Week 6

1 NF :-

2 NF :-

3 NF :-

BCNF :-

4 NF :-

Normalization

1NF: First Normal Form

- A relation is in First Normal Form if and only if all underlying domains contain atomic values only (doesn't have multivalued attributes (MVA))
- **STUDENT(Sid, Sname, Cname)**

Students		
SID	**Sname**	**Cname**
S1	A	C,C++
S2	B	C++, DB
S3	A	DB
SID : Primary Key		

MVA exists ⇒ Not in 1NF

Students		
SID	**Sname**	**Cname**
S1	A	C
S1	A	C++
S2	B	C++
S2	B	DB
S3	A	DB
SID, Cname : Primary Key		

No MVA ⇒ In 1NF

Source: http://www.edugrabs.com/normal-forms/#fnf

2NF: Second Normal Form

- Relation ***R*** is in Second Normal Form (2NF) only iff :
  - ***R*** is in 1NF and
  - ***R*** contains no Partial Dependency

R (A B C D)

PA :- attribute of C.K

NPA :- Not part of C.K

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

D → NPA

AB is C.K, PA :→ A & B

NPA → C & D

Partial Dependency:

Let ***R*** be a relational Schema and ***X*, *Y*, *A*** be the attribute sets over ***R*** where ***X* :** Any Candidate Key, ***Y* :** Proper Subset of Candidate Key, and ***A* :** Non Prime Attribute

If $Y \rightarrow A$ exists in R, then R is not in 2NF.

$(Y \rightarrow A)$ is a Partial dependency only if

- ***Y***: Proper subset of Candidate Key
- ***A***: Non Prime Attribute

A **prime attribute** *of a relation is an attribute that is a part of a candidate key of the relation*

Y → A → NPA

C → D

AB → C

C.K

BCNF

3NF: Third Normal Form

Let R be the relational schema.

- [E. F. Codd,1971] R is in 3NF only if:
  - R should be in 2NF
  - R should not contain transitive dependencies (OR, Every non-prime attribute of R is non-transitively dependent on every key of R)
- [Carlo Zaniolo, 1982] Alternately, R is in 3NF iff for each of its functional dependencies $X \to A$, at least one of the following conditions holds:
  - X contains A (that is, A is a subset of X, meaning $X \to A$ is trivial functional dependency), or
  - X is a superkey, or
  - Every element of $A - X$, the set difference between A and X, is a *prime attribute* (i.e., each attribute in $A - X$ is contained in some candidate key)
- [Simple Statement] A relational schema $\boldsymbol{R}$ is in 3NF if for every FD $\boldsymbol{X} \to \boldsymbol{A}$ associated with R either
  - $\boldsymbol{A} \subseteq \boldsymbol{X}$ (that is, the FD is trivial) or
  - $\boldsymbol{X}$ is a superkey of $\boldsymbol{R}$ or
  - $\boldsymbol{A}$ is part of some candidate key (not just superkey!)
- **A relation in 3NF is naturally in 2NF**

3 NF

→ It should be in 1 NF & 2 NF

→ No transitive Dependency

T.D → NPA → NPA ✗

1 C.K

2 PA & NPA

BCNF

→ It should be in 3NF

→ No P.D

→ No TD

→ No MVA

⇒ LHS →

(key)$^+$ →

⇒ LHS of every FD must be Super key

assume that 1NF

1NF

No MVA

2) Consider the relational schema $\mathbf{R}(A, B, C, D)$, where the domains of A, B, C and D include only atomic values. Identify the sets of functional dependencies satisfied by $\mathbf{R}$ such that $\mathbf{R}$ is in BCNF. **2 points**

$(ABC)^+ = \{ABCD\} = R$

AD

- [ ] FD: $\{ABC \to D, AD \to BC\}$, $AB \to C\}$

NF :-

- [ ] FD: $\{AB \to CD, CD \to B, D \to A\}$ Not in BCNF

$(LHS)^+ \to \{ \quad \}$

C

- [ ] FD: $\{A \to B, B \to C, C \to D, D \to A\}$

$(AB)^+ = ABCD$

- [ ] FD: $\{AB \to D, D \to B, D \to C, C \to A\}$

$CD \to CDBA$

$(D)^+ = \{DA\}$

5) A database designer observes that a relation R is in 2NF, and has transitive functional dependencies. Relation R is then decomposed into relations R1 and R2, such that the decomposition is 3NF and lossless. Now, the designer observes that R2 is in 3NF but not in BCNF. Relation R2 is further decomposed into R3 and R4 such that the decomposition is in BCNF, and it satisfies the following conditions: ***3 points***

Not in 3 NF

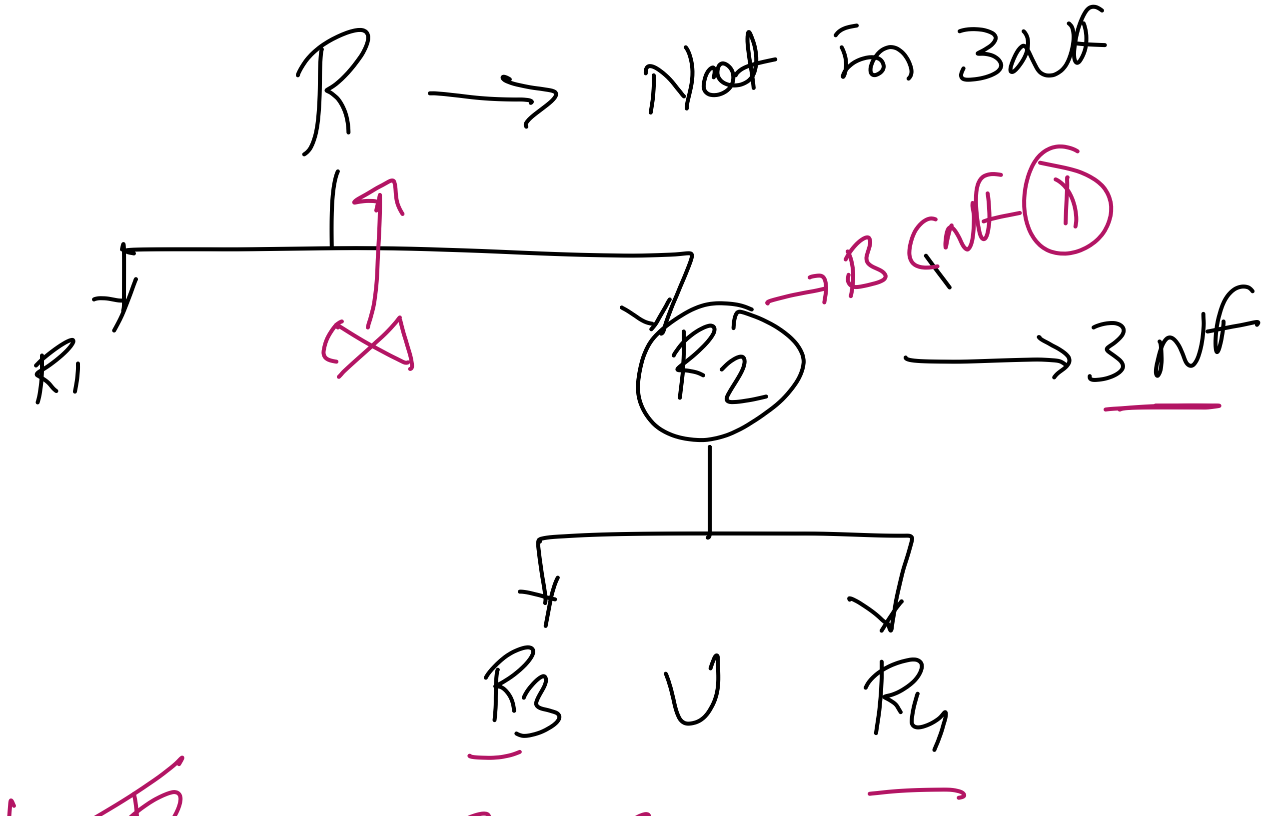
- $R2 = R3 \cup R4$
- $R3 \cap R4 \rightarrow R3$

$R3 \cup R4 = R2$ lossless

Choose the correct options.

③

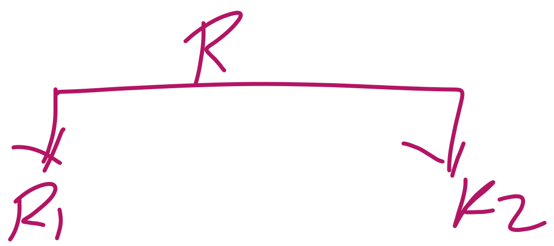
- [ ] The decomposition of R2 into R3 and R4 is lossy.
- [ ] The decomposition of R2 into R3 and R4 is lossless.
- [ ] $R = R1 \bowtie R2$
- [ ] The number of tuples in R2 is more than the number of tuples in $(R3 \bowtie R4)$



$R_2 = R_3 \cup R_4$

$R_3 \cap R_4 \rightarrow R_3$

⇒ lossless



① $R_1 \cup R_2 = R$

② $R_1 \cap R_2 \neq \phi$

③ $R_1 \cap R_2 \rightarrow R_1$ or R_2

→ W6 A4

$\boxed{\text{PA Q10}}$

MVD: Definition

- Let R be a relation schema and let $\alpha \subseteq R$ and $\beta \subseteq R$. The **multivalued dependency**

  $\alpha \twoheadrightarrow \beta$

  holds on R if in any legal relation $r(R)$, for all pairs for tuples t_1 and t_2 in r such that $t_1[\alpha] = t_2\ [\alpha]$, there exist tuples t_3 and t_4 in r such that:

$$t_1[\alpha] = t_2\ [\alpha] = t_3\ [\alpha] = t_4\ [\alpha]$$
$$t_3[\beta] = t_1\ [\beta]$$
$$t_3[R - \beta] = t_2[R - \beta]$$
$$t_4\ [\beta] = t_2[\beta]$$
$$t_4[R - \beta] = t_1[R - \beta]$$

$\alpha \twoheadrightarrow \beta$

$t_1(\alpha)$
$t_2(\alpha)$

Example: A relation of university courses, the books recommended for the course, and the lecturers who will be teaching the course:

- **course $\twoheadrightarrow$ book**
- **course $\twoheadrightarrow$ lecturer**

mvD

Test: **course $\twoheadrightarrow$ book**

Course	Book	Lecturer	Tuples
AHA	Silberschatz	John D	t1
AHA	Nederpelt	William M	t2
AHA	Silberschatz	William M	t3
AHA	Nederpelt	John D	t4
AHA	Silberschatz	Christian G	
AHA	Nederpelt	Christian G	
OSO	Silberschatz	John D	
OSO	Silberschatz	William M	

R = (course, Book, lecturer)
β = Book
R − β = course, lec
Lec → Book

AHA Nede Willi
AHA Silb John

check for all unique α values

⟹ atleast 3 attributes
→ any 2 must be independent attributes

X ⟹ course, R
Y = Book
X →→ Lecturer, lec

Fourth Normal Form

- A relation schema R is in **4NF** with respect to a set D of functional and multivalued dependencies if for all multivalued dependencies in D^+ of the form $\alpha \twoheadrightarrow \beta$, where $\alpha \subseteq$ R and $\beta \subseteq$ R, at least one of the following hold:
  - $\alpha \twoheadrightarrow \beta$ is trivial (that is, $\beta \subseteq \alpha$ or $\alpha \cup \beta =$ R)
  - α is a superkey for schema R   BCNF
- If a relation is in 4NF, then it is in BCNF

No MVD
→ BCNF

A B → B
α
AB
$\beta = B$
$\beta \subseteq \alpha$
AB = R
α B D

A
B → Index C

4NF → It should be BCNF
→ No MVD

MVD . Axioms

16) Consider a relation $\mathbf{R}(A, B, C, D, E)$ with the following multivalued dependencies:

$A \twoheadrightarrow B$

$B \twoheadrightarrow D$

$C \twoheadrightarrow D$

Suppose relation $\mathbf{R}$ contains the tuples (0, 1, 2, 3, 4) and (0, 5, 6, 7, 8). Which of the following tuple(s) must also be in $\mathbf{R}$?

- ☐ (0, 1, 2, 7, 8)
- ☐ ✓ (0, 5, 6, 3, 8)
- ☐ ✓ (0, 1, 6, 7, 8)
- ☐ (0, 1, 6, 3, 4)

$A \twoheadrightarrow B$

$B \twoheadrightarrow D$

A	B	C	D	E
0	1	2	3	4
0	5	6	7	8
0	5	2	3	4
0	1	6	7	8
0				

8

A	B	C	D	E
0	1	2	3 (7)	4
0	1	6	7 (3)	8
0	1	2	7	4
0	1	6	3	8

A	B	C	D	F
0	5	6	7 (3)	8
0	5	2	3 (7)	4
0	5	6	3	8
0	5	2	7	4

$\Rightarrow \quad X \twoheadrightarrow Y$, then $[X \twoheadrightarrow (R - (X \cup Y))]$

$A \twoheadrightarrow B$, $\quad A \to B$

$A \twoheadrightarrow B$ ✗ $\quad A \to B$

$A \to B$

A	B	C
a1	2	c1
~~a1~~	2	c2
a1	2	c2
a1		
a1	2	c2

6) Consider the relational schema: ***2 points***

MVD = Intern_code →→ Project
Intern_code →→ hobby
FD = Intern_code → Intern_name

$\mathbf{Intern}(intern\_code, intern\_name, project, hobby)$.

An intern can work in several projects and can have several hobbies. However, it maintains the FD: $intern\_code \rightarrow intern\_name$.

→ BCNF → No MVD

C.K = {Intern_code, Project, hobby}

Identify the most appropriate 4NF decomposition for the given schema.

- $\mathbf{R1}(intern\_code, intern\_name, project, hobby), \mathbf{R2}(intern\_code, project, hobby)$
- $\mathbf{R1}(intern\_code, intern\_name, project), \mathbf{R2}(intern\_code, hobby)$
- $\mathbf{R1}(intern\_code, intern\_name, hobby), \mathbf{R2}(intern\_code, project)$
- $\mathbf{R1}(intern\_code, intern\_name), \mathbf{R2}(intern\_code, project), \mathbf{R3}(intern\_code, hobby)$

⇒ atleast 3 attribute

9) Consider the relational schema: **2 points**

prescription$(doctor\_id, doctor\_name, patient\_id, patient\_name, medicine\_id, medicine\_name)$, where the domains of all the attributes consist of atomic values. Consider the following FDs for the relation $department$.

lossless

2 ① $R_1 \cup R_2 = R$

2 ② $R_1 \cap R_2 \neq \phi$, ③ $R_1 \cap R_2 \rightarrow R_1$ or R_2

$$doctor\_id \rightarrow doctor\_name$$
$$patient\_id \rightarrow patient\_name,$$
$$medicine\_id \rightarrow medicine\_name,$$
$$doctor\_id \rightarrow\rightarrow patient\_id,$$
$$doctor\_id \rightarrow\rightarrow medicine\_id$$

F.D

3 MVD

From among the decompositions given, identify the one that is in 4NF.

MVD

- R_1 $(doctor\_id, doctor\_name)$, BCNF
  R_2 $(patient\_id, patient\_name)$,
  R_3 $(medicine\_id, medicine\_name)$,
- $(doctor\_id, doctor\_name, patient\_id, patient\_name)$,
  $(doctor\_id, doctor\_name, medicine\_id, medicine\_name)$
- $(doctor\_id, doctor\_name)$, $R_1 \cup R_2 \cup R_3 = R$
  $(patient\_id, patient\_name)$,
  $(medicine\_id, medicine\_name)$,
  $(doctor\_id, patient\_id, medicine\_id)$
  mvd
- $(doctor\_id, doctor\_name)$,
  $(patient\_id, patient\_name)$,
  $(medicine\_id, medicine\_name)$,
  $(doctor\_id, patient\_id)$,
  $(doctor\_id, medicine\_id)$