GALGOTIAS UNIVERSITY

School of Computing Science and Engineering

Program: BCA

Course Code:BCAC2102

Course Name: Database Management System

Lecture-23

Topic- Basic concepts of Data Models

Faculty:-Dr. Satyajee Srivastava

Lecture-22(RECAP)

Topic- Other Normalization Form

Objective : To acquire knowledge about Other Normalization Form

Lecture-22

Multivalued Dependencies

- Let R be a relation schema, and X and Y be disjoint subsets of R (i.e., X ⊆R, Y⊆ R, X∩Y=ϕ), and Z = R- XY.A relation r(R) satisfies X↠ Y if for any two tuples t_1 and t_2,
  - $t_1(X)=t_2(X)$, then there exist t_3 in r such that
  - $t_3(X)=t_1(X)$, $t_3(Y)=t_1(Y)$, $t_3(Z)=t_2(Z)$.
  - By symmetry, there exist t4 in r such that
  - $t_4(X)=t_1(X)$, $t_4(Y)=t_2(Y)$, $t_4(Z)=t_1(Z)$.

	X	Y	Z
t1	x1	y1	z1
t2	x1	y2	z2
t3	x1	y1	z2
t4	x1	y2	z1

- The MVD X↠ Y says that the relationship between X and Y is independent of the relationship between X and R-Y
- For example consider the table **Employee**:

Employee-name	Project-name	Dependant-name
Smith	X	John
Smith	Y	Ann
Smith	X	Ann
Smith	Y	John

Lecture-22

- MVDs Employee-name↠ Project-name and Employee-name↠ Dependant-name hold in the relation
- The employee named Smith works on projects X and Y, and has two dependents John and Ann.
- If we store only the first two tuples in the relation, it would incorrectly show the associations among attributes
- If we have MVDs in a relation, we may have to repeat values redundantly in the tuples. In the Employee relation, values X and Y of Project-name are repeated with each value of Dependant-name--- clearly undesirable
- Problem: Employee schema is in BCNF because no FDs hold for it
- Trivial MVD: If MVD X Y is satisfied by all relations whose schemas include X and Y, it is called trivial MVD.
  - $X \twoheadrightarrow Y$ is trivial whenever $Y \subseteq X$ or $X \cup Y = R$
- If a relation r fails to satisfy a given MVD, a relation r' that satisfies the MVD can be constructed by adding tuples to r.
  - MVD is called "tuple generating dependency"
  - compare it with FD: need to delete tuples to make the relation to satisfy a given FD

Fourth Normal Form (4NF)

- Fourth normal form eliminates independent many-to-one relationships between columns.
- To be in Fourth Normal Form,
  - a relation must first be in Boyce-Codd Normal Form.
  - a given relation may not contain more than one multi-valued attribute.

Example (Not in 4NF)

Scheme → {MovieName, ScreeningCity, Genre)

Primary Key: {MovieName, ScreeningCity, Genre)

1. All columns are a part of the only candidate key, hence BCNF
2. Many Movies can have the same Genre
3. Many Cities can have the same movie
4. Violates 4NF

Movie	ScreeningCity	Genre
Hard Code	Los Angles	Comedy
Hard Code	New York	Comedy
Bill Durham	Santa Cruz	Drama
Bill Durham	Durham	Drama
The Code Warrier	New York	Horror

Fourth Normal Form (4NF)

Example 2 (Not in 4NF)

Scheme → {Manager, Child, Employee}

1. Primary Key → {Manager, Child, Employee}
2. Each manager can have more than one child
3. Each manager can supervise more than one employee
4. 4NF Violated

Manager	Child	Employee
Jim	Beth	Alice
Mary	Bob	Jane
Mary	NULL	Adam

Example 3 (Not in 4NF)

Scheme → {Employee, Skill, ForeignLanguage}

1. Primary Key → {Employee, Skill, Language }
2. Each employee can speak multiple languages
3. Each employee can have multiple skills
4. Thus violates 4NF

Employee	Skill	Language
1234	Cooking	French
1234	Cooking	German
1453	Carpentry	Spanish
1453	Cooking	Spanish
2345	Cooking	Spanish

4NF - Decomposition

1. Move the two multi-valued relations to separate tables
2. Identify a primary key for each of the new entity.

Example 1 (Convert to 3NF)

Old Scheme → {MovieName, ScreeningCity, Genre}

New Scheme → {MovieName, ScreeningCity}

New Scheme → {MovieName, Genre}

Movie	Genre
Hard Code	Comedy
Bill Durham	Drama
The Code Warrier	Horror

Movie	ScreeningCity
Hard Code	Los Angles
Hard Code	New York
Bill Durham	Santa Cruz
Bill Durham	Durham
The Code Warrier	New York

4NF - Decomposition

Example 2 (Convert to 4NF)

Old Scheme → {Manager, Child, Employee}

New Scheme → {Manager, Child}

New Scheme → {Manager, Employee}

Manager	Child
Jim	Beth
Mary	Bob

Manager	Employee
Jim	Alice
Mary	Jane
Mary	Adam

Example 3 (Convert to 4NF)

Old Scheme → {Employee, Skill, ForeignLanguage}

New Scheme → {Employee, Skill}

New Scheme → {Employee, ForeignLanguage}

Employee	Skill
1234	Cooking
1453	Carpentry
1453	Cooking
2345	Cooking

Employee	Language
1234	French
1234	German
1453	Spanish
2345	Spanish

Fifth Normal Form (5NF)

- Fifth normal form is satisfied when all tables are broken into as many tables as possible in order to avoid redundancy. Once it is in fifth normal form it cannot be broken into smaller relations without changing the facts or the meaning.

Domain Key Normal Form (DKNF)

- The relation is in DKNF when there can be no insertion or deletion anomalies in the database.

Lecture-23

Topic- Basic concepts of Data Models

Objective : To acquire Basic concepts of Data Models

Lecture-23

Data Models: Data models are a collection of conceptual tools for describing data, data relationships, data semantics and data constraints.

There are three different groups :

(i) Object-Based Logical Models.

(ii) Record-Based Logical Models.

(iii) Physical Data Models.

Lecture-23

(i) Object-Based Logical Models : Describe data at the conceptual and view levels. Provide fairly flexible structuring capabilities. Allow one to specify data constraints explicitly. Over 30 such models, including :

a) Entity-Relationship Model

b) Object-Oriented Model

c) Binary Model

d) Semantic Data Model

e) Info Logical Model

f) Functional Data Model

Lecture-23

Record-Based Logical Models :

- Also describe data at the conceptual and view levels.
- Unlike object-oriented models, they are used to
  i) Specify overall logical structure of the database, and
  ii)Provide a higher-level description of the implementation.
- Named so because the database is structured in fixed-format records of several types.
- Each record type defines a fixed number of fields, or attributes.
- Each field is usually of a fixed length (this simplifies the implementation).

Lecture-23

- Record-based models do not include a mechanism for direct representation of code in the database.
- Separate languages associated with the model are used to express database queries and updates.

Lecture-23

The three most widely-accepted models are the

1. relational,
2. network,
3. and hierarchical.

Lecture-23

(iii) Physical Data Models :

a) Are used to describe data at the lowest level.

b) Very few models, e.g.

- Unifying model
- Frame memory

Lecture-23

Class –Quiz

A data model is : [ ]

A) Used to describe structure of a database

B) Set of basic operations on the database

C) Both

D)none of these

Lecture-23(Summary)

- A model is a representation of real world events and objects and their association.
- A **data model** is a collection of concepts for describing data.
- A **schema** is a description of a particular collection of data, using the a given data model.
- Three types of data models are :
  1. **Conceptual Data Model :** Description of application that users can understand (e.g. ER model entities and relationships).
  2. **Record based Data Model:** It is collection of table definitions describing the data. (e.g. DBMS specification, Relational model).
  3. **Physical Data Model :** Description of how the logical data model is represented in storage (generated automatically, but can be modified).

Thank You