Multivalued Dependencies

Fourth Normal Form

A New Form of Redundancy

- Multivalued dependencies (MVD's) express a condition among tuples of a relation that exists when the relation is trying to represent more than one many-many relationship.
- Then certain attributes become independent of one another, and their values must appear in all combinations.

Example

Drinkers(name, addr, phones, beersLiked)

- A drinker's phones are independent of the beers they like.
- Thus, each of a drinker's phones appears with each of the beers they like in all combinations.
- This repetition is unlike redundancy due to FD's, of which name->addr is the only one.

Tuples Implied by Independence

If we have tuples:

name	addr	phones	beersLiked
sue	a	p1	b1
sue	a	p2	b2
sue	a	p2	b1
sue	a	p1	b2

Then these tuples must also be in the relation.

Definition of MVD

- A *multivalued dependency* (MVD) *X* ->-> *Y* is an assertion that if two tuples of a relation agree on all the attributes of *X*, then their components in the set of attributes *Y* may be swapped, and the result will be two tuples that are also in the relation.

Example

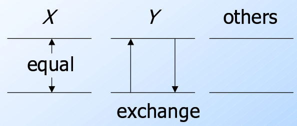
- The name-addr-phones-beersLiked example illustrated the MVD

  name->->phones

  and the MVD

  name ->-> beersLiked.

Picture of MVD *X* ->-> *Y*



7

MVD Rules

- Every FD is an MVD.
  - If *X* -> *Y*, then swapping *Y* 's between two tuples that agree on *X* doesn't change the tuples.
  - Therefore, the "new" tuples are surely in the relation, and we know *X* ->-> *Y*.
- *Complementation* : If *X* ->-> *Y*, and *Z* is all the other attributes, then *X* ->-> *Z*.

8

Splitting Doesn't Hold

- Like FD's, we cannot generally split the left side of an MVD.
- But unlike FD's, we cannot split the right side either --- sometimes you have to leave several attributes on the right side.

9

Example

- Consider a drinkers relation:

Drinkers(name, areaCode, phone, beersLiked, manf)

- A drinker can have several phones, with the number divided between areaCode and phone (last 7 digits).
- A drinker can like several beers, each with its own manufacturer.

10

Example, Continued

- Since the areaCode-phone combinations for a drinker are independent of the beersLiked-manf combinations, we expect that the following MVD's hold:

  name ->-> areaCode phone

  name ->-> beersLiked manf

11

Example Data

Here is possible data satisfying these MVD's:

name	areaCode	phone	beersLiked	manf
Sue	650	555-1111	Bud	A.B.
Sue	650	555-1111	WickedAle	Pete's
Sue	415	555-9999	Bud	A.B.
Sue	415	555-9999	WickedAle	Pete's

But we cannot swap area codes or phones my themselves. That is, neither name ->-> areaCode nor name ->-> phone holds for this relation.

12

Fourth Normal Form

- The redundancy that comes from MVD's is not removable by putting the database schema in BCNF.
- There is a stronger normal form, called 4NF, that (intuitively) treats MVD's as FD's when it comes to decomposition, but not when determining keys of the relation.

13

4NF Definition

- A relation *R* is in 4NF if whenever *X* ->-> *Y* is a nontrivial MVD, then *X* is a superkey.
  - "Nontrivial means that:
    1. *Y* is not a subset of *X*, and
    2. *X* and *Y* are not, together, all the attributes.
  - Note that the definition of "superkey" still depends on FD's only.

14

BCNF Versus 4NF

- 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.
- But *R* could be in BCNF and not 4NF, because MVD's are "invisible" to BCNF.

15

Decomposition and 4NF

- If *X* ->-> *Y* is a 4NF violation for relation *R*, we can decompose *R* using the same technique as for BCNF.
  1. *XY* is one of the decomposed relations.
  2. All but *Y* – *X* is the other.

16

Example

Drinkers(<u>name</u>, addr, <u>phones</u>, <u>beersLiked</u>)

FD: name -> addr
MVD's: name ->-> phones
name ->-> beersLiked

- Key is {name, phones, beersLiked}.
- All dependencies violate 4NF.

17

Example, Continued

- Decompose using name -> addr:
1. Drinkers1(<u>name</u>, addr)
   - In 4NF, only dependency is name -> addr.
2. Drinkers2(<u>name</u>, <u>phones</u>, <u>beersLiked</u>)
   - Not in 4NF. MVD's name ->-> phones and name ->-> beersLiked apply. No FD's, so all three attributes form the key.

18

Example: Decompose Drinkers2

- Either MVD name ->-> phones or name ->-> beersLiked tells us to decompose to:
  - Drinkers3(name, phones)
  - Drinkers4(name, beersLiked)

19