Comments on "Problem Decomposition Method to Compute an Optimal Cover for a Set of Functional Dependencies"

Xiaoning PENG, Zhijun XIAO
Department of Computer Engineering, Huaihua University, China
pxnxzj@gmail.com

Maier, in his seminal paper [1], presents that the problem of finding an optimal cover is NP-complete. Cotelea's paper [2] proposes a problem decomposition method for finding an optimal cover. The basic idea of the method is: (1) a big and intractable problem (optimal cover problem) is broken down into some smaller problems; (2) particular solutions of these smaller problems are combined to construct the initial problem solution. More specifically, the main idea of the method is: (1) a relational schema R with a set F of functional dependencies is partitioned in some subschema; (2) each subschema will be found determinants with the fewest attributes (including repeated); (3) substituting the groups of attributes in F that are determinants in some subschema with the shortest determinants, Cotelea states that this happens only for equivalence classes of functional dependencies containing the determinants in the left or right sides as subsets.

The idea embodied in this method is classic in the algorithm theory area, and Cotelea proposes many valuable results. However, are these sufficient for finding all optimal covers? (Seemingly unlikely)

Please see a counterexample as the following.

Example 1. Given $G=\{AD\to E, A\to BC, B\to A, C\to A\}$ that is a minimum and reduced set of functional dependencies, please find an optimal cover for G. For example, $H=\{AD\to E, A\to B, B\to C, C\to A\}$ is an optimal cover for G.

The contribution graph of G is shown in Figure 1.

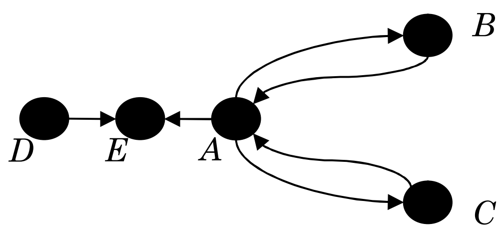

Fig. 1. A contribution graph for G

G is divided into two equivalence classes $G=G_1\cup G_2$, where $G_1=\{AD\to E\}$, and $G_2=\{A\to BC, B\to A, C\to A\}$. Note that G_1 and G_2 satisfy the strict partial order. The equivalence class G_1 precedes the equivalence class G_2.

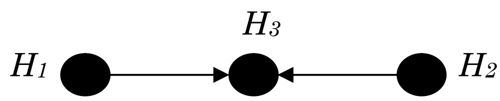

Fig. 2. Condensed graph of the graph from Figure 1

Figure 2 shows that H_1, H_2 and H_3 satisfy the strict partial order. The set of vertices of the graph in Figure 1 can be divided into three equivalence classes of attributes $S_1=\{D\}$, $S_2=\{A, B, C\}$ and $S_3=\{E\}$, and are reduced to the following sequence of non redundant equivalent classes of attributes $T_1=\{D\}$, $T_2=\{A, B, C\}$.

The set G of functional dependencies, below, is projected on the sets of attributes T_1 and T_2, resulting in the following sets of functional dependencies:

$\pi_{T1}(G)=\Phi$,

$\pi_{T2}(G)=\{A\to BC, B\to A, C\to A\}$.

Thus, for the non redundant classes of attributes, there were obtained the following sets of determinants $\{D\}$, $\{A, B, C\}$, respectively.
Now the groups of attributes that are determinants and part of dependencies in G are substituted by those with the smallest length. Substitutions occur in the equivalence classes of dependencies which precede corresponding class that has generated the determinant. Then, we can obtain an optimal cover $F=\{AD\rightarrow E, A\rightarrow BC, B\rightarrow A, C\rightarrow A\}$ for G. However, H rather than F is the optimal cover for G.
Example 1 tells us that the idea embodied in the Cotelea's paper is not sufficient for finding all optimal covers, in spite of Cotelea presents a correct example in his paper. The optimal cover problem is NP-complete, and it seems hard to find a deterministic algorithm for finding an optimal cover.

References

[1] D. Maier, "Minimum Covers in the Relational Database Model", Journal of the ACM, 1980, V.27, N 4, pp.664-674.
[2] Vitalie Cotelea, "Problem Decomposition Method to Compute an Optimal Cover for a Set of Functional Dependencies", Database Systems Journal, 2011, V.2, N 4, pp.17-30.