CS422 Principles of Database Systems
Multivalued Dependency

Chengyu Sun
California State University, Los Angeles

Motivational Example

drinker	address	beerLiked
Sue	123 Main St.	Bud
Sue	321 State St.	Pete's Ale

drinkers

- FD?? Keys??
- 3NF?? BCNF??
- Is this a good design??

A New Form of Redundancy

drinker	address	beerLiked
Sue	123 Main St.	Bud
Sue	321 State St.	Pete's Ale

Sue	321 State St.	Bud
Sue	123 Main St.	Pete's Ale

- Any combination of `address` and `beerLiked` for `Sue` is a valid tuple

Multivalued Dependency (MVD)

- A Multivalued Dependency (MVD) **A** **B** is an assertion that if two tuples of a relation agree on all the attributes of **A**, then their components in the set of attributes **B** may be swapped, and the result will be two tuples that are also in the relation.
- In the `drinkers` example:
  - **A**?? **B**?? **C**=**R**-**AB**??
  - ?? ??

A Couple of Observations about MVD

- MVD characterizes the case where one relation tries to represents more than one *many-to-many* relationships.
- MVD vs. FD (why it's called *multivalued* dependency)

Trivial MVD

MVD: **A** **B**

- MVD is trivial if
  - **B** ⊆ **A**, *or*
  - **A** ∪ **B** = **R**

MVD Rules

MVD Complementation
If **A** **B**, then **A** **R** - **AB**

MVD Transitivity
If **A** **B** and **B** **C**, then **A** **C**-**B**

MVD Augmentation
If **A** **B**, then **AC** **BD** for any **D** $\subseteq$ **C**

MVD + FD Rules

Replication
If **A** **B**, then **A** **B**

Coalescence
If **A** **B**, **C** **D**, **C** $\cap$ **B** = Ø , **D** $\subseteq$ **B**
then **A** **D**

Proof by Chase

- Given a set of FDs and MVDs *D*, does another dependency *d* (FD or MVD) follow from D?
- Procedure
  - Start with the hypotheses of d, and treat them as "seed" tuples in a relation
  - Apply the given dependencies in D repeatedly
    - If we apply an FD, we infer equality of two symbols
    - If we apply an MVD, we infer more tuples
- If we infer the conclusion of d, we have a proof; otherwise we have a counter-example

From Jun Yang's lecture notes at http://www.cs.duke.edu/~junyang

Proof by Chase Example

- In R(A, B, C, D), does A B and B C implies A C?

Have

A	B	C	D
a	b_1	c_1	d_1
a	b_2	c_2	d_2

Need

A	B	C	D
a	b_1	c_2	d_1
a	b_2	c_1	d_2

From Jun Yang's lecture notes at http://www.cs.duke.edu/~junyang

Proof by Chase Example

- In R(A, B, C, D), does A B and B C implies A C?

Have

A	B	C	D
a	b_1	c_1	d_1
a	b_2	c_2	d_2

Need

A	B	C	D
a	b_1	c_2	d_1
a	b_2	c_1	d_2

A B

a	b_2	c_1	d_1
a	b_1	c_2	d_2

??

From Jun Yang's lecture notes at http://www.cs.duke.edu/~junyang

Union and Decomposition

- Union: if **A** **B** and **A** **C**, then **A** **BC**
  - Proof??
- Decomposition rule *no longer holds*
  - Counter-example??

Fourth Normal Form (4NF)

- A relation R is in 4NF if for every *nontrivial* MVD **A** **B**, **A** is a super key.

Decompose into 4NF

- Find a 4NF violation **A** **B**
- Decompose R into:
  - $\mathbf{R_1} = \mathbf{A} \cup \mathbf{B}$
  - $\mathbf{R_2} = (\mathbf{R} - \mathbf{AB}) \cup \mathbf{A}$
- Repeat until all relations are in 4NF

4NF Decomposition Example

- Drinkers(name, addr, beerLiked, favBeer)
  - FD?? Key??
  - MVD??

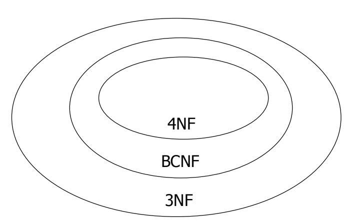
4NF Decomposition vs. BCNF Decomposition

- In 4NF decomposition we do not compute $\mathbf{A}^+$
  - $\mathbf{A}^+$ does not make sense for MVD
  - **A** (**R**-**A**) and **A** **A**
- Inferring MVDs for the projections are very difficult
  - However, we can usually get by using the rules of *transitivity*, *complementation*, and *intersection*.

Exercise: Prove the Intersection Rule

If **A** **B**, and **A** **C**, then **A** $\mathbf{B} \cap \mathbf{C}$

4NF vs. BCNF



- Why??