

Atomicity, Consistency, Isolation, Durability (ACID) DBMS

DBMS v/s file system

- (1) Unstructured \rightarrow handled by operating system
- (2) File Location \rightarrow hierarchical
- (3) Allow pre-determined access to data
- (4) Data Redundancy \rightarrow not present \rightarrow duplicate
- (5) Inconsistency
- (6) Concurrent access not allowed \rightarrow RR, RW, RVD

- (1) Integrity
 - (2) Security \rightarrow Role Based
- Keys in DBMS

Keys / Superkeys

Attributes / fields \rightarrow
keys

Set of attributes or attributes that uniquely identify each record of Relation.

$$\text{max SK} = 2^n - 1$$

we add any key to sk \rightarrow also be a super key

Candidate key - SK whose proper subset is not a sk.
Minimal key - CK is min sk
 $S_2 \subset S_1, [S_2 \notin S_1]$

\rightarrow every CK is SK, vice versa is not true

Primary key - those attributes which are part of candidate key
minimal set of attribute that is req to uniquely identify all attribute in relation

Q. $R(A, B, C)$, A is CK \rightarrow no of SK?

Ans $\rightarrow [A, AB, AC, ABC]$

Functional Dependencies

determinant determines

FD: $x \rightarrow y$ \rightarrow dependent

if $[t_1.x = t_2.x]$ then $t_1.y = t_2.y$

tuples

x	y
1	2
2	3
3	5

Table

Attributes	Roll No.	Dept.	Course
a	78	CS	C1
b	60	EE	C2
a	78	CS	C2
b	60	EE	C3
c	80	IT	C2

Type of functional Dependencies

1. Trivial (Always valid) $x \rightarrow x$
2. Non-trivial $x \rightarrow y$ if $y \subset X$

Trivial $\rightarrow x$ can determine itself

$(R.No., Name) \rightarrow Name$

No need to check always valid

then it is trivial functional dependencies

$x \rightarrow y$ \rightarrow Roll No. \rightarrow Name

if $(X \cap Y = \emptyset)$ \rightarrow Nothing is common

Semi-trivial \rightarrow Roll No., Name \rightarrow Name, Marks

Need to check

3. Superkey / Multi-valued

4. Transitive

DBMS

Armstrong's Axioms / Inference Rules :-

1. Reflexivity

$$[x \rightarrow x], [x \rightarrow y \text{ if } (y \in x)]$$

2. Transitivity

$$\text{if } (x \rightarrow y \text{ \& } y \rightarrow z) \rightarrow \text{if this is true then } x \rightarrow z$$

3. Augmentation

$$\text{if } x \rightarrow y \text{ then } xA \rightarrow yA$$

4. Union

$$\text{if } x \rightarrow y \text{ \& } x \rightarrow z \text{ then } x \rightarrow yz$$

5. Decomposition / Splitting

$$\text{if } x \rightarrow yz \text{ then } x \rightarrow y \text{ \& } x \rightarrow z$$

6. Pseudo transitivity

$$\text{if } x \rightarrow y \text{ \& } yz \rightarrow A \text{ then } xz \rightarrow A$$

7. Composition

$$\text{if } x \rightarrow y \text{ \& } A \rightarrow B \text{ then } xA \rightarrow yB$$

Attribute Closure / Closure Set :-

Set of attributes where closure contains all attributes of given relation

$X^+ \rightarrow$ contains set of attributes determined by X

$$SK ABCDE^+ = \{A, B, C, D, E\}$$

$$ACD^+ = \{A, C, D, B, E\}$$

$$A \leftarrow SK$$

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

$$ABCDE^+ = \{A, B, C, D, E\}$$

$$F.D. :- \{A \rightarrow B, D \rightarrow E\}$$

Prime attributes

part of CK

$$\{A, C, D\}$$

AV on which? No

No more CK

Normalization in DBMS

Problems

- (1) Insertion Anomaly
- (2) Update Anomaly
- (3) Delete Anomaly

Solutions

decompose tables

Reduce Redundancy / duplicate data

Related to each other through foreign key.

Online transaction and analytical processing

Need

- (1) Reduce Redundancy
- (2) Save the space
- (3) Remove anomalies
- (4) Normalises NOW values
- (5) Simplify queries
- (6) used in OLAP & OLTP \rightarrow denormalised
- (7) Searching / Sorting will be easy
- (8)

(OLAP)

Online analytical processing

OLTP / Transaction

(1) First Normal Form

No multivalued attributes

- ER diagram \rightarrow Relation schema
- Entity \rightarrow attributes
- must not contain any multi-valued attributes or composite attributes

each column must contain atomic values

Rules for 1NF:

1. A col should contain value from same domain
2. Each column should have unique name
3. No adding to rows and cols.
4. No duplicates

* 123 \rightarrow atomic

1, 2, 3 \rightarrow not atomic

CS001 \rightarrow Not atomic

E.R \rightarrow relational schema [always in 1NF]
 • Separate table for multivalued attributes, and separate column for composite attributes.

(2) Second Normal Form

Rules to check whether R.D is in 1NF.

1. If it is in 1st normal form
2. No partial dependency in the relation

AB \rightarrow C
 A \rightarrow C
 B \rightarrow C

P.D Proper subset of CK

non-prime attributes

attr. which are not part of CK

eg. R(A, B, C, D, E, F)

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

AB \rightarrow C \rightarrow D \rightarrow E \rightarrow F

AP = {A, B, C, D, E}

AP = {A, B, C, D, E} X JK
 F⁺ = {F}

CK

{A, F} \rightarrow not in CK
 \rightarrow unique CK

AP = prime attr. = {A, B, C, D, E}
 non-prime attr. = {F}

for check P.D

proper subset of AP = {A, F} \rightarrow {B, C, D, E}

P.D is

Not in 2nd Normal form

means no proper subset is possible

* If CK is in single attr \rightarrow that relation is in 2nd Normal form.

* If all R is present in prime attributes \rightarrow 2nd N.F.
 \rightarrow there is no non-prime attributes

Drawbacks in 2nd Normal form
 → still have some Redundancy. (there is update anomaly)

Third Normal form → Remove all anomaly (insert, update & delete)

(1) It is in 2nd NF

(2) No transitive dependency for non-prime attr.

Transitive depend. $\boxed{NPA \rightarrow NPA}$

$\boxed{PA \rightarrow NPA}$ not in 2NF

eg. $R(A, B, C, D, E, F)$

$F, D \rightarrow \{AB \rightarrow CDEF, BD \rightarrow F\}$

$AB \neq \emptyset \neq F^+ = \{A, B, C, D, E, F\}$

Not able to handle the cases when have multiple CK & all CKs are overlapping

$X \rightarrow Y \rightarrow Z$
 CK PK
 no unique non-prime
 non-prime
 SK $AB^+ = \{A, B\}$ SK

2 Method:

A table is in 3NF if & only if for each of its non-trivial functional dependencies at least one of the following cond. holds: —

1. LHS is SK
2. RHS is prime attributes

$\boxed{AB^+ \rightarrow CK}$

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

$NPA \rightarrow \{C, D, E, F\}$

$AB \rightarrow CDEF$
 $PA \rightarrow X$ T.D.

$BD \rightarrow F$ T.D.
 NPA NPA

Hence, Not in 3rd Normal form

BCNF (Boyce - Codd Normal form)

→ If it is in 3NF

→ for each non-trivial F.D $X \rightarrow Y$
 X must be super key or CK

Q. find No. of NP?

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

$F \rightarrow \{AB \rightarrow CDE, D \rightarrow A\}$

$AB \neq \emptyset \neq F^+ = \{A, B, C, D, E\}$

$A \rightarrow CDE, A$
 $B \rightarrow CDE, A, B$
 $\{A\} \rightarrow X$ SK
 $\{B\}$

$\boxed{CK \in AB^+ = \{A, B\}}$
 $CK \in DB^+ = \{D\}$
 $\{D\} \rightarrow X$ SK

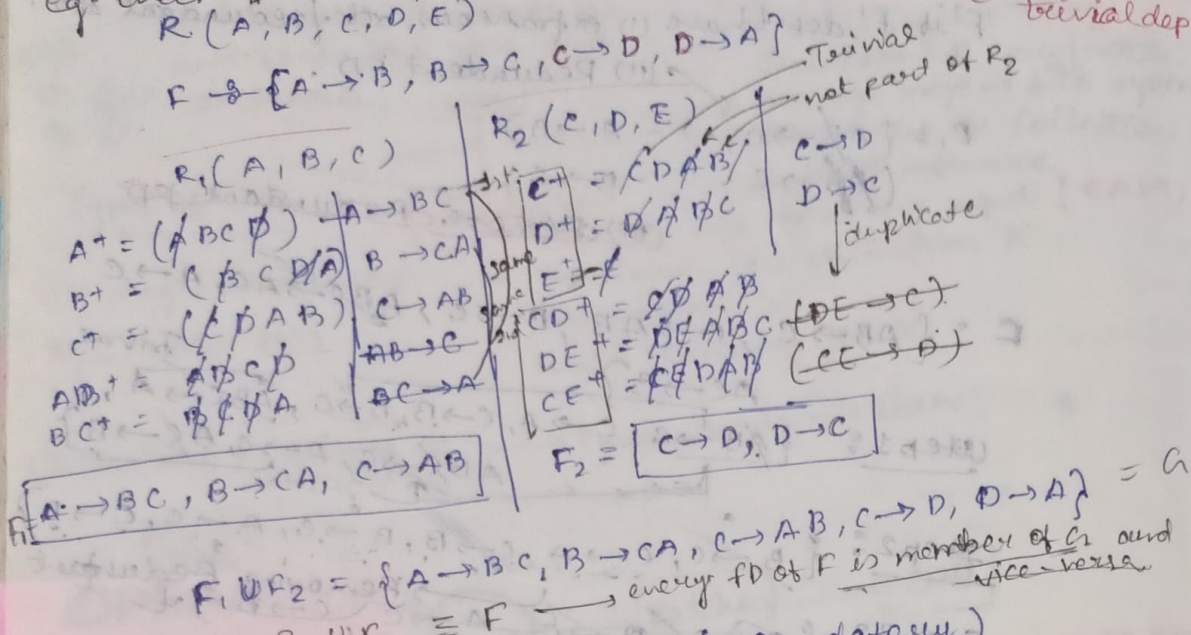
$PA \rightarrow \{A, B\}$
 $D \rightarrow \{D, B\}$

$PA \rightarrow \{A, B, D\}$
 $NPA \rightarrow \{C, E\}$

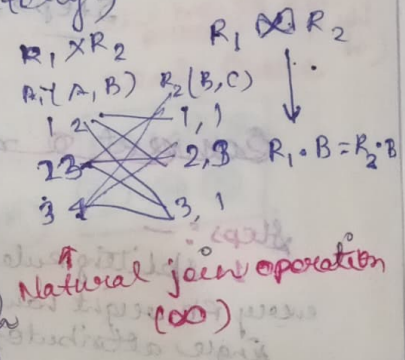
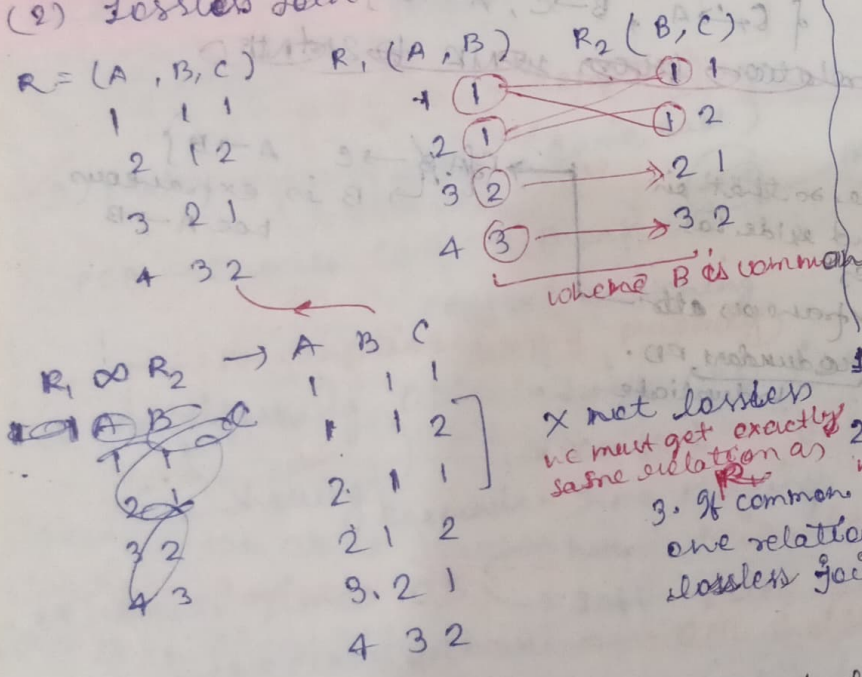
BCNF

3NF → $AB \rightarrow PA$ T.D.
 $D \rightarrow PA$ T.D.

(1) Dependency Preserving Decomposition
 eg: check DP Decomposition
 $R(A, B, C, D, E)$
 $F = \{A \rightarrow B, B \rightarrow C, C \rightarrow D, D \rightarrow A\}$
 (2) Loss less
 (will not take trivial dep.)



check $F_1 \cup F_2 \equiv F$
 Mandatory Property
 (2) Lossless Join Decomposition (Mandatory)



- 1. $attr(R_1) \cup attr(R_2) = attr(R)$
- 2. $attr(R_1) \cap attr(R_2) \neq \emptyset$
- 3. If common attr is sk of atleast one relation then that would be lossless join decomposition.

* Loss less is compulsory but Dependency preserving is not necessary.
 * Sometimes, it is not possible to get dependency preserving decomposition in case BCNF decomp.
 * upto 3NF is always possible to get dependency preserving decomposition.

→ If R is decomposed into R_1 & R_2 then decomposition is lossless if -
 1. $attr(R_1) \cup attr(R_2) = attr(R)$
 2. $attr(R_1) \cap attr(R_2) \neq \emptyset$
 3. $attr(R_1) \cap attr(R_2) = attr(R_1)$
 4. $attr(R_1) \cap attr(R_2) = attr(R_2)$

DB
 OS → memory
 APPS → map
 CN → vol
 COA → mem

Canonical Cover / Minimal Cover / Irreducible set of FDs.

F

F' if F' don't have (1) extraneous attr. / redundant attr.

(11) Redundant FD.

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

$(A \rightarrow C, B \rightarrow C)$

(11) Remove Redundant FD

$F : \{AB \rightarrow C, C \rightarrow AB, B \rightarrow C, ABC \rightarrow AC, A \rightarrow C$

$AC \rightarrow B\}$ Trivial

Step 1: $\{A \rightarrow C, C \rightarrow A, C \rightarrow B, B \rightarrow C, ABC \rightarrow A, A \rightarrow B\}$

Step 2: $\{B \rightarrow C, C \rightarrow A, C \rightarrow B, B \rightarrow C, A \rightarrow C, C \rightarrow B\}$

Step 3: $\{C \rightarrow A, B \rightarrow C, A \rightarrow C\}$ Removing Redundant

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

Convert a relation from 4NF to 2NF

Steps:-

1. Splitting Rule so that in every FD right hand side has single attributes

2. Remove extraneous attr.

3. Remove redundant FD.

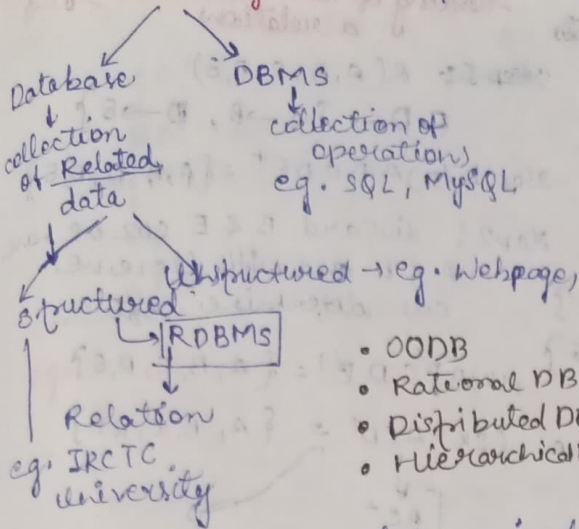
↳ duplicate

$\{A \rightarrow C, A \rightarrow B\}$
 $A \rightarrow B$ is extraneous
 bec $A \rightarrow B$

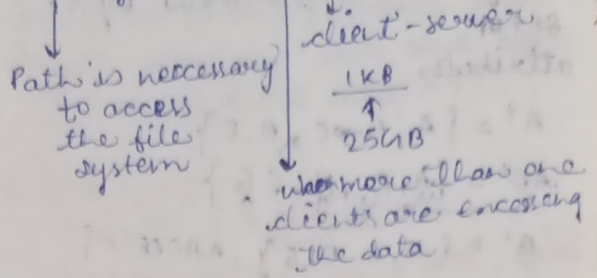
5 1 1
 1 1 5
 5 1 5
 1 5 5
 5 5 5

* on splitting properly we can only split right hand side not left side

Database system \rightarrow comprised of 2 things



File System, Vs DBMS



- OODB
- Rational DB
- Distributed DB
- Hierarchical DB

functional dependencies

\hookrightarrow Every game X there should be game Y

Key To access each data uniquely.

key / Superkey is an attribute, or set of attributes that uniquely identify each record of Relation.

Max superkey \rightarrow only a theoretical concept

(Roll.no, name | dept | course | marks)

$$= 5C_1 + 5C_2 + 5C_3 + 5C_4 + 5C_5$$

$$= 31 = 2^5 - 1$$

If (B, C, D) is not a superkey then subset of it cannot be a superkey.

2. Candidate key \rightarrow sk where proper subsets is not a sk.

eg. If sk A, AB, AC, ABC, BC
ABC \rightarrow proper subsets \rightarrow AB, A, B, C, BC, AC

There are 3 SK.

so, ABC is not a candidate key.

AC \rightarrow proper subsets \rightarrow A, C is a SK.

Minimal key so, not a candidate key

$A \rightarrow \{A\}$ so it is a candidate key.

$BC \rightarrow \{B\}, \{C\}$ \leftarrow proper subsets not a superkey.

so a candidate key.

$\{A\}, \{BC\} \rightarrow$ or Minimal Key.

Roll. No.	Name	Marks	Dep.	Course
1	a	78	CS	C1
2	b	80	EE	C4
3	a	78	CS	C2
4	b	60	EE	C3
5	c	80	IT	C3
6	d	80	EC	C2

R.No. \rightarrow Name

Name \rightarrow Roll.No. \times

R.No. \rightarrow Marks

Dept. \rightarrow course \times

course \rightarrow dept. \times

Marks \rightarrow dept. \times

[R.No., Name] \rightarrow Marks

[Name, Marks] \rightarrow dept.

[Name, Marks] \rightarrow [dept., course] \times

Types \rightarrow

1. Trivial
2. Non-Trivial
3. Multi-valued
4. Transitive

Attribute closure (closed set)

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

FD $\{A \rightarrow B, B \rightarrow C, C \rightarrow D, D \rightarrow E\}$

$A \rightarrow C$

$A \rightarrow A$

$A \rightarrow D$

$A \rightarrow E$

$A \rightarrow ABCDE$

Similarly,

$B \rightarrow BCDE$

$C \rightarrow CDE$

$D \rightarrow DE$

$E \rightarrow E$

X^+ contains set of attributes determined by X .
 ↓
 set of attributes

$A^+ = \{A, B, C, D, E\}$
 $AD^+ = \{A, D, B, C, E\}$
 $B^+ = \{B, C, D, E\}, ABCDE^+ = \{A, B, C, D, E\}$
 $CD^+ = \{C, D, E\}, AB^+ = \{A, B, C, D, E\}$

1. Superkey when FD is given — set of attributes whose closure contains all relation of given relation

$A^+, AD^+, AB^+ \rightarrow A$ anything
 So, No. of superkeys possible
 $R(A, B, C, D, E)$
 ↓
 no. of relation = $2^4 = 16$
 So, 16 → superkeys is possible

2. Candidate key

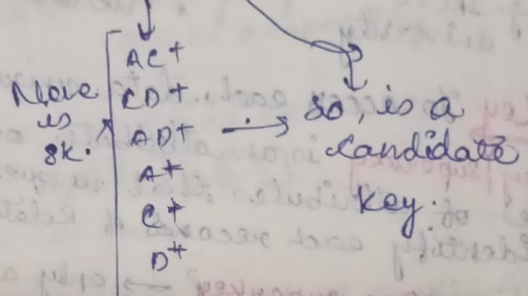
A^+ → only superkey which is a candidate key

eg. $R(A, B, C, D, E)$
 $FD = \{A \rightarrow B, D \rightarrow E\}$
 To find sk,
 → find closure
 $A^+ = \{A, B\}$ ✗
 $BC^+ = \{B, C\}$ ✗
 $AB^+ = \{A, B, C, D, E\}$ ✓
 $AD^+ = \{A, B, D, E\}$ ✗
 $ACD^+ = \{A, C, D, E, B\}$ ✓
 contains all attributes of the relation
 $\{A, C, D, B, E\}$
 To find is it a ck?
 all proper subsets → A AC AD CD
 so it is a candidate key.

Finding of all candidate keys of a relation

step 1: $R(A, B, C, D, E)$
 $FD \rightarrow \{A \rightarrow B, D \rightarrow E\}$
 step 1: (sk) $ABCDE^+ = \{A, B, C, D, E\}$
 step 2: discard B & E coz we have A & D and with them we can determine B & E

$(sk) ABCDE^+ = \{A, B, C, D, E\}$
 $(sk) ACD^+ = \{A, B, C, D, E\}$



step 4: For check whether there is more ck or not.

prime attributes → attributes which are part of ck.
 $\{A, C, D\}$

If any of these attributes are present on the right side of F. Relation.

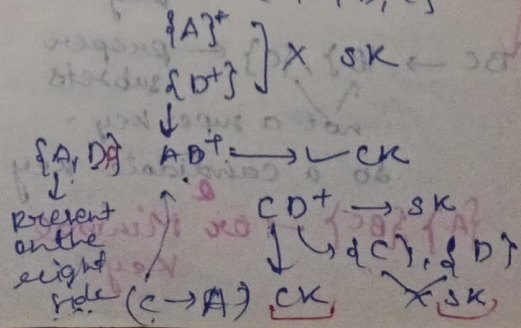
then there is more ck

else if none of prime attributes are present on right side of func dep. then there is no more ck.

eg 2: $R(A, B, C, D)$

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

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



prime attributes $\{A, D, C\}$
 $CD^+ \rightarrow SK$
 $BD^+ \rightarrow SK$
 $\{B, D\} \rightarrow XSK$
 CK
 AD^+

Normalisation
 ↳ process of making the table free from insert, updates and delete anomaly and save space from redundant data.

eg. 3. $R(A, B, C, D, E, F)$
 $P.D = \{AB \rightarrow C, C \rightarrow DE, E \rightarrow F\}$
 $D \rightarrow A, C \rightarrow B$
 $AB C D E F^+ = \{A, B, C, D, E, F\}$
 $AB F^+ = \{A, B, F, C, D, E\}$
 $AB \rightarrow F \quad AB^+ = \{A, B, C, D, E, F\}$
 (SK)
 $for(CK) \quad find \{A^+\} = \{A\} \rightarrow XSK$
 $\{B^+\} = \{B\} \rightarrow XSK$
 $so \quad AB \rightarrow CK$
 prime attributes = $\{A, B\}$
 $D \rightarrow A \rightarrow$ prevention right side
 $DB^+ \rightarrow SK$
 $\Rightarrow \{D\} \rightarrow XSK$
 $\{B\} \rightarrow XSK$

check NF:
 $R(A, B, C, D, E)$
 $FD = \{A \rightarrow BCDE, BC \rightarrow ACE, D \rightarrow E\}$
 $AB C D E^+ = \{A, B, C, D, E\} \Rightarrow A^+ = \{A, B, C, D, E\}$
 $BC = \{ACE, B, D\}$
 $\{B\} \rightarrow XSK$
 $\{C\} \rightarrow XSK$
 $A = CK$
 $BC = CK$

BCNF: X must be super key
 $A \rightarrow SK$
 $BC \rightarrow SK$
 $D \rightarrow XSK$
 Not in BCNF
 3NF: $A \checkmark$
 $BC \checkmark$
 $D \times$
 2NF: $A \checkmark$
 $BC \checkmark$
 $D \checkmark$

prime $A = \{A, B, D\}$
 $for \quad C \rightarrow B$
 $AC^+ = SK$
 $\Rightarrow \{A\} \Rightarrow XSK$
 $\{C\} = \{D, E, A, B, C, F\} \rightarrow XSK$
 $AC^+ = XSK$
 $C \rightarrow SK \rightarrow CK$
 $P.A = \{A, B, D, C\}$
 $\downarrow C \rightarrow D$
 $CB^+ = \{C\} \rightarrow SK$
 XCK
 $CK = \{AB, BD, C\}$
 $P.A = \{A, B, D, C\}$
 $Non-P.A = \{E, F\}$

lecture - 14 example
 $R(A, B, C, D, E)$
 $F = \{AB \rightarrow CDE, D \rightarrow BE\}$
 $CK = AB, AD$
 $PA = (A, B, D)$
 $(X) NPA = (C, E)$
 BCNF: \checkmark X
 3NF: \checkmark X
 $D (PA \cup NPA) = NPA$
 2NF: \checkmark X
 : INF

check dependency preserving decomposition

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

$F = \{A \rightarrow BCD, B \rightarrow AE, BC \rightarrow AED, D \rightarrow E, C \rightarrow DE\}$

$R_1(A, B) \quad R_2(B, C) \quad R_3(C, D, E)$

for R_1
 $\rightarrow A \rightarrow$ check in F

$A^+ = \{A, B, C, D, E\} \Rightarrow A \rightarrow B$
 Trivial \downarrow not part of R_1

$B^+ = \{B, A, E, C, D\}$
 Trivial \downarrow not part of R_1

$B^+ = \{A\} \rightarrow B \rightarrow A$

pick 2 \rightarrow put here it will become trivial
 (AB)

for $R_2 (BC)$

$B^+ = \{B, A, E, C, D\} \Rightarrow B \rightarrow C$

$C^+ = \{C, D, E\} \rightarrow \{ \}$

2 (BC) \rightarrow will become trivial

for $R_3 (C, D, E)$

$C^+ = \{C, D, E\} \rightarrow C \rightarrow DE$

$D^+ = \{D, E\} \rightarrow D \rightarrow E$

$E^+ = E \rightarrow$ duplicate

$CD^+ = \{C, D, E\} \rightarrow CD \rightarrow E$

$F_1 \cup F_2 \cup F_3 = \{A \rightarrow B, B \rightarrow A, B \rightarrow C, C \rightarrow DE, D \rightarrow E\}$

Now will check each F.D of R as member of G

for $A \rightarrow BCD \checkmark \rightarrow A^+ = \{A, B, C, D, E\}$

$B \rightarrow AE \checkmark \rightarrow B^+ = \{A, C, D, E\}$

$BC \rightarrow AED \checkmark \rightarrow BC^+ = \{B, C, A, D, E\}$

$D \rightarrow E \checkmark$

$C \rightarrow DE \checkmark$

ques on. lossless join decomposition
 L10: [Jenny]