

BÀI 5. BIỂU DIỄN TRI THỨC BẰNG CÁC LUẬT SẢN XUẤT

Đặt vấn đề

Phương pháp biểu diễn tri thức nhờ logic (logic mệnh đề và logic vị từ) khá trực quan song chỉ phù hợp khi không có quá nhiều luật suy diễn. Biểu diễn tri thức bằng các luật sản xuất là phương pháp biểu diễn tri thức có cấu trúc

Ví dụ

Ta có các luật về kinh nghiệm dự báo thời tiết:

"Chuồn chuồn bay thấp thì mưa, bay cao thì nắng, bay vừa thì râm"

a: chuồn chuồn bay thấp

b: chuồn chuồn bay cao

c: chuồn chuồn bay vừa

d: trời mưa

e: trời nắng

f: trời râm

Có thể biểu diễn các luật như sau:

$$a \rightarrow d$$

$$b \rightarrow e$$

$$c \rightarrow f$$

1. Khái niệm

Đây là phương pháp biểu diễn tri thức có cấu trúc.

Ý tưởng là tri thức được cấu trúc hóa bằng cặp điều kiện - hành động (if- then)

nghĩa là nếu điều kiện xảy ra thì hành động được thi hành

- Phần điều kiện của quy luật xác định khi nào quy luật có thể được áp dụng với một vấn đề nêu ra
- Phần hành động định nghĩa giải quyết vấn đề liên quan

Luật sản xuất có dạng sau:

Nếu: Điều kiện 1 (giả thiết 1)

Điều kiện 2 (giả thiết 2)

.....

Điều kiện m (giả thiết m)

Thì: Hành động 1 (kết luận 1)

Hành động 2 (kết luận 2)

.....

Hành động n (kết luận n)

Ví dụ 1:

Các sự kiện:

$$F = \{a, b, c, d, e\}$$

Tập luật:

$$R = \{r_1, r_2, r_3\}$$

$$r_1: a \rightarrow d$$

$$r_2: a \wedge b \rightarrow c$$

$$r_3: d \wedge c \rightarrow e$$

Suy diễn là quá trình suy luận xuất phát từ một số sự kiện ban đầu, xác định các sự kiện có thể được "sinh" ra từ sự kiện này.

Ví dụ 2:

Tập sự kiện ban đầu: H, K

Tập các luật R (quy tắc):

$$(R_1): A \rightarrow E$$

$$(R_2): B \rightarrow D$$

$$(R_3): H \rightarrow A$$

$$(R_4): E \wedge G \rightarrow C$$

$$(R_5): E \wedge K \rightarrow B$$

$$(R_6): D \wedge E \wedge K \rightarrow C$$

$$(R_7): G \wedge K \wedge F \rightarrow A$$

Từ sự kiện ban đầu có thể sinh ra các sự kiện nào?

$$(R_3): H \rightarrow A \text{ thì suy ra tập sự kiện } \{A, H, K\}$$

$$(R_1): A \rightarrow E \text{ thì suy ra tập sự kiện } \{A, E, H, K\}$$

$$(R_5): E \wedge K \rightarrow B \text{ thì suy ra tập sự kiện } \{A, B, E, H, K\}$$

$$(R_2): B \rightarrow D \text{ thì suy ra tập sự kiện } \{A, B, D, E, H, K\}$$

$$(R_6): D \wedge E \wedge K \rightarrow C \text{ thì suy ra tập sự kiện } \{A, B, C, D, E, H, K\}$$

Vậy từ sự kiện ban đầu có thể sinh ra các sự kiện $\{A, B, C, D, E, H, K\}$

Ưu điểm của các hệ sản xuất:

- Hệ sản xuất đưa ra khung làm việc tổng quát cho việc thực hiện tìm kiếm chính bởi nhờ sự đơn giản của nó, tính có thể thay đổi được, và tính mềm dẻo trong việc áp dụng tri thức giải quyết vấn đề. Hệ sản xuất đã được chứng minh là một công cụ quan trọng trong việc xây dựng hệ chuyên gia và ứng dụng trí tuệ nhân tạo

- Ưu điểm chính của các hệ sản xuất cho trí tuệ nhân tạo đó là sự phân tách giữa tri thức và điều khiển. Hệ sản xuất là một mô hình của sự phân tách giữa tri thức và điều khiển trong một chương trình máy tính. Điều khiển được cung cấp bởi chu kỳ nhận thức- hành động của vòng lặp hệ sản xuất và tri thức giải quyết vấn đề được mã hóa trong chính các luật. Ưu điểm của sự chia tách này bao gồm sự thoải mái thay đổi cơ sở tri thức mà không yêu cầu một sự thay đổi trong mã hóa cho điều khiển chương trình, và một cách ngược lại, khả năng dễ thay đổi mã hóa cho điều khiển chương trình mà không cần thay đổi tập các luật sản xuất

Suy diễn là quá trình suy luận dựa vào các quy luật đã cho, thiết lập các thông tin mới từ các thông tin đã biết. Suy diễn sẽ sử dụng tập sự kiện làm tiên đề.

Các phương pháp suy diễn dần dần chuyển từ các giả thiết về các kết luận bằng cách thêm vào giả thiết những sự kiện đã được khẳng định đúng, dựa trên 2 phương thức:

- Modus ponens:
$$\frac{A, A \Rightarrow B}{B}$$

nghĩa là nếu A đúng và $A \Rightarrow B$ đúng thì B cũng đúng

- Modus tollens
$$\frac{\neg B, A \Rightarrow B}{\neg A}$$

nghĩa là nếu B sai và biết rằng $A \Rightarrow B$ đúng thì A cũng sai.

Trong quá trình suy diễn, ta cần quan tâm đến các vấn đề sau:

- Xây dựng tập luật, câu hỏi nào được chọn để người sử dụng trả lời
- Chọn quá trình tìm kiếm như thế nào
- Thông tin nhận được có ảnh hưởng đến quá trình tìm kiếm không

Bài toán suy diễn

Cho tập sự kiện $GT = \{gt_1, gt_2, \dots, gt_n\}$ và tập luật $R = \{r_1, r_2, \dots, r_m\}$. Chứng minh tập kết luận $KL = \{q_1, q_2, \dots, q_n\}$ đúng.

Các phương pháp suy diễn

Quá trình suy diễn trong hệ luật sản xuất bao gồm 2 phương pháp cơ bản: suy diễn tiến và suy diễn lùi.

2. Suy diễn tiến đối với logic mệnh đề

(Tư tưởng cơ bản của suy diễn tiến là áp dụng luật suy diễn Modus Ponens tổng quát)

Là quá trình suy diễn bắt đầu từ tập sự kiện đã biết, rút ra những sự kiện mới và cứ như vậy cho đến khi có được sự kiện cần chứng minh hoặc không có luật nào sinh ra các sự kiện mới (tập sự kiện đúng là cực đại).

- **Phương pháp**

Dữ liệu vào: Tập mệnh đề $G_t = \{g_{t1}, g_{t2} \dots g_{tm}\}$

Tập các luật $R = \{r_1, r_2, \dots r_m\}$

với $r_i: P_1 \wedge P_2 \wedge P_3 \wedge \dots \wedge P_n \rightarrow q_i, i=1..m$

Tập kết luận = $\{q_1, q_2, \dots q_k\}$

Dữ liệu ra: Thông báo thành công nếu mọi q_i đều được suy ra từ G_t và tập R

$G_t \rightarrow TG_1 \rightarrow TG_2 \rightarrow \dots \rightarrow TG_n \supset KL$

Sử dụng tập TG (trung gian) là tập các sự kiện đúng cho đến thời điểm đang xét

SAT là tập các luật có vế trái thuộc TG

(SAT là tập các luật có dạng $P_1 \wedge P_2 \wedge P_3 \wedge \dots \wedge P_n \rightarrow q$, sao cho $\forall p_i$ thuộc TG)

- **Giải thuật**

Void SDT()

{

$TG = G_t$

$SAT = Loc\{R, TG\}$

While $KL \not\subset TG$ and $SAT \neq \emptyset$

{

$r \leftarrow get(SAT)$ // lấy luật r trong SAT

$TG = TG \cup \{q\}$ // bổ sung vế phải vào TG

$R = R \setminus \{r\}$

$SAT = Loc(R, TG)$ // tính lại SAT

}

IF $KL \in TG$ dừng “Thành Công” Else “Không thành công”

}

Ví dụ 1

Tập luật

$R = \{r_1, r_2, r_3, r_4, r_5, r_6\}$

$GT = \{a, b\};$

$KL = \{g\}$

$r_1: a \rightarrow c$

$r_2: b \rightarrow d$

$r_3: a \rightarrow e$

$r_4: a \wedge d \rightarrow e$

$r_5: b \wedge c \rightarrow f$

$r_6: e \wedge f \rightarrow g$

Ban đầu $TG = GT = \{a, b\}; SAT = \{r_1, r_2, r_3\}$ // vì a, b thuộc TG

KL chưa thuộc TG và SAT khác Rỗng

Lấy r_1 áp dụng ta có $TG = \{a, b, c\}; R = \{r_2, r_3, r_4, r_5, r_6\}$

$SAT = \{r_2, r_3, r_5\}$ // vì b, c thuộc TG thêm r_5

KL chưa thuộc TG và SAT khác Rỗng

Lấy r_3 áp dụng $TG = \{a, b, c, e\}; R = \{r_2, r_4, r_5, r_6\}; SAT = \{r_2, r_5\}$

KL chưa thuộc TG và SAT khác Rỗng

Lấy r_2 áp dụng $TG = \{a, b, c, e, d\}; R = \{r_4, r_5, r_6\};$

$SAT = \{r_4, r_5\}$ // vì a, d thuộc TG

KL chưa thuộc TG và SAT khác Rỗng

Lấy r_5 áp dụng ta có $TG = \{a, b, c, d, e, f\}; R = \{r_4, r_6\}; SAT = \{r_4, r_6\}$

KL chưa thuộc TG và SAT khác Rỗng

Lấy r_6 áp dụng ta có $TG = \{a, b, c, d, e, f, g\}; R = \{r_4\}; SAT = \{r_4\}$

$KL = \{g\}$ thuộc TG Thông báo thành công

Ví dụ

Cho trước tập sự kiện $GT = \{a, b\}, KL = \{g\}$.

Sử dụng các luật:

$r_1: a \rightarrow c$

$r_2: b \rightarrow d$

$r_3: c \rightarrow e$

$r_4: a \wedge d \rightarrow e$

$r_5: b \wedge c \rightarrow f$

$r_6: e \wedge f \rightarrow g$

r	TG	SAT	R
r_1	a, b	r_1, r_2, r_3	$r_1, r_2, r_3, r_4, r_5, r_6$
r_2	a, b, c	r_2, r_3, r_5	r_2, \dots, r_6
r_3	a, b, c, d	r_3, r_4, r_5	r_3, \dots, r_6
r_4	a, b, c, d, e	r_4, r_5	r_4, r_5, r_6
r_5	a, b, c, d, e	r_5	r_5, r_6
r_6	a, b, c, d, e, f	r_6	r_6
r_6	a, b, c, d, e, f, g		

$g \in T$ nên bài toán được chứng minh (g: true)

Chú ý

- Quá trình suy diễn tiến là quá trình xem xét các luật, với mỗi luật ta xét phần điều kiện (ở vế trái) tới phần kết luận (ở vế phải) và khi mà tất cả các điều kiện của luật đều thỏa mãn thì ta suy ra sự kiện trong phần kết luận. Chính vì lẽ đó mà có tên là suy diễn tiến.
- Trong mỗi bước của thủ tục, người ta xét một luật trong tập luật. So sánh mỗi điều kiện (ở vế trái) của tập luật với các sự kiện trong cơ sở sự kiện, nếu tất cả các điều kiện của luật được thỏa mãn thì sự kiện trong phần kết luận được xem là sự kiện được suy ra. nếu sự kiện này là sự kiện mới (không có trong bộ nhớ làm việc) thì nó được đưa vào bộ nhớ làm việc. Quá trình trên cứ lặp lại cho đến khi nào không có luật nào sinh ra sự kiện mới.
- Quá trình suy diễn tiến không định hướng tới giải quyết một vấn đề nào cả, không hướng tới tìm ra câu trả lời cho một câu hỏi nào cả. Suy diễn tiến chỉ là quá trình suy ra các sự kiện mới từ các sự kiện có trong bộ nhớ làm việc.

Nhận xét về suy diễn tiến:

Ưu điểm:

- Làm việc tốt khi bài toán có bản chất là đi thu thập thông tin rồi thấy điều cần suy diễn
- Cho ra khối lượng lớn các thông tin từ một số thông tin ban đầu. Nó sinh ra nhiều thông tin mới.
- Suy diễn tiến là tiếp cận lý tưởng đối với các loại bài toán cần giải quyết các nhiệm vụ như lập kế hoạch, điều hành, điều khiển và diễn dịch.

Nhược điểm:

- i) Không cảm nhận được rằng chỉ cần một vài thông tin quan trọng. Hệ thống hỏi các câu hỏi có thể hỏi mà không biết rằng chỉ một ít câu đã đi đến kết luận được.
- ii) Hệ thống có thể hỏi cả câu hỏi không liên quan. Có thể các câu trả lời cũng quan trọng nhưng làm người dùng lúng túng khi phải trả lời các câu chẳng dính đến chủ đề.

3. Suy diễn tiến đối với logic vị từ

- Phương pháp

Dữ liệu vào:

Tập các mệnh đề $GT = \{gt1(.), gt2(.), \dots, gtm(.)\}$

Tập các luật $R = \{r_1, r_2, \dots, r_k\}$, với r_i ở dạng chuẩn Horn

$$P_1(.) \wedge P_2(.) \wedge P_3(.) \wedge \dots \wedge P_n(.) \rightarrow q(.), i=1..n$$

Tập kết luận $KL = \{q_1(.), q_2(.), \dots, q_k(.)\}$

Dữ liệu ra:

Thông báo thành công nếu mọi q_i đều được suy ra từ GT và tập R

Một tri thức được thể hiện bằng một **câu Horn dạng chuẩn**:

$$p_1 \wedge p_2 \wedge \dots \wedge p_n \rightarrow q$$

Các câu Horn dạng này còn được gọi là luật if- then và được biểu diễn như sau:

if P_1 and....and P_m then Q

Một câu Horn dạng tổng quát:

$$p_1 \wedge p_2 \wedge \dots \wedge p_n \rightarrow q_1 \vee q_2 \vee \dots \vee q_m$$

Lưu ý:

Nếu có luật dạng: $p_1 \wedge p_2 \wedge \dots \wedge p_n \rightarrow q_1 \vee q_2 \vee \dots \vee q_m$ thì tương đương với m luật sau:

$$p_1 \wedge p_2 \wedge \dots \wedge p_n \wedge \neg q_2 \wedge \dots \wedge \neg q_m \rightarrow q_1$$

$$p_1 \wedge p_2 \wedge \dots \wedge p_n \wedge \neg q_1 \wedge \neg q_3 \wedge \dots \wedge \neg q_m \rightarrow q_2$$

$$p_1 \wedge p_2 \wedge \dots \wedge p_n \wedge \neg q_1 \wedge \dots \wedge \neg q_{m-1} \rightarrow q_m$$

Tuy nhiên ta chỉ xét câu Horn dạng chuẩn ($m=1$)

- Nếu $n=0, m=1$: câu Horn có dạng $\rightarrow q$: gọi là sự kiện (fact) q .
- Nếu $n>0, m=1$: câu Horn có dạng: $p_1 \wedge p_2 \wedge \dots \wedge p_n \rightarrow q$: gọi là luật (rule).

Trong các hệ chuyên gia, cơ sở tri thức gồm 2 phần: tập các sự kiện (facts) và tập luật (rules).

- **Giải thuật**

```

Void SDT_VT()
{
    TG=Gt
    SAT=Loc {R, TG}
    While  $KL \not\subset TG$  and  $SAT \neq \emptyset$ 
    {
         $r \leftarrow \text{get}(SAT)$ 
         $TG = TG \cup \{q\theta\}$ 
         $SAT = \text{Loc}(R, TG)$ 
    }
    IF KL thuộc TG dừng báo “Thành Công”
    Else “ Không thành công”
}

```

Chú ý:

Khi lấy 1 luật trong SAT ta phải xác định luôn phép gán θ thỏa mãn $p_i\theta \in TG$,
 $\forall i$

Ví dụ:

Cho Luật r: $p(x,y,z) \wedge q(y,z) \rightarrow r(x,z)$

$GT = \{p(a,a,b), p(a,c,d), q(c,d)\}$

$TG = GT$,

khi đó phép gán $\theta_1 = \{a/x, a/y, b/z\}$

$\rightarrow r_{\theta_1} : p(a,a,b) \wedge q(a,b) \rightarrow r(a,b)$

Giả thiết $q(a,b)$ chưa thuộc TG nên thực hiện phép gán khác

$\theta_2 = \{a/x, c/y, d/z\}$

$\rightarrow r_{\theta_2} : p(a,c,d) \wedge q(c,d) \rightarrow r(a,d)$

Cả 2 giả thiết đã thuộc GT nên $TG = TG \cup \{r(a,b), r(a,d)\}$

Để minh họa cho giải thuật này ta xét bài toán ví dụ sau:

Ví dụ 2: Hai người gọi là anh em nếu họ có cùng cha mẹ. Hai người cũng được gọi là anh em nếu cha mẹ là anh em. Biết Bắc là cha mẹ của Nam, Tây là cha mẹ của Đông, Bắc và Tây có cùng cha mẹ. Chứng minh rằng Nam và Đông là Anh em

Các bước thực hiện

Bước 1: Xác định vị từ

Bước 2: Xây dựng quy tắc

Bước 3: Chứng minh

Giải:

Bước 1:

$\text{Chame}(x,y)$: x là cha mẹ của y

$\text{Anhem}(x,y)$: x là anh em của y

Bước 2:

Hai người gọi là anh em họ nếu có cùng cha mẹ

$\text{Chame}(x,y) \wedge \text{chame}(x,z) \rightarrow \text{anhem}(y,z)$

Hai người được gọi là anh em nếu cha mẹ là anh em.

$\text{Chame}(x,y) \wedge \text{Chame}(z,t) \wedge \text{Anhem}(x,z) \rightarrow \text{anhem}(y,t)$

$\text{Anhem}(x,y) \rightarrow \text{anhem}(y,x)$

Bước 3:

$\text{GT} = \{\text{Chamme}(\text{Bắc}, \text{Nam}), \text{Chame}(\text{Tây}, \text{Đông}), \text{Chame}(\text{A}, \text{Bắc}), \text{Chame}(\text{A}, \text{Tây})\};$

$\text{KL} = \{\text{Anhem}(\text{Nam}, \text{Đông})\}$

$\text{TG} = \text{GT}$, thực hiện phép gán $r_1 \theta_1$ với $\theta_1 = \{\text{A}/x, \text{Bắc}/Y, \text{Tây}/Z\}$

$r_1: \text{Chame}(x,y) \wedge \text{chame}(x,z) \rightarrow \text{Anhem}(y,z)$

$r_1 \theta_1: \text{Chame}(\text{A}, \text{Bắc}) \wedge \text{chame}(\text{A}, \text{Tây}) \rightarrow \text{Anhem}(\text{Bắc}, \text{Tây})$

Vì $\text{Chame}(\text{A}, \text{Bắc}), \text{chame}(\text{A}, \text{Tây}) \in \text{TG}$ nên $\text{SAT} = \{(r_1, \theta_1)\}$

$\text{KL} \notin \text{TG}$, $\text{SAT} \neq \emptyset$ nên lấy (r_1, θ_1) trong SAT

$\text{TG} = \text{TG} \cup \text{Anhem}(\text{Bắc}, \text{Tây})$

thực hiện phép gán $r_2 \theta_2$ với $\theta_2 = \{\text{Bắc}/x, \text{Nam}/Y, \text{Tây}/Z, \text{Đông}/t\}$

$r_2: \text{Chame}(x,y) \wedge \text{Chame}(z,t) \wedge \text{Anhem}(x,z) \rightarrow \text{anhem}(y,t)$

$r_2 \theta_2 \text{Chame}(\text{Bắc}, \text{Nam}) \wedge \text{Chame}(\text{Tây}, \text{Đông}) \wedge$

$\text{Anhem}(\text{Bắc}, \text{Tây}) \rightarrow \text{anhem}(\text{Nam}, \text{Đông})$

Vì $\text{Chame}(\text{Bắc}, \text{Nam}), \text{Chame}(\text{Tây}, \text{Đông}), \text{Anhem}(\text{Bắc}, \text{Tây}) \in \text{TG}$ nên $\text{SAT} = \{(r_2, \theta_2)\}$

$\text{KL} \notin \text{TG}$, $\text{SAT} \neq \emptyset$ nên lấy (r_2, θ_2) trong SAT

$TG = TG \cup \text{Anhem}(\text{Nam}, \text{Đông})$

$KL \in TG \rightarrow$ Thông báo thành công

4. Suy diễn lùi đối với logic mệnh đề

Là quá trình suy luận ngược xuất phát từ một số sự kiện ban đầu, ta tìm kiếm các sự kiện đã "sinh" ra sự kiện này.

Nói cách khác quá trình xuất phát từ sự kiện cần chứng minh và thay vào đó là những sự kiện ở vế trái của 1 luật có vế phải là sự kiện cần chứng minh. Quá trình này được thực hiện cho đến khi đưa về các sự kiện là tập sự kiện con của tập sự kiện giả thiết.

Nghĩa là: để đưa ra kết luận b , ta thử tìm tất cả các luật có dạng: $a_1 \wedge \dots \wedge a_n \Rightarrow b$, để có b , phải đưa ra các kết luận a_1, \dots, a_n . Quá trình xác định a_i cũng tương tự như đối với b , nếu đến một lúc nào đó phát hiện được rằng có một a_i nào đó không dẫn xuất được từ các giả thiết thì quay lui sang các luật sản xuất khác sinh ra b có dạng $b_1 \wedge \dots \wedge b_m \Rightarrow b$. Ngược lại, nếu mọi a_i đều dẫn xuất được giả thiết thì quá trình dẫn xuất ra b là đúng \rightarrow DỪNG

Giải thuật

T là tập các sự kiện cần chứng minh tại thời điểm đang xét (khởi tạo $T =$ Tập kết luận).

$S(p) = \{r_i \in R \mid \text{right}(r_i) = p\}$ (là tập các luật trong R sao cho vế phải chứa p)
 p sự kiện cần chứng minh

r : luật lấy từ $S(p)$

Ví dụ 1

Cho tập sự kiện $Gt = \{p, r\}$,

Tập luật R : $r_1: p \rightarrow q$

$r_2: q \wedge r \rightarrow s$

Chứng minh s

p	r	T	S(p)
		s	
s	r_2	q, r	r_2
q	r_1	r, p	r

Ví dụ 2:

Cho tập luật sau, vẽ đồ thị AND, OR mô tả suy diễn lùi

$(R_1): A \rightarrow E$

(R₂): $B \rightarrow D$

(R₃): $H \rightarrow A$

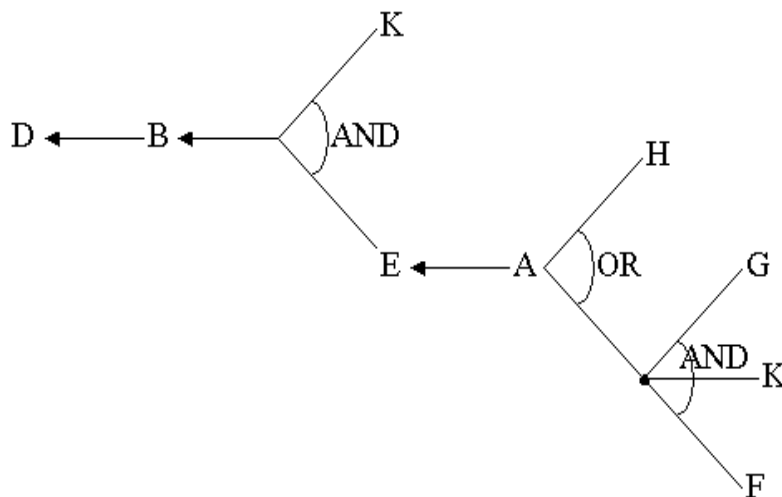
(R₄): $E \wedge G \rightarrow C$

(R₅): $E \wedge K \rightarrow B$

(R₆): $D \wedge E \wedge K \rightarrow C$

(R₇): $G \wedge K \wedge F \rightarrow A$

Giải:



Nhận xét về suy diễn lùi:

Ưu điểm:

- Phù hợp với bài toán đưa ra giả thuyết và liệu giả thuyết đó có đúng hay không?
- Tập trung vào đích đã cho. Nó tạo ra một loạt câu hỏi chỉ liên quan đến vấn đề đang xét, thuận tiện đối với người dùng.
- Khi suy diễn một điều gì từ thông tin đã biết, nó chỉ tìm trên một phần của cơ sở tri thức thích đáng đối với bài toán đang xét.
- Suy diễn lùi được đánh giá cao trong các bài toán như là chẩn đoán, dự đoán và tìm lỗi.

Nhược điểm:

Nhược điểm cơ bản của loại suy diễn này là nó thường tiếp theo dòng suy diễn thay vì đúng ra phải dừng ở đó mà sang nhánh khác.

Kết luận biểu diễn tri thức bằng luật

Biểu diễn tri thức bằng luật đặc biệt hữu hiệu trong những tình huống hệ thống cần đưa ra những hành động dựa vào những sự kiện có thể quan sát được. Nó có những ưu điểm chính yếu sau đây:

- Các luật rất dễ hiểu nên có thể dễ dàng dùng để trao đổi với người dùng (vì nó là một trong những dạng tự nhiên của ngôn ngữ).
- Có thể dễ dàng xây dựng được cơ chế suy luận và giải thích từ các luật.
- Việc hiệu chỉnh và bảo trì hệ thống là tương đối dễ dàng.
- Có thể cải tiến dễ dàng để tích hợp các luật mờ.
- Các luật thường ít phụ thuộc vào nhau.