Solutions for Exercise No. 8

2nd May 2017

1 Normal Forms

(a) Consider relation schema $R_1(A, B, C, D)$, where all attributes are atomic, and the following set of functional dependencies:

$$F_1 = \{A \to B,\ A \to C,\ C \to D\},$$

(b) Consider relation schema $R_2(A, B, C, D)$, where all attributes are atomic, and the following set of functional dependencies:

$$F_2 = \{\ CB \to AD,\ AB \to CD,\ D \to A\}.$$

For each relation schema and its set of functional dependencies (R_1 and F_1; R_2 and F_2) perform the following tasks:

(a)
1. Determine the set S of all candidate keys for the relation schema.

   $$S = \{A\}.$$

2. Determine the highest normal form (no NF, 1NF, 2NF, 3NF, BCNF) that the relation schema is in. Explain your answer.

   R_1 is in 2NF. As all attributes are atomic, the relation schema is at least in 1NF. As the relation has only one candidate key, that contains only one attribute, there is no partial function dependency on a candidate key of R_1, and R_1 is in 2NF. The relation is not in 3NF, because the dependency $C \to D$ violates 3NF because all of the following statements are false for $C \to D$: it is trivial, C is a super key, D contained in a candidate key. Since 3NF is a prerequisite for BCNF, the relation is not in BCNF.

3. Decompose the relation schema into new relation schemas that satisfy the next higher normal form.

   The two new relation schemas: $R_{11}(A, B, C)$, $R_{12}(C, D)$.

4. Show the minimal set of functional dependencies for the decomposed relation schemas.

   FD of R_{11}: $FD_1 = \{A \to BC\}$, FD of R_{12}: $FD_2 = \{C \to D\}$.

5. List the candidate keys of the decomposed relation schemas.

   CK of R_{11}: $\{A\}$, CK of R_{12}: $\{C\}$.

6. Determine if the decomposition is a lossless join decomposition or/and dependency preserving. Explain your answer.

   The decomposition is dependency preserving. $\{\{A \to B,\ A \to C\} \cup \{C \to D\}\} = \{A \to B,\ A \to C,\ C \to D\}$, i.e., $(F_1|R_{11} \cup F_1|R_{12})$ is equal to F_1, thus $(F_1|R_{11} \cup F_1|R_{12})^+ = F_1^+$.

   The decomposition is a lossless join decomposition, because $(R_{12} \cap R_{11}) \to (R_{12} - R_{11})$ is equivalent to $C \to D$, which is in F_1^+.

(b) 1. Determine the set S of all candidate keys for the relation schema.

$$S = \{AB,\ BC,\ BD\}.$$

2. Determine the highest normal form (no NF, 1NF, 2NF, 3NF, BCNF) that the relation schema is in. Explain your answer.

   R_2 is in 3NF. As all attributes are atomic, the relation schema is at least in 1NF. As all the attributes are contained in candidates keys, R_2 is in 2NF. As each functional dependency from the $\{AB \to CD,\ CB \to AC\}$ has a superkey in the left side, and $\{D \to A\}$ has A in the right side, which is a part of the candidate key, the relation is in 3NF. The dependency $D \to A$ violates BCNF, because it's not trivial and D is not a superkey.

3. Decompose the relation schema into new relation schemas that satisfy the next higher normal form.

   The two new relation schemas: $R_{21}(B, C, D)$, $R_{22}(D, A)$.

4. Show the minimal set of functional dependencies for the decomposed relation schemas.

   FD of R_{21}: $FD_1 = \{BC \to D,\ BD \to C\}$, FD of R_{22}: $FD_2 = \{D \to A\}$.

5. List the candidate keys of the decomposed relation schemas.

   CK of R_{21}: $\{BC,\ BD\}$, CK of R_{22}: $\{D\}$.

6. Determine if the decomposition is a lossless join decomposition or/and dependency preserving. Explain your answer.

   The decomposition is not dependency preserving, because $AB \to CD$ is not in $(F_2|R_{21} \cup F_2|R_{22})^+$. To prove it, we compute the $\{A, B\}^+$ in $(F_2|R_{21} \cup F_2|R_{22})$.

   $(F_2|R_{21} \cup F_2|R_{22}) = \{BC \to D,\ BD \to C,\ D \to A\}$. The $\{A\}^+ = \{A\}$, the $\{B\}^+ = \{B\}$, and there are no formulas in $\{BC \to D,\ BD \to C,\ D \to A\}$, where $\{A, B\}$ is in the left side, thus $\{A, B\}^+ = \{A, B, AB\}$

   The decomposition is a lossless join decomposition, because $(R_{22} \cap R_{21}) \to (R_{22} - R_{21})$ is equivalent to $D \to A$, which is in F_2^+.

2 Functional Dependencies

1. Consider relation schema $R_3(A, B, C, D, F, E)$ and the set of functional dependencies:

$$F = \{A \to BC,\ B \to C,\ AD \to B,\ DF \to E,\ F \to D\}.$$

   Determine the minimal cover of F.

   The $F_{min} = \{A \to B,\ B \to C,\ F \to DE\}$.

2. Consider relation schema $R_4(A, B, C, D)$ and the two sets of functional dependencies:

$$H = \{A \to B,\ A \to BC,\ AB \to C,\ AC \to D,\ B \to C\},$$

$$G = \{CA \to B,\ BA \to D,\ B \to D,\ DB \to C\}.$$

Determine if $H \models G$, $G \models H$, or H equivalent to G ($H \Leftrightarrow G$). Explain your answer.

All statements $H \Leftrightarrow G$, $H \models G$, $G \models H$ are not correct.

$H \models G$ is not correct, because $B \to D$ can not be inferred from H. H is equivalent to $H' = \{A \to B,\ A \to C,\ A \to D,\ B \to C\}$, where D can be inferred only form A and A can not be inferred.

As the $H \models G$ is a prerequisite for $H \Leftrightarrow G$ by definition, $H \Leftrightarrow G$ is also not correct.

$G \models H$ is not correct, because G is equivalent to $G' = \{CA \to B,\ B \to D,\ B \to C\}$, and $A \to B$ can not be inferred from G'.

3 Multivalued Dependencies

Relation $catWinner$ with schema $CatWinner(CatName, Color, Award)$ stores information about cats: their names, their colors, and their awards. All cats have different names. All attributes in $CatWinner$ are atomic. The set of functional dependencies F for the relation schema $CatWinner$ is **empty**.

Consider the following assumptions:

(a) Assumption: a cat might have one color only.

(b) Assumption: a cat might have several colors.

Answer the following questions for each of assumptions:

1. How does the assumption change F?
2. What is the set of candidate keys of relation schema $CatWinner$ under the stated assumption?
3. What is the set of all multivalued dependencies from the attribute $CatName$ in relation schema $CatWinner$ under the stated assumption?

Explain your answers.

(a) Assumption: a cat might have one color only.

1. Since the color is unique for each cat and all cats have different names, $CatName \to Color$ holds. Thus, the new set of functional dependencies $F' = \{CatName \to Color\}$.
2. The set of the candidate keys is $\{\{CatName, Award\}\}$.
3. There are three non-trivial multivalued dependencies $CatName \twoheadrightarrow Color$, $CatName \twoheadrightarrow Award$, and $CatName \twoheadrightarrow Color, Award$. The first one is inferred from $CatName \to Color$, the second one is an augmentation of the first one. Additionally, there are two trivial multivalued dependencies $CatName \twoheadrightarrow CatName$ and $CatName \twoheadrightarrow CatName, Color, Award$.

(b) Assumption: a cat might have several colors.

1. This assumption does not change the set of functional dependencies.
2. The set of the candidate keys is $\{\{CatName, Color, Award\}\}$.
3. The non-trivial multivalued dependencies $CatName \twoheadrightarrow Color$, $CatName \twoheadrightarrow Award$, and $CatName \twoheadrightarrow Color, Award$ still hold. Assume, that a cat has several colors, then all tuples with all colors must be repeated for each new award. There are two trivial multivalued dependencies $CatName \twoheadrightarrow CatName$ and $CatName \twoheadrightarrow CatName, Color, Award$.