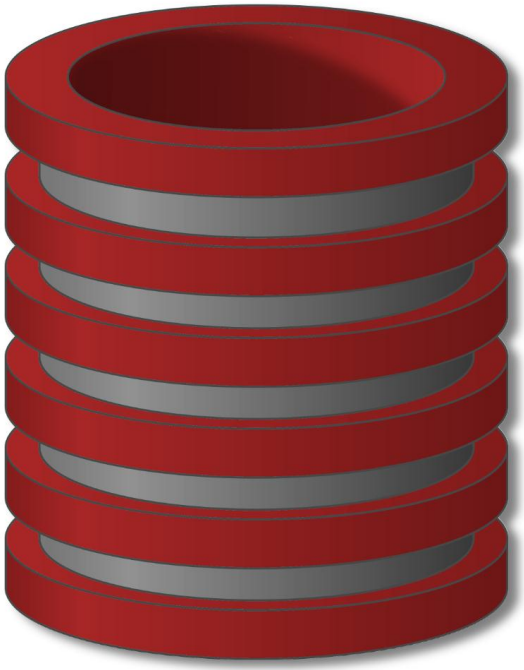


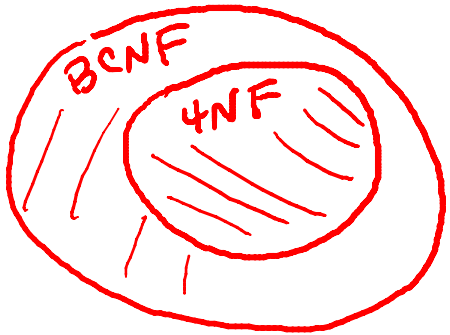
Relational Design Theory

Multivalued Dependencies & 4th Normal Form

Jennifer Widom

Relational design by decomposition

- "Mega" relations + properties of the data
- System decomposes based on properties
- Final set of relations satisfies normal form
  - No anomalies, no lost information
- Functional dependencies ⇒ Boyce-Codd Normal Form
- Multivalued dependences ⇒ Fourth Normal Form



Example: College application info.

Apply(SSN, cName, hobby)

FDs? No.

Keys? All attrs.

BCNF? Yes.

Good design? No.

5 colleges, 6 hobbies → 30 tuples.

Multivalued Dependency

- Based on knowledge of real world
- All instances of relation must adhere

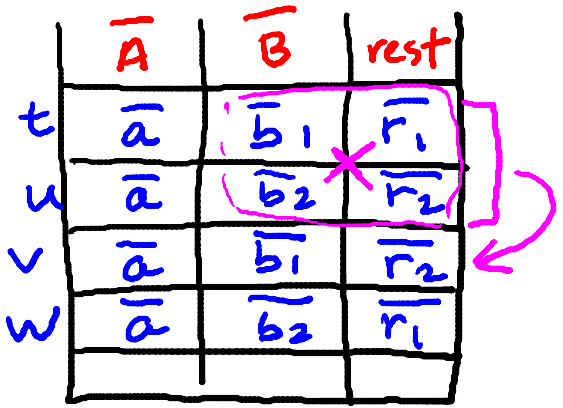
$R \quad \bar{A} \twoheadrightarrow \bar{B} \quad A_1, \ldots, A_n \quad B_1, \ldots, B_n$

$\forall t, u \in R : t[\bar{A}] = u[\bar{A}]$ then

$\exists v \in R : v[\bar{A}] = t[\bar{A}]$ and

$v[\bar{B}] = t[\bar{B}]$ and

$v[\text{rest}] = u[\text{rest}]$



tuple-generating dependencies

Apply(SSN, cName, hobby)

SSN →→ cName SSN →→ hobby

	SSN	cName	hobby
t	123	Stanford	trumpet
u	123	Berkeley	tennis
v	123	Stanford	tennis
w	123	Berkeley	trumpet
	⋮	⋮	⋮

Modified example

Apply(SSN, cName, hobby) ★

Reveal hobbies to colleges selectively ★★

MVDs? None

Good design? Yes.

Expanded example

Apply(SSN, cName, date, major, hobby)

Reveal hobbies to colleges selectively ✓
Apply once to each college one day
May apply to multiple majors ✓

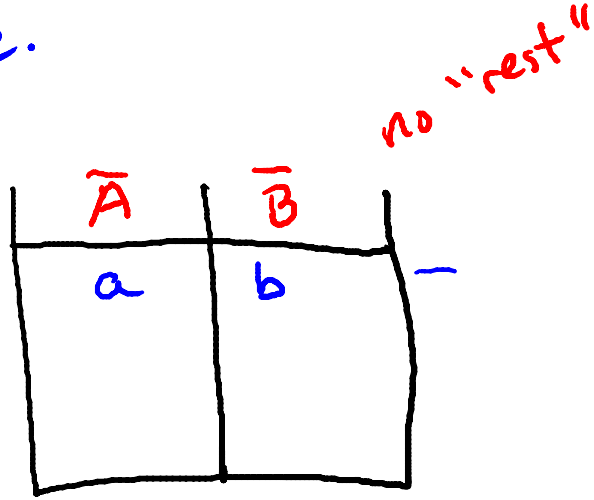
SSN, cName → date
SSN, cName, date →→ major "rest" hobby

Trivial Multivalued Dependency

$\bar{A} \twoheadrightarrow \bar{B}$ $\bar{B} \subseteq \bar{A}$ or $\bar{A} \cup \bar{B}$ = all attributes

Nontrivial MVD

otherwise.

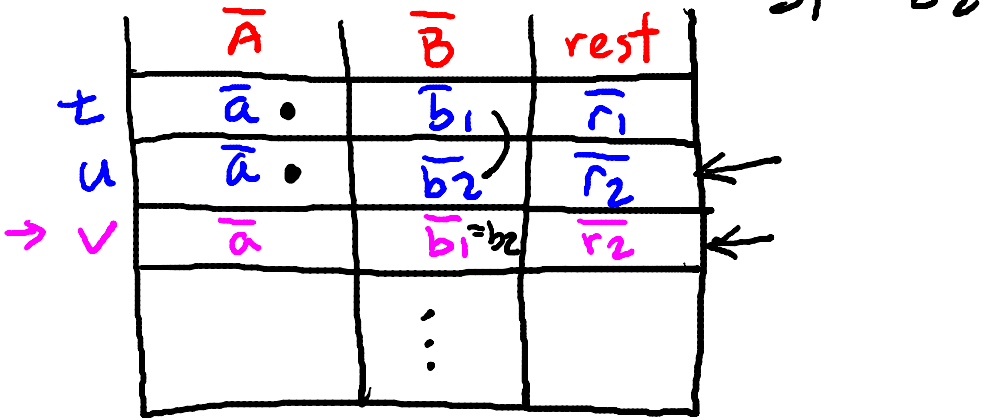


Rules for Multivalued Dependencies

FD-is-an-MVD rule

$\bar{A} \rightarrow \bar{B}$ then $\bar{A} \twoheadrightarrow \bar{B}$

$\bar{b_1} = \bar{b_2}$



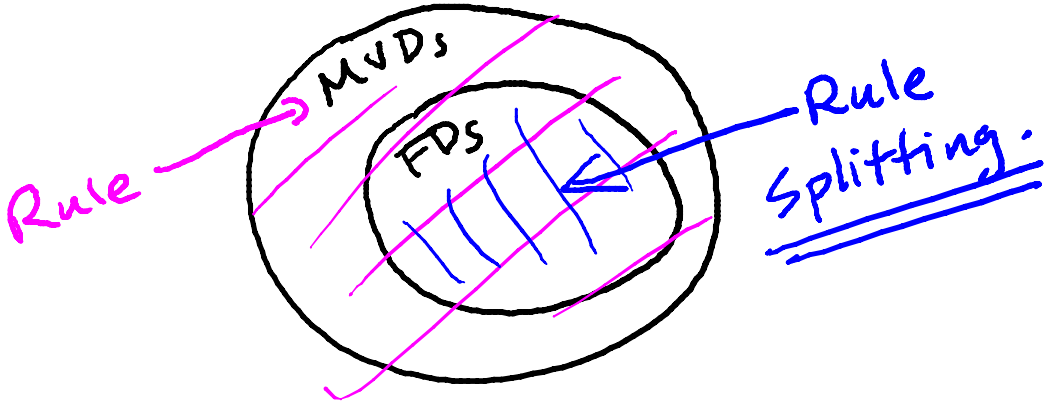
Rules for Multivalued Dependencies

Intersection rule

$\bar{A} \twoheadrightarrow \bar{B}$ $\bar{A} \twoheadrightarrow \bar{C}$ then $\bar{A} \twoheadrightarrow \bar{B} \cap \bar{C}$

Transitive rule

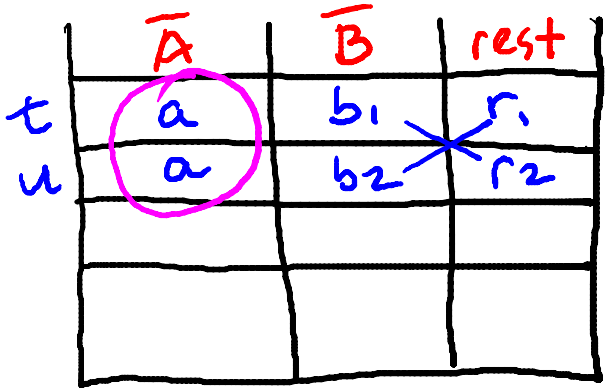
$\bar{A} \twoheadrightarrow \bar{B}$ $\bar{B} \twoheadrightarrow \bar{C}$ then $\bar{A} \twoheadrightarrow \bar{C} - \bar{B}$



Fourth Normal Form

Relation R with MVDs is in 4NF if:

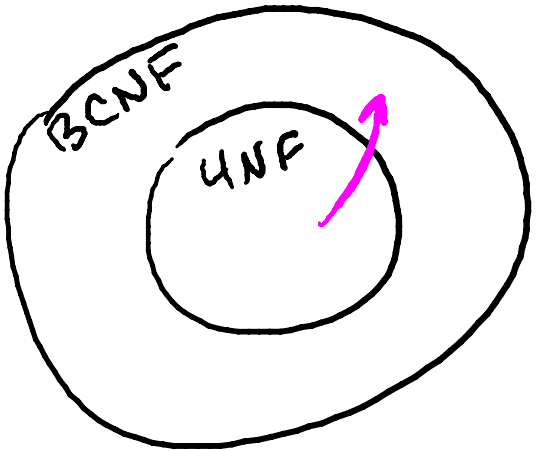
For each nontrivial $A \twoheadrightarrow B$, A is a key



Fourth Normal Form ⇒ BCNF

Relation R with MVDs is in 4NF if:

For each nontrivial $A \twoheadrightarrow B$, A is a key



A → B A is a key

↳ A ↠ B

4NF decomposition algorithm

Input: relation R + FDs for R + MVDs for R

Output: decomposition of R into 4NF relations with "lossless join"

Compute keys for R

Repeat until all relations are in 4NF:

  Pick any R' with nontrivial $A \twoheadrightarrow B$ that violates 4NF

  Decompose R' into $R_1(A, B)$ and $R_2(A, rest)$

  Compute FDs and MVDs for R_1 and R_2

  Compute keys for R_1 and R_2

4NF Decomposition Example #1

Apply(SSN, cName, hobby)

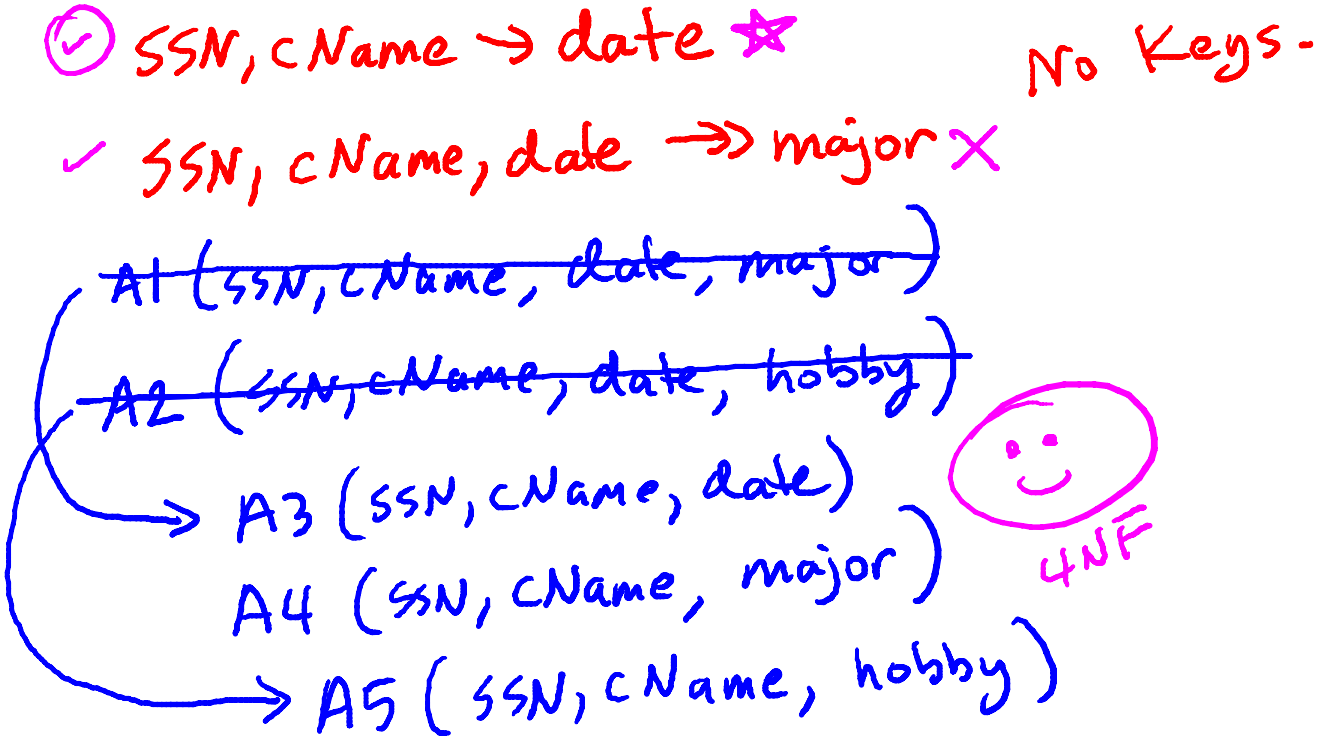
SSN ↠ cName No keys.

A1(SSN, cName) No FDs, No MVDs

A2(SSN, hobby)

4NF Decomposition Example #2

Apply(SSN, cName, date, major, hobby)



Relational design

- Functional dependencies & Boyce-Codd Normal Form

  $R(A, B, C)$ $A \rightarrow B$

- Multivalued dependences & Fourth Normal Form

  $R(A, B, C, D)$ $A \twoheadrightarrow B$