COMPUTER SCIENCE

Database Management System

FD's & Normalization

Lecture_15

Vijay Agarwal sir

01 Normal Forms

02 Normal Form Decomposition

Normal Form

1NF : No Multivalued Attribute

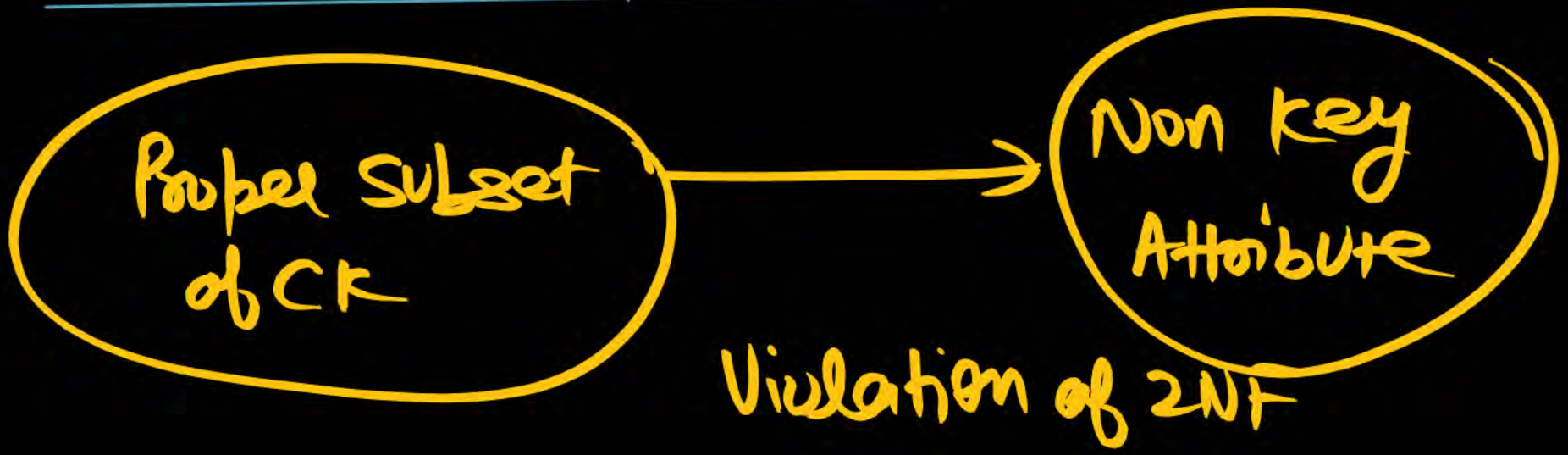
2NF :

3NF : X: super key (or) Y: key/Prime Attribute

BCNF → X: Super key.

Normal Forms

2NF Decomposition



Normal Forms

3NF Decomposition

$X \rightarrow y$

Every Non Trivial FD

X : Super key

OR

y : key / Prime Attribute.

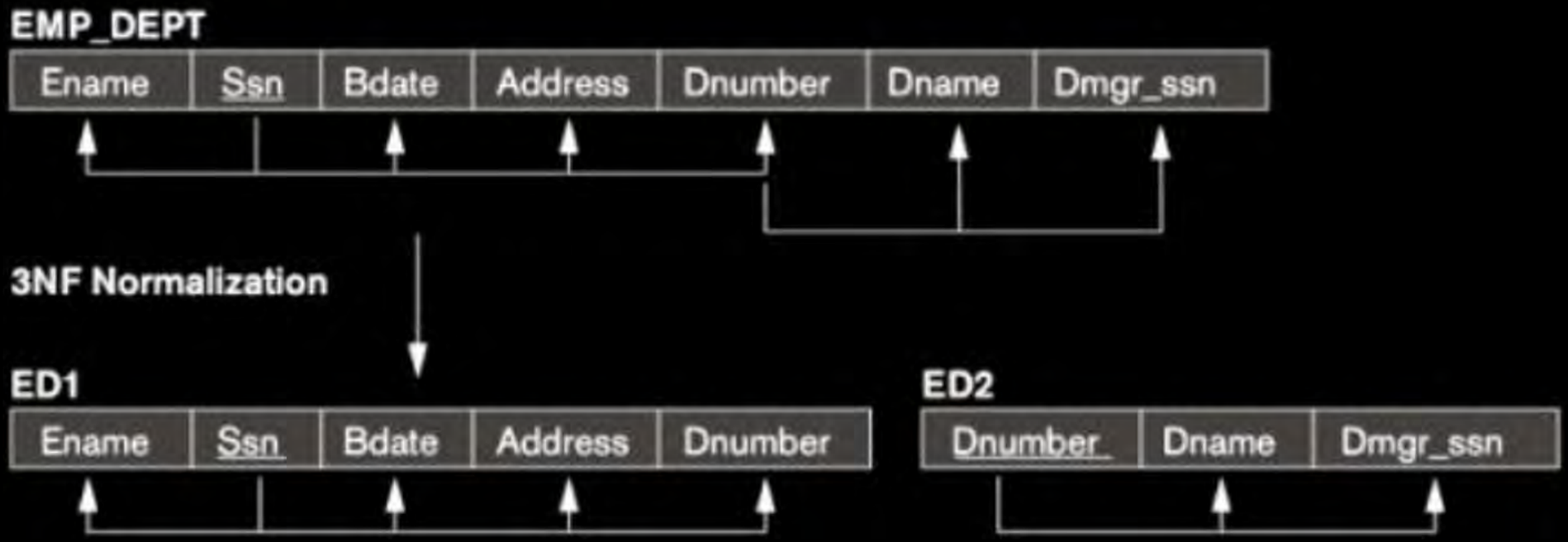
Normal Forms

Third Normal Form

Definition: According to Codd's original definition, a relation schema R is in 3NF if it satisfies 2NF and no nonprime attribute of R is transitively dependent on the primary key.

OR

Definition: A relation schema R is in third normal form (3NF) if, whenever a nontrivial functional dependency $X \rightarrow A$ holds in R either (a) X is a superkey of R, or (b) A is a prime attribute of R.

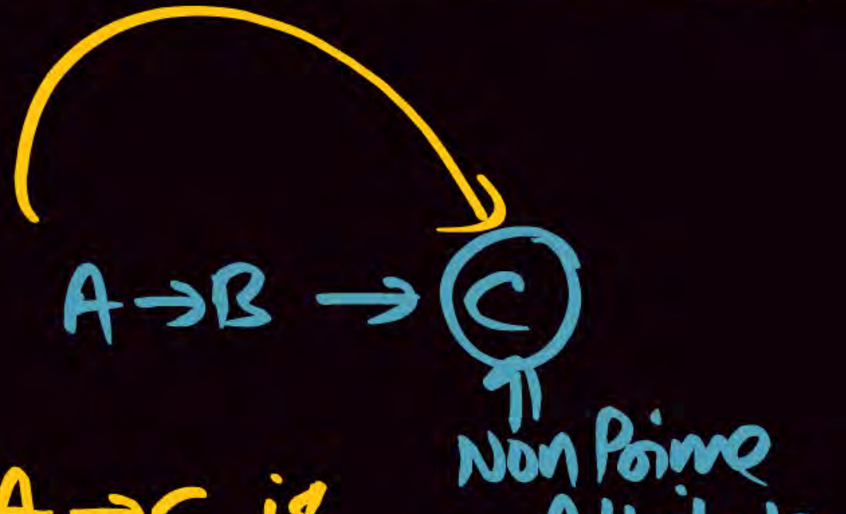

EMP_DEPT
Ename
Ssn
Bdate
Address
Dnumber
Dname
Dmgr_ssn
3NF Normalization
ED1
Ename
Ssn
Bdate
Address
Dnumber
ED2
Dnumber
Dname
Dmgr_ssn

R(ABC) [A→B, B→C]

Candidate key = [A]

Non key Attribute = [B, C]

B→C : N.K, N.K



CHECK 3NF ?

A→B ✓3NF A: Super key

B→C X 3NF [B is Not Super key (OR) Not in 3NF C is Not Prime Attribute]

A→C is Transitive FD.

Not in 3NF

(Non key Attribute transitively determine by Primary key)

3NF Decomposition

Q.1 R(ABC) $[A \rightarrow B, B \rightarrow C]$

3NF Decomposition

3NF + Lossless + D.P.

3NF Decomposition

Q.4 R (ABCD) {AB→CD, D→A}

Candidate key = [AB, DB]

CHECK 2NF ?

AB→CD ✓2NF

D→A ✓2NF

D → key Attribute

R is in 2NF

CHECK 3NF ?

AB→CD; AB is Super key.

D→A; D is Not Super key

But

A is Prime/key Attribute

So R is in 3NF

CHECK BCNF ?

AB→CD ✓BCNF

D→A; ✗BCNF fail

D is Not Super key

Not in BCNF.

3NF Decomposition

Q.5 R (ABCDEFGH) {A→BC, B→DEF, DE→AGH}

Candidate key = [A, DE, B]

CHECK 3NF ?

A→BC ✓ 3NF

B→DEF ✓ 3NF

DE→AGH ✓ 3NF

X: Super key

R is in 3NF

CHECK BCNF ?

A→BC

B→DEF

DE→AGH

} BCNF (∵ X is super key)

BCNF

Q.6 R(ABCDE) $\{AB \to C, C \to D, B \to E\}$

Decompose into 2NF, 3NF, BCNF

Candidate key = [AB]

Non key Attribute = [C, D, E]

CHECK 2NF ?

$B \to E$ ← Non key Attribute

↑ Proper Subset of CK

Not in 2NF

2NF Decomposition

R(ABCDE)

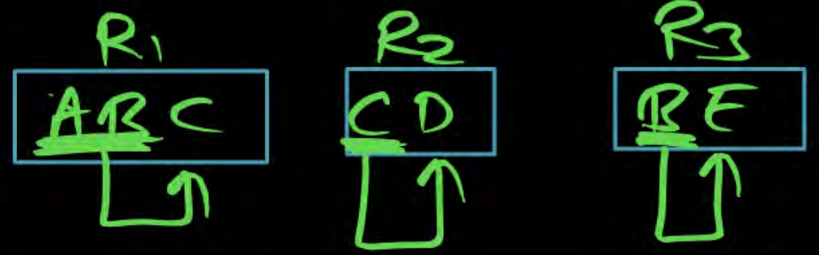
$(B)^+ = [BE]$

Now R is in 2NF

CHECK 3NF ?

2NF ✓ & $C \to D$ Violation of 3NF

3NF Decomposition

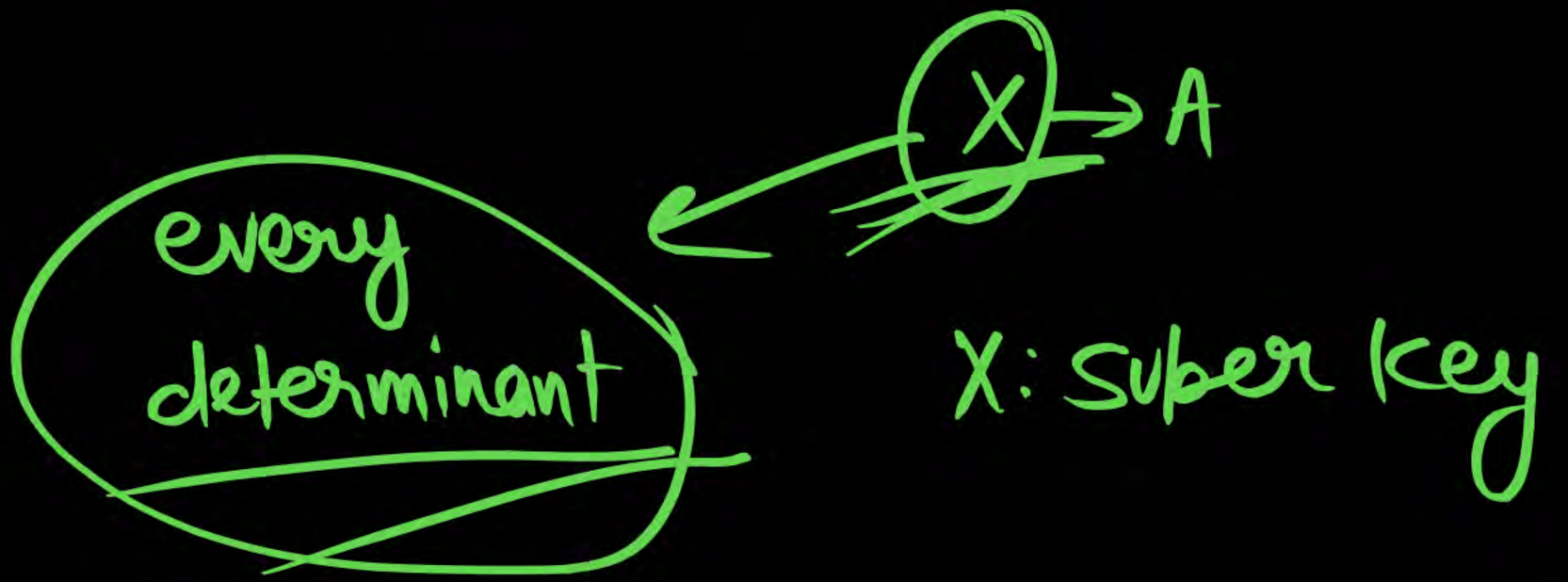


3NF + Lossless Join + Dependency Preserving

R is in BCNF

Boyce – Codd Normal Form

Definition: A relation schema R is in BCNF if whenever a nontrivial functional dependency X → A holds in R, then X is a superkey of R.



BCNF Decomposition

Q.1 R(ABCDE) $\{A \to B, B \to C, C \to D, D \to E\}$

Candidate key = [A]

Non key Attribute = [B, C, D, E]

CHECK 2NF ?

R is in 2NF

Proper subset of CK → Non key Attribute

CHECK 3NF ?

$X \to Y$

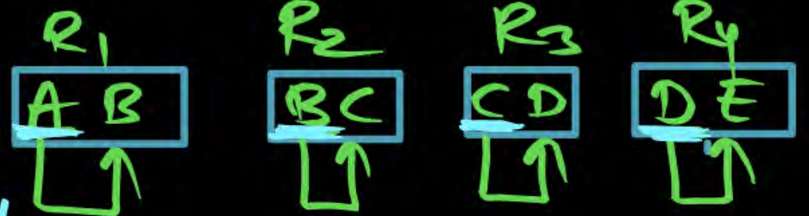
$B \to C$ X: Not Super key

$C \to D$

$D \to E$ Y: Not key Attribute

Not in 3NF

3NF Decomposition



3NF + Dep. Preserved + Lossless Join.

R is in BCNF ✓ [X: super key]

3NF ✓ BCNF ✓/✗

R(ABCD) [AB→CD , D→A]

Candidate key = [AB , DB]

D → A ; D is Not Super key
But
A is key / Prime Attribute

} 3NF But Not in BCNF.

BCNF Decomposition

Q.1 R(ABCDE) $\{A \to B, B \to C, C \to D, D \to E\}$

Candidate key = [A]

Non key Attribute = [B, C, D, E]

BCNF Checking ?

$A \to B$

$B \to C$

$C \to D$

$D \to E$

X is Not Superkey.

BCNF Violation

Not in BCNF.

BCNF Decomposition

$R_1(A\,B\,C\,D\,E)$

① $B \to C$

R_2 [B C]

× $C \to D$

② $D \to E$

R_3 [D E]

[A B D] [B C] [D E]

BCNF Decomposition

Q.1 R(ABCDE) $\{A \to B, B \to C, C \to D, D \to E\}$

Candidate key = [A]

Non key Attribute = [B, C, D, E]

BCNF Checking ?

$A \to B$

$B \to C$

$C \to D$

$D \to E$

X is Not Superkey.

BCNF Violation

Not in BCNF.

BCNF Decomposition.

R(ABCDE)

② $B \to C$

① $C \to D$

× $D \to E$

R_1 (ABE)

R_2 (BC)

R_3 (CD)

BCNF Decomposition

Q.1 R(ABCDE) $\{A \to B, B \to C, C \to D, D \to E\}$

Candidate key = [A]

Non key Attribute = [B, C, D, E]

BCNF Checking ?

$A \to B$

$B \to C$

$C \to D$

$D \to E$

X is Not Superkey.

BCNF Violation

Not in BCNF.

BCNF Decomposition.

R(ABCDE)

③ $B \to C$ R_2 BC

② $C \to D$ R_3 CD

① $D \to E$ R_4 DE

AB

BCNF Decomposition

Q.2 R(ABCD) $\{AB \rightarrow CD, D \rightarrow A\}$

Candidate key = (AB, DB)

CHECK BCNF ?

$D \rightarrow A$ fail BCNF

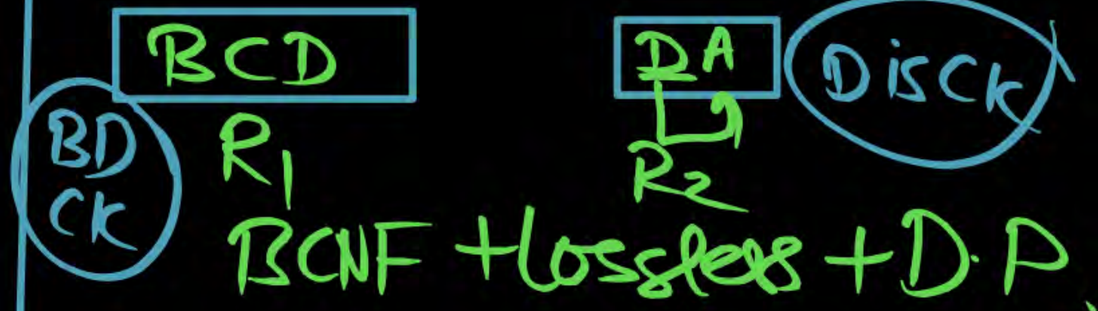
∵ D is Not Subset key

Not in BCNF

BCNF Decomposition

R(ABCD)

$D \rightarrow A$



BCNF + lossless + D.P.

$R_1(BCD) \cap R_2(DA) = D$

$(D)^+ = [DA]$ Super key of R_2

Lossless Join

BCNF Decomposition

Q.3 R(ABCDE) $\{A \to B, BC \to D, D \to E\}$

Candidate key = [AC]

CHECK BCNF ?

$X \to Y$ X is Not Super key

$A \to B$

$BC \to D$ So R Not in BCNF

$D \to E$

BCNF Decomposition

$R_1(ABCDE)$

① $A \to B$

✗ $BC \to D$

② $D \to E$

R_2 AB

R_3 DE

R_1 ACD

R_2 AB

R_3 DE

BCNF Decomposition

$R_1(ABCDE)$

② $A \to B$

① $BC \to D$

✗ $D \to E$

R_2 BCD

R_3 AB

R_1 ACE

R_2 BCD

R_3 AB

BCNF Decomposition

$R_1(ABCDE)$

③ $A \to B$

② $BC \to D$

① $D \to E$

R_2 DE

R_3 BCD

R_4 AB

$R_1(AC)$
$R_2(DE)$
$R_3(BCD)$
$R_4(AB)$

BCNF Decomposition

Q.4 $R(ABCDEFG) \quad \{A \rightarrow BF, F \rightarrow DEG, A \rightarrow D\}$

Note

In BCNF Dependency May/may Not be Preserved.

→ Lossless Join Guaranteed But

Note

Till 3NF Lossless Join & Dependency Preserving Must be Satisfied.

R(ABCD) [AB → CD , D → A]

C.k : [AB. DB)

3NF But Not in BCNF.

Q.5 $R(ABCDEFGHIJ)$ $\{AB \rightarrow C, A \rightarrow DE, B \rightarrow F, F \rightarrow GH, D \rightarrow IJ\}$

Candidate key = [AB]

Non key Attribute = [C, D, E, F, G, H, IJ]

CHECK 2NF ?

$A \rightarrow DE$, $B \rightarrow F$ } Fail 2NF

2NF Decomposition

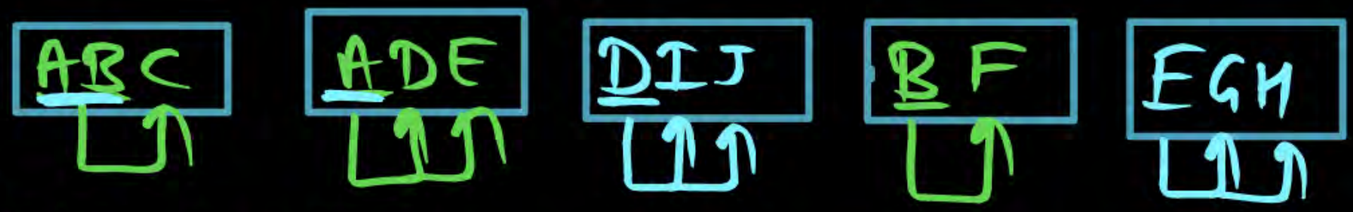
$(A)^+ = (ADEIJ)$

$(B)^+ = (BFGH)$

CHECK 3NF ?

$D \rightarrow IJ$, $F \rightarrow GH$ } Fail 3NF

3NF Decomposition



3NF + BCNF + Lossless Join + Dep. Preserved

Actual Implementation of Normal Form.

R(ABC) [A→B, B→C)

A	B	C
1	b1	x
2	b1	x
3	b1	x
4	b1	x
5	b1	x
6	c1	y
7	c1	y
8	c1	y
9	c1	y
10	c1	y

R1(AB)

A	B
1	b1
2	b1
3	b1
4	b1
5	b1
6	c1
7	c1
8	c1
9	c1
10	c1

R2(BC)

B	C
b1	x
c1	y

Q.5

Roll Name CGPA Daddy

$R(ABCDEFGHIJ) \quad \{AB \rightarrow C, A \rightarrow DE, B \rightarrow F, F \rightarrow GH, D \rightarrow IJ\}$

ABC ADE BF FGH DIJ

STUDENT DB

RollNo Name CGPA Branch FatherName Semester F.Contact F.Address Area City PinCode

{8 Semester}

Unique & 1 2 3 & Semester CGPA of each semester Different

F.Name Branch DOB M.Name 10th 12th , Pin, City, Area State Contact , Address

Arity : 50

Assume

50 Attribute

⇓

40 Attribute Repeat

For each Student

eg

Like FN, MN, 10th, 12th Area, City, State

University

1 Lakh Student

1 X 280 = 280 Lakh

⇓

2.8 Crore

7 times unnecessary Repeat.

for 1 Student = 40 X 7 = 280 entries Repeat

Q

Relation R is decomposed using a set of functional dependencies, F, and relation S is decomposed using another set of functional dependencies, G. One decomposition is definitely BCNF, the other is definitely 3NF, but it is not know which is which. To make a guaranteed identification, which one of the following tests should be used on the decompositions? (Assume that the closures of F and G are available).

[2002: 2 Marks]

3NF BCNF

A Dependency-preservation

B Lossless-join

C BCNF definition

D 3 NF definition

WHY Normalization ?

Insertion Anomalies

Updation Anomalies

Deletion Anomalies.

Q

Which of the following relational schema with given FD's follows is/are in BCNF?

A. R(ABCDE) and FD's are $\{A \rightarrow B, B \rightarrow C, C \rightarrow D, D \rightarrow E, C \rightarrow A\}$

B. R(ABCDE) and FD's are $\{A \rightarrow B, C \rightarrow D, D \rightarrow E\}$

C. R(ABCD) and FD's are $\{A \rightarrow B, B \rightarrow C, C \rightarrow D, D \rightarrow A\}$

D. R(ABCD) and FD's are $\{A \rightarrow B, B \rightarrow C, C \rightarrow A\}$

Single Valued Functional Dependency

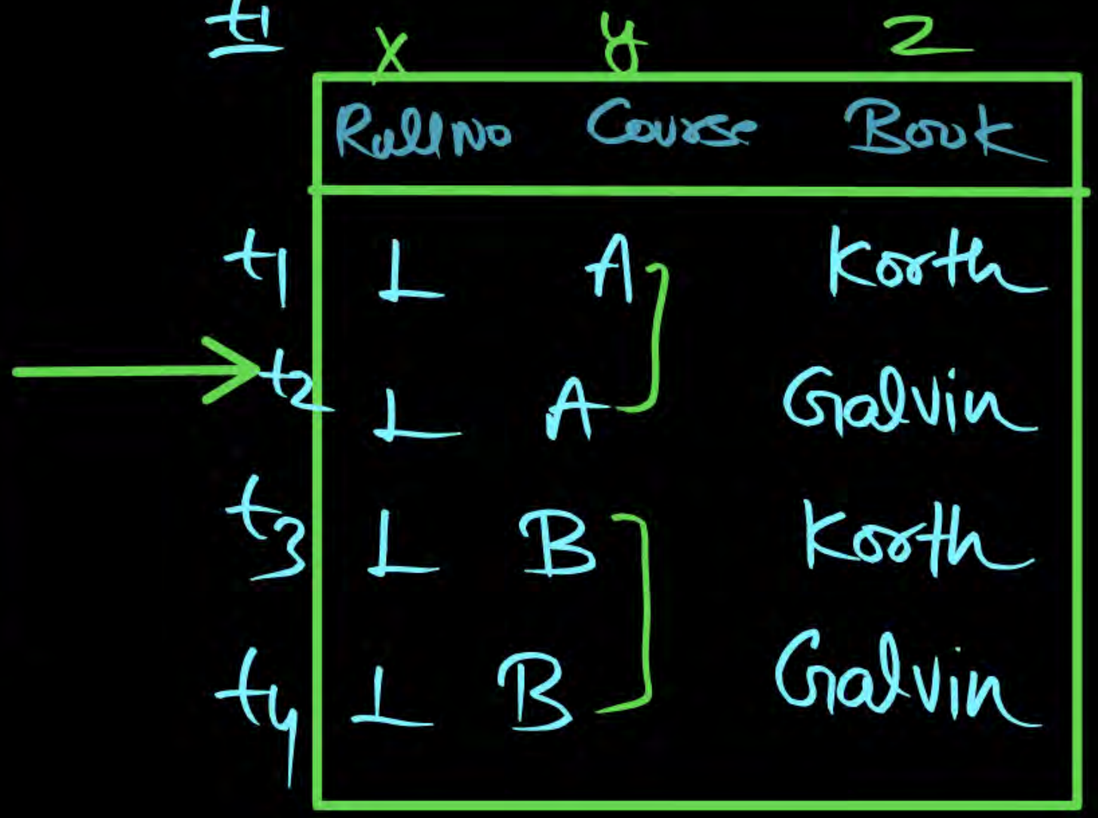
$X \rightarrow y$

If $t_1.X = t_2.X$ then $t_1.y = t_2.y$ must be same.

Multivalued Functional Dependency:

$X \rightarrow\rightarrow Y$

RollNo	Course	Book
1	A/B	Korth/Galvin



Multivalued Functional Dependency:

$$X \rightarrow\rightarrow Y$$

If $t_1.X = t_2.X = t_3.X = t_4.X$

&

$t_1.Y = t_2.Y$ && $t_3.Y = t_4.Y$

&

$t_1.Z = t_3.Z$ && $t_2.Z = t_4.Z$

t	X	Y	Z
t_1	x_1	y_1	z_1
t_2	x_1	y_1	z_2
t_3	x_1	y_2	z_1
t_4	x_1	y_2	z_2

Q.

Let R(A, B, C, D, E, P, G) be a relational schema in which the following functional dependencies are known to hold:

$AB \rightarrow CD$, $DE \rightarrow P$, $C \rightarrow E$, $P \rightarrow C$ and $B \rightarrow G$.

The relational schema R is

A. In BCNF

B. In 3NF, but not in BCNF

C. In 2NF, but not in 3NF

D. Not in 2NF

Q

The relation scheme student Performance (name, courseNO, rollNo, grade) has the following functional dependencies:

[2004: 2 Marks]

name, courseNo → grade

RollNo, courseNo → grade

name → rollNo

rollNO → name

Candidate key = [Name CourseNo, RollNo CourseNo]

key Attribute

The highest normal form of this relation scheme is

A 2 NF

B 3 NF

C BCNF

D 4 NF

Ans (B)

MCQ

In a relational data model, which one of the following statements is TRUE? GATE-2022-CS: 1M]

Binary Relation

A. A relation with only two attributes is always in BCNF.

B. If all attributes of a relation are prime attributes, then the relation is in BCNF.

C. Every relation has at least one non-prime attribute.

Ans (A)

D. BCNF decompositions preserve functional dependencies.

NAT

Consider a relation R(A, B, C, D, E) with the following three functional dependencies.

$AB \rightarrow C$; $BC \rightarrow D$; $C \rightarrow E$;

The number of super keys in the relation R is 8 Ans

[GATE-2022-CS: 1M]

MCQ

Consider a relational table R that is in 3 NF, but not in BCNF. Which one of the following statements is TRUE?

[GATE-2020-CS: 2M]

A. R has a non-trivial functional dependency $X \rightarrow A$, where X is not a superkey and A is a prime attribute.

B. R has a non-trivial functional dependency $X \rightarrow A$, where X is not a superkey and A is a non-prime attribute and X is not a proper subset of any key.

C. R has a non-trivial functional dependency $X \rightarrow A$, where X is not a superkey and A is a non-prime attribute and X is a proper subset of some key.

D. A cell in R holds a set instead of an atomic value.

Q.

Consider the following four relational schemas. For each schema, all non-trivial functional dependencies are listed. The underlined attributes are the respective primary keys.

Schema I: Registration (rollno, courses)
Field 'courses' is a set-valued attribute containing the set of courses a student has registered for.
Non-trivial functional dependency:
rollno → courses → BCNF

Schema II: Registration (rollno, courseid, email)
Non-trivial functional dependencies:
rollno, courseid → email
email → rollno → key Attribute → 3NF But Not in BCNF

Schema III: Registration (rollno, courseid, marks, grade)
Non-trivial functional dependencies:
rollno, courseid → marks, grade
marks → grade → 2NF But Not in 3NF
Marks is Not Super key
grade is Non key
So Not in 3NF

Schema IV: Registration (rollno, courseid, marks, credit)
Non-trivial functional dependencies:
rollno, courseid → credit
courseid → credit → Not in 2NF.

Proper Subset of CK → Non key Attribute

Which one of the relational schemas above is in 3NF but not in BCNF?

(A) Schema I
(B) Schema II
(C) Schema III
(D) Schema IV

[MCQ: 2018: 2M]

NAT

Given an instance of the STUDENTS relation as shown below:

Student ID	Student Name	Student Email	Student Age	CPI
2345	Shankar	shankar@math	X	9.4
1287	Swati	swati@ee	19	9.5
7853	Shankar	shankar@cse	19	9.4
9876	Swati	swati@mech	18	9.3
8765	Ganesh	ganesh@civil	19	8.7

For (Student Name, Student Age) to be a key for this instance, the value X should NOT be equal to 19 Ans .

[GATE-2014-CS: 1M]

NAT

The maximum number of superkeys for the relation schema R (E, F, G, H) with E as the key is 8 Superkey.

[GATE-2014-CS: 1M]

MCQ

Given the following two statements:

S1: Every table with two single-valued attributes is in 1 NF, 2 NF, 3 NF and BCNF.

S2: AB→C, D→E, E→C is a minimal cover for the set of functional dependencies AB→C, D→E, AB→E, E→C.

Which one of the following is CORRECT?

[GATE-2014-CS: 2M]

(A) S1 is TRUE and S2 is FALSE.

(B) Both S1 and S2 are TRUE.

(C) S1 is FALSE and S2 is TRUE

(D) Both S1 and S2 are FALSE.

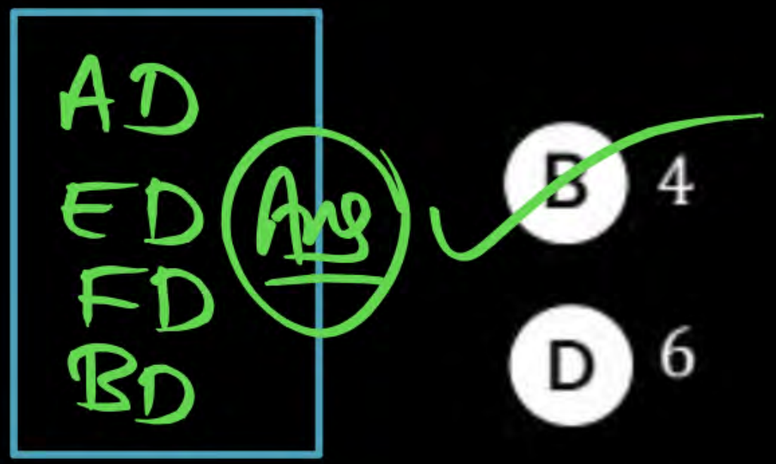
Ans (A)

MCQ

Relation R has eight attributes ABCDEFGH.
Fields of R contain only atomic values.
$F = \{CH \rightarrow G, A \rightarrow BC, B \rightarrow CFH, E \rightarrow A, F \rightarrow EG\}$ is a set of functional dependencies (FDs) so that F^+ is exactly the set of FDs that hold for R.
How many candidate keys does the relation R have?

[GATE-2013-CS: 2M]

4CK



A 3

B 4

C 5

D 6

MCQ

4 CK: AD, ED, FD, BD

Key/Prime Attribute = (A, B, D, E, F)

Non key Attribute = (C, G, H)

Relation R has eight attributes ABCDEFGH. Fields of R contain only atomic values. $F = \{CH \rightarrow G, A \rightarrow BC, B \rightarrow CFH, E \rightarrow A, F \rightarrow EG\}$ is a set of functional dependencies (FDs) so that F^+ is exactly the set of FDs that hold for R. The relation R is

Proper subset of CK → Non key Attribute

$A \rightarrow BC$, $A \rightarrow B$, $A \rightarrow C$

$F \rightarrow EG$, $F \rightarrow E$, $F \rightarrow G$

$B \rightarrow CFH$, $B \rightarrow C$, $B \rightarrow F$, $B \rightarrow H$

[GATE-2013-CS: 2M]

Not in 2NF

(A) in 1 NF, but not in 2 NF.

(B) in 2 NF, but not in 3 NF.

(C) in 3NF, but not in BCNF.

(D) in BCNF.

Ans (A)

MCQ

Which of the following is TRUE?

[GATE-2012-CS: 1M]

A. Every relation in 3 NF is also in BCNF

B. A relation R is in 3 NF if every non-prime attribute of R is fully functionally dependent on every key of R

C. Every relation in BCNF is also in 3 NF

D. No relation can be in both BCNF and 3 NF

MCQ

Consider the following relational schemes for a library database:
Book (Title, Author, Catalog_no, Publisher, Year, price)
Collection (Title, Author, Catalog_no)
With the following functional dependencies:

I. TitleAuthor ⟶ Catalog_no
II. Catalog_no ⟶ Title Author Publisher Year
III. Publisher Title Year → Price

Assume { Author, Title} is the key for both schemes.
Which of the following statements is true?

[GATE-2008-CS: 2M]

A Both Book and Collection are in BCNF

B Both Book and Collection are in 3 NF only

C Book is in 2 NF and Collection is in 3NF

D Both Book and Collection are in 2 NF only

MCQ

Let R(A, B, C, D, E, P, G) be a relational schema in which the following functional dependencies are known to hold:
AB→CD, DE→P, C→E, P→C and B→G.
The relational schema R is

[GATE-2008-CS: 2M]

A in BCNF

B in 3NF, but not in BCNF

C in 2 NF, but not in 3 NF

D not in 2 NF

Q

Consider the following Relation:

R(ABCDEFG) with FD set of Relation R {A → B, C → D, E → FG}. What is the minimum number of relations required to decompose into BCNF which satisfy lossless join and Dependency preserving decomposition _____

MCQ

Let the set of functional dependencies F= {QR→S, R→P, S→Q} hold on a relation schema X=(PQRS) . X is not in BCNF. Suppose X is decomposed into two schemas Y and Z, where Y=(P R) and Z=(Q R S).
Consider the two statements given below.
I. Both Y and Z are in BCNF
II. Decomposition of X into Y and Z is dependency preserving and lossless
Which of the above statements is/are correct?

[GATE-2019-CS: 2M]

A Both I and II

B I only

C II only

D Neither I nor II

Q.

A database of research articles in a journal uses the following schema.

(VOLUME, NUMBER, STARTPAGE, ENDPAGE, TITLE, YEAR, PRICE)

The primary key is (VOLUME, NUMBER, STARTPAGE, ENDPAGE) and the following functional dependencies exist in the schema

FD I (VOLUME, NUMBER, STARTPAGE, ENDPAGE) → TITLE

FD II (VOLUME, NUMBER) → YEAR

FD III (VOLUME, NUMBER, STARTPAGE, ENDPAGE) → PRICE.

The database is redesigned to use the following schemas.

R_1 (VOLUME, NUMBER, STARTPAGE, ENDPAGE, TITLE, PRICE)

R_2 (VOLUME, NUMBER, YEAR)

Which of the weakest normal form that the new database satisfies, but the old one does not?

[MCQ: 2016: 1M]

(A) 1NF

(B) 3NF

(C) 2NF

(D) BCNF

Any Doubt ?

THANK YOU!