MTA SZTAKI Közlemények 35/1986 p.47-52

SOME REMARKS ON THE ALGORITHM OF LUCCHESI AND OSBORN

HO THUAN

Computer and Automation Institute, Hungarian Academy of Sciences

Let $S = \langle \Omega, F \rangle$ be a relation scheme where $\Omega = \{A_1, A_2, \ldots, A_n\}$ the universe of attributes, and $F = \{L_i \rightarrow R_i \mid L_i, R_i \subseteq \Omega,\ i = 1,2,\ldots,m\}$ - the set of functional dependencies.

In [2] C.L. Lucchesi and S.L. Osborn provided a very interesting algorithm to determine the set of all keys for any relation scheme $S = \langle \Omega, F \rangle$.

The algorithm has time complexity

$$O(|F|\ |K_S|\ |\Omega|(|K_S| + |\Omega|)),$$

following our notation, i.e. in time polynomial in $|\Omega|, |F|$ and $|K_S|$,
where

$|F|$ - the cardinality of F, and

K_S - the set of all keys for S.

We reproduce here this algorithm with some modifications in accordance to our notation.

Algorithm OL1. Set of all keys for $S = \langle \Omega, F \rangle$

Comment K_S is the set of keys being accumulated in a sequence which can be scanned in the order in which the keys are entered;

$K_S \leftarrow \{key^{*)}(\Omega,F,\Omega)\}$;

for each K in K_S *do*

for each FD $(L_i \rightarrow R_i)$ in F *do*

$T \leftarrow L_i \cup (K \setminus R_i)$;

test ← true;

for each J in K_S *do*

if T includes J *then* test ← false;

if test *then* $K_S \leftarrow K_S \cup \{key(\Omega,F,T)\}$

end

end;

return K_S.

The following simple remarks in some cases can be used to improve the performance of the algorithm of Lucchesi and Osborn.

Remark 1 To find the first key for $S = \langle\Omega,F\rangle$, instead of Ω, it is better to use the superkey $(\Omega\setminus R) \cup (L \cap R)$ and algorithm 1 in [1], and instead of the algorithm key (Ω,F,T), it is better to use algorithm 2 [1] for finding one key for S included in a given superkey T.

Remark 2 In [1] we have proved that

$$R \setminus L \subseteq \Omega \setminus H, \quad H = \bigcup_{K \in K_S} K$$

i.e., $R \setminus L$ consists only of non-prime attributes. Therefore if $R_i \subseteq R \setminus L$ then

$$R_i \cap K = \emptyset, \quad \forall K \in K_S.$$

and $$L_i \cup (K \setminus R_i) \supseteq K.$$

*) key (Ω,F,X) is the algorithm which determines a key for $S = \langle\Omega,F\rangle$ that is a subset of a specified superkey X.

That means, when computing $T = L_i \cup (K \setminus R_i)$, we can neglect all FDS $L_i \to R_i$ with $R_i \subseteq R \setminus L$, for every $K \in K_S$.
Let us denote

$$\bar{F} = F \setminus \{L_j \to R_j \mid L_j \to R_j \in F \text{ and } R_j \subseteq R \setminus L\}$$

Remark 3 With a fixed K in K_S, it is clear that if $K \cap R_i = \emptyset$ then $L_i \cup (K \setminus R_i) \supseteq K$. In that case it is not necessary to continue to check whether T includes J for each J in K_S.
So, it is better to compute T by the following order:

$$T = (K \setminus R_i) \cup L_i .$$

Remark 4. The algorithm of Lucchesi and Osborn is particularly effective when the number of keys for $S = <\Omega, F>$ is small. But basing on what information one can conclude that the number of keys for $S = <\Omega, F>$ is small? There is no general answer for all the cases, and it is shown in [3] that the number of keys for a relation scheme $S = <\Omega, F>$ can be factorial in $|F|$ or exponential in $|\Omega|$, and that both of these upper bounds are attainable.
However, it is shown (in [1], Corollary 1) that

$$|K_S| \leq C_h^{[h/2]}$$

where h is the cardinality of $L \cap R$.
Thus if $L \cap R$ has only a few elements then it is a good criterion for saying that S has a small number of keys.
In the case $L \cap R = \emptyset$, $\Omega \setminus R$ is the unique key for $S = <\Omega, F>$ as pointed out in [1, Corollary 4].

Example We take up the example in [2, Appendix 1]

$$\Omega = \{a,b,c,d,e,f,g,h\}$$

$$F = \{a \to b,\ c \to d,\ e \to f,\ g \to h\}$$

It is clear that for this relation scheme

$$L \cap R = \emptyset,$$

and it has exactly one key, namely aceg. Taking into account the remarks 1, 2, 3 just mentioned, the above algorithm can be modified as follows.

Algorithm OL2. Set of all keys for $S = \langle\Omega,F\rangle$;

```
K_S ← {Algo1*) (Ω,F,(Ω \R)(L ∩ R))}
for each K in K_S do
   for each FD (L_i → R_i) in F̄ such that  K \ R_i ≠ K do

         T ← (K \ R_i) ∪ L_i;
         test ← true;
         for each J in K_S do
            if T includes J then test ← false;
         if test then K_S ← K_S ∪ {Algo2*)(Ω,F,T)}
   end
end;
return K_S.
```

Remark 5 The Algorithm OL2 now has time complexity

$$O(|K_S|\,|\Omega|\,(|K_S|\,|\bar{F}| + |F|\ |L \cap R|))$$

---

*) Algo 1 and Algo 2 refer to Algorithm 1 and Algorithm 2 in [1] respectively.

R E F E R E N C E S

[1] Ho Thuan and Le van Bao, Some results about keys of relational schemas, Acta Cybernetica, Tom. 7, Fasc. 1, Szeged, 1985, pp. 99-113.

[2] Lucchesi, C.L., Osborn, S.L., Candidate keys for relations, J. of Computer and System Sciences, 17. 1978, pp. 270-279.

[3] Osborn, S.L., Normal forms for relational databases. Ph.D. Dissertation, University of Waterloo, 1977.

Néhány megjegyzés a Lucchesi-Osborn algoritmushoz

Ho Thuan

Összefoglaló

Az [1]-ben két algoritmust javasoltunk az adott szuper-kulcsban levő kulcsok megkeresésére. Ebben a cikkben, felhasználva a két algoritmust és néhány egyszerü megjegyzést, megvitatjuk a Lucchesi-Osborn féle algoritmust [2].

Несколько замечаний об алгоритме Луччеси и Осборна

Хо Тхуан

Р е з ю м е

В статье [1] мы дали два алгоритма для нахождения ключа реляционной схемы, который содержится в данном супер-ключе. В этой статье, используя эти алгоритмы и несколько простых замечаний, улучшается алгоритм Луччеси и Осборна [2].