CS 377: Database Systems

Homework #3 Solutions

1. **Functional Dependencies via Question 14.26** (10 + 8 points):

   Consider the following relation:

Tuple #	A	B	C
1	10	b1	c1
2	10	b2	c2
3	11	b4	c1
4	12	b3	c4
5	13	b1	c1
6	14	b3	c4

   (a) Given the above database content, which of the following functional dependencies **may hold** in the above relation. If the functional dependency is invalid, explain why by specifying the tuples that cause the violation.

   i. A $\rightarrow$ B
   ii. B $\rightarrow$ C
   iii. C $\rightarrow$ B
   iv. B $\rightarrow$ A
   v. C $\rightarrow$ A

   (b) Does the above relation have a potential candidate key that does not include all attributes in the relation? If it does, what is it? If it does not, why not?

   (**ANSWER**)

   (a) i. CANNOT HOLD. Tuples 1 and 2 have the same value for attribute A but different values for B.
   ii. MAY HOLD. For the tuples that have the same value for attribute B (1 and 5, 4 and 6), the same values are present for C.
   iii. CANNOT HOLD. Tuples 1,3, and 5 have the same value for attribute C (c1), but different values for B (b1 and b4).
   iv. CANNOT HOLD. Tuples 1 and 6 have the same value for attribute B (b1), but different values of A (10, 13).
   v. CANNOT HOLD. Tuples 1 and 5 have the same value for attribute C (c1), but different values of A (10, 13).

   (b) Potential keys are (A, B) and (A, C). (B,C) is not a potential key because of the tuples: (10, b1, c1) and (13, b1, c1).

2. **Closures adapted from Question 14.27** (6 + 6 + 6 points):

   Consider a relation:

   `R(A, B, C, D, E)`

with the following dependencies:

- A, B → C
- C, D → E
- D, E → B

(a) Compute the closure $\{A, B\}^+$

(b) Compute the closure $\{C, D\}^+$

(c) Compute the closure $\{D, E\}^+$

(**ANSWER**)

(a) $\{A, B\}^+ = \{A, B, C\}$

(b) $\{C, D\}^+ = \{C, D, E, B\}$

(c) $\{D, E\}^+ = \{D, E, B\}$

3. **Keys and BCNF Normalization via Question 14.24** (10 + 15 points):

Consider the relation:

`R(A, B, C, D, E, F, G, H, I, J)`

and the following dependencies:

- A, B → C
- A → D, E
- B → F
- F → G, H
- D → I, J

(a) What is the key(s) in R?

(b) Decompose R losslessly into BCNF and identify the keys for each new relation.

(**ANSWER**)

(a) Using heuristic #2, we can see that the base set must be $\{A, B\}$. If we compute the closure, $\{A, B\}^+ = \{A, B, C, D, E, F, G, H, I, J\}$, we can see that all relations are encompassed by these two. For safety, check that each one individually would not be a key: $\{A\}^+ = \{A, D, E, I, J\}$ and $\{B\}^+ = \{B, F, G, H\}$.

(b) Check each of the functional dependencies to see if it breaks the BCNF rules.

- A, B → C: okay as (A, B) is a key.
- A → D, E: violation as A is a subset of the key and D is a non-key attribute.

Decompose based on the violation A → D, E using the closure set $\{A\}^+$, we have two new relations $R1(\underline{A}, D, E, I, J)$ and $R2(\underline{A}, \underline{B}, C, F, G, H)$. Check the individual relation $R1$ via the functional dependencies:

- D → I, J: violation as D is not a key, so it's an embedded entity.

Decompose $R1$ further into 2 parts based closure set $\{D\}^+$ yields two new relations $R11(\underline{A}, D, E)$ and $R12(\underline{D}, I, J)$. Now $R11$ and $R12$ are in BCNF. Time to check individual relation $R2$ via the functional dependencies:

- B → F: violation as B is a subset of the key and F is a non-key attribute.

Decompose $R2$ further into 2 parts based closure set $\{B\}^+$ to obtain two new relations $R21(\underline{B}, F, G, H)$ and $R22(\underline{A}, \underline{B}, C)$. $R22$ is in BCNF so we need to check $R21$:

- F $\rightarrow$ G, H: violation as F is not a key.

Decompose $R21$ further into 2 relations based on closure $\{F\}^+$ to get two new relations $R211(\underline{F}, G, H)$ and $R212(\underline{B}, F)$, both of which are BCNF. Thus the final answer is:

- $R11(\underline{A}, D, E)$
- $R12(\underline{D}, I, J)$
- $R211(\underline{F}, G, H)$
- $R212(\underline{B}, F)$
- $R22(\underline{A}, \underline{B}, C)$

4. **BCNF** (10 + 7 + 15 + 7 points):

Consider the following relation:

```
R(A, B, C, D, E, F, G, H)
```

and the following dependencies:

- B $\rightarrow$ C, D
- B, F $\rightarrow$ H
- C $\rightarrow$ A, G
- C, E, H $\rightarrow$ F
- C, H $\rightarrow$ B

(a) What are the key(s) of the relation?

(b) Which of these functional dependencies violate BCNF?

(c) Decompose the relation to obtain a lossless decomposition of R that are in BCNF. Make sure it is clear what the keys are for each relation.

(d) Is the resulting decomposition functional dependency-preserving? Explain why it does or does not.

(**ANSWER**)

(a) We will illustrate the first heuristic to figure out the keys. Find the closure of each of the FDs.

- {B}$^+$ = B, C, D, A, G
- {B, F}$^+$ = B, C, D, A, G, H, F
- {C}$^+$ = C, A, G
- {C, E, H}$^+$ = C, E, H, F, A, G, B, D
- {C, H}$^+$ = C, H, B, A, G, D

Second step is to add / subtract based on the closure of each set. So for the first one {B}$^+$, we are missing E, F, and H.

- Add E: {B, E}$^+$ = B, C, D, A, G, E
- Add F: {B, F}$^+$ = B, C, D, A, G, H, F
- Add H: {B, H}$^+$ = B, C, D, A, G, H
- Add E, F: {B, E, F}$^+$ = B, C, D, A, G, H, F, E which is a key!
- Add E, H: {B, E, H}$^+$ = B, C, D, A, G, E, H, F which is also a key!

We can skip the second closure set (covered by the previous one) and move onto the third one $\{C\}^+$, which is missing B, D, E, F, H.

- Add B: $\{C, B\}^+ = C, A, G, B, D$
- Add D: $\{C, D\}^+ = C, A, G, D$
- Add E: $\{C, E\}^+ = C, A, G, E$
- Add F: $\{C, F\}^+ = C, A, G, F$
- Add H: $\{C, H\}^+ = C, H, B, A, G, D$
- Add B, E: $\{C, B, E\}^+ = C, A, G, B, D, E$
- Add B, F: $\{C, B, F\}^+ = C, A, G, B, D, F, H$
- Add D, E: $\{C, D, E\}^+ = C, A, G, D, E$
- Add D, F: $\{C, D, F\}^+ = C, A, G, D, F, B, H$
- Add D, H: $\{C, D, H\}^+ = C, A, G, D, H, B$
- Add E, F: $\{C, E, F\}^+ = C, A, G, E, F$
- Add E, H: $\{C, E, H\}^+ = C, E, H, F, A, G, B, H$ which means this is a key!
- Add F, H: $\{C, F, H\}^+ = C, H, B, A, G, D, F$
- Add B, E, F: $\{C, B, E, F\}^+$ which would be a superkey from (B,E,F), so we ignore this one.
- Add D, E, F: $\{C, D, E, F\}^+ = C, A, G, D, E, F$
- Add D, E, H: $\{C, D, E, H\}^+$ which would be a superkey from (C, E, H) so skip.

Since the fourth one was proved to be a key, we can skip this and goto the last one. Note that since we enumerated C, H and the various ones, we can also skip this as well.

Thus we found 3 keys, (B, E, F), (B, E, H), and (C, E, H).

(b) BCNF requires that we check all the functional dependencies.

- B is not a super key, but part of a key, so it violates BCNF.
- B, F is a part of a key, but not a key itself, so it too violates BCNF.
- C is a part of a key, but A is not part of a key, so this is also a violation of BNC.
- (C, E, H) is okay since it is a key
- C, H, is also part of a key and thus violates BCNF.

(c) We will use the previous part to decompose our relation.

- $B \rightarrow C, D$ is a violation and $\{B\}^+ = B, C, D, A, G$ so we obtain two relations R1(A, $\underline{B}$, C, D, G) and R2(B, E, F, H)
- Check the FDs against R1. $C \rightarrow A, G$ is a violation of BCNF form. Thus we split R1 into two further relations: R11($\underline{C}$, A, G) and R12 ($\underline{B}$, C, D).
- Check R11 against the FDs, note that they are okay, so R11 is BCNF.
- Check R12 against the FDs, and since they are also okay, R12 is also BCNF.
- Check R2 against the FDs and note that $B, F \rightarrow H$ is a violation of BCNF. Decompose R2 into two further relations, R21($\underline{B}$, $\underline{F}$, H) and R22($\underline{B}$, $\underline{E}$, $\underline{F}$).
- Check R21 against the FDs, note that they are okay, so R21 is BCNF.
- Check R22 against the FDs, note that they are okay, so R22 is BCNF.

Thus, the resulting decomposition is: R11($\underline{C}$, A, G), R12 ($\underline{B}$, C, D), R21($\underline{B}$, $\underline{F}$, H) and R22($\underline{B}$, $\underline{E}$, $\underline{F}$).

(d) Note that our BCNF decomposition from above does not preserve all the functional dependencies. For the first three functional dependencies, there is a relation that includes all of the functional dependencies attributes, ensuring that the functional dependency is preserved. However, at least one functional dependency is not preserved: C, E, $H \rightarrow F$. Note that we can construct a valid instance of the relations that when joined does not preserve the functional dependency. For example:

B	C	D
b	1	c
f	1	g

C	A	G
1	a	d

E	B
2	b
2	f

B	F	H
b	4	3
f	5	3

The natural join of the four tables yields the following:

A	B	C	D	E	F	G	H
a	b	1	c	2	4	d	3
a	f	1	g	2	5	d	3

Note that in the resulting table, the relation violates the functional dependency C, E, H $\rightarrow$ F and C, H $\rightarrow$ B.