343A3

Siwei Tang, Zhenyu Wang

December 2020

BCNF

1. BCNF requires that the LHS of an FD be a superkey
   ✗ AEG^+ = AEGFHI, so AEG is not a superkey and AEG $\rightarrow$ F violates BCNF
   ✗ B^+ = ABD, so ABD is not a superkey and B $\rightarrow$ AD violates BCNF
   ✗ AG^+ = AGHI, so AG is not a superkey and AG $\rightarrow$ HI violates BCNF
   ✗ BG^+ = ABDGHI, so BG is not a superkey and BG $\rightarrow$ D violates BCNF

2. Decompose R using FD $AEG \rightarrow F$. AEG^+ = AEGFHI, so this yields two relations: R_1=AEGFHI and R_2=ABCDEG
   Project the FDs on to R_1=AEGFHI

A	E	G	F	H	I	closure	FDs
✓						A^+ = A	nothing
	✓					E^+ = E	nothing
		✓				G^+ = G	nothing
			✓			F^+ = F	nothing
				✓		H^+ = H	nothing
					✓	I^+ = I	nothing
✓	✓					AE^+ = AE	nothing
✓		✓				AG^+ = AGHI	$AG \rightarrow$ HI lex¿lex

   We must decompose R_1 further.
   Decompose R_1 using FD $AG \rightarrow$ HI. AG^+ = AGHI, so this yields two relations: R_3=AGHI and R_4=AEGF
   Project the FDs onto R_3=AGHI

A	G	H	I	closure	FDs
✓				A^+ = A	nothing
	✓			G^+ = G	nothing
		✓		H^+ = H	nothing
			✓	I^+ = I	nothing
✓	✓			AG^+ = AGHI	$AG \rightarrow$ HI
✓		✓		AH^+ = AH	nothing
✓			✓	AI^+ = AI	nothing
	✓	✓		GH^+ = GH	nothing
	✓		✓	GI^+ = GI	nothing
		✓	✓	HI^+ = HI	nothing
supersets of AG				irrelevant	can only generate weaker FDs than what we already have
✓		✓	✓	AHI^+ = AHI	nothing
	✓	✓	✓	GHI^+ = GHI	nothing

This relation satisfies BCNF.
Project the FDs onto R_4 = ADEF.

A	E	G	F	closure	FDs
✓				A^+ = A	nothing
	✓			E^+ = E	nothing
		✓		G^+ = G	nothing
			✓	F^+ = F	nothing
✓	✓			AE^+ = AE	nothing
✓		✓		AG^+ = AGHI	nothing
✓			✓	AF^+ = AF	nothing
	✓	✓		EG^+ = EG	nothing
	✓		✓	EF^+ = EF	nothing
		✓	✓	GF^+ = GF	nothing
✓	✓	✓		AEG^+ = AEGFHI	$AEG \rightarrow$ F
✓	✓		✓	AEF^+ = AEF	nothing
✓		✓	✓	AGF^+ = AGFHI	nothing
	✓	✓	✓	EGF^+ = EGF	nothing

This relation satisfies BCNF.
Return to R_2 = ABCDEG and project the FDs onto it.

A	B	C	D	E	G	closure	FDs
✓						A^+ = A	nothing
	✓					B^+ = ABD	$B \rightarrow$ AD

We must decompose R_2 further.
Decompose R_2 using FD $B \rightarrow$ AD. B^+ = ABD, so this yields two relations: R_5=ABD and R_6=BCEG
Project the FDs onto R_5=ABD

A	B	D	closure	FDs
✓			A^+ = A	nothing
	✓		B^+ = ABD	$B \rightarrow$ AD
✓			D^+ = D	nothing
supersets of B			irrelevant	can only generate weaker FDs than what we already have
✓		✓	AD^+ = AD	nothing

This relation satisfies BCNF.
Project the FDs onto R_6 = BCEG

B	C	E	G	closure	FDs
✓				B^+ = ABD	nothing
	✓			C^+ = C	nothing
		✓		E^+ = E	nothing
			✓	G^+ = G	nothing
✓	✓			BC^+ = ABCD	nothing
✓		✓		BE^+ = ABDE	nothing
✓			✓	BG^+ = ABDG	nothing
	✓	✓		CE^+ = CE	nothing
	✓		✓	CG^+ = CG	nothing
		✓	✓	EG^+ = EG	nothing
✓	✓	✓		BCE^+ = ABCDE	$AEG \to$ F
✓	✓		✓	BCG^+ = ABCDGHI	nothing
✓		✓	✓	BEG^+ = ABDEFGHI	nothing
	✓	✓	✓	CEG^+ = CEG	nothing

This relation satisfies BCNF.
Final decomposition:
(a) R_3=AGHI with $AG \to$ HI
(b) R_4=AEGF with $AEG \to$ F
(c) R_5=ABD with $B \to$ AD
(d) R_6=BCEG with no FDs.

3. Yes, it is. For each of the first three of the original FDs in set S, there is a relation that includes all of the FD's attributes. This ensures that they are preserved.
For the rest FD, $BG \to$ D, since $B \to$ AD is preserved, so $B \to$ D is preserved. Since BG is superets of B, so $BG \to$ D is preserved.

4. The following valid instance of the relation demonstrates that it is not a lossless-join decomposition.

A	B	C	D	E	F	G	H	I
a	3	3	3	3	3	g	h	i
a	3	3	3	e	f	g	3	3
a	b	3	d	3	3	3	3	3
a	b	c	d	e	f	g	h	i

Minimal basis

1. Step 1: split RHSs
   a: LNOP $\rightarrow$ M
   b: M $\rightarrow$ N
   c: M $\rightarrow$ Q
   d: NO $\rightarrow$ L
   e: NO$\rightarrow$ Q
   f: MNQ $\rightarrow$ L
   g: MNQ $\rightarrow$O
   h: LMQ $\rightarrow$ N
   i: LMQ $\rightarrow$ O
   j: LMQ $\rightarrow$ S

   step 2:
   a: L^+ = L, N^+ = N, O^+ = O, P^+ = P
   LN^+ = LN, LO^+ = LO, LP^+ = LP, NO^+ = NOQL, NP^+ = NP, OP^+=OP
   LNO^+ = LNOQ, LNP^+ = LNP, NOP^+ = NOPL
   Simplify LHS to NOP
   b: no simplify on LHS
   c: no simplify on LHS
   d: no simplify on LHS
   e: no simplify on LHS
   f: M^+ = MNQOLS, so we simplify f as M $\rightarrow$ L
   g: M^+ = MNQOLS, so we simplify f as M $\rightarrow$ O
   h: M^+ = MNQOLS, so we simplify f as M $\rightarrow$ N
   i: M^+ = MNQOLS, so we simplify f as M $\rightarrow$ O
   j: M^+ = MNQOLS, so we simplify f as M $\rightarrow$ S

   Since some of FDs are the same, so we con delete some FDs. B2 = { a: NOP $\rightarrow$ M, b: M $\rightarrow$N, c: M$\rightarrow$Q, d: NO $\rightarrow$ L, e: NO $\rightarrow$ Q, f: M $\rightarrow$ L, g: M $\rightarrow$O, h: M $\rightarrow$S}

   Step 3: try to eliminate entire FDs
   a) $NOP^+_{B2-\{a\}}$= LNOP - we need this FD
   b) $M^+_{B2-\{b\}}$= MQOLS - we need this FD
   c) $M^+_{B2-\{c\}}$= MNLOSQ - we can remove this FD
   d) $NO^+_{B2-\{c,d\}}$= NOQ - we need this FD
   e) $NO^+_{B2-\{c,e\}}$= NOL - we need this FD
   f) $M^+_{B2-\{c,f\}}$= MNOSLQ - we can remove this FD
   g) $M^+_{B2-\{c,f,g\}}$= MNS - we need this FD
   h) $M^+_{B2-\{c,f,h\}}$= MNOLQ - we need this FD

   We found this minimal basis:
   NOP $\rightarrow$ M
   M $\rightarrow$ N
   M $\rightarrow$ O
   M $\rightarrow$ S
   NO $\rightarrow$ L
   NO $\rightarrow$ Q

2.

	on LHS	on RHS	conclude?
R	✓	✓	in every key
P	✓	✓	in every key
L,M,N,O	✓	✓	must check
Q, S	✓	✓	in no key

So we need to check L, M, N, O

L	M	N	O	closure
✓				LPR^+ = LPR
	✓			MPR^+ = MPRNOSLQ
		✓		NPR^+ = NPR
			✓	OPR^+ = OPR
any supersets of MPR contains all alphabet				
✓		✓		$LNPR^+$ = LNPR
✓			✓	$LOPR^+$ = LOPR
		✓	✓	$NOPR^+$ = NOPRLQMS
✓		✓	✓	$LNOPR^+$ = NOPRLQM

So we found that MPR and NOPR are keys.

3. Following the 3NF synthesis algorithm, we could get one relation for each FD. however, we can merge the right-hand sides before doing so. This will yield a smaller set of relations and they will still form a lossless and dependency-preserving decomposition of relation R into a collection of relations that are in 3NF.
Let's called the revised FDs B3:
NOP → M
M → NOS
NO → LQ
The set of relations that would result would have these attributes:
R1(M, N, O, P), R2(M, N, O , S), R3(L, N, O, Q)
Since there is no super key for B, so we need add a relation that includes a key, R4(M, P, R)
So the final set of relation is:
R1(M, N, O, P), R2(M, N, O , S), R3(L, N, O, Q), R4(M, P, R)

4. Because we formed each relation from an FD, the LHS of those FDs are indeed superkeys for their relations. However, there may be other FDs that violate BCNF and therefore allow redundancy.The only way to find out is to project the FDs onto each relation.
We can quite quickly find a relation that violates BCNF without doing all the full projections:Clearly R4 does not have superkeys since M, P, R all are not superkeys.
For example, Consider relation $R1(M, N, O, P)$, after apply FD project, we have, $M^+ = LMNOQS$, thus, $M \to NO$, and M is not a superkey in this relation. So yes, these schema allows redundancy.