Kỷ yếu Hội nghị Quốc gia lần thứ X về Nghiên cứu cơ bản và ứng dụng Công nghệ thông tin (FAIR), Đà Nẵng, ngày 17-18/08/2017
DOI: 10.15625/vap.2017.00074

MỘT SỐ KẾT QUẢ VỀ KHAI PHÁ LUẬT QUYẾT ĐỊNH TRÊN KHỐI DỮ LIỆU CÓ GIÁ TRỊ THUỘC TÍNH THAY ĐỔI

Trịnh Đình Thắng[1], Trần Minh Tuyến[2], Đỗ Thị Lan Anh[3], Nguyễn Thị Quyên[3]

[1] ĐHSP Hà Nội 2, [2] ĐH Công đoàn, [3] ĐHSP Hà Nội 2,

thangsp2@yahoo.com, tuyentm@dhcd.edu.vn, lananh.cntt.sp2@gmail.com, quyen.cntt.sp2@gmail.com

TÓM TẮT: *Báo cáo đề xuất và chứng minh một số tính chất của việc làm mịn, làm thô các giá trị của thuộc tính chỉ số điều kiện hoặc quyết định đối với khối thông tin và lát cắt của nó. Mỗi khi các lớp tương đương điều kiện hoặc quyết định trên khối quyết định được làm mịn, làm thô thì chúng sẽ cảm sinh một phần hoặc cảm sinh việc làm mịn, làm thô các lớp tương ứng trên lát cắt. Từ các kết quả tìm được của việc làm mịn, làm thô các lớp tương đương điều kiện hoặc quyết định được cảm sinh một phần hoặc cảm sinh trên lát cắt thì việc tính các ma trận độ hỗ trợ trên lát cắt sẽ được đơn giản hơn giống như việc tính các ma trận đó trên khối quyết định khi làm mịn, làm thô các giá trị của thuộc tính chỉ số điều kiện hoặc quyết dịnh.*

Từ khóa: *Khối thông tin, khai phá luật quyết định, khối dữ liệu.*

I. CÁC KHÁI NIỆM

1. Khối, lát cắt của khối [1]

Định nghĩa I.1

Gọi R = (id; $A_1, A_2,..., A_n$) *là một bộ hữu hạn các phần tử, trong đó* id *là tập chỉ số hữu hạn khác rỗng,* A_i (i=1..n) *là các thuộc tính. Mỗi thuộc tính* A_i (i=1..n) *có miền giá trị tương ứng là* dom(A_i). *Một khối* r *trên* R, *kí hiệu* r(R) *gồm một số hữu hạn phần tử mà mỗi phần tử là một họ các ánh xạ từ tập chỉ số* id *đến các miền trị của các thuộc tính* A_i (i=1..n).

Nói một cách khác:

$$t \in r(R) \Leftrightarrow t = \{ t^i : id \to dom(A_i)\}_{i=1..n} .$$

Ta kí hiệu khối đó là r(R) hoặc r(id; $A_1, A_2,..., A_n$), đôi khi nếu không gây nhầm lẫn ta kí hiệu đơn giản là r.

Định nghĩa I.2

Cho R = (id; $A_1, A_2,..., A_n$), r(R) *là một khối trên* R. *Với mỗi* $x \in$ id *ta kí hiệu* r(R_x) *là một khối với* R_x = ({x}; $A_1, A_2,..., A_n$) *sao cho*:

$$t_x \in r(R_x) \Leftrightarrow t_x = \{t^i_x = t^i\big|_x\}_{i=1..n}, \text{ ở đây } t \in r(R), t = \{ t^i : id \to dom(A_i)\}_{i=1..n},$$

Khi đó r(R_x) *được gọi là một lát cắt trên khối* r(R) *tại điểm* x, *đôi khi ta kí hiệu là* r_x.

Sau đây, để cho đơn giản ta sử dụng các kí hiệu:

$$x^{(i)} = (x; A_i) \text{ ; } id^{(i)} = \{x^{(i)} \mid x \in id\}.$$

Ta gọi $x^{(i)}$ ($x \in$ id, i = 1..n) là các thuộc tính chỉ số của lược đồ khối R = (id; $A_1,A_2,...,A_n$).

2. Khối thông tin [5]

Định nghĩa I.3: *Cho lược đồ khối R = (id; $A_1, A_2, \ldots , A_n$), r là một khối trên R. Khi đó khối thông tin là một bộ bốn IB = (U, A, V, f) với U là tập các đối tượng thuộc r gọi là không gian các đối tượng, A =* $\bigcup_{i=1}^{n} id^{(i)}$ *là tập các thuộc tính chỉ số của đối tượng, V =* $\bigcup_{x^{(i)} \in A} V_{x^{(i)}}$, $V_{x^{(i)}}$ *là tập giá trị của các đối tượng ứng với thuộc tính chỉ số* $x^{(i)}$, *f là hàm thông tin* $U \times A \to V$ *thỏa mãn:* $\forall u \in U$, $\forall x^{(i)} \in A$ *ta có* $f(u, x^{(i)}) \in V_{x^{(i)}}$.

Ta gọi $f(u, x^{(i)})$ là giá trị của đối tượng u tại thuộc tính chỉ số $x^{(i)}$.

Nếu V chứa giá trị thiếu ở ít nhất một thuộc tính chỉ số $x^{(i)} \in A$ thì ta nói IB là khối thông tin không đầy đủ, ngược lại IB là khối thông tin đầy đủ, hay nói một cách đơn giản là khối thông tin.

Định nghĩa I.4: *Cho lược đồ khối R = (id; $A_1, A_2, \ldots , A_n$), r là một khối trên R, r_x là lát cắt của khối r tại điểm*

$x \in id$. Khi đó lát cắt của khối thông tin tại x là một bộ bốn $IB_x = (U, A_x, V_x, f_x)$ với U là tập các đối tượng thuộc r gọi là không gian các đối tượng, $A_x = \bigcup_{i=1}^{n} x^{(i)}$ là tập các thuộc tính chỉ số của đối tượng trên lát cắt tại x, $V_x = \bigcup_{x^{(i)} \in A_x} V_{x^{(i)}}$, $V_{x^{(i)}}$ là tập giá trị của các đối tượng ứng với thuộc tính chỉ số $x^{(i)}$, f_x là hàm thông tin $U x A_x \to V_x$ thỏa mãn: $\forall u \in U, \forall x^{(i)} \in A_x$ ta có $f(u, x^{(i)}) \in V_{x^{(i)}}$.

3. Quan hệ không phân biệt được [5]

Định nghĩa I.5

Cho khối thông tin IB = (U, A, V, f). Khi đó với mỗi tập thuộc tính chỉ số $P \subseteq A$ ta xác định một quan hệ tương đương, kí hiệu IND(P) định nghĩa như sau:

$$IND(P) = \{(u,v) \in U x U \mid \forall x^{(i)} \in P: f(u, x^{(i)}) = f(v, x^{(i)})\},$$

và gọi là quan hệ không phân biệt được:

Từ định nghĩa ta có: IND(P) = $\bigcap_{x^{(i)} \in P} IND(x^{(i)})$.

Quan hệ IND(P) chia U thành các lớp tương đương, tạo nên một phân hoạch của U, kí hiệu U/IND(P) hay đơn giản hơn là U/P.

Với mỗi $u \in U$, lớp tương đương chứa u theo quan hệ IND(P), kí hiệu $[u]_P$ được định nghĩa như sau:

$$[u]_P = \{v \in U \mid (u,v) \in IND(P)\}.$$

Theo định nghĩa này ta thấy: hai phần tử u, v $\in$ U thuộc cùng một lớp tương đương khi và chỉ khi chúng có giá trị như nhau trên mọi thuộc tính chỉ số trong P.

Định nghĩa I.6

Cho khối thông tin IB = (U, A, V, f), P, $Q \subseteq A$ là tập các thuộc tính chỉ số, $U/P = \{P_1, P_2, \dots, P_m\}$, $U/Q = \{Q_1, Q_2, \dots, Q_n\}$ là các phân hoạch được sinh bởi P, Q tương ứng. Khi đó ta nói phân hoạch theo Q thô hơn phân hoạch theo P, hay phân hoạch theo P mịn hơn phân hoạch theo Q khi và chỉ khi:

$\forall P_i \in U/P, \exists Q_j \in U/Q: P_i \subseteq Q_j$, $i = 1..m$, $j = 1..n$.

4. Khối quyết định [5]

Định nghĩa I.7

Cho khối thông tin IB = (U,A,V,f) với U là không gian các đối tượng, $A = \bigcup_{i=1}^{n} id^{(i)}$. Khi đó nếu A được chia thành 2 tập C và D sao cho:

$$C = \bigcup_{i=1, x \in id}^{k} x^{(i)}, \; D = \bigcup_{i=k+1, x \in id}^{n} x^{(i)},$$

thì khối thông tin IB gọi là khối quyết định và kí hiệu là $DB=(U, C \cup D, V, f)$, với C là tập các thuộc tính chỉ số điều kiện và D là tập các thuộc tính chỉ số quyết định.

Từ định nghĩa của khối quyết định, ta thấy:

$$C \cup D = A, \; C \cap D = \varnothing,$$

Ta có thể kí hiệu khối quyết định một cách đơn giản là: DB=(U, $C \cup D$).

Định nghĩa I.8: *Cho khối quyết định $DB=(U, C \cup D, V, f)$, với C là tập các thuộc tính chỉ số điều kiện và D là tập các thuộc tính chỉ số quyết định. Khi đó lát cắt của khối quyết định tại x ($x \in id$) là một bộ bốn $DB_x = (U, C^x \cup D^x, V_x, f_x)$ với U là tập các đối tượng thuộc r gọi là không gian các đối tượng,*

$$C^x = \bigcup_{i=1}^{k} x^{(i)}, \; D^x = \bigcup_{i=k+1}^{n} x^{(i)}, A_x = C^x \cup D^x,$$

$V_x = \bigcup_{x^{(i)} \in A_x} V_{x^{(i)}}$, $V_{x^{(i)}}$ là tập các giá trị của các đối tượng ứng với thuộc tính chỉ số $x^{(i)}$, f_x là hàm thông tin $U x A_x \to V_x$ thỏa mãn: $\forall u \in U, \forall x^{(i)} \in A_x$ ta có $f(u, x^{(i)}) \in V_{x^{(i)}}$.

Nhận xét:

Cho khối quyết định DB=(U,C∪D,V,f). Khi đó, nếu id={x} thì khối quyết định DB suy biến thành bảng quyết định như đã biết.

Khi nghiên cứu khối quyết định, người ta muốn rút ra từ đó các luật quyết định. Trong các luật quyết định này, phần điều kiện sẽ tương ứng với các thuộc tính chỉ số điều kiện, phần kết luận sẽ tương ứng với các thuộc tính chỉ số quyết định.

Các luật quyết định được rút ra từ khối quyết định được chia thành hai loại:

i) Các luật đúng trên toàn khối.

ii) Các luật chỉ đúng trên từng lát cắt cụ thể.

5. Luật quyết định [5]

Định nghĩa I.9

Cho khối quyết định DB = (U, C∪D), *với U là không gian các đối tượng:*

$$C = \bigcup_{i=1,x\in id}^{k} x^{(i)},\ D = \bigcup_{i=k+1,x\in id}^{n} x^{(i)},\ \text{và } C^x=\bigcup_{i=1}^{k} x^{(i)},\ D^x=\bigcup_{i=k+1}^{n} x^{(i)},\ x\in id.$$

Khi đó:

$U/C=\{C_1,C_2,...,C_m\}$, $U/C^x=\{C_{x1},C_{x2},...,C_{xt_x}\}$,

$U/D=\{D_1,D_2,...,D_k\}$, $U/D^x=\{D_{x1},D_{x2},...,D_{xh_x}\}$, *tương ứng là các phân hoạch được sinh ra bởi C, C^x, D, D^x. Một luật quyết định trên khối có dạng:*

$C_i \rightarrow D_j$, *i=1..m, j=1..k* ,

và trên lát cắt tại điểm x có dạng:

$C_{xi} \rightarrow D_{xj}$, *$i=1..t_x$, $j=1..h_x$* .

Mệnh đề I.1

Cho khối quyết định DB = (U, C∪D), *với U là không gian các đối tượng:*

$$C= \bigcup_{i=1,x\in id}^{k} x^{(i)},\ D= \bigcup_{i=k+1,x\in id}^{n} x^{(i)},\ \text{và } C^x=\bigcup_{i=1}^{k} x^{(i)},\ D^x=\bigcup_{i=k+1}^{n} x^{(i)},\ x\in id.$$

$U/C=\{C_1,C_2,...,C_m\}$, $U/C^x=\{C_{x1},C_{x2},...,C_{xt_x}\}$, $U/D=\{D_1,D_2,...,D_k\}$, $U/D^x=\{D_{x1},D_{x2},...,D_{xh_x}\}$,

Khi đó: $\forall C_i\in U/C$, $\forall D_j\in U/D$ ta có:

$$C_i = \bigcap_{x\in id} C_{xp_x},\ D_j = \bigcap_{x\in id} D_{xq_x}\ \text{ với } p_x\in\{1,2,...,t_x\},\ q_x\in\{1,2,...,h_x\}.$$

Định nghĩa I.10

Cho khối quyết định DB=(U,C∪D), $C_i\in U/C$, $D_j\in U/D$, $C_{xp}\in U/C^x$, $D_{xq}\in U/D^x$, *i=1..m, j=1..n, $p\in\{1,2,...,t_x\}$, $q\in\{1,2,...,h_x\}$, $x\in id$. Khi đó, độ hỗ trợ, độ chính xác và độ phủ của luật quyết định $C_i\rightarrow D_j$ trên khối là:*

- Độ hỗ trợ: $Sup(C_i,D_j) = |C_i\cap D_j|$,

- Độ chính xác: $Acc(C_i,D_j) = \dfrac{|C_i \cap D_j|}{|C_i|}$,

- Độ phủ: $Cov(C_i,D_j) = \dfrac{|C_i \cap D_j|}{|D_j|}$,

còn đối với luật quyết định $C_{xp} \rightarrow D_{xq}$ trên lát cắt của khối tại điểm x là:

- Độ hỗ trợ: $Sup(C_{xp},D_{xq}) = |C_{xp}\cap D_{xq}|$,

- Độ chính xác: $\mathrm{Acc}(C_{xp},D_{xq}) = \frac{|C_{xp} \cap D_{xq}|}{|C_{xp}|}$,

- Độ phủ: $\mathrm{Cov}(C_{xp},D_{xq}) = \frac{|C_{xp} \cap D_{xq}|}{|D_{xq}|}$.

Từ định nghĩa này, ta có:

$0 \le \mathrm{Acc}(C_i,D_j) \le 1$, $0 \le \mathrm{Acc}(C_{xp},D_{xq}) \le 1$, $\sum_{j=1}^{n} Acc(C_i, D_j) = 1$, $\sum_{q=1}^{h_x} Acc(C_{xp}, D_{xq}) = 1$,

$0 \le \mathrm{Cov}(C_i,D_j) \le 1$, $0 \le \mathrm{Cov}(C_{xp},D_{xq}) \le 1$,

$\sum_{i=1}^{m} Cov(C_i, D_j) = 1$, $\sum_{p=1}^{t_x} Cov(C_{xp}, D_{xq}) = 1$.

Ta có thể biểu diễn độ đo của các luật quyết định trên khối dưới dạng các ma trận độ đo như sau:

- Ma trận độ hỗ trợ:

$$\mathrm{Sup}(C_i,D_j)_{mxn} = \begin{pmatrix} Sup(C_1, D_1) & \ldots & Sup(C_1, D_n) \\ & \ldots & \\ Sup(C_m, D_1) & \ldots & Sup(C_m, D_n) \end{pmatrix}$$

- Ma trận độ chính xác:

$$\mathrm{Acc}(C_i,D_j)_{mxn} = \begin{pmatrix} Acc(C_1, D_1) & \ldots & Acc(C_1, D_n) \\ & \ldots & \\ Acc(C_m, D_1) & \ldots & Acc(C_m, D_n) \end{pmatrix}$$

- Ma trận độ phủ:

$$\mathrm{Cov}(C_i,D_j)_{mxn} = \begin{pmatrix} Cov(C_1, D_1) & \ldots & Cov(C_1, D_n) \\ & \ldots & \\ Cov(C_m, D_1) & \ldots & Cov(C_m, D_n) \end{pmatrix}$$

Với các luật quyết định trên các lát cắt của khối thì ta cũng có các ma trận độ hỗ trợ, độ chính xác và độ phủ tương tự.

Định nghĩa I.11

Cho khối quyết định DB=(U,C∪D), $C_i \in U/C$, $D_j \in U/D$ tương ứng là các lớp tương đương điều kiện và các lớp tương đương quyết định được sinh bởi C, D, $C_i \rightarrow D_j$ là một luật quyết định trên khối DB, i=1..m, j=1..n.

- *Nếu $Acc(C_i \rightarrow D_j) = 1$ thì $C_i \rightarrow D_j$ gọi là một luật quyết định chắc chắn.*
- *Nếu $0 < Acc(C_i \rightarrow D_j) < 1$ thì $C_i \rightarrow D_j$ gọi là một luật quyết định không chắc chắn.*

Mệnh đề I.2

Cho khối quyết định DB = (U, C∪D), *với U là không gian các đối tượng:*

$C = \bigcup_{i=1, x \in id}^{k} x^{(i)}$, $D = \bigcup_{i=k+1, x \in id}^{n} x^{(i)}$.

Khi đó $\forall C_i \in U/C$, $\forall D_j \in U/D$, (i=1..m, j=1..n) ta có:

i) $\mathrm{Acc}(C_i \rightarrow D_j) = \frac{Sup(C_i, D_j)}{\sum_{q=1}^{n} Sup(C_i, D_q)}$,

ii) $\mathrm{Cov}(C_i \rightarrow D_j) = \frac{Sup(C_i, D_j)}{\sum_{p=1}^{m} Sup(C_p, D_j)}$.

Định nghĩa I.12

Cho khối quyết định DB=(U,C∪D), $C_i \in U/C$, $D_j \in U/D$, i=1..m, j=1..n tương ứng là các lớp tương đương điều kiện và các lớp tương đương quyết định được sinh bởi C, D; α, β là hai ngưỡng cho trước (α, β∈(0,1)). Khi đó, nếu $Acc(C_i,D_j) \geq \alpha$ và $Cov(C_i,D_j) \geq \beta$ thì ta gọi $C_i \rightarrow D_j$ là luật quyết định có ý nghĩa.

Định nghĩa I.13

Cho khối quyết định DB=(U,C∪D,V,f), *với U là không gian các đối tượng, a∈ C∪D, V_a là tập các giá trị hiện có của thuộc tính chỉ số a. Giả sử $Z=\{x_s \in U \mid f(x_s,a) = z\}$ là tập các đối tượng có giá trị z trên thuộc tính chỉ số a. Nếu Z được phân hoạch thành hai tập W và Y sao cho: Z=W∪Y, W∩Y=∅ với $W=\{x_p \in U \mid f(x_p,a) = w, w \notin V_a\}$, $Y==\{x_q \in U \mid f(x_q,a) = y, y \notin V_a\}$, thì ta nói giá trị z của thuộc tính chỉ số a được làm mịn thành hai giá trị mới w và y.*

Định nghĩa I.14

Cho khối quyết định DB=(U,C∪D,V,f), *với U là không gian các đối tượng, a∈ C∪D, V_a là tập các giá trị hiện có của thuộc tính chỉ số a. Giả sử $f(x_p,a)=w$, $f(x_q,a)=y$ tương ứng là giá trị của x_p, x_q trên thuộc tính chỉ số a (p≠q). Nếu tại thời điểm nào đó ta có: $f(x_p,a)= f(x_q,a)=z$, $(z \notin V_a)$ thì ta nói hai giá trị w, y của a được làm thô thành giá trị mới z.*

Định lý I.1

Cho khối quyết định DB = (U, C∪D, V, f), *với U là không gian các đối tượng, a∈ C∪D, V_a là tập các giá trị hiện có của thuộc tính chỉ số a. Khi đó, hai lớp tương đương E_p, E_q nào đó (E_p, $E_q \in U/E$, E∈{C,D}) được làm thô thành lớp tương đương mới E_s khi và chỉ khi $\forall a_j \neq a$: $f(E_p,a_j) = f(E_q,a_j)$.*

Định lý I.2

Cho khối quyết định DB = (U, C∪D, V, f), *với U là không gian các đối tượng, a∈ C∪ D, V_a là tập các giá trị hiện có của thuộc tính chỉ số a. Khi đó, lớp tương đương E_s ($E_s \in U/E$, E∈{C,D}) được làm mịn thành hai lớp tương đương mới E_p, E_q nào đó khi và chỉ khi ta có thể đặt: $f(E_p,a)=w$, $f(E_q,a)=y$ và $E_p \cup E_q= E_s$, w, $y \notin V_a$, w≠y.*

Định lý I.3

Cho khối quyết định DB = (U, C∪D), *với U là không gian các đối tượng:*

$$C=\bigcup_{i=1,x\in id}^{k} x^{(i)},\ D=\bigcup_{i=k+1,x\in id}^{n} x^{(i)},\ \text{và } C^x=\bigcup_{i=1}^{k} x^{(i)},\ D^x=\bigcup_{i=k+1}^{n} x^{(i)},\ x\in id.$$

$$U/C=\{C_1,C_2,...,C_m\},\ U/C^x=\{C_{x1},C_{x2},...,C_{xt_x}\},\ U/D=\{D_1,D_2,...,D_k\},\ U/D^x=\{D_{x1},D_{x2},...,D_{xh_x}\},$$

α, β là hai ngưỡng cho trước (α, β∈(0,1)).

Khi đó: nếu $C_i \rightarrow D_j$ là một luật quyết định có ý nghĩa trên khối quyết định thì nó cũng là một luật quyết định có ý nghĩa trên một lát cắt bất kì của khối quyết định tại x∈id.

II. KẾT QUẢ NGHIÊN CỨU

1. Làm mịn, thô các lớp tương đương điều kiện trên khối quyết định và trên lát cắt.

Mệnh đề II.1

Cho khối quyết định DB = (U, C∪D, V, f), *$a=x^{(i)} \in C$, V_a là tập các giá trị hiện có của thuộc tính chỉ số điều kiện a, giá trị z của a được làm mịn thành hai giá trị mới w và y.*

$$C=\bigcup_{i=1,x\in id}^{k} x^{(i)},\ D=\bigcup_{i=k+1,x\in id}^{n} x^{(i)},\ \text{và } C^x=\bigcup_{i=1}^{k} x^{(i)},\ D^x=\bigcup_{i=k+1}^{n} x^{(i)},\ x\in id.$$

$$U/C=\{C_1,C_2,...,C_m\},\ U/C^x=\{C_{x1},C_{x2},...,C_{xt_x}\},\ U/D=\{D_1,D_2,...,D_k\},\ U/D^x=\{D_{x1},D_{x2},...,D_{xh_x}\},$$

Khi đó, nếu lớp tương đương điều kiện $C_s \in U/C$, ($f(C_s,a)=z$) được làm mịn thành hai lớp tương đương điều kiện mới C_p,C_q ($f(C_p,a)=w$, $f(C_q,a)=y$, với $w,y \notin V_a$) nào đó thì trên lát cắt r_x, tồn tại lớp tương đương C_{xi} thỏa mãn: $C_s \subseteq C_{xi}$, cũng được làm mịn thành hai lớp tương đương điều kiện mới $C_{xi'}$ và $C_{xi''}$ sao cho: $C_p \subseteq C_{xi'}$, $C_q \subseteq C_{xi''}$ ($f(C_{xi'},a)=w$, $f(C_{xi''},a)=y$).

Ta nói trên lát cắt r_x thì C_{xi} được làm mịn cảm sinh một phần *thành hai lớp tương đương điều kiện mới $C_{xi'}$ và $C_{xi''}$* bởi sự làm mịn của C_s *thành hai lớp tương đương điều kiện mới C_p,C_q.*

Chứng minh

Theo giả thiết ta có: $C_s \in U/C$, *(f(*C_s*,a)=z) được làm mịn thành hai lớp tương đương điều kiện mới* C_p, C_q *(f(*C_p*,a)=w, f(*C_q*,a)=y, với* $w,y \notin V_a$ *).* Vì $C_s \in U/C$ nên áp dụng kết quả của mệnh đề I.1 ta có: $C_s = \bigcap_{x\in id} C_{xp_x}$ *,từ đó suy ra* $\exists\, C_{xi} \in U/C^x$ thỏa mãn: $C_s \subseteq C_{xi}$. Mặt khác, do C_s được làm mịn thành hai lớp C_p *và* C_q *nên theo định lý I.2 ta có:* $C_s = C_p \cup C_q \Rightarrow C_p, C_q \subseteq C_{xi}$ *với f(*C_p*,a)=w, f(*C_q*,a)=y.*

Sau cùng, ta gán cho mỗi phần tử $u \in C_{xi} \setminus C_s$ tại thuộc tính chỉ số a một trong hai giá trị w hoặc y thì ta được một phân hoạch của C_{xi} thành hai lớp tương đương điều kiện mới là $C_{xi'}$ *và* $C_{xi''}$ *thỏa mãn f(*$C_{xi'}$*,a)=w, f(*$C_{xi''}$*,a)=y và* $C_{xi} = C_{xi'} \cup C_{xi''}$.

Kết quả là trên lát cắt r_x thì *lớp tương đương điều kiện* C_{xi} *thỏa mãn:* $C_s \subseteq C_{xi}$ *, cũng được làm mịn thành hai lớp tương đương điều kiện mới* $C_{xi'}$ *và* $C_{xi''}$ *sao cho:* $C_p \subseteq C_{xi'}$*,* $C_q \subseteq C_{xi''}$ *(f(*$C_{xi'}$*,a)=w, f(*$C_{xi''}$*,a)=y) và* $C_{xi} = C_{xi'} \cup C_{xi''}$.

Mệnh đề II.2

Cho khối quyết định DB = (U, C∪D), a=$x^{(i)} \in C$, V_a *là tập các giá trị hiện có của thuộc tính chỉ số điều kiện a, giá trị z của a được làm mịn thành hai giá trị mới w và y.*

$$C=\bigcup_{i=1,x\in id}^{k} x^{(i)},\ D=\bigcup_{i=k+1,x\in id}^{n} x^{(i)},\ \text{và } C^x=\bigcup_{i=1}^{k} x^{(i)},\ D^x=\bigcup_{i=k+1}^{n} x^{(i)},\ x\in id.$$

$$U/C=\{C_1,C_2,\ldots,C_m\},\ U/C^x=\{C_{x1},C_{x2},\ldots,C_{xt_x}\},\ U/D=\{D_1,D_2,\ldots,D_k\},\ U/D^x=\{D_{x1},D_{x2},\ldots,D_{xh_x}\},$$

$C_s \in U/C$, $C_{xi} \in U/C^x$, $C_s \subseteq C_{xi}$, $D_{xj} \in U/D^x$, *s=1..m, i=1..*t_x*, j=1..*h_x*. Khi đó, nếu* C_s *(f(*C_s*,a)=z) được làm mịn thành hai lớp tương đương điều kiện mới* C_p *và* C_q *(f(*C_p*,a)=w, f(*C_q*,a)=y và* trên lát cắt r_x, C_{xi} được làm mịn cảm sinh một phần *thành hai lớp tương đương điều kiện mới* $C_{xi'}$ *và* $C_{xi''}$ *thì:*

i) $C_{xi} = C_{xi'} \cup C_{xi''}$,

ii) $\forall D_{xj} \in U/D^x$: $Sup(C_{xi},D_{xj}) = Sup(C_{xi'},D_{xj}) + Sup(C_{xi''},D_{xj})$, *với j=1,2,…,*$h_x$.

Chứng minh

i) Do cách làm mịn của lớp tương đương điều kiện C_{xi} ta thấy rằng: $C_{xi} = C_{xi'} \cup C_{xi''}$.

ii) Theo giả thiết ta có: C_{xi} được làm mịn cảm sinh một phần *thành hai lớp tương đương điều kiện mới* $C_{xi'}$ *và* $C_{xi''}$

$\Rightarrow C_{xi} = C_{xi'} \cup C_{xi''}$ và $C_{xi'} \cap C_{xi''} = \varnothing$..

Mặt khác: $\forall D_{xj} \in U/D^x$: $Sup(C_{xi},D_{xj}) = |C_{xi} \cap D_{xj}| = |(C_{xi'} \cup C_{xi''}) \cap D_{xj}| = |(C_{xi'} \cap D_{xj}) \cup (C_{xi''} \cap D_{xj})|$.

Ta có: $C_{xi'} \cap C_{xi''} = \varnothing \Rightarrow (C_{xi'} \cap D_{xj}) \cap (C_{xi''} \cap D_{xj}) = \varnothing$.

Suy ra: $Sup(C_{xi},D_{xj}) = |(C_{xi'} \cap D_{xj}) \cup (C_{xi''} \cap D_{xj})| = |(C_{xi'} \cap D_{xj})| + |(C_{xi''} \cap D_{xj})| = Sup(C_{xi'},D_{xj}) + Sup(C_{xi''},D_{xj})$.

Vậy ta suy ra: $\forall D_{xj} \in U/D^x$: $Sup(C_{xi},D_{xj}) = Sup(C_{xi'},D_{xj}) + Sup(C_{xi''},D_{xj})$, *với j=1,2,…,*$h_x$.

Từ kết quả này ta thấy: *dòng tương ứng với lớp tương đương điều kiện* C_{xi} *trong ma trận độ hỗ trợ đối với lát cắt* r_x *sẽ được tách thành hai dòng mới tương ứng với hai lớp tương đương điều kiện mới* $C_{xi'}$ *và* $C_{xi''}$.

Do đó, để tính giá trị của các phần tử của hai dòng mới này trong ma trận độ hỗ trợ với lát cắt r_x thì đầu tiên ta tính các giá trị $Sup(C_{xi}, D_{xj})$ với j=1,2,…,h_x. Sau đó, ta suy ra các giá trị $Sup(C_{xi''}, D_{xj})$ là hiệu giữa $Sup(C_{xi}, D_{xj})$ và $Sup(C_{xi'}, D_{xj})$ với j=1,2,…,h_x.

Mệnh đề II.3

Cho khối quyết định DB = (U, C∪D, V, f), *a=*$x^{(i)} \in C$, V_a *là tập các giá trị hiện có của thuộc tính chỉ số điều kiện a, các giá trị w và y của a được làm thô thành giá trị mới z.*

$$C=\bigcup_{i=1,x\in id}^{k} x^{(i)},\ D=\bigcup_{i=k+1,x\in id}^{n} x^{(i)},\ \text{và } C^x=\bigcup_{i=1}^{k} x^{(i)},\ D^x=\bigcup_{i=k+1}^{n} x^{(i)},\ x\in id.$$

U/C={$C_1,C_2,\ldots,C_m$}, U/C^x= $\{C_{x1}, C_{x2},\ldots,C_{xt_x}\}$, *U/D={$D_1,D_2,\ldots,D_k$}, U/$D^x$=* $\{D_{x1}, D_{x2},\ldots,D_{xh_x}\}$,

Khi đó, nếu hai lớp tương đương điều kiện $C_p,C_q \in U/C$, ($f(C_p,a)=w$, $f(C_q,a)=y$) nào đó được làm thô thành lớp tương đương điều kiện mới $C_s \in U/C$ ($f(C_s,a)=z$) thì trên lát cắt r_x tồn tại hai lớp tương đương điều kiện C_{xi}, C_{xj} thỏa mãn: $C_p \subseteq C_{xi}$, $C_q \subseteq C_{xj}$, cũng được làm thô thành lớp tương đương điều kiện mới C_{xk} sao cho: $C_s \subseteq C_{xk}$.

Ta nói trên lát cắt r_x thì *hai lớp tương đương điều kiện C_{xi}, C_{xj}* được làm thô cảm sinh thành C_{xk} bởi sự làm thô của *hai lớp tương đương điều kiện C_p,C_q thành* C_s.

Chứng minh

Theo giả thiết ta có: *C_p, $C_q \in U/C$, ($f(C_p,a)=w$, $f(C_q,a)=y$)*, áp dụng kết quả của mệnh đề I.1 ta suy ra trên lát cắt *r_x tồn tại hai lớp tương đương điều kiện C_{xi}, C_{xj} thỏa mãn: $C_p \subseteq C_{xi}$, $C_q \subseteq C_{xj}$*. Từ đó ta có: *$f(u,a)=w$ với $u \in C_p \subseteq C_{xi} \Rightarrow f(C_{xi},a)=w$, tương tự ta cũng có: $f(u,a)=y$ với $u \in C_q \subseteq C_{xj} \Rightarrow f(C_{xj},a)=y$.*

Mặt khác ta lại có theo giả thiết: *hai lớp tương đương điều kiện $C_p,C_q \in U/C$ được làm thô thành lớp tương đương điều kiện mới $C_s \in U/C$*, do đó theo kết quả của định lý I.1 thì ta có:

$\forall a_j \neq a,\ a_j \in C$: $f(C_p,a_j) = f(C_q,a_j) \Rightarrow \forall a_j \neq a,\ a_j \in C^x$: $f(C_p,a_j) = f(C_q,a_j)$ (1)

Trong lát cắt r_x thì từ $C_p \subseteq C_{xi} \in U/C^x \Rightarrow \forall a_j \neq a,\ a_j \in C^x$: $f(C_p,a_j) = f(C_{xi},a_j)$ (2)

Tương tự, ta có: $C_q \subseteq C_{xj} \in U/C^x \Rightarrow \forall a_j \neq a,\ a_j \in C^x$: $f(C_q,a_j) = f(C_{xj},a_j)$ (3)

Từ (1), (2) và (3) ta suy ra: $\forall a_j \neq a,\ a_j \in C^x$: $f(C_{xi},a_j) = f(C_{xj},a_j)$.

Do vậy, áp dụng điều kiện cần và đủ trong phát biểu của định lý I.1 ta có *hai lớp tương đương điều kiện C_{xi}, C_{xj}* được làm thô cảm sinh thành C_{xk} bởi sự làm thô của *hai lớp tương đương điều kiện C_p,C_q thành* C_s.

Từ tính chất của việc làm thô *hai lớp tương đương điều kiện C_{xi}, C_{xj}* thành C_{xk} ta có:

$C_{xk} = (C_{xi} \cup C_{xj}) \supseteq (C_p \cup C_q) = C_s$.

Do đó: $C_s \subseteq C_{xk}$.

Mệnh đề II.4

Cho khối quyết định DB = (U, C∪D), $a=x^{(i)} \in C$, V_a là tập các giá trị hiện có của thuộc tính chỉ số điều kiện a, các giá trị w và y của a được làm thô thành giá trị mới z.

$$C=\bigcup_{i=1,x\in id}^{k} x^{(i)},\ D=\bigcup_{i=k+1,x\in id}^{n} x^{(i)},\ \text{và } C^x=\bigcup_{i=1}^{k} x^{(i)},\ D^x=\bigcup_{i=k+1}^{n} x^{(i)},\ x\in id.$$

U/C={$C_1,C_2,\ldots,C_m$}, U/C^x= $\{C_{x1}, C_{x2},\ldots,C_{xt_x}\}$, *U/D={$D_1,D_2,\ldots,D_k$}, U/$D^x$=* $\{D_{x1}, D_{x2},\ldots,D_{xh_x}\}$,

C_p, $C_q \in U/C$, ($f(C_p,a)=w$, $f(C_q,a)=y$), $D_{xh} \in U/D^x$, $h=1..h_x$. Khi đó, nếu C_p, C_q được làm thô thành lớp tương đương điều kiện mới C_s, ($f(C_s,a)=z$) và trên lát cắt r_x hai lớp tương đương điều kiện C_{xi}, C_{xj} ($C_p \subseteq C_{xi}$, $C_q \subseteq C_{xj}$) được làm thô cảm sinh thành *C_{xk} thì:*

i) $C_{xi} \cup C_{xj} = C_{xk}$

ii) $\forall D_{xh} \in U/D^x$: $Sup(C_{xi},D_{xh})+Sup(C_{xj},D_{xh}) = Sup(C_{xk},D_{xh})$, với $h=1,2,\ldots,h_x$.

Chứng minh

i) Giả sử ta có: $x \in C_{xi} \cup C_{xj} \Rightarrow x \in C_{xi}$ hoặc $x \in C_{xj}$. Nếu $x \in C_{xi}$ thì do hai lớp tương đương C_{xi}, C_{xj} được làm thô thành lớp tương đương $C_{xk} \Rightarrow f(x,a) = f(C_{xi},a)=f(C_{xk},a)=z$.

Mặt khác, áp dụng kết quả của định lý 2.1 ta có $\forall a_j \neq a$: $f(C_{xi},a_j) = f(C_{xj},a_j)= f(C_{xk},a_j) \Rightarrow f(x,a_j) = f(C_{xi},a_j) = f(C_{xj},a_j)= f(C_{xk},a_j) \Rightarrow x \in C_{xk}$. Hoàn toàn tương tự, khi $x \in C_{xj}$ ta cũng chứng minh được $x \in C_{xk}$.

Vậy suy ra: $(C_{xi} \cup C_{xj}) \subseteq C_{xk}$. (5)

Ngược lại, giả sử $x \in C_{xk}$, vì C_{xi} và C_{xj} được làm thô thành C_{xk} nên áp dụng kết quả của định lý 2.1 ta có: $\forall a_j \neq a$: $f(C_{xi},a_j)= f(C_{xj},a_j)= f(C_{xk},a_j) \Rightarrow f(x,a_j) =f(C_{xi},a_j) = f(C_{xj},a_j)$. Mặt khác, do $x \in C_{xk} \Rightarrow f(x,a)=z$ mà z được làm thô từ w và y $\Rightarrow f(x,a)=w$ hoặc $f(x,a)=y$.

- Nếu $f(x,a)=w \Rightarrow f(x,a)=f(C_{xi},a)= w \Rightarrow x \in C_{xi}$.
- Nếu $f(x,a)=y \Rightarrow f(x,a)=f(C_{xj},a)= y \Rightarrow x \in C_{xj}$.

Vậy $x \in C_{xi}$ hoặc $x \in C_{xj} \Rightarrow x \in C_{xi} \cup C_{xj}$.

Do đó, từ $x \in C_{xk} \Rightarrow x \in C_{xi} \cup C_{xj}$. Vậy: $C_{xk} \subseteq (C_{xi} \cup C_{xj})$ (6)

Kết hợp (5) và (6) ta có: $C_{xi} \cup C_{xj} = C_{xk}$.

ii) Vì C_{xi}, C_{xj} là các lớp tương đương điều kiện, nên ta có: $C_{xi} \cap C_{xj} = \emptyset$.

Mặt khác: $\forall D_{xh} \in U/D^x$: $Sup(C_{xk},D_{xh}) = |C_{xk} \cap D_{xh}| = |(C_{xi} \cup C_{xj}) \cap D_{xh}| = |(C_{xi} \cap D_{xh}) \cup (C_{xj} \cap D_{xh})|$.

Ta có: $C_{xi} \cap C_{xj} = \emptyset \Rightarrow (C_{xi} \cap D_{xh}) \cap (C_{xj} \cap D_{xh}) = \emptyset$.

Suy ra: $Sup(C_{xk},D_{xh}) = |(C_{xi} \cap D_{xh}) \cup (C_{xj} \cap D_{xh})| = |(C_{xi} \cap D_{xh})| + |(C_{xj} \cap D_{xh})| = Sup(C_{xi},D_{xh}) + Sup(C_{xj},D_{xh})$.

Vậy suy ra: $\forall D_{xh} \in U/D^x$: $Sup(C_{xi},D_{xh}) = Sup(C_{xi},D_{xh}) + Sup(C_{xj},D_{xh})$ *với* $h=1,2,\ldots,h_x$.

Như vậy, ta thấy hai dòng của ma trận độ hỗ trợ trên lát cắt r_x tương ứng với hai lớp *tương đương điều kiện* C_{xi}, C_{xj} được kết hợp thành một dòng mới tương ứng với lớp *tương đương điều kiện* C_{xk}. Giá trị mỗi phần tử của dòng mới tương ứng với C_{xk} là tổng giá trị hai phần tử của hai dòng tương ứng với C_{xi} và C_{xj}.

2. Làm mịn, thô các lớp tương đương quyết định trên khối và trên lát cắt.

Mệnh đề II.5

Cho khối quyết định DB = (U, C∪D, V, f), $a=x^{(i)} \in D$, *V_a là tập các giá trị hiện có của thuộc tính chỉ số quyết định a, giá trị z của a được làm mịn thành hai giá trị mới w và y.*

$$C=\bigcup_{i=1,x\in id}^{k} x^{(i)},\ D=\bigcup_{i=k+1,x\in id}^{n} x^{(i)},\ \text{và } C^x=\bigcup_{i=1}^{k} x^{(i)},\ D^x=\bigcup_{i=k+1}^{n} x^{(i)},\ x\in id.$$

$U/C=\{C_1,C_2,\ldots,C_m\}$, $U/C^x=\{C_{x1},C_{x2},\ldots,C_{xt_x}\}$, $U/D=\{D_1,D_2,\ldots,D_k\}$, $U/D^x=\{D_{x1},D_{x2},\ldots,D_{xh_x}\}$,

Khi đó, nếu lớp tương đương quyết định $D_s \in U/D$ ($f(D_s,a)=z$) được làm mịn thành hai lớp tương đương quyết định mới D_p,D_q ($f(D_p,a)=w$, $f(D_q,a)=y$, với $w,y \notin V_a$) nào đó thì trên lát cắt r_x, tồn tại lớp tương đương D_{xi} thỏa mãn: $D_s \subseteq D_{xi}$, cũng được làm mịn thành hai lớp tương đương quyết định mới $D_{xi'}$ và $D_{xi''}$ sao cho: $D_p \subseteq D_{xi'}$, $D_q \subseteq D_{xi''}$ ($f(D_{xi'},a)=w$, $f(D_{xi''},a)=y$).

Ta nói trên lát cắt r_x thì *lớp tương đương quyết định* D_{xi} được làm mịn cảm sinh một phần *thành hai lớp tương đương quyết định mới $D_{xi'}$ và $D_{xi''}$* bởi sự làm mịn của D_s *thành hai lớp tương đương quyết định mới D_p,D_q.*

Việc chứng minh mệnh đề này tương tự như chứng minh của mệnh đề II.1.

Mệnh đề II.6

Cho khối quyết định DB = (U, C∪D), $a=x^{(i)} \in D$, V_a là tập các giá trị hiện có của thuộc tính chỉ số quyết định a, giá trị z của a được làm mịn thành hai giá trị mới w và y.

$$C=\bigcup_{i=1,x\in id}^{k} x^{(i)},\ D=\bigcup_{i=k+1,x\in id}^{n} x^{(i)},\ \text{và } C^x=\bigcup_{i=1}^{k} x^{(i)},\ D^x=\bigcup_{i=k+1}^{n} x^{(i)},\ x\in id.$$

$U/C=\{C_1,C_2,\ldots,C_m\}$, $U/C^x=\{C_{x1},C_{x2},\ldots,C_{xt_x}\}$, $U/D=\{D_1,D_2,\ldots,D_k\}$, $U/D^x=\{D_{x1},D_{x2},\ldots,D_{xh_x}\}$,

$D_s \in U/D$, $D_{xi} \in U/D^x$, $D_s \subseteq D_{xi}$, $C_{xj} \in U/C^x$, $s=1..k$, $i=1..h_x$, $j=1..t_x$. Khi đó, nếu D_s ($f(D_s,a)=z$) được làm mịn thành hai lớp tương đương quyết định mới D_p, D_q ($f(D_p,a)=w$, $f(D_q,a)=y$ và trên lát cắt r_x, D_{xi} được làm mịn cảm sinh một phần *thành hai lớp tương đương quyết định mới $D_{xi'}$ và $D_{xi''}$ thì:*

i) $D_{xi} = D_{xi'} \cup D_{xi''}$,

ii) $\forall C_{xj} \in U/C^x$: $Sup(C_{xj},D_{xi}) = Sup(C_{xj},D_{xi'}) + Sup(C_{xj},D_{xi''})$, *với* $j=1,2,\ldots,t_x$.

Chứng minh

i) Do cách làm mịn của lớp tương đương *quyết định* D_{xi} ta thấy rằng: $D_{xi} = D_{xi'} \cup D_{xi''}$.

ii) Theo giả thiết ta có: D_{xi} được làm mịn cảm sinh một phần *thành hai lớp tương đương quyết định mới $D_{xi'}$ và $D_{xi''}$*

$\Rightarrow D_{xi} = D_{xi'} \cup D_{xi''}$ và $D_{xi'} \cap D_{xi''} = \emptyset$..

Mặt khác: $\forall C_{xj} \in U/C^x$: $Sup(C_{xj},D_{xi}) = |C_{xj} \cap D_{xi}| = |C_{xj} \cap (D_{xi'} \cup D_{xi''})| = |(C_{xj} \cap D_{xi'}) \cup (C_{xj} \cap D_{xi''})|$.

Ta có: $D_{xi'} \cap D_{xi''} = \varnothing \Rightarrow (C_{xj} \cap D_{xi'}) \cap (C_{xj} \cap D_{xi''}) = \varnothing$.

Suy ra: $Sup(C_{xj},D_{xi}) = |(C_{xj} \cap D_{xi'}) \cup (C_{xj} \cap D_{xi''})| = |(C_{xj} \cap D_{xi'})| + |(C_{xj} \cap D_{xi''})| = Sup(C_{xj},D_{xi'}) + Sup(C_{xj},D_{xi''})$.

Vậy ta suy ra: $\forall C_{xj} \in U/C^x$: $Sup(C_{xj},D_{xi}) = Sup(C_{xj},D_{xi'}) + Sup(C_{xj},D_{xi''})$, *với* $j=1,2,\dots,t_x$.

Từ kết quả này ta thấy: cột tương ứng với lớp tương đương quyết định D_{xi} trong ma trận độ hỗ trợ đối với lát cắt r_x sẽ được tách thành hai cột mới tương ứng với hai lớp tương đương quyết định mới $D_{xi'}$ và $D_{xi''}$.

Do đó, để tính giá trị của các phần tử của hai cột mới này trong ma trận độ hỗ trợ đối với lát cắt r_x thì đầu tiên ta tính các giá trị $Sup(C_{xj}, D_{xi'})$ với $j=1,2,\dots,t_x$. Sau đó, ta suy ra các giá trị $Sup(C_{xj}, D_{xi''})$ là hiệu giữa $Sup(C_{xj}, D_{xi})$ và $Sup(C_{xj}, D_{xi'})$ với $j=1,2,\dots,t_x$.

Mệnh đề II.7

Cho khối quyết định DB = (U, C∪D, V, f), $a=x^{(i)} \in D$, V_a *là tập các giá trị hiện có của thuộc tính chỉ số quyết định a, các giá trị w và y của a được làm thô thành giá trị mới z.*

$$C=\bigcup_{i=1,x\in id}^{k} x^{(i)},\ D=\bigcup_{i=k+1,x\in id}^{n} x^{(i)},\ \text{và } C^x=\bigcup_{i=1}^{k} x^{(i)},\ D^x=\bigcup_{i=k+1}^{n} x^{(i)},\ x\in id.$$

$$U/C=\{C_1,C_2,\dots,C_m\},\ U/C^x=\{C_{x1},C_{x2},\dots,C_{xt_x}\},\ U/D=\{D_1,D_2,\dots,D_k\},\ U/D^x=\{D_{x1},D_{x2},\dots,D_{xh_x}\},$$

Khi đó, nếu hai lớp tương đương quyết định D_p,D_q, $(f(D_p,a)=w, f(D_q,a)=y)$ *nào đó được làm thô thành lớp tương đương quyết định mới* $D_s \in U/D$ $(f(D_s,a)=z)$ *thì trên lát cắt* r_x *tồn tại hai lớp tương đương quyết định* D_{xi}, D_{xj} *thỏa mãn:* $D_p \subseteq D_{xi}$, $D_q \subseteq D_{xj}$, *cũng được làm thô thành lớp tương đương quyết định mới* D_{xk} *sao cho:* $D_s \subseteq D_{xk}$.

Ta nói trên lát cắt r_x thì *hai lớp tương đương quyết định* D_{xi}, D_{xj} được làm thô cảm sinh thành D_{xk} bởi sự làm thô của *hai lớp tương đương quyết định* D_p,D_q *thành lớp tương đương quyết định* D_s.

Việc chứng minh mệnh đề này tương tự như chứng minh của mệnh đề II.3.

Mệnh đề II.8

Cho khối quyết định DB = (U, C∪D), $a=x^{(i)} \in D$, V_a *là tập các giá trị hiện có của thuộc tính chỉ số quyết định a, các giá trị w và y của a được làm thô thành giá trị mới z.*

$$C=\bigcup_{i=1,x\in id}^{k} x^{(i)},\ D=\bigcup_{i=k+1,x\in id}^{n} x^{(i)},\ \text{và } C^x=\bigcup_{i=1}^{k} x^{(i)},\ D^x=\bigcup_{i=k+1}^{n} x^{(i)},\ x\in id.$$

$$U/C=\{C_1,C_2,\dots,C_m\},\ U/C^x=\{C_{x1},C_{x2},\dots,C_{xt_x}\},\ U/D=\{D_1,D_2,\dots,D_k\},\ U/D^x=\{D_{x1},D_{x2},\dots,D_{xh_x}\},$$

D_p, $D_q \in U/D$, $(f(D_p,a)=w, f(D_q,a)=y)$, $C_{xh} \in U/C^x$, $h=1..t_x$. *Khi đó, nếu* D_p, D_q *được làm thô thành lớp tương đương quyết định mới* D_s, $(f(D_s,a)=z)$ *và trên lát cắt* r_x *hai lớp tương đương quyết định* D_{xi}, D_{xj} $(D_p \subseteq D_{xi}, D_q \subseteq D_{xj})$ *được làm thô cảm sinh thành* D_{xk} *thì:*

i) $D_{xi} \cup D_{xj} = D_{xk}$

ii) $\forall C_{xh} \in U/C^x$: $Sup(C_{xh},D_{xi})+Sup(C_{xh},D_{xj}) = Sup(C_{xh},D_{xk})$, *với* $h=1,2,\dots,t_x$.

Chứng minh

i) Giả sử ta có: $u \in D_{xi} \cup D_{xj} \Rightarrow u \in D_{xi}$ hoặc $u \in D_{xj}$. Nếu $u \in D_{xi}$ thì do hai lớp tương đương D_{xi}, D_{xj} được làm thô thành lớp tương đương $D_{xk} \Rightarrow f(u,a) = f(D_{xi},a)=f(D_{xk},a)=z$.

Mặt khác, áp dụng kết quả của định lý 2.1 ta có $\forall a_r \neq a$: $f(D_{xi},a_r) = f(D_{xj},a_r) = f(D_{xk},a_r) \Rightarrow f(u,a_r) = f(D_{xi},a_r) = f(D_{xj},a_r) = f(D_{xk},a_r) \Rightarrow u \in D_{xk}$. Hoàn toàn tương tự, khi $u \in D_{xj}$ ta cũng chứng minh được $u \in D_{xk}$.

Vậy suy ra: $(D_{xi} \cup D_{xj}) \subseteq D_{xk}$. (7)

Ngược lại, giả sử $u \in D_{xk}$, vì D_{xi} và D_{xj} được làm thô thành D_{xk} nên áp dụng kết quả của định lý 2.1 ta có: $\forall a_r \neq a$: $f(D_{xi},a_r)= f(D_{xj},a_r)= f(D_{xk},a_r) \Rightarrow f(u,a_r) = f(D_{xi},a_r) = f(D_{xj},a_r)$. Mặt khác, do $u \in D_{xk} \Rightarrow f(u,a)=z$ mà z được làm thô từ w và y $\Rightarrow f(u,a)=w$ hoặc $f(u,a)=y$.

- Nếu $f(u,a)=w \Rightarrow f(u,a)=f(D_{xi},a)= w \Rightarrow u \in D_{xi}$.
- Nếu $f(u,a)=y \Rightarrow f(u,a)=f(D_{xj},a)= y \Rightarrow u \in D_{xj}$.

Vậy $u \in D_{xi}$ hoặc $u \in D_{xj} \Rightarrow u \in D_{xi} \cup D_{xj}$.

Do đó, từ $u \in D_{xk} \Rightarrow u \in D_{xi} \cup D_{xj}$. Vậy: $D_{xk} \subseteq (D_{xi} \cup D_{xj})$ (8)

Kết hợp (7) và (8) ta có: $D_{xi} \cup D_{xj} = D_{xk}$.

ii) Vì D_{xi}, D_{xj} là các lớp tương đương *quyết định*, nên ta có: $D_{xi} \cap D_{xj} = \emptyset$.

Mặt khác: $\forall C_{xh} \in U/C^x$: $Sup(C_{xh},D_{xk}) = |C_{xh} \cap D_{xk}| = |(D_{xi} \cup D_{xj}) \cap C_{xh}| = |(D_{xi} \cap C_{xh}) \cup (D_{xj} \cap C_{xh})|$.

Ta có: $D_{xi} \cap D_{xj} = \emptyset \Rightarrow (D_{xi} \cap C_{xh}) \cap (D_{xj} \cap C_{xh}) = \emptyset$.

Suy ra: $Sup(C_{xh},D_{xk}) = |(C_{xh} \cap D_{xi}) \cup (C_{xh} \cap D_{xj})| = |(C_{xh} \cap D_{xi})| + |(C_{xh} \cap D_{xj})| = Sup(C_{xh},D_{xi}) + Sup(C_{xh},D_{xj})$.

Vậy suy ra: $\forall C_{xh} \in U/C^x$: $Sup(C_{xh},D_{xi}) + Sup(C_{xh},D_{xj}) = Sup(C_{xh},D_{xk})$, *với* $h=1,2,\ldots,t_x$.

Như vậy, ta thấy hai cột của ma trận độ hỗ trợ trên lát cắt r_x tương ứng với hai lớp *tương đương quyết định* D_{xi}, D_{xj} được làm thô cảm sinh thành một cột mới tương ứng với lớp *tương đương quyết định* D_{xk}. Giá trị mỗi phần tử của cột mới tương ứng với D_{xk} là tổng giá trị hai phần tử của hai cột tương ứng với hai lớp *tương đương quyết định* D_{xi} và D_{xj}.

III. KẾT LUẬN

Từ các kết quả ban đầu về khối quyết định, bài báo đã đề xuất và chứng minh một số kết quả về mối quan hệ giữa việc làm thô, làm mịn các giá trị của thuộc tính điều kiện hoặc quyết định đối với các lớp tương đương điều kiện hoặc quyết định trên khối quyết định và trên lát cắt. Việc làm mịn lớp tương đương điều kiện hoặc quyết định trên khối đã cảm sinh một phần việc làm mịn lớp tương đương điều kiện hoặc quyết định tương ứng trên lát cắt. Việc làm thô hai lớp tương đương điều kiện hoặc quyết định trên khối đã cảm sinh việc làm thô hai lớp tương đương điều kiện hoặc quyết định tương ứng trên lát cắt. Nhờ các kết quả này mà việc tính ma trận độ hỗ trợ trên lát cắt cũng được xác định tương tự như việc tính ma trận độ hỗ trợ trên khối khi làm thô hoặc mịn các giá trị của thuộc tính điều kiện hoặc quyết định.

Trong trường hợp tập chỉ số id = {x}, khối thông tin suy biến thành hệ thông tin thì các kết quả này lại trùng với các kết quả đã được nhiều tác giả đưa ra đối với hệ thông tin. Trên cơ sở của các kết quả này ta có thể nghiên cứu tiếp mối quan hệ ngược lại giữa các lát cắt của khối thông tin với chính khối đó, trường hợp các đối tượng của khối thông tin bị thay đổi..., một số kết quả khác có thể được xét trong các trường hợp riêng của khối thông tin..., góp phần bổ sung thêm các kết quả về mặt lí thuyết của việc khai phá luật quyết định trên các khối thông tin.

TÀI LIỆU THAM KHẢO

[1] Nguyễn Xuân Huy, Trịnh Đình Thắng, *Mô hình cơ sở dữ liệu dạng khối,* Tạp chí Tin học và Điều khiển học, T.14, S.3 (52-60), 1998.

[2] Trịnh Đình Thắng, Trần Minh Tuyến, *Ánh xạ đóng và phép dịch chuyển lược đồ khối*, Kỷ yếu Hội nghị quốc gia lần thứ VI về Nghiên cứu cơ bản và ứng dụng Công nghệ Thông tin (FAIR), (174-179), Thừa Thiên-Huế 20-21/6/2013.

[3] Trịnh Đình Thắng, Trần Minh Tuyến, *Phép dịch chuyển lược đồ khối và vấn đề biểu diễn bao đóng, khóa trong mô hình dữ liệu dạng khối*, Kỷ yếu Hội thảo quốc gia lần thứ XIII "Một số vấn đề chọn lọc của Công nghệ Thông tin và Truyền thông", (276-286), Hưng Yên, 19-20/08/2010.

[4] Trịnh Đình Thắng, Trần Minh Tuyến, Trịnh Ngọc Trúc, *Phụ thuộc boolean dương đa trị trong mô hình dữ liệu dạng khối,* Kỷ yếu Hội nghị quốc gia lần thứ IX FAIR, Nghiên cứu cơ bản và ứng dụng Công nghệ Thông tin, (602-609), Cần Thơ 04-05/08/2016.

[5] Trịnh Đình Thắng, Trần Minh Tuyến, Đỗ Thị Lan Anh, *Khai phá luật quyết định trên khối dữ liệu có giá trị thuộc tính thay đổi*, Kỷ yếu Hội thảo quốc gia lần thứ XIX "Một số vấn đề chọn lọc của Công nghệ Thông tin và Truyền thông", (163-169), Hà Nội 01-02/10/2016.

SOME RESULTS ON THE RECLAIM OF DECISION LAW ON THE DATA BLOCK HAS VARIABLE ATTRIBUTE VALUES

Trinh Dinh Thang, Tran Minh Tuyen, Do Thi Lan Anh, Nguyen Thi Quyen

ABSTRACT: *The report proposes and demonstrates some of the properties of smoothing, roughening the values of a conditional index attribute or decision index attribute for its information block and slice. Whenever the conditional or decisions equivalence classes on the decision blocks are smoothed or roughened, they will partially feel or feel smoothing, roughing the corresponding layers on the slice. From the results of smoothing, roughening, the conditional or decisions equivalence classes to be partially feel or feel on the slice then the computation of the support matrices on the slice will be as simple as compiling those matrices on the decision block when smoothing, roughening the values of the conditional or decision index attributes.*