation- R (V , W , X , Y , Z) with functional dependencies-

VW $\rightarrow$ XY

Y $\rightarrow$ V

WX $\rightarrow$ YZ

Is this relation in 2NF?

Solution:

Here candidate keys are VW, WX and WY. Therefore non-prime attributes are Z. Clearly Z is fully dependent on candidate keys. Therefore R is in 2NF.

Example: Consider relation R(A, B, C, D, E) with set of following functional dependencies

$A \rightarrow BC$,

$CD \rightarrow E$,

$B \rightarrow D$,

$E \rightarrow A$

Is this relation in 2NF?

Solution:

Here candidate keys are A, E, CD, BC. Clearly all the attributes belong into some candidate keys, therefore no non-prime attribute. Therefore, R is in 2NF.

Third normal form(3NF)

A relation schema R is in third normal form (3NF) with respect to a set F of functional dependencies if, for all functional dependencies in F^+ of the form $\alpha \to \beta$, where $\alpha \subseteq R$ and $\beta \subseteq R$, at least one of the following holds:

1. $\alpha \to \beta$ is a trivial functional dependency.
2. α is a super key for R.
3. Each attribute A in $\beta - \alpha$ is contained in a candidate key for R.

In other words, we can define 3NF as following:-
A relation schema R with set of functional dependencies F is said to in 3NF if
(1) It is in 2NF.
(2) No non-prime attribute transitively depends on any candidate key.

Example: Consider a relation- R (V , W , X , Y , Z) with functional dependencies-

$VW \rightarrow XY$

$Y \rightarrow V$

$WX \rightarrow YZ$

Is this relation in 3NF?

Solution:

Here candidate keys are VW, WX and WY. Therefore non-prime attributes are Z. Clearly Z is fully dependent on candidate keys. Therefore R is in 2NF.

Now, Z is not transitively dependent on any candidate key. Therefore, it is in 3NF.

Example: Consider relation R(A, B, C, D, E) with set of following functional dependencies

A $\rightarrow$ BC,

C $\rightarrow$ E,

B $\rightarrow$ D,

Is this relation in 3NF?

Solution:

Here candidate key is A. Therefore, non-prime attributes are B, C, D, E. Since candidate key contains single attribute, therefore R is in 2NF.

Now, Clearly attributes D and E are transitively dependent on candidate key A, therefore R is not in 3NF.

Example: The relation schema Student_Performance (name, courseNo, rollNo, grade) has the following FDs:

name,courseNo $\rightarrow$ grade

rollNo,courseNo $\rightarrow$ grade

name $\rightarrow$ rollNo

rollNo $\rightarrow$ name

Is this relation in 3NF?

Solution:

Here candidate keys are {name, courseNo} and {rollNo, courseNo}. Therefore non-prime attributes are grade. Clearly, grade is fully functional dependent on candidate keys, therefore this relation schema is in 2NF.

Now, clearly grade is not transitively dependent on candidate keys, therefore this relation schema is in 3NF.