Solution to Questions of Fall-02's Test

1. For the relation schema T(A,B,C,D,E) with the following FDs

$$\mathcal{F} = AB \rightarrow C, DE \rightarrow C, B \rightarrow D$$

do the following:

a. Indicate at least 10 non-trivial BCNF violations

b. Decompose T into a collection of BCNF relations

c. Among the 10 non-trivial BCNF violations of your answer to (a) indicate those that violate the 3NF condition.

Answer:

a. We have that $AB^+ = \{A, B, C, D\}$, $DE^+ = \{D, E, C\}$, $B^+ = \{B, D\}$. This shows that none of the given FDs has its left hand side as a superkey. Furthermore, none of the given FDs are trivial (a trivial FD is a FD whose left hand side is a superset of the right hand side). As such, we have three non-trivial FDs

$$AB \rightarrow C, DE \rightarrow C, B \rightarrow D$$

which violate the BCNF conditions.

Since we need to find 10 violations, we need to find other FDs that can be derive from the given FDs and violate the BCNF condition.

Based on the computation of the three closures, we know that $AB \rightarrow D$ is another FD that violates the BCNF condition as well.

This is not enough either. We need to find more! Now we can use the Amstrong's axioms to derive the FDs. For example, the augmentation rule say that $ABD \rightarrow C$ is another FD of the relation. Because $ABD^+ = \{A, B, C, D\}$ we can conclude that this is another non-trivial FD that violates the BCNF condition.

Similar to the above, we have $DEB \rightarrow C$ is another non-trivial BCNF violation.

So far, we find 6. How can we find another 4? For the same reason, we can see that $BC \rightarrow D, BE \rightarrow D$ are another two non-trivial BCNF violations.

Still, 2 more? Remember that non-trivial FDs require only that the left hand side is not a superset of the right hand side. As such, we can also use the augmentation rule to get $BA \rightarrow DA$ or $BC \rightarrow DC$, for example, which we can check that they are non-trivial BCNF violations.

Note: The above is my reasoning. It is enough if you provide the answer as follows.

- $AB \rightarrow C$ is a non-trivial FD since $C \not\subseteq AB$ and it violates the BCNF because $AB^+ = \{A, B, C, D\}$ which implies that AB is not a superkey.
- $BA \rightarrow DA$ is a FD entailed by $\mathcal{F}$ (augmentation rule on $B \rightarrow D$) and violates the BCNF because $AB^+ = \{A, B, C, D\}$ which implies that AB is not a superkey.
- etc.

b. Because $AB \rightarrow C$ violates the BCNF condition, we can decompose T using $AB \rightarrow C$ into T_1 and T_2 as follows

- $T_1 = (ABC, \{AB \rightarrow C\})$
- $T_2 = (ABDE, \{AB \rightarrow D, B \rightarrow D\})$

T_1 has only one FD which satisfies the BCNF condition. So, T_1 is in BCNF.

For T_2, we have that $B^+ = \{B, D\}$. Furthermore, D is not a subset of B. Thus, $B \rightarrow D$ violates the BCNF condition. We need to decompose T_2. Using $B \rightarrow D$, we have the decomposition $T_3 = (BD, \{B \rightarrow D\})$ and $T_4 = (ABE, \emptyset)$. Both of these relations are in BCNF. So, the final decomposition of T is T_1, T_3, T_4.

c. To answer this question, we need to find the keys of T. Observe the FDs in $\mathcal{F}$, we see that A, B, and E do not belong to the right hand side of any FDs. This implies that ABE has to be a part of any key of T. Furthermore, $ABE^+ = \{A, B, E, C, D\}$. This means that ABE is the only key of the relation.

Insert: A FD satisfies the 3NF condition if (i.) it is a trivial FD (right hand side is subset of left hand side); or (ii) left hand side is a super key; or (iii) every attribute in the right hand side is part of a key).

Looking at the answers from (a), every FD has either C or D on the right hand side which is not a part of any key. As such, none of them will satisfy the 3NF condition.

2. Suppose that we have a relation $R(B, C, D)$ with the MVD $D \twoheadrightarrow B$. If we know that the current instance of R contains the tuples (b_1, c_1, d_1), (b_2, c_2, d_1), (b_3, c_3, d_2), and (b_4, c_4, d_2). List all other tuples that must be contains in this instance.

The tuples that need to be in this instance (besides the given four) are:

- (b_1, c_2, d_1) (because of (b_1, c_1, d_1) and (b_2, c_2, d_1))
- (b_2, c_1, d_1) (because of (b_1, c_1, d_1) and (b_2, c_2, d_1))
- (b_3, c_4, d_2) (because of (b_3, c_3, d_2) and (b_4, c_4, d_2))
- (b_4, c_3, d_2) (because of (b_3, c_3, d_2) and (b_4, c_4, d_2))