*Tạp chí Tin học và Điều khiển học, T.17, S.2 (2001), 20–26*

# THUẬT TOÁN LÀM MỊN TẬP LUẬT VÀ XÂY DỰNG HỆ LUẬT CHÍNH QUI CỦA HỆ CHUYÊN GIA

LÊ HẢI KHÔI

**Abstract.** In this paper we give some algorithms for refining the rules set and building the regular rule-based system of the expert system.

**Tóm tắt.** Bài báo cung cấp một số thuật toán liên quan đến việc làm mịn tập luật và xây dựng hệ luật chính qui của hệ chuyên gia. Tính đúng đắn của thuật toán được chứng minh chặt chẽ dưới góc độ toán học.

## 1. MỞ ĐẦU

Như chúng ta đều biết, trong năm thành phần chính của hệ chuyên gia thì cơ sở tri thức và mô tơ suy diễn đóng vai trò quan trọng nhất. Vì thế, người ta còn nói rằng "Hệ chuyên gia = Cơ sở tri thức + Mô tơ suy diễn". Cơ sở tri thức được biểu diễn bằng nhiều phương pháp: phương pháp lôgic, phương pháp mạng ngữ nghĩa, phương pháp mô hình, phương pháp hệ luật, phương pháp thông qua khung, phương pháp bộ ba OAV (đối tượng - thuộc tính - giá trị ), v.v.. Trong số các phương pháp này thì phương pháp biểu diễn bằng hệ luật là tương đối phổ biến, nhờ một số ưu điểm như: tính trực quan, tính mở, khả năng kiểm tra và xử lý mâu thuẫn cũng như dư thừa,... Vì thế, các bài toán liên quan đến hệ luật được nhiều người quan tâm. Độc giả có thể tìm trong [1, 2, 4, 5] những kiến thức cơ sở về hệ chuyên gia cũng như các phương pháp biểu diễn tri thức.

Ở một bài báo trước [3], khi đề cập các vấn đề liên quan đến việc biểu diễn tri thức bằng hệ luật, chúng tôi đã trình bày thuật toán tìm bao đóng của tập sự kiện và các thuật toán về loại bỏ luật dư thừa của tập luật cũng như sự dư thừa của hệ luật. Một câu hỏi tự nhiên đặt ra là: có thể nói gì về bản thân các luật? Cụ thể hơn, liệu trong nội tại của từng luật trong tập luật có sự dư thừa nào không, và nếu có thì làm thế nào để loại bỏ dư thừa đi? Trong bài báo này chúng tôi nghiên cứu vấn đề đó.

Cấu trúc của bài báo như sau. Trong mục 2 chúng tôi nêu lại các thuật toán về tìm bao đóng của tập sự kiện và loại bỏ luật dư thừa của tập luật mà đã được trình bày trong bài [3], bởi vì chúng cần thiết cho việc xây dựng thuật toán tổng hợp sau này. Ngoài ra, để các thuật toán đó thực sự có ý nghĩa, chúng tôi đưa ra ví dụ minh họa cho các thuật toán. Mục 3 trình bày thuật toán vét cạn tập luật không dư thừa, như là hệ quả của thuật toán loại bỏ luật dư thừa của tập luật. Mục 4 đề cập thuật toán mịn hóa một luật cũng như mịn hóa tập luật. Trong mục 5 chúng tôi đưa ra khái niệm hệ luật chính qui và trên cơ sở tổng hợp các thuật toán trước đó, trình bày thuật toán xây dựng hệ luật chính qui từ một hệ luật bất kỳ cho trước.

## 2. CÁC KHÁI NIỆM CƠ BẢN

Người ta dùng hệ luật bao gồm các câu "NẾU ..., THÌ ..." để biểu diễn tri thức theo cấu trúc sau:

$$\begin{cases} \text{NẾU } \langle\text{điều kiện 1}\rangle, \langle\text{điều kiện 2}\rangle, \ldots, \langle\text{điều kiện m}\rangle, \\ \text{THÌ } \langle\text{kết luận 1}\rangle, \langle\text{kết luận 2}\rangle, \ldots, \langle\text{kết luận n}\rangle. \end{cases}$$

Hệ luật này còn có tên gọi là hệ luật dạng 1 (khác với hệ luật dạng 2 là hệ luật mà ở đó trong phần "THÌ" các "kết luận" được thay bằng các "thực hiện"). Trong hệ luật trên các điều kiện và kết luận được thể hiện tương đối tự do.

Chúng ta có thể hình thức hóa cao hơn để thể hiện toàn bộ tri thức trong một hệ luật. Cụ thể như sau.

**Định nghĩa 2.1.** Hệ luật, kí hiệu là $L = \langle F, R \rangle$, gồm hai thành phần $F = \{f_1, \dots, f_p\}$ là tập các sự kiện, $R = \{r_1, \dots, r_q\}$ là tập các luật.

Theo các qui tắc biến đổi của Vương Hạo, luôn có thể coi rằng hệ luật chỉ gồm các luật với vế trái toàn phép "và" ($\wedge$) và vế phải có đúng một sự kiện, tức là hệ luật chỉ bao gồm các luật dạng

$$P_1 \wedge P_2 \wedge \dots \wedge P_n \to Q,$$

ở đây $P_1, P_2, \dots, P_n$ và $Q$ là các sự kiện. Để đơn giản, chúng ta thay dấu $\wedge$ trong vế trái bằng dấu phẩy (,), khi đó luật được viết dưới dạng

$$P_1, P_2, \dots, P_n \to Q.$$

Đối với luật $r$ chúng ta kí hiệu $Left(r)$ là tập các sự kiện ở vế trái, $Right(r)$ là sự kiện ở vế phải của luật.

Giả sử có hệ luật $L = \langle F, R \rangle$, trong đó $F = \{f_1, \dots, f_p\}$ là tập các sự kiện, $R = \{r_1, \dots, r_q\}$ là tập các luật. Kí hiệu $F^*$ là tập các sự kiện $f \in F$ thỏa mãn đồng thời hai điều kiện:

(i) $f$ có mặt ở vế trái,

(ii) $f$ không có mặt ở vế phải,

trong tất cả các luật thuộc $R$. Tập $F^*$ này được gọi là *tập các sự kiện gốc*.

Dưới đây nêu lại thuật toán tìm bao đóng của tập sự kiện và loại bỏ luật dư thừa của tập luật. Những thuật toán này sẽ được sử dụng trong quá trình xây dựng những thuật toán khác ở các mục tiếp theo. Chúng ta sử dụng ký hiệu $\langle\ .\ , \dots ,\ .\ \rangle$ để chỉ dãy (tức là có thứ tự) các phần tử.

Nếu $F' \subseteq F$, thì *bao đóng* của $F'$ đối với $R$, kí hiệu $(F'_R)^+$, được định nghĩa là tập thu được từ $F'$ sau khi áp dụng tất cả các luật có thể có của $R$. Chúng ta luôn giả thiết là các phép suy diễn không bị lặp (tức là không có chu trình).

**Thuật toán 2.2.** (tính $(F'_R)^+$)

Input: $L = \langle F, R \rangle$ với $F = \langle f_1, \dots, f_p \rangle$, $R = \langle r_1, \dots, r_q \rangle$ và $F' \subseteq F$.

Output: $(F'_R)^+$.

- Bước 0: đặt $K_0 = F'$;

- Bước $i$: nếu có luật $r \in R$ thỏa mãn điều kiện $Left(r) \subseteq K_{i-1}$ và $Right(r) \notin K_{i-1}$, thì đặt $K_i = K_{i-1} \cup Right(r)$.

- Quá trình được lặp lại cho đến khi $K_i = K_{i+1}$.

Lúc đó đặt $(F'_R)^+ = K_i$.

**Mệnh đề 2.3.** *Thuật toán 2.2 có độ phức tạp là đa thức theo lực lượng của F và R.*

**Ví dụ 2.4.** (minh họa Thuật toán 2.2)

Xét hệ luật $L = \langle F, R \rangle$, trong đó $F = \{A, B, C, D, E, F, G, H, I, J, K\}$, $R = \langle r_1, \dots, r_5 \rangle$, với $r_1 = AB \to C$, $r_2 = CD \to E$, $r_3 = EF \to G$, $r_4 = DH \to I$, $r_5 = IJ \to K$ và $F' = \{A, B, D, H\}$. Khi đó $F^* = \{A, B, D, F, H, J\}$ và $F' \subset F^*$.

Áp dụng thuật toán, chúng ta có:

- Bước 0: $K_0 = F' = \{A, B, D, H\}$.

- Bước 1: luật $r_1$ cho thêm sự kiện $C \notin K_0 = F'$, nên ta có $K_1 = \{A, B, C, D, H\}$.

- Bước 2: luật $r_2$ cho thêm sự kiện $E \notin K_1$, nên ta có $K_2 = \{A, B, C, D, E, H\}$.

- Bước 3: luật $r_4$ cho thêm sự kiện $I \notin K_2$, nên ta có $K_3 = \{A, B, C, D, E, H, I\}$.

- Bước 4: do không có luật nào nữa mà cho thêm sự kiện mới không thuộc $K_3$, nên $K_4 = K_3$.

Vậy, $\left(F'_R\right)^+ = K_3 = \{A, B, C, D, E, H, I\}$.

Bây giờ chúng ta chuyển sang vấn đề luật dư thừa. Với $F^*$ là tập các sự kiện gốc của hệ luật $L = \langle F, R\rangle$, nếu có $r \in R$ sao cho $\left(F^*_R\right)^+ = \left(F^*_{R\setminus\{r\}}\right)^+$, thì luật $r$ được coi là *luật thừa* và về nguyên tắc chúng ta có thể loại bỏ luật này đi.

**Thuật toán 2.5.** (loại bỏ luật thừa)

Input: $L = \langle F, R\rangle$ với $F = \langle f_1, \ldots, f_p\rangle$ và $R = \langle r_1, \ldots, r_q\rangle$.

Output: $R'$ thỏa mãn $R' \subseteq R$, $\left(F^*_{R'}\right)^+ = \left(F^*_R\right)^+$ và $\forall r \in R' : R'' = R' \setminus \{r\}$ luôn có $\left(F^*_{R''}\right)^+ \neq \left(F^*_{R'}\right)^+$.

- Bước 0: Đặt $K_0 = R$, tính $\left(F^*_R\right)^+$.
- Bước $i$ $(1 \leq i \leq q-1)$:

$$K_i = \begin{cases} K_{i-1} \setminus \{r_i\}, & \text{nếu } \left(F^*_{K_{i-1}\setminus\{r_i\}}\right)^+ = \left(F^*_R\right)^+, \\ K_{i-1}, & \text{nếu ngược lại.} \end{cases}$$

- Bước $q$:

Nếu $K_{q-1}$ chỉ còn $r_q$, thì đặt $K_q = K_{q-1}$.

Nếu $K_{q-1}$ chứa không chỉ có $r_q$, thì đặt

$$K_q = \begin{cases} K_{q-1} \setminus \{r_q\}, & \text{nếu } \left(F^*_{K_{q-1}\setminus\{r_q\}}\right)^+ = \left(F^*_R\right)^+, \\ K_{q-1}, & \text{nếu ngược lại.} \end{cases}$$

- Bước $q+1$: Đặt $R' = K_q$.

**Mệnh đề 2.6.** *Thuật toán 2.5 có độ phức tạp là đa thức theo lực lượng của $F$ và $R$.*

**Ví dụ 2.7.** (minh họa Thuật toán 2.5)

Xét hệ luật $L = \langle F, R\rangle$, trong đó $F = \{A, B, C, D, E, F, G, H, I, J, K\}$, $R = \langle r_1, \ldots, r_6\rangle$, với $r_1 = AB \to C$, $r_2 = CD \to E$, $r_3 = EF \to G$, $r_4 = DH \to I$, $r_5 = IJ \to K$, $r_6 = CE \to I$. Khi đó $F^* = \{A, B, D, F, H, J\}$.

Áp dụng thuật toán, chúng ta có:

- Bước 0: Đặt $K_0 = R$, khi đó $\left(F^*_R\right)^+ = \{A, B, C, D, E, F, G, H, I, J, K\}$.
- Bước 1: do $\left(F^*_{K_0\setminus\{r_1\}}\right)^+ = \{A, B, D, F, H, I, J, K\} \neq \left(F^*_R\right)^+$, nên $K_1 = K_0$.
- Bước 2: do $\left(F^*_{K_1\setminus\{r_2\}}\right)^+ = \left(F^*_{K_0\setminus\{r_2\}}\right)^+ = \{A, B, C, D, F, H, I, J, K\} \neq \left(F^*_R\right)^+$, nên $K_2 = K_1$.
- Bước 3: do $\left(F^*_{K_2\setminus\{r_3\}}\right)^+ = \left(F^*_{K_0\setminus\{r_3\}}\right)^+ = \{A, B, C, D, E, F, H, I, J, K\} \neq \left(F^*_R\right)^+$, nên $K_3 = K_2$.
- Bước 4: do $\left(F^*_{K_3\setminus\{r_4\}}\right)^+ = \left(F^*_{K_0\setminus\{r_4\}}\right)^+ = \{A, B, C, D, E, F, G, H, I, J, K\} = \left(F^*_R\right)^+$, nên $K_4 = K_3 \setminus \{r_4\} = \langle r_1, r_2, r_3, r_5, r_6\rangle$.
- Bước 5: do $\left(F^*_{K_4\setminus\{r_5\}}\right)^+ = \{A, B, C, D, E, F, G, H, I, J\} \neq \left(F^*_R\right)^+$, nên $K_5 = K_4$.
- Bước 6: do $\left(F^*_{K_5\setminus\{r_6\}}\right)^+ = \{A, B, C, D, E, F, G, H, J\} \neq \left(F^*_R\right)^+$, nên $K_6 = K_5$.
- Bước 7: Chúng ta được $R' = K_6 = \langle r_1, r_2, r_3, r_5, r_6\rangle$ và luật $r_4$ là luật thừa.

## 3. TẬP LUẬT KHÔNG DƯ THỪA

Trong mục này chúng ta xem xét khả năng vét cạn tất cả các luật không dư thừa trong tập luật $R$.

Để làm điều này, ta xếp các luật thành một dãy $R = \langle r_1, \ldots, r_q\rangle$, rồi sử dụng thuật toán loại bỏ luật dư thừa 2.5 cho tất cả các hoán vị của dãy $R$. Nói cách khác, với mỗi hoán vị của dãy này, chúng ta áp dụng Thuật toán 2.5 để loại luật dư thừa đi. Do tập luật $R = \{r_1, \ldots, r_q\}$ có $q$ phần tử, nên dãy $(r_1, \ldots, r_q)$ có $q!$ hoán vị khác nhau tương ứng với các hoán vị của dãy $(1, \ldots, q)$. Ký hiệu các dãy luật có được từ tập luật $R$ là $R_1, R_2, \ldots, R_{q!}$.

**Thuật toán 3.1.** (vét cạn tập luật không dư thừa)

Input: $L = \langle F, R\rangle$ với $F = \{f_1, \ldots, f_p\}$ và $R = \{r_1, \ldots, r_q\}$, $R_1, R_2, \ldots, R_{q!}$.

Output: $R'_1, R'_2, \ldots, R'_{q!}$ thỏa mãn với mọi $i = 1, 2, \ldots, q!$ luôn có các điều kiện sau:

- $R'_i \subseteq R_i$,
- $(F^*_{R'_i})^+ = (F^*_{R_i})^+$,
- $\forall r \in R'_i : R''_i = R'_i \setminus \{r\}$ luôn có $(F^*_{R''_i})^+ \neq (F^*_{R'_i})^+$.
- Bước 1: thực hiện Thuật toán 2.5 đối với $L_1 = \langle F, R_1\rangle$, được $L'_1 = \langle F, R'_1\rangle$.
- Bước 2: thực hiện Thuật toán 2.5 đối với $L_2 = \langle F, R_2\rangle$, được $L'_2 = \langle F, R'_2\rangle$.
- ......
- Bước $q!$: thực hiện Thuật toán 2.5 đối với $L_{q!} = \langle F, R_{q!}\rangle$, được $L'_{q!} = \langle F, R'_{q!}\rangle$.

**Định lý 3.2.** *Thuật toán 3.1 là dừng và cho tất cả các tập luật không dư thừa có thể có được từ tập luật* $R = \{r_1, \ldots, r_q\}$.

*Chứng minh.* Chúng ta phải chứng minh rằng mọi dãy luật không dư thừa của $R$ đều sinh ra từ dãy nào đó trong số các dãy $R_1, R_2, \ldots, R_{q!}$ qua thuật toán nêu trên.

Thật vậy, lấy dãy luật không dư thừa bất kỳ $R' = \langle r_{i_1}, r_{i_2}, \ldots, r_{i_s}\rangle$. Dãy này tương ứng với dãy chỉ số $(i_1, i_2, \ldots, i_s)$.

Có hai khả năng xảy ra:

- Trường hợp tầm thường khi $s = q$, vì lúc đó $(i_1, i_2, \ldots, i_s)$ là một hoán vị của $(1, 2, \ldots, q)$, và bản thân $R$ là tập luật không dư thừa.

- Trường hợp $s < q$. Xét dãy chỉ số $(j_1, j_2, \ldots, j_{q-s}, i_1, i_2, \ldots, i_s)$ sao cho dãy này là một hoán vị nào đó của dãy $(1, 2, \ldots, q)$. Khi đó, rõ ràng rằng dãy luật $\langle r_{j_1}, r_{j_2}, \ldots, r_{j_{q-s}}, r_{i_1}, r_{i_2}, \ldots, r_{i_s}\rangle$ trùng với một trong các $R_1, R_2, \ldots, R_{q!}$, chẳng hạn, $R_t$ nào đó. Thuật toán trên khi áp dụng cho dãy $R_t$ này sẽ cho $R'_t$. Theo cách làm của thuật toán loại bỏ luật thừa 2.5, chúng ta có $R'_t = \langle r_{i_1}, r_{i_2}, \ldots, r_{i_s}\rangle$.

**Nhận xét 3.3.** 1) Theo công thức Stirling với những $n$ lớn

$$n! = \sqrt{2\pi n}\; n^n\; e^{-n+\frac{\theta_n}{12n}}\; (0 < \theta_n < 1),$$

do đó thuật toán nêu trên sẽ có độ phức tạp lớn. Vì thế nó chỉ thực sự có ý nghĩa với những $n$ nhỏ.

2) Có thể xảy ra trường hợp là các dãy luật không dư thừa nhận được tuy khác nhau, nhưng nếu xét từ góc độ tập hợp thì có những tập hợp trùng nhau. Do đó, chúng ta phải so sánh các phần tử (theo quan điểm tập hợp) của các $R'_i$ $(1 \le i \le q!)$ để loại đi những dãy có các phần tử như nhau, sẽ được tất cả các tập luật không dư thừa từ $R$.

## 4. MỊN HÓA TẬP LUẬT

Trong mục này chúng ta nghiên cứu việc làm mịn tập luật $R$ đã được xếp theo thứ tự thành dãy luật, cụ thể là $R = \langle r_1, r_2, \ldots, r_q\rangle$.

Để thuận tiện cho việc trình bày thuật toán, ta quay lại cú pháp với phép "và" $(\wedge)$ trong mô tả một luật, cụ thể với luật $r : P_1, P_2, \ldots, P_t \to Q$ chúng ta viết tắt dưới dạng $r = \wedge_i P_i \to Right(r)$, trong đó $P_1, P_2, \ldots, P_t$ và $Q = Right(r)$ là các sự kiện. Từ luật $r$ này sau khi bỏ đi sự kiện $P_i$ sẽ được luật mới, ký hiệu $r'_i$, có dạng $r'_i = P_1 \wedge \ldots \wedge P_{i-1} \wedge P_{i+1} \wedge \ldots \wedge P_t \to Right(r'_i)$. Có thể chấp nhận $Right(r'_i) = Q$, tuy nhiên trong việc loại bỏ sự kiện $P_i$ cần cân nhắc đến ngữ nghĩa của sự kiện này, đến trọng số (nếu có) của sự kiện đó trong vế trái của luật $r$. Về mặt suy diễn lôgic thì có thể bỏ sự kiện $P_i$ trong vế trái của $r$ đi, xong ý nghĩa của sự kiện này trong luật $r$ trên thực tế cần phải được xem xét rất thận trọng. Như vậy, việc mịn hóa một luật cho chúng ta khả năng bớt đi các sự kiện dư thừa về mặt suy diễn lôgic để luật được gọn hơn.

Việc làm mịn một luật hiểu theo nghĩa sau: đối với mỗi luật $r \in R$, $r = \wedge_i P_i \to Right(r)$, kiểm tra xem liệu có thể loại bỏ một số sự kiện nào đó trong số các sự kiện $P_1, P_2, \ldots, P_t$ sao cho bao đóng của $F^*$ trong tập luật mới không thay đổi (điều đó có nghĩa là việc loại bỏ một sự kiện nào đó cũng không được làm thay đổi tập $F^*$). Những sự kiện đó, nếu có, coi như là thừa. Nói cách khác, chúng ta gọi luật $r = \wedge_i P_i \to Right(r)$ với $P = \langle P_1, P_2, \ldots, P_t \rangle$ là *luật mịn*, nếu như $\left(F^*_{R\setminus\{r\}\cup\{r'_i\}}\right)^+ \neq \left(F^*_R\right)^+$, $\forall i = 1, 2, \ldots, t$. Việc xây dựng luật mịn từ một luật cho trước được gọi là *mịn hóa* luật đó. Để cho gọn chúng ta viết $\wedge_{p\in P}$ thay cho $\wedge_i P_i$.

**Thuật toán 4.1.** (mịn hóa một luật)

Input: $L = \langle F, R \rangle$, $r \in R$ và $r = P_1 \wedge P_2 \wedge \ldots \wedge P_t \to Right(r)$ với dãy các sự kiện ở vế phải $P = \langle P_1, P_2, \ldots, P_t \rangle$.

Output: $r' = \wedge_{p\in P'} \to Right(r')$ thỏa mãn $P' \subseteq P$, $\left(F^*_{R\setminus\{r\}\cup\{r'\}}\right)^+ = \left(F^*_R\right)^+$ và $\forall p \in P' : P'' = P' \setminus \{p\}$, $r'' = \wedge_{p\in P''} \to Right(r'')$, luôn có $\left(F^*_{R\setminus\{r\}\cup\{r''\}}\right)^+ \neq \left(F^*_R\right)^+$.

- Bước 0: đặt $K_0 = P$.

- Bước $i$ $(1 \leq i \leq t-1)$:

$$K_i = \begin{cases} K_{i-1} \setminus \{P_i\}, & \text{nếu } \left(F^*_{R\setminus\{r\}\cup\{r'=\wedge_{p\in K_{i-1}\setminus\{P_i\}}\to Righ.(r')\}}\right)^+ = \left(F^*_R\right)^+, \\ K_{i-1}, & \text{nếu ngược lại.} \end{cases}$$

- Bước $t$:

Nếu $K_{t-1}$ chỉ còn $P_t$, thì đặt $K_t = K_{t-1} = \{P_t\}$.

Nếu $K_{t-1}$ chứa không chỉ $P_t$, thì đặt

$$K_t = \begin{cases} K_{t-1} \setminus \{P_t\}, & \text{nếu } \left(F^*_{R\setminus\{r\}\cup\{r'=\wedge_{p\in K_{t-1}\setminus\{P_t\}}\to Right(r')\}}\right)^+ = \left(F^*_R\right)^+, \\ K_{t-1}, & \text{nếu ngược lại.} \end{cases}$$

- Bước $t+1$: Đặt $P' = K_t$ và $r' = \wedge_{p\in P'} \to Right(r')$.

**Định lý 4.2.** *Thuật toán 4.1 là dừng và cho kết quả là luật $r' = \wedge_{p\in P'} \to Right(r')$ mịn.*

*Chứng minh.* Chúng ta chứng minh bằng phương pháp phản chứng.

Lưu ý rằng để có được $K_{t-1}$ chúng ta đã kiểm tra tính dư thừa của $t-1$ sự kiện ở vế phải của luật $r$ là $P_1, \ldots, P_{t-1}$, do đó, như thuật toán đã chỉ rõ, có thể xảy ra hai khả năng sau:

- Khả năng thứ nhất: $K_{t-1}$ chỉ chứa một phần tử. Thế thì phần tử này chính là $P_t$ và do đó $K_{t-1}$ không thể giảm đi được nữa. Vậy thì $K_t = K_{t-1}$, tức là $P' = K_t = \{P_t\}$ và $r' = \wedge_{p\in P'} \to Right(r')$ là luật không dư thừa.

- Khả năng thứ hai: $K_{t-1}$ có ít nhất hai phần tử. Thế thì ngoài $P_t$ ra, $K_{t-1}$ còn chứa thêm ít nhất một phần tử nữa. Khi đó, theo thuật toán chúng ta có:

$$K_t = \begin{cases} K_{t-1} \setminus \{P_t\}, & \text{nếu } \left(F^*_{R\setminus\{r\}\cup\{r'=\wedge_{p\in K_{t-1}\setminus\{P_t\}}\to Right(r')\}}\right)^+ = \left(F^*_R\right)^+, \\ K_{t-1}\ , & \text{nếu ngược lại.} \end{cases}$$

Giả sử rằng $K_t$ chưa phải là tối ưu, tức là $P' \subset K_t$ nhưng $P' \neq K_t$. Điều đó có nghĩa là trong $K_t$ vẫn còn sự kiện thừa, nói cách khác, $\exists p \in K_t$ sao cho với $P'' = K_t \setminus \{p\}$ thì $\left(F^*_{R\setminus\{r\}\cup\{r''=\wedge_{p\in P''}\to Right(r'')\}}\right)^+ = \left(F^*_R\right)^+$.

Xét từng trường hợp đối với $K_t$:

(1) $K_t = K_{t-1} \setminus \{P_t\}$: thế thì mọi sự kiện trong tập $P$ đã được kiểm tra hết, điều này mâu thuẫn với việc trong $K_t$ ngoài $P_t$ ra vẫn còn ít nhất một sự kiện nào đó chưa kiểm tra.

(2) $K_t = K_{t-1}$: trong trường hợp này, theo thuật toán thì

$$\left(F^*_{R\setminus\{r\}\cup\{r'=\wedge_{p\in K_{t-1}\setminus\{P_t\}}\to Right(r')\}}\right)^+ \neq \left(F^*_R\right)^+,$$

tức là $P_t$ không phải là sự kiện thừa và như vậy tất cả các sự kiện thuộc $P$ đã được kiểm tra. Điều này lại mâu thuẫn với việc trong $K_t$ vẫn còn sự kiện chưa kiểm tra.

Như vậy điều giả thiết rằng $P' \subset K_t$ là sai. Vậy, $r' = \wedge_{p \in P'} \rightarrow Right(r')$ là luật mịn có được từ luật $r$. Thuật toán được chứng minh.

Dễ dàng chứng minh kết quả sau đây.

**Mệnh đề 4.3.** *Thuật toán 4.1 có độ phức tạp là đa thức theo lực lượng của tập sự kiện ở vế phải của luật đó.*

**Ví dụ 4.4.** (minh họa Thuật toán 4.1)

Xét hệ luật $L = \langle F, R \rangle$, trong đó $F = \{P_1, \ldots, P_6, Q_1, Q_2, Q_3\}$, $R = \langle r_1, \ldots, r_5 \rangle$, với $r_1 = P_1P_2P_3 \rightarrow Q_1$, $r_2 = P_2 \rightarrow Q_2$, $r_3 = P_4P_6 \rightarrow P_5$, $r_4 = P_1P_4P_6 \rightarrow Q_3$, $r_5 = P_3P_4P_5 \rightarrow Q_1$. Khi đó $F^* = \{P_1, P_2, P_3, P_4, P_6\}$.

Áp dụng Thuật toán 4.1 đối với luật $r_1$, chúng ta có:

- Bước 0: đặt $K_0 = P = \langle P_1, P_2, P_3 \rangle$.

- Bước 1: Xét $K_0 \setminus \{P_1\}$, khi đó luật $r_1' = \wedge_{p \in K_0 \setminus \{P_1\}} \rightarrow Right(r_1')$; do trong trường hợp này $\left(F^*_{R \setminus \{r_1\} \cup \{r_1'\}}\right)^+ = \left(F^*_R\right)^+$, nên được $K_1 = K_0 \setminus \{P_1\} = \langle P_2, P_3 \rangle$.

- Bước 2: Xét $K_1 \setminus \{P_2\}$, khi đó luật $r_1'' = \wedge_{p \in K_1 \setminus \{P_2\}} \rightarrow Right(r_1'')$; do trong trường hợp này $\left(F^*_{R \setminus \{r_1\} \cup \{r_1''\}}\right)^+ = \left(F^*_R\right)^+$, nên được $K_2 = K_1 \setminus \{P_2\} = \langle P_3 \rangle$.

- Bước 3: Vì trong $P$ chỉ còn sự kiện $P_3$, nên đặt $K_3 = K_2 = \langle P_3 \rangle$.

Như vậy, luật $r_1$ đã được mịn hóa bởi luật $r_1'' = P_3 \rightarrow Q_1$.

Bây giờ chúng ta chuyển sang việc mịn hóa các luật của tập luật $R$ trong hệ luật.

**Thuật toán 4.5.** (mịn hóa tập luật)

Input: $L = \langle F, R \rangle$, $R = \langle r_1, r_2, \ldots, r_q \rangle$.

Output: $R' = \langle r_1', r_2', \ldots, r_q' \rangle$ thỏa mãn $r_i'$ là mịn hóa của $r_i$ với mọi $i = 1, 2, \ldots, q$.

- Bước 1: thực hiện Thuật toán 4.1 cho luật $r_1$ được $r_1'$.

- Bước 2: thực hiện Thuật toán 4.1 cho luật $r_2$ được $r_2'$.

- ..........

- Bước q: thực hiện Thuật toán 4.1 cho luật $r_q$ được $r_q'$.

Dễ dàng chứng minh kết quả sau đây.

**Mệnh đề 4.6.** *Thuật toán 4.5 có độ phức tạp đa thức.*

## 5. HỆ LUẬT CHÍNH QUI

Cho hệ luật $L = \langle F, R \rangle$. Chúng ta nói rằng hệ luật $L$ là *chính qui*, nếu $L$ thỏa mãn các điều kiện sau:

(i) không có sự kiện dư thừa trong tập các sự kiện $F$,

(ii) không có luật dư thừa trong tập luật $R$,

(iii) không có sự kiện dư thừa trong các luật của $R$.

Tổng hợp các thuật toán đã trình bày chúng ta có thuật toán sau đây về việc xây dựng hệ luật chính qui từ một hệ luật bất kỳ cho trước.

**Thuật toán 5.1.** (xây dựng hệ luật chính qui)

Input: $L = \langle F, R \rangle$ với $F = \langle f_1, \ldots, f_p \rangle$, $R = \langle r_1, \ldots, r_q \rangle$.

Output: $L''' = \langle F', R'' \rangle$ là hệ luật chính qui có được từ hệ luật $L$.

- Thực hiện thuật toán loại bỏ luật dư thừa để loại các luật không cần thiết trong $R$: từ $L = \langle F, R\rangle$ có $L' = \langle F, R'\rangle$, trong đó $R'$ là tập luật không dư thừa.

- Thực hiện thuật toán tính bao đóng để loại bỏ các sự kiện dư thừa trong $F$, được hệ luật không dư thừa: $L'' = \langle F', R'\rangle$, với $F' = \left(F^*_{R'}\right)^+$.

- Thực hiện thuật toán mịn hóa của tập luật để loại các sự kiện không cần thiết trong tất cả các luật của $R'$, được $R''$ đã được mịn hóa.

- Hệ luật $L''' = \langle F', R''\rangle$ là hệ luật chính qui cần xây dựng.

Dễ dàng chứng minh kết quả sau đây.

**Mệnh đề 5.2.** *Thuật toán 5.1 có độ phức tạp là đa thức.*

**Lời cảm ơn.** Tác giả xin chân thành cảm ơn PGS TS Vũ Đức Thi đã đóng góp những ý kiến quí báu trong quá trình hoàn thành bài báo này. Tác giả cũng xin cảm ơn người phản biện về những nhận xét góp ý cho bài báo được tốt hơn.

## TÀI LIỆU THAM KHẢO

[1] Bạch Hưng Khang, Hoàng Kiếm, *Trí tuệ nhân tạo: các phương pháp và ứng dụng*, NXB Khoa học và Kỹ thuật, 1989.

[2] Durkin J., *Expert Systems*, Prentice Hall, 1994.

[3] Lê Hải Khôi, Thuật toán tìm bao đóng của tập sự kiện và loại bỏ luật dư thừa của tập luật trong hệ luật của hệ chuyên gia, *Tạp chí Tin học và Điều khiển học* **16** (4) (2000) 79–84.

[4] Sundermeyer K., *Knowledge Based Systems*, Wissenschafts Verlag, 1991.

[5] Turban E., *Decisions Support & Expert Systems - Management Support Systems*, Prentice Hall, 1998.

*Nhận bài ngày 26 tháng 10 năm 2000*
*Nhận bài sau khi sửa ngày 25 tháng 4 năm 2001*

*Viện Công nghệ thông tin*