Tạp chí Tin học và Điều khiển học, T. 19, S. 4 (2003), 353–358

CƠ SỞ DỮ LIỆU QUAN HỆ THÔ VÀ VẤN ĐỀ TỐI ƯU HOÁ CÂU HỎI

NGUYỄN ĐĂNG KHOA

Học viện Hành chính Quốc gia

Abstract. In 1994, on the basis of using the discernibility concept (equivalence relation) and approximation space of rough sets theory, Beaubouef, T. and Petry, F. E. have extended classical relational Database into rough relational Database. The authors have defined the operators: difference, union, intersection, selection, projection, join for rough relations. In this paper, some properties of these operators are proved. These results give us a basis for application of query optimization techniques in rough relational database.

Tóm tắt. Năm 1994 T. Beaubouef, và F. E. Petry trên cơ sở sử dụng các khái niệm quan hệ không phân biệt được (quan hệ tương đương) và không gian xấp xỉ của lý thuyết tập thô đã mở rộng CSDL quan hệ kinh điển thành CSDL quan hệ thô. Các tác giả trên đã định nghĩa các phép toán: hiệu, hợp, giao, chọn, chiếu và kết nối đối với các quan hệ thô. Trong bài báo này, một số tính chất của tổ hợp các phép toán nêu trên đã được chứng minh và dùng làm cơ sở để có thể áp dụng các kỹ thuật tối ưu hoá câu hỏi trong CSDL quan hệ thô.

1. MỞ ĐẦU

Trong [1] theo tiếp cận của lý thuyết tập thô, trên cơ sở sử dụng khái niệm tính không phân biệt được của các phần tử của một tập khi chúng thuộc cùng một lớp tương đương, và ý tưởng biểu diễn một tập bằng các xấp xỉ trên và dưới của nó, các tác giả Beaubouef T. và Petry F. E. đã đề xuất mô hình cơ sở dữ liệu (CSDL) quan hệ thô, là sự mở rộng của mô hình CSDL quan hệ của Codd.

CSDL quan hệ thô có nhiều đặc điểm chung với CSDL quan hệ thông thường: cả hai mô hình đều biểu diễn dữ liệu dưới dạng một tập các quan hệ chứa các bộ. Các bộ trong một quan hệ là không có thứ tự và không được giống nhau (vì một quan hệ là một tập hợp).

Trong CSDL quan hệ thông thường, một bộ t_i của một quan hệ r được xác định trên tập thuộc tính $\{A_1, A_2 ... A_n\}$ có dạng $t_i = (d_{i1}, d_{i2} ... d_{in})$ với $d_{ij} \in \mathrm{Dom}(A_j)$ trong đó $\mathrm{Dom}(A_j)$ là miền trị của thuộc tính A_j, $j = 1, 2 ... n$.

Do đó hai bộ $t_i = (d_{i1} ... d_{in})$ v à $t_k = (d_{k1} ... d_{kn})$ là dư thừa của nhau nếu $d_{ij} = d_{kj}$, $j = 1, 2 ... n$ có nghĩa hai bộ đó là hoàn toàn giống nhau và trong quan hệ r ta chỉ giữ lại một bộ.

Còn trong CSDL quan hệ thô, một bộ thô $t_i \in r(A_1, A_2 ... A_n)$ có dạng

$$t_i = (d_{i1}, d_{i2} ... d_{in}) \quad \text{với} \quad d_{ij} \subseteq \mathrm{Dom}(A_j),\ d_{ij} \neq \emptyset\ \forall j = 1, 2 ... n.$$

Gọi D_j là $\mathrm{Dom}(A_j)$ và ký hiệu $P(D_j) = 2^{D_j} - \emptyset$, trong đó 2^{D_j} là tập luỹ thừa của D_j (là họ tất cả các tập con của D_j).

Từ đó, có thể định nghĩa các khái niệm: quan hệ thô, một thể hiện của một bộ thô, các bộ dư thừa của nhau và các phép toán của đại số quan hệ thô.

Bài báo được tổ chức như sau: Mục 2 giới thiệu ngắn gọn về mô hình CSDL quan hệ thô được trình bày trong [1], bao gồm các định nghĩa về quan hệ thô, bộ thô, thể hiện của một bộ thô, hai bộ thô là dư thừa của nhau và các phép toán quan hệ thô. Trong Mục 3, một số tính chất của các phép toán nêu trong Mục 2 và của tổ hợp các phép toán đó sẽ được chứng minh và làm cơ sở cho việc áp dụng các kỹ thuật tối ưu hoá câu hỏi trên CSDL quan hệ thô. Về các khái niệm cơ bản của tập thô và không gian xấp xỉ, có thể tham khảo [2].

2. MÔ HÌNH CƠ SỞ DỮ LIỆU QUAN HỆ THÔ

Với các ký hiệu được dùng như trong Mục 1, ta có các định nghĩa sau:

Định nghĩa 1. Một quan hệ thô $r(A_1, A_2...A_n)$ là một tập con của tích Đềcác $P(D_1) \times P(D_2) \times \cdots \times P(D_n)$.

Từ đó, nếu t_i là một bộ của r, $t_i = (d_{i1}, d_{i2}...d_{in})$ trong đó $d_{ij} \subseteq D_j$, $\forall j = 1, 2...n$.

Định nghĩa 2. Một thể hiện $\alpha = (a_1...a_n)$ của một bộ thô $t_i = (d_{i1}, d_{i2}...d_{in})$ là một phép gán giá trị bất kỳ sao cho $a_j \in d_{ij}$ với mọi $j = 1, 2...n$.

Trên miền trị của mỗi thuộc tính A_j xác định một quan hệ tương đương (được chỉ định bởi người thiết kế CSDL hay người dùng). Như vậy, trong mỗi miền trị, các giá trị thuộc cùng một lớp tương đương là không phân biệt được và trong tìm kiếm cơ chế hỏi sẽ sử dụng lớp tương đương thay cho sự bằng nhau của các giá trị như đã được dùng trong CSDL quan hệ thông thường.

Gọi $[d_{ij}]$ là lớp tương đương chứa d_{ij}. Nếu $d_{ij} = \{b_1...b_m\}$ thì $[d_{ij}] = [b_1] \cup [b_2] \cup \cdots \cup [b_m]$ là hợp của các lớp tương đương của các phần tử thuộc tập d_{ij}. Trường hợp nếu $b_1...b_m$ cùng thuộc một lớp tương đương (được xác định bởi một quan hệ tương đương nào đó trên D_j) thì $[d_{ij}] = [b_l]$ với $l \in \{1, 2...m.\}$

Định nghĩa 3. Hai bộ $t_i = (d_{i1}, d_{i2}...d_{in})$ và $t_k = (d_{k1}, d_{k2}...d_{kn})$ được gọi là dư thừa của nhau nếu $[d_{ij}] = [d_{kj}]$ với mọi $j = 1, 2...n$.

Trong một quan hệ thô không được chứa các bộ dư thừa của nhau. Nếu có ta chỉ giữ lại một bộ và ưu tiên cho bộ thuộc xấp xỉ dưới.

Định nghĩa 4. (Hiệu của hai quan hệ thô)

Cho X và Y là hai quan hệ thô khả hợp. Hiệu thô giữa X và Y, ký hiệu $X - Y$ là quan hệ thô T, trong đó:

$\underline{R}T = \{t|(t \in \underline{R}X) \wedge (t \not\in \underline{R}Y)\}$, trong đó $\underline{R}T$ là xấp xỉ dưới của T,

$\overline{R}T = \{t|(t \in \overline{R}X) \wedge (t \not\in \overline{R}Y)\}$, trong đó $\overline{R}T$ là xấp xỉ trên của T.

Như vậy $\underline{R}T$ chứa các bộ thuộc xấp xỉ dưới của X mà không dư thừa với một bộ nào trong xấp xỉ dưới của Y.

Cũng tương tự cho $\overline{R}T$.

Định nghĩa 5. (Hợp của hai quan hệ thô)

Cho X và Y là hai quan hệ thô khả hợp. Hợp thô của X và Y, ký hiệu $X \cup Y$, là một quan hệ thô T trong đó $\underline{R}T = \{t|(t \in \underline{R}X) \cup \underline{R}Y)\}$ và $\overline{R}T = \{t|(t \in \overline{R}X) \cup \overline{R}Y)\}$.

Định nghĩa 6. (Giao của hai quan hệ thô)

Cho X và Y là hai quan hệ thô khả hợp. Giao thô của X và Y, ký hiệu $X \cap Y$ là một quan hệ thô T trong đó:

$\underline{R}T = \{t|(t \in \underline{R}X) \wedge (t \in \underline{R}Y)\}$,

$\overline{R}T = \{t|(t \in \overline{R}X) \wedge (t \in \overline{R}Y)\}$,

Định nghĩa 7. (Phép chọn trên một quan hệ thô)

Cho X là một quan hệ thô xác định trên một tập thuộc tính có chứa A, $a = \{a_i\}$ trong đó $a_i,\ b_j \in \mathrm{Dom}(A)$.

Phép chọn thô trên X, ký hiệu $\sigma_{A=\boldsymbol{a}}(X)$ là một quan hệ thô T có cùng lược đồ như X, trong đó:

$$\begin{cases} \underline{R}T = \{t|(t \in \underline{R}X) \wedge (\bigcup_i [a_i] = \bigcup_j [b_j])\} \quad a_i \in \boldsymbol{a},\ b_j \in t[A] \\ \overline{R}T = \{t|(t \in \overline{R}X) \wedge (\bigcup_i [a_i] \subseteq \bigcup_j [b_j])\} \quad a_i \in \boldsymbol{a},\ b_j \in t[A] \end{cases} \tag{2.1}$$

Định nghĩa 8. (Phép chiếu)

Cho X là một quan hệ thô xác định trên tập thuộc tính $\{A_1, A_2...A_n\}$ và $B \subseteq \{A_1, A_2...A_n\}$.

Phép chiếu thô của X lên B, ký hiệu $\Pi_B(X)$ là một quan hệ thô T có lược đồ $T(B)$ trong đó $T(B) = \{t[B]|t \in X\}$.

Trường hợp, sau khi chiếu nếu có hai bộ dư thừa trong đó một bộ thuộc xấp xỉ dưới và một bộ thuộc xấp xỉ trên thì bộ thuộc xấp xỉ dưới được giữ lại.

Định nghĩa 9. (Phép kết nối bằng trên các thuộc tính giống nhau)

Cho $X(A_1, A_2...A_n)$ và $Y(B_1, B_2...B_m)$ là hai quan hệ thô, với $A = \{A_1, A_2...A_m\}$ và $B = \{B_1, B_2...B_n\}$.

Kết nối bằng của hai quan hệ X và Y, ký hiệu $X \bowtie Y$ là quan hệ thô $T(A_1, A_2...A_m,\ B_1, B_2...B_n)$ trong đó

$$\underline{R}T = \{t = (u, v)|u \in \underline{R}X),\ v \in \underline{R}Y),\ u[A \cap B] = v[A \cap B]\},$$

$$\overline{R}T = \{t = (u, v)|u \in \overline{R}X),\ v \in \overline{R}Y),\ (u[A \cap B] \subseteq v[A \cap B]) \vee (v[A \cap B] \subseteq u[A \cap B])\}.$$

Chú ý:

+ Tính không phân biệt được trong CSDL quan hệ thô có thể được biểu diễn bằng một quan hệ phụ trợ. Các bộ của quan hệ này biểu diễn tất cả các giá trị một phần tử có thể d_{ij} cho mỗi miền trị D_j và còn chứa một định danh tuỳ ý của tính không phân biệt được, liên kết giá trị d_{ij} với lớp tương đương chứa nó.
+ Với mỗi quan hệ thô, cần có một cơ chế đánh dấu để phân biệt các bộ thuộc xấp xỉ dưới với các bộ thuộc xấp xỉ trên.
+ Trong [1] phép chọn $\sigma_{A=\boldsymbol{a}}(X)$ trên quan hệ X được định nghĩa là:

$$\begin{cases} \underline{R}T = \{t|(t \in X) \wedge (\bigcup_i^{\sigma_{A=\boldsymbol{a}}=T} [a_i] = \bigcup_j [b_j])\} \quad a_i \in \boldsymbol{a},\ b_j \in t[A] \\ \overline{R}T = \{t|(t \in X) \wedge \bigcup_i [a_i] \subseteq [b_j])\} \quad a_i \in \boldsymbol{a},\ b_j \in t[A] \end{cases}$$

Rõ ràng, định nghĩa theo (2.1) trong Định nghĩa 7 là chính xác hơn. Cũng nhận xét tương tự cho Định nghĩa 9 (phép kết nối bằng).

3. MỘT SỐ TÍNH CHẤT CỦA CÁC TOÁN TỬ QUAN HỆ THÔ

3.1. Tính chất đóng của các toán tử quan hệ thô

Theo định nghĩa của các toán tử quan hệ thô (các toán tử một ngôi và các toán tử hai ngôi) đều cho kết quả là một quan hệ thô.

3.2. Tính chất phân bố của phép chọn đối với các phép hợp, giao và hiệu

Cho $r(A_1, A_2...A_n)$ và $s(A_1, A_2...A_n)$ thì

$$\sigma_{A_{j_0}=\boldsymbol{a}}(r\theta s) = \sigma_{A_{j_0}=\boldsymbol{a}}(r)\theta\sigma_{A_{j_0}=\boldsymbol{a}}(s) \tag{3.1}$$

với $\theta \in \{\cup, \cap, -\}$; $A_{j_0} \in \{A_1...A_n\}$.

Chứng minh. Trong [1] đã chứng minh (3.1) cho trường hợp $\theta = \cap$. Ở đây ta chứng minh tiếp cho $\theta = \cup$ và $\theta = -$

Trường hợp $\theta = \cup$

Ta có: $\sigma_{A_{j_0}=\boldsymbol{a}}(r \cup s) = \sigma_{A_{j_0}=\boldsymbol{a}}(T)$ với $T = r \cup s$. Theo định nghĩa của phép hợp, có

$$\underline{R}T = \{t | t \in \underline{R}r \cup \underline{R}s\} \quad \text{và} \quad \overline{R}T = \{t | t \in \overline{R}r \cup \overline{R}s\}$$

Từ đó, theo định nghĩa của phép chọn: $\sigma_{A_{j_0}=\boldsymbol{a}}(T) = Q$, trong đó

$$\underline{R}Q = \{t | (t \in \underline{R}r \cup \underline{R}s) \wedge (\bigcup_i [a_i] = \bigcup_j [b_j])\},\ a_i \in \boldsymbol{a},\ b_j \in t[A_{j_0}]$$

và
$$\overline{R}Q = \{t | (t \in \overline{R}r \cup \overline{R}s) \wedge (\bigcup_i [a_i] \subseteq \bigcup_j [b_j]\}\ a_i \in \boldsymbol{a},\ b_j \in t[A_{j_0}]$$

Dễ thấy là: $\underline{R}Q = \{t | (t \in \underline{R}r) \wedge (\bigcup_i [a_i] = \bigcup_j [b_j])\} \cup \{t | (t \in \underline{R}r) \wedge (\bigcup_i [a_i] = \bigcup_j [b_j])\}$

và
$$\overline{R}Q = \{t | (t \in \overline{R}r) \wedge (\bigcup_i [a_i] \subseteq [b_j])\} \cup \{t | (t \in \overline{R}s) \wedge (\bigcup_i [a_i] \subseteq [b_j])\}.$$

Suy ra: $\sigma_{A_{j_0}=\boldsymbol{a}}(r \cup s) = Q = \sigma_{A_{j_0}=\boldsymbol{a}}(\{t | t \in r\}) \cup \sigma_{A_{j_0}=\boldsymbol{a}}(\{t | t \in s\}) = \sigma_{A_{j_0}=\boldsymbol{a}}(r) \cup \sigma_{A_{j_0}=\boldsymbol{a}}(s)$.

Trường hợp $\theta = -$

Ta có: $\sigma_{A_{j_0}=\boldsymbol{a}}(r - s) = \sigma_{A_{j_0}=\boldsymbol{a}}(T)$ với $T = r - s$. Theo định nghĩa của phép - , có

$$\underline{R}T = \{t | (t \in \underline{R}r) \wedge (t \notin \underline{R}s)\} \quad \text{và} \quad \overline{R}T = \{t | (t \in \overline{R}r) \wedge (t \notin \overline{R}s)\}$$

Từ đó, theo định nghĩa của phép chọn: $\sigma_{A_{j_0}=\boldsymbol{a}}(T) = Q$, trong đó

$$\underline{R}Q = \{t | (t \in \underline{R}r - \underline{R}s) \wedge (\bigcup_i [a_i] = \bigcup_j [b_j])\},\ a_i \in \boldsymbol{a},\ b_j \in t[A_{j_0}]\}$$

và
$$\overline{R}Q = \{t | (t \in \overline{R}r - \overline{R}s) \wedge (\bigcup_i [a_i] \subseteq \bigcup_j [b_j])\},\ a_i \in \boldsymbol{a},\ b_j \in t[A_{j_0}]$$

Dễ thấy là
$$\underline{R}Q = \{t | (t \in \underline{R}r) \wedge (\bigcup_i [a_i] = \bigcup_j [b_j])\} - \{t | (t \in \underline{R}s) \wedge (\bigcup_i [a_i] = \bigcup_j [b_j])\},$$

và
$$\overline{R}Q = \{t | (t \in \overline{R}r) \wedge (\bigcup_i [a_i] \subseteq \bigcup_j [b_j])\} - \{t | (t \in \overline{R}s) \wedge (\bigcup_i [a_i] \subseteq \bigcup_j [b_j])\}.$$

Suy ra
$$\sigma_{A_{j_0}=\boldsymbol{a}}(r - s) = Q = \sigma_{A_{j_0}=\boldsymbol{a}}\{t | t \in r\}) - \sigma_{A_{j_0}=\boldsymbol{a}}(\{t | t \in s\}) = \sigma_{A_{j_0}=\boldsymbol{a}}(r - s).$$

3.3. Tính chất của một dãy liên tiếp các phép chiếu một quan hệ thô trên các tập thuộc tính lồng nhau

Cho $r(A_1, A_2...A_n)$ là một quan hệ thô và $X_1 \subseteq X_2 \subseteq \cdots \subseteq X_m$ là một dãy các tập con của $\{A_1, A_2...A_n\}$. Khi đó

$$\Pi_{X_1}(\Pi_{X_2}(\Pi_{X_2}(...(\Pi_{X_m}(r))...)) = \Pi_{X_1}(r) \tag{3.2}$$

Chứng minh. Trước hết ta chứng minh (3.2) cho trường hợp $m = 2$. Cụ thể

$$\Pi_{X_1}(\Pi_{X_2}(r)) = \Pi_{X_1}(r) \quad \text{với} \quad X_1 \subseteq X_2 \subseteq \{A_1, A_2...A_n\}. \tag{3.3}$$

Gọi $\Pi_{X_2}(r) = s$. Theo định nghĩa của phép chiếu, s có lược đồ $s(X_2)$ và $s(X_2) = \{t[X_2] | t \in r\}$ đồng thời, sau khi loại bỏ khỏi r các thuộc tính (cột) không thuộc s, nếu có hai bộ $t_1 \in \underline{R}r$, $t_2 \in \overline{R}r$ sao cho $t_1[X_2]$ và $t_2[X_2]$ là các bộ dư thừa của nhau thì bộ $t_1[X_2]$ được giữ lại trong hình chiếu.

Với cách thao tác như vậy, rõ ràng vế trái và vế phải của (3.3) là hai quan hệ thô sao cho mỗi bộ của vế trái là dư thừa với một và chỉ một bộ của vế phải và ngược lại. Ngoài ra, vết của các bộ thuộc xấp xỉ dưới của r vẫn được bảo toàn. Đẳng thức (3.3) được chứng minh. Việc mở rộng (3.3) để có (3.2) hiển nhiên.

3.4. Tính chất của phép chọn theo một điều kiện hội

Cho $r(A_1, A_2...A_n)$ là một quan hệ thô. Khi đó

$$\sigma_{(A_{j_0}=\boldsymbol{a}^{(0)} \wedge A_{j_1}=\boldsymbol{a}^{(1)})}(r) = \sigma_{(A_{j_0}=\boldsymbol{a}^{(0)}}(\sigma_{(A_{j_1}=\boldsymbol{a}^{(1)}}(r)) = \sigma_{(A_{j_1}=\boldsymbol{a}^{(1)}}(\sigma_{(A_{j_0}=\boldsymbol{a}^{(0)}}(r)) \tag{3.4}$$

trong đó $A_{j_0},\ A_{j_1},\ \in \{A_1, A_2...A_n\}$ và $A_{j_0} \neq A_{j_1}$.

Chứng minh. Gọi $\sigma_{(A_{j_0}=\boldsymbol{a}^{(0)} \wedge A_{j_1}=\boldsymbol{a}^{(1)})}(r) = T$. Theo định nghĩa của phép chọn,

$$\underline{R}T = \{t | (t \in \underline{R}r) \wedge (\bigcup_i [a_i^{(0)}] = (\bigcup_j [b_i^{(0)}]) \wedge (\bigcup_i [a_i^{(1)}] = \bigcup_j [b_j^{(1)}])\},$$

$$a_i^{(0)} \in \boldsymbol{a}^{(0)},\ a_i^{(1)} \in \boldsymbol{a}^{(1)},\ b_j^{(0)} \in t[A_{j_0}],\ b_j^{(1)} \in t[A_{j_1}];$$

và

$$\overline{R}T = \{t | (t \in \overline{R}r) \wedge (\bigcup_i [a_i^{(0)}] \subseteq (\bigcup_j [b_i^{(0)}]) \wedge (\bigcup_i [a_i^{(1)}] \subseteq \bigcup_j [b_j^{(1)}])\}.$$

$$a_i^{(0)} \in \boldsymbol{a}^{(0)},\ a_i^{(1)} \in \boldsymbol{a}^{(1)},\ b_j^{(0)} \in t[A_{j_0}],\ b_j^{(1)} \in t[A_{j_1}].$$

Từ đó

$$T = \sigma_{(A_{j_0}=\boldsymbol{a}^{(0)}}(\sigma_{(A_{j_1}=\boldsymbol{a}^{(1)}}(r)) = \sigma_{(A_{j_1}=\boldsymbol{a}^{(1)}}(\sigma_{(A_{j_0}=\boldsymbol{a}^{(0)}}(r)).$$

Đẳng thức (3.4) được chứng minh.

3.5. Ta sẽ chứng minh tính chất sau

Cho $r(A_1, A_2...A_m)$ và $s(B_1, B_2...B_n)$ là các quan hệ thô với $A = \{A_1, A_2...A_m\}$ và $B = \{B_1, B_2...B_n\}$. Khi đó

$$\sigma_{(A_{j_0}=\boldsymbol{a}^{(0)})} \wedge (B_{k_1=\boldsymbol{a}^{(1)}})(r \bowtie s) = \sigma_{(A_{j_0}=\boldsymbol{a}^{(0)}}(r) \bowtie \sigma_{B_{k_1}=\boldsymbol{a}^{(1)}}(s) \tag{3.5}$$

trong đó $A_{j_0} \in A - A \cap B,\ B_{k_1} \in B - A \cap B$, còn $\bowtie$ là phép kết nối bằng trên $A \cap B \neq \emptyset$.

Chứng minh. Gọi vế trái của (3.5) là T. Ta có

$$\underline{R}T = \{t | (t = (u, v)) \wedge (u \in \underline{R}r) \wedge (v \in \underline{R}s) \wedge (u[A \cap B]) = v[A \cap B]) \wedge (\bigcup_i [a_i^{(0)}]$$
$$= \bigcup_j [b_j^{(0)}] \wedge (\bigcup_i [a_i^{(1)}]) = \bigcup_j [b_j^{(1)}]\}$$

với $a_i^{(0)} \in \boldsymbol{a}^{(0)},\ b_j^{(0)} \in t[A_{j_0}],\ a_i^{(1)} \in \boldsymbol{a}^{(1)},\ b_j^{(1)} \in t[B_{k_1}]$.

và

$$\overline{R}T = \{t|(t = (u, v)) \wedge (u \in \overline{R}r) \wedge (v \in \overline{R}s) \wedge (u[A \cap B]) \subseteq v[A \cap B]) \vee (v[A \cap B] \subseteq u[A \cap B] \wedge \bigcup_i [a_i^{(0)}] \subseteq \bigcup_j [b_j^{(0)}]) \wedge (\bigcup_i [a_i^{(1)}] \subseteq \bigcup_j [b_j^{(1)}])\}$$

với

$$a_i^{(0)} \in \boldsymbol{a}^{(0)},\ b_j^{(0)} \in t[A_{j_0}],\ a_i^{(1)} \in \boldsymbol{a}^{(1)},\ b_j^{(1)} \in t[B_{k_1}].$$

Gọi vế phải của (3.5) là Z. Dễ thấy là $\underline{R}Z = \underline{R}T$ và $\overline{R}Z = \overline{R}T$.

Ở đây sự bằng nhau của hai quan hệ thô được hiểu theo nghĩa mỗi bộ của quan hệ thứ nhất dư thừa với một và chỉ một bộ của quan hệ thứ hai và ngược lại. Đẳng thức (3.5) đã được chứng minh. ∎

Từ các tính chất (3.1) – (3.5) đã được chứng minh, dễ thấy là các kỹ thuật tối ưu hoá heuristic các biểu thức hỏi của đại số quan hệ trong CSDL quan hệ thông thường vẫn còn áp dụng được cho CSDL thô như: thực hiện phép chọn sớm nhất có thể, thực hiện sớm các phép chiếu, tránh thực hiện tích Đềcat...

TÀI LIỆU THAM KHẢO

[1] Beaubouef, F. E. Petry, A rough set model for relational Databases, *Rough sets, Fuzzy sets and Knowledge Discovery,* Wojciech P. Ziarko (Ed.) Springer Verlag, 1994.

[2] Nguyễn Đăng Khoa, Cách tiếp cận tập thô trong việc phát hiện tri thức trong cơ sở dữ liệu, *Tạp chí Tin học và Điều khiển học* **18** (4) (2002) 309 – 316.

[3] Nguyễn Đăng Khoa, Rút gọn tập thô, Tuyển tập “Báo cáo tại Hội nghị Khoa học nhân dịp kỷ niệm 40 năm ngày thành lập trường ĐH Bách Khoa Hà Nội” (2001).

Nhận bài ngày 01 - 9 - 2003