

MC 302 – DBMS: Functional Dependencies

Goonjan Jain

Department of Applied Mathematics

Delhi Technological University

Functional Dependencies

- Motivation is – create ‘good’ tables
- For example:

Table1(roll_no, course_id, grade, name, address)

Is this table good or bad?

Roll_no	Course_id	Grade	Name	Address
2017/MC/24	MC302	A	Aman	Prime
2017/MC/24	MC304	A+	Aman	Prime
2017/MC/24	MC306	A	Aman	Prime

- Table1 is bad. Why?
- Answer – Redundancy-
 - Space
 - Inconsistency
 - Updation anomalies
- What caused the problem?
- Answer – name depends on roll_no

Functional Dependencies

- Definition – $a \rightarrow b$
- a functionally determines b
- If you know 'a', there is only one 'b' to match

Formally:

$X \rightarrow Y$ implies $(t1[x1] = t2[x1] \text{ then } t1[y1] = t2[y1])$

if two tuples agree on the “X” attribute, the *must* agree on the “Y” attribute, too

- 'X' and 'Y' can be set of attributes
- Other examples of functional dependencies:

Roll_no	Course_id	Grade	Name	Address
2017/MC/24	MC302	A	Aman	Prime
2017/MC/24	MC304	A+	Aman	Prime
2017/MC/24	MC306	A	Aman	Prime

$\text{Roll_no} \rightarrow \text{name, address}$

$\text{Roll_no, course_id} \rightarrow \text{grade}$

Closure

- **Closure** of a set of FD: all implied FDs

- For example –

$\text{Roll_no} \rightarrow \textit{name, address}$

$\textit{Roll_no, course_id} \rightarrow \textit{grade}$

Imply

$\text{Roll_no, course_id} \rightarrow \textit{grade, name, address}$

$\textit{Roll_no, course_id} \rightarrow \textit{roll_no}$

- How to find all the implied ones, systematically?

Armstrong's Axioms

- “Armstrong's axioms” guarantee soundness and completeness:

- ***Reflexivity:***

$$Y \subseteq X \Rightarrow X \rightarrow Y$$

eg., roll_no, name \rightarrow roll_no

- ***Augmentation***

$$X \rightarrow Y \Rightarrow XW \rightarrow YW$$

- eg., roll_no \rightarrow name then roll_no, grade \rightarrow name, grade

Armstrong's Axioms

- *Transitivity*

$$X \rightarrow Y \text{ and } Y \rightarrow Z \Rightarrow X \rightarrow Z$$

For example, roll_no \rightarrow address, and

address \rightarrow HRA_rate

THEN:

roll_no \rightarrow HRA_rate

Armstrong's Axioms

Reflexivity:

$$Y \subseteq X \Rightarrow X \rightarrow Y$$

Augmentation

$$X \rightarrow Y \Rightarrow XW \rightarrow YW$$

Transitivity

$$X \rightarrow Y \text{ and } Y \rightarrow Z \Rightarrow X \rightarrow Z$$

‘sound’ and ‘complete’

Armstrong's axioms

- Additional rules:

- Union

$$X \rightarrow Y \text{ and } X \rightarrow Z \Rightarrow X \rightarrow YZ$$

- Decomposition

$$X \rightarrow YZ \Rightarrow X \rightarrow Y \text{ and } X \rightarrow Z$$

- Pseudo-transitivity

$$X \rightarrow Y \text{ and } YW \rightarrow Z \Rightarrow XW \rightarrow Z:$$

Prove 'Union', 'Decomposition' and 'pseudo-transitivity' from Armstrong's axioms.

Prove ‘Union’ from three axioms:

$$\left. \begin{array}{l} X \rightarrow Y \quad (1) \\ X \rightarrow Z \quad (2) \end{array} \right\}$$

$$(1) + \text{augm. w/ } Z \Rightarrow XZ \rightarrow YZ \quad (3)$$

$$(2) + \text{augm. w/ } X \Rightarrow XX \rightarrow XZ \quad (4)$$

but XX is X ; thus

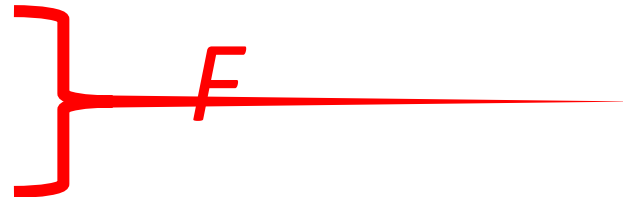
$$(3) + (4) \text{ and transitivity} \Rightarrow X \rightarrow YZ$$

FDs – Closure F^+

- Given a set F of FD (on a schema)
- F^+ is the set of all implied FD. Eg.,
table1(roll_no, course_id, grade, name, address)

roll_no, course_id \rightarrow grade

roll_no \rightarrow name, address



Closure F+

Roll_no, course_id \rightarrow grade

Roll_no \rightarrow name, address

Roll_no \rightarrow roll_no

Roll_no, course_id \rightarrow address

Course_id, address \rightarrow course_id



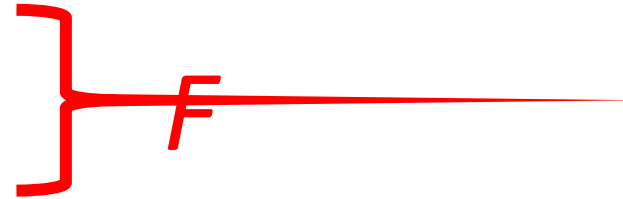
F+

FDs – Closure A+

- Given a set F of FD (on a schema)
- A+ is the set of all attributes determined by A:
table1(roll_no, course_id, grade, name, address)

roll_no, course_id \rightarrow grade

Roll_no \rightarrow name, address



{roll_no}+ = ??

{roll_no}+ = {roll_no, name, address}

{course_id}+ = ??

{course_id, roll_no}+ = ??

FDs – A+ closure

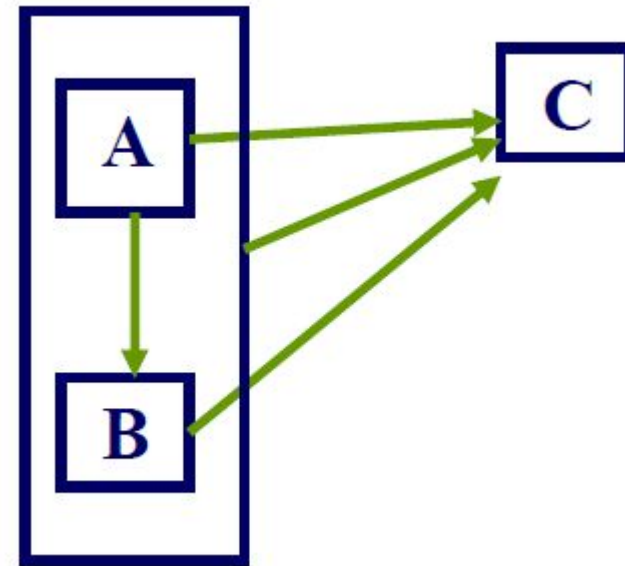
if $A^+ = \{\text{all attributes of table}\}$
then 'A' is a **superkey**

$AB \rightarrow C$ (1)

$A \rightarrow BC$ (2)

$B \rightarrow C$ (3)

$A \rightarrow B$ (4)



Canonical Cover F_c

Given a set F of FD (on a schema)

F_c is a minimal set of equivalent FD. Eg.,

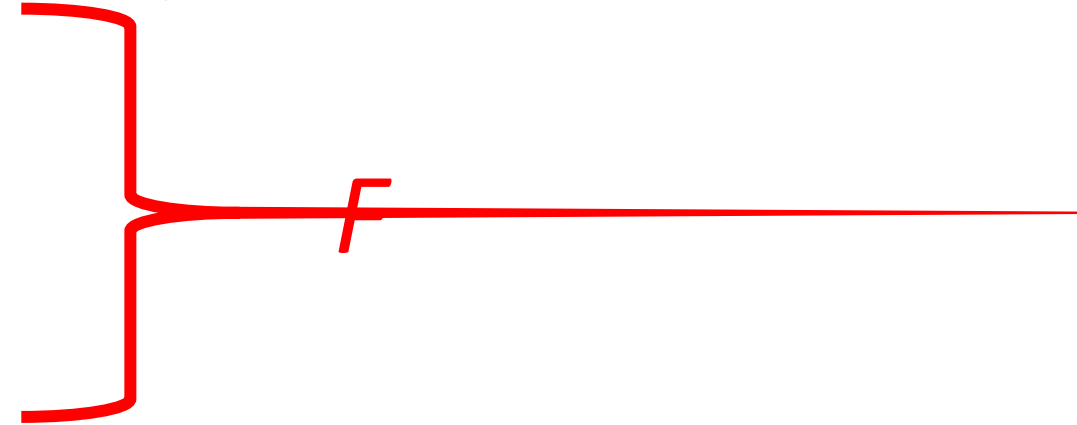
table1(roll_no, course_id, grade, name, address)

roll_no, course_id \rightarrow grade

Roll_no \rightarrow name, address

Roll_no,name \rightarrow name, address

roll_no, course_id \rightarrow grade, name



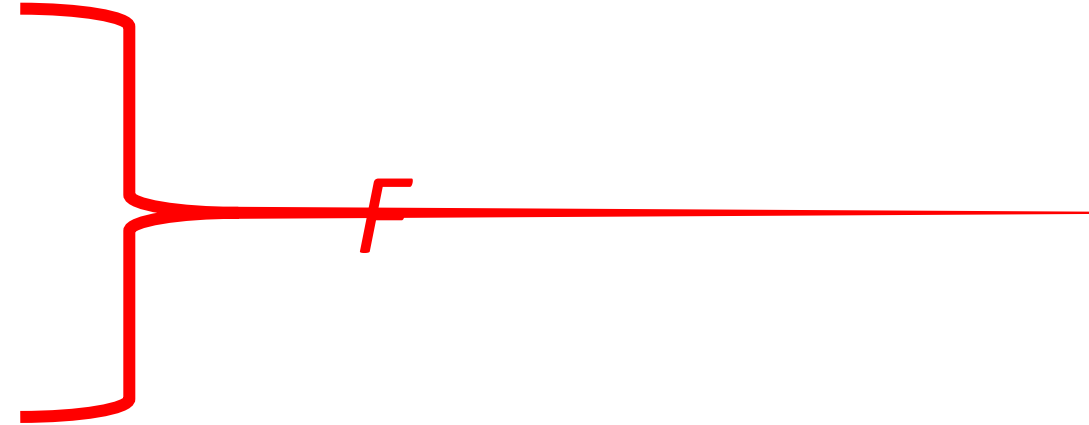
F_c

roll_no, course_id \rightarrow grade

Roll_no \rightarrow name, address

Roll_no, name \rightarrow name, address

roll_no, course_id \rightarrow grade, name



FDs – Canonical cover F_c

- Why do we need it?
- define it properly
- compute it efficiently

FDs – Canonical cover F_c

- Why do we need it?
 - easier to compute candidate keys
- Define it properly – three properties
 - 1) the RHS of every FD is a single attribute
 - 2) the closure of F_c is identical to the closure of F
 - i.e., F_c and F are equivalent
 - 3) F_c is minimal
 - i.e., if we eliminate any attribute from the LHS or RHS of a FD, property #2 is violated
 - we need to eliminate 'extraneous' attributes.
- An attribute is 'extraneous' if
 - the closure is the same, before and after its elimination
 - or if F -before implies F -after and vice-versa

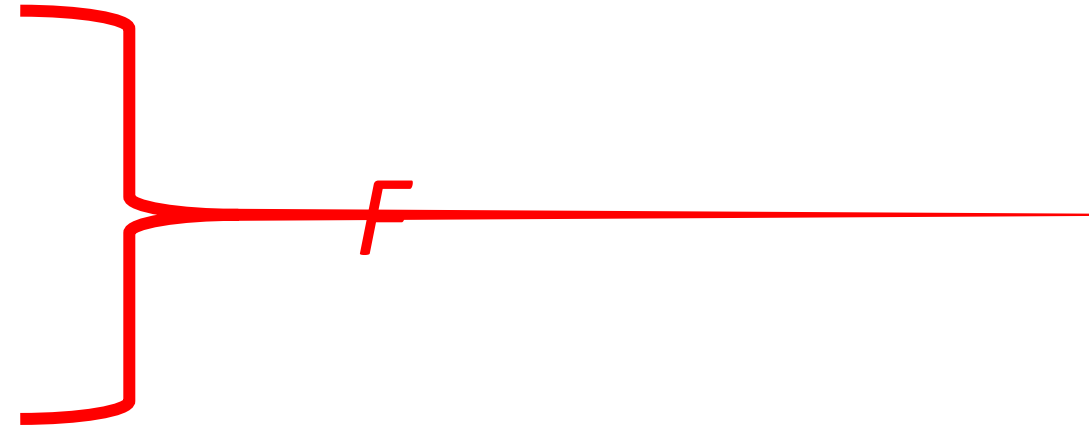
Canonical cover F_c

roll_no, course_id \rightarrow grade

Roll_no \rightarrow name, address

~~Roll_no, name \rightarrow name, address~~

~~roll_no, course_id \rightarrow grade, name~~



Algorithm for Canonical cover F_c

- examine each FD; drop extraneous LHS or RHS attributes; or redundant FDs
- make sure that FDs have a single attribute in their RHS
- repeat until no change

For example: Trace algorithm for

$$AB \rightarrow C \quad (1)$$

$$A \rightarrow BC \quad (2)$$

$$B \rightarrow C \quad (3)$$

$$A \rightarrow B \quad (4)$$

Canonical cover F_c

- Step 1: Split (2)

$$AB \rightarrow C \quad (1)$$

$$A \rightarrow B \quad (2')$$

$$A \rightarrow C \quad (2'')$$

$$B \rightarrow C \quad (3)$$

$$A \rightarrow B \quad (4)$$

Step 2: Delete redundant FDs

$$AB \rightarrow C \quad (1)$$

$$A \rightarrow B \quad (2')$$

$$A \rightarrow C \quad (2'')$$

$$B \rightarrow C \quad (3)$$

$$A \rightarrow B \quad (4)$$

$$AB \rightarrow C \quad (1)$$

$$A \rightarrow C \quad (2'')$$

$$B \rightarrow C \quad (3)$$

$$A \rightarrow B \quad (4)$$

- Step 3: (2'') is redundant (implied by (4), (3) and transitivity)

$$AB \rightarrow C \quad (1)$$

$$AB \rightarrow C \quad (1)$$

$$A \rightarrow C \quad (2'')$$

$$B \rightarrow C \quad (3)$$

$$B \rightarrow C \quad (3)$$

$$A \rightarrow B \quad (4)$$

$$A \rightarrow B \quad (4)$$

- Step 4: in (1), 'A' is extraneous:

$$AB \rightarrow C \quad (1)$$

$$B \rightarrow C \quad (1')$$

$$B \rightarrow C \quad (3)$$

$$B \rightarrow C \quad (3)$$

$$A \rightarrow B \quad (4)$$

$$A \rightarrow B \quad (4)$$

Canonical Cover F_c

- Step 5: (1') is redundant

~~$B \rightarrow C$~~ (1')

$B \rightarrow C$ (3)

$A \rightarrow B$ (4)

$B \rightarrow C$ (3)

$A \rightarrow B$ (4)

- nothing is extraneous
- all RHS are single attributes
- final and original set of FDs are **equivalent** (same closure)

Before

$AB \rightarrow C$ (1)

$A \rightarrow BC$ (2)

$B \rightarrow C$ (3)

$A \rightarrow B$ (4)

After

$B \rightarrow C$ (3)

$A \rightarrow B$ (4)

Equivalence of set of FDs

- Given - 2 sets of FDs, E and F
- If every dependency of E can be inferred from F, then E is covered by F, and vice versa.
- If $E^+ = F^+$, E covers F and F covers E
- Check – F covers E
 - For all $X \rightarrow Y$ in E, calculate X^+
 - Check if X^+ includes the attributes in Y.
- For example: $E = \{A \rightarrow C, AC \rightarrow D, E \rightarrow AD, E \rightarrow H\}$
 $F = \{A \rightarrow CD, E \rightarrow AH\}$

Check – F covers E?

$E = \{A \rightarrow C, AC \rightarrow D, E \rightarrow AD, E \rightarrow H\}$

$F = \{A \rightarrow CD, E \rightarrow AH\}$

In E:

$A \rightarrow C$

Compute A^+ using FDs in F

$A^+ = \{A, C, D\}$

$AC \rightarrow D$

Compute $\{AC\}^+$ using FDs in F

$\{AC\}^+ = \{A, C, D\}$

$E \rightarrow AD$

Compute $\{E\}^+$ using FDs in F

$\{E\}^+ = \{E, A, H, C, D\}$

$E \rightarrow H$

Compute $\{E\}^+$ using FDs in F

$\{E\}^+ = \{E, A, H, C, D\}$

- Thus, F covers E.

Check – E covers F?

$E = \{A \rightarrow C, AC \rightarrow D, E \rightarrow AD, E \rightarrow H\}$

$F = \{A \rightarrow CD, E \rightarrow AH\}$

In F:

$A \rightarrow CD$

Compute A^+ using FDs in E

$A^+ = \{A, C, D\}$

$E \rightarrow AH$

Compute $\{E\}^+$ using FDs in E

$\{E\}^+ = \{E, A, D, H, C\}$

Thus, E covers F.

Is E equivalent to F?

- E is covered by F, and
- F is covered by E
- Thus, E is equivalent to F.

Question 1:

- Find canonical cover for the following sets of dependencies:
 - $E = \{B \rightarrow A, D \rightarrow A, AB \rightarrow D\}$
 - $F = \{A \rightarrow BCDE, CD \rightarrow E\}$