TẠP CHÍ KHOA HỌC TÍNH TOÁN VÀ ĐIỀU KHIỂN TẬP IV, (1988) Số 2 (13–16)

HỆ CHUYÊN GIA THEO MÔ HÌNH PHÂN CẤP CƠ SỞ TRI THỨC

LÊ HỒ KHÁNH
Đại học kỹ thuật quân sự

Khi xây dựng một lý thuyết nào đó, thông thường chúng ta tiến hành theo các bước sau: Từ tập các khái niệm cơ bản T_0 (tập các tiên đề hoặc các định nghĩa), nhờ vào các quy tắc lập luận ta có được tập khái niệm mới T_1 dựa trên T_0. Từ T_0 và T_1, bằng các lập luận chúng ta rút được tập T_2 -- tập các khái niệm ở mức cao hơn. Quá trình xây dựng để hoàn thiện lý thuyết về cơ bản là tuân theo nguyên tắc trên. Dưới một góc độ nào đó, sự phát triển tư duy con người dường như cũng tuân thủ nguyên tắc này. Bản thân nội tại các KB có một cấu trúc trong tương tự như vậy không? Đây là vấn đề chúng ta sẽ giải quyết.

1. Mô hình phân cấp của cơ sở tri thức (CSTT) bao gồm các luật dẫn

1.1. Các khái niệm cơ bản

Xét một cơ sở tri thức biểu diễn theo luật dẫn:

$$\begin{cases} R_i: \bigwedge\limits_{j=1}^{n_i} a_{ij} \rightarrow \bigwedge\limits_{j=1}^{m_i} b_{ij} \\ i = \overline{1,K} \end{cases}$$

$R = \{ R_i \mid i = \overline{1, k} \}$ - tập các luật dẫn

$L(R_i) = \{ a_{ij} \mid j = \overline{1,n_i} \}$ tập các mệnh đề giả thiết của luật R_i

$R(R_i) = \{ b_{ij} \mid j = \overline{1,m_i} \}$ tập các mệnh đề kết luận của luật R_i

Chú ý rằng ở đây chúng ta không đòi hỏi có sự phân biệt nào giữa các mệnh đề giả thiết và mệnh đề kết luận giữa các luật. Trong CSTT có thể tồn tại những mệnh đề là giả thiết của luật này lại là kết luận của luật khác.

Ký hiệu $F = \bigcup\limits_{i=1}^{k} (L(R_i) \cup R(R_i))$

tập các mệnh đề tham gia vào CSTT đang xét. Ta sẽ ký hiệu các phần tử của F là f.

Định nghĩa 1: Ta gọi một xích suy diễn trong CSTT là một dãy các luật $R_{i_1}, R_{i_2}, \ldots, R_{i_t}$ thỏa mãn các điều kiện sau:

a) $\{ i_1, i_2, \ldots, i_t \} \subset \{ 1, 2, \ldots, k \}$,

b) $R(R_{i_q}) \cap L(R_{i_{q+1}}) \neq \Phi$.

Định nghĩa 2: CSTT K được gọi là xoay vòng tại $f \in F$ nếu tồn tại một xích suy diễn trong K: $R_{i_1}, R_{i_2}, \ldots, R_{i_q}$ sao cho $f \in L(R_{i_1})$ và $f \in L(R_{i_q})$.

Định nghĩa 3: CSTT K được gọi là không xoay vòng nếu nó không xoay vòng tại $\forall f \in F$.

1.2. Mô hình phân cấp của CSTT không xoay vòng

Xét CSTT K cho như trong 1.1.

Xây dựng tập $A^{(0)}$ như sau:

$$A^{(o)} = \{ f \in F \mid \neg\exists\ R_i \text{ mà } f \in R(R_i)\ \forall i = \overline{1,k} \}$$

$A^{(o)}$ là tập các mệnh đề giả thiết thực sự của K. Đó là những giả thiết không được suy ra từ bất cứ giả thiết nào trong K.

Với những CSTT nào thì $A^{(o)} \neq \Phi$? Ta có mệnh đề sau:

Mệnh đề 1: Trong một CSTT không xoay vòng ta luôn có $A^{(o)} \neq \Phi$.

Chứng minh: Xét quan hệ $\rightarrow$ trên R như sau: $R_i \rightarrow R_j$ nếu tồn tại một xích suy diễn trong K: $R_{j1}, R_{j2}, \ldots, R_{jl}$ mà:

$$R_i = R_{j1} \text{ và } R_j = R_{jl}.$$

Dễ thấy rằng trong K không xoay vòng quan hệ $\rightarrow$ trên R là một thứ tự bộ phận.

Gọi $R_q \in R$ là một phần tử nhỏ nhất trong thứ tự bộ phận $\rightarrow$ trên R. Ta có:

$\forall R_p \in R \setminus \{R_q\} : R(R_p) \cap L(R_q) = \Phi$.

Vậy $L(R_q) \subset A^{(o)} \Rightarrow A^{(o)} \neq \Phi$.

Bằng truy hồi chúng ta xây dựng các tập T_p như sau:

$T_o = A^{(o)}$

$T_{p+1} = \{ f \in F \mid \exists R_i : L(R_i) \cap S_p \neq \Phi,\ L(R_i) \subset S'_p,\ f \in R(R_i) \}$

Trong đó:

$$S_p = T_p \setminus \bigcup_{h=0}^{p-1} T_h$$

$$S'_p = \bigcup_{h=0}^{p} T_h \qquad \forall p = 0,1,2,\ldots$$

Việc xây dựng các tập T_p dừng lại khi gặp $M \in N$ sao cho $T_{M+1} = \Phi$. Khi đó ta có ngay $T_{M+2} = T_{M+3} = \ldots = \Phi$. Ta dễ dàng chứng minh được mệnh đề sau:

Mệnh đề 2: Với K không xoay vòng luôn tồn tại một số tự nhiên M mà từ đó trở đi ta có:

$$T_{M+1} = T_{M+2} = \ldots = \Phi.$$

Phải chăng với $\forall f \in F$ luôn tồn tại $p \in \{0,1,\ldots,M\}$ mà $f \in T_p$? Nếu $F \in T_p$ thì chỉ số p có ý nghĩa gì? Ta có mệnh đề sau.

Mệnh đề 3: a) Với một CSTT K không xoay vòng với $\forall f \in F$ luôn tồn tại $p \in \{0,1,2,\ldots,M\}$ sao cho $f \in T_p$

b) Nếu $f \in T_p$ $(p \geq 1) \rightarrow$ tồn tại một xích suy diễn có độ dài p: $R_{i1}, R_{i2}, \ldots, R_{ip}$ thỏa mãn các điều kiện sau:

— $L(R_{i1}) \subset A^{(o)}$,

— $f \in R(R_{ip})$.

Chứng minh a) Giả sử f là phần tử tùy ý của F. Nếu $f \in A^{(o)} \Rightarrow f \in T_o$. Xét trường hợp $f \notin A^{(o)}$. Khi đó $\exists R_{i1} \in R$ mà $f \in R(R_i)$. Giả sử với $\forall p = 1,2,\ldots,M$ $f \notin T_p$. Theo định nghĩa của T_p điều này xảy ra khi $\exists f_1 \in L(R_i)$ mà $f_1 \notin T_p$ với $\forall p = 0,1,\ldots,M$.

Do $f \notin A^{(o)} \Rightarrow \exists R_{i1}$ mà $f_1 \in R(R_{j1})$ mà

$f_1 \notin T_p$ $\forall p = 1,2\ldots,M \Rightarrow \exists f_2 \in L(R_{i1})$ mà

$f_2 \notin T_p$ với $\forall p = 0,1,2,\ldots,M$.

Ta tiếp tục quá trình trên đến khi nào không làm được nữa thì dừng lại. Thực chất là ta xây dựng một xích suy diễn có phần tử phải nhất là R_i nhưng trong mỗi phần tử của xích đó tồn tại một mệnh đề giả thiết và một mệnh đề kết luận không được chứa trong bất kỳ T_p nào. Do

CSTT của chúng ta là không xoay vòng nên quá trình trên sẽ dừng lại sau hữu hạn bước. (Cực đại là bằng k), giả sử tại bước thứ N xét R_{i_N} tương ứng ta thấy xảy ra những sự kiện sau:

i) $\exists f_{i_{N+1}} \in L(R_{i_N})$ mà $f_{N+1} \overline{\in} T_p$

với $\forall p = 0,1,2,\ldots,M$

ii) $\neg \exists R_i \in R$ mà $R_i \rightarrow R_{i_N}$

Do ii) $\Rightarrow L(R_{i_N}) \subset A^{(o)}$ $(= T_0)$. Điều này mâu thuẫn với i). Vậy phần a) của mệnh đề đã được chứng minh.

b) Giả sử $f \in T_p$, từ định nghĩa của T_p ta có thể dựng được ngay một xích có độ dài p thỏa mãn yêu cầu nói trên.

Với $f \in F$ tùy ý, ta xác định tập $\mathcal{B}(f) \subset 2^{A^{(o)}}$ được gọi là tập các tập cơ sở của f như sau:

– Nếu $f \in T_0 \Rightarrow \mathcal{B}(f) = \{\{f\}\}$.

– Nếu $f \overline{\in} T_0 \Rightarrow \exists p : f \in T_p$.

Giả sử ta đã xây dựng cơ sở của mọi phần tử thuộc các tập $T_0, T_1, \ldots, T_{p-1}$.

Do $f \in T_p \Rightarrow \exists R_l \in R$ mà $f \in R(R_l)$

và $L(R_l) \subset \bigcup_{h=0}^{p-1} T_h$, ta định nghĩa:

$$\bigcup_{\substack{f' \in L(R_l) \\ B(f') \in \mathcal{B}(f')}} B(f') \in \mathcal{B}(f).$$

Quá trình xây dựng tập các tập cơ sở có thể tiến hành theo thủ tục sau: Xây dựng $\mathcal{B}(f)$ của $\forall f \in T_1$, sau đó xây dựng $\mathcal{B}(f)$ $\forall f \in T_2, \ldots$
Cuối cùng xây dựng $\mathcal{B}(f)$ với $\forall f \in T_M$.

Định nghĩa 4: Giả sử $f_1, f_2 \in F$. Ta nói rằng chúng thuộc cùng lớp với cơ sở $\mathcal{B}_k \in 2^{A^{(o)}}$ nếu các điều kiện sau đây thỏa mãn:

a) $\exists p\ (p = \overline{1, M})$ mà $f_1, f_2 \in T_p$

b) $\mathcal{B}_k \in \mathcal{B}(f_1) \cap \mathcal{B}(f_2)$

Đặt: $\mathcal{B} = \cup \mathcal{B}(f)$ – Tập tất cả các tập cơ sở phân biệt của K. $\mathcal{B}$ – hữu hạn.

$\forall f : \exists R_l$ mà $f \in R(R_l)$

Đánh số các phần tử của tập $\mathcal{B}$ là $\mathcal{B}_1, \mathcal{B}_2, \ldots, \mathcal{B}_D$

Chúng ta xây dựng ma trận quan hệ như sau:

$$\begin{array}{c|ccccc}
 & T_1 & T_2 \cdots & T_p & \cdots & T_M \\
\hline
\mathcal{B}_1 & & & \vdots & & \\
\mathcal{B}_2 & & & \vdots & & \\
\vdots & & & \vdots & & \\
\mathcal{B}_q & \cdots & \cdots & T_{pq} & & \\
\vdots & & & & & \\
\mathcal{B}_D & & & & &
\end{array}$$

Trong đó $T_{pq} \subset T_p$ gồm tất cả các phần tử của T_p cùng lớp với cơ sở $\mathcal{B}_q$.

Xét trên dòng thứ q ta thấy các phần tử của các tập $T_{1q}, T_{2q}, \ldots, T_{pq}, \ldots, T_{Mq}$ đều có chung một tập cơ sở là $\mathcal{B}_q$. Mối liên hệ này nói lên điều gì? Ta sẽ làm rõ thêm mối liên hệ này.

Định nghĩa 5: Cho $f \in F$ và $G \subset F$. Ta nói rằng G khẳng định f (hoặc G chứng minh F) nhờ CSTT đã cho nếu các điều kiện sau đây được thỏa mãn:

i) Tồn tại trong K xích suy diễn $R_{i_1}, R_{i_2}, \ldots, R_{i_t}$

ii) $f \in R_{i_t}$

iii) $L(R_{i_1}) \subset G$

iv) $L(R_{i_h}) \subset G \cup \left[\bigcup_{d=1}^{h-1} R(R_{i_d})\right] \quad \forall h = 2, 3, \ldots t.$

Định lý: Nếu $f \in T_{pq} \Rightarrow \mathcal{B}_q$ khẳng định f với $\forall p = \overline{1, M}$.

Chứng minh: Từ định nghĩa 5 ta có nhận xét sau:

Muốn chứng minh $f \in F$ được khẳng định bởi G ta cần chứng tỏ: $\exists R_l \in R$ mà $f \in R(R_l)$ và $\forall f \in L(R_l)$ đều được khẳng định bởi G.

Bây giờ ta chứng minh định lý trên bằng quy nạp theo p.

Với $p = 1 \Rightarrow f \in T_1$ và f có cơ sở là $\mathcal{B}_q$.

Vậy:

$$\left.\begin{array}{r} \exists R_l : f \in f(R_l) \\ L(R_l) \subset \mathcal{B}_q \end{array}\right\} \rightarrow \mathcal{B}_q \text{ khẳng định f theo định nghĩa}$$

Giả sử định lý đúng với $\forall n < p$. ta chứng minh định lý đúng với $n = p$. Thật vậy:

Vì $f \in T_{pq} \Rightarrow \exists$ xích suy diễn trong KB $R_{l_1}, R_{l_2}, \ldots, R_{l_t}$ thỏa:

– $L(R_{l_1}) \subset \mathcal{B}_q$ (do f có một tập cơ sở là $\mathcal{B}_q$)

– $L(R_{l_t}) \subset \bigcup_{h=0}^{p-1} T_{h_q} \quad (*)$

– $f \in R(R_{l_t})$,

Từ (*) ta suy ra: $\forall f \in L(R_{l_t})$ được khẳng định bởi $\mathcal{B}_q$ (theo giả thiết quy nạp). Do nhận xét ở trên ta suy ra f được khẳng định bởi $\mathcal{B}_q$. Định lý được chứng minh.

2. Một số nhận xét chung

– Với phương pháp đề cập ở trên, CSTT K đã được thể hiện thông qua cấu trúc trong của nó. Đó là ma trận liên hệ giữa các lớp trong T_p và các $\mathcal{B}_q \subset A^{(o)}$ là cơ sở của chúng. Dựa vào mối liên hệ này vấn đề lập luận trong K hầu như đã được giải quyết một cách có hiệu quả. Hơn nữa với ma trận liên hệ này chúng ta còn có hy vọng giải quyết được ở một mức độ nào đó các vấn đề lý thuyết khác đang được đặt ra với hệ cơ sở tri thức như tính nhất quán, độ phức tạp của hệ cơ sở tri thức v.v...

– Với mô hình phân cấp tri thức này chúng ta có thể khai thác các điểm mạnh của hệ quản trị cơ sở dữ liệu trong khi cài đặt hệ cơ sở tri thức.

TÀI LIỆU THAM KHẢO

1. Минский М., Фреймы для представления знаний. Знегия 1979.
2. Хант Э., Искусственный интеллект. Мир 1976.
3. Нильсон Н., Искусственный интеллект. Мир 1973.
4. Laurière J. L., « Representation et utilisation des connaissances », TSI V1, N° 1–2, 1982.