Section 3

CSE 444
Introduction to Databases
Announcements

• Project 1 was due yesterday (10/14/2009)
• Homework 1 was released, due 10/28/2009
From Last time...

• DELETE FROM Table WHERE column = value
  – Don’t forget the WHERE clause
  – Otherwise this empties the content of the table
Today

• E/R Diagrams (Brief overview)
  – English requirements to E/R Diagram
  – E/R diagram to Tables

• BCNF
  – FDs, Closure
  – Examples
E/R basics

• Know and symbols
  – Entity
  – Attributes
  – Relationship
  – Arrows

• ISA
  – Difference from OOP in C++/Java
E/R (English requirements to diagram)

- Each project is managed by one professor (principal investigator)
- Professor can manage multiple projects
E/R (English requirements to diagram)

- Each project is managed by one professor (principal investigator)
- Professor can manage multiple projects

Example courtesy: Database Management Systems, 3rd E, R. Ramakrishnan and J. Gehrke
E/R (English requirements to diagram)

- Each project is managed by one professor (principal investigator)
- Professor can manage multiple projects

Example courtesy: Database Management Systems, 3rd E, R. Ramakrishnan and J. Gehrke
E/R (English requirements to diagram)

- Each project is worked on by one or more professors
- Professors can work on multiple projects

Example courtesy: Database Management Systems, 3rd E, R. Ramakrishnan and J. Gehrke
E/R (English requirements to diagram)

• Each project is worked on by one or more professors
• Professors can work on multiple projects

Example courtesy: Database Management Systems, 3rd E, R. Ramakrishnan and J. Gehrke
Convert to tables

- Professor(ssn, age, rank, specialty)
- Project(pid, sponsor, start_date, end_date, budget)
- Work_in(ssn, pid)
- Manages(ssn, pid)

Example courtesy: Database Management Systems, 3rd E, R. Ramakrishnan and J. Gehrke
Professor

- $ssn$
- $age$
- $rank$
- $specialty$

Project

- $pid$
- $sponsor$
- $start_date$
- $end_date$
- $budget$

Work_in

- $ssn$
- $pid$

Example courtesy: Database Management Systems, 3rd E, R. Ramakrishnan and J. Gehrke
CREATE TABLE Professor (  
    ssn  INT PRIMARY KEY,  
    age  INT,  
    urank VARCHAR(30),  
    specialty VARCHAR(30)  
);  

CREATE TABLE Project (  
    pid  INT PRIMARY KEY,  
    sponsor  INT,  
    start_date DATE,  
    end_date DATE,  
    budget  FLOAT,  
    ssn  INT REFERENCES Professor(ssn)  
);  

CREATE TABLE Work_In (  
    ssn  INT REFERENCES Professor(ssn),  
    pid  INT REFERENCES Project(pid),  
    PRIMARY KEY (ssn, pid)  
);
Data Anomalies

• Redundancy is Bad, why?

• Redundancy

• Update

• Delete
### Functional Dependencies

- Dependencies for this relation:
  - \( A \rightarrow B \)
  - \( A \rightarrow D \)
  - \( B,C \rightarrow E,F \)

- Do they all hold in this instance of the relation \( R \)?

<table>
<thead>
<tr>
<th>( R )</th>
<th>( A )</th>
<th>( B )</th>
<th>( C )</th>
<th>( D )</th>
<th>( E )</th>
<th>( F )</th>
</tr>
</thead>
<tbody>
<tr>
<td>a1</td>
<td>b1</td>
<td>c1</td>
<td>d1</td>
<td>e1</td>
<td>f1</td>
<td></td>
</tr>
<tr>
<td>a1</td>
<td>b1</td>
<td>c2</td>
<td>d1</td>
<td>e2</td>
<td>f3</td>
<td></td>
</tr>
<tr>
<td>a2</td>
<td>b1</td>
<td>c2</td>
<td>d3</td>
<td>e2</td>
<td>f3</td>
<td></td>
</tr>
<tr>
<td>a3</td>
<td>b2</td>
<td>c3</td>
<td>d4</td>
<td>e3</td>
<td>f2</td>
<td></td>
</tr>
<tr>
<td>a2</td>
<td>b1</td>
<td>c3</td>
<td>d3</td>
<td>e4</td>
<td>f4</td>
<td></td>
</tr>
<tr>
<td>a4</td>
<td>b1</td>
<td>c1</td>
<td>d5</td>
<td>e1</td>
<td>f1</td>
<td></td>
</tr>
</tbody>
</table>

- How would you go by finding these in an unknown table?

- Functional dependencies are specified by the database programmer based on the intended meaning of the attributes.
Keys

• Keys, what?
  – Superkey
  – Key
BCNF

• What is it?
BCNF Decomposition Algorithm

BCNF_Decompose(R)

find X s.t.: X ≠X⁺ ≠ [all attributes]

if (not found) then “R is in BCNF”

let Y = X⁺ - X
let Z = [all attributes] - X⁺

decompose R into R1(X ∪ Y) and R2(X ∪ Z)
continue to decompose recursively R1 and R2
A table $R(A, B, C, D, E)$: Example 1

Consider the following FDs:

- $CD \rightarrow E$  \hspace{1cm} BAD
- $D \rightarrow B$  \hspace{1cm} BAD
- $A \rightarrow CD$

Which one are the bad dependences?

- $CD^+ = BCDE$
- $D^+ = BD$
- $A^+ = ABCDE$

Note: a set of attributes $X$ is a superkey if $X^+ = ABCDE$
A table $R(A,B,C,D,E)$: Example 1

Consider the following FDs:
- $CD \rightarrow E$ BAD
- $D \rightarrow B$ BAD
- $A \rightarrow CD$

Note: a set of attributes $X$ is a superkey if $X^+ = ABCDE$
A table $R(A,B,C,D)$: Example 2

Consider the following FDs:

- $C \rightarrow D$, $C^+ = AD$  BAD
- $C \rightarrow A$, $C^+ = AD$  BAD
- $B \rightarrow C$, $B^+ = ABCD$

Note: a set of attributes $X$ is a superkey if $X^+ = ABCDE$
A table $S(A,B,C,D,E)$ : Example 3

Consider the following FDs:

- $AB \rightarrow C$, $AB^+ = ABCD$  BAD
- $DE \rightarrow C$, $DE^+ = CDE$  BAD
- $B \rightarrow D$, $B^+ = BD$  BAD

1st Solution:

<table>
<thead>
<tr>
<th>Table</th>
<th>FD</th>
<th>BCNF</th>
</tr>
</thead>
<tbody>
<tr>
<td>$S(A,B,C,D,E)$</td>
<td>$AB^+ = ABCD \neq ABCDE$</td>
<td></td>
</tr>
<tr>
<td>$S2(A,B,C,D)$</td>
<td>$B^+ = BD \neq ABCD$</td>
<td>[BCNF]</td>
</tr>
<tr>
<td>$S3(A,B,E)$</td>
<td></td>
<td>[BCNF]</td>
</tr>
<tr>
<td>$S4(B,D)$</td>
<td></td>
<td>[BCNF]</td>
</tr>
<tr>
<td>$S5(A,B,C)$</td>
<td></td>
<td>[BCNF]</td>
</tr>
</tbody>
</table>

Note: a set of attributes $X$ is a superkey if $X^+ = ABCDE$
A table $S(A,B,C,D,E)$ : Example 3

Consider the following FDs:

- $AB \rightarrow C$, $AB^+ = ABCD$  BAD
- $DE \rightarrow C$, $DE^+ = CDE$  BAD
- $B \rightarrow D$, $B^+ = BD$  BAD

$S(A,B,C,D,E)$  
$[DE^+ = CDE \neq ABCDE]$

2nd Solution:

- $S2(A,B,D,E)$  
  $[B^+ = BD \neq ABDE]$

- $S3(C,D,E)$  
  $[BCNF]$

- $S4(B,D)$  
  $[BCNF]$

- $S5(A,B,E)$  
  $[BCNF]$

Note: a set of attributes $X$ is a superkey if $X^+ = ABCDE$
A table $S(A,B,C,D,E)$ : Example 3

Consider the following FDs:

- $AB \rightarrow C$, $AB^+ = ABCD$  BAD
- $DE \rightarrow C$, $DE^+ = CDE$  BAD
- $B \rightarrow D$, $B^+ = BD$  BAD

3rd Solution:

- $S(A,B,C,D,E)$
  $[B^+ = BD \neq ABCDE]$
- $S2(A,B,C,E)$
  $[AB^+ = ABC \neq ABCE]$
- $S3(B,D)$
  $[BCNF]$
- $S4(A,B,C)$
  $[BCNF]$
- $S5(A,B,E)$
  $[BCNF]$

Note: a set of attributes $X$ is a superkey if $X^+ = ABCDE$