Relational Design Theory

Multivalued Dependencies & 4th Normal Form
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 $\Rightarrow$ Boyce-Codd Normal Form
- Multivalued dependences $\Rightarrow$ Fourth Normal Form
Example: College application info.

\text{Apply}(\text{SSN}, \text{cName}, \text{hobby})

FDs? No.

Keys? All attrs.

BCNF? Yes.

Good design? No.

5 colleges, 6 hobbies \rightarrow 30 tuples.
Multivalued Dependency

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

\[
R \overset{\bar{A}}{\rightarrow} \bar{B} \quad A_1, \ldots, A_n \quad B_1, \ldots, B_n
\]

\[
\forall t, u \in R: t[A] = u[A] \text{ then } \exists v \in R: v[A] = t[A] \quad \text{and} \quad v[B] = t[B] \quad \text{and} \quad v[\text{rest}] = u[\text{rest}]
\]

- tuple-generating dependencies
Apply($SSN$, $c\text{Name}$, $hobby$)

$$SSN \rightarrow c\text{Name} \quad SSN \rightarrow hobby$$

<table>
<thead>
<tr>
<th></th>
<th>$SSN$</th>
<th>$c\text{Name}$</th>
<th>$hobby$</th>
</tr>
</thead>
<tbody>
<tr>
<td>t</td>
<td>123</td>
<td>Stanford</td>
<td>trumpet</td>
</tr>
<tr>
<td>u</td>
<td>123</td>
<td>Berkeley</td>
<td>tennis</td>
</tr>
<tr>
<td>v</td>
<td>123</td>
<td>Stanford</td>
<td>tennis</td>
</tr>
<tr>
<td>w</td>
<td>123</td>
<td>Berkeley</td>
<td>trumpet</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
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

\[ \overline{A} \Rightarrow \overline{B} \quad \overline{B} \subseteq \overline{A} \quad \text{or} \quad \overline{A} \cup \overline{B} = \text{all attributes} \]

Nontrivial MVD

otherwise.

\[ \begin{array}{cc}
\overline{A} & \overline{B} \\
\hline
a & b \\
\end{array} \]

no "rest"
Rules for Multivalued Dependencies

**FD-is-an-MVD rule**

\[ \overline{A} \rightarrow \overline{B} \text{ then } \overline{A} \rightarrow \overline{B} \]

\[ b_1 = b_2 \]
Rules for Multivalued Dependencies

Intersection rule
\[ \overline{A} \rightarrow \overline{B} \quad \overline{A} \rightarrow \overline{C} \quad \text{then} \quad \overline{A} \rightarrow \overline{B} \cap \overline{C} \]

Transitive rule
\[ \overline{A} \rightarrow \overline{B} \quad \overline{B} \rightarrow \overline{C} \quad \text{then} \quad \overline{A} \rightarrow \overline{C} - \overline{B} \]
Fourth Normal Form

Relation \( R \) with MVDs is in 4NF if:

For each nontrivial \( A \rightarrow B \), \( A \) is a key

![Diagram showing a relation R with attributes A and B, and a key indicated by a circled 'a'.]
Fourth Normal Form (4NF) \implies BCNF

Relation $R$ with MVDs is in 4NF if:

For each nontrivial $A \rightarrow B$, $A$ is a key.
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 → 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)

\[ SSN \rightarrow cName \]

No keys.

\[ A_1(SSN, cName) \]

No FDs
No MVDs

\[ A_2(SSN, hobby) \]

MVDs & 4NF
BCNF Decomposition Example #2

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

\[ \text{No Keys -} \]

- SSN, cName $\rightarrow$ date *

- SSN, cName, date $\rightarrow$ major *

\[ A_1 (\text{SSN, cName, date, major}) \]

\[ A_2 (\text{SSN, cName, date, hobby}) \]

\[ A_3 (\text{SSN, cName, date}) \]

\[ A_4 (\text{SSN, cName, major}) \]

\[ A_5 (\text{SSN, cName, hobby}) \]
Relational design

- Functional dependencies & Boyce-Codd Normal Form
  \[ R(A, B, C) \quad A \rightarrow B \]
- Multivalued dependences & Fourth Normal Form
  \[ R(A, B, C, D) \quad \bar{A} \rightarrow \bar{B} \]