Tạp chí Tin học và Điều khiển học, T. 18, S. 1 (2002), 73–79

MỘT SỐ PHƯƠNG PHÁP TIẾP CẬN TRONG VIỆC XÁC ĐỊNH NGỮ NGHĨA CỦA CƠ SỞ DỮ LIỆU TUYỂN

LÊ MẠNH THẠNH, TRẦN NGUYÊN PHONG

Abstract. There are some different approaches that overcome the problems of deductive databases; such as Closed Word Assumption (CWA), Generalized Closed World Assumption (GCWA), Disjunctive Database Rule (DDR),... These approaches concerned with negative information in database. In this paper, we introduce some approaches that define semantics of deductive database and their remained problems.

Tóm tắt. Hiện nay đã có nhiều cách tiếp cận được đưa ra nhằm mục đích giải quyết các vấn đề tồn tại trong cơ sở dữ liệu duy diễn như giả thiết thế giới đóng (CWA), giả thiết thế giới đóng mở rộng (GCWA), các qui tắc cơ sở dữ liệu tuyển (DDR),... Các phương pháp này tập trung vào việc xử lí các thông tin âm (negative information) xuất hiện trong cơ sở dữ liệu. Trong bài báo này, chúng tôi đề cập đến một số phương pháp tiếp cận trong xử lí ngữ nghĩa của cơ sở dữ liệu suy diễn và xem xét đến những tồn tại trong các cách tiếp cận đó.

1. CÁC KHÁI NIỆM

Trước hết, chúng tôi đề cập đến một số khái niệm sẽ được sử dụng trong các phần còn lại. Các khái niệm được dựa trên cơ sở của logic vị từ cấp một và cơ sở dữ liệu quan hệ. Tuy nhiên, trong bài báo này chúng tôi chỉ đề cập đến những cơ sở dữ liệu trong đó không có sự xuất hiện của các kí hiệu hàm; tức là các đối của các vị từ chỉ là các biến hoặc là hằng.

Một *mệnh đề* là một công thức có dạng:

$$A_1 \vee \ldots \vee A_m \leftarrow B_1 \wedge \ldots \wedge B_n .$$

Trong đó các A_i $(i = 1, , m)$ và B_j $(j = 1, \ldots , n)$ là các công thức nguyên tử. $A_1 \vee \ldots \vee A_m$ được gọi là *phần đầu* của mệnh đề và $B_1 \wedge \ldots \wedge B_n$ được gọi là *thân* của mệnh đề. Nếu phần đầu của mệnh đề chỉ có duy nhất một nguyên tử (tức là $m = 1$) thì mệnh đề được gọi là *mệnh đề Horn*. Một mệnh đề có thể có phần đầu hoặc phần thân rỗng (nhưng không thể là cả hai). Một mệnh đề được gọi là *mệnh đề âm* nếu phần đầu của nó là rỗng, khi đó mệnh đề còn có thể được viết dưới dạng:

$\neg B_1 \vee \ldots \vee \neg B_n$

hoặc $\neg(B_1 \wedge \ldots \wedge B_n)$.

Các mệnh đề âm được xem như là các ràng buộc toàn vẹn trong cơ sở dữ liệu. Trong trường hợp một mệnh đề có phần thân là rỗng thì mệnh đề đó được gọi là *mệnh đề dương*. Một mệnh đề được gọi là *đầy đủ* nếu cả phần thân và phần đầu đều khác rỗng.

Một *cơ sở dữ liệu* là một tập hữu hạn các mệnh đề. Một cơ sở dữ liệu được xem là ở dạng Horn nếu tất cả các mệnh đề trong nó đều là mệnh đề Horn, ngược lại là *cơ sở dữ liệu tuyển*.

Tập tất cả các nguyên tử cơ sở của một cơ sở dữ liệu liệu được gọi là *cơ sở Herbrand* của cơ sở dữ liệu đó. Nếu gọi H là cơ sở Herbrand thì một tập con bất kì của H được gọi là *thể hiện Herbrand* (hay *thể hiện*) của cơ sở dữ liệu.

Gọi DB là một tập các mệnh đề và M là thể hiện Herbrand của DB. Ta nói M là một *mô hình* của DB nếu DB đúng trong M. M được gọi là *mô hình cực tiểu* nếu không tồn tại bất kì một mô hình M' nào của DB sao cho M' là tập con thực của M. DB được gọi là *nhất quán* nếu tồn tại ít nhất một mô hình của DB, nếu không DB được gọi là không nhất quán.

Một mệnh đề C được gọi là *mệnh đề được suy dẫn* từ DB (kí hiệu $DB \vdash C$) nếu mọi mô hình của DB cũng là mô hình của C. Một mệnh đề cơ sở $C = A_1 \vee \ldots \vee A_n$ được gọi là một *mệnh đề cực tiểu dương* được suy dẫn từ DB nếu thỏa mãn các điều kiện sau:

(1) C dương.

(2) $DB \vdash C$.

(3) $DB \not\vdash A_1 \vee \ldots \vee A_{i-1} \vee A_{i+1} \vee \ldots \vee A_n$ $(\forall i = 1, \ldots, n)$.

Trong các phần còn lại, chúng tôi đề cập đến một cách tiếp cận trong việc giải quyết ngữ nghĩa của một tập các mệnh đề. Để đơn giản trong việc trình bày, chúng tôi giả sử một cơ sở dữ liệu chỉ bao gồm các mệnh đề cơ sở, tức là các mệnh đề được biểu diễn với các hằng xuất hiện trong cơ sở dữ liệu.

2. GIẢ THIẾT THẾ GIỚI ĐÓNG SUY RỘNG (GCWA)

Trong phần này chúng tôi bàn luận đến giả thiết thế giới đóng suy rộng GCWA (Generalized Closed World Assumption). Trước hết, chúng tôi đề cập đến giả thiết thế giới đóng CWA, được sử dụng trong trường hợp các mệnh đề trong cơ sở dữ liệu là các mệnh đề Horn và một số vấn đề mà CWA gặp phải trong trường hợp các mệnh đề không phải là mệnh đề Horn.

Trong trường hợp các mệnh đề trong cơ sở dữ liệu là các mệnh đề Horn (chẳng hạn như chương trình Datalog), Reiter đã đưa ra CWA nhằm xử lí ngữ nghĩa của các literal âm. Theo Reiter, nếu một công thức nguyên tử cơ sở $p(a_1, \ldots, a_n)$ không thể suy ra được từ những qui tắc và sự kiện đã biết trong cơ sở dữ liệu thì $\neg p(a_1, \ldots, a_n)$ sẽ được xem là đúng [6]. Như vậy, bản thân CWA cho phép suy ra những sự kiện có dạng $\neg p(a_1, \ldots, a_n)$ khi các phương pháp suy diễn không thể khẳng định được giá trị chân lý của $p(a_1, \ldots, a_n)$.

Khi các mệnh đề trong cơ sở dữ liệu là các mệnh đề Horn thì CWA đóng một vai trò khá quan trọng. Tuy nhiên, trong trường hợp các mệnh đề không ở dạng Horn (tức là cơ sở dữ liệu tuyển) thì bản thân CWA lại dẫn đến những mâu thuẫn. Chẳng hạn, gọi $DB = \{p \vee q\}$. Khi đó theo CWA thì cả p và q đều không thể suy ra từ DB và do đó $\mathrm{CWA}(DB) = \{\neg p, \neg q\}$. Điều này dẫn đến $DB \cup \mathrm{CWA}(DB)$ là không nhất quán.

Để giải quyết những mâu thuẫn đối với CWA, J. Minker đã đề xuất cách giải quyết khác mở rộng từ CWA gọi là *giả thiết thế giới đóng mở rộng* (GCWA) [4]. GCWA được hình thành dựa trên cơ sở mô hình cực tiểu. Gọi DB là một cơ sở dữ liệu nhất quán và p là một nguyên tử cơ sở. Theo Minker, $\neg p$ được xem là đúng nếu p không xuất hiện trong bất kì một mô hình cực tiểu nào của DB.

Ví dụ. Giả sử $DB = \{p \vee q,\ q \vee r \vee u,\ u \vee v,\ \neg(p \wedge v)\}$. Tập các mô hình cực tiểu của DB là $\{\{p, u\}, \{q, v\}, \{q, u\}\}$. Như vậy, r không thuộc vào một mô hình cực tiểu nào của DB nên $\neg r$ được coi là đúng.

Xét cơ sở dữ liệu nhất quán DB. Gọi H là cơ sở Herbrand của DB và kí hiệu $\mathrm{PMGC}(DB)$ là tập tất cả các mệnh đề cực tiểu dương được suy dẫn từ DB. Kí hiệu ATOM(PMGC) là tập tất cả các nguyên tử cơ sở $A \in C$ (với $C \in \mathrm{PMGC}(DB)$). Khi đó GCWA còn được phát biểu như sau:

Định nghĩa 2.1. Gọi DB là một cơ sở dữ liệu nhất quán và A là một nguyên tử cơ sở. $\neg A$ được xem là đúng nếu $A \in H - \mathrm{ATOM(PMCG)}$.

Định lý 2.1. [4] *Gọi DB là một cơ sở dữ liệu nhất quán và A là một nguyên tử cơ sở. Khi đó*

(i) *$A \in H - ATOM(PMGC)$ nếu và chỉ nếu A không thuộc vào bất kỳ một mô hình cực tiểu nào của DB.*

(ii) *$DB \cup GCWA(DB)$ là nhất quán.*

(iii) *$DB \vdash C$ nếu và chỉ nếu $DB \cup GCWA(DB) \vdash C$ (với C là một mệnh dương).*

(iv) *$DB \cup GCWA(DB) \vdash \neg A$ nếu và chỉ nếu $\neg A \in GCWA(DB)$.*

3. QUI TẮC CƠ SỞ DỮ LIỆU TUYỂN

Trong phần này, chúng tôi đề cập đến *qui tắc cơ sở dữ liệu tuyển* (DDR) được Ross và Topor đề xuất [5][(*)].

Ta định nghĩa *tập đóng* của một cơ sở dữ liệu là một tập các nguyên tử cơ sở có thể được thừa nhận là sai. Gọi DB là một cơ sở dữ liệu, H là cơ sở Herbrand và S là một tập con của H. Khi đó S là một tập đóng của DB nếu với mọi nguyên tử cơ sở $A \in S$ và với mọi mệnh đề cơ sở $C \in DB$ sao cho A nằm trong phần đầu của C, tồn tại một nguyên tử B trong phần thân của C sao cho $B \in S$. Gọi *tập đóng lớn nhất* của DB là hợp của tất cả các tập đóng của DB và kí hiệu tập này là gcs(DB).

Ví dụ. Xét $DB = \{p \vee q,\ r \leftarrow p \wedge q,\ u \leftarrow v\}$, ta nhận thấy gcs$(DB) = \{v, u\}$.

Theo Ross và Topor, nếu DB là một cơ sở dữ liệu và A là một nguyên tử cơ sở thì $\neg A$ được xem là đúng nếu $A \in$ gcs(DB). Trước khi bàn luận đến định nghĩa điểm cố định của DDR, ta xét ánh xạ T_{DB} được định nghĩa như sau:

Gọi DB là cơ sở dữ liệu và I là một thể hiện Herbrand của DB. Khi đó $T_{DB}(I)$ là tập tất cả các nguyên tử cơ sở $A \in H$ sao cho: với C là một mệnh đề cơ sở của DB, A xuất hiện trong phần đầu của C và với mọi nguyên tử B trong phần thân của C ta có $B \in I$. Ta định nghĩa chuỗi $T_{DB}^{\ k}$ như sau:

$$T_{DB}^{\ 0} = \emptyset$$
$$T_{DB}^{\ i+1} = T_{DB}(T_{DB}^{\ i})$$
$$\ldots$$
$$\text{và} \qquad T_{DB}^{\ \omega} = \bigcup_{i=1}^{\infty} T_{DB}^{\ i}$$

Ví dụ: Gọi $DB = \{p \vee q,\ r \vee s \vee v \leftarrow p,\ u \leftarrow r \wedge s\}$. Khi đó, ta có:

$$T_{DB}^{\ 0} = \emptyset$$
$$T_{DB}^{\ 1} = \{p, q\}$$
$$T_{DB}^{\ 2} = \{p, q, r, s, v\}$$
$$T_{DB}^{\ 3} = \{p, q, r, s, v, u\}$$
$$T_{DB}^{\ 4} = T_{DB}^{\ 3}$$
$$\text{Do đó} \qquad T_{DB}^{\ \omega} = \{p, q, r, s, v, u\}$$

Định nghĩa 3.1. Gọi DB là một cơ sở dữ liệu nhất quán, H là cơ sở Herbrand của DB và A là một nguyên tử cơ sở. $\neg A$ được xem là đúng nếu $A \in H - T_{DB}^{\ \omega}$.

Kí hiệu DDR$(DB) = \{\neg A \mid A \in H - T_{DB}^{\ \omega}\}$. Khi đó, ta có một số tính chất sau:

Định lý 3.1. [5] *Gọi DB là một cơ sở dữ liệu nhất quán, khi đó:*

(i) $gcs(DB) = H - T_{DB}^{\ \omega}$.

(ii) *$DB \cup DDR(DB)$ là nhất quán.*

(iii) *Với C là một mệnh đề dương, $DB \vdash C$ nếu và chỉ nếu $DB \cup DDR(DB) \vdash C$.*

(*) Một cách tiếp cận khác tương tự DDR được gọi là giả thiết thế giới đóng tổng quát yếu (WGCWA) được trình bày trong [5].

(iv) *Nếu* $C = B_1 \vee \ldots \vee B_m \leftarrow A_1 \wedge \ldots \wedge A_n$ *là một mệnh đề không dương sao cho* $DB \cup DDR(DB) \vdash C$ *nhưng* $DB \not\vdash C$ *thì tồn tại* A_i *nào đó sao cho* $\neg A_i \in DDR(DB)$.

(v) *Với* A *là một nguyên tử cơ sở,* $DB \cup \mathrm{DDR}(DB) \not\vdash \neg A$ *nếu và chỉ nếu* $DB \not\vdash \neg A$ *hay* $A \in H - T_{DB}^{\omega}$.

4. NGỮ NGHĨA THẾ GIỚI KHẢ HỮU (PWS) CỦA CƠ SỞ DỮ LIỆU TUYỂN

DDR được đề xuất nhằm khắc phục những vấn đề tồn tại trong GCWA. Tuy nhiên, DDR vẫn còn những cái cần phải xem xét. PWS (Possible World Semantics) được Edward P. F. Chan đề xuất nhằm khắc phục những bất thường tồn tại trong CGWA và DDR [2].

Ví dụ. Gọi $DB = \{p, q \vee r \leftarrow p,\ u \leftarrow q \wedge r,\ \neg(q \wedge r)\}$. Theo DDR ta có $T_{DB}^{\omega} = \{p, q, r, u\}$. Ta nhận thấy, mệnh đề $\neg(q \wedge r)$ không ảnh hưởng gì đến các trực kiện âm được suy từ DB trong khi lẽ ra việc tồn tại hay không $\neg(q \wedge r)$ trong DB phải tác động đến điều này.

Ta định nghĩa một *thế giới khả hữu* của một cơ sở dữ liệu là một tập các nguyên tử cơ sở được thừa nhận là đúng. Như vậy, tập các nguyên tử này sẽ tạo thành một mô hình của cơ sở dữ liệu. Xét trường hợp của ví dụ trên, ta nhận thấy p luôn xuất hiện trong một thế giới khả hữu của DB, do đó q hoặc r (nhưng không thể cả hai) cũng có thể xuất hiện trong thế giới khả hữu của DB. Do q và r không thể đồng thời đúng nên u không thể xuất hiện trong bất kì thế giới khả hữu nào của DB. Vậy tập các thế giới khả hữu của DB là $\{\{p, q\}, \{p, r\}\}$.

Cho C là một mệnh đề cơ sở có dạng $A_1 \vee \ldots \vee A_m \leftarrow B_1 \wedge \ldots \wedge B_n$ và S là một tập con khác rỗng của $\{A_1, \ldots, A_m\}$. *Phép tách* của C theo S là một tập các mệnh đề Horn được định nghĩa như sau:

$$\mathrm{Split}(S) = \{A_i \leftarrow B_1 \wedge \ldots \wedge B_n \mid A_i \in S\}.$$

Cho $DB = PC \cup MC \cup NC^{(*)}$, ta kí hiệu Horn$(DB)$ là tập tất cả các chương trình Horn sao cho mỗi chương trình DB' trong Horn(DB) có được bằng cách:

(i) Thay mỗi mệnh đề tuyển trong DB bởi các mệnh đề trong phép tách của nó.

(ii) Giữ nguyên các mệnh đề khác.

Khi đó, tập các thế giới khả hữu của DB là tập các mô hình Herbrand nhỏ nhất của các chương trình nhất quán trong Horn(DB).

Ví dụ. Xét $DB = \{\{p \vee q,\ q \vee r,\ \neg(q \wedge r)\}$. Khi đó Horn$(DB) = \{\{p, q, \neg(q \wedge r)\}, \{p, r, \neg(q \wedge r)\}, \{q, \neg(q \wedge r)\}, \{q, r, \neg(q \wedge r)\}, \{p, q, r, \neg(q \wedge r)\}\}$. Ta nhận thấy $\{q, r, \neg(q \wedge r)\}$, $\{p, q, r, \neg(q \wedge r)\}$ là không nhất quán, do đó tập các thế giới khả hữu của DB là $\{\{p, q\}, \{p, r\}, \{q\}\}$.

Gọi A là một công thức nguyên tử cơ sở của DB, khi đó ta có:

- A được xem là *đúng* nếu A thuộc tất cả các thế giới khả hữu của DB.
- A được gọi là *sai* nếu A không thuộc bất kì thế giới khả hữu nào của DB.
- A được gọi là *có khả năng đúng* nếu A thuộc một số thế giới khả hữu nào đó của DB (nhưng không phải là tất cả).

Ta kí hiệu:

PW$(DB) = \{W \mid W$ là một thế giới khả hữu của $DB\}$.

True$(DB) = \{A \mid A$ là một nguyên tử cơ sở thuộc tất cả các thế giới khả hữu của $DB\}$.

Possibly_True$(DB) = \{A \mid A$ là một nguyên tử cơ sở có khả năng đúng trong $DB\}$.

(*) PC, MC và NC lần lượt là tập các mệnh đề dương, các mệnh đề đầy đủ và các mệnh đề âm trong cơ sở dữ liệu.

Khi đó ta có:

$$\cup \mathrm{PW}(DB) = \mathrm{True}(DB) \cup \mathrm{Possibly\_True}(DB).$$

Định nghĩa của PWS. Gọi DB là một cơ sở dữ liệu nhất quán và A là một nguyên tử cơ sở. $\neg A$ được suy ra từ DB nếu $A \in H - \cup \mathrm{PW}(DB)$.

5. NGỮ NGHĨA ĐÁNG TIN CẬY CỦA CƠ SỞ DỮ LIỆU TUYỂN

Trong phần trước, chúng tôi đã đề cập đến một số cách tiếp cận trong việc xử lí ngữ nghĩa của các cơ sở dữ liệu tuyển. Các cách tiếp cận nói chung tập trung vào việc xác định giá trị chân lý của các nguyên tử cơ sở là *true* hay *flase*. Tuy nhiên, không phải trường hợp nào các thông tin hiện có cũng đầy đủ để có thể cho phép chúng ta xác định chính xác giá trị chân lý của một sự kiện.

Ngữ nghĩa đáng tin cậy (WFS: Well-Founded Semantics) cung cấp cho chúng ta cái nhìn tự nhiên hơn đối với ngữ nghĩa của các cơ sở dữ liệu. Trong phần này, chúng tôi trước hết đề cập đến các thể hiện 3 giá trị, trong đó giá trị chân lý của các sự kiện có thể là *true* (đúng), *false* (sai) hoặc *unknown* (không xác định). Tiếp đó sẽ bàn luận đến mô hình 3 giá trị và ngữ nghĩa đáng tin cậy (WFS) của cơ sở dữ liệu xác định.

Ví dụ. Xét $DB = \{u,\ s \leftarrow \neg u \wedge p,\ p \leftarrow \neg q,\ r \leftarrow \neg p,\ q \leftarrow \neg r\}$. Ta nhận thấy những thông tin hiện có không đủ để kết luận giá trị chân lý của p, q và r. Hay nói cách khác, ta có giá trị chân lý của u là *true*, s là *false* và p, q và r là *unknown*.

Để giải quyết những trường hợp như trên, ta mở rộng miền giá trị chân lý từ hai giá trị *true* và *false* thành 3 giá trị là *true, false* và *unknown*. Ta kí hiệu *true* là 1, *false* là 0 và *unknown* là 1/2.

Gọi DB là một cơ sở dữ liệu và H là cơ sở Herbrand của DB. Một thể hiện 3 giá trị I của DB là một ánh xạ toàn phần từ H vào $\{0, 1/2, 1\}$. Ta kí hiệu I^1, $I^{1/2}$ và I^0 là tập các nguyên tử cơ sở trong H có giá trị chân lý lần lượt là 1, 1/2 và 0. Gọi I_1 và I_2 là hai thể hiện, ta định nghĩa quan hệ thứ tự trên chúng như sau: $I_1 < I_2$ *nếu và chỉ nếu với mỗi* $A \in H$ *ta có* $I_1(A) \leq I_2(A)$.

Gọi I là một thể hiện 3 giá trị, ta định nghĩa hàm $\hat{I}$ từ tập các công thức cơ sở vào $\{0, 1/2, 1\}$ như sau:

- Nếu A là nguyên tử cơ sở thì $\hat{I}(A) = I(A)$.
- Nếu S và V là các công thức cơ sở thì:
  - $\hat{I}(\neg S) = 1 - \hat{I}(S)$.
  - $\hat{I}(S \vee V) = \max(\hat{I}(S), \hat{I}(V))$.
  - $\hat{I}(S \wedge V) = \min(\hat{I}(S), \hat{I}(V))$.
  - $\hat{I}(S \leftarrow V) = \begin{cases} 1 \text{ nếu } \hat{I}(S) \geq \hat{I}(V), \\ 0 \text{ trong các trường hợp còn lại.} \end{cases}$

Ta nhận thấy một thể hiện Herbrand I sẽ là mô hình của DB nếu mọi mệnh đề của DB đúng trong I. Tức là với mọi mệnh đề cơ sở $A \leftarrow A_1 \wedge \ldots \wedge A_n$ ta có:

$$\hat{I}(A) \geq \min\{\hat{I}(A_i) : i = 1, 2, \ldots, n\}.$$

Tiếp đến, chúng ta sẽ bàn luận đến *mô hình 3 giá trị bền* dựa trên thể hiện 3 giá trị đã đề cập đến ở trên. Gọi DB là một cơ sở dữ liệu và I là thể hiện 3 giá trị của DB. Ta kí hiệu $\mathrm{pg}(DB, I)$ là tập các mệnh đề cơ sở có được bằng cách thay mỗi giả thiết âm A bởi $\hat{I}(A)$. Như vậy $\mathrm{pg}(DB, I)$ sẽ không còn chứa những trực kiện âm trong nó. Ta gọi $\Psi_{DB}(I)$ là một phép thể hiện được định nghĩa như sau:

$$\Psi_{DB}(I)(A) = \begin{cases} 1 & \text{nếu tồn tại trong } DB \text{ mệnh đề } A \leftarrow A_1 \wedge \ldots \wedge A_n \text{ sao cho} \\ & \hat{I}(A_1 \wedge \ldots \wedge A_n) = 1 \ (\forall i \leq n), \\ 0 & \text{nếu với mỗi mệnh đề } A \leftarrow A_1 \wedge \ldots \wedge A_n \text{ trong } DB \\ & \hat{I}(A_1 \wedge \ldots \wedge A_n) = 0 \text{ (hoặc không tồn tại mệnh đề nào có } A \text{ ở} \\ & \text{phần đầu),} \\ \frac{1}{2} & \text{trong các trường hợp còn lại.} \end{cases}$$

Cho DB cơ sở dữ liệu và I là thể hiện 3 giá trị của DB. Gọi $\perp$ là thể hiện trong đó tất cả các nguyên tử cơ sở đều có giá trị là 0. Kí hiệu $\Gamma_P(I)$ là điểm bất động nhỏ nhất của $\Psi_{\mathrm{pg}(P,I)}(\perp)$. I được gọi là *mô hình 3 giá trị bền* của P nếu và chỉ nếu:

$$\Gamma_P(I) = I.$$

Ví dụ. Xét $DB = \{p \leftarrow \neg r;\ q \leftarrow \neg r \wedge p;\ s \leftarrow \neg t,\ t \leftarrow q \wedge \neg s;\ u \leftarrow \neg t \wedge p \wedge s\}$ và $I = \{p, q, \neg r\}$. Khi đó, $\mathrm{pg}(DB, I) = \{p \leftarrow 1;\ q \leftarrow 1 \wedge p;\ s \leftarrow 1/2;\ t \leftarrow q \wedge 1/2;\ u \leftarrow 1/2 \wedge p \wedge s\}$, $\perp = \{\neg p, \neg q, \neg r, \neg s, \neg t, \neg u\}$.
Ta có:

$\Psi_{\mathrm{pg}(P,I)}{}^{(1)}(\perp) = \{p, \neg q, \neg r, \neg t, \neg u\}$,
$\Psi_{\mathrm{pg}(P,I)}{}^{(2)}(\perp) = \{p, q, \neg r, \neg t\}$,
$\Psi_{\mathrm{pg}(P,I)}{}^{(3)}(\perp) = \{p, q, \neg r\}$,
$\Psi_{\mathrm{pg}(P,I)}{}^{(4)}(\perp) = \{p, q, \neg r\}$.

Vậy I là mô hình 3 giá trị bền của DB do $\Gamma_p(I) = \{p, q, \neg r\} = I$.

Định nghĩa ngữ nghĩa đáng tin cậy. Gọi DB là một cơ sở dữ liệu. Ngữ nghĩa đáng tin cậy của DB là một thể hiện 3 giá trị chứa tất cả các sự kiện âm và dương thuộc tất cả các mô hình 3 giá trị bền của DB.

Trên đây là định nghĩa hình thức về ngữ nghĩa đáng tin cậy của cơ sở dữ liệu. Tuy nhiên, sẽ rất khó khăn trong việc xác định WFS bằng cách kiểm tra tất cả các thể hiện 3 giá trị, xác định xem thể hiện nào là mô hình 3 giá trị bền và tiếp đó lấy giao của chúng. Phương pháp được nêu dưới đây giúp chúng ta xác định được WFS của một sơ sở dữ liệu được dễ dàng hơn [7].

Ta kí hiệu $\perp$ là thể hiện 3 giá trị trong đó tất cả các sự kiện được xem là có giá trị 0 (false). Quá trình tính WFS của một cơ sở dữ liệu được tiến hành thông qua các bước như sau:

Bước 1: Khởi tạo $I_0 = \perp$.

Bước 2: Bước lặp

Tính chuỗi $\{I_i\}_{i \geq 0}$ theo công thức:

$$I_{i+1} = \Gamma_P(I_i).$$

Quá trình này tiến hành lặp đi lặp lại cho đến khi không còn sự thay đổi nào trên chuỗi $\{I_{2j}\}_{j \geq 0}$ và $\{I_{2j+1}\}_{j \geq 0}$.

Bước 3: Đặt $I_* = \mathrm{limit}\,(\{I_{2j}\}_{j \geq 0})$ và $I^* = \mathrm{limit}\,(\{I_{2j+1}\}_{j \geq 0})$. Gọi I_*^* là thể hiện 3 giá trị chứa tất cả các sự kiện được biết trong I_* và I^* được xác định như sau:

$$I_*^* = \begin{cases} 1 & \text{nếu } I_*(A) = I^*(A) = 1, \\ 0 & \text{nếu } I_*(A) = I^*(A) = 0, \\ \frac{1}{2} & \text{trong các trường hợp còn lại.} \end{cases}$$

Bước 4: Kết luận WFS của DB là I_*^*.

Ví dụ. Xét $DB = \{p \leftarrow \neg r;\ q \leftarrow \neg r \wedge p;\ s \leftarrow \neg t,\ t \leftarrow q \wedge \neg s;\ u \leftarrow \neg t \wedge p \wedge s\}$. Áp dụng phương pháp trên, ta có chuỗi các thể hiện 3 giá trị I_i như sau:

$I_0 = \perp = \{\neg p, \neg q, \neg r, \neg s, \neg t, \neg u\}$,

$I_1 = \{p, q, \neg r, s, t, u\}$,

$I_2 = \{p, q, \neg r, \neg s, \neg t, \neg u\}$,

$I_3 = \{p, q, \neg r, s, t, u\}$,

$I_4 = \{p, q, \neg r, \neg s, \neg t, \neg u\}$.

Vậy $I_* = I_4 = \{p, q, \neg r, \neg s, \neg t, \neg u\}$ và $I^* = I_3 = \{p, q, \neg r, s, t, u\}$.

Do đó $I_*^* = \{p, q, \neg r\}$ hay WFS của DB là $\{p, q, \neg r\}$.

TÀI LIỆU THAM KHẢO

[1] A. Rajasekar, J. Lobo, J. Minker, Weal generalized closed world assumption, *Journal Automated Reasoning* **5** (1989) 293–307.

[2] Edward P. F. Chan, A possible world semantics for disjunctive databases, *IEEE Transactions on Knowledge and Data Engineering* **5** (1993) 282–292.

[3] J. D. Ullman, *Principle of Database and Knowledgebase Systems*, Computer Sciences Press, 1988.

[4] J. Minker, On indefinite databases and the closed world assumption, *Proc. 6th Int. Conf. on Automated Deduction, Lecture Notes in Computer Science* 310, Springer-Verlag, 1982, 292–308.

[5] K. A. Ross, R. W. Topor, Inferring negative information from disjunctive databases, *Journal of Automated Reasoning* **4** (1998) 397–424.

[6] R. Reiter, *On Closed Worls Databases*, Logic and Database, Plenum Press, New York, 1978.

[7] Serge Abiteboul, Richard Hull, Victor Vianu, *Foundations of Databases*, Addison-Wesley, 1995.

[8] Skama C., Possible model semantics for disjunctive databases, *Proc. 1st Int. Conf. On Deductive and Object-Oriented Databases*, 1989, 337–351.

Nhận bài ngày 5 - 10 - 2001
Nhận lại sau khi sửa ngày 7 - 1 - 2002

Trường Đại học Khoa học Huế