Fourth and fifth Normal Forms
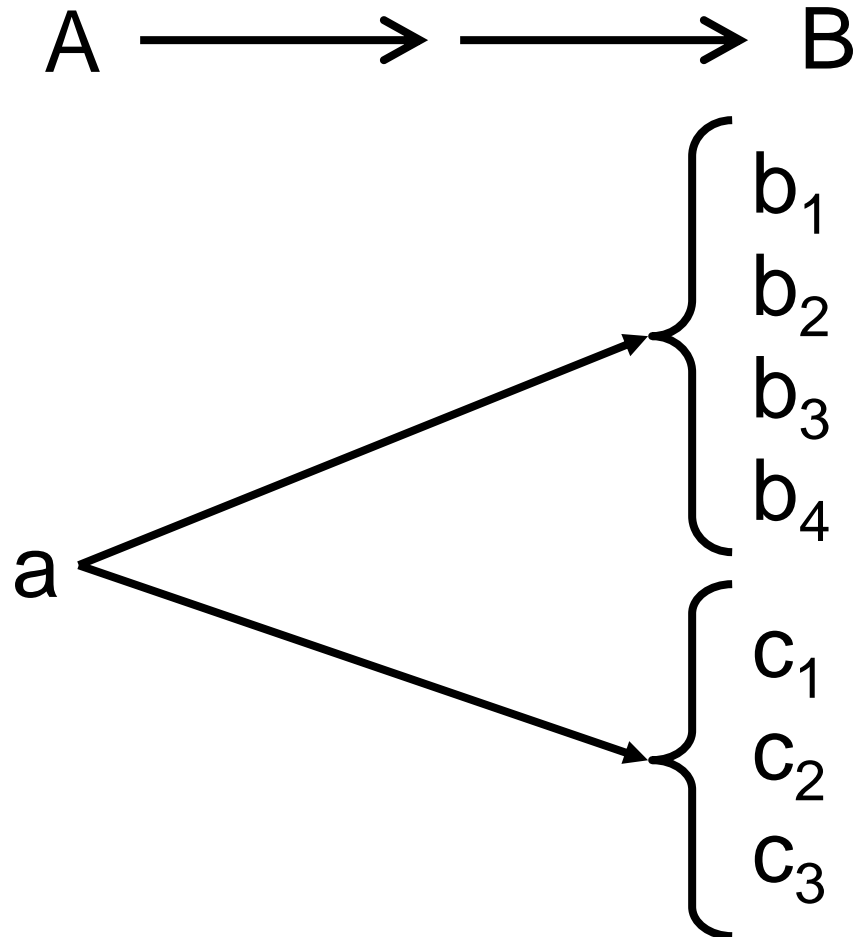
Dr. Seema Gupta Bhol

Multivalue dependency

• Multivalue dependency occurs when the presence of one or more rows in a table implies the presence of one or more other rows in that same table.

• Two attributes in a table are independent of one another, but both depend on a third attribute.

• A **multivalued dependency** always requires at least three attributes because it consists of at least two attributes that are dependent on a third.

• For a dependency A -> B, if for a single value of A, multiple values of B exist, then the table may have a multi-valued dependency.

• The table should have at least 3 attributes and B and C should be independent for A ->> B multivalued dependency.

B is multidependent on A

$$A \longrightarrow \longrightarrow B$$

$$a \Big\langle \begin{array}{l} b_1 \\ b_2 \\ b_3 \\ b_4 \end{array}$$

$$\begin{array}{l} c_1 \\ c_2 \\ c_3 \end{array}$$

4th Normal Form

- No multivalued dependencies

  A multivalued dependency of column B on column A occurs when a table has a key with three or more attributes, (A, B, C) and

  - ☐ *each value of A is associated with a collection of values of B*

  - ☐ *this collection of values is independent of C*

Example

Person	Mobile	Food_Likes
Mahesh	9893/9424	Burger/Pizza
Ramesh	9191	Pizza

Person->->mobile , i.e. "person multi determines mobile"
Person ->-> food_likes i.e. "person multi determines food_likes"

Note that a functional dependency is a special case of multivalued dependency. In a functional dependency X -> Y, every x determines exactly one y, never more than one.

A relation R is in 4NF if and only if the following conditions are satisfied:

1. It should be in the Boyce-Codd Normal Form (BCNF).
2. The table should not have any Multi-valued Dependency.

A table with a multivalued dependency violates the normalization standard of the Fourth Normal Form (4NF) because it creates unnecessary redundancies and can contribute to inconsistent data.

To bring this up to 4NF, it is necessary to break this information into two tables.

SID	SNAME	CID	CNAME
S1	A	C1	C
S1	A	C2	D
S2	B	C1	C
S2	B	C2	D

Multivalued dependencies (MVD) are:
SID->->CID; SID->->CNAME; SNAME->->CNAME
Decomposed into two relations R1 and R2

SID	SNAME
S1	A
S2	B

CID	CNAME
C1	C
C2	D

4th Normal Form

- No multivalued dependencies

  A multivalued dependency of column B on column A occurs when a table has a key with three or more attributes, (A, B, C) and

  - *each value of A is associated with a collection of values of B*

  - *this collection of values is independent of C*

Example

Restaurant	Pizza	DeliveryArea
Pizza Milano	Thin crust	SW Houston
Pizza Milano	Thick crust	SW Houston
Pizza Firenze	Thin crust	NW Houston
Pizza Firenze	Thick crust	NW Houston
Pizza Milano	Thin crust	NW Houston
Pizza Milano	Thick crust	NW Houston

Discussion

- The table has no non-key attributes
  - ☐ Key is { Restaurant, Pizza, DeliveryArea}
- Two non-trivial multivalued dependencies
  - ☐ Restaurant ->-> Pizza
  - ☐ Restaurant ->->DeliveryArea

  since each restaurant delivers the same pizzas to all its delivery areas

4NF Solution

Restaurant	DeliveryArea
Pizza Milano	SW Houston
Pizza Firenze	NW Houston
Pizza Milano	NW Houston

- Two separate tables

Restaurant	Pizza
Pizza Milano	Thin crust
Pizza Milano	Thick crust
Pizza Firenze	Thin crust
Pizza Firenze	Thick crust

Example

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

4NF - Decomposition

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

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

Join dependency

- A table T is subject to a join dependency if it can always be recreated by joining multiple tables each having a subset of the attributes of T

- The join dependency is said to be trivial if one of the tables in the join has all the attributes of the table T

Join dependency (JD)

- If the join of R1 and R2 is equal to relation R, then we can say that a join dependency (JD) exists.

- Where R1 and R2 are the decompositions R1(A, B, C) and R2(C, D) of a given relations R (A, B, C, D).

- Alternatively, R1 and R2 are a lossless decomposition of R.

- A JD ⋈ {R1, R2,..., Rn} is said to hold over a relation R if R1, R2,....., Rn is a lossless-join decomposition.

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.

- A table T is said to be 5NF iff

  - Every non-trivial join dependency in it is implied by its candidate keys

An example

Store	Brand	Product
Circuit City	Apple	Tablets
Circuit City	Apple	Phones
Circuit City	Toshiba	Laptops
CompUSA	Apple	Laptops

- Note that Circuit City sells Apple tablets and phones but only Toshiba laptops

Decomposition

Store	Product
Circuit City	Tablets
Circuit City	Phones
Circuit City	Laptops
CompUSA	Laptops

Product	Brand
Tablets	Apple
Phones	Apple
Laptops	Apple
Laptops	Toshiba

- Let see what happens when we do a natural join

The result of the join

Store	Product	Brand
Circuit City	Tablets	Apple
Circuit City	Phones	Apple
Circuit City	Laptops	Apple
Circuit City	Laptops	Toshiba
CompUSA	Laptops	Apple
CompUSA	Laptops	Toshiba

- Introduces two spurious tuples

Examples of Join Dependency

Stu_Name	Stu_Skills	Stu_Job (Assigned Work)
Tag	Marketing	GK001
Barry	PR	GK002
Paulo	Graphic Designing	GK003

We can decompose the table given above into these three tables. And thus, it is not in the Fifth Normal Form.

- Student_Skills
- Student_Job
- Job_Skills

Stu_Name	Stu_Skills
Tag	Marketing
Barry	PR
Paulo	Graphic Designing

Stu_Name	Stu_Job
Tag	GK001
Barry	GK002
Paulo	GK002

Stu_Skills	Stu_Job
Marketing	GK001
PR	GK002
Graphic Designing	GK003