Tạp chí Tin học và Điều khiển học, T. 18, S. 1 (2002), 9–14

THE RELATIONSHIP BETWEEN DIRECT DETERMINATION AND FD-GRAPH

HO THUAN, NGUYEN VAN DINH

Abstract. The notion of direct determination was introduced by D. Maier [5] to study the structure of minimum covers. Using direct determination he showed that it is possible to find covers with the smallest number of FDs (Functional Dependencies) in polynomial time. In [2], G. Ausiello et al. presented an approach which is based on the representation of the set of FDs by FD-graph (considered as a special case of the hypergraph formalism introduced in [7]). Such a representation provides a unified framework for the treatment of various properties and for the manipulation of FDs.

In this paper, we establish the relation between FD-graph and direct determination, and prove some well-known and new properties concerning direct determination.

Tóm tắt. Khái niệm *xác định trực tiếp* đã được trình bày bởi D. Maier [5] để nghiên cứu cấu trúc các phủ cực tiểu. Sử dụng khái niệm này, ông đã chỉ ra rằng có thể tìm được các phủ với số phụ thuộc hàm là ít nhất trong thời gian đa thức. Trong [2], G. Ausiello và các tác giả khác đã đưa ra một cách tiếp cận mới trên cơ sở biểu diễn tập các phụ thuộc hàm bằng một FD-đồ thị (xem như một trường hợp đặc biệt của siêu đồ thị, được giới thiệu trong [7]). Cách biểu diễn như vậy cho một khung thống nhất để xử lý nhiều tính chất khác nhau và thao tác trên các FD.

Trong bài báo này, chúng tôi xác định mối liên hệ giữa FD-đồ thị và khái niệm *xác định trực tiếp*, chứng minh một số tính chất quen biết và những tính chất mới liên quan đến khái niệm này.

1. BASIC NOTIONS AND RESULTS

In this section we recall some notions and results which will be needed in the sequel. The reader is required to know the basic notions of the relational model and functional dependency [8]. As usual, we will only consider sets of FD in natural reduced form [4] and we assume that all attributes are chosen from some fixed universe Ω. That means for any $F = \{X_i \to Y_i \mid i = 1, 2, \dots, m\}$

$$X_i \cap Y_i = \emptyset, \ \forall i = 1, 2, \dots, m;$$
$$X_i \neq X_j \text{ for } i \neq j;$$
$$X_i, Y_i \subseteq \Omega, \ \forall i = 1, 2, \dots, m.$$

Let F^+ be the closure of F, i.e. the set of all FDs that can be inferred from the FDs in F by repeated application of the Armstrong's axioms [1].

Definition 1.1.

(a) Two sets F_1, F_2 of FDs over Ω are said equivalent, written $F_1 \equiv F_2$ if ${F_1}^+ = {F_2}^+$. If $F_1 \equiv F_2$ then F_1 is a cover for F_2 and vice versa.

(b) A set F of FDs is nonredundant if there is no proper subset F' of F with $F' \equiv F$. F_1 is a nonredundant cover for F_2 if F_1 is a cover for F_2 and F_1 is nonredundant.

(c) Let F be a set of FDs over Ω and let $X \to Y$ be a FD in F. Attribute $A \in \Omega$ is said *extraneous* in $X \to Y$ if
$$((F \setminus \{X \to Y\}) \cup \{X \setminus A \to Y \setminus A\})^+ = F^+.$$

(d) Two set of attributes X and Y are equivalent under a set of FDs, written $X \leftrightarrow Y$, if $X \to Y$ and $Y \to X$ are in F^+.

Definition 1.2. [5] Given a set of FDs F with $X \to Y$ in F^+. X direct determines Y under F, written $X \xrightarrow{\cdot} Y$ if $(X \to Y) \in [F \setminus E_F(X)]^+$, where $E_F(X)$ is the set of all FDs in F with left sides equivalent to X. That is, no FDs with left sides equivalent to X are used to derive $X \to Y$.

Definition 1.3. [5] A set of FDs F is *minimum* if there is no set G with fewer FD than F such that $G \equiv F$.

Theorem 1.1. [5] *Given equivalent minimum set of FDs F and G*

$$|E_F(X)| = |E_G(X)| \text{ for any } X.$$

Thus the size of equivalence classes in $\overline{E}_F$ is the same for all minimum F with the same closure (where $\overline{E}_F$ is the collection of all non empty $E_F(X)$).

Definition 1.4. [2] Given a set of FDs on Ω, the FD-graph $G_F = \langle V, E \rangle$ associated with F is the graph with node labeling function $w : V \to P(\Omega)$ and are labeling function $w' : E \to \{0, 1\}$ such that:

(i) for every attribute $A \in \Omega$, there is a node in V labeled A (called simple node);

(ii) for every dependency $X \to Y$ in F where $|X| > 1$, there is a node in V labeled X (called a compound node);

(iii) for every dependency $X \to Y$ in F where $Y = A_1 \ldots A_k$, there are arcs labeled 0 (full arcs) from the node labeled X to the nodes labeled $A_1, \ldots, A_k$;

(iv) for every compound node i in V labeled $A_{i_1} \ldots A_{i_p}$ there are arcs labeled 1 (dotted arcs) from the node i to all simple nodes (component nodes of i) labeled $A_{i_1}, \ldots, A_{i_p}$.

The set of full arcs (dotted arcs, respectively) is denoted E_0 (E_1, respectively).

Example 1.1. Given a set of attributes $\Omega = \{A, B, C, D, E, F, H\}$, let F be a set of FDs over Ω, $F = \{A \to BCF,\ C \to D,\ FBD \to H,\ BD \to E\}$ the corresponding FD-graph $G_F = \langle V, E \rangle$ is shown in Fig. 1.1.

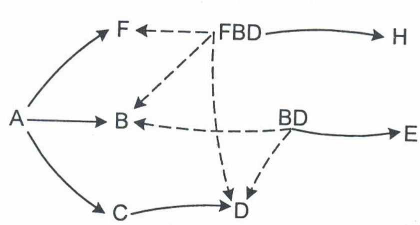

Fig. 1.1. An FD-graph

Definition 1.5. [2] Given an FD-graph $G_F = \langle V, E \rangle$ and two nodes $i, j \in V$, a (directed) FD-path $\langle i, j \rangle$ from i to j is a *minimal subgraph* $\overline{G}_F = \langle \overline{V}, \overline{E} \rangle$ of G_F such that $i, j \in \overline{V}$ and either $(i, j) \in \overline{E}$ or one of the following possibilities holds:

(a) j is a simple node and there exists a node k such that $(k, j) \in \overline{E}$ and there is an FD-path $\langle i, k \rangle$ included in $\overline{G}_F$ (graph transitivity).

(b) j is a compound node with component nodes $m_1, \ldots, m_r$ and there dotted arcs $(j, m_1), \ldots, (j, m_r)$ in $\overline{G}_F$ and r FD-paths $\langle i, m_1 \rangle, \ldots, \langle i, m_r \rangle$ included in $\overline{G}_F$ (graph union).

Further more, an FD-path $\langle i, j \rangle$ is *dotted* if all its arcs leaving i are *dotted*; otherwise it is *full*.

Example 1.2. For the FD-graph of the Example 1.1: (a) full FD-path $\langle A, E \rangle$, (b) full FD-path $\langle A, D \rangle$, and dotted FD-path $\langle FBD, E \rangle$ are given in Fig. 1.2.

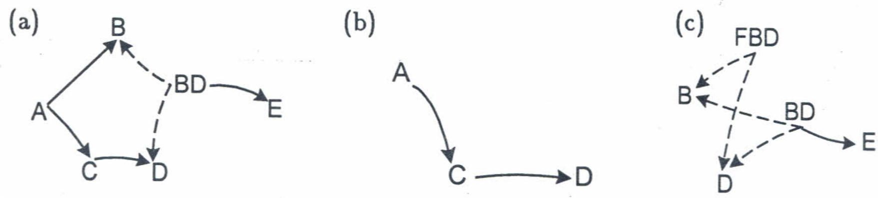

Fig. 1.2. FD-paths

Definition 1.6. [2]

(a) The closure of an FD-graph $G_F = \langle V, E\rangle$ is the graph $G_F{}^+ = \langle V, E^+\rangle$, labeled on the nodes and on the arcs, where the set V is the same as in G_F, while the set $E^+ = (E^+)_0 \cup (E^+)_1$ is defined in the following way

$(E^+)_1 = \{(i, j) \mid i, j \in V$ and there exists a dotted FD-path $\langle i, j\rangle\}$;

$(E^+)_0 = \{(i, j) \mid i, j \in V, (i, j \notin (E^+)_1$ and there exists a full FD-path $\langle i, j\rangle\}$.

(b) Two nodes i, j in an FD-graph are said equivalent if the arcs (i, j) and (j, i) both belong to the closure of G_F. Further more, a node i of G_F is said to be equivalent to node j of $G_{\overline{F}}$ where $G_{\overline{F}}$ is a cover of G_F (i.e. $F^+ = \overline{F}^+$) if i, j are equivalent in some cover of G_F.

(c) Given two FD-graphs G_{F_1}, G_{F_2}; G_{F_2} is a cover of G_{F_1} if F_2 is a cover of F_1.

(d) An FD-graph G_F is nonredundant if F is nonredundant.

Theorem 1.2. [2] *Let $G_F = \langle V, E\rangle$ be the FD-graph associated with the set F of FDs, and let $G_F{}^+ = \langle V, E^+\rangle$ be its closure. An arc (i, j) is in E^+ if and only if $w(i) \to w(j)$ is in F^+.*

Theorem 1.3. [2] *A nonredundant FD-graph $G_F = \langle V, E\rangle$ is minimum if and only if it has no superfluous node.*

Recall that a node $i \in V$ is superfluous if there exists a dotted FD-path $\langle i, j\rangle$ where j is a node of V equivalent to i.

2. DIRECT DETERMINATION AND FD-GRAPH

In this section, we establish the relation between FD-graph and direct determination by proving some well-known and new properties of direct determination.

First it is worth giving a few comments on the definition of an FD-graph.

Remark 2.1. Definition 1.4 is reasonable and concise in the sense that the FD-graph G_F includes all the "meaning part" of the closure of the set of FDs. On the other hand, with the formalism of FD-graph, we can provide a simple and unified treatment of all properties of sets of FDs.

Following the definition of a FD-graph, it is clear that every compound node has at least one outgoing full arc. However, according to the necessity, we can freely add to an FD-graph some new coumpound nodes without outgoing full arcs if it makes easy to prove a certain required property.

So, a natural way is to think that an FD-graph $G_F = \langle V, E\rangle$ associated with F is defined by Definition 1.4 precisely to an arbitrary finite number of different compound nodes which do not correspond to the left side of any FD in F, together with the dotted arcs from each of them to their corresponding component nodes.

Definition 2.1. [2] Given an FD-graph $G_F = \langle V, E\rangle$ and a node $i \in V$ with at least a full outgoing arc. A strong component of G_F with representative node i is a maximal set of pairwise equivalent nodes which contains i, denoted by SC(i). Notice that every node in SC(i) has at least one full outgoing arc.

The following lemma is obvious.

Lemma 2.1. *Given an FD-graph $G_F = \langle V, E\rangle$, a node $i \in V$, its corresponding strong component $SC(i)$ and two nodes j, k such that j is equivalent to i. (j not necessarily belong to $SC(i)$, i.e. j can be a compound node without outgoing full arc that we add it to the FD-graph. The same situation can happen with the node k too).*

Then $w(j) \stackrel{\cdot}{\rightarrow} w(k)$ if and only if there exists a dotted FD-path $\langle j, k\rangle$ containing no full outgoing arc from any node of $SC(i)$.

In other words, the dotted FD-path $\langle j, k\rangle$ contains no intermediate node that is node of $SC(i)$.

In that case, for sake of simplicity, we write $\langle j \xrightarrow{SC(i)} k\rangle$.

Example 2.1. Given $\Omega = ABCDEIH$, $F = \{A \to BCH, BC \to A, AD \to EI, EA \to ID\}$. It is easy to verify that:

$$E_F(AD) = \{AD \to EI, AE \to DI\} \text{ and } BCD \leftrightarrow AD.$$

The corresponding FD-graph G_F with an added node BCD (without outgoing full arc) is shown in Fig. 2.1.

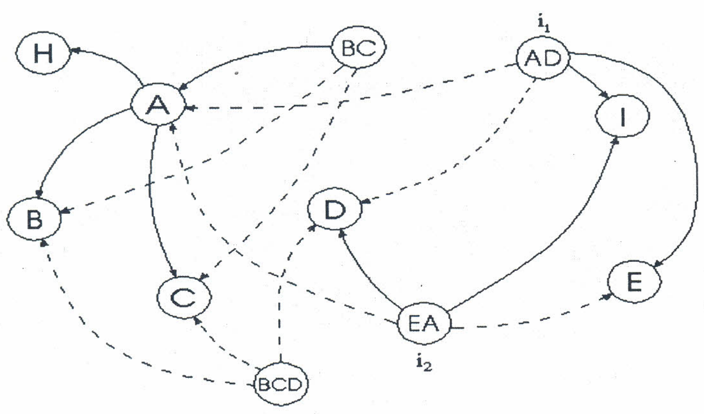

Fig. 2.1. FD-graph with added node BCD

We have

$$SC(i_1) = \{i_1, i_2\} \text{ where } w(i_1) = AD,\ w(i_2) = EA,$$

we find that

$$BCD \stackrel{\cdot}{\rightarrow} H$$

and

$$BCD \stackrel{\cdot}{\rightarrow} AD.$$

Lemma 2.2. *Given an FD-graph $G_F = \langle V, E\rangle$, two equivalent nodes $i, j \in V$ and i_q, j_q are two nodes equivalent to i and j respectively.*

If $\langle i_q \xrightarrow{SC(i)} j_q\rangle$ and $\langle j_q \xrightarrow{SC(j)} k\rangle$ then $\langle i_q \xrightarrow{SC(i)} k\rangle$.

Proof. By merging two FD-paths $\langle i_q \xrightarrow{SC(i)} j_q\rangle$ and $\langle j_q \xrightarrow{SC(j)} k\rangle$ appropriately at compound nodes of j_q which are intermediate nodes of the FD-path $\langle j_q \xrightarrow{SC(j)} k\rangle$ we obtain the FD-path $\langle i_q \xrightarrow{SC(i)} k\rangle$.

In other words, from $w(i) \leftrightarrow w(i_q)$, $w(j) \leftrightarrow w(j_q)$ and $w(i_q) \stackrel{\cdot}{\rightarrow} w(j_q)$, $w(j_q) \stackrel{\cdot}{\rightarrow} w(k)$, we have $w(i_q) \stackrel{\cdot}{\rightarrow} w(k)$.

Notice that the above lemma corresponds to [5, Lemma 5].

Example 2.2. Take up again Example 2.1 (Fig. 2.1), we have $BCD \stackrel{\cdot}{\rightarrow} AD$ and $AD \stackrel{\cdot}{\rightarrow} H$.

Since A is the unique component node of AD that is an intermediate node on the FD-path

$\langle AD \overset{SC(i_1)}{\not\longrightarrow} H\rangle$, we will merge two FD-paths $\langle BCD, AD\rangle$ and $\langle AD, H\rangle$ at A to obtain the FD-path $\langle BCD, H\rangle$ such that $BCD \overset{\bullet}{\rightarrow} H$.

Lemma 2.3. *Given an FD-graph $G_F = \langle V, E\rangle$, $i \in V$ is a node having at least one outgoing full arc and i_0 is equivalent to i (i_0 can be an added node to the FD-graph without outgoing full arc). Then there exists $j \in SC(i)$ such that $\langle i_0 \overset{SC(i)}{\not\longrightarrow} j\rangle$.*

Proof. Suppose that $i_0 \notin \mathrm{SC}(i)$. Otherwise, take $j \equiv i_0$ and the lemma is proved. Consider the dotted FD-path $\langle i_0, i\rangle$. In the case there is no intermediate node in $\langle i_0, i\rangle$ that is node of $\mathrm{SC}(i)$ then i is the node to be found.

Otherwise, suppose that $i_1 \in \mathrm{SC}(i)$ is an intermediate node of $\langle i_0, i\rangle$. Now we have only to consider the FD-path $\langle i_0, i_1\rangle$. Repeat the above reasoning for $\langle i_0, i_1\rangle$. Finally, we will find the required j such that $\langle i_0 \overset{SC(i)}{\not\longrightarrow} j\rangle$. □

Notice that the above lemma corresponds to [5, Lemma 6].

Lemma 2.4. *Let $G_F = \langle V, E\rangle$, be a minimum FD-graph (i.e. F is minimum), and $i \in V$ is a node with at least one outgoing full arc. Then in $SC(i)$ there exist no j_1, j_2; $j_1 \neq j_2$ such that $\langle j_1 \overset{SC(i)}{\not\longrightarrow} j_2\rangle$.*

Proof. Assume the contrary that there exist $j_1, j_2 \in \mathrm{SC}(i)$, $j_1 \neq j_2$ such that there is a dotted FD-path from j_1 to j_2. Since j_1 is equivalent to j_2, j_1 is a superfluous node. We arrive to a contradiction. (See Theorem 1.3). □

Notice that the above lemma corresponds to [5, Lemma 7].

Lemma 2.5. *Given two nonredundant FD-graph $G_{F_1} = \langle V_1, E_1\rangle$, $G_{F_2} = \langle V_2, E_2\rangle$, wherein G_{F_2} is a cover of G_{F_1}. Let i_1 and i_2 be two equivalent nodes in V_1 and V_2, respectively, with at least one outgoing full arc, (p_2, q_2) be a full arc of E_2 with $p_2 \neq SC^{(2)}(i_2)$.(*) If $(i_1, p_2) \in {E_2}^+$, then $\langle p_2 \overset{SC^{(1)}(i_1)}{\not\longrightarrow} q_2\rangle$.*

Proof. Since $(i_1, p_2) \in {E_2}^+$, by Theorem 1.2, there exists a FD-path in G_{F_1} from i_1 to p_2. Now assume the contrary that the FD-path in G_{F_1} from p_2 to q_2 has an intermediate node $j_1 \in \mathrm{SC}^{(1)}(i_1)$. The presence of the FD-path $\langle j_1, i_1\rangle$ shows that p_2 is equivalent to i_1, i.e. $p_2 \in \mathrm{SC}^{(2)}(i_2)$, a contradition. □

Theorem 2.6. *With the same assumptions as in Lemma 2.5, if we replace in G_{F_1} all nodes belonging to $SC^{(1)}(i_1)$ together with their corresponding outgoing arcs by all nodes in $SC^{(2)}(i_2)$ together with their corresponding outgoing arcs, then the new FD-graph is a cover of G_{F_1}.*

Proof. We have only to prove that for every full arc $(j_1, k_1) \in E_1$ with $j_1 \in \mathrm{SC}^{(1)}(i_1)$ there exists a FD-path $\langle j_1, k_1\rangle$ in the new FD-graph. By Lemma 2.5 we have just the required result. □

Remark 2.2. Theorem 2.6 can be formulated in another form as follows:

If F_1, F_2 are nonredundant and equivalent sets of FDs, then

$$F_1 \equiv \{F_1 \setminus E_{F_1}(X)\} \cup E_{F_2}(X) \equiv \{F_2 \setminus E_{F_2}(X)\} \cup E_{F_1}(X).$$

Let us close the paper with the following useful proposition:

Proposition 2.7. *Let $U \rightarrow W$ be an FD in F^+ and let $X \rightarrow Y$ be an FD in F that participates in the Armstrong's derivation sequence for $U \rightarrow W$. Then we have:*

$$U \rightarrow X,\ UY \rightarrow W \in (F \setminus \{X \rightarrow Y\})^+.$$

$\mathrm{SC}^{(1)}$ and $\mathrm{SC}^{(2)}$ refer to G_{F_1} and G_{F_2}, respectively

Proof. Let $G_F = \langle V, E\rangle$ be the FD-graph associated with F. From $U \to W$ in F^+ it follows that there is an FD-path $\langle i, j\rangle$ from i to j, where $w(i) = U$, $w(j) = W$. Since $X \to Y \in F$ takes part in the derivation sequence for $U \to W$, the nodes p and q with $w(p) = X$ and $w(q) = Y$ are intermediate nodes on $\langle i, j\rangle$. It is clear that the FD-paths $\langle i, p\rangle$ and $\langle q, j\rangle$ contain no outgoing full arc from node p. □

Example 2.3. Reconsider the Example 2.1 (Fig. 2.1). We have $BCD \to H \in F^+$, $(BC \to A) \in F$ participates in the derivation sequence for $BCD \to H$.

It is clear that:
$BCD \to BC \in (F \setminus \{BC \to A\})^+$ and corresponds to the FD-path $\langle BCD, BC\rangle$;
$BCDA \to H \in (F \setminus \{BC \to A\})^+$ and corresponds to the FD-path $\langle BCDA, H\rangle$.

CONCLUSIONS

An FD-graph approach for the representation of functional dependencies (FDs) in relational databases. It also supports the studies of FDs. This approach allow a homogeneous treatment of several problems (closure, minimization, etc.), which leads to simpler proofs and, in some cases, more efficient algorithms than in the current literature. Therefore, the studies of FD-graph is a middle step to further study Database Hypergraphs in which directed hyperedges represent FDs and undirected hyperedges represent the join dependency.

REFERENCES

[1] Armstrong W. W., Dependency structures of database relationships, *Information Processing 74*, North Holland Publishing Company, 1974, 580–583.

[2] Ausiello G. et al., Graphs algorithms for functional dependency manipulation, *J. ACM* **30** (1983) 752–766.

[3] Fagin R., Ling Ling Yan, Renee J. Miller, and Laura M. Haas, Data-driven understanding and refinement of schema mappings, *Proc. 2001 ACM SIGMOD Symposium*, Santa Barbara, 485–496.

[4] Ho Thuan, *Contribution to the Theory of Relational Database*, Tanulmányok, 184/1986, Budapest, Hungary.

[5] Maier D., Minimum covers in the relational database model, *J. ACM* **27** (1980) 664–674.

[6] S. Nguyen, D. Pretolani, and L. Markenzon, Some path problems on oriented hypergraphs, *Theoretical Informatics and Applications* (Elsevier-Paris) **32** (1998), No. 1, 2, 3.

[7] Saccá D., Closures of database hypergraphs, *J. ACM* **32** (1985) 774–803.

[8] Ullman Jeffrey D., *Principles of Database and Knowledge-Base Systems*, Computer Science Press, USA, 1989.

Received October 25, 2001

Ho Thuan, National Institute of Information Technology, Hanoi.
Nguyen Van Dinh, United Nations International School of Hanoi.