DBMS I — Homework 8

December 7, 2007

The purpose of this homework is for you to practice the 3NF and 4NF decomposition algorithm.

Let $R(A, B, C, D, E)$ be a relation schema with the following set of FDs and MVDs:

$$F = \left\{ \begin{array}{l} AB \rightarrow C \\ DE \rightarrow C \\ B \twoheadrightarrow E \end{array} \right\}$$

1. Given an example of a legal instance of R.

   Answer: There are many ways to answer this question. The table must satisfy the first two conditions, which is easy to get. The third condition implies that if two rows agree on B then every combination of (ACD) and (E) needs to be present in the database. Notice that the FD and MVD only say that "if the left hand side of two rows are equal then ...". As such, a *possible answer* is to have an instance with **ONE** row like the following:

A	B	C	D	E
1	2	3	1	0

   Well, if you feel not satisfy, you can add something to it, make sure that it satisfies the first two conditions and the B's are different. If the B's are different, then the first FD satisfies trivially; The second one can be made satisfied by making DE different on different rows.

A	B	C	D	E
1	2	3	1	0
1	3	3	4	0

   Still, if you insist on having some rows with B equals, here is an example; we start with the following:

A	B	C	D	E
1	2	3	1	1
2	2	3	4	0

   This one satisfies $AB \rightarrow C$ (because (A, B) is different on two rows). It also satisfies $DE \rightarrow C$ for the same reason. Let us look at $B \twoheadrightarrow E$. B is the same on two rows. So, we should have two more combinations: $(1, 2, 3, 1, 0)$, $(2, 2, 3, 4, 1)$. This gives us:

A	B	C	D	E
1	2	3	1	1
2	2	3	4	0
2	2	3	4	1
1	2	3	1	0

2. Find a canonical cover for the set of FDs belonging to F.

   Answer: First, notice that the set of FDs is $\mathcal{F} = \{AB \to C, DE \to C\}$.

   The canonical cover for the set $\mathcal{F}$ is itself since there are no extraneous attributes in either $AB \to C$ or $DE \to C$.

3. Decompose R into a set of 3NF.

   Answer: To answer this question, we need to find the keys of the relations. Observe that none of the FDs contain A, B, D, E on the right hand side, any key will need to contain these four attributes. On the other hand, $ABDE^+ = ABCDE$. So, we can conclude that $ABDE$ is the unique key of the relation.

   Applying the algorithm to decompose the relation into 3NF, we go through the following iterations:

   - $i = 0$; $R_1 = ABC$ (the first FD in the canonical cover); $i = i + 1 = 1$
   - $i = 1$; $R_2 = DEC$ (the second FD in the canonical cover); $i = i + 1 = 2$
   - Both R_1 and R_2 do not contain the key; so, we add $R_3 = ABDE$ to the decomposition set.

   The result is: $ABC, DEC, ABDE$.

4. Decompose R into a set of 4NF relations.

   Answer: First, we need to find all FDs and MVDs that satisfied by the relation. We know that if $\alpha \to \beta$ then $\alpha \twoheadrightarrow \beta$. So, we first have:

$$D = \left\{ \begin{array}{lll} AB \to C & DE \to C & \\ AB \twoheadrightarrow C & DE \twoheadrightarrow C & B \twoheadrightarrow E \end{array} \right\}$$

   Transitivity does not add anything. Complementation adds

$$\left\{ \; AB \twoheadrightarrow DE \quad DE \twoheadrightarrow AB \quad B \twoheadrightarrow ABC \; \right\}$$

   So, we have the closure of D

$$D^+ = \left\{ \begin{array}{lll} AB \to C & DE \to C & \\ AB \twoheadrightarrow C & DE \twoheadrightarrow C & B \twoheadrightarrow E \\ AB \twoheadrightarrow DE & DE \twoheadrightarrow AB & B \twoheadrightarrow ADC \end{array} \right\}$$

   Clearly, the relation is not in 4NF. For example, the MVD $B \twoheadrightarrow E$ does not satisfy the 4NF condition.

   The first decomposition (using $B \twoheadrightarrow E$) yields $R_1 = BE$ and $R_2 = BACD$. The first one is already in 4NF. So, we will look at R_2.

   We need to find the FDs and MVDs in D^+ that R_2 needs to satisfy. By definition, all FDs in D^+ consisting of only attributes in R_2 and all MVDs that contain only attributes on the left in R_2 should be considered. The right hand side is the result of the intersection between it and R_2. So, the restriction of D on R_2 is

$$D^+ = \left\{ \begin{array}{ll} AB \to C & \\ AB \twoheadrightarrow C & \\ AB \twoheadrightarrow D & B \twoheadrightarrow AC \end{array} \right\}$$

   R_2 satisfies $AB \to C$. Thus, ABD is the key of this relation. $AB \twoheadrightarrow C$ is a MVD for this relation. So, this leads to the decomposition of R_2 to $R_{21} = ABC$ and $R_{22} = ABD$.

   R_{21} is in 4NF (the MVD $AB \twoheadrightarrow C$ satisfy the 4nf condition).

   R_{22} is also in 4NF since there is no non-trivial MVD related to ABD.

   So, the decomposition is BE, ABC, ABD.