Outline

- Schema design
  - Introduction
  - Notation
- Functional dependencies
- Armstrong's axioms
- Minimal cover
- Keys
- Questions

Schema Design

Customer:

<table>
<thead>
<tr>
<th>Cust_Id</th>
<th>Faculty</th>
<th>Track</th>
</tr>
</thead>
<tbody>
<tr>
<td>12345</td>
<td>CS</td>
<td>Software</td>
</tr>
<tr>
<td>45678</td>
<td>EE</td>
<td>Hardware</td>
</tr>
<tr>
<td>11111</td>
<td>IE</td>
<td>IS</td>
</tr>
<tr>
<td>22222</td>
<td>IE</td>
<td>Accounting</td>
</tr>
</tbody>
</table>

Ordered:

<table>
<thead>
<tr>
<th>Cust_Id</th>
<th>Book_Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>12345</td>
<td>Database Systems</td>
</tr>
<tr>
<td>45678</td>
<td>Anatomy</td>
</tr>
<tr>
<td>12345</td>
<td>Database And Knowledge</td>
</tr>
<tr>
<td>11111</td>
<td>Anatomy</td>
</tr>
<tr>
<td>22222</td>
<td>Intro. To Economy</td>
</tr>
</tbody>
</table>

Which design is better?

Good design – cont’d

- No duplications of data
- Enables simple updates
- Simple
- Not too many tables

We saw a way to design DB’s
  - ERD – has its limitations…
This lesson we will focus on an alternative way
  - Functional dependencies (FD’s)

Disadvantages of one table:
  - Redundant storage
    - Harder to update
    - Consistency issues
  - How should one represent a customer that has not ordered any book?
Notations

• Attributes - A,B,C,…
• Sets of attributes – X,Y,…
• We will often replace
  – {A} with A
  – {A,B} with AB
• Relational Schemas – R,S,T,…
  – R(A,B,C) or R={A,B,C} or R[A,B,C]
• The content of the relations – r, s, t
  – r={(1,2,3), (2,1,4)}
• Set of functional dependencies – F
  – A single functional dependency - f

Functional Dependencies - definitions

• Let
  – R(A,…,A_n) be a relational schema, let
  – r be a relation over R and let
  – X,Y ⊆ R be sets of attributes
• r is said to satisfy the functional dependency X→Y if
  – every two tuples that have the same value in X have the same values in Y
  – We denote this by r ⊨ X→Y

Outline

• Schema design
  – Introduction
  – Notation
• Functional dependencies
  – Armstrong’s axioms
  – Minimal cover
  – Keys
• Questions

Armstrong’s Axioms

• Three axioms for proving existence of dependencies. Given X,Y,Z⊆R:
  – Reflexivity: if X⊆Y then Y→X
  – Inclusion: if X→Y then XZ→YZ
  – Transitivity: if X→Y and Y→Z then X→Z
• Other modus ponens:
  – Union: if X→Y and X→Z then X→YZ
  – Decomposition: if X→Y and Z⊆Y then X→Z
  – Semi-transitivity: if X→Y and WY→Z then WX→Z

Armstrong’s Axioms

• Let F be a set of functional dependencies and let f be a FD.
  – f is provable from F (f ⊢ F) if f can be deduced from F by Armstrong’s axioms.
  – That is, we can formally prove f from F
Armstrong’s Axioms - example

• F={Cust_Id→Track, Track→Faculty}
• We show that
  – F⊦ Cust_Id→{Track, Faculty}

  1. Track → Faculty ∈ F
  2. Track → {Track, Faculty} ∈ Inclusion, 1
  3. Cust_Id → Track ∈ F
  4. Cust_Id → {Track, Faculty} ∈ 2,3,Trans.

Outline

• Schema design
  – Introduction
  – Notation
• Functional dependencies
• Armstrong’s axioms
• Minimal cover
• Keys
• Questions

Closure of FD’s

• Let F be a set of FD’s
  – The closure of F (F⁺) is the set
    \[ \{X \rightarrow Y \mid F \vdash X \rightarrow Y\} \]

• Example: F={A→B, B→C}, the set F⁺ contains the following FD’s:
  – A→C, AB→C, AC→C, B→B, A→B, Ø→Ø, C→Ø
• Note that the set F⁺ is exponential and therefore we will try to avoid computing it.

Closure of a property

• Let
  – X be a set of properties and let
  – F be a set of FD’s
• The closure of X with respect to F (X⁺ₓ) is the set \( \{A \mid F \vdash X \rightarrow A\} \)
  – Note that A is a single attribute

Minimal Cover

• A set of FD’s might have “redundant” information, for instance the sets
  – F={A→B, B→C, A→C}
  – G={A→B, B→C}

  are equivalent in the sense that F⁺=G⁺
  – The dependency A→C is redundant
• Our goal: A unified form for FD’s

Minimal set of FD’s

• Let F be a set of FD’s, F is minimal if for every FD X→Y∈ F the following hold:
  – |Y|=1
  – F⁺ does not equal (F \ {X→Y})⁺
  – For all Z⊂X the following hold:
    • F⁺ does not equal ((F \ {X→Y}) \cup {Z→Y})⁺
    • i.e., F does not contain an FD X→A such that X includes redundant attributes
Minimal Cover

- Let $F$ and $G$ be sets of FD's
  - $G$ is a cover for $F$ (and vice versa) if $F^+ = G^+$
  - In this case we can use $F$ instead of $G$

- A set of FD's $F_C$ is a minimal cover of $F$ if
  - It is a cover for $F$
  - It is minimal

- Note that there might be more than one minimal cover

Algorithm for finding a minimal cover

- Let $F$ be a set of FD's we define
  - $G \leftarrow \{ (X \rightarrow A) \mid \exists Y ((X \rightarrow Y) \in F \land A \in Y) \}$

Repeat:
1. For each $f = X \rightarrow A \in G$ do:
   - if $A \in X^+_G \setminus \{f\}$ then $G \leftarrow G \setminus \{f\}$
2. For each $f = X \rightarrow A \in G$ and $B \in X$ do
   - if $A \in (X \setminus \{B\})^+_G$ then
     - $G \leftarrow (G \setminus \{X \rightarrow A\}) \cup \{X \setminus B \rightarrow A\}$

Until no more changes to $G$

Finding a minimal cover - example

- Let
  - $R = \{A, B, C, D\}$
  - $F = \{A \rightarrow B, BC \rightarrow A, ABC \rightarrow D, D \rightarrow A\}$

- Find a minimal cover of $F$

1st stage $G \leftarrow F$

Step 1: no change

Step 2: We omit the $A$ from the FD $ABC \rightarrow D$ (since $D \in (\{A, B, C\}\setminus \{A\})^+_G$ and obtain
  - $G = \{A \rightarrow B, BC \rightarrow A, BC \rightarrow D, D \rightarrow A\}$

Step 1: We omit the FD $BC \rightarrow A$ (since $A \in BC^+_G \setminus \{BC \rightarrow A\}$ and obtain
  - $G = \{A \rightarrow B, BC \rightarrow D, D \rightarrow A\}$

- There are no more changes and therefore $G$ is the minimal cover of $F$.

Outline

- Schema design
  - Introduction
  - Notation
- Functional dependencies
- Armstrong’s axioms
- Minimal cover
- Keys
- Questions

Keys

- Let
  - $R$ be a relational schema and let
  - $X \subseteq R$ be a subset of attributes and let
  - $F$ be a set of FD's

- $X$ is a superkey of $R$ if and only if $F = X \rightarrow R$, or equivalently:
  - $X$ is a superkey of $R$ if $X \rightarrow^* F$
  - $X$ is a superkey of $R$ if $X^+ = R$

Keys – cont’d

- Let $R = \{A, B, C, D\}$ and $F = \{A \rightarrow C, B \rightarrow D\}$
  - $ABC$ is a superkey of $R$
    - However, it is not unique
    - And not minimal

- $X$ is a key of $R$ if
  - It is a superkey of $R$
  - There does not exist $Y \subset X$ such that $Y$ is also a superkey
Question

Let $U$ be a schema and let $F$ be a non empty set of non trivial FDs over $U$.
Assume that the FDs in $F$ are of the form $X \rightarrow A$ where $A$ is a single attribute.

- True/False
  - For every non-trivial FD $X \rightarrow A$ in $F$ there exists a key $K$ such that $A \in K$
  - No!
  - $U=\{A,B\}$
  - $F=\{A \rightarrow B\}$
  - The only key is $A$ and $B$ does not belong to the set $\{A\}$

Question – cont’d

- The number of keys is at most the number of attributes in $U$?
  - False.
  - Let $U=\{A,B,C,D\}$ and let
  - $F=\{AB \rightarrow CD, AC \rightarrow BD, AD \rightarrow BC, BC \rightarrow AD, BD \rightarrow AC, CD \rightarrow AB\}$
  - Then $AB$, $AC$, $AD$, $BC$, $BD$, $CD$ are keys
  - But there are only 4 attributes

Question – cont’d

- For every key $K$ that does not equal $U$, there exists a non trivial FD $X \rightarrow A$ in $F$ such that $A \in K$?
  - This statement is true.
  - There exists $A \in U$ such that $A \notin K$
  - Since $K$ is a key, it holds that $A \in K$
  - Since $A \in K$, there exists $X$ such that $A \notin X$ and $X \rightarrow A \in F$.

Question – cont’d

- Let $F$ be a minimal cover and
  - $L$ be the set of attributes that appear in the l.h.s of FD’s in $F$, and
  - $R$ be the set of attributes that appear in the r.h.s of FD’s in $F$
  - If $R\cap L=\emptyset$ then there is a unique key?
    - True.
    - $U\setminus R$ is a superkey.
    - In fact, $X=U\setminus R$ is a key since every $A \in U\setminus R$ cannot be omitted without affecting its closure.
    - Let $X$ be a key, we contend that $X=U\setminus R$.
    - If $X$ does not contain $B \in U\setminus R$ then $B \notin (X)+F$.
    - Therefore $X$ contains $U\setminus R$.
    - Since $X$ is a key it is minimal, so $X=U\setminus R$.
    - Thus, $U\setminus R$ is a unique key