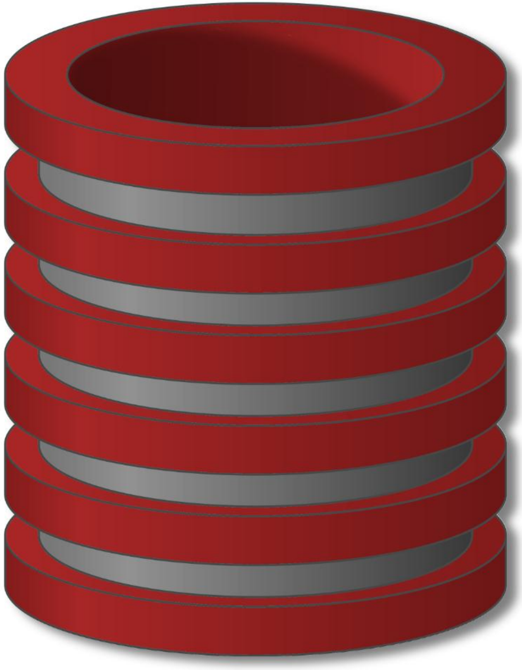


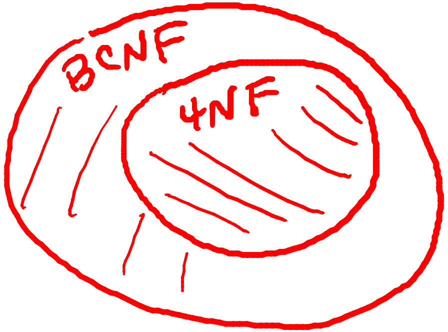
Relational Design Theory

Multivalued Dependencies & 4^{th} 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

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

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

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

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

$v[rest] = u[rest]$

	$\bar{A}$	$\bar{B}$	rest
t	$\bar{a}$	$\bar{b}_1$	$\bar{r}_1$
u	$\bar{a}$	$\bar{b}_2$	$\bar{r}_2$
v	$\bar{a}$	$\bar{b}_1$	$\bar{r}_2$
w	$\bar{a}$	$\bar{b}_2$	$\bar{r}_1$

tuple-generating dependencies

Apply(SSN, cName, hobby)

SSN $\twoheadrightarrow$ cName   SSN $\twoheadrightarrow$ 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

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

Nontrivial MVD

otherwise.

no "rest"

$\bar{A}$	$\bar{B}$
a	b

Rules for Multivalued Dependencies

FD-is-an-MVD rule

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

$\bar{b}_1 = \bar{b}_2$

	$\bar{A}$	$\bar{B}$	rest
t	$\bar{a}$	$\bar{b}_1$	$\bar{r}_1$
u	$\bar{a}$	$\bar{b}_2$	$\bar{r}_2$
$\to v$	$\bar{a}$	$\bar{b}_1 = b_2$	$\bar{r}_2$
		⋮	

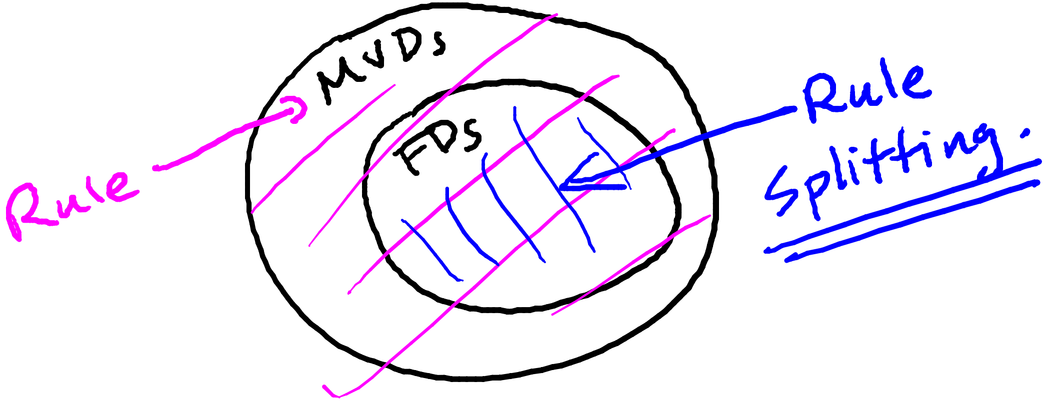
Rules for Multivalued Dependencies

Intersection rule

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

Transitive rule

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



Fourth Normal Form

Relation R with MVDs is in 4NF if:

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

	$\bar{A}$	$\bar{B}$	rest
t	a	b_1	r_1
u	a	b_2	r_2

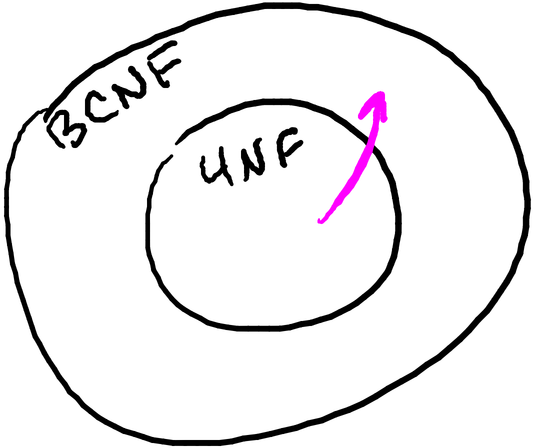
Fourth Normal Form $\Rightarrow$ BCNF

Relation R with MVDs is in 4NF if:

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

$A \rightarrow B$ A is a key

$\hookrightarrow A \twoheadrightarrow B$



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 ✓

BCNF Decomposition Example #1

`Apply(SSN, cName, hobby)`

$SSN \twoheadrightarrow cName$ No keys.

A1(SSN, cName) No FDs No MVDs

A2(SSN, hobby)

BCNF Decomposition Example #2

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

✓ SSN, cName → date ☆

No Keys.

✓ SSN, cName, date →→ major ✗

~~A1 (SSN, cName, date, major)~~

~~A2 (SSN, cName, date, hobby)~~

→ A3 (SSN, cName, date)

A4 (SSN, cName, major)

→ A5 (SSN, cName, hobby)

4NF

Relational design

- Functional dependencies & Boyce-Codd Normal Form

$R(A,B,C)$ $\underline{A} \rightarrow \underline{B}$

- Multivalued dependences & Fourth Normal Form

$R(A,B,C,D)$ $\underline{\bar{A} \twoheadrightarrow \bar{B}}$