CS482/502 Database Management Systems I

Assignment 4: Theory of Database Design

A. (40%) Consider a relation schema R(A,B,C,D,E) that satisfies the set of functional dependencies F = {BC → D, D → E, A → C, B → C}.

(i) (10%) Is BC→ E in F^+? Please justify. (Justification carries 7%.

Yes. We have: BC -> D and D -> E, thus by transitivity, get BC ->E.

(ii) (10%) Calculate $(AB)^+$.

Take AB, get:

AB

ABC (A->C, B->C)

ABCD (BC ->D, BC ⊆ ABC)
ABCDE (D -> E, D⊆ ABCD)

(iii) (15%) Find all the candidate keys of R and show your steps to find them. (Steps carry 10%)

Step 1: Calculating closures for all individual attributes

A+ = AC,

B+ = BCDE,

C+ = C

D+ = DE,

E+ = E

No candidate key found.

Step2: Calculate the closure of sets with two attributes

(AB)+ = ABCDE (from the previous question)

(AC)+ = AC (A->C)

(AD)+ = AD -> ACDE (A->C, D->E)

(AE)+ = ACE (A ->C)

(BC)+ = BC

->BCD (BC -> D)

-> BCDE (D-> E and D⊆BCD)

(BD)+ = BD -> BCDE (B-> C, D-> E)

(BE)+ = BE

->BCE (B -> C)

-> BCDE (BC-> D and BC⊆BCE)

(CD)+ = CD -> CDE (D->E)

(CE)+ = CE

(DE)+ = DE

(AB)+ has all the attributes of R

We get AB is a super key since (AB)+ = ABCDE contains all attributes of R.

And A or B individually was not super key.

Thus, AB is a candidate key.

Step 3: Calculate the closure of sets with 3 attributes. AB did not included since AB is a candidate key.

(ACD)+= ACD -> ACDE (D -> E, D ⊆ ACD)

(ACE)+ = ACE

(ADE)+ = ADE -> ACDE (A -> C, A ⊆ ADE)
(BCD)+ = BCD -> BCDE (B -> C, B⊆BCD)
(BDE)+ =BDE -> BCDE (B -> C, B ⊆ BDE)
(BCE)+ = BCE -> BCDE (BC -> D, BC ⊆ BCE)

(CDE)+ =CDE

No candidate key found.

Step 4: Calculate the closure of sets with 4 attributes, AB did not include.

(ACDE)+ = ACDE

(BCDE)+ = BCDE

No candidate key found.

Step 5: Calculate the closure of sets with 5 attributes, AB did not include.
Since the only closure is ABCDE which includes AB, we did not find new candidate key.

So, the candidate key is AB.

B. (20%) Consider a relation schema R(A, B,C,D,E) that satisfies the set of functional dependencies F = $\{AB \rightarrow D, D \rightarrow C\}$.

1. (10%) Is R in BCNF? Please justify your answer. (Justification carries 8%.)

The requirement of BCNF is, **for all** functional dependencies in F^+ of the form $\alpha \rightarrow \beta$ ($\alpha \subseteq R$ and $\beta \subseteq R$), fits

$\alpha \rightarrow \beta$ is trivial (i.e., $\beta \subseteq \alpha$) or

α is a superkey for R

Here we have F= { AB -> D, D -> C}

Since (D)+ = DC (D -> C), D is not a superkey of R.

Also, D->C is not trivial.

So R is not in BCNF.

2. (10%) Is R in 3NF? Please justify your answer. (Justification carries 8%.)

The requirement of 3NF is, for all $\alpha \rightarrow \beta$ in F+, at least one of the following holds:
$\alpha \rightarrow \beta$ is trivial (i.e., $\beta \in \alpha$)
α is a superkey for R
Each attribute A in $\beta - \alpha$ is contained in a candidate key for R.

As previous question, D->C did not fit the first 2 requirements.
Also, C – D = C, C+ = C did not contained any candidate key.
So the relation schema R is not in 3NF.

C. (30%) Consider a relation schema R(A,B,C,D) that satisfies the set of functional dependencies F = {A → B, BC → D}.

1. (15%) Find all the candidate keys of R (Steps will carry 10 points).

Step 1: Calculating closures for all individual attributes

A+ = AB

B+ = B

C+ = C

D+ = D

No candidate key found.

Step 2: Calculate the closure of sets with 2 attributes

(AB)+ = AB

(AC)+= AC

-> ABC (A-> B)

-> ABCD (BC->D, BC $\subseteq$ ABC)

(AD)+ =AD -> ABD(A ->D)

(BC)+ = BC -> BCD(BC -> D)

(BD)+ = BD

(CD)+ = CD

We get (AC)+ has all the attributes of R.

To check if AC is a candidate key:

We have AC is a super key since (AC)+ = ABCD

Then, A or C individually is not a super key.

Thus AC is a candidate key.

Step 3: Calculate the closure of sets with 3 attributes, not included AC.

(ABD)+ = ABD

(BCD)+ = BCD

No candidate key found.

Step 4: Calculate the closure of sets with 4 attributes, not included AC.

We get ABCD which include AC, thus no candidate key found.

So, AC is the candidate key.

2. (20%) Is R in BCNF? If R is not in BCNF, give a lossless-join BCNF decomposition of R.

For relation BC → D is not trivial. And BC+ = BCD is not a super key. Thus, Schema R is not in BCNF.

The lossless decomposition of R:

Let

R1: (A U B) = (A ,B)

R2: (R - B) = (A,C,D)

We can also have:

R1: (BC U D) = (B,C,D)

R2: (R-D) = (A,B,C)

and R1 U R2 gives A,B,C,D

Considering second functional dependency BC → D

3. (10%) Is R in 3NF? Please justify your answer. (Justification carries 8%.)

No.

For relation A -> B, A->B is non-trivial. And A is not a super key of R.

B - A = B, B+ = B, so B is not a candidate key.

Therefore, R is NOT in 3NF.