

# VỀ ĐỘ PHỨC TẠP TÍNH TOÁN CỦA MỘT BÀI TOÁN LIÊN QUAN ĐẾN TẬP RÚT GỌN TRÊN BẢNG QUYẾT ĐỊNH

Nguyễn Ngọc Cương<sup>1</sup>, Vũ Đức Thi<sup>2</sup>

<sup>1</sup>Khoa Toán-Tin - Học viện an ninh nhân dân  
<sup>2</sup>Viện Công nghệ thông tin - Đại học quốc gia Hà Nội

**TÓM TẮT** - Trên thực tiễn, các vấn đề liên quan đến tập rút gọn trên bảng quyết định đã được nhiều tác giả đề cập và nghiên cứu. Trong bài báo này, chúng tôi trình bày một bài toán co-NP - đầy đủ liên quan đến các tập rút gọn trên bảng quyết định. Chúng ta gọi  $A$  là tập tựa rút gọn trên bảng quyết định nhất quán  $DS$  với tập thuộc tính  $CU\{d\}$  và  $d$  là thuộc tính quyết định nếu  $A$  chứa một tập rút gọn nào đó. Tương tự chúng ta gọi  $A$  là tập tựa tối thiểu của thuộc tính  $d$  trên sơ đồ quan hệ  $s = \langle C U \{d\}, F \rangle$  nếu  $A$  chứa một tập tối thiểu của thuộc tính  $d$ . Gọi  $Q$  là tập tất cả các tập tựa rút gọn của bảng quyết định  $DS$ ,  $P$  là tập tất cả các tập tựa tối thiểu của thuộc tính  $d$  trên  $s$ . Khi đó bài toán xác định  $Q$  có là tập con của  $P$  hay không là co-NP - đầy đủ.

Việc tìm kiếm các tập rút gọn đóng vai trò quan trọng trong việc xử lý thông tin trên bảng quyết định. Mục tiêu của rút gọn thuộc tính là loại bỏ các thuộc tính dư thừa để tìm ra các thuộc tính cơ bản phục vụ cho việc xử lý thông tin. Về thực chất, việc rút gọn thuộc tính là tìm tập con nhỏ nhất của tập các thuộc tính để bảo toàn thông tin phân lớp trên bảng quyết định. Trong bài báo này chúng tôi chỉ đề cập tới các bảng quyết định nhất quán. Trên thực tiễn, tùy theo từng bài toán cụ thể, chúng ta có thể chuyên bảng quyết định không nhất quán về bảng quyết định nhất quán.

Bài báo này trình bày về một bài toán co-NP - đầy đủ liên quan đến các tập rút gọn trên bảng quyết định. Chúng ta gọi  $A$  là tập tựa rút gọn trên bảng quyết định  $DS$  với tập thuộc tính  $C U \{d\}$  và  $d$  là thuộc tính quyết định nếu  $A$  chứa một tập rút gọn nào đó. Tương tự chúng ta gọi  $A$  là tập tựa tối thiểu của thuộc tính  $d$  trên sơ đồ quan hệ

$s = \langle C U \{d\}, F \rangle$  nếu  $A$  chứa một tập tối thiểu không tầm thường của thuộc tính  $d$ . Gọi  $Q$  là tập tất cả các tập tựa rút gọn của bảng quyết định  $DS$ ,  $P$  là tập tất cả các tập tựa tối thiểu của thuộc tính  $d$  trên  $s$ . Khi đó bài toán xác định  $Q$  có là tập con của  $P$  hay không là co-NP - đầy đủ.

## I. NHỮNG KHÁI NIỆM CƠ BẢN

Mục này cung cấp một số khái niệm cơ bản được dùng trong bài báo. Các khái niệm này có thể xem trong [2,3,5].

**Định nghĩa 1.1.** Hệ thông tin là một bộ bốn  $S = (U, A, V, f)$  trong đó  $U$  là tập hữu hạn, khác rỗng các đối tượng;  $A$  là tập hữu hạn, khác rỗng các thuộc tính;  $V = \bigcup_{a \in A} V_a$  với  $V_a$  là tập giá trị của thuộc tính  $a \in A$ ;  $f: U \times A \rightarrow V_a$  là hàm thông tin,  $\forall a \in A, u \in U f(u, a) \in V_a$ .

Với mọi  $u \in U, a \in A$ , ta ký hiệu giá trị thuộc tính  $a$  tại đối tượng  $u$  là  $a(u)$  thay vì  $f(u, a)$ . Nếu  $B = \{b_1, b_2, \dots, b_k\} \subseteq A$  là một tập con các thuộc tính thì ta ký hiệu bộ các giá trị  $b_i(u)$  bởi  $B(u)$ . Như vậy, nếu  $u$  và  $v$  là hai đối tượng, thì ta viết  $B(u) = B(v)$  nếu  $b_i(u) = b_i(v)$  với mọi  $i = 1, \dots, k$ .

**Định nghĩa 1.2.** Bảng quyết định là một hệ thông tin  $S = (U, A, V, f)$  với  $A = C \cup D$  và  $C \cap D = \emptyset$ . Bảng quyết định  $S$  được gọi là nhất quán nếu  $D$  phụ thuộc hàm vào  $C$ , tức là với mọi  $u, v \in U, C(u) = C(v)$  kéo theo  $D(u) = D(v)$ . Ngược lại thì gọi là không nhất quán hay mâu thuẫn.  $C$  được gọi là tập thuộc tính điều kiện và  $D$  là tập thuộc tính quyết định

Thông thường  $D = \{d\}$  chứa một thuộc tính

**Định nghĩa 1.3.** Cho bảng quyết định nhất quán  $DS = (U, C \cup D, V, f)$  và tập thuộc tính  $R \subseteq C$ .  $R$  được gọi là tập rút gọn nếu:

- với mọi cặp đối tượng  $u, v$  thì  $R(u) = R(v)$  kéo theo  $D(u) = D(v)$
- với mọi  $E$  là tập con thực sự của  $R$  thì tồn tại cặp  $u, v$  để  $E(u) = E(v)$  không kéo theo  $D(u) = D(v)$

Tập rút gọn định nghĩa như trên còn gọi là tập rút gọn Pawlak. Ký hiệu  $PRED(C)$  là họ tất cả các tập rút gọn của  $C$ .

Cho  $R = \{a_1, \dots, a_n\}$  là tập hữu hạn, khác rỗng các thuộc tính, mỗi thuộc tính  $a_i$  có miền giá trị là  $D(a_i)$ . Quan hệ  $r$  trên  $R$  là tập các bộ  $\{h_1, \dots, h_m\}$  với  $h_j: R \rightarrow \bigcup_{a_i \in R} D(a_i), 1 \leq j \leq m$  là một hàm sao cho  $h_j(a_i) \in D(a_i)$ .

Cho  $r = \{h_1, \dots, h_m\}$  là một quan hệ trên tập thuộc tính  $R = \{a_1, \dots, a_n\}$ . Phụ thuộc hàm (PTH) trên  $R$  là một dãy ký tự có dạng  $A \rightarrow B$  với  $A, B \subseteq R$ . PTH  $A \rightarrow B$  thỏa mãn quan hệ  $r$  trên  $R$  nếu  $(\forall h_i, h_j \in r) ((\forall a \in A)(h_i(a) = h_j(a)) \Rightarrow (\forall b \in B)(h_i(b) = h_j(b)))$ . Đặt  $F_r = \{(A, B): A, B \subseteq R, A \rightarrow B\}$  là họ đầy đủ các PTH thỏa mãn quan hệ  $r$ . Ký hiệu  $P(R)$  là tập các tập con của  $R$ . Cho  $F = P(R) \times P(R)$ . Ta nói rằng  $F$  là một họ  $f$  trên  $R$  nếu với mọi  $A, B, C, D \subseteq R$

- (1)  $(A, A) \in F$ .
- (2)  $(A, B) \in F, (B, C) \in F \Rightarrow (A, C) \in F$ .
- (3)  $(A, B) \in F, A \subseteq C, D \subseteq B \Rightarrow (C, D) \in F$ .
- (4)  $(A, B) \in F, (C, D) \in F \Rightarrow (A \cup C, B \cup D) \in F$ .

Rõ ràng là  $F_r$  là một họ  $f$  trên  $R$ . Nếu  $F$  là một họ  $f$  trên  $R$  thì có một quan hệ  $r$  trên  $R$  sao cho  $F_r = F$ . Ký hiệu  $F^+$  là tập tất cả các PTH được dẫn xuất từ  $F$  bằng việc áp dụng các quy tắc (1)–(4).

Sơ đồ quan hệ (SDQH)  $s$  là một cặp  $\langle R, F \rangle$  với  $R$  là tập thuộc tính và  $F$  là tập các phụ thuộc hàm trên  $R$ . Ký hiệu  $A^+ = \{a: A \rightarrow \{a\} \in F^+\}$ ,  $A^+$  được gọi là bao đóng của  $A$  trên  $s$ . Dễ thấy  $A \rightarrow B \in F^+$  khi và chỉ khi  $B \subseteq A^+$ . Tương tự ký hiệu  $A_r^+ = \{a: A \rightarrow \{a\} \in F_r^+\}$ ,  $A_r^+$  được gọi là bao đóng của  $A$  trên quan hệ  $r$ .

Cho  $s = \langle R, F \rangle$  là SDQH trên  $R$ ,  $a \in R$ . Đặt  $\mathcal{K}_a^s = \{A \subseteq R: A \rightarrow \{a\}, \nexists B: (B \rightarrow \{a\})(B \subset A)\}$ . Khi đó,  $\mathcal{K}_a^s$  được gọi là họ các tập tối thiểu của thuộc tính  $a$  trên  $s$ . Tương tự, cho  $r$  là một quan hệ trên  $R$  và  $a \in R$ . Đặt  $\mathcal{K}_a^r = \{A \subseteq R: A \rightarrow \{a\}, \nexists B: (B \rightarrow \{a\})(B \subset A)\}$ . Khi đó,  $\mathcal{K}_a^r$  được gọi là họ các tập tối thiểu của thuộc tính  $a$  trên  $r$ .

Gọi  $\mathcal{K} \subseteq P(R)$  là một hệ Sperner trên  $R$  nếu với mọi  $A, B \in \mathcal{K}$  kéo theo  $A \not\subset B$ . Dễ thấy  $\mathcal{K}_a^s, \mathcal{K}_a^r$  là các hệ Sperner trên  $R$ . Với tập  $\mathcal{K}$  là một hệ Sperner trên  $R$ , ta định nghĩa tập  $\mathcal{K}^{-1}$  như sau:

$$\mathcal{K}^{-1} = \{A \subset R: (B \in \mathcal{K}) \Rightarrow (B \not\subset A)\} \text{ và nếu } (A \subset C) \Rightarrow (\exists B \in \mathcal{K})(B \subseteq C)$$

Dễ thấy  $\mathcal{K}^{-1}$  cũng là một hệ Sperner trên  $R$ . Nếu  $\mathcal{K}$  là một hệ Sperner trên  $R$  đóng vai trò là tập các khóa tối thiểu của quan hệ  $r$  (hoặc SDQH  $s$ ) thì  $\mathcal{K}^{-1}$  là họ tất cả các tập không phải khóa lớn nhất của  $r$  (hoặc của  $s$ ), gọi là tập các phân khóa. Nếu  $\mathcal{K}$  là một hệ Sperner trên  $R$  đóng vai trò là họ các tập tối thiểu của thuộc tính  $a$  trên  $r$  (hoặc trên  $s$ ), hay  $\mathcal{K} = \mathcal{K}_a^r$  (hoặc  $\mathcal{K} = \mathcal{K}_a^s$ ), thì  $\mathcal{K}^{-1} = \{\mathcal{K}_a^r\}^{-1}$  (hoặc  $\mathcal{K}^{-1} = \{\mathcal{K}_a^s\}^{-1}$ ) là họ tất cả các tập lớn nhất không phải là tập tối thiểu của thuộc tính  $a$ , được định nghĩa như sau [1]:

$$\{\mathcal{K}_a^r\}^{-1} = \{A \subseteq R: A \rightarrow \{a\} \notin F_r^+, A \subset B \Rightarrow B \rightarrow \{a\} \in F_r^+\},$$

$$\{\mathcal{K}_a^s\}^{-1} = \{A \subseteq R: A \rightarrow \{a\} \notin F^+, A \subset B \Rightarrow B \rightarrow \{a\} \in F^+\}.$$

## II. KẾT QUẢ

Trong các bài toán thực tế, bảng quyết định thường chứa các đối tượng không nhất quán (là các đối tượng bằng nhau trên tập thuộc tính điều kiện nhưng khác nhau trên tập thuộc tính quyết định), gọi là bảng quyết định không nhất quán. Tuy nhiên, tùy thuộc vào lớp bài toán cần giải quyết mà ta có thể chuyển bảng quyết định không nhất quán về bảng quyết định nhất quán qua bước tiền xử lý số liệu bằng cách loại bỏ các đối tượng không nhất quán.

**Bổ đề 2.1.** [4] Cho bảng quyết định nhất quán  $DS = (U, C \cup \{d\}, V, f)$  với  $C = \{c_1, c_2, \dots, c_n\}$ ,  $U = \{u_1, u_2, \dots, u_m\}$ .

Xét quan hệ  $r = \{u_1, u_2, \dots, u_m\}$  trên tập thuộc tính  $R = C \cup \{d\}$ .

Đặt  $\mathcal{E}_r = \{E_{ij} : 1 \leq i < j \leq m\}$  với  $E_{ij} = \{a \in R : a(u_i) = a(u_j)\}$

Đặt  $\mathcal{M}_d = \{A \in \mathcal{E}_r : d \notin A, \nexists B \in \mathcal{E}_r : d \notin B, A \subset B\}$ .

Thì  $\mathcal{M}_d = (\mathcal{K}_d^r)^{-1}$ . Ở đây  $\mathcal{K}_d^r$  là họ các tập tối thiểu của thuộc tính  $\{d\}$  trên quan hệ  $r$

**Bổ đề 2.2.** Cho trước bảng quyết định  $DS = (U, C \cup \{d\}, V, f)$  thì  $(K_d^f)^{-1}$  là hệ Sperner trên  $C$ . Ngược lại nếu  $K$  là hệ Sperner trên  $C$  thì tồn tại một bảng quyết định nhất quán  $DS = (U, C \cup \{d\}, V, f)$  để  $K = (K_d^f)^{-1}$

### Chứng minh

Theo định nghĩa  $\{\mathcal{K}_a^r\}^{-1} = \{A \subseteq R : A \rightarrow \{a\} \notin F_r^+, A \subset B \Rightarrow B \rightarrow \{a\} \in F_r^+\}$ . Dương

nhiên  $(K_d^f)^{-1}$  là hệ Sperner trên  $C$ . Ngược lại, nếu  $K$  là hệ Sperner trên  $C$ .

Giả sử  $K = \{A_1, \dots, A_m\}$ . Chúng ta xây dựng bảng quyết định  $DS = (U, C \cup \{d\}, V, f)$  như sau:

$U = \{u_0, u_1, \dots, u_m\}$  với mọi  $c \in C : c(u_0) = 0$  và  $d(u_0) = 0$ . Với mọi  $i, i = 1, \dots, m$  và  $c$  là phần tử của  $C$ . Chúng ta đặt  $c(u_i) = 0$  nếu  $c \in A_i$ . Ngược lại  $c(u_i) = 1$ . Đặt  $d(u_i) = i$ . Với  $R = C \cup \{d\}$ .

Đặt  $\mathcal{E}_r = \{E_{ij} : 1 \leq i < j \leq m\}$  với  $E_{ij} = \{a \in R : a(u_i) = a(u_j)\}$

Đặt  $\mathcal{M}_d = \{A \in \mathcal{E}_r : d \notin A, \nexists B \in \mathcal{E}_r : d \notin B, A \subset B\}$ .

Có thể thấy  $\mathcal{M}_d = \{A_1, \dots, A_m\}$ . Theo Bổ đề 2,1, ta có  $\mathcal{M}_d = (K_d^f)^{-1}$ . Như vậy

$K = (K_d^f)^{-1}$ .

### Bổ đề 2.3.

Giả sử  $K = \{K_1, K_2, \dots, K_t\}$  là một hệ Sperner trên  $C$ . Chúng ta xây sơ đồ quan hệ  $s_d = \langle R, F \rangle$ , ở đây  $R = C \cup \{d\}$ ,  $F = \{K_i \rightarrow \{d\} : i=1, \dots, t\}$ . Khi đó

$K = K_d^s - \{d\}$

### Chứng minh

1) Với  $K \in K$  ta có  $K \rightarrow \{d\}, K \neq \{d\}$  và vì  $K$  là hệ Sperner nên không tồn tại  $K' \subset K$  sao cho  $K' \rightarrow \{d\}$ . Do đó, theo định nghĩa  $K$  là một tập tối thiểu của thuộc tính  $\{d\}$  trên  $s_d$ , nghĩa là  $K \in \mathcal{K}_d^s - \{d\}$ .

2) Ngược lại, với  $K \in \mathcal{K}_d^s - \{d\}$  ta có  $K \rightarrow \{d\}, K \neq \{d\}$  và không tồn tại  $K' \subset K$  sao cho  $K' \rightarrow \{d\}$ .

Dễ thấy với mọi  $K_i$  là phần tử của hệ Sperner  $K$  ta có  $K \not\subset K_i$ .

Vì nếu  $K \subset K_i$  thì theo định nghĩa sơ đồ quan hệ  $s_d$ ,  $K$  không xác định hàm với  $\{d\}$

Hơn nữa, với mọi  $K_i$  là phần tử của hệ Sperner  $K$  chúng ta có thể thấy  $K_i$  không là tập con thực sự của  $K$ .

Vì nếu  $K_i \subset K$  thì  $K$  không phải là một tập tối thiểu của thuộc tính  $\{d\}$  trên  $s_d$ . Từ đó suy ra  $T = \{K, K_1, \dots, K_r\}$  là hệ Sperner trên  $C$ . Theo định nghĩa bao đóng của tập  $K$  trên sơ đồ quan hệ  $s_d$  ta có  $K^+ = K$ . Có nghĩa là  $K$  không xác định hàm với  $\{d\}$ . Điều này trái với  $K$  là một tập tối thiểu sơ đồ quan hệ  $s_d$ . Vậy  $K$  là phần tử của hệ Sperner  $K$

Từ 1) và 2) chúng ta có  $K = K_d^s - \{d\}$

Chúng ta gọi một bài toán  $A$  là co-NP-đầy đủ nếu bài toán phủ định  $A$  là NP - đầy đủ.

Chúng ta biết rằng [1] bài toán phân bù giới hạn các tập con (Subset delimiter complementarity – SDC) sau đây là co – NP – đầy đủ: Cho một tập hữu hạn  $T$ , hai họ  $P = \{P_1, \dots, P_n\}$ ,  $Q = \{Q_1, \dots, Q_m\}$  là các tập con của  $T$ . Kiểm tra xem với mọi  $A \subseteq T$  có tồn tại  $P_i$  để  $P_i \subseteq A$  hoặc có  $Q_j$  để  $A \subseteq Q_j$  với mọi  $i$  và với mọi  $j$ . Trong [1] cũng chỉ ra rằng nếu  $Q$  là một hệ Sperner thì vấn đề trên vẫn là co-NP- đầy đủ

Chúng ta sẽ trình bày về một bài toán co-NP- đầy đủ liên quan đến các tập rút gọn trên bảng quyết định.

**Định nghĩa 2.4.** Chúng ta gọi  $A$  là tập tựa rút gọn trên bảng quyết định DS với tập thuộc tính  $C \cup \{d\}$  và  $d$  là thuộc tính quyết định nếu  $A$  chứa một tập rút gọn nào đó. Tương tự chúng ta gọi  $A$  là tập tựa tối thiểu của thuộc tính  $d$  trên sơ đồ quan hệ  $s = \langle C \cup \{d\}, F \rangle$  nếu  $A$  chứa một tập tối thiểu không tầm thường  $B$  của thuộc tính  $d$ . Ở đây  $B \in K_d^s - \{d\}$ .

Gọi  $Q_d$  là tập tất cả các tập tựa rút gọn của bảng quyết định DS với  $d$  là thuộc tính quyết định,  $P_d$  là tập tất cả các tập tựa tối thiểu của thuộc tính  $d$  trên  $s$ . Khi đó bài toán xác định  $Q_d$  có là tập con của  $P_d$  hay không là co-NP - đầy đủ.

**Định lý 2.5.** Vấn đề sau đây là co - NP - đầy đủ

Cho trước bảng quyết định DS với tập thuộc tính  $C \cup \{d\}$  và  $d$  là thuộc tính quyết định và sơ đồ quan hệ  $s = \langle C \cup \{d\}, F \rangle$  kiểm tra xem  $Q_d$  có là tập con của  $P_d$  hay không. Có nghĩa rằng mỗi tập tựa rút gọn của DS có là tập tựa tối thiểu của thuộc tính  $d$  trên  $s$  hay không.

### Chứng minh

Đối với mỗi  $A \subseteq R = C \cup \{d\}$  Chúng ta kiểm tra rằng  $A$  là hoặc không là một tựa rút gọn của DS bằng một thuật toán đa thức. Bởi một thuật toán tìm bao đóng  $A^+$  và định nghĩa tựa tối thiểu của thuộc tính  $d$  trên sơ đồ quan hệ, chúng ta cũng có thể kiểm tra trong thời gian đa thức  $A$  là hoặc không là tập tựa tối thiểu của  $d$  trên  $s$ . Do đó, chúng ta chọn tùy ý một tập con  $A$  của  $R$  sao cho  $A$  là tựa rút gọn của DS nhưng không là tựa tối thiểu của  $s$ . Hiển nhiên thuật toán này là một thuật toán bất định đa thức. Như vậy, vấn đề của chúng ta thuộc co – NP. Bây giờ chúng ta khảo sát vấn đề được nêu ra ở [1] với tập  $T$  và hay họ  $P_1, P_2, \dots, P_n$  và  $Q_1, Q_2, \dots, Q_m$  ở đây tập  $Q$  là hệ Sperner trên  $T$ .

Ký pháp  $P' = \{P_i \in P: \text{Không có } P_j \text{ để } P_j \text{ là tập con thực sự của } P_i, i = 1, \dots, n \text{ và } j = 1, 2, \dots, n\}$

Rõ ràng rằng  $P'$  là tập tất cả các phần tử nhỏ nhất của  $P$  và  $P'$  là hệ Sperner trên  $T$ . Từ  $P$  chúng ta có thể tính  $P'$  trong thời gian đa thức với kích thước của  $P$  và  $T$ . Dễ thấy rằng  $\{T, P', Q\}$  là một thể hiện tương đương của  $\{T, P, Q\}$ . Từ đó chúng ta có thể giả thiết rằng có  $P$  là một hệ Sperner trên  $T$ . Chúng ta có thể chứng minh rằng bài toán SDC được dẫn về bài toán của chúng ta bằng một thuật toán thời gian đa thức.

Đặt  $R = T \cup d$ ,  $s = \langle R, F \rangle$ , ở đây  $F = \{P_1 \rightarrow d, \dots, P_n \rightarrow d\}$

Chúng ta xây dựng bảng quyết định DS =  $(U, T \cup \{d\}, V, f)$  như sau:

$U = \{u_0, u_1, \dots, u_m\}$  với mọi  $c \in T: c(u_0) = 0$  và  $d(u_0) = 0$ . Với mọi  $i, i = 1, \dots, m$  và  $c$  là phần tử của  $T$ . Chúng ta đặt  $c(u_i) = 0$  nếu  $c \in Q_i$ . Ngược lại  $c(u_i) = i$ . Đặt  $d(u_i) = i$ .

Rõ ràng  $r$  và  $s$  được xây dựng trong thời gian đa thức theo  $|T|$  và  $|P|, |Q|$ .

Theo Bổ đề 2.3, chúng ta có  $P = K_d^s - \{d\}$ . Do định nghĩa của tập tựa tối thiểu  $A$  theo thuộc tính  $d$  trên sơ đồ quan hệ  $s$  đều tồn tại  $P_i$  sao cho  $P_i \subseteq A$ , ở đây  $1 \leq i \leq n$ . Theo Bổ đề 2.2, chúng ta có  $Q = (K_d^d)^{-1}$  Có nghĩa là  $Q$  là tập các phân khóa của bảng quyết định DS. Do định nghĩa tập phân khóa của DS, chúng ta thấy  $A$  là tập tựa rút gọn của DS khi và chỉ khi nếu đối với mọi  $i = 1, \dots, m$  ta có  $A$  không là tập con của  $Q_i$ .

Gọi  $Q_d$  là tập tất cả các tập tựa rút gọn của bảng quyết định DS với  $d$  là thuộc tính quyết định,  $P_d$  là tập tất cả các tập tựa tối thiểu của thuộc tính  $d$  trên  $s$ . Chúng ta có thể thấy

$Q_d \subseteq P_d$  khi và chỉ khi với mỗi  $A \subseteq T$  và với mọi  $i = 1, \dots, m$  và  $A$  không là tập con của  $Q_i$  thì tồn tại  $P_j$  sao cho  $P_j \subseteq A$ . Từ đó chúng ta thấy rằng vấn đề SDC được dẫn về bài toán của chúng ta bằng một thuật toán có thời gian tính toán là đa thức. Kết quả đã được chứng minh xong

### III. TÀI LIỆU THAM KHẢO

- [1] Gottlob G. Libkin L. Investigation on Armstrong relations dependency inference and excluded functional dependencies. *Acta Cybernetica* (1990), pp. 385-402
- [2] Demetrovics J. and Thi V.D. (1995), “Some remarks on generating Armstrong and inferring functional dependencies relation”, *Acta Cybernetica* 12, pp. 167-180.
- [3] Nguyen Long Giang, Vu Duc Thi (2011), “Some Problems Concerning Condition Attributes and Reducts in Decision Tables”, *Proceeding of the Fifth National Symposium “Fundamental and Applied Information Technology Research” (FAIR)*, Bien Hoa, Dong Nai, pp. 142-152.
- [4] Nguyễn Long Giang, Vũ Đức Thi (2011), “Thuật toán tìm tất cả các rút gọn trong bảng quyết định”, *Tạp chí Tin học và Điều khiển học*, T.27, S.3, tr. 199-205.
- [5] Pawlak Z. (1991), “*Rough sets: Theoretical Aspects of Reasoning About Data*”, Kluwer Academic Publishers.