CSC343 Assignment 3

Angela Zhu, Xinyi Ji, Zhuozi Zou

April 5, 2019

1 Database Design

1. a. Since G is not on the RHS of any FD, it must be included in any superkey of R.

   $G^+ = G$
   No one-attribute superkeys found, need to calculate closures of all possible two-attribute combinations.

   $AG^+ = AG$
   $BG^+ = BDG$
   $CG^+ = CG$
   $DG^+ = DG$
   $EG^+ = EFG$
   $FG^+ = FG$
   No two-attribute superkeys found, need to calculate closures of all possible three-attribute combinations.

   $ABG^+ = ABCDEFG$, therefore ABG is a superkey.
   $ACG^+ = ABCDEFG$, therefore ACG is a superkey.
   $ADG^+ = ADEFG$
   $AEG^+ = AEFG$
   $AFG^+ = AFG$
   $BCG^+ = ABCDEFG$, therefore BCG is a superkey.
   $BDG^+ = BDG$
   $BEG^+ = BDEFG$
   $BFG^+ = BDFG$
   $CDG^+ = CDG$
   $CEG^+ = CEFG$
   $CFG^+ = CFG$
   $DEG^+ = DEFG$
   $DFG^+ = DFG$
   $EFG^+ = EFG$

   Since we have finished calculating the closures of all possible three-attribute combinations, and we have found 3 superkeys, this means that ABG, ACG, and BCG must be the minimal sets of attributes with the property that it functionally determines all the other attributes in a relation.
   Therefore ABG, ACG, and BCG are the candidate keys of R.

   b. **Step 1:**
   Since all FDs have only one attribute on the RHS, there is no need to split the FDs.

   Step 2:
   For the FDs that have ≥ 2 attributes on the LHS, try to reduce the LHS.
   $B \to D$ and $E \to F$: cannot simplify.

$BC \to A$: we cannot reduce the LHS of this FD since $B^+ = BD$ and $C^+ = C$, and none of them yield A.
$AB \to C$: we cannot reduce the LHS of this FD since $A^+ = A$ and $B^+ = BD$, and none of them yield C.
$AC \to B$: we cannot reduce the LHS of this FD since $A^+ = A$ and $C^+ = C$, and none of them yield B.
$AD \to E$: we cannot reduce the LHS of this FD since $A^+ = A$ and $D^+ = D$, and none of them yield E.
The current set of FDs, which is still FD, is:
1. $B \to D$
2. $BC \to A$
3. $E \to F$
4. $AB \to C$
5. $AC \to B$
6. $AD \to E$

Step 3:
Try to eliminate each FD.
1) $B^+_{FD-1} = B$, we need this FD.
2) $BC^+_{FD-2} = BCD$, we need this FD.
3) $E^+_{FD-3} = E$, we need this FD.
4) $AB^+_{FD-4} = ABDEF$, we need this FD.
5) $AC^+_{FD-5} = AC$, we need this FD.
6) $AD^+_{FD-6} = AD$, we need this FD.

Therefore, the current FD is already the minimal cover

c. R is not in BCNF. BCNF requires that the LHS of all the FDs to be a superkey:
1) $B^+ = BD$, so B is not a superkey and $B \to D$ violates BCNF.
2) $BC^+ = ABCDEF$, so BC is not a superkey and $BC \to A$ violates BCNF.
3) $E^+ = EF$, so E is not a superkey and $E \to F$ violates BCNF.
4) $AB^+ = ABCDEF$, so AB is not a superkey and $AB \to C$ violates BCNF.
5) $AC^+ = ABCDEF$, so AC is not a superkey and $AC \to B$ violates BCNF.
6) $AD^+ = ADEF$, so AD is not a superkey and $AD \to E$ violates BCNF.
BCNF Decomposition:

- Decompose R using FD $B \to D$. $B^+ = BD$ so this yields two relations $R_1(B, D)$ and $R_2(A, B, C, E, F, G)$.
  Project the FDs onto R_1:

B	D	closure	FDs
✓		$B^+ = BD$	$B \to D$; B is a superkey of $R1$
	✓	$D^+ = D$	Nothing

  This relation satisfies BCNF.
- Project the FDs onto R_2:

A	B	C	E	F	G	closure	FDs
✓						$A^+ = A$	Nothing
	✓					$B^+ = BD$	Nothing
		✓				$C^+ = C$	Nothing
			✓			$E^+ = EF$	$E \to F$; violates BCNF; abort projection

  We need to decompose R_2 further.
- Decompose R_2 using FD $E \to F$. $E^+ = EF$ so this yields two relations $R_3(E, F)$ and $R_4(A, B, C, E, G)$.

- Project the FDs onto R_3:

E	F	closure	FDs
✓		$E^+ = EF$	$E \to F$; E is a superkey of $R3$
	✓	$F^+ = F$	Nothing

This relation satisfies BCNF.

- Project the FDs onto R_4:

A	B	C	E	G	closure	FDs
✓					$A^+ = A$	Nothing
	✓				$B^+ = BD$	Nothing
		✓			$C^+ = C$	Nothing
			✓		$E^+ = EF$	Nothing
				✓	$G^+ = G$	Nothing
✓	✓				$AB^+ = ABCDEF$	$AB \to CE$; violates BCNF; abort projection

We need to decompose R_4 further.

- Decompose R_4 using FD $AB \to CE$. $AB^+ = ABCDEF$ so this yields two relations $R_5(A, B, C, E)$ and $R_6(A, B, G)$.
- Project the FDs onto R_5:

A	B	C	E	closure	FDs
✓				$A^+ = A$	Nothing
	✓			$B^+ = BD$	Nothing
		✓		$C^+ = C$	Nothing
			✓	$E^+ = EF$	Nothing
✓	✓			$AB^+ = ABCDEF$	$AB \to CE$; AB is a superkey of R_5
✓		✓		$AC^+ = ABCDEF$	$AC \to BE$; AC is a superkey of R_5
✓			✓	$AE^+ = AEF$	Nothing
	✓	✓		$BC^+ = ABCDEF$	$BC \to AE$; BC is a superkey of R_5
	✓		✓	$BE^+ = BDEF$	Nothing
		✓	✓	$CE^+ = CEF$	Nothing
	✓	✓	✓	$BCE^+ = ABCDEF$	Nothing, since $BCE \to A$ is weaker than $BC \to A$ which we already have
✓		✓	✓	$ACE^+ = ABCDEF$	Nothing, since $ACE \to B$ is weaker than $AC \to B$ which we already have
✓	✓		✓	$ABE^+ = ABCDEF$	Nothing, since $ABE \to C$ is weaker than $AB \to C$ which we already have
✓	✓	✓		$ABC^+ = ABCDEF$	Nothing, since $ABC \to E$ is weaker than $AB \to E$ which we already have

This relation satisfies BCNF.

- Project the FDs onto R_6:

A	B	G	closure	FDs
✓			$A^+ = A$	Nothing
	✓		$B^+ = BD$	Nothing
		✓	$G^+ = G$	Nothing
✓	✓		$AB^+ = ABCDEF$	Nothing
	✓	✓	$BG^+ = BDG$	Nothing
✓		✓	$AG^+ = AG$	Nothing

This relation satisfies BCNF.

Final decomposition:
$R_1(B, D)$ with FD $B \rightarrow D$
$R_3(E, F)$ with FD $E \rightarrow F$
$R_5(A, B, C, E)$ with FDs
$\{AB \rightarrow CE, AC \rightarrow BE, BC \rightarrow AE\}$
$\mathrm{R}_6(A, B, G)$ with no FDs.

d. R is not in 3NF. $E \rightarrow F$ is a FD that violates 3NF since E is not a superkey ($E^+ = EF$) and F is not a prime (since we know that F is not a part of the three candidate keys ABG, ACG, BCG as found in a).
3NF decomposition:
The minimal basis for FD is FD itself.

The set of relations that result from the FDs is:
$B \rightarrow D$: define $R_1(B, D)$.
$BC \rightarrow A$: define $R_2(A, B, C)$.
$AB \rightarrow C$: define $R_3(A, B, C)$.
$AC \rightarrow B$: define $R_4(A, B, C)$.
$E \rightarrow F$: define $R_5(E, F)$.
$AD \rightarrow E$: define $R_6(A, D, E)$.
Since R_2, R_3, R_4 are the same relations, we only need to keep one. We will keep R_2.
Since none of the relations is a superkey for R, use the superkey ABG (found in part a) and define $R_7(A, B, G)$.

The final result of 3BF decomposition of R:
$R_1(B, D)$, $R_2(A, B, C)$, $R_5(E, F)$, $R_6(A, D, E)$, and $R_7(A, B, G)$.

2. *Proof.* Recall the definition of BCNF and 3NF:
We say a relation R is in BCNF if for every nontricial FD $X \rightarrow Y$ that holds in R, X is a superkey.
We say a relation R is in 3NF if for every FD $X \rightarrow A$, X is a superkey or A is a prime.
The proof contains two parts.

Part 1: BCNF $\implies$ 3NF.
Since 3NF is less strict than BCNF, this implication must be true. For every FD $X \rightarrow Y$, X must be a superkey, therefore it does not violate 3NF.

Part 2: 3NF $\implies$ BCNF.
Since S is 3NF, we know that for every FD $X \rightarrow Y$, X is a superkey or Y is a prime.
Suppose X is a superkey, then this FD satisfies BCNF.
Suppose X is not a superkey, then Y must be a prime. This means that Y must be a member of any key. Since S has only one-attribute keys, Y must be a key itself, and Y^+ contains all attributes. However, since $X \rightarrow Y$, X^+ must then contain all attributes too. This implies that X is also a superkey, which violates with the assumption of X not being a superkey.
Therefore, for every FD $X \rightarrow Y$, X must be a superkey, thus S is in BCNF.

Since BCNF $\implies$ 3NF and 3NF $\implies$ BCNF both hold, it is true that if S has only one-attribute keys, S is in BCNF if and only if it is in 3NF.

■

2 Entity-Relationship Model