Normalization Questions and Answers

Database Systems, CSCI 4380-01
Sibel Adalı

October 28, 2002

Question 1 Suppose you are given a relation $R = (A, B, C, D, E)$ with the following functional dependencies: $\{CE \rightarrow D, D \rightarrow B, C \rightarrow A\}$.
a. Find all candidate keys.
b. Identify the best normal form that R satisfies (1NF, 2NF, 3NF, or BCNF).
c. If the relation is not in BCNF, decompose it until it becomes BCNF. At each step, identify a new relation, decompose and re-compute the keys and the normal forms they satisfy.

Answer.
a. The only key is $\{C, E\}$
b. The relation is in 1NF
c. Decompose into R1=(A,C) and R2=(B,C,D,E). R1 is in BCNF, R2 is in 2NF. Decompose R2 into, R21=(C,D,E) and R22=(B,D). Both relations are in BCNF.

Question 2 Suppose you are given a relation R=(A,B,C,D,E) with the following functional dependencies: $\{BC \rightarrow ADE, D \rightarrow B\}$.
a. Find all candidate keys.
b. Identify the best normal form that R satisfies (1NF, 2NF, 3NF, or BCNF).
c. If the relation is not in BCNF, decompose it until it becomes BCNF. At each step, identify a new relation, decompose and re-compute the keys and the normal forms they satisfy.

Answer.
a. The keys are $\{B, C\}$ and $\{C, D\}$
b. The relation is in 3NF
c. It cannot be put into BCNF, even if I remove D and put into a relation of the form (B,C,D) (I need C for the functional dependency), the resulting relation would not be in BCNF.

Question 3 Suppose you are given a relation R=(A,B,C,D,E) with the following functional dependencies: $BD \rightarrow E, A \rightarrow C$.
a. Show that the decomposition into R1=(A,B,C) and R2=(D,E) is lossy. You can show using any method. My suggestion is to show how spurious tuples result from this decomposition with respect to the table below:

A	B	C	D	E
1	2	3	4	5
1	8	3	4	4

b. Find a single dependency from a single attribute X to another attribute Y such that when you add the dependency $X \rightarrow Y$ to the above dependencies, the decomposition in part a is no longer lossy.

Answer.
a. If we were to decompose the relations into:

A	B	C
1	2	3
1	8	3

D	E
4	5
4	4

and then join the two (in this case with a cartesian product), we would get:

A	B	C	D	E
1	2	3	4	5
1	8	3	4	5
1	2	3	4	4
1	8	3	4	4

Tuples 2 and 3 are not in the original relation. Hence, this decomposition is lossy.

b. This decomposition cannot be made lossless. The problem is there is no longer a way to make sure $BD \rightarrow E$ holds across two relations since they do not share any attributes. However, a lossy decomposition of the form (A,B,C), (C,D,E) can be made lossless by adding an FD $B \rightarrow C$.

Question 4 You are given the following set of functional dependencies for a relation R(A,B,C,D,E,F), $\mathcal{F} = \{AB \rightarrow C, DC \rightarrow AE, E \rightarrow F\}$.
a. What are the keys of this relation?
b. Is this relation in BCNF? If not, explain why by showing one violation.
c. Is the decomposition (A,B,C,D) (B,C,D,E,F) a dependency preserving decomposition? If not, explain briefly.

Answer.
a. What are the keys of this relation?
$\{A, B, D\}$ and $\{B, C, D\}$.
b. Is this relation in BCNF? If not, explain why by showing one violation.
No, all functional dependencies are actually violating this. No dependency contains a superkey on its left side.
c. Is the decomposition (A,B,C,D) (B,C,D,E,F) a dependency preserving decomposition? If not, explain briefly.
Yes, $AB \rightarrow C$ and $DC \rightarrow A$ are preserved in the first relation. $DC \rightarrow E$ and $E \rightarrow F$ are preserved in the second relation.

Question 5 You are given the below functional dependencies for relation R(A,B,C,D,E), $\mathcal{F} = \{AB \rightarrow C, AB \rightarrow D, D \rightarrow A, BC \rightarrow D, BC \rightarrow E\}$.
a. Is this relation is in BCNF? If not, show all dependencies that violate it.
b. Is this relation in 3NF? If not, show all dependencies that violate it.

c. Is the following dependency implied by the above set of dependencies? If so, show how using the Amstrong's Axioms given in the book (p. 362-363): $ABC \to AE$

Answer.
Keys for the relation: $\{A, B\}$, $\{B, D\}$, $\{B, C\}$.
a. Not in BCNF since $D \to A$ does have a superkey on the left hand side.
b. In 3NF since in $D \to A$, A is part of a key.
c. $BC \to E$ (given)
$ABC \to AE$ by the augmentation rule.

Question 6 You are given the table below for a relation R(A,B,C,D,E). You do not know the functional dependencies for this relation. This question is independent of Question 2 above.

A	B	C	D	E
'a'	122	1	's1'	'a'
'e'	236	4	'e2'	'b'
'a'	199	1	'b5'	'c'
'b'	213	2	'z8'	'd'

Suppose this relation is decomposed into the following two tables: R1(A,B,C,D) and R2(A,C,E). Is this decomposition lossless? Explain your reasoning.

Answer.

R1

A	B	C	D
'a'	122	1	's1'
'e'	236	4	'e2'
'a'	199	1	'b5'
'b'	213	2	'z8'

R2

A	C	E
'a'	1	'a'
'e'	4	'b'
'a'	1	'c'
'b'	2	'd'

$R1 \bowtie R2$

A	B	C	D	E
'a'	122	1	's1'	'a'
'e'	236	4	'e2'	'b'
'a'	199	1	'b5'	'c'
'b'	213	2	'z8'	'd'
'a'	122	1	's1'	'a'
'a'	199	1	'b5'	'c'

Since the last two rows are not in the original relation, then this decomposition is lossy.

Question 7 You are given the below set of functional dependencies for a relation R(A,B,C,D,E,F,G), $\mathcal{F} = \{AD \to BF, CD \to EGC, BD \to F, E \to D, F \to C, D \to F\}$.
a. Find the minimal cover for the above set of functional dependencies using the algorithm described in class. Give sufficient detail to show your reasoning, but be succinct. You do not have to list all the cases you test/consider for the algorithm. Show all steps where you make changes to the above set in detail.
b. Using the functional dependencies that you computed in step a, find the keys for this relation. Is it in BCNF? Explain your reasoning.
c. Suppose we decompose the above relation into the following two relations:
R1(A,B,C,D,E) R2(A,D,F,G)
Use the functional dependencies in the minimal cover. For each relation, write down the functional dependencies that fall within that relation (you can decompose a dependency of the form $AD \to BF$ into two i.e. $AD \to B$ and $AD \to F$ when computing this).

Using these functional dependencies, determine if this decomposition is lossless and/or dependency preserving. Explain your reasoning.

Answers.
a.
Step 1.
$\{AD \to B, AD \to F, CD \to E, CD \to G, CD \to C, BD \to F, E \to D, F \to C, D \to F\}$
Step 2. $removeCD \to C$, $AD \to F$, and $BD \to F$.
$\{AD \to B, CD \to E, CD \to G, F \to C, D \to F, E \to D\}$
Step 3. remove D from $CD \to E$ and $CD \to G$
$\{AD \to B, D \to E, D \to G, F \to C, D \to F, E \to D\}$
Finally recombine
$\{AD \to B, D \to EGF, F \to C, E \to D\}$.
b. Keys: $\{A, D\}$, $\{A, E\}$. Not in BCNF since the last three functional dependencies do not have a superkey on the left hand side.
c. R1(A,B,C,D,E) Dependencies: $AD \to B, D \to E, E \to D$ R2(A,D,F,G) Dependencies: $D \to GF$.
Not functional dependency preserving, the dependency $F \to C$ is not preserved.
$head(R1) \cap head(R2) = \{A, D\}$
R1: $AD \to ABCDE$ is not true since C is not implied by A,D
R2: $AD \to ADFG$ is true since this is implied by $D \to GF$ as follows:
$AD \to AD$ inclusion rule, since $D \to GF$, use set accumulation rule, $AD \to ADGF$. Hence, this is a lossless decomposition.

Question 8 You are given the following set F of functional dependencies for a relation R(A,B,C,D,E,F):
$\mathcal{F} = \{ABC \to D, ABD \to E, CD \to F, CDF \to B, BF \to D\}$.
a. Find all keys of R based on these functional dependencies.
b. Is this relation in Boyce-Codd Normal Form? Is it 3NF? Explain your answers.
c. Can the set F be simplified (by removing functional dependencies or by removing attributes from the left hand side of functional dependencies) without changing the closure of F (i.e. F+)? Hint. Consider the steps of the minimal cover algorithm. Do any of them apply to this functional dependency?

Answer.
a. Keys: $\{A, B, C\}$ and $\{A, C, D\}$
b. It is not in BCNF. Counterexample $ABD \to E$ and ABD is not a superkey.
It is not in 3NF. Counterexample $ABD \to E$, and ABD is not a superkey and E is not prime attribute (part of a key).
c. Let F' be obtained by replacing $CDF \to B$ with $CD \to B$.
According to F and F', $CD+ = \{C, D, B, F\}$. Hence, we can remove F from this functional dependency without changing the meaning of the system.

Question 9 Consider relation R(X, Y, Z). Relation R currently has three tuples: (6, 4, 2), (6, 6, 8) and (6, 4, 8). Which of the following three functional dependencies can you infer do not hold for relation R? Explain your answer.
$Y \to X$

$Z \to Y$
$XY \to Z$

Answer. The first functional dependency holds, but the rest do not hold. The second and third tuples both have 8 for Z but different values of Y. The first and third tuples both have 6 and 4 for X and Y but different values for Z.

Question 10 Consider the relation R(V, W, X, Y, Z) with functional dependencies $\{Z \to Y, Y \to Z, X \to Y, X \to V, VW \to X\}$.
a) List the possible keys for relation R based on the functional dependencies above.
b) Show the closure for attribute X given the functional dependencies above.
c) Suppose that relation R is decomposed into two relations, R1(V, W, X) and R2(X, Y, Z). Is this decomposition a lossless decomposition? Explain your answer.

Answer.
a. $\{V, W\}$, $\{X, W\}$
b. $X^+ = \{X, V, Y, Z\}$
c. Yes it is lossless. To be lossless the attributes in common between the two relations must functionally determine all the attributes in one of the two relations. The only attribute in common is X and it functionally determines all the attributes in R2.

Question 11 Given relation R(W, X, Y, Z) and set of functional dependencies $\mathcal{F} = \{X \to W, WZ \to XY, Y \to WXZ\}$. Compute the minimal cover for $\mathcal{F}$.

Answer.
Step 1: $X \to W, WZ \to X, WZ \to Y, Y \to W, Y \to X, Y \to Z$
Step 2: Don't need $WZ \to X$, since $WZ \to Y$ and $Y \to X$
Don't need $Y \to W$, since $Y \to X$ and $X \to W$
This leaves $\{X \to WWZ \to Y, Y \to X, Y \to Z\}$
Step 3: Only need to consider $WZ \to Y$. Can't eliminate W or Z. So nothing is eliminated.
Step 4: $\{X \to WWZ \to Y, Y \to XZ\}$ is the minimal cover

Question 12 Given relation R(W, X, Y, Z) and set of functional dependencies $G = \{Z \to W, Y \to XZ, XW \to Y\}$, where G is a minimal cover:
a) Decompose R into a set of relations in Third Normal Form.
b)Is your decomposition in part a) also in Boyce Codd Normal Form? Explain your answer.

Answer.
a. Possible keys: $\{Y\}, \{X, Z\}, \{W, X\}$
R1=(Z, W), R2=(X, Y, Z), R3=(X, Y, W)
b.Yes. In each of the three relations, the left side of the funcational dependencies that apply are superkeys for the relation. Hence, all three relations satisfy the definition of BCNF.

Question 13 Consider a relation named EMP_DEPT with attributes: ENAME, SSN, BDATE, ADDRESS, DNUMBER, DNAME, and DMGRSSN. Consider also the set G of functional dependencies for EMP_DEPT:

$G = \{SSN \rightarrow ENAME\,BDATE\,ADDRESS\,DNUMBER, DNUMBER \rightarrow DNAME, DMGRSSM\}$.
a) Calculate the closures SSN+ and DNAME+ with respect to G.
b) Is the set of functional dependences G minimal? If not, find a minimal set of functional dependencies that is equivalent to G.
c) List an update anomaly that can occur for relation EMP_DEPT.
d) List an insertion anomaly that can occur for relation EMP_DEPT.
e) List a deletion anomaly that can occur for relation EMP_DEPT.

Answer.
a) $SSN^+ = \{SSN, ENAME, BDATE, ADDRESS, DNUMBER, DNAME, DMGRSSN\}$
$DNAME^+ = \{DNAME\}$
b) It is minimal.
c) Since every member of a department has a reference to the manager of that department (i.e., Dmgrssn), when the department manager changes this reference must be changed multiple places. This leads to the possibility of an inconsistency in the database if they are not all changed.
d) You cannot enter data about a department until you have employees for the department.
e) If you delete the last employee for a department, you lose all information about the department.

Question 14 You are given the following functional dependencies for the "EMPLOYEE" relation. Explain whether the relation "EMPLOYEE" is BCNF and 3NF?
Database:

```
EMPLOYEE(ssn, first-name, last-name, address, date-joined, supervisor-ssn)
DEPARTMENT(dept-no, name, manager-ssn)
WORKS-IN(employee-ssn, dept-no)
INVENTORY(dept-no, item-id, quantity)
ITEMS(item-id, item-name, type)
```

Foreign keys:
1. EMPLOYEE.supervisor-ssn and WORKS-IN.employee-ssn point to EMPLOYEE.ssn.
2. WORKS-IN.dept-no and INVENTORY.dept-no point to DEPARTMENT.dept-no.
3. INVENTORY.item-id points to ITEMS.item-id.
$\{ssn \rightarrow supervisor - ssn, ssn \rightarrow first - name, ssn \rightarrow last - name, ssn \rightarrow date - joined, ssn \rightarrow address, address \rightarrow ssn\}$.

Answer. In BCNF, since ssn and address are both keys of EMPLOYEE.