



GALGOTIAS
UNIVERSITY

**School of Computing
Science and Engineering**

Program: BCA

Course Code: BCAC2102

Course Name: Database Management System

Lecture-18

Topic- 1NF

Faculty:-Dr. Satyajee Srivastava

Lecture-17

Topic- Normal Forms

Objective :

To acquire knowledge about Normal Forms

Lecture-17

Topic- Normal Forms

Objective :

To acquire knowledge about Normal Forms

Lecture-17

Topic- Normal Forms

Objective :

To acquire knowledge about Normal Forms

Lecture-17

Topic- Normal Forms

Objective :

To acquire knowledge about Normal Forms

Lecture-17 Recap

Topic- Normal Forms

Objective :

To acquire knowledge about Normal Forms

Lecture-17

Topic- Normal Forms

Objective :

To acquire knowledge about Normal Forms

Lecture-17

Topic- Normal Forms

Objective :

To acquire knowledge about Normal Forms

Lecture-17

Topic- Normal Forms

Objective :

To acquire knowledge about Normal Forms

Lecture-17

Normalization

Normalization: is the process of “fixing” relational schemata so that they avoid three closely related kinds of problems.

Storage redundancy: The same information is repeated many times.

Unnecessary information dependency: Information about some x cannot be represented without having at least corresponding instance of y .

Update anomalies: The way in which data is represented complicates the support of certain kinds of updates.

Lecture-17

Illustration of Problems of an Unnormalized Schema

Firm

<u>SSN</u>	Name	Dept	Bldg
000112222	Alice	3	8
000113333	Bruce	3	8
000114444	Carol	3	8
000115555	David	5	7
000116666	Alice	4	7

$SSN \rightarrow \{Name, Dept\}$

$Dept \rightarrow Bldg$

- The FD $Dept \rightarrow Bldg$ does not define a key and leads to problems.

Lecture-17

Storage redundancy: The information about Department 3 is repeated three times.

Update anomaly: If the building of Department 3 is to be changed, three updates are necessary.

Unnecessary information dependency:

- Information about an employee who does not have a department requires null values.
- Information about a department cannot be represented unless at least one employee works in it.

Lecture-17

Approaches to Normalization

Approaches to normalization: There are two principle approaches to normalization, and each will be considered in these slides.

Decomposition: Break larger relations into smaller ones.

Synthesis: Begin with a set of dependencies (usually FDs), and construct a corresponding relational schema.

The changes forced by normalization: Generally speaking, by forcing FDs to define (super)key dependencies, the problems identified above are minimized or disappear completely... but the devil is in the details.

Lecture-17

Normal Forms

Normal forms: In early research on the relational model, a number of so called *normal forms* were developed.

- The principal ones which are based upon FDs were developed in the following order:

1NF → 2NF → 3NF → BCNF

- There are some others which are based upon other types of dependencies: 4NF, 5NF, DKNF.



Lecture-17

- For pedagogical reasons, they will be considered in the reverse order of development:

BCNF → 3NF → 2NF → 1NF

- The main focus will be upon BCNF and 3NF, as 2NF is largely of historical interest and 1NF is just a constraint on domains.

Lecture-17(Normal Forms)

Design "Anomalies"

This Apply relation exhibits three types of anomalies:

- 1.Redundancy
- 2.Update Anomaly
- 3.Deletion Anomaly

Normalization

Design "Anomalies"

This Apply relation exhibits three types of anomalies:

- 1.Redundancy
- 2.Update Anomaly
- 3.Deletion Anomaly

Lecture-17

(Class-Assignment)

Explain Normalization with design issues.

Lecture-18

Topic- 1NF

Objective :

To acquire knowledge about 1-Normal Forms

Lecture-18

Normalization is a process that “improves” a database design by generating relations that are of higher normal forms.

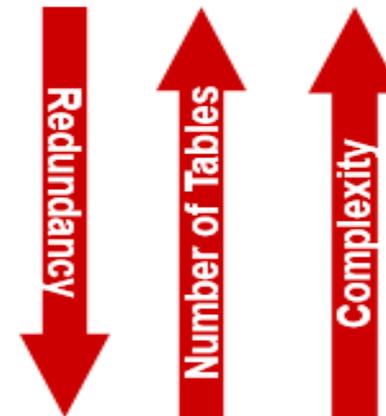
The *objective* of normalization:

“to create relations where every dependency is on the key, the whole key, and nothing but the key”.

Lecture-18

Levels of Normalization

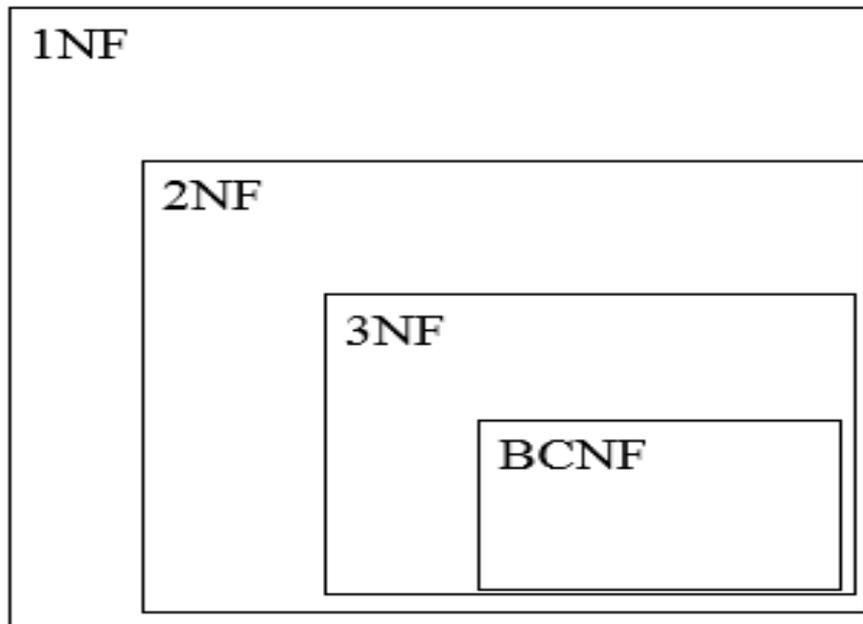
- Levels of normalization based on the amount of redundancy in the database.
- Various levels of normalization are:
 - First Normal Form (1NF)
 - Second Normal Form (2NF)
 - Third Normal Form (3NF)
 - Boyce-Codd Normal Form (BCNF)
 - Fourth Normal Form (4NF)
 - Fifth Normal Form (5NF)
 - Domain Key Normal Form (DKNF)



Most databases should be 3NF or BCNF in order to avoid the database anomalies.

Lecture-18

Normalization



a relation in BCNF, is also in 3NF

a relation in 3NF is also in 2NF

a relation in 2NF is also in 1NF

Lecture-18

- According to Date's definition of 1NF, a table is in 1NF if and only if it is "isomorphic to some relation", which means, specifically, that it satisfies the following five conditions:
 1. There's no top-to-bottom ordering to the rows.
 2. There's no left-to-right ordering to the columns.
 3. There are no duplicate rows.
 4. Every row-and-column intersection contains exactly one value from the applicable domain (and nothing else).
 5. All columns are regular [i.e. rows have no hidden components such as row IDs, object IDs, or hidden timestamps].

Lecture-18

First Normal Form

We say a relation is in **1NF** if all values stored in the relation are single-valued and atomic.

1NF places restrictions on the structure of relations. Values must be simple.

Lecture-18

First Normal Form

The following is **not** in 1NF

<u>EmpNum</u>	EmpPhone	EmpDegrees
123	233-9876	
333	233-1231	BA, BSc, PhD
679	233-1231	BSc, MSc

EmpDegrees is a multi-valued field:

employee 679 has two degrees: *BSc* and *MSc*

employee 333 has three degrees: *BA*, *BSc*, *PhD*

Lecture-18

To obtain 1NF relations we must, without loss of information, replace the above with two relations

Lecture-18

First Normal Form

Employee

EmpNum	EmpPhone
123	233-9876
333	233-1231
679	233-1231

EmployeeDegree

EmpNum	EmpDegree
333	BA
333	BSc
333	PhD
679	BSc
679	MSc

An outer join between Employee and EmployeeDegree will produce the information we saw before

Lecture-18

A relation is in 1NF if and only if all underlying domains contain atomic or scalar values.

For example

An attribute **customer-id** has a atomic domain but **address** attribute may not have a atomic domain as it can be divided further into city, street and road number etc.

Lecture-18

Example 2.

Unnormalized relation:

Defn: An unnormalized relation contains non atomic values.

Each row may contain multiple set of values for some of the columns, these multiple values in a single row are also called non atomic values.

<i>Order no.</i>	<i>Order date</i>	<i>Item code</i>	<i>Quantity</i>	<i>Price/unit</i>
1456	260289	3687	52	50.40
		4627	38	60.20
		3214	20	17.50
1886	040389	4629	45	20.25
		4627	30	60.20
1788	040489	4627	40	60.20

Lecture-18

<i>Order no.</i>	<i>Order date</i>	<i>Item code</i>	<i>Quantity</i>	<i>Price/unit</i>
1456	260289	3687	52	50.40
1456	260289	4627	38	60.20
1456	260289	3214	20	17.50
1886	040389	4629	45	20.25
1886	040389	4627	30	60.20
1788	040489	4627	40	60.20



Thank You