COMPUTER SCIENCE

Database Management System

FD's & Normalization

Lecture_13

Vijay Agarwal sir

TOPICS TO BE COVERED

01 Normal Forms

02 Normal Form Decomposition

1NF ⇒ atomic. OR No Multivalued Attribute

2NF
3NF
BCNF

Checking Condition

Partial FD

Full FD

Partial FD

$\boxed{AB \to C \text{ is Partial FD}}$

if $A \to C$

OR

if $B \to C$

$\boxed{X \to y \text{ is Partial FD}}$

$A \in X$

$(X - A) \longrightarrow y$

FULL FD

$\boxed{AB \to C \text{ is FULL FD}}$

if $A \not\to C$

$B \not\to C$.

$\boxed{X \to Y \text{ is FULL FD}}$

$A \in X$

$(X - A) \not\to Y$

Normal Forms

1NF

- R is in 1NF if all attribute of R are atomic
- Default RDBMS is in 1NF
- 1NF Ensured by Candidate key.

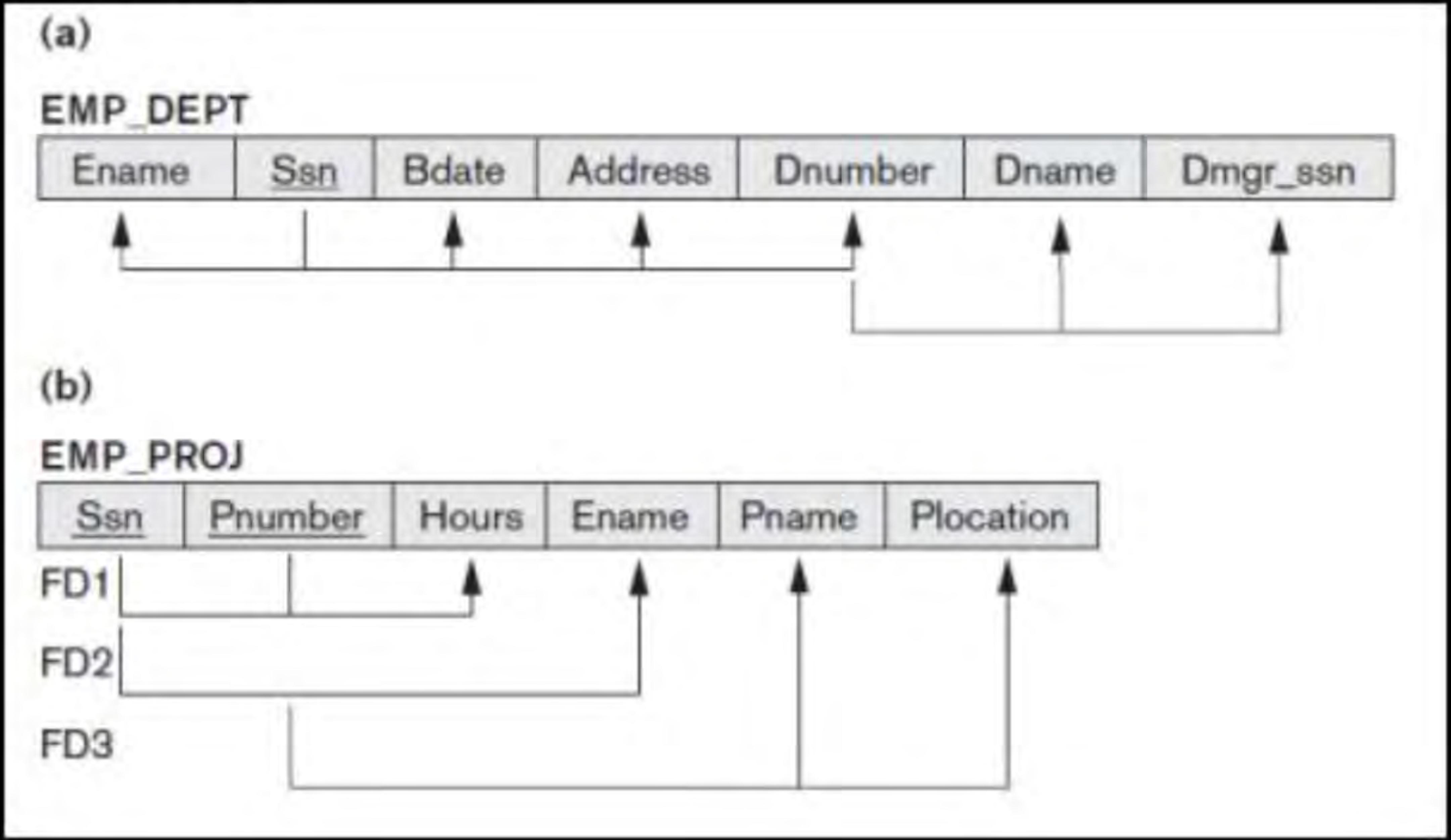
Normal Forms

15.3.5 Second Normal Form

Second normal form (2NF) is based on the concept of *full functional dependency*. A functional dependency $X \rightarrow Y$ is a **full functional dependency** if removal of any attribute A from X means that the dependency does not hold any more; that is, for any attribute $A \varepsilon X$, $(X - \{A\})$ does *not* functionally determine Y. A functional dependency $X \rightarrow Y$ is a **partial dependency** if some attribute $A \varepsilon X$ can be removed from X and the dependency still holds; that is, for some $A \varepsilon X$, $(X - \{A\}) \rightarrow Y$. In Figure 15.3(b), {Ssn, Pnumber} → Hours is a full dependency (neither Ssn → Hours nor Pnumber → Hours holds). However, the dependency {Ssn, Pnumber} → Ename is partial because Ssn → Ename holds.

> **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.

Figure 15.3


(a)
EMP_DEPT
Ename
Ssn
Bdate
Address
Dnumber
Dname
Dmgr_ssn
(b)
EMP_PROJ
Ssn
Pnumber
Hours
Ename
Pname
Plocation
FD1
FD2
FD3

Normal Forms

8.17 A functional dependency $\alpha \rightarrow \beta$ is called a partial dependency if there is a proper subset γ of α such that $\gamma \rightarrow \beta$. We say that β is *partially dependent* on α. A relation schema *R* is in second normal form (2NF) if each attribute *A* in *R* meets one of the following criteria:

- It appears in a candidate key.
- It is not partially dependent on a candidate key.

Normal Forms

Second Normal Form

Ename, Pname, Plocation is Partially Dep on P.K (C.K)

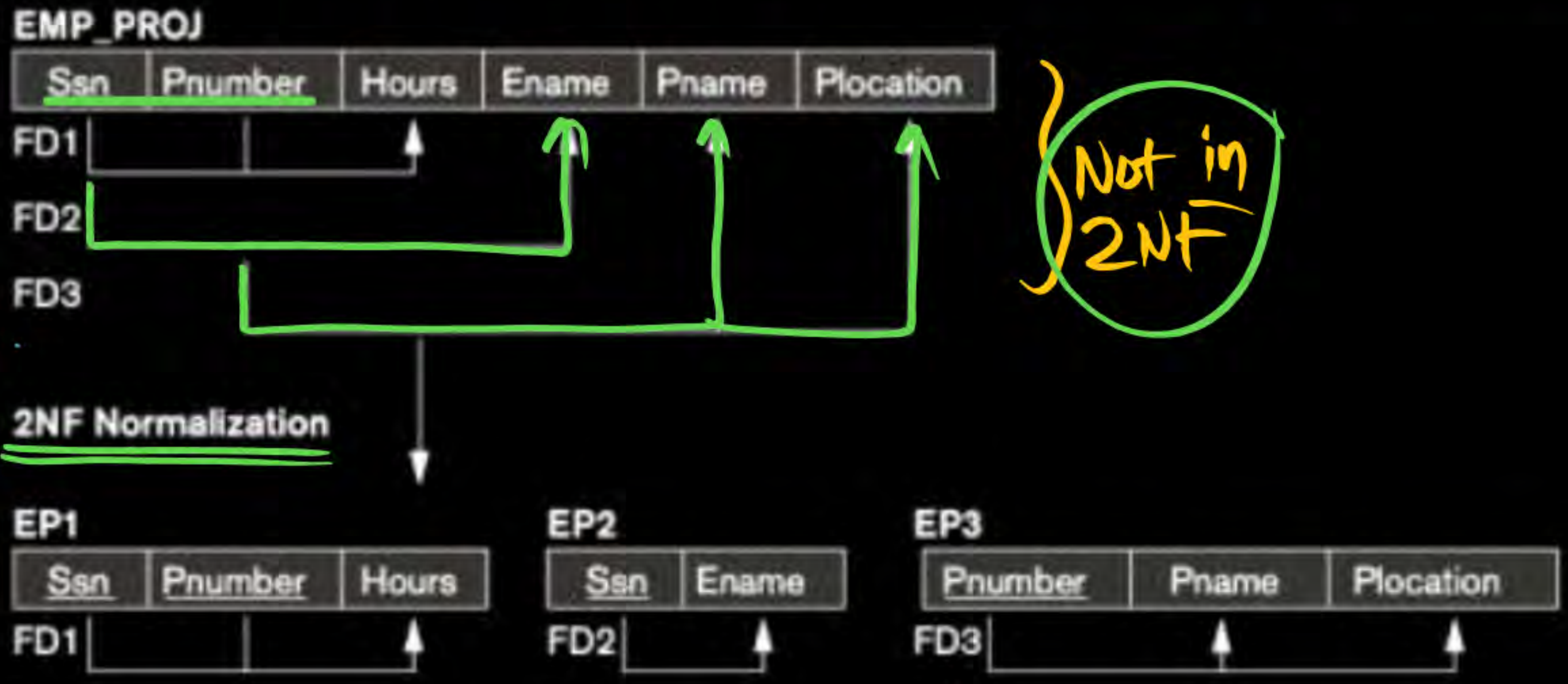
Definition: A relation schema R is in 2NF if every nonprime attribute A in R is fully functionally dependent on the primary key of R.

$\boxed{\text{Ssn Pnumber}} \rightarrow$ Hours

Ssn → Ename

Pnumber → Pname Plocation

$\boxed{\text{Ssn Pnumber is CK}}$



Not in 2NF

Partial FD

FD I: $\boxed{\text{Ssn Pnumber} \rightarrow \text{Hours}}$

FD II: Ssn → Ename

FD III: Pnumber → Pname Plocation

$\boxed{\text{Candidate key} = [\text{Ssn Pnumber}]}$

Ssn Pnumber is key

Ssn → Ename

Pnumber → Pname Plocation

Ssn Pnumber → Ename

Non key Attribute is Not Fully Dependent on P.K

Ssn → Ename

Proper subset of Cand. key → Non key Attribute ✗

R(ABCDE) [AB→C, C→D, B→E]

Candidate key = [AB]

Non key Attribute = [C, D, E]

AB→E

B → E

Proper subset of CK

Non key Attribute

} Not in 2NF

FD I : Ssn Pnumber → Hours

FD II : Ssn → Ename

FD III : Pnumber → Pname Plocation

$\boxed{\text{Candidate key} = [\text{Ssn Pnumber}]}$

Ssn → Ename

Pnumber → Pname Plocation

Proper subset of C.K → Non key Attribute

Not in 2NF.

Proper subset of Cand. key → Non key Attribute ✗

R(ABCDE) [AB→C, C→D, B→E]

Candidate key = [AB]

Non key Attribute = [C, D, E]

B → E

Proper subset of CK

Non key Attribute

Not in 2NF

Normal Forms

2NF

R is in 2NF

(i) R is in 1NF.

(ii) R does not contain this below type FD

Condition checking

Proper subset of Candidate key $\rightarrow$ Non key Attribute

Violation of 2NF

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

Candidate key = (AB)

Non key Attribute = (C, D, E, P, G)

$B \rightarrow G$

Proper Subset of CK

Non key Attribute

Violation of 2NF

Not in 2NF

(A) In BCNF

(B) In 3NF, but not in BCNF

(C) In 2NF, but not in 3NF

(D) Not in 2NF

Normal Forms

Second Normal Form

Ename, Pname, Plocation is Partially Dep on P.K (C.K).

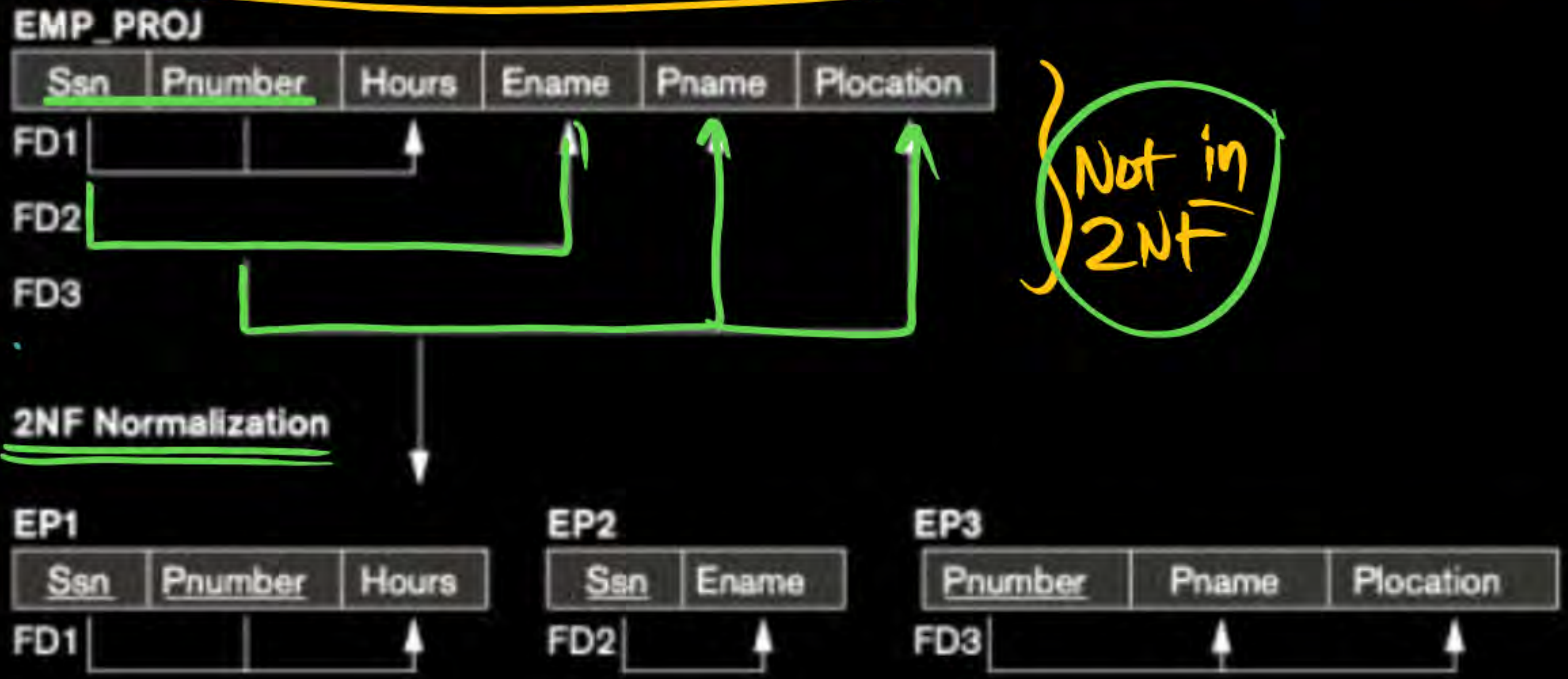
Definition: A relation schema R is in 2NF if every nonprime attribute A in R is fully functionally dependent on the primary key of R.

$\boxed{\text{Ssn Pnumber}} \rightarrow$ Hours

Ssn $\rightarrow$ Ename

Pnumber $\rightarrow$ Pname Plocation

$\boxed{\text{Ssn Pnumber is CK}}$



Not in 2NF

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

Candidate key = (AB)

Non key Attribute = (C, D, E, P, G)

AB is a key, ie AB Determine all attribute of Relation

A. In BCNF

B. In 3NF, but not in BCNF

C. In 2NF, but not in 3NF

D. Not in 2NF

$\boxed{AB \longrightarrow G}$

$B \rightarrow G$

Partial Dependency

Here Non Prime Attribute G is Not fully Dependent on Primary key AB.

Not in 2NF

Third Normal Form

R is in 3NF every Non Trivial FD Must Satisfy Following Condition

R is in 2NF
No Transitive Dep.

X : Super key
OR
y : key/Prime Attribute

Third Normal Form

R is in 3NF every Non Trivial FD Must Satisfy Following Condition

X : Super key

OR

y : key/Prime Attribute

$R(ABC)$

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

then $A \to C$ is Transitive Dep.

<u>Transitive Dep.</u>

$$X \to Y \longrightarrow Z$$

then $X \to Z$ is transitive Dep.

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.

$X \rightarrow A$ is in 3NF

X: Super key

OR

A: Prime / key Attribute.

R(ABC) [A → B, B → C]

Candidate key = [A]

Non key Attribute = [B, C]

every Non Trivial FD $X \to Y$ is in 3NF

either X: super key. OR

Y: key | Prime Attribute

$A \to B$ ✓ 3NF [A is Super key]

X $B \to C$ X 3NF [Bcz Neither B is Super key nor C is key | Prime Attribute]

B ↓ Not Super key

C ↓ C Non Prime

Not in 3NF

Non key → Non key

$B \to C$

Not in 3NF

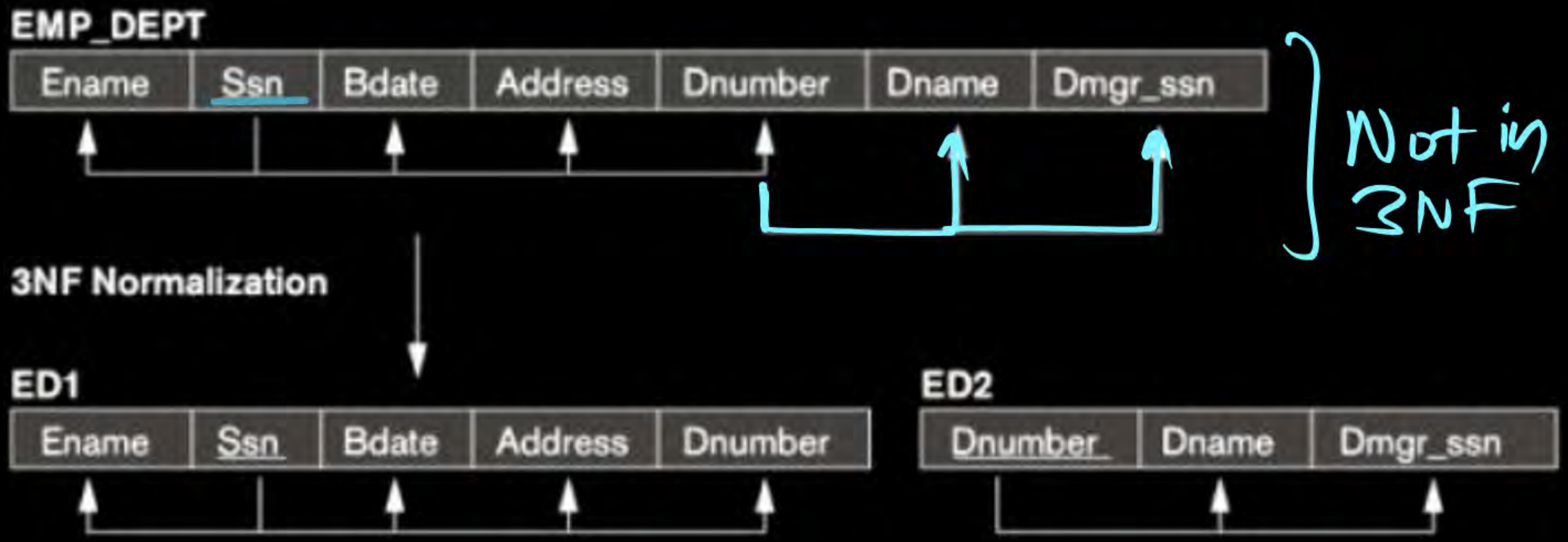
CODD Defination

$A \to B \to C$

$A \to C$ is Transitive FD

Here Non Prime Attribute C is Transitively Dependant on Primary key [A]

so, Not in 3NF


EMP_DEPT
Ename
Ssn
Bdate
Address
Dnumber
Dname
Dmgr_ssn
Not in 3NF
3NF Normalization
ED1
Ename
Ssn
Bdate
Address
Dnumber
ED2
Dnumber
Dname
Dmgr_ssn

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 : Super key.

Important Points [Checking condition]

2NF Checking.

Proper subset of C.K → Non key Attributes

Violation of 2NF

∴ Not in 2NF

3NF Checking.

$X \to A$

R is in 3NF if every Non Trivial FD must satisfy the following condition.

either

X : Super key

OR

y : key/Prime Attribute

BCNF Checking.

R is in BCNF if every $X \to A$ Non Trivial FD must satisfy the following condition

X : Super key.

Important Points

R(ABCD) {AB → CD, D → A}

Candidate key = [AB, DB]

3NF Checking ?

AB → CD ✓3NF AB is Super key

D → A ✓3NF D is Not super key
But A is key/Prime Attribute

So R is in 3NF

BCNF Checking

AB → CD ✓BCNF (AB super key)

D → A fail BCNF Bcz
D is Not Super key

Not in BCNF

Important Points [1NF]

If a Relation R has only One Candidate key then R always in 1NF But May/May Not in 2NF, 3NF & BCNF

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

Candidate key = [AB]

Non key Attribute = [C, D, E]

$B \to E$

Proper subset of CK

Non key Attribute

Not in 2NF

Important Points [2NF]

If In a Relation R all Candidate keys are simple [Single Attribute] Candidate key [No Composite key] then Relation R always is in 2NF But may or may Not in 3NF & BCNF.

eg R(ABCDE) [$A \rightarrow B$, $B \rightarrow C$, $C \rightarrow D$, $D \rightarrow E$, $C \rightarrow A$]

Candidate key = [A, C, B]

Here all C.K are simple C.K then R always in 2NF.

But Not in 3NF

$D \rightarrow E$

D: Not Super key
E: Not Prime/Not key Attribute

So R Not in 3NF.

Important Points [3NF]

If In a Relation R, all attributes are key/Prime Attribute then R always is in 3NF But May or May Not in BCNF.

eg R(ABCD) [A→B, B→C, C→A]

Candidate keys = [AD, BD, CD]

Key/Prime Attribute = [A, B, C, D]

Non Prime /Non key Attribute = [Empty]

All attributes are Prime/key Attribute So R is 3NF But Not in BCNF.

A → B

B → C

C → A

X: Not Super key.

∴ Not in BCNF

Important Points [BCNF]

If a Relation R is in 3NF, && All Candidate keys are Simple Candidate key then R always is in BCNF.

eg R (ABC) [A → B, B → C, C → A]

Candidate key = [A, B, C]

key | Prime Attribute = [A, B, C]

Here all Attribute are key | Prime Attribute. So R is in 3NF

&& all C.K are simple C.K So R is in BCNF

A → B ; A is super key

B → C ; B is super key

C → A ; C is super key

Important Points [BCNF]

Binary Relation (Relation with 2 Attribute) is always in BCNF

eg

R(AB) [$A \to B$]

Candidate key = [A]

$A \to B$; A is super key

So R is in BCNF

OR

eg

R(AB) [$B \to A$]

Candidate key = [B]

$B \to A$; B is super key

So R is in BCNF

OR

eg

R(AB) [$A \to B$, $B \to A$]

Candidate key = [A, B]

$A \to B$: A is super key

$B \to A$; B is super key

So R is in BCNF

Important Points [BCNF]

A Relation R with No Non Trivial FD is always is in BCNF

(Trivial FD) is in BCNF.

$XYZ \rightarrow XY$

superkey BCNF

Important Points

$X \rightarrow Y$

Design Goal	1NF	2NF	3NF	BCNF
0% Redundancy	✗	✗	✗	✓
Lossless Join	✓	✓	✓	✓
Dependency Preserving	✓	✓	✓	May/may Not

✗ Deal 4NF 5NF

Suffer from multivalued FD $X \rightarrow\rightarrow Y$

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 $\rightarrow$ grade

RollNo, courseNo $\rightarrow$ grade

name $\rightarrow$ rollNo

rollNO $\rightarrow$ 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.

$X \to A$

3NF

× X : Super key

OR

✓ A : key / Prime Attribute

$X \to A$

BCNF

X : Super key.

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: Binary Relation

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.
$F = \{CH \rightarrow G, A \rightarrow BC, B \rightarrow CFH, E \rightarrow A, F \rightarrow EG\}$ is a set of functional dependencies (FDs) so that F^+ is exactly the set of FDs that hold for R.
The relation R is

[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 $\longrightarrow$ Catalog_no
II. Catalog_no $\longrightarrow$ Title Author Publisher Year
III. Publisher Title Year $\rightarrow$ Price

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

[GATE-2008-CS: 2M]

MCQ

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

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

The relational schema R is

[GATE-2008-CS: 2M]

(A) in BCNF

(B) in 3NF, but not in BCNF

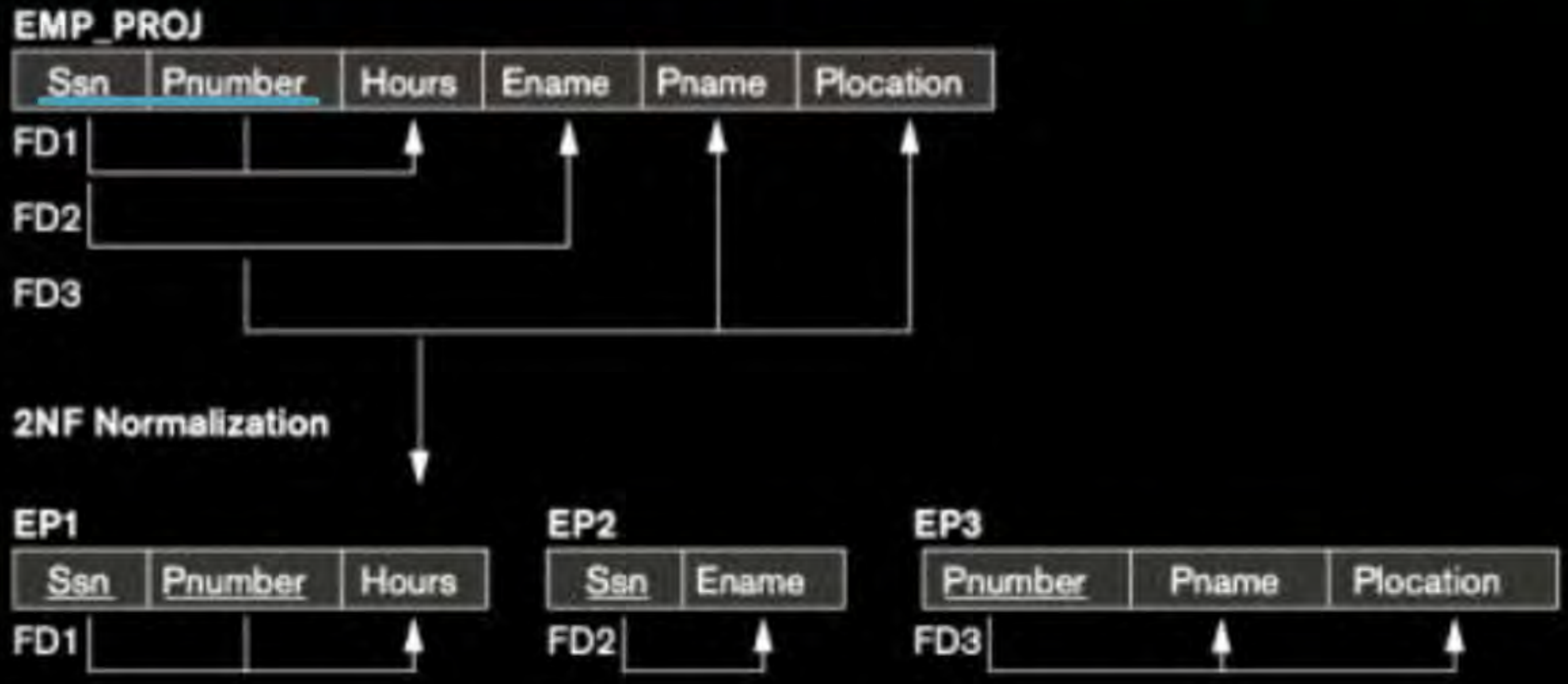
(C) in 2 NF, but not in 3 NF

(D) not in 2 NF

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.



2NF Decomposition

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

Candidate key = [AB]

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

2NF Checking ?

B → E

Proper subset of C.K

Non key Attribute

So, R Not in 2NF.

$[B]^+ = [BEFGH]$

R_1

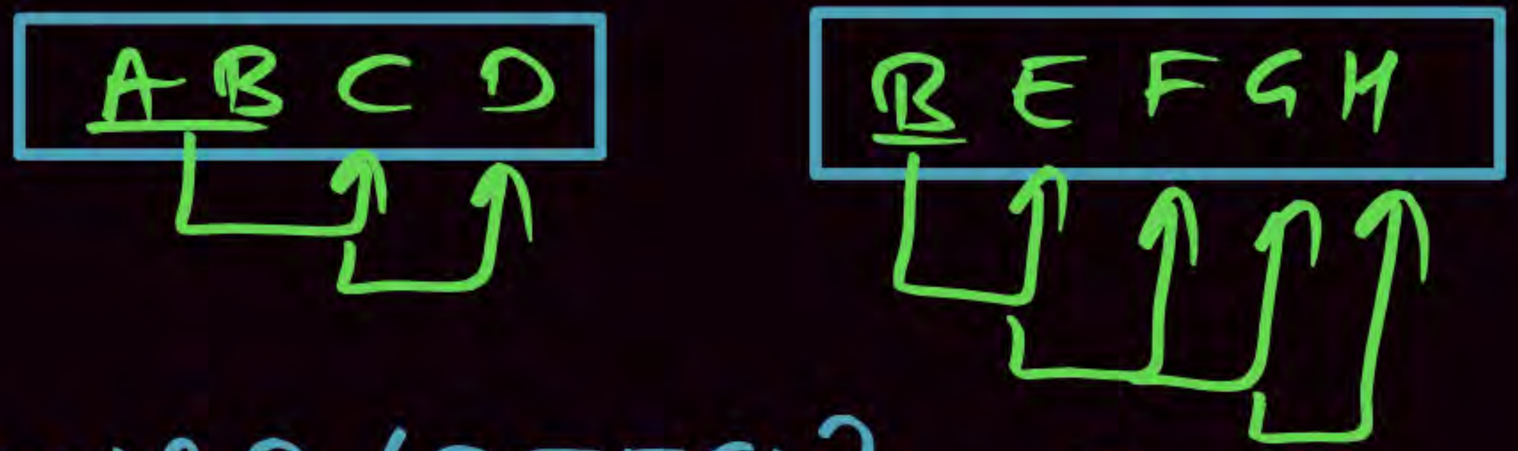
A	B	C	D

2NF Decomposition.

R(ABCD~~EFGH~~)

R_2

B	E	F	G	H



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

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

2NF + Lossless + Dependency Preserving.

2NF Decomposition

Q.2 R(ABCDE) F: [$A \rightarrow B$, $B \rightarrow E$, $C \rightarrow D$]

Decompose it into 2NF.

2NF Decomposition

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

2NF Decomposition

Q.4 $R(ABCDEF)\ \{AB \rightarrow C, C \rightarrow D, B \rightarrow EF\}$

2NF Decomposition

Q.5 $R(ABCDEFGH)\ \{AB \rightarrow C,\ C \rightarrow D,\ B \rightarrow E,\ E \rightarrow F,\ A \rightarrow GH\}$

THANK YOU!