TRƯỜNG ĐẠI HỌC BÁCH KHOA HÀ NỘI
HANOI UNIVERSITY OF SCIENCE AND TECHNOLOGY

Database
Lesson 8. Functional Dependency

Ba Lam DO

Learning Map

Sequence	Title
1	Introduction to Databases
2	Relational Databases
3	Relational Algebra
4	Structured Query Language – Part 1
5	Structured Query Language – Part 2
6	Constraints and Triggers
7	Entity Relationship Model
8	Functional Dependency
9	Normalization
10	Storage - Indexing
11	Query Processing
12	Transaction Management – Part 1
13	Transaction Management – Part 2

Outline

- Functional Dependency
- Armstrong 's Axioms and secondary rules
- closure of a FD set, closure of a set of attributes
- A minimal key
- Equivalence of sets of functional dependencies
- Minimal Sets of FDs

Objectives

- Upon completion of this lesson, students will be able to:
  - Recall the concepts of functional dependency, Armstrong 's axioms and secondary rules
  - Identify closure of a FD set, closure of a set of attributes
  - Find a minimal key of a relation under a set of FDs
  - Identify the equivalence of sets of FDs and find the minimal cover of a set of FDs

1. Functional Dependency (FD)

- Introduction
- Definition

1.1. Introduction

- FD is the single most important concept in relational schema design theory
  - We have to deal with the problem of database design: anomalies, redundancies
  - **Redundancy** means having multiple copies of same data in the database.

Suppliers(sid, sname, city, Number of Employee, product, quantity)

Sid	Sname	City	NOE	Product	quantity
S1	Smith	London	100	Screw	50
S1	Smith	London	100	Nut	100
S2	J&J	Paris	100	Screw	78
S3	Blake	Tokyo	75	Bolt	100

1.1. Introduction: Anomalies in database design

- Insertion Anomaly
- Deletion Anomaly
- Updation Anomaly

Sid	Sname	City	NOE	Product	quantity
S1	Smith	London	100	Screw	50
S1	Smith	London	100	Nut	100
S2	J&J	Paris	100	Screw	78
S3	Blake	Tokyo	75	Bolt	100

1.2. Definition

- Suppose that R(U), U = {A_1, A_2, ... , A_n} , X and Y are non-empty subsets of U.
- A functional dependency (FD), denoted by X → Y, specifies a constraint on the possible tuples that can form a relation state r of R. The constraint is that, for any two tuples t_1 and t_2 in r that have $t_1[X] = t_2[X]$, they must also have $t_1[Y] = t_2[Y]$.
  - X: the left-hand side of the FD
  - Y: the right-hand side of the FD

1.2. Definition (cont.)

- This means that the values of the X component of a tuple uniquely (or functionally) determine the values of the Y component.
- A FD $X \rightarrow Y$ is trivial if $X \supseteq Y$
- If X is a candidate key of R, then $X \rightarrow R$

1.2. Definition (cont.)

- Examples
  - $AB \rightarrow C$

A	B	C	D
a1	b1	c1	d1
a1	b1	c1	d2
a1	b2	c2	d1
a2	b1	c3	d1

  - subject_id $\rightarrow$ name,
  - subject_id $\rightarrow$ credit,
  - subject_id $\rightarrow$ percentage_final_exam,
  - subject_id $\rightarrow$ {name, credit}

<u>subject_id</u>	name	credit	percentage_final_exam
IT3090	Databases	3	0.7
IT4843	Data integration	3	0.7
IT4868	Web mining	2	0.6
IT2000	Introduction to ICT	2	0.5
IT3020	Discrete Mathematics	2	0.7
IT3030	Computer Architectures	3	0.7

2. Armstrong 's axioms

- Armstrong 's axioms
- Secondary rules
- An example

2.1. Armstrong axioms

- $U = \{A_1, A_2, \dots, A_n\}$, X, Y, Z, W are subsets of U.
- XY denoted for $X \cup Y$

- Reflexivity
  - If $Y \subseteq X$ then $X \to Y$
- Augmentation
  - If $X \to Y$ then $XZ \to YZ$
- Transitivity
  - If $X \to Y$, $Y \to Z$ then $X \to Z$

2.2. Secondary rules

- Union
  - If $X \rightarrow Y$, $X \rightarrow Z$ then $X \rightarrow YZ$.
- Pseudo-transitivity
  - If $X \rightarrow Y$, $WY \rightarrow Z$ then $XW \rightarrow Z$.
- Decomposition
  - If $X \rightarrow Y$, $Z \subseteq Y$ then $X \rightarrow Z$

2.3. An example

- Given a set of FDs: $F = \{AB \to C, C \to A\}$
- Prove: $BC \to ABC$
  - From $C \to A$, we have $BC \to AB$ (Augmentation)
  - From $AB \to C$, we have $AB \to ABC$ (Augmentation)
  - And we can conclude $BC \to ABC$ (Transitivity)

3. Closure of a FD set, closure of a set of attributes

- Closure of a FD set
- Closure of a set of attributes
- Discussion

3.1. Closure of a FD set

- Suppose that $F = \{A \rightarrow B, B \rightarrow C\}$ on $R(A, B, C,...)$. We can infer many FD such as: $A \rightarrow C$, $AC \rightarrow BC$,...
- Definition
  - Formally, the set of all dependencies that include F as well as all dependencies that can be inferred from F is called the closure of F, denoted by F^+.
- $F \models X \rightarrow Y$ to denote that the FD $X \rightarrow Y$ is inferred from the set of FDs F.

3.2. Closure of a set of attributes

- Problem
  - We have F, and $X \rightarrow Y$, we have to check if $F \models X \rightarrow Y$ or not
- Should we calculate F^+ ?
  - $\Rightarrow$closure of a set of attributes
- Definition
  - For each such set of attributes X, we determine the set X^+ of attributes that are functionally determined by X based on F; X^+ is called the closure of X under F.

3.2. Closure of a set of attributes (cont.)

- To find the closure of an attribute set X^+ under F

Input: A set F of FDs on a relation schema R, and a set of attributes X, which is a subset of R.

$X^0 := X;$

repeat

  for each functional dependency $Y \rightarrow Z$ in F do

    if $X^{i-1} \supseteq Y$ then $X^i := X^{i-1} \cup Z;$

    else $X^i := X^{i-1}$

until (X^i unchanged);

$X^+ := X^i$

3.2. Closure of a set of attributes (cont.)

- Example
  - Given R(A, B, C, D, E, F) and F = {AB $\rightarrow$ C, BC $\rightarrow$ AD, D $\rightarrow$ E, CF $\rightarrow$ B}. Calculate $(AB)^+_F$
    - X^0 = AB
    - X^1 = ABC (from AB $\rightarrow$ C)
    - X^2 = ABCD (from BC $\rightarrow$ AD)
    - X^3 = ABCDE (from D $\rightarrow$ E)
    - X^4 = ABCDE
    - $(AB)^+_F$=ABCDE

3.3. Discussion

- $X \rightarrow Y$ can be inferred from F if and only if $Y \subseteq X^+_F$
  - $F \vDash X \rightarrow Y \Leftrightarrow Y \subseteq X^+_F$
  - An example
    - Let R(A, B, C, D, E), $F = \{A \rightarrow B, B \rightarrow CD, AB \rightarrow CE\}$. Consider whether or not $F \vDash A \rightarrow C$
    - $(A)^+_F = ABCDE \supseteq \{C\}$

4. Minimal key

- Definition
- An algorithm to find a minimal key
- Example

4.1. Definition

- Given R(U), U = $\{A_1, A_2, \dots, A_n\}$, a set of FDs F
- K is considered as a minimal key of R if:
  - $K \subseteq U$
  - $K \rightarrow U \in F^+$
  - For every $\forall K' \subset K$, then $K' \rightarrow U \notin F^+$
- Discussion
  - $K^+ = U$ and $K \setminus \{A_i\} \rightarrow U \notin F^+$

4.2. An algorithm to find a minimal key

- Input: $R(U)$, $U = \{A_1, A_2, \dots , A_n\}$, a set of FDs F
  - - Step^0 $K^0 = U$
  - - Step^i If $(K^{i-1} \setminus \{A_i\}) \rightarrow U$ then $K^i = K^{i-1} \setminus \{A_i\}$
    - else $K^i = K^{i-1}$
  - - Step^{n+1} $K = K^n$

4.3. Example

- Given R(U), U= {A, B, C, D, E}, F = {AB $\rightarrow$ C, AC $\rightarrow$ B, BC $\rightarrow$ DE}.
- Find a minimal key
- Step 0: K^0= U = ABCDE
- Step 1: Check if or not $(K^0\backslash\{A\}) \rightarrow U$ (i.e, BCDE $\rightarrow$ U).

  $(BCDE)^+$= BCDE $\neq$ U. Then, $K^1 = K^0$ = ABCDE
- Step 2: Check if $(K^1\backslash\{B\}) \rightarrow U$ (i.e, ACDE $\rightarrow$ U).

  $(ACDE)^+$ = ABCDE = U. Then, $K^2 = K^1\backslash\{B\}$ = ACDE
- Step 3: K^3 = ACDE
- Step 4: K^4 = ACE
- Step 5: K^5 = AC
- We conclude that AC is a minimal key

5. Equivalence of Sets of FDs

- Definition
- Example

5.1. Definition

- Definition.
  - A set of FDs F is said to cover another set of FDs G if every FD in G is also in F^+ (every dependency in G can be inferred from F).
  - Two sets of FDs F and G are equivalent if $F^+ = G^+$. Therefore, equivalence means that every FD in G can be inferred from F, and every FD in F can be inferred from G; that is, G is equivalent to F if both the conditions - G covers F and F covers G - hold.

5.2. Example

- Prove that $F = \{A \rightarrow C, AC \rightarrow D, E \rightarrow AD, E \rightarrow H\}$ and $G = \{A \rightarrow CD, E \rightarrow AH\}$ are equivalent
  - For each FD of F, prove that it is in G^+
    - $A \rightarrow C$: $(A)^+_G = ACD \supseteq C$, so $A \rightarrow C \in G^+$
    - $AC \rightarrow D$: $(AC)^+_G = ACD \supseteq D$, so $AC \rightarrow D \in G^+$
    - $E \rightarrow AD$: $(E)^+_G = EAHCD \supseteq AD$, so $E \rightarrow AD \in G^+$
    - $E \rightarrow H$: $(E)^+_G = EAHCD \supseteq H$, so $E \rightarrow H \in G^+$
    - $\Rightarrow F^+ \subseteq G^+$
  - For each FD of G, prove that it is in F^+ (the same)
    - $\Rightarrow G^+ \subseteq F^+$
  - $\Rightarrow F^+ = G^+$

6. A minimal cover of a set of FDs

- Definition
- An algorithm to find a minimal cover of a set of FDs
- Example

6.1. Definition

- Minimal Sets of FDs
  - A set of FDs F to be minimal if it satisfies:
    - Every dependency in F has a single attribute for its right-hand side.
    - We cannot replace any dependency X → A in F with a dependency Y → A, where Y is a proper subset of X, and still have a set of dependencies that is equivalent to F.
    - We cannot remove any dependency from F and still have a set of dependencies that is equivalent to F.

→ a set of dependencies in a standard form with no redundancies in number of dependencies and left, right-hand side of dependencies.

6.2. An algorithm to find a minimal cover of a set of FDs

- Finding a Minimal Cover F for a Set of FDs G

Input: A set of FDs G.

1. Set F := G.

2. Replace each functional dependency $X \rightarrow \{A_1, A_2, ..., A_n\}$ in F by n independent FDs: $X \rightarrow A_1$, $X \rightarrow A_2$, ..., $X \rightarrow A_n$.

3. For each FD $X \rightarrow A$ in F

   for each attribute B that is an element of X

   if $\{\{F - \{X \rightarrow A\}\} \cup \{(X - \{B\}) \rightarrow A\}\}$ is equivalent to F

   then replace $X \rightarrow A$ with $(X - \{B\}) \rightarrow A$ in F.

4. For each remaining functional dependency $X \rightarrow A$ in F if $\{F - \{X \rightarrow A\}\}$ is equivalent to F, then remove $X \rightarrow A$ from F.

6.3. Example

- G = {B → A, D → A, AB → D}. We have to find the minimal cover of G.
- All above dependencies right-hand side are single attributes
- In step 2, we need to determine if AB → D has any redundant attribute
  - on the left-hand side; that is, can it be replaced by B → D or A → D?
  - Since B → A then AB → D may be replaced by B → D.
  - We now have a set equivalent to original G, say G1: {B → A, D → A, B → D}.
- In step 3, we look for a redundant FD in G1. Using the transitive rule on B → D and D → A, we conclude B → A is redundant.
- Therefore, the minimal cover of G is {B → D, D → A}

Remarks

- Functional dependencies
- Armstrong axioms and their' secondary rules
- Closure of a set of FDs,
- Closure of a set of attributes under a set of FDs
- An algorithm to find a minimal key
- Equivalence of sets of FDs
- Finding a minimal set of a set of FDs

Next lesson: Normalization

- Hector Garcia-Molina, Jeffrey D. Ullman, Jennifer Widom. Database Systems: The Complete Book. Pearson Prentice Hall. the 2nd edition. 2008: Chapter 7
- Nguyen Kim Anh, Nguyên lý các hệ cơ sở dữ liệu, NXB Giáo dục. 2004: Chương 7