Tạp chí Tin học và Điều khiển học, T. 13, S. 4 (1997) (98–102)

ON THE HYPERGRAPHS AND CANDIDATE KEYS

VU DUC THI

Abstract. The combinatorial problems are interesting in the relational datamodel. The theory of hypergraphs was a very useful tool for the solution of combinatorial problems. The transversal and the minimal transversal are important concepts in this theory.

In this paper, base on hypergraph we give some new characterizations of the set the candidate keys in the relational datamodel.

Key words and Phrases: hypergraph, transversal, minimal tranversal, independent set, relation datamodel, functional dependency, relation scheme, closure, key, candidate key.

1. INTRODUCTION

Let us give some necessary definitions that are used in the next sections. The concepts give in this session can be found in [1, 2, 3, 4, 7, 9, 10, 15, 16, 17].

Let R be a nonempty finite set and $P(R)$ is power set. The family $H = \{E_i : E_i \in P(R),\ i = 1, ..., m\}$ is called a hypergraph over R if $E_i \neq \emptyset$. (In [4] author requires that the union of E_{ts} is R. In this paper we do not).

A hyperaph H is simlpe if $E_t \subset E_j$ implies $i = j$.

The elements of R are called vertices, and the sets $E_t, ..., E_m$ are the edges of the hypergraph H.

It is easy to seen that a simple graph is simple hypergraph with $|E_t| = 2$.

Let $H = \{E_1, ..., E_m\}$ be a hypergraph over R. Set

$$m(H) = \{E_t \in H : E_j \in H : E_j \subset E_t\}$$

It can be seen that $m(H)$ is simple hypergraph and the family H uniquely determines the family $m(H)$.

Let H be a hypergraph over R. A set $A \subseteq R$ is called a stransversal of H (Sometime it is called a hitting set) if $E \in H$ implies $A \cap E \neq \emptyset$.

The family of all minimal transversals of H is called the transversal hypergraph of H, and denoted by $\mathrm{tr}(H)$. Clearly, $\mathrm{tr}(H)$ is a simple hypergraph.

Let $R = \{a_1, ..., a_n\}$ be a nonempty finite set of attributes. A functional dependency is a statement of the from $A \to B$, where $A, B \subseteq R$. The FD $A \to B$

holds in a relation $r = \{h_1, ..., h_m\}$ over R if $\forall h_i,\ h_j \in r$ we have $h_i(a) = h_j(a)$ for all $a \in A$ impiles $h_i(b) = h_j(b)$ for all $b \in B$. We also say that r satisfies the FD $A \to B$.

Let F_r be a family of all FDs that hold in r. Then $F = F_r$ satisfile

(1) $A \to A \in F$,

(2) $(A \to B \in F,\ B \to C \in F) \Rightarrow (A \to C \in F)$,

(3) $(A \to B \in F,\ A \subseteq C,\ D \subseteq B) \Rightarrow (C \to D \in F)$,

(4) $(A \to B \in F,\ C \to D \in F) \Rightarrow (A \cup C \to B \cup D) \in F$.

A family of FDs satisflying (1) - (4) is called an f-family (some times it is called the full family) over R.

Clearly, F_r is an F-family over R. It is known [1] that if F is an arbitrary f-family, then there is a relation r over R such that $F_r = F$.

Given a family F of FDs, there exists an unique minimal f-family F^+ that contains F. It can be seen that F^+ contain all FDs which can be derived from F by the rules (1) - (4).

A relation scheme s is a pair $\langle R,\ F\rangle$ where R is a set of attributes, and F is a set of FDs over R. Denote $A^+ = \{a : A \to \{a\} \in F^+\}$. A is called the closure of A over s. It is clear that $A \to B \in F^+$ iff $B \subseteq A^+$.

Clearly, if $s = \langle R,\ F\rangle$ be a relation scheme, then there is a relation r over R such that $F_r = F^+$ (see [1]).

Let r be a relation, $s = \langle R,\ F\rangle$ be a relation scheme. Then A is a key of r (a key of s) if $A \to B \in F_r$ ($A \to R \in F^+$). A is a candidate key of $r(s)$ if A is a key of $r(s)$ and any proper subset of A is not a key of $r(s)$.

Denote K_r (K_s) the set of all candidate keys of $r(s)$.

It can be seen that K_r, K_s are simple hypergraph over R.

Let $I \subseteq P(R)$, $R \in I$, and $A, B \in I \Rightarrow A \cap B \in I$. I is called a meet-semilattice over R. Let $M \subseteq P(R)$. Denote $M^+ = \{\cap M : M \subseteq M\}$. We say that M is generator of I if $M^+ = I$. Note that $R \in M^+$ but not in M by convention it is the intersection of the empty collection of sets.

Dentote $N = \{A \in I : A \neq \cap\{A \in I : A \subset A'\}\}$. It can be seen that N is the unique minimal generator of I.

2. HYPERGRAPHS AND CANDIDATE KEYS

The keys and the candidate keys play an essential role in the relational data-model. They are used to distinguish, find, magage records in relations.

Base on serults, prensented in [18], in this section, we give some new charac-terrizations and the properties on the candidate keys.

Let $s = \langle R, F\rangle$ be a relation scheme and r a relation over R. For every $A \subseteq R$, set $I(A) = \{a \in R : A \to \{a\} \notin F^+\}$. Then $I(A)$ is called the independent set of s. For r, put $I(A) = \{a : A \to \{a\} \notin F_r\}$.

Denote by I_s the family of all independent sets of s.

Set $m(s) = \{B \in I_s : B \neq \emptyset, \exists C \in I_s n : C \subset B\}$. $m(s)$ is called the family of all independent sest of s.

It can be seen that A is a key of s if and only if $I(A) = \emptyset$.

Denote by I_r and $m(r)$ the family of all independent sets of s.

Set $m(s) = \{B \in I_s : B \neq \emptyset, \exists C \in I_s : C \subset B\}$. $m(s)$ is called the family of all independent sets of s.

It can be seen that A is a key of s and only if $I(A) = \emptyset$

Denote by I_r and $m(r)$ the family of all independents sedt and the family of all independent sets of r.

Theorem 2.1. *Let $s = \langle R, F\rangle$ be a relation scheme over R. Then*

$$tr(K_s) = m(s)\,.$$

Proof: Clearly, $m(s)$ is a simple hypergraph over R. It can be seen that from the definitions of I_s and keys, if $D \in I_s$ and $D \neq \emptyset$, then $R - D$ is closure over s. Consequently, $R - A$ is not a keys of s (*).

Assume that A is an element of $\mathrm{tr}(K_s)$, i.e., for all $B \in K_s : A \cap B \neq \emptyset$, and it is minimal for this property. From these, we can see that $R - A$ is not a key od s. Clearly, $A \neq \emptyset$. Hence, $(R - A)^+ \neq R$ holds. If $R - A \subset (R - A)^+$, then $R - (R - A)^+ \cap B \neq \emptyset$ for all $B \in K_s$. This contradicts $A \in \mathrm{tr}(K_s)$. Thus, $R - A = (R - A)^+$ holds. According to the definition of the independent set there is a C such that $I(C) = A$. Thus, $A \in I_s$ holds.

Suppose that there exists a $D \neq \emptyset$, $D \in I_s$ and $D \subset A$. From (*) $R - D$ is not a key of s. Consequently, D is a transversal of K_s. This contracdicts $A \in \mathrm{tr}(K_s)$. Hence, $A \in m(s)$ holds.

Conversely, assuma that $a \in m(s)$. It is obvious that $A \neq 0$. According to (*) we obtain $A \cap B \neq \emptyset$ for all $B \in K_s$. Thus, A is transversal of K_s.

Suppose that there is $D \in \mathrm{tr}(K_s)$ and $D \subset A$. By the above proof we obtain $D \in m(s)$. This conflicts with the fact that $m(s)$ is a simple hypergraph. Hence, $A \in \mathrm{tr}(K_s)$ holds. Our proof is complete.

It is know [4] that if H, H' are two simple hypergaphs over R, then $H = \mathrm{tr}(H')$ if and only if $H' = \mathrm{tr}(H)$. From this, we obtain

Corollary 2.2. *Let $s = \langle R, F\rangle$ be a relation scheme over R. Then $K_s = tr(m(s))$.*

Remark 2.3. Let H be a simple hypergraph over R. We define the next family of H, denoted H^{-1}, as follows:

$$H^{-1} = \{A \subset R : (B \in H) \Rightarrow (N \notin A) \text{ and } (A \subset C) \Rightarrow (\exists B \in H)(B \subseteq C)\}.$$

It is easy to see that H^{-1} is also simple hypergraph over R.

It can be seen that if H is a simple hypergraph over R, then from the definition of $\mathrm{tr}(H)$ we obtain $H^{-1} = \{R - A : A \in \mathrm{tr}(H)\}$.

Remark 2.4. Let $s = \langle R, F\rangle$ be a relation scheme over R.
Set $Z_s = \{A^+ : A \subset R\}$, i.e., Z_s is the set of all closure of s. Put $T_s = \{A \in Z_s : A \neq R \; \exists \in Z_s : A \subset B\}$. Thus, T_s is the set of all maximal element of $Z_s - R$. By the definition of the independent set of s, we can see that $T_s = \{R - B : B \in m(s)\}$.

REFERENCES

1. Armstrong W. W. *Dependency structures of Database Relationship*, Information Processing, Holland Publ. Co,, 74 (1994) 580-583.
2. Beeri C., Bernstein P. A., *Comutainonal problems related to the design of normal form relations schemas*, ACM Trans on Database Syst., 4, 1 (1979) 30-59.
3. Beeri C., Dowd M., Fagin R., Staman R., *On the structure of Armstrong relations for functional dependencies*, J. ACM, 31, 1 (1984) 30-46.
4. Berge C., *Hypergraph: combinatories of finite sets*, Amsterdam, 1989.
5. Demetrovics J., Thi V. D., *Some result about functional dependencies*, Acta Cybernetica, 8, 3 (1988) 273-278.
6. Demetrovics J., Thi V. D., *Relations and minimal keys*, Acta Cybernetica, 8, 3 (1988) 279-285.
7. Demetrovics J., Thi V. D., *On keys in the relational datamodel*, Inform. Process. Cybern. EIK, 24, 10 (1988) 515-519.
8. Demetrovics J., Thi V. D., *Algorithm for generating Armstrong relations and inferring functional dependencies in the relational datamodel*, Computers and Mathematics Applications, Great Britain, 26, 4 (1993) 43-45.
9. Demetrovics J., Thi V. D., *Some problems concerning keys for relations scheme and relations in the datamodel*, Information Processing letters, North Holland, 46, 4 (1993) 179-183.
10. Demetrovics J., Thi V. D., *Some computational problems related to the functional dependency in the relational datamodel*, Acta Acientiarum Mathematicarum, 57, 1-4 (1993) 627-628.
11. Demetrovics J., Thi V. D., *Armstrong relation, functional dependencies and strong dependencies*, Comput. and AI. (submitted for publication).
12. Gare M. R., Jonhson D. S., *Computers and intractability: a guide to theory of NP-completeness*, Bell Laboratories, W. H. Freeman and Company, San Francisco, 1979.
13. Gottlob G., Libkin L., *Investigations on armstrong relations dependency inference, and excluded functional dependencies*, Acta Cybernetica Hungary, IX/4 (1990) 385-402.
14. Lucchesi C. L., Osborn S. L., *Canditante keys for relations*, J. Comput. Syst. Scien., 17, 2 (1978) 270-279.

15. Thi V. D., *Investigation on combinatorial characterizations related to functional dependency inj the relational datamodel*, MTA - AZTAKI Tanulmanyok, Budapest, Ph. D. Dissertation, 191 (1996) 1-157.

16. Thi V. D., *Minimal keys and antikeys*, Acta Cybernetica, 7, 4 (1986) 361-371.

17. Demetrovics J., Thi V. D., *Some result about normal for functional dependency in the relation datamodel*, Discrete Applied Mathematics, 69 (1996) 61-74.

18. Thi V. D., *Some results about hypergraph*, J. Informatics and Cybernetics, Hanoi, Vietnam, 13, 2 (1997) 8-15.

Institute of Information Technology
Nghia Do, Tu Liem, Hanoi.

Received: March 11, 1997