Normalization

Normalization used to reduce or eliminate redundancy.

⇒ Redundancy in DB Table occurs if two or more independent relations stored in one relation.

Functional Dependencies: { ① Sid → Sname DoB ② Cid → Cname Instr ③ Sid Cid → Fee

Independent Relation

{Sid Cid} PK

{ I D U } XXX

Sid	Sname	DoB	Cid	Cname	Instr	Fee
S1	B	2005	C1	DB	Korth	5000
S2	B	2005	C2	DB	Korth	6000
S3	A	2000	C1	DB	Korth	5000
S3	A	2000	C2	Algo	Coreman	4000
S3	A	2000	C3	Algo	Ullman	6000
—	—	—	C4	OS	Galwin	—

⇒ Problem because of Redundancy :- (DB Anamalies).

① Insertion Anamaly :-

In order to insert some data also required to insert other independent data.

② Deletion Anamaly :-

Because of deletion of some data also deletes other independent data

Date________
Page________

③ Updation Anamaly :-
If some redundant copy updated && other redundant copy failed to update causes incosistency.

Normalization of DB Table :-
Decompose relation into two or more sub relations to reduce /elliminate redundancy. && DB anamalies.

R_1 Sid → Sname DoB

Sid	Sname	DoB
S_1	B	2005
S_2	B	2005
S_3	A	2000

R_2 Sid Cid → fee

FK FK

Sid	Cid	fee
S_1	C_1	-
S_2	C_1	-
S_3	C_1	-
S_3	C_2	-
S_3	C_3	-

R_3 Cid → Cname Inst

Cid	Cname	Instr
C_1	DB	K
C_2	Algo	C
C_3	Algo	U
C_4	OS	G

Normarlized DB design : 0% redundancy
No DB anamalies.

Functional Dependency [FD] ($X \to Y$)

FD's are constraints that are derived from the meaning and inter relationships of data attributes.

A set of attributes x functionally determines a set of attributes y if the values of x determines a unique value of y.

$x \to y$ holds if whenever two tuples have same value for x, they must have the same value for y

for any two tuples t_1, t_2 in rel R
if $t_1.x = t_2.x$ then $t_1.y = t_2.y$

rck

R

x	y	z
x_1	y_5	z_1
x_1	y_5	z_2
x_1	y_5	z_3
x_2	y_1	z_4
x_3	y_1	z_5
x_4	y_2	z_6

$x \to y$ exists R

$\left.\begin{array}{l} z \to x \\ z \to y \end{array}\right\}$?/ exists ✓

Date
Page

If K is a key of R, then K functionally determines all attributes in R.

R	A	B	C	D
	a_1			
	a_2			
	a_3			
	a_4			

$A \rightarrow BCD$ guaranted. exists in R

A : Key

R	Sid	Sname	Cid
	S_1	A	C_1
	S_1	A	C_2
	S_2	B	C_2
	S_3	B	C_3
	S_4	C	C_1

If x, y, z Attribute sets over Rel R :

① Trivial FD	② Non Trivial FD	③ Semi Non Trivial FD
If y is sub set of x then $x \rightarrow y$ is trivial FD. [$x \supseteq y$]	If Attributes Set x determine other attribute y set [$x \cap y = \Phi$]	Combination of Trivial FD and Non Trivial FD
Sid → Sid Sname → Sname Cid → Cid Sid Sname → Sname Sid Sname → Sid Sname	Sid → Sname Sid Cid → Sname	Sid → Sid Sname { Sid → Sid, Sid → Sname } Sid Cid → Cid Sname { Sid Cid, Sid Cid → Sname }

NOTE – Every Possible trivial FD over attributes of R is always members of Rel R.

Date
Page

R(AB)	$A \to A$	$A \to B$	$A \to AB$
	$B \to B$	$B \to A$	$B \to AB$
	$AB \to A$		
	$AB \to B$		
	$AB \to AB$		
	Trivial FD's	non Trivial FD's	semi non trivial

X, Y, Z. are some attributes sets over R.

Armstrong's inference rules over FD's

① Reflexivity → [Trivial FD]
If Y is subset of X then $X \to Y$
$Sid \to Sid$
$SidCid \to Cid$

② Augmentation →
If $X \to Y$ then $XZ \to YZ$

1:40

③ Transitivity →
If $X \to Y$ and $Y \to Z$ then $X \to Z$

completeness of FD inferences rules

④ Decomposition
If $X \to YZ$ then $X \to Y$, $X \to Z$

⑤ Union
If $X \to Y$, $X \to Z$ then $X \to YZ$

⑥ Psudo transitivity :-
If $X \to Y$, $WY \to Z$ then $WX \to Z$

can be derived by using ① ② ③

Date
Page

Q. Find all Possible Non Trivial FD's which are satisfied for given instance.

(I) R

A	B	C
2	3	4
2	3	5
2	6	5
4	7	2
4	7	3

$x \rightarrow y$

$t_1 \; x_1 \rightarrow y_1$

$t_2 \; x_1 \rightarrow y_1$

Possible Non trivial FD's

× $A \rightarrow B$ ✓ $B \rightarrow A$ ✓ $C \rightarrow A$

× $A \rightarrow C$ × $B \rightarrow C$ × $C \rightarrow B$

× $A \rightarrow BC$ × $B \rightarrow AC$ × $C \rightarrow AB$

× $AB \rightarrow C$

✓ $BC \rightarrow A$

× $AC \rightarrow B$

$\{B \rightarrow A,\; C \rightarrow A,\; BC \rightarrow A\}$ Ans

(II) R

A	B	C
2	3	4
2	3	5
2	4	4
2	4	5
2	3	4

No non Trivial FD's. Ans

Date
Page

Attribute closure (x^+)

x^+ = Set of all attributes that are functionally determined by x.

$\{AB \to CD, AF \to D, DE \to F, C \to G, F \to E, G \to A\}$ // Given FD

$(AB)^+ = \{A, B, C, D, G\}$ $AB \to ABCDG$

$(BG)^+ = \{B, G, A, C, D\}$ $BG \to ABCDG$

$(ABF)^+ = \{A, B, F, C, D, G, E\}$ $ABF \to ABCDEFG$

$(BFG)^+ = \{B, F, G, E, A, C, D\}$ $BFG \to ABCDEFG$

Closure set of F of FD's :- $[F^+]$

$F^+ = \{$ All FD's that can be inferred from F. $\}$

$F = \{A \to B, B \to C\}$ (Given set) $F^+ = \{A \to A \quad A \to B \quad B \to C\}$

$AB \to AB$ etc etc

etc

$F = F^+$

Super Key :-

x is super key of Relation R

iff x^+ must determine all attributes of R

R	A	B	C	D
	a_1			
	a_2			
	a_3			
	a_4			

{A : SK}

$A \rightarrow ABCD$

R(ABCDE)

$\{AB \rightarrow C,\ C \rightarrow D,\ B \rightarrow E\}$.

$(AB)^+ = \{ABCDE\}$ SK ✓✓

$(ABC)^+ = \{ABCDE\}$ SK ✓✓

$(ABCDE)^+ = \{ABCDE\}$ SK ✓✓

$(BD)^+ = \{BDE\}$ ✗✗ SK

40:51 Cand key - {A B} (minimal super key.)

Candidate Key [x] :-

x is candidate key of Relation R

iff ① x must be super key of R

$x^+ = \{\text{all attribute of } R\}$

and ②

No Proper sub set of x is super key of R

$\forall\ y \subset x$ such that $y^+ =$ not all attributes of R.

Not SK.

***** Find candidate keys of the given Relations.**

① R(ABC)

With no non trivial FD's

CK of R : ABC

$(ABC)^+ = \{ABC\}$

② R(ABCD) $\{AB \rightarrow C\}$

CK of R: ABD

$(ABD)^+ = \{ABCD\}$

③ R(ABCDEF) $\{A \rightarrow C,\ E \rightarrow B,\ F \rightarrow E\}$

CK of R : ADF

$(ADF)^+ = \{ABCDEF\}$

④ R(ABCDEF) $\{B \rightarrow F,\ E \rightarrow A,\ D \rightarrow B\}$

CK of R : CDE

$(CDE)^+ = \{ABCDEF\}$

Date
Page

$X \rightarrow \underbrace{y}_{\text{Prime}}$

{ Any non trivial FD $X \rightarrow y$ with y is Prime Attribute exists in R }

$\equiv$ { then there exist More than one ck for relation R }

(5) $R(ABCDEF)$ $\{AB \rightarrow \overset{\text{Prime}}{C},\ C \rightarrow \overset{\text{Prime}}{DE},\ E \rightarrow \overset{\text{Prime}}{F},\ F \rightarrow \overset{\text{Prime Attr}}{A}\}$

$(AB)^+ = \{ABCDEF\}$ $F \rightarrow A$

$\cancel{(BF)^+ = \{ABCDEF\}}$

$(BF)^+ = \{ABCDEF\}$ $(FB/BF) \rightarrow$ other ck

$\{AB, BF\}$

$E \rightarrow F_{\text{prime}}$

$(BE)^+ = \{ABCDEF\}$ $E^+ = EFA$

$C \rightarrow E_{\text{prime}}$

$(BC)^+ = \{ABCDEF\}$ $C^+ = CDEFA$

$AB \rightarrow C$

$\{AB, BF, BE, BC\}$: Cand keys

Date
Page

(4A)

8 R(ABCDEFG)

{AB→CD, AF→D, DE→F, C→G, F→E, G→A}

Ans – 6 candidate keys

Finding CK's of Rel R with n attribute is NPC Problem
[exponential TC Problem]

$TC = O(2^n)$.

Membership test of FD :-

$x \rightarrow y$ Functional dependency is member of FD set (F)
if and only if x^+ must determine y in FD set (F).

Q. Consider given FD set

{AB→CD, AF→D, DE→F, C→G, F→E, G→A}

which FD is not member of given FD set?

a) BEG → CF $(BEG)^+ = \{BEG\ ACDF\}$

b) BG → D $(BG)^+ = \{BG\cancel{A}\ \cancel{C}D\}$

d) AB → E $(AB)^+ = \{ABCDG\}$

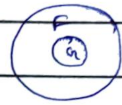
$F^+ \equiv G^+$

$\underbrace{F = \{\ldots\} \qquad G = \{\ldots\}}_{F \equiv G}$?

FD sets F and G are equivalent :-

F & G FD sets are equal if and only if

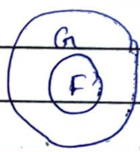
F covers G — (I) every FD of G set must member of F set



$\boxed{F \supseteq G}$

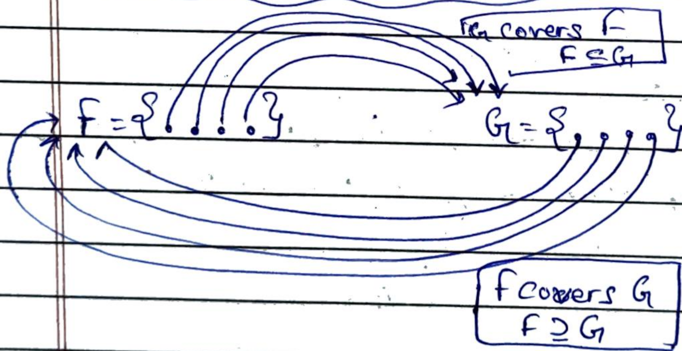
and (II) G covers F :-

Every FD of F set must member of G set



$\boxed{F \subseteq G}$

F & G FD sets equal iff
$F^+ = G^+$



F cover G	G covers F	
True	True	⇒ $F = G$
True	false	⇒ $F \supset G$
false	True	⇒ $F \subset G$
false	False	⇒ F & G not comparable.

Q1 consider given FD sets

$F = \{A \to B, B \to C, C \to A\}$

$G = \{A \to BC, B \to AC, AB \to C, BC \to A\}$

which is ~~true~~ for given FD set?

a) $F \subset G$ b) $F \supset G$ c) $F = G$ d) None.

1:03

F covers $\boxed{G}$

in FD "G"

(i) $A^+ = ABC$ } True

(ii) $B^+ = BCA$

$AB^+ = ABC$

$BC^+ = BCA$

from FD F

G covers $\boxed{F}$

false $C^+ = C$

$C^+ \to A$ of F set is not member of G set.

Canonical cover or minimal cover { Minimal Possible FD's which are equal to given FD set }

Non reducible equal FD set for given FD set.

$F = \{A \to B, B \to C, A \to C, A \to A, AC \to B\}$ // Given FD set.

Trivial FD

$A \to B$ { $AC \to BC$ { $AC \to B$, $AC \to C$

equal

$F_m = \{A \to B, B \to C\}$ // minimal cover.

Date ____
Page ____

Procedure to find minimal cover :-

Assume given FD Set F.

$x \rightarrow y$ — x: Determinant, y: Determines

step 1) Remove extraneous attributes from each determinant of given FD set.

$\{xyz \rightarrow w,\ x \rightarrow y\} \equiv \{xz \rightarrow w,\ x \rightarrow y\}$ — y: extraneous

Extraneous Attribute :-

If $xyz \rightarrow w$, $x^+ = \{\ldots y \ldots\}$ in FD set (F)

Then y extraneous From $xyz \rightarrow w$

$\{xyz \rightarrow w,\ x \rightarrow y\} \equiv \{xz \rightarrow w,\ x \rightarrow y\}$

step 2) Remove redundant FD's from result of step 1.

Redundant FD:- $x \rightarrow y$ FD from FD set (F) is redundant iff $x \rightarrow y$ must be member of $\{F - (x \rightarrow y)\}$ FD set.

$F = \{\ldots \cancel{x \rightarrow y} \ldots\}$ $x^+ = \ldots y \ldots$ → redundant.

ex :- $\{A \rightarrow B,\ B \rightarrow C,\ \cancel{A \rightarrow C}\}$ → redundant.

Q. Find Minimal cover of given FD set?

① $\{$ ✓ $A \rightarrow BC$ — Leave the single determinant

$CD \rightarrow E$ $= D^+ = \ldots C \ldots$

✓ $E \rightarrow C$

✓ $D \rightarrow AEH$

$ABH \rightarrow BD$ $= A^+ = \ldots B \ldots$

$DH \rightarrow BC$ $= \underbrace{D^+ = \ldots H \ldots}_{\text{step 1}}$

$\}$

$\Rightarrow$ Decomposition:

$\{A \rightarrow B$ ✓
$A \rightarrow C$ ✓
$D \rightarrow E$ ✓
$E \rightarrow C$ ✓
$D \rightarrow A$ ✓
~~$D \rightarrow E$~~
$D \rightarrow H$ ✓
$AH \rightarrow B$ ✗
$AH \rightarrow D$ ✓
$D \rightarrow B$ ✗
$D \rightarrow C$ ✗
$\}$

step 2

$\Rightarrow$ Minimal cover

$A \rightarrow BC$
$D \rightarrow AEH$
$E \rightarrow C$
$AH \rightarrow D$

Minimal cover of FD set F May not Unique.

Q. $\{AB \to C, A \to B, B \to A\} \equiv \{A \to C, A \to B, B \to A\} \equiv \left\{\begin{matrix} A \to BC \\ B \to A \end{matrix}\right\}$
$\equiv \{B \to C, A \to B, B \to A\} \equiv \{B \to AC, A \to B\}$

→ Minimal cover of FD set F May not unique.
but all minimal covers are logically equal & also equal to given FD set.

$F_{M_1} = F_{M_2} = F$

4B

Properties of decomposition.

1) Loss less Join decomposition
2) Dependency Preserving decomposition.

Loss less Join decomposition :-

Relational schema R with instance r
decomposed into sub relation, $R_1\ R_2\ R_3 \ldots R_n$

$\boxed{\bowtie : \text{Natural Join}}$

* In General $(R_1 \bowtie R_2 \bowtie \cdots \bowtie R_n) \supseteq R$

* If $R_1 \bowtie R_2 \bowtie \ldots \bowtie R_n = R$ Then,
Loss less Join decomposition

* If $R_1 \bowtie R_2 \bowtie \cdots \bowtie R_n \supset R$
Then Lossy Join decomposition

R
$R_1 \quad R_2$

$(R_1 \bowtie R_2) = R$: LLJ
$(R_1 \bowtie R_2) \supset R$: Not LLJ
× $R_1 \bowtie R_2 \subset R$ ✗
Not possible

R

Sid	Sname	Cid
S_1	A	C_1
S_1	A	C_2
S_2	B	C_2
S_3	B	C_1
S_4	C	C_2

$\{Sid \rightarrow Sname\}$

Cand key : Sid Cid.

1) Decompose R into (LLD)

R_1 (Sid Sname) R_2 (Sid Cid)

R_1

Sid	Sname
S_1	A
S_2	B
S_3	B
S_4	C

R_2

Sid	Cid
S_1	C_1
S_1	C_2
S_2	C_2
S_3	C_1
S_4	C_2

$\Rightarrow$ $(R_1 \bowtie R_2) = R$

Sid	Sname	Cid
S_1	A	C_1
S_1	A	C_2
S_2	B	C_2
S_3	B	C_1
S_4	C	C_2

Loss less Join decomp.

2) Decompose R into

R_1 (Sid Sname) R_2 (Sname Cid)

R_1

Sid	Sname
S_1	A
S_2	B
S_3	B
S_4	C

R_2

Sname	Cid
A	C_1
A	C_2
B	C_2
B	C_1
C	C_2

$\Rightarrow$ $R_1 \bowtie R_2 > R$

Sid	Sname	Cid
S_1	A	C_1
S_1	A	C_2
S_2	B	C_2
S_2	B	C_1
S_3	B	C_2
S_3	B	C_1
S_4	C	C_2

lossy Join decomp.

Spurious tuples [causes inconsisting]

FD set

Loss less decomposition:

→ Relational Schema R with FD set decomposed into R_1, R_2

Decomposition is loss less Join

iff (1) $R_1 \cup R_2 = R$ // every attribute of R relation must be in Sub relation.

and

(2) $(R_1 \cap R_2) \rightarrow R_1$

or

$(R_1 \cap R_2) \rightarrow R_2$

$(R_1 \cap R_2)$ must be super key of R_1 or R_2

Q R(ABCDE) {AB → C, C → D, B → E}

Test given decomposition is LLJ or not ?

(I) {ABC, CD}

$R_1 \cup R_2 \neq R$

Att E not in Sub Rel

∴ Lossy Join.

(II) {ABC, DE}

· $R_1 \cup R_2 = R$

- $R_1 \cap R_2 = \phi$

∴ lossy Join.

(III) {ABC, CDE}

→ $R_1 \cup R_2 = R$

$(R_1 \cap R_2)^+ = C^+ = \{C, D\}$

↓

not SK of R_1 (ABC), R_2 (CDE)

∴ Lossy Join decomp

(IV) {ABCD, BE}

→ $R_1 \cup R_2 = R$ ✓

$(R_1 \cap R_2)^+ = B^+ = \{BE\}$

lossy Join

V) {ABC, ABDE}

$R_1 \cup R_2 = R$ ✓

$(R_1 \cap R_2)^+ = \{ABCDE\} = AB^+$

↓

∴ loss less decomp.

Date
Page

~~4A~~ $R(ABCDEFG)$ $\{AB \to C, BC \to A, B \to D, D \to E, E \to FG\}$

(i) $\{\overset{R_1}{ABC}, \overset{R_2}{ADE}, \overset{R_3}{EFG}\}$

$\to R_1 \cup R_2 \cup R_3 = R$ ✓

$\to R_1(ABC)$ $R_2(ADE)$ $R_3(EFG)$

$\bowtie$ $E^+ = EFG$

$\underbrace{R_1(ABC)\ R_{23}(ADEFG)}_{A^+ = A}$ SK of R_3

Lossy Join decomp.

Dependency Preserving Decomposition :-

Relational Schema R with FD set F decomposed into sub relations R_1 R_2 R_3 ... R_n with FD sets F_1 F_2 F_3 .... F_n respectively.

① In General

$$\{F_1 \cup F_2 \cup F_3 \ldots \cup F_n\}^+ \subseteq F^+$$

② if $\{F_1 \cup F_2 \cup F_3 \ldots \cup F_n\}^+ = F^+$

Then dependency Preserving decomposition.

③ if $\{F_1 \cup F_2 \cup F_3 \ldots \cup F_n\}^+ \subset F^+$

Then Not dependency Preserving decomposition

$R \ldots F = \{\ldots\}$

R_1 $F_1 = \{\}$ R_2 $F_2 = \{\}$

$[F_1 \cup F_2] = F$ DP decomp

$[F_1 \cup F_2] \subset F$ Not DP decomp.

$[F_1 \cup F_2] \supset F$ Not Possible

Date
Page

R(. . . .)   F = { x→y . . . }

// Find FD set of sub relation

R1 (xyz) $F_1 = \{x \to y, xz \to y, yz \to x\}$

$x^+ = \dots y \dots$

$y^+ = y$

$z^+ = z$

$xy^+ = xy$

$xz^+ = \dots y \dots$

$yz^+ = \dots x \dots$

Q Test given decompositions dependency Preserving or not ?

① R(ABCDE) {A→B, B→C, C→D, D→BE}

decomposed into R1 (AB) R2 (BC) R3 (CD) R4 (DE)

~~DP Test~~

R1 (AB)	R2 (BC)	R3 (CD)	R4 (DE)
A→B	B→C	C→D	D→E
	C→B	D→C	
(F1)	(F2)	(F3)	(F4)
A^+ = ABCDE	B^+ = BCDE	C^+ = BCDE	D^+ = BCDE
B^+ = BCDE.	C^+ = CDBE	D^+ = BCDE	E^+ = E

Test all FD's of F are member of subrel?

A→B

B→C

C→D

D→BE

$\therefore \{F_1 \cup F_2 \cup F_3 \cup F_4\} = F$

DP decomposition.

$\boxed{\text{(LLJ \& DP) are independent to each other}}$