# COMPUTER SCIENCE

## Database Management System

## FD's & Normalization

Lecture_14



Vijay Agarwal sir


01 Normal Forms

02 Normal Form Decomposition

1NF

2NF

3NF

BCNF

1 NF ⟶ atomic

# Normal Forms

## 2NF :


Violation of 2NF

# Normal Forms

3NF : Every Non Trivial $X \rightarrow Y$ Must Satisfy the following Condition

X : Super key

OR

Y : key / Prime Attribute

# Normal Forms

BCNF: $X \rightarrow Y$ every Non Toivial FD

X: Super key.

# Boyce – Codd Normal Form

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

X: superkey.

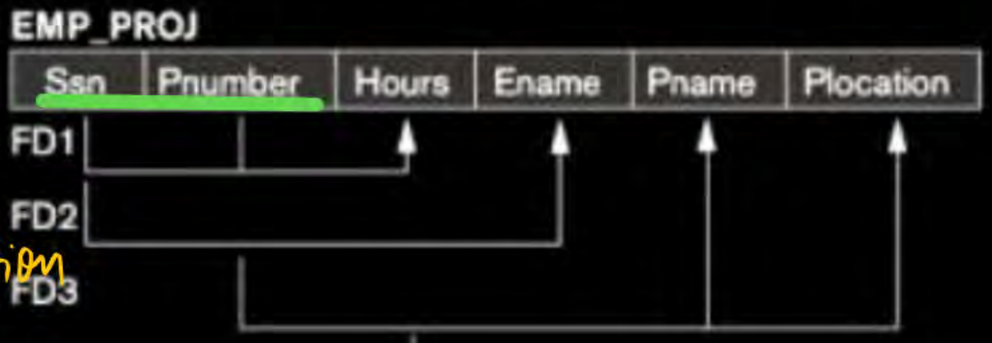
# Normal Forms

## Second Normal Form

**Definition:** A relation schema R is in 2NF if every nonprime attribute A in R is fully functionally dependent on the primary key of R.

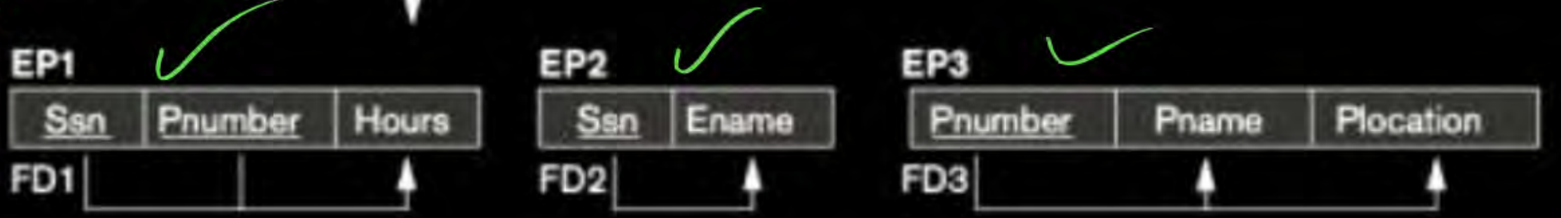
FD2 : Ssn ⟶ Ename

FD3 : Pnumber ⟶ Pname Plocation




2NF Normalization

2NF ⟵



# 2NF Decomposition

**Q.1** R(ABCDEFGH) {AB → C, C → D, B → E, E → FG, G → H}

Candidate key = [AB]

Non key / Non Prime Attribute = [C, D, E, F, G]

Check 2NF ?

B → E

Proper subset of CK

Non key Attribute

Not in 2NF

## 2NF Decomposition

R(ABCDEFGH)

$(B)^+ = [BEFGH]$

R1

| AB | C | D |
|---|---|---|

CK: AB

2NF

CK: B R2

| B | E | F | G | H |
|---|---|---|---|---|

- ✓ AB → C
- ✓ C → D
- ✓ B → E
- ✓ E → FG
- ✓ G → H

Dependency Preserving.

$R_1(ABCD)$ $R_2(BEFGH)$

$R_1(ABCD) \cap R_2(BEFGH) = B$

$[B]^+ = [BEFGH]$ Super key of $R_2$.

$R_{12}(ABCDEFGH)$ : Lossless Join

2NF + Lossless Join + Dep. Preserved.

# 2NF Decomposition

**Q.2** R(ABCDE) F: [A → B, B → E, C → D]

Decompose it into 2NF.

Candidate key = [AC]

Non key Attribute = [B, D, E]

Check 2NF ?

A → B
C → D
} Proper Subset of CK → Non key Attribute

Not in 2NF

## 2NF Decomposition

$(A)^+ = [ABE]$

$(C)^+ = [CD]$

$R_1$ = ABE

$R_2$ = CD

Lossy

$R_1(ABE) \cap R_2(CD)$

No Common Attribute

Lossy Join.

# 2NF Decomposition

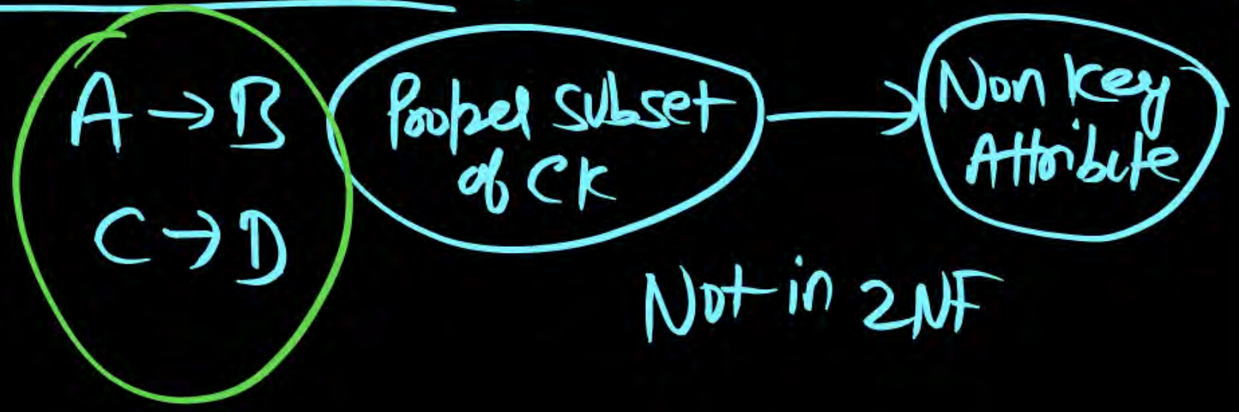
**Q.2** R(ABCDE) F: [A → B, B → E, C → D]

Decompose it into 2NF.

Candidate key = [AC]

Non key Attribute = [B, D, E]

Check 2NF ?




Not in 2NF

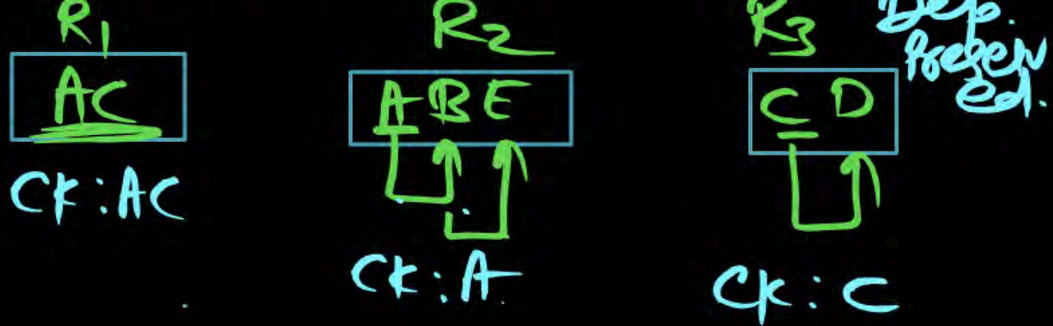
## 2NF Decomposition

R(ABCDE)

A → B
B → E
C → D
Dep. preserved

$(A)^+ = (ABE)$

$(C)^+ = (CD)$



$R_1(AC)$ $R_2(ABE)$ $R_3(CD)$

$R_1(AC) \cap R_2(ABE) = A$

$[A]^+ = [ABE]$ Super key of $R_2$

$R_{12}(ABCE) \cap R_3(CD) = C$

$(C]^+ = [CD]$ Super key of $R_3$

$R_{123}(ABCDE)$ Lossless Join

2NF + Lossless + D.P.

# 2NF Decomposition

① Find Minimal Cover.

# 2NF Decomposition

**Q.3** R (ABCDEFGHIJ) {AB→C, BD→EF, AD→GH, A→I, H→J}

Candidate key = [ABD]

Non key Attribute = [C, E, F, G, H, I, J]

Checking 2NF ?

AB→C
BD→EF
AD→GH
A→I

Not in 2NF

Proper subset of CK → Non key Attribute

$(AB)^+ = (ABCI)$
$(BD)^+ = (BDEF)$
$(AD)^+ = (ADGHIJ)$
$(A)^+ = (AI)$

2NF Decomposition

R(ABCDEFGHIJ)

$R_2$ ABCI

$R_3$ BDEF

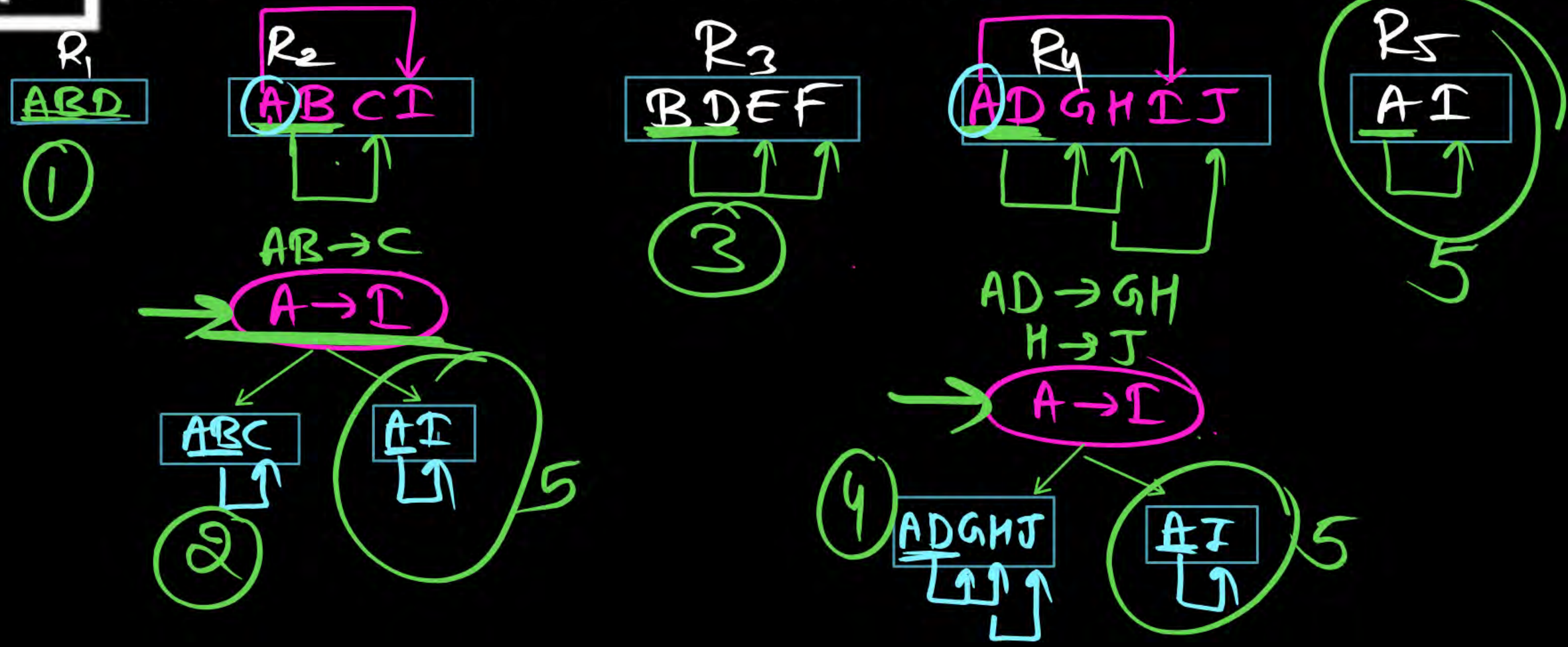
$R_4$ ADGHIJ

$R_5$ AI

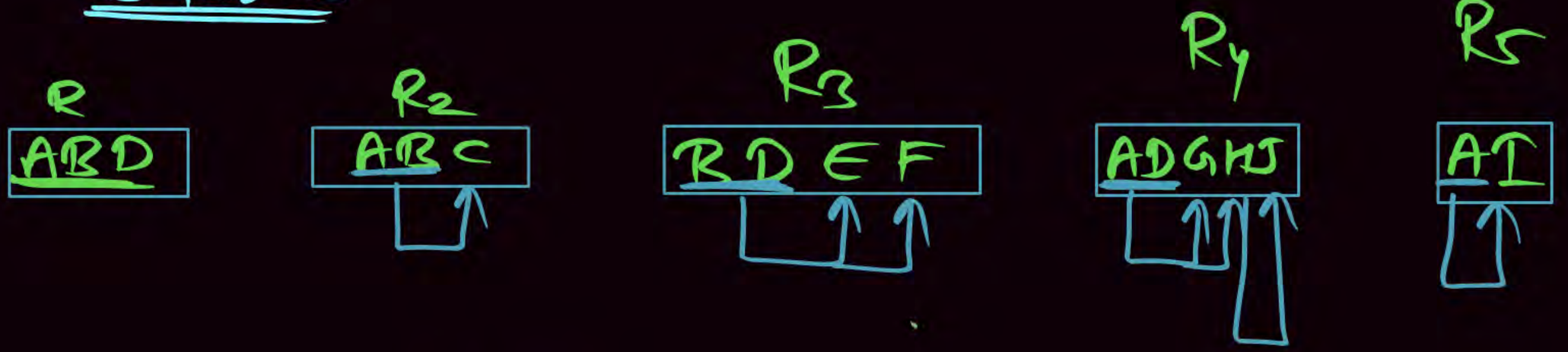
ABD.

# 2NF Decomposition

**Q.3** R (ABCDEFGHIJ) {AB→C, BD→EF, AD→GH, A→I, H→J}



5 Table




2NF
+ Lossless
+ Dep. Preserving.

# 2NF Decomposition

**Q.4** R(ABCDEF) {AB → C, C → D, B → EF}

Candidate key = [AB]

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

Checking 2NF ?

B → EF

B: Proper subset of CK

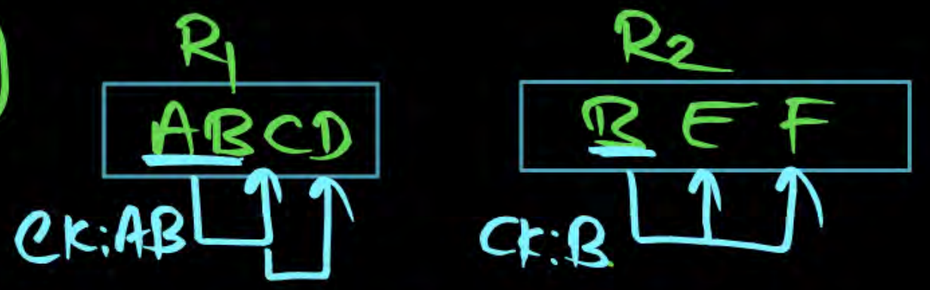
EF: Non key Attribute

Not in 2NF

$(B)^+ = (BEF)$

2NF Decomposition

R(ABCDEF)




2NF + Dep. Preserved + Lossless Join

$R_1(ABCD) \cap R_2(BEF) = B$

$(B)^+ = (BEF)$ Super key of $R_2$

$R_{12}(ABCDEF)$ Lossless

# 2NF Decomposition

**Q.5** R(ABCDEFGH) {AB → C, C → D, B → E, E → F, A → GH}

Candidate key = [AB]

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

Checking 2NF ?

B → E

A → GH

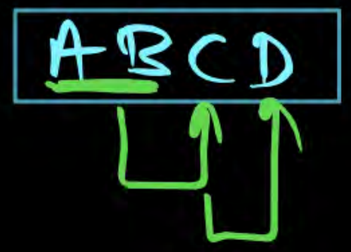
Proper Subset of CK → Non key Attribute

$(B)^+ = [BEF]$

$(A)^+ = [AGH]$

2NF Decomposition.

R(ABCD EF GH)






2NF + D.P + Lossless Join.

$R_1(ABCD)$ $R_2(BEF)$ $R_3(AGH)$

$R_1(ABCD) \cap R_2(BEF) = B$

$(B)^+ = [BEF]$ Super key of $R_2$

$R_{12}(ABCDEF) \cap R_3(AGH) = A$

$[A]^+ = [AGH]$ Super key of $R_3$

$R_{123}(ABCDEFGH)$

Lossless Join

Class Notes

+ PYQ [PW (Gate Wallah)
With video solution.

+ D.P.P

+ W.T

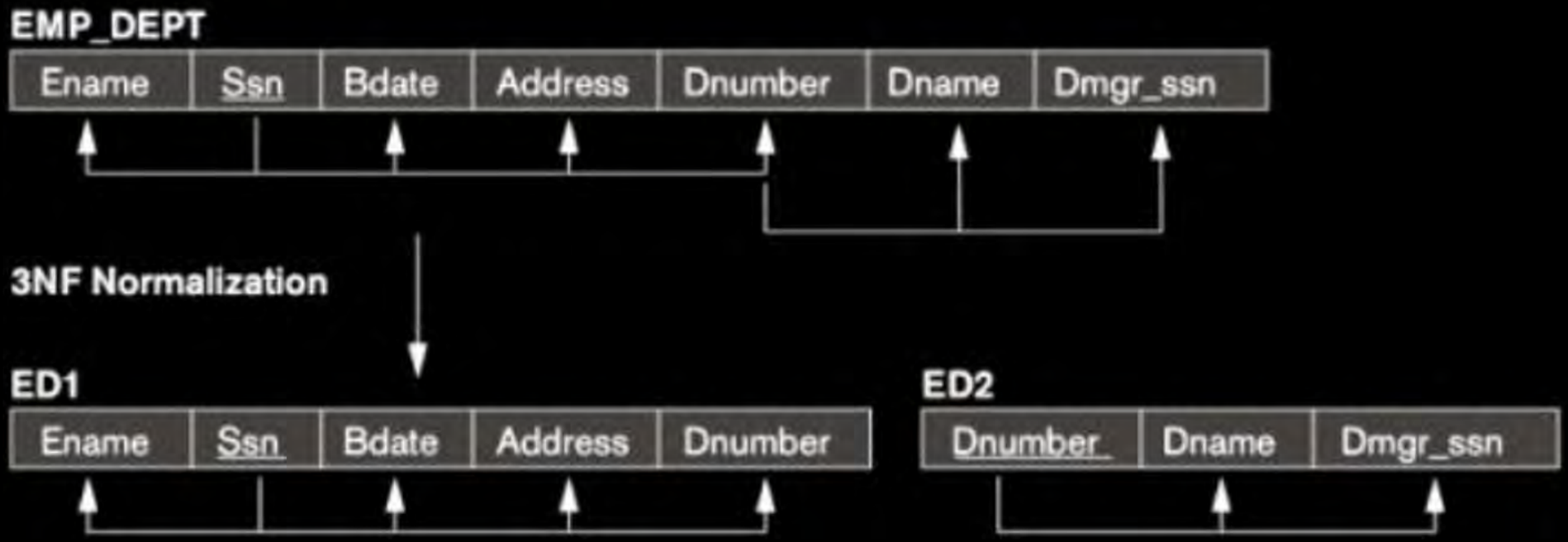
+ PW [Gate Wallah Test series]

# 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.

**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

# 3NF Decomposition

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

Candidate key = (A)

Non key Attribute = (B, C)

CHECKING 2NF ?

R is in 2NF.

Check 3NF ?

$B \rightarrow C$

N.K → N.K

B is Not Super key

C is Non Prime Attribute.

Not in 3NF

## 3NF Decomposition


A : CK

B : CK

3NF

+ Dependency Preserved

+ Lossless Join

$R_1(AB) \cap R_2(BC)$

$(B)^+ = [BC]$

S.K of Re

Lossless.

# ⑧ Advantage of Normalization.

# 3NF Decomposition

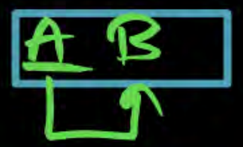
**Q.1** R(ABC) $[A \rightarrow B, B \rightarrow C]$

| A | B | C |
|---|---|---|
| 1 | $b_1$ | X |
| 2 | $b_1$ | X |
| 3 | $b_1$ | X |
| 4 | $b_1$ | X |
| 5 | $b_1$ | X |
| 6 | $c_1$ | y |
| 7 | $c_1$ | y |
| 8 | $c_1$ | y |
| 9 | $c_1$ | y |
| 10 | $c_1$ | y |
| 11 | $c_1$ | y |

✓ $A \rightarrow B$

✓ $B \rightarrow C$

3NF Decomposition





| A | B |
|---|---|
| 1 | $b_1$ |
| 2 | $b_1$ |
| 3 | $b_1$ |
| 4 | $b_1$ |
| 5 | $b_1$ |
| 6 | $c_1$ |
| 7 | $c_1$ |
| 8 | $c_1$ |
| 9 | $c_1$ |
| 10 | $c_1$ |
| 11 | $c_1$ |

| B | C |
|---|---|
| $b_1$ | X |
| $c_1$ | y |

3NF

3NF + Dep. Preserving + Lossless Join

# 3NF Decomposition

**Q.2** R(ABCDEF) [AB → C, C → D, D → E, E → F]

Candidate key = {AB}

Non key Attribute = {C, D, E, F}

Checking 2NF ?

R is in 2NF

Checking 3NF

X → Y

C → D

D → E

E → F

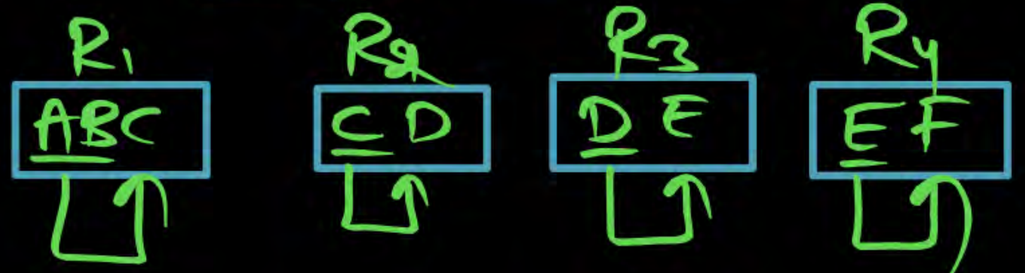
X: is Not Subset key

OR

Y: Not Prime/Not key Attribute

Not in 3NF

3NF Decomposition




3NF + Lossless

+ Dep. Preserved.

# 3NF Decomposition

**Q.2** R(ABCDEF) [AB → C, C → D, D → E, E → F]

Candidate key = {AB}

Non key Attribute = {C, D, E, F}

Checking 2NF ?

R is in 2NF

Checking 3NF

X → Y

C → D

D → E

E → F

X: is Not Subset key

OR

Y: Not Prime/Not key Attribute

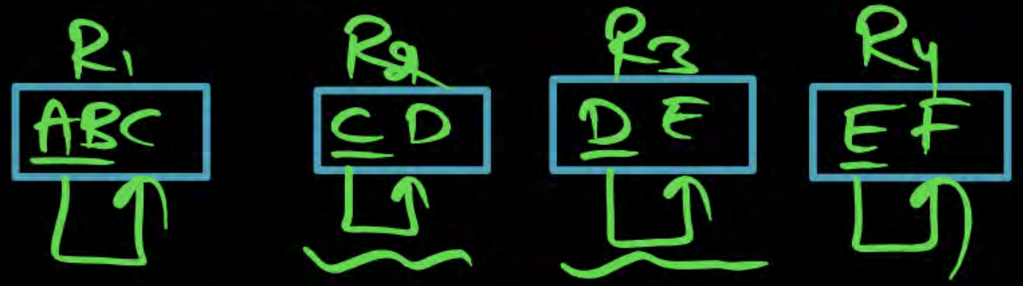
Not in 3NF


NK → N.K

OR

D: Not Subset key, E Non key

3NF Decomposition



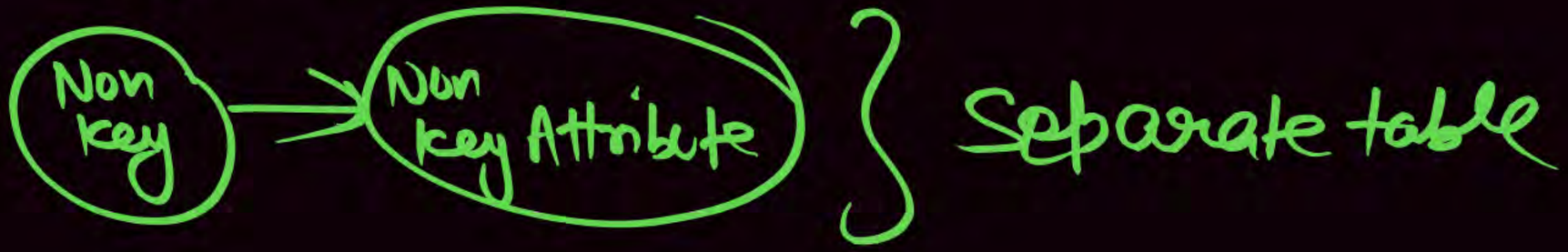

So

3NF + Lossless

+ Dep. Preserved.

# 3NF Decomposition

2NF +



# 3NF Decomposition

**Q.2** R(ABCDEF) [AB → C, C → D, C → E, E → F]

Candidate key = {AB}

Non key Attribute = {C, D, E, F}

Checking 2NF ?

R is in 2NF

Checking 3NF

X → Y

C → D

C → E

E → F

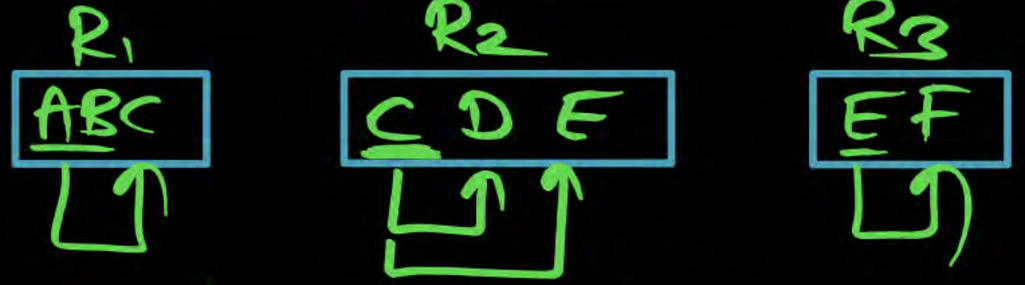
X: is Not Subset key

OR

Y: Not Prime/Not key Attribute

Not in 3NF

3NF Decomposition




3NF + Lossless

+ Dep. Preserved.

R2: C is Subset key [3NF]

# 3NF Decomposition

**Q.3** R (ABCDEFGHIJ) {AB→C, BD→EF, AD→GH, A→I, H→J}

Candidate key = ABD

Non key Attribute = [C, E, F, G, H, I, J]

Directly checking 3NF ?

3NF $X \to y$: X: Super key (OR) y: key | Prime Attribute

$X \to y$ ?

AB→C
BD→EF
AD→GH
A→I
H→J

X: is Not Super key (OR) y: is Non key | Non Prime Attribute

So Not in 3NF

# 3NF Decomposition

**Q.3** R (ABCDEFGHIJ) {AB→C, BD→EF, AD→GH, A→I, H→J}

Step by step

Candidate key = ABD

Non key Attribute = [C, E, F, G, H, I, J]

CHeck 2NF ?

AB→C
BD→EF
AD→GH
A→I

Proper subset of CK → Non key Attribute

Not in 2NF

# 2NF Decomposition

① AD → H → J ; Non key Attribute J is transitively Depend on Ck : AD

## 5 Table

R: ABD

R2: ABC

R3: BDEF

R4: ADGHJ

R5: AI

② H → J
H: sub part of key
OR
J: Not key Attribute

③ H → J (N.K → N.K)

# 3NF Decomposition

Ck: ABD — ABD

Ck: A — ABC

Ck: BD — BDEF

Ck: AD — ADGH

Ck: H — HJ

Ck: A — AI

3NF + Lossless + D.P.

# 3NF Decomposition

**Q.4** R (ABCD) {AB→CD, D→A}

# 3NF Decomposition

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

**Q.6**

R(ABCDE) $\{AB \rightarrow C, C \rightarrow D, B \rightarrow E\}$

Decompose into 2NF, 3NF, BCNF

# Boyce – Codd Normal Form

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

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]

A Dependency-preservation

B Lossless-join

C BCNF definition

D 3 NF definition

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

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

The highest normal form of this relation scheme is

A 2 NF

B 3 NF

C BCNF

D 4 NF

# MCQ

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

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.

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 _____.

[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.

# 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 _____.

[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 _____.

[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: $AB \rightarrow C$, $D \rightarrow E$, $E \rightarrow C$ is a minimal cover for the set of functional dependencies $AB \rightarrow C$, $D \rightarrow E$, $AB \rightarrow E$, $E \rightarrow C$.

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.

# 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]

A 3

B 4

C 5

D 6

# MCQ

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

[GATE-2013-CS: 2M]

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.

# 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.

(VOLUME, NUMBER, STARTPAGE, ENDPAGE) $\rightarrow$ TITLE

(VOLUME, NUMBER) $\rightarrow$ YEAR

(VOLUME, NUMBER, STARTPAGE, ENDPAGE) $\rightarrow$ PRICE.

The database is redesigned to use the following schemas.

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

(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!