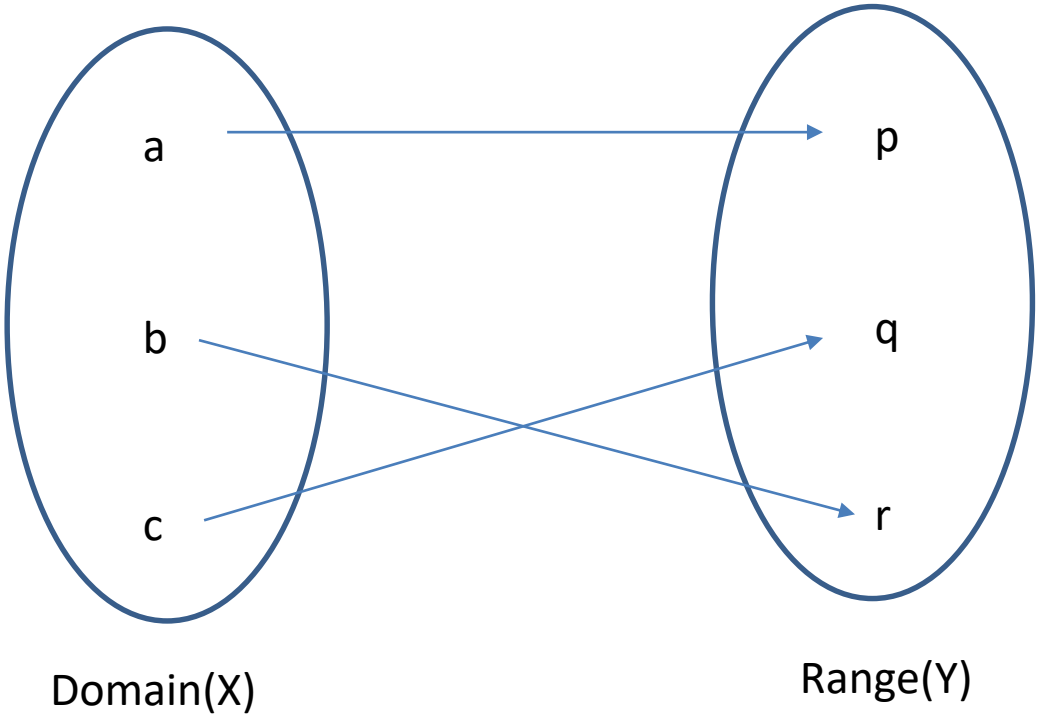
Multi Valued Dependencies and Fourth Normal Form

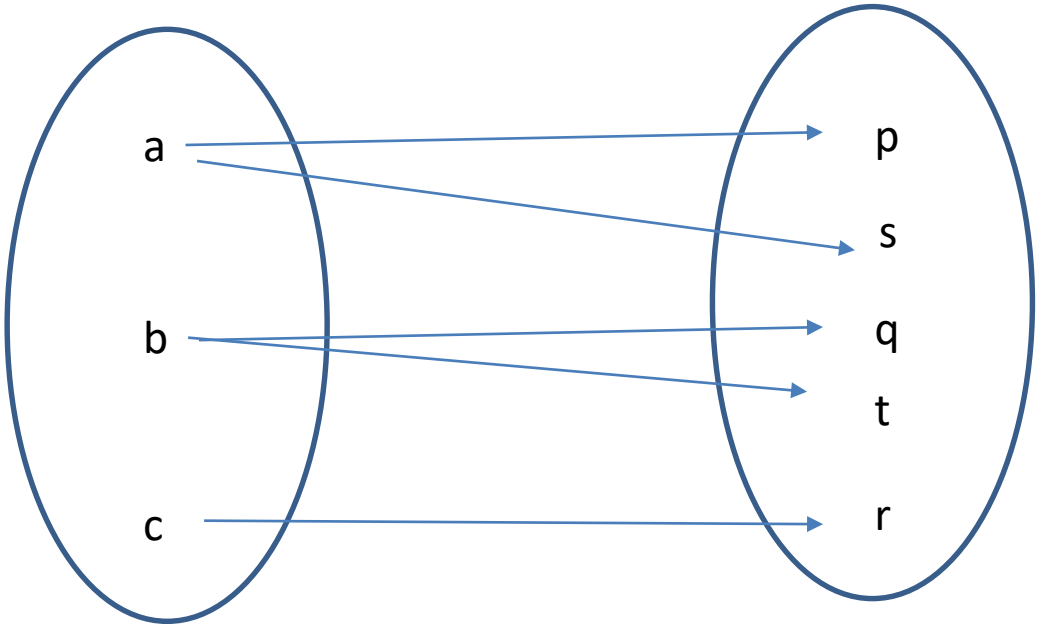
COURSE : 23CAT- Database Management System
UNIT III : Database Design
CLASS : I Semester / I MCA

Recall Functional Dependency



- **Domain** is a set of all possible values of X
- **Range** is a set of all possible values of Y

Multi Valued Dependencies (MVD)



Multivalued dependency

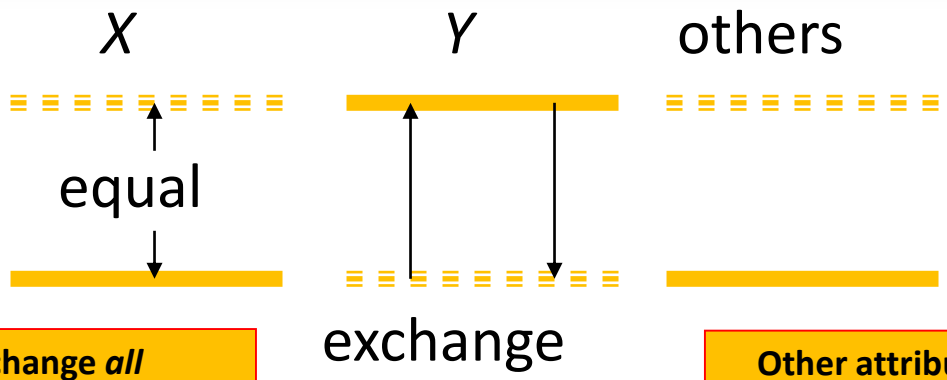
- Multivalued dependency would occur whenever two separate attributes in a given table happen to be independent of each other.
- And yet, both of these depend on another third attribute.
- The multivalued dependency contains at least two of the attributes dependent on the third attribute.
- It is represented by x –> –> y

for some single value of the attribute ‘x’, multiple values of attribute ‘y’ can exist.

Multi Valued Dependencies (MVD)

- Multi-valued dependencies (MVDs) express a condition among tuples of a relation that exists when the table (relation) is trying to represent
  - more than one many-many relationship.
  - then certain columns (attributes) become independent of one another
  - and their values must appear in all combinations.

- A multi-valued dependency (MVD) X →→ Y :
  - if an assertion that if two rows of a table agree on all the attributes of X,
  - then their components may be swapped in the set of attributes Y,
  - and the result will be two tuples that are also in the relation



We must exchange *all* components of Ys (not just some)

Other attributes get copied remain the same

If there is an FD X → Y, then swapping Ys components doesn't change anything.

Every FD is an MVD.

4NF Definition

- A relation R is in 4NF if whenever $X \rightarrow\rightarrow Y$ is a nontrivial MVD, then X is a superkey.
  - Nontrivial means that:
    1. Y is not a subset of X

       (swapping components does not change tuples)

    2. X and Y are not, together, all the attributes.

       (swapping components yields the same tuples)
- Note that the definition of "superkey" still depends on FDs only.

1.It should be in BCNF.
2.There should be no Multi-valued Dependency

4NF Definition

- A relation R is in 4NF if whenever $X \rightarrow\rightarrow Y$ is a nontrivial MVD, then X is a superkey.
  - Nontrivial means that:
    1. Y is not a subset of X

       (swapping components does not change tuples)

    2. X and Y are not, together, all the attributes.

       (swapping components yields the same tuples)
- Note that the definition of "superkey" still depends on FDs only.

1.It should be in BCNF.
2.There should be no Multi-valued Dependency

- Remember that
  - Every FD X → Y is also an MVD, X →→ Y.
- Thus, if R is in 4NF, it is certainly in BCNF.
  - Because any BCNF violation is a 4NF violation.
- However, R could be in BCNF and not 4NF, because MVDs are “invisible” to BCNF.

- Customers(<u>name</u>, addr, <u>phones</u>, <u>sodasLiked</u>)
  - FD: name → addr
  - MVDs: name →→ phones, name →→ sodasLiked

- Key is {<u>name</u>, <u>phones</u>, <u>sodasLiked</u>}.
- All dependencies violate 4NF because name is not a superkey.

- Decompose using name → addr:
  1. Customers1(<u>name</u>, addr)
     - In 4NF, only dependency is name → addr.
  2. Customers2(<u>name</u>, <u>phones</u>, <u>sodasLiked</u>)
     - Not in 4NF because MVDs
       - name →→ phones and name →→ sodasLiked apply.
     - No FD's, so all three attributes form the key.

- Recall
  - Customers2(name, phones, sodasLiked)
- Either MVD
  - name →→ phones
  - name →→ sodasLiked
- tells us to decompose to:
  - Customers3(name, phones)
  - Customers4(name, sodasLiked)

- 4th Normal Form Consists of:
  - Customers1(name, addr)
    - Contact
  - Customers3(name, phones)
    - Phone
  - Customers4(name, sodasLiked)
    - Likes
- The following FD and MVDs are satisfied:
- FD: name → addr
- MVDs: name →→ phones
  name →→ sodasLiked

Boyce – Codd Normal Form (BCNF)

- Developed in 1974 by Raymond F.Boyce and Edgar F.Codd
- BCNF is a stricter form of 3NF
- It applies to tables with more than one candidate key
- A relation is in BCNF if every determinant in the table is a candidate key. Ie LHS is super key
- If a table contains single candidate key, the 3NF and BCNF are equvalent

Reasoning About FDs

Reasoning About FDs

Architectural Design Challenges