CSC343

"Assignment 3"

Student: Yiyang Hua, Jiazhi Chang and

April 5, 2019

Database Design Q1

a. Find the set of attributes not on the RHS of any FD, which is $NotOnRHS = \{\text{G}\}$. Every candidate keys must contain this attribute. Since it only contains G, the closure set of $NotOnRHS$ is also G.
And then, find the set of attributes that appeared on the RHS of some FD, but not on the LHS of any FD, which is $OnRHSNotOnLHS = \{\text{F}\}$. So F cannot be in any candidate key.
AB+ = ABCDEF
AC+ = ACBDEF
AD+ = ADEF
B+ = BD
BC+ = BCADEF
E+ = EF
We can see that AB+ = ABCDEF and since G is the attribute not on the RHS of any FD, which means that GAB+ = ABCDEFG. So GAB is a candidate key and no superset of GAB can be a key.
Similar as GAB, we can see that AC+ = ACBDEF and since G is the attribute not on the RHS of any FD, which means that GAC+ = ABCDEFG. So GAC is a candidate key and no superset of GAB can be a key.
And BC+ = BCADEF and since G is the attribute not on the RHS of any FD, which means that GBC+ = ABCDEFG. So GBC is a candidate key and no superset of GAB can be a key.
And there is no key that has G but not AB. We know this because even if we use every other attribute except AB (which we know cant be in any other key), we dont have a candidate key: CDEFG+ = CDEFG.
Similarly, there is also no key that has G but not AC. We know this because even if we use every other attribute except GB (which we know cant be in any other key), we dont have a candidate key: BDEFG+ = BDEFG.
And also, there is also no key that has G but not BC. We know this because even if we use every other attribute except GB (which we know cant be in any other key), we dont have a candidate key: ADEFG+ = ADEFG.
So we have 3 candidate keys: GAB, GAC. GBC.

b. Well first eliminate redundant FDs.
Well rewrite the FD into those with only one attribute on RHS, and call this set $S1$ (but there is nothing changed since all the RHS is only one attribute):

1. $B \hookrightarrow D$
2. $BC \hookrightarrow A$
3. $E \hookrightarrow F$
4. $AB \hookrightarrow C$
5. $AC \hookrightarrow B$
6. $AD \hookrightarrow E$

After it, we'll do the closure test to see whether we can remove any of them, and call this set $S2$, but since they are all the only way to get their RHS, they must be kept, so there is nothing changed:
1. $B \hookrightarrow D$
2. $BC \hookrightarrow A$
3. $E \hookrightarrow F$
4. $AB \hookrightarrow C$
5. $AC \hookrightarrow B$
6. $AD \hookrightarrow E$

Step 3, we'll try reducing the LHS of any FDs with multiple attributes on the LHS, and call this set $S3$. But there is still no change:
1. $B \hookrightarrow D$

2. $BC \hookrightarrow A$: Since B+ = BD, B+ doesn't contain C, so we cannot reduce this FD to B. And since C+ = C, C+ doesn't contain B, so we cannot reduce this FD to C. Therefore, this FD remains as it is.

3. $E \hookrightarrow F$

4. $AB \hookrightarrow C$: Since A+ = A, A+ doesn't contain B, so we cannot reduce this FD to A. And since B+ = BD, B+ doesn't contain A, so we cannot reduce this FD to B. Therefore, this FD remains as it is.

5. $AC \hookrightarrow B$: Since A+ = A, A+ doesn't contain C, so we cannot reduce this FD to A. And since C+ = C, C+ doesn't contain A, so we cannot reduce this FD to C. Therefore, this FD remains as it is.

6. $AD \hookrightarrow E$: Since A+ = A, A+ doesn't contain D, so we cannot reduce this FD to A. And since D+ = D, D+ doesn't contain A, so we cannot reduce this FD to D. Therefore, this FD remains as it is.

At the end, we'll do the closure test again to check, and call this set $S4$, since $S4$ is same as $S2$, so this closure test solution is also same as $S2$, so there is still nothing changed...:
1. $B \hookrightarrow D$
2. $BC \hookrightarrow A$
3. $E \hookrightarrow F$
4. $AB \hookrightarrow C$
5. $AC \hookrightarrow B$
6. $AD \hookrightarrow E$

c. A relation is in BCNF if and only if for every non-trivial FD, the LHS is a superkey.
From a), we know the 3 candidate keys are GAB, GAC, and GBC.
The FD $B \hookrightarrow D$ is non-trivial and its LHS is not a superkey. It violates BCNF.
We can begin to decompose R into R1 and R2.
R1 = B+ = BD; FD:$\{B \hookrightarrow D\}$
R2 = R - (B+) $\cup$ B = ABCEFG; FD:$\{BC \hookrightarrow A, E \hookrightarrow F, AB \hookrightarrow C, AC \hookrightarrow B\}$
Check R1 for BCNF violations. There is no violation.
Then check R2 for BCNF violations. The FD $BC \hookrightarrow A$ violates BCNF. So we decompose R2 into R21 and R22.
R21 = BC+ = BCA; FD:$\{BC \hookrightarrow A, AB \hookrightarrow C, AC \hookrightarrow B\}$
R22 = R2 - (BC+) $\cup$ BC = BCEFG; FD:$\{E \hookrightarrow F\}$
Check R21 for BCNF violations. There is no violation.
Then check R22 for BCNF violations. The FD $E \hookrightarrow F$ violates BCNF. So we decompose R22 into R221 and R222.
R221 = E+ = EF; FD:$\{E \hookrightarrow F\}$
R222 = R22 - (E+) $\cup$ E = BCEG; FD:$\{\}$
Check R221 for BCNF violations. There is no violation.
Then check R222 for BCNF violations. There is no violation.
Final decomposition:
R1 = B+ = BD; FD:$\{B \hookrightarrow D\}$
R21 = BC+ = BCA; FD:$\{BC \hookrightarrow A, AB \hookrightarrow C, AC \hookrightarrow B\}$
R221 = E+ = EF; FD:$\{E \hookrightarrow F\}$
R222 = R22 - (E+) $\cup$ E = BCEG; FD:$\{\}$
This decomposition is not dependency preserving, as $AD \hookrightarrow E$ is lost.

d. From a), we know the 3 candidate keys are GAB, GAC, and GBC. So A, B, C, G are members of keys of this relation R.
We need to check every FD that its RHS is prime, which is a member of any key, or its LHS is a superkey.
The FD $B \hookrightarrow D$ violates definition of 3NF: it is non-trivial, LHS is not superkey, and RHS is not prime.
From b), we know the minimal cover is $\{B \hookrightarrow D, BC \hookrightarrow A, E \hookrightarrow F, AB \hookrightarrow C, AC \hookrightarrow B, AD \hookrightarrow E\}$.
So we will have 3NF decomposition:
R1 = (B,D); FDs:$\{B \hookrightarrow D\}$
R2 = (A, B, C); FDs:$\{BC \hookrightarrow A, AB \hookrightarrow C, AC \hookrightarrow B\}$.
R3 = (E, F); FDs:$\{E \hookrightarrow F\}$.
R4 = (A, D, E); FDs: $\{AD \hookrightarrow E\}$.
For a lossless decomposition, we add a relation corresponding to one of the candidate keys of R,
R5 = (A, B, G); FDs: $\{\}$

Database Design Q2

Let F(F+) denote the closure of the set of functional dependencies satisfied by the relation S, has only one-attribute keys, which is assumed to be in 3NF.
We need to show that for each nontrivial dependency $X \hookrightarrow A$ in F+, X is a superkey.
To show this, consider such a dependency. If X is not a superkey, the 3NF property guarantees that the attribute A must be prime. Since all keys are one-attribute by assumption, $A \hookrightarrow$ {every other attribute}. Then by transitivity, $X \hookrightarrow$ {every other attribute}, which implies X is a superkey. There is a contradiction. So the relation S is in BCNF if and only if it is in 3NF.

Entity-Relationship Model Q1