

Multivalued Dependencies

Fourth Normal Form

Definition of MVD

- ◆ A *multivalued dependency* (MVD) on R , $X \twoheadrightarrow Y$, says that if two tuples of R agree on all the attributes of X , then their components in Y may be swapped, and the result will be two tuples that are also in the relation.
- ◆ i.e., for each value of X , the values of Y are independent of the values of $R-X-Y$.

Example

Drinkers(name, addr, phones, beersLiked)

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

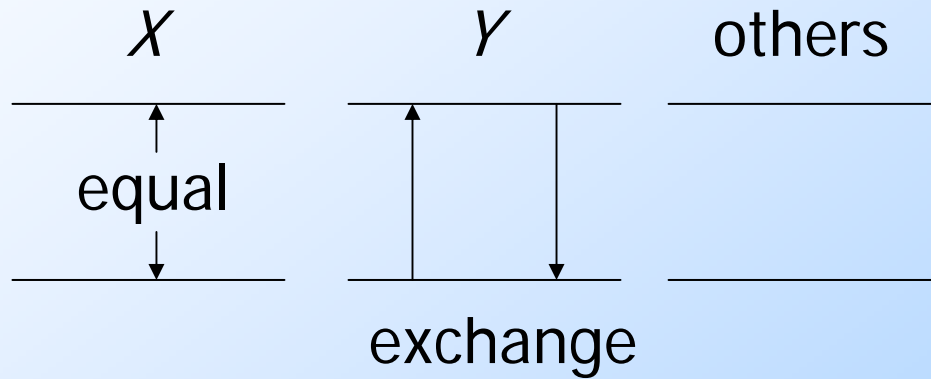
Tuples Implied by $\text{name} \twoheadrightarrow \text{phones}$

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.

Picture of MVD $X \rightarrow - \rightarrow Y$



MVD Rules

- ◆ Every FD is an MVD (*promotion*).
 - ◆ If $X \rightarrow Y$, 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 $X \twoheadrightarrow Y$.
- ◆ *Complementation* : If $X \twoheadrightarrow Y$, and Z is all the other attributes, then $X \twoheadrightarrow Z$.

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.

Example

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.

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 \multimap areaCode phone

name \multimap beersLiked manf

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 by themselves.
That is, neither $\text{name} \twoheadrightarrow \text{areaCode}$ nor $\text{name} \twoheadrightarrow \text{phone}$ holds for this relation.

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.

4NF Definition

- ◆ A relation R is in *4NF* if: whenever $X \twoheadrightarrow Y$ is a nontrivial MVD, then X is a superkey.
- ◆ *Nontrivial MVD* 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.

BCNF Versus 4NF

- ◆ Remember that every FD $X \rightarrow Y$ is also an MVD, $X \twoheadrightarrow Y$.
- ◆ Thus, if R is in 4NF, it is certainly in BCNF.
 - ◆ Because any BCNF violation is a 4NF violation (after conversion to an MVD).
- ◆ But R could be in BCNF and not 4NF, because MVD's are "invisible" to BCNF.

Decomposition and 4NF

- ◆ If $X \twoheadrightarrow Y$ 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.

Example

Drinkers(name, addr, phones, beersLiked)

FD: name -> addr

MVD's: name ->-> phones

 name ->-> beersLiked

◆ Key is {name, phones, beersLiked}.

◆ All dependencies violate 4NF.

Example, Continued

◆ Decompose using $\text{name} \rightarrow \text{addr}$:

1. Drinkers1(name, addr)

◆ In 4NF; only dependency is $\text{name} \rightarrow \text{addr}$.

2. Drinkers2(name, phones, beersLiked)

◆ Not in 4NF. MVD's $\text{name} \twoheadrightarrow \text{phones}$ and $\text{name} \twoheadrightarrow \text{beersLiked}$ apply. No FD's, so all three attributes form the key.

Example: Decompose Drinkers2

◆ Either MVD $\text{name} \twoheadrightarrow \text{phones}$ or $\text{name} \twoheadrightarrow \text{beersLiked}$ tells us to decompose to:

- ◆ $\text{Drinkers3}(\underline{\text{name}}, \underline{\text{phones}})$
- ◆ $\text{Drinkers4}(\underline{\text{name}}, \underline{\text{beersLiked}})$