Assignment 3 – Sample Solution

Q1. [6 marks] Parts (a), (b), and (c) are 0.5 mark each meaning 1.5 marks for each of (i), (ii), (iii), and (iv).

R (ABCDEF)

i. AD → B, C → D, BC → A, B → D

a) Candidate keys are: ACEF and BCEF.

b) All FDs are BCNF violations.

c) C → D violates => decompose R into R1(CD) and R2(ABCEF)
R1 is BCNF since it has 2 attributes

For R2, FD is BC → A only, candidate key is BCEF
BC → A violates => decompose R2 into R3(BCA) and R4(BCEF)

R3 is in BCNF since FD is BC → A only and candidate key is BC and BC → A does not violate
R4 is in BCNF since no FDs

Hence the final decomposition is: CD, ABC, and BCEF

If you start with a different FD, you will get the following decompositions:
AB, BD, CD, and ACEF
or
BD, ABC, and BCEF

ii. BC → E, C → AB, AF → CD

a) Candidate keys are: AF and CF.

b) BC → E and C → AB are the BCNF violations.

c) C → AB violates => decompose R to R1(CAB) and R2(CDEF)
No more violation => final decomposition: ABC, CDEF

If you start with BC → E, your final decomposition will be BCE, CAB, and CDF.

iii. ABF → D, CD → E, BD → A

a) Candidate keys are: ABCF and BCDF.

b) All FDs are BCNF violations.

c) CD → E violates => decompose R to R1(CDE) and R2(ABCDF)
R1 is BCNF since it has two attributes

For R2, ABF → D and BD → A are the FDs and candidate keys are ABCF BCDF
BD → A violates => decompose R2 to R3(ABD) and R4(BCDF)

R3 is in BCNF since FD is BD → A only and candidate key is BD and BD → A does not violate
R4 is in BCNF since no FDs

Hence, final decomposition is CDE, ABD, and BCDF

If you start with a different FD, you will get following decompositions:
CDE, ABD, BDF, and ABCF
or
ABD, BDF, and ABCEF

iv. AB → D, BCD → EF, B → C

a) Candidate key is: AB.

b) BCD → EF and B → C are the BCNF violations.

c) BCD → EF violates => decompose R to R1(BCDEF) and R2(ABCD)

B → C violates => decompose R1 and R2 to: BC, BDEF, and ABD
No more violation => final decomposition: BC, ABD, and BDEF

If you choose B → C to start with, the final decomposition will be BC, ABD, and **ABEF**.

Q2. [4 marks] One (1) mark for each of the four RA queries.

Note that there are many correct answers to this question. The following are just sample answers.

i.

$Proj_{[name]}(Sel_{[category.sector = 'Technology']}(category\ Join\ company))$

ii.

R1 = executive
R2 = executive
R3 = executive
R4 = executive
R5 = executive
R6 = executive (since we need more than 5)

$Proj_{[R1.code]}(Sel_{[R1.code=R2.code=R3.code=R4.code=R5.code=R6.code\ and\ R1.person!=R2.person!=R3.person!=R4.person!=R5.person!=R6.person]}(R1xR2xR3xR4xR5xR6))$

Or

$Proj_{[code]}\ (\ Sel_{[count>5]}\ (\ GroupBy_{[code,\ Count[code]]}(executive))\ )$

iii.

R1 = executive
R2 = executive
$Proj_{[R1.person]}\ (\ Sel_{[R1.person = R2.person\ and\ R1.code\ !=\ R2.code]}(R1\ x\ R2)\ )$

Or

$Proj_{[person]}\ (\ Sel_{[count>1]}\ (\ GroupBy_{[person,\ Count[person]]}(executive))\ )$

iv.

R1 = $Proj_{[code,\ industry]}(category)$
R2 = $Proj_{[code,\ industry]}(category)$
R3 = $Proj_{[R1.code,\ R1.industry]}\ (\ Sel_{[R1.industry = R2.industry\ \&\ R1.code\ !=\ R2.code]}(R1\ x\ R2)\ )$
Result = R1 - R3

Or

R1 = $Proj_{[industry]}\ (\ Sel_{[count=1]}\ (\ GroupBy_{[industry,\ Count[industry]]}(category))\ )$
Result = $Proj_{[code,\ industry]}$ (R1 Join category)

Q3. [3 marks] One (1) mark for each expression (0.5 mark for min and 0.5 mark for max)

i. R UNION (S INTERSECT T)

Min:

Minimum of INTERSECT is 0

Minimum of A UNION B is max(|A|,|B|)

=> min(expression) = max(r,0) = **r**

Max:

Maximum of A INTERSECT B is min(|A|,|B|)

Maximum of A UNION B is |A|+|B|

=> max(expression) = **r + min(s,t)**

ii. $SEL_{[c]}$ (R × S), for some condition c

Min:

Minimum is selecting nothing
=> min(expression) = 0

Max:

Maximum is selecting all R x S rows
=> max(expression) = r*s

iii. $PROJ_{[a]}$ (R) - $PROJ_{[a]}$ (R JOIN S), for some list of attributes a

Min:

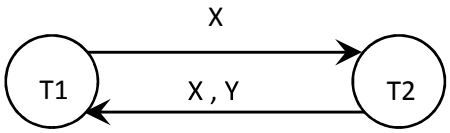
Minimum of A JOIN B is 0
Maximum of A JOIN B is **|A|*|B|**
=> min(expression) = r - (max of JOIN)
=> min(expression) = **r – r*s**
=> min(expression) = r – r = **0** (if **r*s** = r)

Max:

Maximum of expression is r - (min of JOIN)
=> max(expression) = r – 0 = **r**

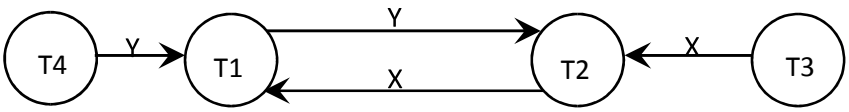
Q4. [2 marks] One (1) mark for each of (i) and (ii) (0.5 mark for the precedence graph and 0.5 mark for the explanation)

i. T1:R(X) T2:R(X) T1:W(X) T2:W(X) T2:R(Y) T1:R(Y) T1:W(Y) T2:W(X)



There is a cycle: [T1] ----X-----> [T2] ----X-----> [T1] (or [T1] ----X-----> [T2] ----Y-----> [T1]) in the precedence graph, hence the schedule is not serialisable.

ii. T3:R(X) T4:W(Y) T4:W(Z) T1:W(Y) T2:R(Y) T3:R(D) T2:W(X) T1:R(X)



There is a cycle: [T1] ----Y-----> [T2] ----X-----> [T1] in the precedence graph, hence the schedule is not serialisable.