Database III

Database Design

- The major pitfalls which should be avoided are:-
  (I) Redundancy - It means repetition of the info.
  (II) Incompleteness - It means certain aspects of the enterprise may not be modeled due to difficulty or complexity.

- Types of Anomoly
  (I) Insertion Anomoly - The user can not insert a fact about an entity until he/she has an additional fact about another entity.

  (II) Deletion Anomoly - Deletion of facts about automatically deleted the fact of another entity.

  (III) Update Anomoly - Modification in the value of specific attribute required modification in all records in which that value occurs.

☆ If these Anomolies are not present than we say that our database is free of redundancy and incompleteness.

☆ We can understand & remove anomaly / Incompleteness if we understand how attributes can be dependent on each other -

- Functional
- Fully - Functional
- Transitive
- Multivalued
- Partial

Functional Dependency

It is a constraint betⁿ 2 sets of attributes where one set can uniquely identify another set of attribute uniquely. $X \rightarrow Y$
(called determinate)

ex {name, address} → {cgpa}
{dob} → {age}
→ not need to be a candidate key.

Types ($X \rightarrow Y$)

Trivial
ex: $Y \subseteq X$ i.e. $AB \rightarrow A$

Non-trivial
ex $X \cap Y = \phi$ i.e. $AB \rightarrow C$

Logical Implication

If a fⁿ $X \rightarrow Y$ can be derived from a set of fⁿ dependencies F on R then F is said to logically imply X determined Y.
$F \models X \rightarrow Y$.

or

When ever a tuple or instance 'r' or 'R' satisfied all FD's in F then r also satisfy $X \rightarrow Y$. If $F \models X \rightarrow Y$.

Armstrong's Axiom

- Reflexive rule – $X \rightarrow X$
  or $AB \rightarrow A$ and $AB \rightarrow B$.

- Augmentation rule – $X \rightarrow Y$
  $XZ \rightarrow Y$

- Transitivity rule – $X \rightarrow Y$, $Y \rightarrow Z$
  $X \rightarrow Z$.

- Addition — $X \to Y,\ X \to Z$

  $X \to YZ$

- Composition — $X \to Y,\ W \to Z$

  $XW \to YZ$

- Decomposition — $X \to YZ$

  $X \to Y$ and $X \to Z$

- Psedotransitivity — $X \to Y,\ YZ \to W$

  $XZ \to W$

$R = \{A, B, C, D\}$

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

derived one

$F \models \{A \to C,\ A \to BD,\ A \to CD,\ A \to BCD,\ AB \to CD,$
$AB \to BCD \ldots\}$

$F^{+} = F \cup F'$

i.e. all f. Dependent present in F + whatever which can be derived.

Closure of attributes

$R = \{A, B, C, D\}$

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

X^{+} include which is determined from X.

ex

Closure of $A = \{A, B, C, D\}$

$B^+ = \{B, C\}$

$C^+ = \{C\}$

$D^+ = \{D\}$

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

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

$(AD)^+ = \{A, B, C, D\}$

To Note

There are not candidate key they are super key because A is sufficient to find others.

Use – Finding the key like primary key, candidate key etc.

– Test for the validation of FD.

TO NOTE – $Y \subseteq X^+$ then $X \rightarrow Y$.

Equivalence of FD

- A set of FD F is said to cover an other set of FD G if every FD in G is present in F+ and vise versa.
- That is each FD in G is derived from F and vise versa.

ex $F = \{ - - \}$

$G = \{ - - \}$

F cover G $\therefore G \subset F$

G cover F $\therefore F \subset G$

so, $F = G$

If a set of FD's F cover another set of FD's G and vise versa i.e. $F^+ = G^+$ then $F = G$.

Normalization

- It is the process of decomposing or breaking a relation Scema R into fragments (i.e. smaller schemas) $R_1, R_2, \ldots R_n$ such that the following condition hold:
  1. Min^m redundancy
  2. Min^m anomanies

  i.e. – Lossless decomposition (all inform preserved + Contain same inform from original relation)
  Dependency preservation → All f^c dependencies should be preserved within each fragment R_i.
  Good form. → Each fragment Ri should be free from any type of redundancy.

- **Lossless Decomposition**
  The base relation is said to be lossless if the original relation can be recovered back by joining the fragment relations. (not more not less)

  $R = R_1 \bowtie R_2 \bowtie \ldots \bowtie R_n.$

ex1 R =

A	B
a	1
a	2
b	1
b	2

R1 = R2 =

A	B
a	1
b	2

$R1 \bowtie R2 = R$ Decomposition is lossless.

ex2 R =

A	B
a	1
a	2
b	1

R_1 =

A
a
b

R2 =

B
1
2

$R1 \bowtie R2 \neq R$ ∴ the decomposition is loosy.

☆ When we divide, for performing join we need primary and foreign key pair, that's why we are refering one key as foreign key in other relations.

→ We need to check whether the common column is the key of any relation or not?

- If yes, decomposition is lossless
- if no, decomposition is loosy.

ex1 R(A, B, C, D) with F = {A → B, B → C, C → D} Key = A

is decomposed to -

R_1 (A, B, C) and R2 (C, D) with F1 = {A → B, B → C} Key = A

and F2 = {C → D} Key = C

'C' is common and key of R2. ∴ lossless

→ with F = {A → B, B → C, D → C} Key = A, D

F1 = {A → B, B → C} Key = A & F2 = {D → C} Key = D

'C' is common but not key of any relation.

~~ex~~1 $R(XYZ)$ $F = \{X \rightarrow Y, Y \rightarrow Z\}$

$\downarrow \qquad \searrow$

$R_1(XY) \qquad R_2(YZ)$

How to check if it lossy or not?

Attribute / Relations	X (A1)	Y (A2)	Z (A3)
R1	α_1	α_1	$\beta_{31} \rightarrow$ row no.
R2	β_{12}	α_2	α_2

☆ Our aim is see can be make any one row full of α using given FD.
- If present fill with α otherwise fill with β.

Decomposition is lossless.

↳ Because using $X \rightarrow Y$ and $Y \rightarrow Z$

$\equiv X \rightarrow Z$

	Z (A3)
R1	~~β_{31}~~ α_1

~~ex~~2 $R(ABC)$, $F = \{A \rightarrow B, C \rightarrow B\}$

$R_1(AB)$, $R_2(BC)$

Attributes / relations	A (A1)	B (A2)	C (A3)
R1	α_1	α_1	β_{31}
R2	β_{12}	α_2	α_2

No, its lossy decomposition.

△ Key attribute – Attributes which are part of some Candidate Key.

△ Non Key attributes – Attributes which are not part of any candidate Key.

Dependency Preservation

The decomposition of a relation schema R with FD F into R_i with FDs F_i: All fe dependencies should be preserved within each fragment R_i.

i.e. $(\cup_i F_i)^+ = F^+$ i.e. $(F_1 \cup F_2 \ldots \cup F_n)^+ = F^+$

ex1

$R = (A, B, C)$ and $F = \{A \rightarrow B, B \rightarrow C, A \rightarrow C\}$

$R1 = (A, B)$ and $F1 = \{A \rightarrow B\}$

$R2 = (B, C)$ and $F2 = \{B \rightarrow C\}$

but $A \rightarrow C$ is not preserved.

- Normal forms – Provide stepwise progression towards the construction of normalized relation schemas, which are free from data redundancies.

Types –

① 1NF (First Normal form)

It states that the domain of an attribute must include only atomic (simple indivisible) values [or that is value of any attribute in a tuple must be a single value from the domain of that attribute.]

- To convert non-1NF relation into a 1NF relation, split up the non-atomic.
- ER relation is always 1NF.

ex

roll	address
01	ABC1 XYZ1
02	ABC2 XYZ2
...	

$\Longrightarrow$ converting

roll	city	state
01	ABC1	XYZ1
02	ABC2	XYZ2
...		

not atomic
∴ not in 1NF

atomic
∴ now in 1NF

(2) 2NF (Second Normal Form)

Condition –

(i) Every non-prime attribute 'A' in R is full functionality dependent on the primary key of R.
(partially is not allowed)

(ii) It is in 1NF.

Explanation –

ex1 Let relation R(A,B,C,D) with PK → AB.
∴ prime attributes are A, B
Non-prime attributes are C, D.

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

C is fully dependent
D is partially dependent ∴ not in 2NF

For converting in 2NF

☆ We need to seperate (i) satisfying one
(ii) Non-satisfying one

$R(A,B,C,D)$

$R1(A,B,D)$ $\quad$ $R2(A,B,C)$ Key AB

$A \rightarrow D$ $\quad$ $AB \rightarrow C$

$B \rightarrow D$

$R_{11}(AD)$ $\quad$ $R_{12}(BD)$

$A \rightarrow D$ $\quad$ $B \rightarrow D$

Key A $\quad$ Key B

Final relation set is $R2, R_{11}, R_{12}$.

ex2 $R(A,B,C,D,E)$

$F = \{AB \rightarrow C, B \rightarrow D, C \rightarrow E\}$ $\quad$ N.P.A = C, D, E.

Key = AB $\quad$ → FD $\quad$ → PD

$R(A,B,C,D,E)$

$R1(A,B,C,E)$ $\quad$ $R2(BD)$

$AB \rightarrow C$ $\quad$ $B \rightarrow D$

$C \rightarrow E$

Key AB $\quad$ Key B

- When Key consist of only one attribute it will always in 2NF.

③ 3NF (Third Normal Form)

When a relation is already in 2NF and there is no non-Prime attribute which is transitively dependent on the primary Key.

ex1 R(ABCDE)

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

Key = A

∴ Condition 1 — ✓

∴ Condition 2 — non prime attribute : B, C, D, E

$A \rightarrow B$ directly dependent ✓

$B \rightarrow C$ transitivily dependent ✗ not in 3NF

$A \rightarrow D$ d.d ✓

$B \rightarrow E$ t.d ✗ not in 3NF

R_1 (ABD)	R_2 (BCE)
$A \rightarrow B$	$B \rightarrow C$
$A \rightarrow D$	$B \rightarrow E$
in 3NF	in 3NF
Key A	Key B
✓	✓

★ $X \rightarrow A$ will satisfy 3NF if X is a super Key or A is a prime attribute.

ex2 R(ABCD) F = {AB → C, C → D}

Key = AB

∴ condition 1 - ✓

∴ Condition 2 - non prime attribute C, D

AB → C d.d ✓

C → D t.d ✗

R1 (ABC)

now in 3NF

Key AB

AB → C

R2 (CD)

now in 3NF

Key C.

C → D

ex3 R(ABCDE)

F = {AB → C, B → D, C → E}

Key = AB

∴ Condition 1 - ✗

R(ABCDE)

R1(ABCE)

AB → C

C → E

Key = AB

R2(BD)

B → D

Key = B

∴ Condition 2 (I) Key = AB

non-prime attribute - C, E

AB → C d.d ✓

C → E t.d ✗

R11 (ABC)

R12 (CE)

(II) Key = B

B → D d.d ✓

R2 (BD)

4) Boyce Codd NF (BCNF)

→ Also Known as 3.5NF because it is slightly stricter than ($ stronger) 3NF.

A relation is said to be in BCNF if

(i) Already in 3NF

(ii) For each FD $X \to A$, X is a super Key.

★ Only 1 condition as that of 3NF

ex

$R(A, B, C, D, E)$

$F = \{AB \to CE,\ C \to D,\ C \to A\}$

Key = AB

non-prime attribute - C, D, E

1st we have checked for 2NF

Condition (1) $AB \to CE$ d.d ✓

$C \to D$ t.d X

$C \to A$ ✓

$R_1(ABCE)$ — Key AB

$AB \to CE$ ✓

$C \to A$ X

$R2(CD)$ — Key C

$C \to D$ ✓

$R_{11}(ABCE)$ — Key AB

$AB \to CE$

$R_{12}(CA)$ — Key C

$C \to A$

∴ Relations are R_{11}, R_{12} $ $R2$.

PAGE NO.:
DATE: / /

Now Advanced NF

(5) Forth normal form (4NF)

A relation is said to be in 4NF when it is already in 3NF and there is not more than one MVD (Multi valued dependency) associated to one attribute set present in the relation.

Multivalued Dependency (MVD) ($\twoheadrightarrow$)

When for one value of an attribute there are multi values exist for another attribute, the dependency (relation) bet^n them is referred as MVD.

ex

Book	Author
DBMS	A
DBMS	B
DBMS	C

∴ Book $\twoheadrightarrow$ Author

ex

Book	Author
DBMS	A
DBMS	B
DBMS	C
DBMS	D

Book	Publisher
DBMS	P_1
DBMS	P_2
DBMS	P_3
DBMS	P_4

We can put them in 1 table but seperatly it will take - 4+4 rows and together in one table it will take - 4×4 rows.

∴ Book $\twoheadrightarrow$ Author
Book $\twoheadrightarrow$ Publisher

★ For violating 4NF rule you atleast need 3 attribute out of which 2 are MVD. on 3rd one.

⑥ Fifth NF (5NF) — Also called PJNF (project join NF)

Conditions :

(i) Must be in 4NF

(ii) There is no join dependency and decomposition is lossless.

★ This is just to check, no further decomposition. In check if till whatever we have decomposed must be join independent & lossless

R

/ | \

R1 R2 ... Rn

$R_1 \bowtie R_2 \ldots R_n = R$

Join dependency means when the relation are needed to be joined in a particular order otherwise it may not give the same result everytime.

1805182
Jayushi

Assignment

1. Consider a relation schema R(A,B, C,D,E,F,G) with F = {AB → DEC, E → FG, C → BE, F → A}

a) List all candidate keys of R.

A = {A}
B = {B}
C = {A B C E F G} ✓
D = {D}
E = {A E F G}
F = {A F}
G = {G}
AB = {A B C D E F G} ✓
AC = {A, B, C, D, E, F, G} but it is super key.
AD = {A D}
AE = {A E F G}
AF = {A F}
AG = {A G}
...

∴ Candidate keys are C and AB.
So we will take C as primary key of table R.
non-prime attribute ABDEFG

b) Convert the relation R(ABCDEFG) into BCNF.

1NF ✓

2NF ✓ as my primary key consist of only one attribute.

1805182
Ayushi

3NF –

AB → DEC ✓
E → FG
C → BE ✓
F → A

R1 (ABCDE) R2 (AEFG)
Key C Key E
AB → DEC ✓ E → FG ✓
C → BE ✓ F → A ✗

R21 (EFG) R22 (FA)
Key E Key F
E → FG F → A

3.5NF

AB → DEC ✗
C → BE ✓
~~need to decompose~~

E → FG ✓

F → A ✓

R11 (ABCDE) R12 (CBE)
AB → DEC C → EB
Key AB Key C

∴ The decomposed relation is as follows
R11 (ABCDE), R12 (CBE), R21 (EFG), R22 (FA).

1805182
Aayushi

2. Convert the relation R(ABCDEG) with
$F = \{AB \to C, AC \to E, EG \to D, AB \to G, D \to E\}$
Consider the decomposition R1(DEG) and R2(ABCD)
a) Check whether the decomposition is lossless or not.

R(ABCDEG)
Key = AB

R1(DEG)
Key EG

R2(ABCD)
Key ABD.

relation \ Attribute	A(A1)	B(A2)	C(A3)	D(A4)	E(A5)	G(A6)
R1	β_{11}	β_{21}	β_{31}	$\alpha 1$	$\alpha 1$	$\alpha 1$
R2	$\alpha 2$	$\alpha 2$	$\alpha 2$	$\alpha 2$	~~β_{52}~~	~~β_{62}~~
					$\alpha 2$	$\alpha 2$ using $AB \to G$

using $AC \to E$

∴ The decomposition is lossless.

b) Check whether the decomposition is dependency preserving or not.

R1(DEG) FD = $\{EG \to D, D \to E\}$
R2(ABCD) FD = $\{AB \to C\}$

$\cup_i F_i = \{EG \to D, D \to E, AB \to C\}$
but we can see we can't preserve $AC \to E$ and $AB \to G$.
∴ no decomposition is not dependency preserving.

3a) Which normal form is generally considered adequate for any relation as the decomposition is always lossless and dependency preserving?

Upto 3NF the decomposition is always lossless and dependency preserving.

b) Consider R(ABCDEF) with FD set F = { AB → CDEF, C → A, D → B, C → D, E → F, B → E }. The above relation is in which normal form.

R(ABCDEF)

Key = C

1NF - ✓

2NF - ✓ as our primary key consist of only one attribute.

3NF - AB → CDEF ✓

C → A ✓

D → B ✗

C → D ✓

E → F ✗

B → E ✗

∴ The above relation is in 2NF.

1805182

Tayushi

4a) R(ABCDEFGH)

F = {AB → C, A → DE, B → F, F → GH}

Check for BCNF if not then convert it into BCNF.

R(ABCDEFGH)

Key = AB

1NF ✗

2NF
AB → C
A → DE ✗
B → F ✗
F → GH

R1 (ABCFGH) Key ABF
AB → C ✗
F → GH ✗

$\boxed{R_{11}\ (ABC)}$ Key AB
AB → C

$\boxed{R_{12}\ (FGH)}$ Key F
F → GH

R2 (ABDEF) Key AB
A → DE ✗
B → F ✗

$\boxed{R_{21}\ (ADE)}$
A → DE
Key A

$\boxed{R_{22}\ (BF)}$
B → F
Key B

★ relations are R11, R12, R21, R22.

3NF
(i) R11 (ABC) Key AB
AB → C ✓

(ii) R12 (FGH) Key F
F → GH ✓

(iii) R21 (ADE) Key A
A → DE ✓

(iv) R22 (BF) Key B
B → F ✓

1805182
Aayushi.

It is in 3NF and also in BCNF.

So relations are R11, R12, R21, R22.

b) R(ABCDE)

F = {CE → D , D → B, C → A}

Check for 3NF and if not convert it into 3NF.

1. 1NF ✓

2. 2NF Key = CE

CE → D	✓
D → B	✓
C → A	X

R1 (BCDE)

CE → D ✓

D → B ✓

Key CE

R2 (CA)

C → A ✓

Key C

3. 3NF R1 (BCDE)

CE → D ✓

D → B X

R2 (CA)

C → A ✓

Key C

R11 (CDE)

CE → D

Key CE

R12 (BD)

D → B

Key D

After decomposition R11, R12, R2.

1805182

Ayushi

4c) R(ABCDEF) $F = \{AB \rightarrow C, CD \rightarrow AE, E \rightarrow F\}$

Check for 3NF and if not then convert into 3NF.

Candidate key = ABD, CDB.

R(ABCDEF)

Key = ABD

1NF ✓

2NF $AB \rightarrow C$ ✗

$CD \rightarrow AE$ ✗

$E \rightarrow F$ ✓

R1(ABC)	R2(ACDE)	R3(EF)
$AB \rightarrow C$	$CD \rightarrow AE$	$E \rightarrow F$
Key AB	Key CD	Key E

3NF $\boxed{R1(ABC)}$ $\boxed{R2(ACDE)}$ $\boxed{R3(EF)}$

$AB \rightarrow C$ ✓ $CD \rightarrow AE$ ✓ $E \rightarrow F$ ✓

So the final decomposition is R1, R2, R3.

18US182
Aayushi.

☆ TO NOTE

If primary key = ABD

$CD \rightarrow AE$

We will see it as non-prime attribute E is partly dependent on the primary key.

5) Consider relational schema R = {EFGHIJKLMN}
and FD = { EF → G, F → IJ, EH → KL, K → M
L → N }
Find the Key of R?

EF → G
F → IJ
EH → KL
K → M
L → N

E = { E }
F = { F I J } — (I)
G = { G }
H = { H }
I = { I }
J = { J }
K = { K M }
L = { L N }
M = { M }
N = { N }

EF = { E, F, G, I, J } (need HKLMN)
EH = { E H K L M N } (need FGIJ) — (II)

if we combine (I) and (II) we have

EHF = { E F G H I J K L M N }

∴ The Key of relation R is EHF.

1805182
Aayushi.

1805182

Jayashi

6.

X	Y	Z
1	4	2
1	5	3
1	6	3
3	2	2

Which of the f^n dependencies are satisfied by the instance?

a) $XY \to Z$ and $Z \to Y$

↓

not satisfied

as for value 3 of Z we can't able to uniquely identify Y.

3 → 5, 6

b) $YZ \to X$ and $Y \to Z$

↓

Yes as we can uniquely identify Z value with Y (all unique)

↓

Yes as all YZ pair are unique.

c) $XZ \to Y$ and $Y \to X$

↓

not unique because we have 2

X	Z
1	3

tuple.

d) $YZ \to X$ and $X \to Z$

↓

not unique values in 1

1 → 2, 3

180518
Aayushi

7. R(ABCDE)

$F = \{A \to B, A \to C, CD \to E, B \to D, E \to A\}$

Which FD is not implied by the above set.

a) $CD \to AC$

b) $BD \to CD$

c) $BC \to CD$

d) $AC \to BC$

8. Find the minimal cover for

$F = \{A \to C, AB \to C, C \to DI, CD \to I, EC \to AB, EI \to C\}$.

(I) $A \to C$

(II) $AB \to C$ ✗

(III) $C \to DI$

(IV) $CD \to I$ ✗

(V) $EC \to AB$

(VI) $EI \to C$

★ NO trivial dependency is allowed.

(II) $AB \to C$ no need of this as $A \to C$ (I)

(IV) $CD \to I$ no need of this as $C \to DI$ both (III)

Minimal cover of $F = \{A \to C, C \to DI, EC \to AB, EI \to C\}$

9. Repeat question 8 with

$F = \{AB \to C, C \to A, BC \to D, ACD \to B, D \to E, D \to G,$
$BE \to C, CG \to B, CG \to D, CE \to A; CE \to G\}$

(i) $AB \to C$

(ii) $C \to A$

(iii) $BC \to D$

(iv) $ACD \to B$. (X) $D \to G$ and $CG \to B$

(v) $D \to E$

(vi) $D \to G$

(vii) $BE \to C$

(viii) $CG \to B$.

(ix) $CG \to D$ (X) $CG \to B$ and $BC \to D$

(x) $CE \to A$ (X) as $C \to A$ (ii)

(xi) $CE \to G$

1805182
Aayushi.

$F = \{ AB \to C, C \to A, BC \to D, D \to E, D \to G, BE \to C, CG \to B, CE \to G \}$

10. R(ABCDE), $F = \{ A \to B, B \to C, C \to D, C \to E \}$
Check the above relation is in which normal form if it given that there is multiple values of D and E for single value of C. Convert the relation into 4NF.

R(ABCDE)
Key = A

1NF ✓

2NF ✓ as primary key consist of only one attribute.

3NF
- $A \to B$ ✓
- $B \to C$ X
- $C \to D$ X
- $C \to E$ X

R1 (AB) R2 (BC) R3 (CDE)

$A \to B$ ✓ $B \to C$ ✓ $C \to D$ ✓

Key A Key B $C \to E$

Key C

4NF $\boxed{R1 (AB)}$ $\boxed{R2 (BC)}$ R3 (CDE)

$A \to B$ ✓ $B \to C$ ✓ $C \to D$ X

$C \to E$ X

$\boxed{R31 (CD)}$ $\boxed{R32 (CE)}$

$C \to D$ ✓ $C \to E$ ✓

∴ Final relations are R1, R2, R31, R32.

1805182
Aayushi