

# Homework 2

Yepei Zhang 912465089

February 3, 2016

I.

(a).  $R(A, B, C, D, E)$   $AB \rightarrow E, B \rightarrow C, CE \rightarrow A, AE \rightarrow D$   
 $A^+ = \{A\}$   $B^+ = \{B, C\}$   $C^+ = \{C\}$   $D^+ = \{D\}$   $E^+ = \{E\}$   $AB^+ = \{A, B, C, D, E\}$   
 $AC^+ = \{A, C\}$   $AD^+ = \{A, D\}$   $AE^+ = \{A, D, E\}$   $BC^+ = \{BC\}$   
 $BD^+ = \{B, C, D\}$   $BE^+ = \{A, B, C, D, E\}$   $CD^+ = \{C, D\}$   $CE^+ = \{A, C, D, E\}$   
 $DE^+ = \{D, E\}$   $ABC^+ = \{A, B, C, D, E\}$   $ABD^+ = \{A, B, C, D, E\}$   
 $ABE^+ = \{A, B, C, D, E\}$   $ACD^+ = \{A, C, D\}$   $ACE^+ = \{A, C, D, E\}$   
 $ADE^+ = \{A, D, E\}$   $BCD^+ = \{B, C, D\}$   $BCE^+ = \{A, B, C, D, E\}$   
 $BDE^+ = \{A, B, C, D, E\}$   $CDE^+ = \{A, C, D, E\}$   $ABCD^+ = \{A, B, C, D, E\}$   
 $BCDE^+ = \{A, B, C, D, E\}$

keys:  $\{A, B\}$   $\{B, E\}$

(b).  $AB^+ = \{A, B, C\}$ , by using FDs except  $AB \rightarrow E$ , this closure doesn't include E. Thus,  $AB \rightarrow E$  is not implied by other three FDs.  
Similarly, we can get  $B^+ = \{B\}$  doesn't include C,  $CE^+ = \{C, E\}$  doesn't include A,  $AE^+ = \{A, E\}$  doesn't include D. Thus, the set of the FDs is a minimal basis.

(c). FDs:  $AB \rightarrow E, B \rightarrow C, CE \rightarrow A, AE \rightarrow D, AB \rightarrow C, AB \rightarrow D, BE \rightarrow A, BE \rightarrow C, BE \rightarrow D, BD \rightarrow C, CE \rightarrow D, ABC \rightarrow D, ABC \rightarrow E, ABD \rightarrow C, ABD \rightarrow E, ABE \rightarrow C, ABE \rightarrow D, ACE \rightarrow D, BCE \rightarrow A, BCE \rightarrow D, BDE \rightarrow A, BDE \rightarrow C, CDE \rightarrow A, ABCD \rightarrow E, BCDE \rightarrow A$   
BCNF violations:  $B \rightarrow C, CE \rightarrow A, AE \rightarrow D, BD \rightarrow C, CE \rightarrow D, ACE \rightarrow D, CDE \rightarrow A$

(d). Start with the BCNF violation  $B \rightarrow C$ , we get  $R_1(B, C)$ , with key  $\{B\}$ , FD  $B \rightarrow C$  which is a minimal basis.

$R_2(A, B, D, E)$ , key  $\{A, B\}$ , FDs  $AB \rightarrow E$ ,  $AE \rightarrow D$ , the second FD violates BCNF, continue decomposing.

$R_{21}(A, E, D)$ , key  $\{A, E\}$ , FD  $AE \rightarrow D$  which is a minimal basis and doesn't violate BCNF.

$R_{22}(A, B, E)$ , key  $\{A, B\}$ , FD  $AB \rightarrow E$  which is a minimal basis and doesn't violate BCNF.

Thus, the decomposed relations:  $\{B, C\}$ ,  $\{A, E, B\}$ ,  $\{A, E, D\}$

(e). 3NF violations:  $B \rightarrow C$ ,  $AE \rightarrow D$ ,  $BD \rightarrow C$ ,  $CE \rightarrow D$ ,  $ACE \rightarrow D$

(f). We can see from (b) that the given set FDs is a minimal basis. Thus, the 3NF decomposition is  $S_1(A, B, E)$ ,  $S_2(B, C)$ ,  $S_3(A, C, E)$ ,  $S_4(A, D, E)$ . Here,  $S_1(A, B, E)$  includes the key  $\{A, B\}$ , thus, there is no need adding another relation.

(g). From (c), we have known all FDs and BCNF violations. According to FD promotion, we know some of the 4NF MVD violations:  $B \twoheadrightarrow C$ ,  $CE \twoheadrightarrow A$ ,  $AE \twoheadrightarrow D$ ,  $BD \twoheadrightarrow C$ ,  $CE \twoheadrightarrow D$ ,  $ACE \twoheadrightarrow D$ ,  $CDE \twoheadrightarrow A$

Then, according to complementation rule, we get other 4NF MVD violations.  $B \twoheadrightarrow ADE$ ,  $CE \twoheadrightarrow BD$ ,  $AE \twoheadrightarrow BC$ ,  $BD \twoheadrightarrow AE$ ,  $CE \twoheadrightarrow AB$ ,  $ACE \twoheadrightarrow B$ ,  $CDE \twoheadrightarrow B$

(h). Start with the violation  $B \twoheadrightarrow C$ , we get  $R_1(B, C)$ . key:  $\{B\}$ , FD:  $B \rightarrow C$ , MVD:  $B \twoheadrightarrow C$ , there is no 4NF violation.

$R_2(A, B, D, E)$ . key  $\{A, B\}$ , FDs:  $AB \rightarrow E$ ,  $AE \rightarrow D$ , MVDs:  $AB \twoheadrightarrow E$ ,  $AE \twoheadrightarrow D$ , etc., one of the violation is  $AE \twoheadrightarrow D$ .

$R_{21}(A, D, E)$ . key  $\{A, E\}$ , FD:  $AE \rightarrow D$ , MVD:  $AE \twoheadrightarrow D$

We can use chase test to check if  $A \twoheadrightarrow D$  holds.

A	D	E
a	d	e <sub>1</sub>
a	d <sub>1</sub>	e

However, there is no useful FD we can apply further to make it agree on D column. Thus,  $A \twoheadrightarrow D$  doesn't hold. Similarly,  $A \twoheadrightarrow E$  doesn't hold.

$R_{22}(A, B, E)$ . key  $\{A, B\}$ , FD:  $AB \rightarrow E$ , MVD:  $AB \twoheadrightarrow E$

If we use chase test to check if  $A \twoheadrightarrow E$  or  $B \twoheadrightarrow E$  hold, similar situation as  $R_{21}$  shows that there is no more FD making any change to make it agree on some column. Thus, the 4NF decomposition:  $\{B, C\}$ ,  $\{A, D, E\}$ ,  $\{A, B, E\}$ . FDs and MVDs are listed above.

(i).  $S(A, B, C)$   $A^+ = \{A\}$ ,  $B^+ = \{B, C\}$ ,  $C^+ = \{C\}$ ,  $AB^+ = \{A, B, C\}$ ,  $BC^+ = \{B, C\}$ ,  $AC^+ = \{A, C\}$  clearly the key is  $\{A, B\}$ ,  $B \rightarrow C$  holds

Thus, the nontrivial FDs:  $B \rightarrow C$ ,  $AB \rightarrow C$  nontrivial MVDs:  $B \twoheadrightarrow C$ ,  $B \twoheadrightarrow A$

Apply chase test to check if there is other nontrivial MVDs.

If we want to check if  $A \twoheadrightarrow C$  holds:

A	B	C
a	b <sub>1</sub>	c
a	b	c <sub>1</sub>

We can see that given FDs and MVDs cannot be applied anymore, because it agrees only on A column. However, there is no FDs or MVDs with A on the left hand side alone. Similarly, we cannot find  $A \twoheadrightarrow B$  hold.

Then, check if  $C \twoheadrightarrow A$  holds:

A	B	C
a	b <sub>1</sub>	c
a <sub>2</sub>	b	c

We can see that given FDs and MVDs cannot be applied anymore, because it agrees only on C column. However, there is no FDs or MVDs with C on the left hand side alone. Similarly, we cannot find  $C \twoheadrightarrow B$  hold.

In brief: the nontrivial FDs:  $B \rightarrow C$ ,  $AB \rightarrow C$  nontrivial MVDs:  $B \twoheadrightarrow C$ ,  $B \twoheadrightarrow A$

2.

(a). Start with the tableau:

C	T	H	R	S	G
c <sub>1</sub>	t <sub>1</sub>	h	r <sub>1</sub>	s	g <sub>1</sub>
c <sub>2</sub>	t <sub>2</sub>	h	r <sub>2</sub>	s	g <sub>2</sub>

Apply  $HR \rightarrow R$ , then  $r_2 = r_1$

C	T	H	R	S	G
c <sub>1</sub>	t <sub>1</sub>	h	r <sub>1</sub>	s	g <sub>1</sub>
c <sub>2</sub>	t <sub>2</sub>	h	r <sub>1</sub>	s	g <sub>2</sub>

Apply  $HR \rightarrow C$ , then  $c_2 = c_1$

C	T	H	R	S	G
$c_1$	$t_1$	$h$	$r_1$	$s$	$g_1$
$c_1$	$t_2$	$h$	$r_1$	$s$	$g_2$

Apply  $CS \rightarrow G$ , then  $g_2 = g_1$

C	T	H	R	S	G
$c_1$	$t_1$	$h$	$r_1$	$s$	$g_1$
$c_1$	$t_2$	$h$	$r_1$	$s$	$g_1$

Thus two tuples now agree in the G column.  $HG \rightarrow G$  holds.

(b). Start with the tableau:

C	T	H	R	S	G
$c_1$	$t$	$h$	$r$	$s_1$	$g_1$
$c_2$	$t$	$h$	$r$	$s_2$	$g_2$

Apply  $HR \rightarrow C$ , then  $c_2 = c_1$

C	T	H	R	S	G
$c_1$	$t$	$h$	$r$	$s_1$	$g_1$
$c_1$	$t$	$h$	$r$	$s_2$	$g_2$

When we apply  $C \rightarrow T$ ,  $HT \rightarrow R$ ,  $HS \rightarrow R$ , there is nothing changed. We cannot get two tuples agree on S. So  $THR \rightarrow S$  doesn't hold.

(c). Start with the tableau:

C	T	H	R	S	G
$c$	$t$	$h$	$r_1$	$s_1$	$g_1$
$c$	$t_2$	$h$	$r$	$s_2$	$g$
$c$	$t_3$	$h_3$	$r$	$s$	$g_3$

The first two rows agree on C and H, apply  $C \rightarrow T$ , so they agree on T.  $t_2 = t$

C	T	H	R	S	G
$c$	$t$	$h$	$r_1$	$s_1$	$g_1$
$c$	$t$	$h$	$r$	$s_2$	$g$
$c$	$t_3$	$h_3$	$r$	$s$	$g_3$

Then the first two rows agree on T,H,C, apply  $HT \rightarrow R$ , then  $r_1 = r$

C	T	H	R	S	G
$c$	$t$	$h$	$r$	$s_1$	$g_1$
$c$	$t$	$h$	$r$	$s_2$	$g$
$c$	$t_3$	$h_3$	$r$	$s$	$g_3$

Then, the first two rows agree on C,T,H,R. No more change can be made because of the given FDs. There is no row that is fully unsubscripted. Thus, this decomposition doesn't have a lossless join.

(d). Firstly, check whether the set of FDs is a minimal basis.

$HR^+ = H, R$  doesn't include C after applying FDs except  $HR \rightarrow C$ . Similarly,  $HT^+ = H, T$ ,  $HS^+ = H, S$ ,  $CS^+ = C, S, T$ ,  $C^+ = C$ . Thus, the given set of FDs is a minimal basis.

Then, we can get relations:  $S_1(C, T)$ ,  $S_2(H, R, C)$ ,  $S_3(H, R, T)$ ,  $S_4(H, S, R)$ ,  $S_5(C, S, G)$ . We note that the key is  $\{H, S\}$  which is included into  $S_4(H, S, R)$ . Thus, we are done.

(e).

Decompose R into BCNF.

Begin from  $C \rightarrow T$ , we get  $R_1(C, T)$  key:  $\{C\}$  FD:  $C \rightarrow D$

$R_2(C, H, R, S, G)$ , key:  $\{H, S\}$ , FDs:  $HR \rightarrow C$ (violation),  $HS \rightarrow R$ ,  $HS \rightarrow G$ ,  $CS \rightarrow G$ (violation).

$R_{21}(H, R, C)$ , key:  $\{H, R\}$ , FDs:  $HR \rightarrow C$

$R_{22}(H, R, S, G)$ , key:  $\{H, S, G\}$ , FD:  $HS \rightarrow R$  (violation)

$R_{221}(H, S, R)$ , key:  $\{H, S\}$ , FD:  $HS \rightarrow R$

$R_{222}(H, S, G)$ , key:  $\{H, S, G\}$  no FD

Thus, BCNF decomposition is  $R_1(C, T)$ ,  $R_{21}(H, R, C)$ ,  $R_{221}(H, S, R)$ ,  $R_{222}(H, S, G)$ , keys and FDs are listed above.

Compare with 3NF decomposition,  $\{H, R, T\}$ ,  $\{C, S, G\}$  are in 3NF but not in BCNF

(f). Decompose R into 4NF.

Start with  $C \twoheadrightarrow T$ , we get  $R_1(C, T)$ , MVD:  $C \twoheadrightarrow D$

$R_2(C, H, R, S, G)$ , key:  $\{H, S\}$ , FDs:  $HR \twoheadrightarrow C$ (violation),  $HS \twoheadrightarrow R$ ,  $HS \twoheadrightarrow G$ ,  $CS \twoheadrightarrow G$ (violation).

$R_{21}(H, R, C)$ , key:  $\{H, R\}$ , MVD:  $HR \twoheadrightarrow C$

$R_{22}(H, R, S, G)$ , key:  $\{H, S, G\}$ , MVDs:  $HS \twoheadrightarrow R$  (violation),  $HS \twoheadrightarrow G$ (violation)

$R_{221}(H, S, R)$ , key:  $\{H, S\}$ , MVD:  $HS \twoheadrightarrow R$

$R_{222}(H, S, G)$ , key:  $\{H, S, G\}$

Thus, 4NF decomposition is  $R_1(C, T)$ ,  $R_{21}(H, R, C)$ ,  $R_{221}(H, S, R)$ ,  $R_{222}(H, S, G)$ , keys and MVDs are listed above.

Compare with 3NF decomposition,  $\{H, R, T\}$ ,  $\{C, S, G\}$  are in 3NF but not in 4NF

3.

Assume they are not getting multiple PhDs. To avoid the situation that some people have the same name, I set the key  $\{name, dissertation\ title, year\ of\ graduate\}$ . Academics can be students and instructors. One student have one instructor as dissertation chair, but one instructor can chair many students. That's why the relationship of "chair" is "one to many". Two instructors can be the dissertation committee members for many students. That's why the relationship of "are committee members of" is "two to many". In addition, one can have many progenies and progenitors. Thus, the relationship of "academic progeny" is "many to many". In this model, the one academic cannot play different roles at the same time, which one cannot chair himself or chair and become committee member of same person. Thus, it is a reasonable model.

