Functional Dependencies (II)

R&G Chapter 19

Annoucement

- **Final exam**
  - 9:30am – 12noon, Saturday, December 17, 2016
  - For both CS442 and CPE442 students

In Last Lecture

- **Data redundancy**
  - Update, insertion, deletion anomaly
- **Source for redundancy: functional dependencies (FDs)**
  - $X \rightarrow Y$: Given any two tuples in *r*, if the X values are the same, then the Y values must also be the same. (but not vice versa)
- **To remove data redundancy**
  - FD-driven schema decomposition

Rules of Inference

- **Armstrong's Axioms (AA) (X, Y, Z are <u>sets</u> of attributes):**
  - *Reflexivity*: If $X \supseteq Y$, then $X \rightarrow Y$
  - *Augmentation*: If $X \rightarrow Y$, then $XZ \rightarrow YZ$ for any Z
  - *Transitivity*: If $X \rightarrow Y$ and $Y \rightarrow Z$, then $X \rightarrow Z$

- **Some additional rules (that follow from AA):**
  - *Union*: If $X \rightarrow Y$ and $X \rightarrow Z$, then $X \rightarrow YZ$
  - *Decomposition*: If $X \rightarrow YZ$, then $X \rightarrow Y$ and $X \rightarrow Z$

Example

Given:

–Relation R = {A,B,C,G,H,I}

–FDs: $A \rightarrow B$, $A \rightarrow C$, $CG \rightarrow H$, $CG \rightarrow I$, $B \rightarrow H$

Questions: Prove the following FDs by using AA rules

(1) A->H,

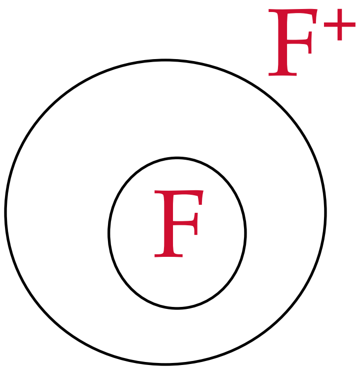
(2) $AG \rightarrow I$

(3) $CG \rightarrow HI$

Closure of FDs

- **An FD *f* is *implied by* a set of FDs *F* if *f* holds whenever all FDs in *F* hold.**

- **F^+ = *closure of F* is the set of all FDs that are implied by *F*. (includes "trivial dependencies")**



Computing FD Closure

- **Typically we want to check if a given FD $X \rightarrow Y$ can be implied from a given set of FDs F.**

- **It is equivalent to checking whether $X \rightarrow Y$ is in F^+.**

- **An efficient check:**
  - Compute *attribute closure* of X (denoted X^+) wrt *F*.
    X^+ = Set of attributes A such that $X \rightarrow A$ is in F^+
    - Initialize $X^+ := X$
    - Repeat until no change: if there is an FD $U \rightarrow V$ in *F* such that U is in X^+, then add V to X^+
  - Check if Y is in X^+ (i.e., $X \rightarrow Y$ is in F^+)

Attribute Closure (Example 1)

- **R(ABCDE)**
- **F={A->D, D->B, B->C, E->B}**
- **What's A^+, D^+, E^+, ACE^+ ?**

Attribute Closure (Example 2)

- **R = {A, B, C, D, E}**
- **F = { B $\rightarrow$CD, D $\rightarrow$ E, B $\rightarrow$ A, E $\rightarrow$ C, AD $\rightarrow$B }**
- **Is B $\rightarrow$ E in F^+ ?**

B^+ = B

B^+ = BCD

B^+ = BCDA

B^+ = BCDAE ... Yes!

FD Closure VS. Finding Key

- **Computing FD closure can be used to find the keys of a relation.**
  - If X^+ = {all attributes of R} , then X is a superkey for R.
  - Question: How to check if X is a *candidate* key?
  - Answer: check whether any subset Y of X satisfies:

$$Y^+ = \{\text{all attributes of } R\}$$

Key Calculation

- R = {A, B, C, D, E}
- F = { B $\rightarrow$CD, D $\rightarrow$ E, B $\rightarrow$ A, E $\rightarrow$ C, AD $\rightarrow$B }
- **Is D a superkey of R?**
- **Is B a superkey of R?**
- **Is B a candidate key of R?**
- **Is AD a superkey of R?**
- **Is AD a *candidate* key of R?**
- **Is ADE a *candidate* key of R?**

How to Determine Candidate Keys?

- **An efficient solution:**
  - When computing FD closure, we distinguish attributes into three categories
    - L: attributes <u>only</u> appear at the <u>left</u> side of all given FDs
    - R: attributes <u>only</u> appear at the <u>right</u> side of all given FDs
    - M: attributes that appear on <u>both</u> sides of all given FDs
  - The principle:
    - Attributes in L: each candidate key should include ***ALL*** attributes in L;
    - Attributes in M may be part of keys
    - Attributes in R will ***NOT*** be part of any key

Determine the keys (example 1)

- **Database : R(A, B, C, D)**
- **FDs: (AB -> C, C-> B, C-> D)**
- **What are the candidate keys of R?**

Determine the keys (example 2)

- **Database : R(A, B, C)**
- **FDs: (A->B, B-> C, C-> A)**
- **What are the candidate keys of R?**

L	M	R
	A, B, C	

– A^+ = {ABC}

– B^+ = {BCA}

– C^+ = {CAB}

– Keys: A, B, C