

UNIT \Rightarrow '4'

Normalization

Abhay Kumar Mishra
Lecturer, Dept. of B.C.A.
Maharaja College, Ara (V.K.S.U.)

Relation Database design Issue

A relation scheme R is a plan that the attributes involved in one or more relation. attributes $\{A_1, \dots, A_n\}$ where A_i is defined as domain D_i for $1 \leq i \leq n$.

Relation R have set tuples $\{t_1, \dots, t_j\}$ such such that $t_j \in R$.

The attribute value $t_j(A_i)$ must be in the corresponding domain D_i .

* Anomalies in Database

1) Redundancy

The aim of database system is to reduce redundancy, meaning that information is to be stored only once. Do not stored information several.

2) Update Anomalies

Multiple copies of same fact must be lead to update or inconsistencies when an update is made.

Abhay Kumar Mishra

Lecturer, Dept. of B.C.A.

(University of Rajasthan, Jaipur)

Insertion Anomalies :-)

Always information stored in these relation, been consistent.

Deletion :-)

If we want to delete certain information then always we delete the information.

Keys :-)

A key is that data item that exclusively identifies a record.

Example :->

A/c-no , Emp-no.

Primary Key :-)

The Primary key uniquely identifies each recording in a table & must never be the same for two record.

Alternate Key :-> Other key accept primary key.

• Candidate key :-

A candidate key is an attribute or set of attributes that uniquely identifies a record.

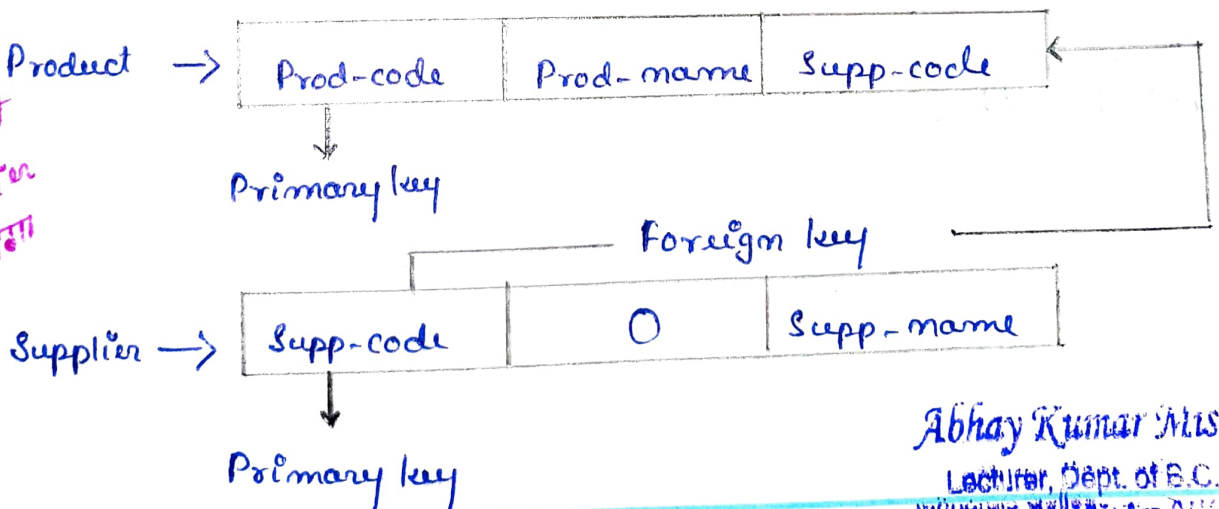
- Composite key :- when we will use more than one column as part of the primary key is called composite (concatenated key) key.
- Concatenated key

Table Super key
 में Primary key रहेगा
 Table में Comp. key
 में Primary नहीं होगा.

• Foreign key :-

The column whose data values correspond to the values of a key column in another relation is called a foreign key.

Primary key
 के साथ कोई
 attribute
 Unique Identification
 हो तो Super key होगा



• Super key :-)

If we add additional attributes to a Primary key, the resulting combination would still uniquely identify an instance of the entity set.

Such keys are called Super keys:

A primary key is, therefore, a minimal super key.

• Secondary key :-)

Secondary key is an attribute or combination of attributes that may not be a but classifies the entity set on a particular characteristics.

For example :->

The entity set EMPLOYEE having the attribute Department, which identifies by its value which means all instances of EMPLOYEE who belong to a given department.

Normalization :-)

It is the process of efficiently organizing data in a database.

✘ There are two goals of Normalization Process.

1) Eliminate Redundant data.

(For ex → Storing the same data in more than one table).

2) Removes duplicated data from the relational tables.

Ensure data dependencies make sense.

(only storing related data in a table).

OR

✘ Normalization is essentially a two step process & puts data into tabular form by removing repeating groups & then removes duplicated data from the relation table.

✘ The database community has developed a series of guidelines for ensuring that database are normalized. These are referred to as normal forms & numbered from one (the lowest form of normalization, referred to as 1NF through Five (5NF).
First Normal Form

i.e. 1NF, 2NF, 3NF -----).

Abhay Kumar Mishra

Lecturer, Dept. of B.C.A.
Maharaja College, Ara, Bihar

* Introduction to Functional Dependency

Given a Relation R, attribute Y of R is functionally dependent on attribute X of R if & only if each value X-value in R is associated with its precisely one Y-value in R.

It is denoted by $X \rightarrow Y$.

Example:-

S.No	Name	Status	City
S ₁	Rom	20	Ara
S ₂	ROM	10	Patna
S ₃	TON	30	Delhi
S ₄	MON	20	Kolkata

S.S.No \rightarrow S.Name

S.S.No \rightarrow S.City

S.S.No \rightarrow S.Status

That S.S.No \rightarrow S.city is read as

"attribute S.city is functionally dependent on attribute by S.S.No."

Example:-

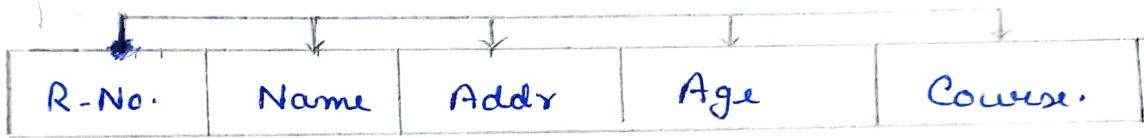
prof \rightarrow course

emp-code \rightarrow Address;

(Each employee has an unique address)-

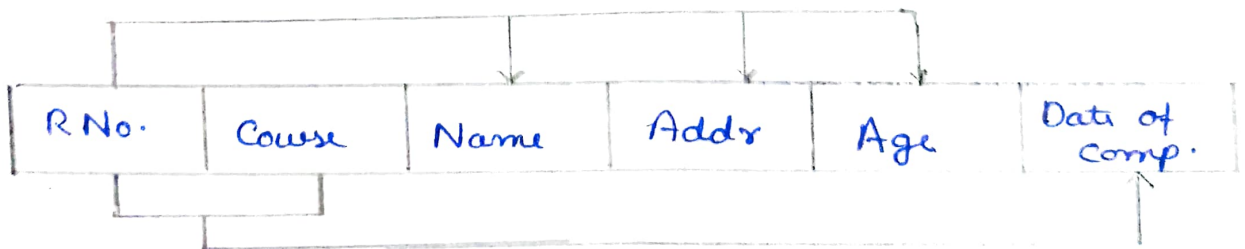
* Full Functional Dependency \Rightarrow

When all non key attributes are dependent on key attribute it is called Full Functional dependency.



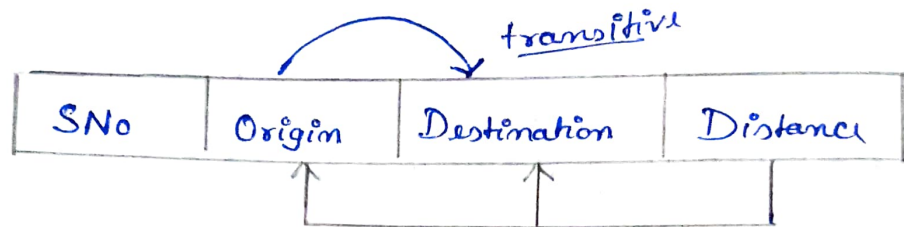
* Partial Dependency \Rightarrow

Partial Dependency in a record type occurs when some non-key attribute depends on key attribute & on one or more non-key attributes.



* Transitive Dependency ^{o²}

$X \rightarrow Y$, $Y \rightarrow Z$
So, $X \rightarrow Z$



* 1NF \rightarrow (First Normal Form)

- * No NULL values in any attributes of any Relation.
- * A table is in the 1NF when it contains no repeating groups.
i.e. there is not NULL value.
- * The repeating column or field present in an unnormalized table are removed for the table & put into separate table or tables.
- * The key to these tables must also be a part of the parent table.

Emp-No.	Emp-Name	Store branch	Dept.	Item Number	Item Description	Sale Price (₹)
211	RON	Downtown	Hardware	TR10	Router	35.00
				SA 1	Saw	19.00
				PT 6	Drill	21.00
301	MON	Dadeland	Home Appliances	TT 1	Humidifier	114.00
				DS10	Dishwasher	262.00
401	TON	Cutler point	Auto parts	MC16	Dishwasher	262.00
				AC146	Snow fire	85.00
				BB100	Alternator	65.00
501	JON	Fashion spot	Men's clothing	HS10	Scarf	215.00

Unnormalized files for sales.

X	Sum	Y	Z
Emp-no.	Emp-name	Store	Dept.
✓	✓	✓	✓
✓	✓	✓	✓
✓	✓	✓	✓
✓	✓	✓	✓

* Emp-no.	* Item no.	Item description	Sales Price.
211	TR10	Router	35.00
211	SA 1	Saw	19.00
211	PT 6	Drill	21.00

Emp-no.	Item-no	Sale price
211	TR10	35.00
301	SA1	19.00
302	PT6	21.00

Item-no.	Item Descripti-
TR10	Router
SA1	Saw
PT6	Drill

and also primary key that should be

2NF → (Second Normal Form)

* A table is in 2NF if all its ~~non~~ in 1NF every non-key column is or fully dependent upon the primary key.

* In 1NF a partial dependency is not present.

All other non-key are depend on primary key & higher any key depend on another non-primary key.

3NF → (Third Normal Form)

In 2NF a this should be no transitive depend between attributes.

* / emp#	emp-name	Store branch

* / Store branch	Dept -

In 2NF & Remove columns that are not dependent upon the Primary key.

Emp-no.	Emp-name	Store branch	Department	Item No.	Item Description	Sale Price (₹)
211	Abh	Downtown	Hardware	TR10	Router	35.00
				SA1	Saw	19.00
				PT6	Drill	21.00
				AB 6	lawn	245.00
311	Ajh	Dadiland	Home application	T11	Humidifier	114.00
				DS10	Dish washer	262.00
411	Akh	Cutlerpoint	Auto parts	MC16	Snow tire	85.00
				AC146	Alternator	65.00
				BB100	Battery	49.00
511	Ach	Fashion Spot	Men's clothes	HS10	Suit	215.

Un-normalized Table.

Abhay Kumar Mishra

Lecturer, Dept. of B.C.A.
 VIT-AP, VIT-AP, VIT-AP

1NF → There is no NULL value.

2NF → The 1NF & full functional dependency.

3NF → There is 2nd & there is no transitive dependency between two attribute.

BCNF → The 3NF & there should be only one candidate key.

1NF

* Emp-no.	Emp-name	Store Branch	Department
-----------	----------	--------------	------------

* Emp-no.	* Item-no.	Sale price	Item Description
-----------	------------	------------	------------------

2NF ✓

Item-no.	Dept.
----------	-------

Ist table से

X		Y	Z
* Emp.	Emp-no.	Store Branch	Department

Emp-name	Emp-no.	Store Branch
----------	---------	--------------

Store Description	Department
-------------------	------------

BCNF → (BOYCE - CODD NORMAL Form).

* BCNF was proposed as a simpler form of 3NF, but it is much more strict than 3NF.

Means that every relation in BCNF is also in 3NF. However a relation in 3NF is not necessarily be in BCNF.

* In BCNF should be only one candidate key in the attribute.

The Grade Relation.

<u>Name of Student</u>	<u>Student # no.</u>	<u>Course</u>	<u>Grade</u>
James	23	353	A
Nulam	24	329	A
James	23	328	in prog
Mohan	25	456	C
Dilip	26	293	B
Deepak	27	491	C
Deepak	27	353	in prog
James	23	491	C
Raj	28	353	A+
Vikash	29	399	in prog.

Abhay Kumar Mishra

Lecturer, Dept. of B.C.A.
V.K.S.U.

Func dep

Name course \rightarrow Grade

Student # course \rightarrow Grade

i.e.

Name \rightarrow Student #

Student # \rightarrow Name?

* Thus each student has a unique name & unique student no.

* The relation has two candidate key
(Name, course) & (Student #, course)

These are two candidate key so ~~it~~ it is in BCNF.

==