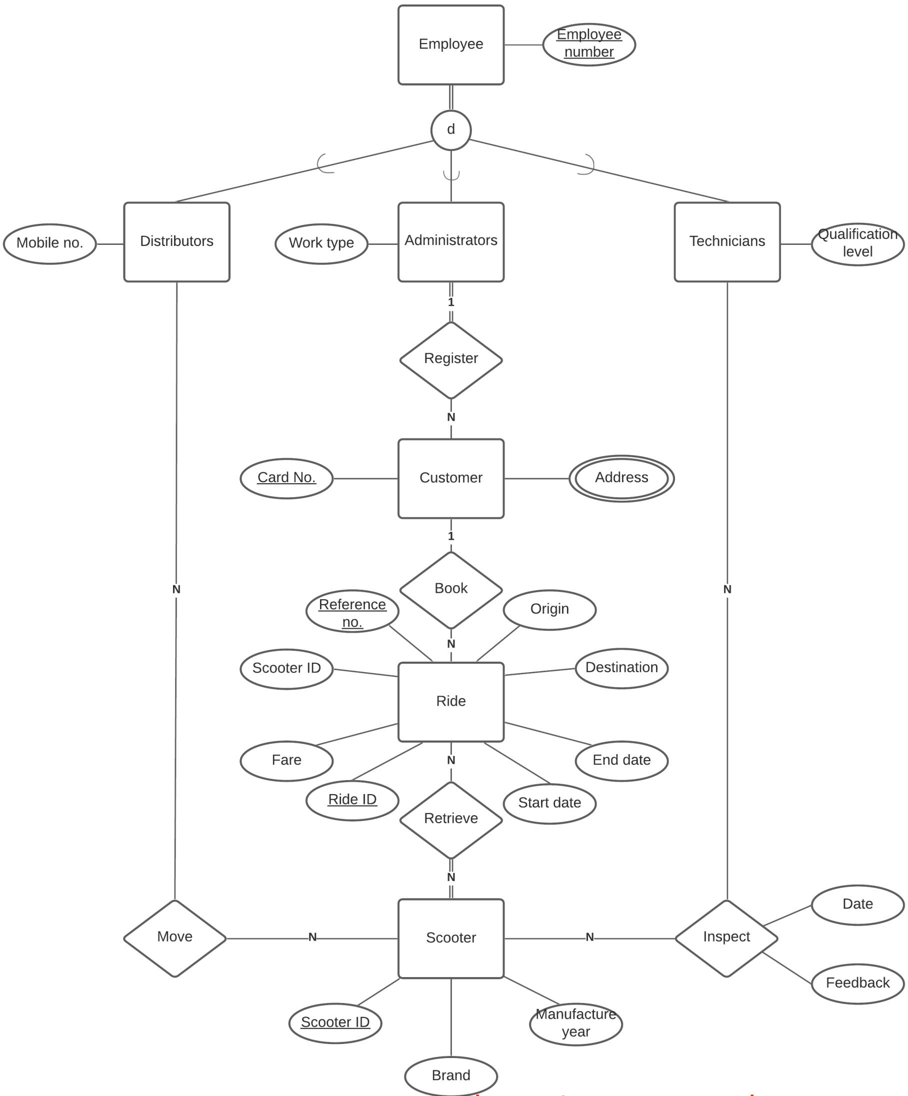
COMP2400/6240 Database Theory

TOTAL = 13.75

1. (Unfortunately, my laptop cannot run TerraER, Lucidchart is therefore the replacement)



Ride retrieve scooter does not make sense -0.5
scooter id already in Ride so its not needed

wrong participation in customer book ride -0.25

marks 3.25

2

2.1

AB→C, C→DE, DE→B, BC→A

Compute the closures for all possible combinations of the attributes in R:

marks 5

$(A)^+ = A$, $(B)^+ = B$, $(C)^+ = C$, $(D)^+ = D$, $(E)^+ = E$;

Notice that C is a single attribute, so firstly build the closure of C:

$(C)^+ \supseteq C$ (initialisation)

$\supseteq CDE$ (Using $C \rightarrow DE$)

$\supseteq CDEB$ (Using $DE \rightarrow B$)

$\supseteq CDEBA$ (Using $BC \rightarrow A$)

$= CDEBA$

If no proper subset Y of X with $Y^+ = R$, therefore, C is the candidate key.

Compute the closures for all possible combinations of the attributes in R that contain two attributes: AB, AD, AE, BD, BE, DE

$(AB)^+ \supseteq AB$ (initialisation)

$\supseteq ABC$ (Using $AB \rightarrow C$)

$\supseteq ABCDE$ (Using $C \rightarrow DE$)

$= ABCDE$

$(AD)^+ = AD$

$(AE)^+ = AE$

$(BD)^+ = BD$

$(BE)^+ = BE$

$(DE)^+ = DEB$ (Using $DE \rightarrow B$)

Therefore, AB is also a candidate key.

Compute the closures for all possible combinations of the attributes in R that contain three attributes: ADE, BDE

$(ADE)^+ \supseteq ADE$ (initialisation)

$\supseteq ADEB$ (Using $DE \rightarrow B$)

$\supseteq ADEBC$ (Using $AB \rightarrow C$)

$= ADEBC$

$(BDE)^+ = BDE$

Therefore, ADE is a candidate key too.

2.2

In order to find the minimal cover, firstly ensure that all functional dependencies only determine one attribute: C→<u>DE</u>, <u>AB</u>→C, <u>DE</u>→B, <u>BC</u>→A

Check for redundant attributes by repeating the above steps in 2.1:

C→DE?	D ∈ $(C)^+$ under Σ?: $(C)^+$ = (ABCDE)	✓	C→D
C→DE?	E ∈ $(C)^+$ under Σ?: $(C)^+$ = (ABCDE)	✓	C→E
AB→C?	C ∈ $(A)^+$ under Σ?: $(A)^+$ = (A)	✘	AB→C
AB→C?	C ∈ $(B)^+$ under Σ?: $(B)^+$ = (B)	✘	AB→C
DE→B?	B ∈ $(D)^+$ under Σ?: $(D)^+$ = (D)	✘	DE→B
DE→B?	B ∈ $(E)^+$ under Σ?: $(E)^+$ = (E)	✘	DE→B
BC→A?	A ∈ $(B)^+$ under Σ?: $(B)^+$ = (B)	✘	BC→A
BC→A?	A ∈ $(C)^+$ under Σ?: $(C)^+$ = (ABCDE)	✓	C→A

Minimal cover: Σ = (AB→C, C→D, C→E, DE→B, C→A)

2.3

Σ = (AB→C, C→DE, DE→B, BC→A)

Σ_1 = (AB→CDE, DE→B, C→AB)

So firstly build the closures of Σ_1:

$(C)^+$ ⊇ C (initialisation)

⊇ CAB (Using C→AB)

⊇ CABDE (Using AB→CDE)

= CABDE

$(AB)^+$ $\supseteq$ AB (initialisation)
$\supseteq$ ABCDE (Using AB→CDE)
= ABCDE

$(DE)^+$ $\supseteq$ DE (initialisation)
$\supseteq$ DEB (Using DE→B)
= DEB

Σ: AB→C (ABCDE) = Σ_1:AB→CDE (ABCDE)
Σ: C→DE (CABDE) = Σ_1:C→AB (CABDE)
Σ: DE→B (DEB) = Σ_1:DE→B (DEB)
Σ: BC→A (ABCDE) = Σ_1:C→AB (ABCDE)

They are equivalent because $\Sigma \models \Sigma_1$ and $\Sigma_1 \models \Sigma$.

3.

Q3: 1.75/2 marks

Alias the attributes to their first letter of their names for clarity.
A = (P, G, D, T, C, R)
PCD→T
P→G
G→C
CDTR→P
PDT→CR

Firstly check the closure of the left-hand side (determinant):

$(PCD)^+$ $\supseteq$ PCD (initialisation)
$\supseteq$ PCDT (Using PCD→T)
$\supseteq$ PCDTR (Using PCD→T)
$\supseteq$ PCDTRG (Using P→G)
= PCDTRG ✓

$(P)^+$ $\supseteq$ P (initialisation)
$\supseteq$ PG (Using P→G)
$\supseteq$ PGC (Using G→C)
= PGC ✘

$(G)^+$ $\supseteq$ G (initialisation)
$\supseteq$ GC (Using G→C)
= GC ✘

$(CDTR)^+$ $\supseteq$ CDTR (initialisation)
$\supseteq$ CDTRP (Using CDTR→P)
$\supseteq$ CDTRPG (Using P→G)
= CDTRPG ✓

$(PDT)^+$ $\supseteq$ PDT (initialisation)
$\supseteq$ PDTCR (Using PDT→CR)
$\supseteq$ PDTCRG (Using P→G)
= PDTCRG ✓

As indicated above, only three closures included all the attributes, therefore, Appointment is not in BCNF.

Then select FDs that violate BCNF requirements and decompose them. Firstly select P→G and produce the below relations and FDs.

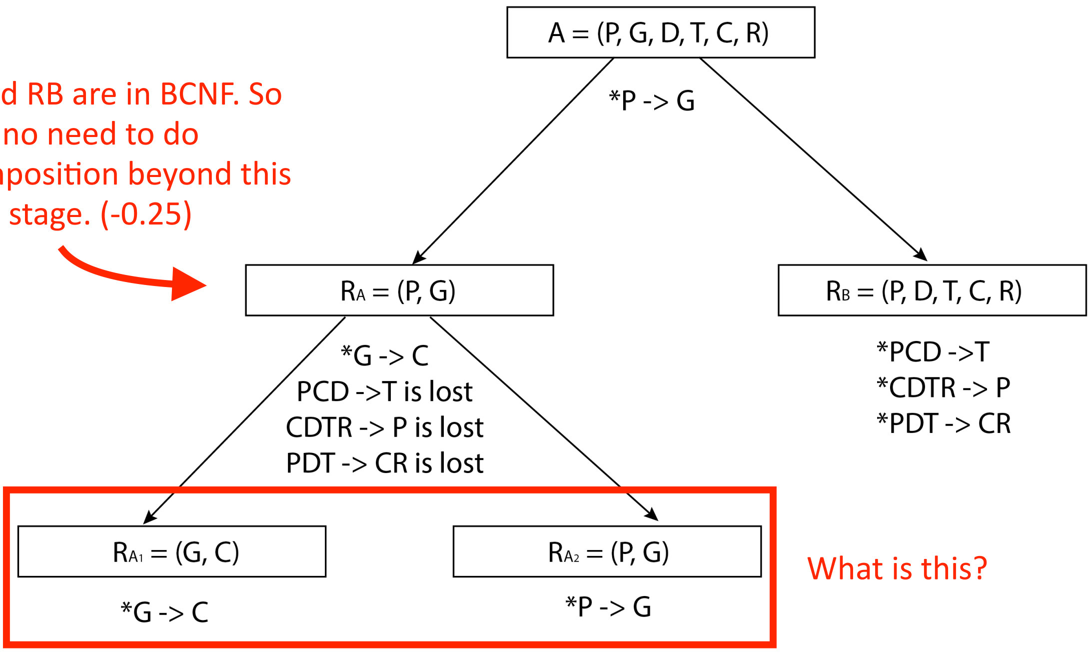
R_A = (P, G)
Σ_A = (P→G)
R_B = (P, D, T, C, R)
Σ_B = (PCD→T, CDTR→P, PDT→CR)

Σ_B = (PCD→T, CDTR→P, PDT→CR) is now in BCNF since both of them are superkeys for Σ_B. But is R_A not in BCNF because G→C and G is not a superkey. Hence select G→C and create the below decomposition of R_A.

Wrong

R_{A_1} = (G, C)
Σ_{A_1} = (G→C)
R_{A_2} = (P, G)
Σ_{A_2} = (P→G)
R_B = (P, D, T, C, R)
Σ_B = (PCD→T, CDTR→P, PDT→CR)
Note that PCD→T, CDTR→P, PDT→CR are lost.

RA and RB are in BCNF. So no need to do decomposition beyond this stage. (-0.25)



What is this?

Now R_{A_1}, R_{A_2} and R_B are all in BCNF, they are then the decomposition of Appointment. Since the lost FDs cannot be recovered, the decomposition is not dependency-preserving.

Wrong justification for why the BCNF decomposition is not dependency-preserving. (-0.25)
GP —> Clinic is lost.

Q4: 3.75/4 marks

4.

4.1

(a)

$R_1 = \pi_{TID}(\sigma_{Semester = 'S22021' \wedge CourseNo = 'COMP2400'}(TUTOR))$

$R_2 = \pi_{SID}(\sigma_{Semester = 'S22020' \wedge CourseNo = 'COMP2400}(ENROL))$

$R_3 = \pi_{SID}(\sigma_{TID = SID}(R_1 \bowtie R_2)$

$R = \pi_{phone}(STUDENT \bowtie R_3)$

1 mark

OR

$R = \pi_{phone}(\pi_{SID}(\pi_{SID}(\sigma_{semester = 'S22020' \wedge CourseNo = 'COMP2400'} ENROL) = \pi_{TID}(\sigma_{semester = 'S22021' \wedge CourseNo = 'COMP2400'} TUTOR) \bowtie STUDENT)$

(b)

$R_1 = (\sigma_{semester = 'S22021'}(TUTOR))$

$Tutor' = \sigma_{semester = 'S22021'}(TUTOR)$

$R_2 = \pi_{TID}(\sigma_{\varphi}(\rho_{RA}(Tutor') \times \rho_{RB}(Tutor'))$

RA.TID or RB.TID?

where

$\varphi = ((R_A.CourseNo \neq R_B.CourseNo) \wedge (R_A.TID = R_B.TID))$

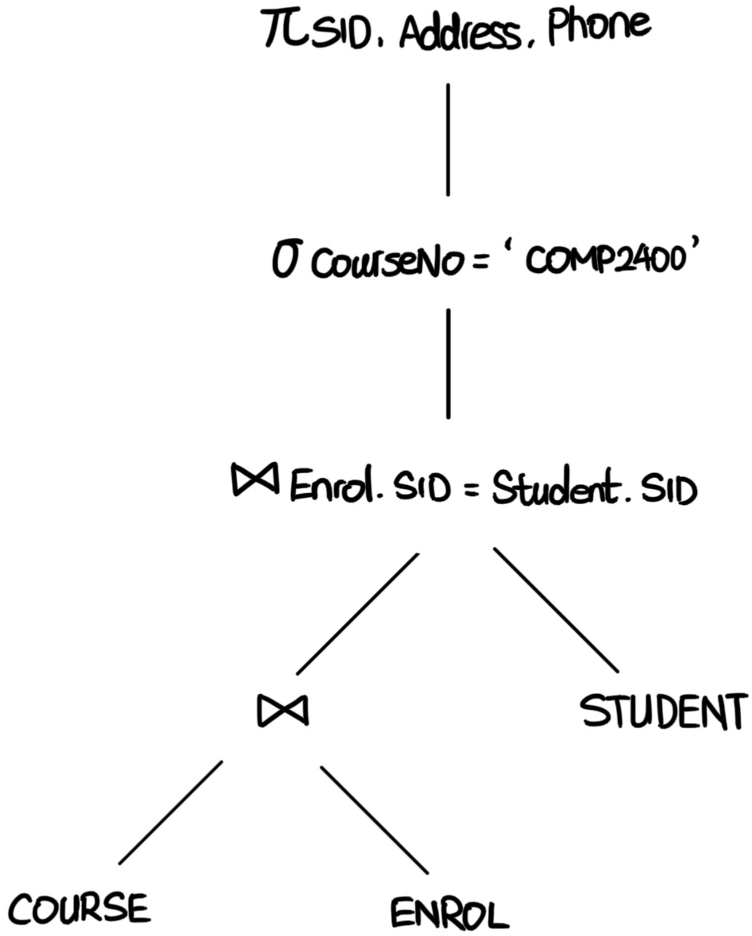
$R = R_1 - R_2$

Missing projection of TID from R1 before set difference

0.75 marks

4.2

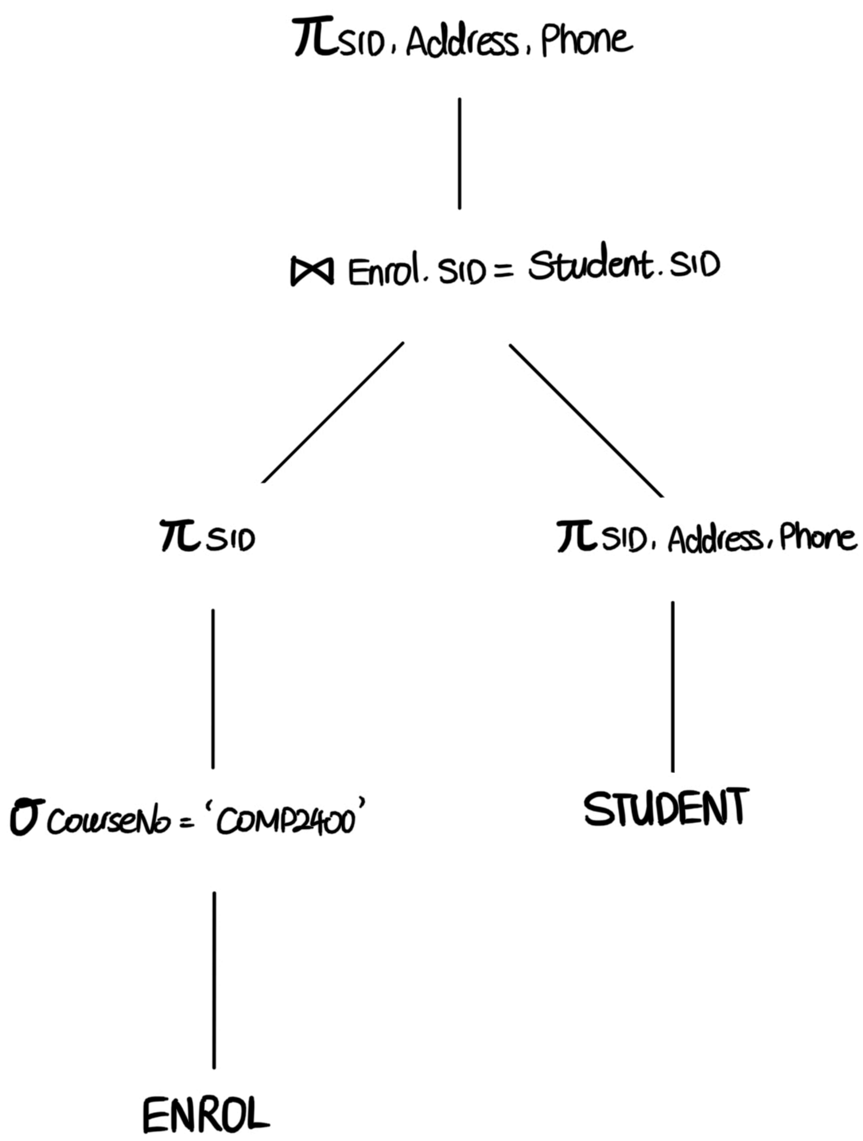
Original Query Tree:



The intention is to list SID, addresses and phone numbers of students who enrolled in COMP2400, therefore the optimised solution is:

$$\pi_{SID,\ Address,\ Phone}((\pi_{SID}(\sigma_{CourseNo='COMP2400'}(ENROL))) \bowtie_{ENROL.SID=STUDENT.SID} (\pi_{SID,\ Address,\ Phone}(STUDENT)))$$

Optimised Query Tree:



2 marks