Week-5

lec 5.1

Module outline:
- features of good relational design
- Atomic domains & First Normal Form.

Good Relational Design:

→ Reflect real-word structure of problem.
→ Can represent all expected data over time
→ avoid redundant storage
→ Provide efficient storage to data.
→ maintainance of data integrity.
→ Clean, consistant & easy to understand.

Redundancy:

→ repetation of data in multiple copies
- this problem arise when database is not NORMALIZE.
- it leads to anomalies.

Anomaly:

→ Inconsistency ~~aries~~ that can arise due to data changes in a database with insertion, deletion and update.

→ This problem occurs in poorly planned, un-normalised dataset.

* Three (3) kinds of Anomaly:
i) Insertion anomaly
ii) Deletion anomaly
iii) update "

DATE : / /

1) INSERTION Anomaly:

→ When insertion of data is not possible without adding some additional data to record.

eg → we have 3 table, course, student, enrollment.

- we can't add enrollment record for a course which is not present in course table
  मतलब enrollment table course से depend है। जो उसमें होगा वही course इसमें आ सकता है।

2) DELETION Anomaly:

→ When deletion of a data record result in loosing some more info from table.

eg → same as previus.
we can't delete couse table otherwise it will also delete enrollment table.

3) Updation Anomaly:

→ when a data is changed, which may involve many records to be changed,

eg → we can't update enrollment table without updating couse table following observatⁿ:

Redundancy ⟹ Anomaly

- What causes redundancy?
  Dependancy ⟹ redundancy

Good decompositⁿ ⟹ minimizatⁿ of dependency

NORMALIZATⁿ ⟹ Good decompositⁿ ⟹ Minimizatⁿ of dependency ⟹ Redundancy ⟹ Anomoly

DECOMPOSITION :

→ process of breaking down a relation into smaller relatⁿ, based on functional dependency.

FUNCTIONAL dependency:

→ relationship b/w two attributes in a table where the value of one attribute (determinant) determines the value of another attribute (dependent)

eg: in table costumer, Custumer_ID determines the value of name, address, city...

How to make good decompost^n :-

→ not all decompost^n are good.
→ a decompost^n is good if it is lossless & less dependent.

Loosey Decompost^n :- if we can't reconstruct original table from decomposed tables.

→ may be some extra table ~~addege~~ get added during reconstruction.

~~Looless~~
Lossless decompost^n :- If we can reconstruct original table from decomposed tables.

ATOMIC DOMAIN:

→ a domain is atomic if it contains values can't be divided further.

FIRST NORMAL FORM (1NF)

→ No repeating groups (rows)
→ domain of all attribute are atomic.
→ value of each attribute - atomic.

ID	F_name	L_name	phone
123	pooja	Sing	555297, 275432
456	san	Zhang	555403, 182976
789	John	doe	12765432

This is not a 1NF becoz.
→ table phone is composite.

lec 5.2 → Decide whether a particular relation 'R' in "good" form.

→ if Relatⁿ R is not in "good" form, decompose it into set of relatⁿ {R₁, R₂, Rn} such that:
i) each relatⁿ is in good form
ii) decompostⁿ is lossless-join

Theory is based on:-
i) Functional dependency
ii) Multivalued dependency
iii) other dependency

FUNCTIONAL DEPENDENCY.

↳ generalization of notation of Key.

valid

id → name	
1	Ram
2	Sym

valid

id → name	
1	Ram
1	Ram

valid

id → name	
1	Ram
2	Ram

↳ confusion:- is both Ram same or different.

not valid

id → name	
1	Ram
1	Sym

ID: clears the confusion so it is valid

TRIVIAL F.D :→ always TRUE

↳ $X \longrightarrow Y$

↳ if X goes to Y then Y is subset of X.

↳ $X \cap Y$ can't be null.

→ trivial FD don't play role in Normalizatⁿ

NON-TRIVIAL F.D

$X \longrightarrow Y$

↳ if X goes to Y then Y is not subset of X.

↳ $X \cap Y = \phi$ null

Armstrong's Axioms:-

1) Reflixivity: if $Y \subseteq X$ then $X \rightarrow Y$

2) Agumentat^n: if $Y \rightarrow Z$ then

$\cancel{XZ \rightarrow XY}$. $X.Y \rightarrow XZ$

3) Transivity:

$X \rightarrow Y$ and, $Y \rightarrow Z$

then $X \rightarrow Z$.

→ process of generation of FDs terminate after finitely number of steps and we call it Closure set F^+ & $F \subseteq F^+$

These axioms are:-

- Sound: generate only functional dependency that are true in all cases.
- Complete: eventually generate all functional dependencies that are true in all cases.

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

$$F^+ = \{A \rightarrow B, B \rightarrow C, A \rightarrow C\}$$

4) Union:

- $X \rightarrow Y$ & $X \rightarrow Z$ then $X \rightarrow YZ$.

$15 = 3 \times 5$

5) Decomposit^n:

if $X \rightarrow YZ$, then $X \rightarrow Y$ and $X \rightarrow Z$

• Pseudotransitivity :

• if $X \to Y$ & $WY \to Z$ then
$WX \to Z$

lec 5.3

CLOSER METHOD :

→ helps in finding all candicate key in a table.

ex: $R(ABCD) \to$ Relation

F.D $\{A \to B, B \to C, C \to D\}$

Closer of A means what A can determine

$A^+ = BCDA$ — $A \to B$, $B \to C$
$A \to C$ & $C \to D$
$A \to D$

→ Here A can detemine all keys (B, C, D) of relation so A can be candicate key

Similarly
$B^+ = BCD$
$C^+ = CD$
$D^+ = D$
} these can't be candicate key.

$(AB^+) = ABCD$

$(AB)^+$ is determing all keys of relat^n but can't be candicate key (it is minimal)

→ it can be Super key.
(AB^+)

Q R (ABCD) → relatⁿ.

$FD = \{A \to B, B \to C, C \to D, D \to A\}$.

$A^+ = ABCD$
$B^+ = BCDA$ all (A,B,C,D) are candicate key.
$C^+ = CDAB$
$D^+ = DABC$

PRIME ATTRIBUTE: those attributes which are Candicate keys.

→ in above ex. all (A,B,C,D) are Prime attribute.

→ in ~~last~~ 2nd last ex. only (A) is prime attribute.

ex R (ABCDE) – relation

$FD = \{A \to B, BC \to D, E \to C, D \to A\}$

→ Check ~~eig~~ right side में कौन ~~कौन~~ key है।
{B, D, C, A}.
Relatⁿ में {ABCDE}.
E right side में नहीं है।
means:- जो भी canticate key होगा उसके साथ 'E' जुड़ा होगा।
so we'll check {AE, BE, CE, DE} for Candicate key.

$AE^+ = AEBCD$ ✓ yes candicate key

C.K = {AE, DE}

AE is C.K, now check if 'A' or 'E' is present at right side: yes D → A. then replace ~~this~~ 'A' with 'D' in C.K which are use to form C.K

~~C.K = {AE, DE, BDE, CE~~

C.K = {AE, DE, BDE}.

Prime attribute: {A, B, D, E}.

Q. $R = (A, B, C, G, H, I)$ → Relation

$F.D = \{A \to B, A \to C, CG \to H, CG \to I, B \to H\}$

$(AG^+) = AG.$

$AG^+ = ABG.$

$AG^+ = ABCGHI$ → contains all attributes of Relation so AG^+ is C.K.

Q. $R = \{A, B, C, D, E, G\}.$

$F.D = \{A \to BCD,\ BC \to DE,\ B \to D,\ D \to A\}$

i) Compute B^+

$B^+ = BD$

$B^+ = BDACE$ → Contain all attributes so B^+ is C.K.

ii) prove that AG is super key:
↳ hint uniquely identify ever tuple.

AG^+ = if AG^+ covers all (ABCDEG) then AG will be super key.

$AG^+ = AGBCDE$ Yes AG is S.K.

→ left side of FD must be C.K.

2) BCNF : Boyce - Codd Normal Form.

→ a relation is in BCNF if every determinant (left-side) is a superkey or Candidate key of the relation.

Super key: set of attribute that uniquely identifies each attribute

↳ if a relation not in BCNF :
↳ then decompose it into multiple smaller relatⁿ that satisfy BCNF.

→ decompositⁿ should be lossless.

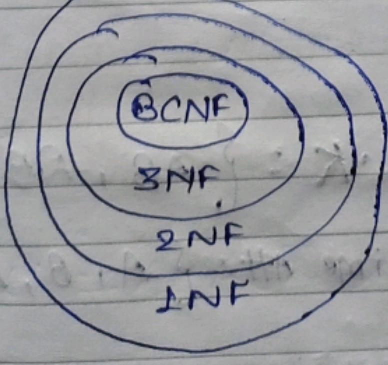
eg R = { A, B, C, D }.

C.K = { A , B }.

FD = { A → C, A → B, B → D, B → A }.

yes it is C.K

left side of FD must be C.K then it is BCNF



3NF :

↳ table must be in 2NF
and
↳ there should not be any transitive dependency in table.

Cond^n of 3NF
→ मतलब left-side of FD should be C.K
OR
right-side of FD should be PRIME ATTRibute.

Qu $R = (ABCD)$
$FD = \begin{pmatrix} AB \to C \\ C \to D \end{pmatrix}$

find C.K
$CK = \{AB\}$
& Prime att = $\{A, B\}$.

let see each FD.

- $AB \to C$
  ↳ it is CK ✓ BCNF ✓

- $C \to D$ ✗ BCNF
  not CK, not Prime att

Qu $R(ABCD)$
$FD: \begin{Bmatrix} AB \to CD \\ D \to A \end{Bmatrix}$

$C.K = \{AB, DB\}$.
Prime Att = $\{A, B, D\}$.

→ $AB \to CD$ ✓ BCNF
↳ CK

→ $D \to A$. ✓ BCNF
↳ Prime Att

DECOMPOSITION :

i) if any relation ~~violates~~ not follows BCNF, decompose it :

ex: $a \longrightarrow \beta$ not follows BCNF.

decompose it :

- $a \cup \beta$
- $R - (\beta - a)$

Lossless Join :

- if we decompose a relat^n R into R_1 & R_2:
  - decompost^n is lossy if $R_1 \bowtie R_2 \neq R$.
  - decompost^n is lossless if $R_1 \bowtie R_2 = R$

- To check lossless join :- $R < \begin{matrix} R_1 \\ R_2 \end{matrix}$

i) union of $R_1 \cup R_2 = R_2$.

ii) intersect^n of $R_1 \cap R_2 \neq \phi$

iii) common attribute must be a key for at least one relat^n

$$R_1 \cap R_2 \longrightarrow R_1 \quad \text{or} \quad R_1 \cap R_2 \longrightarrow R_2$$

Problems with decompost^n :-

i) may be impossible to reconstruct original table

ii) dependency may require join

iii) some queries are more expensive.

lec 5.4

EXTRANEOUS ATTRIBUTE:

i) $\{AB \to C, \; A \to B\}$.

this B is not mandatory here

so 'B' can be excluded from relation,

CONICAL COVER:

Watch gatesmash Lec-30

→ set of functional dependencies that is minimal:

→ Steps to find conical cover :-

i) Remove redundant dependencies
ii) decompose dependencies
iii) Remove Extraneous Attributes:

Q. find correct minimal cover?

$FD: \{A \to B, \; C \to B, \; D \to ABC, \; AC \to D\}$

Step 1: convt. right side into single :

$FD = \{A \to B, \; C \to B, \; D \to A, \; D \to B, \; D \to C, \; AC \to D\}$

ACD → ABE

$A^+ = A$
$C^+ = C$
$D^+ = DBC$,

Step-2 ~~Minimal~~ Set = { A → B, C → B, D → A, D → C, AC → D }.

अगर हम A ~~को~~ A → B को remove करते हैं, तो बाकी बचे FD से A^+ में B नहीं आ रहा, so A → B रहेगा।

Step-3 now try to make left-side single:

{ A → B, C → B, D → A, D → C, AC → D }

already single (left side)

AC → D

→ अगर A^+ में C आ जाएगा तो C को हटा देंगे

$A^+ = AB$ (C नहीं आ रहा तो C नहीं हटा सकते)

→ अगर C^+ में A आ जाएगा तो A को हटा देंगे।

$C^+ = CB$ नहीं है तो A नहीं हटेगा

~~find~~ final minimal set = { A → B, C → B, D → A, D → C, AC → D }.

CK
PK
SK

Q FD: {A → BC, CD → E, E → C, D → AEH,
ABH → BD, DH → BC}.

Check: BCD → H

$BCD^+ = BCDEA\underline{H}$ yes

ii) check: AED → C

$AED^+ = AEDB\underline{C}$ yes.

Q FD: {A → BC, B → E, CD → EF}

i) check: AD → F

$AD^+ = ADBCE\underline{F}$ yes

Q find super Key using:
R(ABCDE)
FD = {AB → C, DE → B, CD → E}

right side {BCE}
Relatⁿ {$\underline{A}$BC$\underline{D}$E}

Q R(EFGHIJKLMN)

FD: {EF → G, F → IJ, EH → KL, K → M
L → N}

✓ a) EFH
✓ b) MNEFH
P c) GHIK
P d) JKLMN
✓ e) EFGH
f) KLMNEFGH

EFH is not present in right side relatⁿ:
So C.K must contain EFH

step:1 find CK.

$(EFH)^+ = EFHGIJKLMN.$

this is CK

step2 So optⁿ check. जिस optⁿ में EFH होगा
वो S.K होगा
because

SK
C.K

Q find minimal cover / canonical cover

a) $AB \rightarrow CD$, $BC \rightarrow D$

st 1 make right side single

$AB \rightarrow C$, $AB \rightarrow D$, $BC \rightarrow D$

minimal set ($AB \rightarrow C$, $AB \rightarrow D$, $BC \rightarrow D$)
~~($AB \rightarrow CD$, $BC \rightarrow D$~~

st2 make left side single
– not possible

this is final minimal set

b) $ABCD \rightarrow E$, $E \rightarrow D$, $AC \rightarrow D$, $A \rightarrow B$

right side is already single.

($ABCD \rightarrow E$, $E \rightarrow D$, $AC \rightarrow D$, $A \rightarrow B$)

right side can't be reduce further.

this is final canonical cover

AQ. 5.4:

1) R (ABCGH)

FD = (A → H, GC → A).

(BCG) not in right side

so BCG must in C.K.

$(BCG)^+$ = BCGAH yes C.K.

only one CK.

2) R (ABCGH)

FD = (AB → C, C → G, B → AH).

B not in right side

B^+ = AHBCG yes C.K

only one C.K. ∴ B^+

no. of S.K = 2^{n-1}

$2^4 = 16$

Q R = (ABC....n) & C.K is "ABC"

No. of S.K = 2^{n-3}

Q R(a, b, c ... n) C.K = 'a, b

two C.K

then S.K = (S.K possible with a) + (SK with b)

− (common S.K from both a & b)

$= 2^{n-1} + 2^{n-1} - 2^{n-2}$

Q $R(a_1, a_2 \dots a_n)$ C.K are "a_1" & "$a_2 a_3$"

$S.K = 2^{n-1} + 2^{n-2} - 2^{n-3}$

4) $F = \{E \rightarrow FI,\ FJ \rightarrow H,\ I \rightarrow J,\ EH \rightarrow G\}$

Can't be remove

~~remove~~ make Right side single

$\left\{ \begin{array}{l} E \rightarrow F,\ FJ \rightarrow H,\ I \rightarrow J,\ EH \rightarrow G \\ E \rightarrow I \end{array} \right\}$

none can't be remove $EH^+ = EH.FIJ\,H.$

now try to make left side single

$E \rightarrow F,\ E \rightarrow I,\ FJ \rightarrow H,\ I \rightarrow J,\ E \rightarrow G$

$E^+ = EFIJH$

E^+ में H आ रहा तो H को remove कर सकते

So H is extraneous

(5) all relat^n in F_1 present in F_2 so F_2 covers F_1

$F_1 \cong F_2$ TRUE

PAGE No.

DATE : / /

Lec 5-5

LOSSLESS JOIN:

to find whether a decompos^n is lossless or n

i) $R_1 \cup R_2 = R$.

ii) $R_1 \cap R_2 \neq \phi$

iii) $R_1 \cap R_2 \rightarrow R_1$ or $R_1 \cap R_2 \rightarrow R_2$.

Q

$R(ABCDE)$

$FD = (AB \rightarrow C, DE \rightarrow B, CD \rightarrow E)$.

find all C.K. & prime attribute.

$AD^+ = AD$

$ADB = ADBCE$ ✓ C.K.

$ADC = ADCEB$ ✓ C.K.

$ADE = ADEBC$ ✓ C.K.

prime attribute = A, B, C, D, E