Hệ sinh ánh xạ đóng và bài toán biểu diễn phản cơ sở

Generation Systems for Closure Mapping and The Problem of Antibase Representation

Bùi Đức Minh

***Abstract**: Closure mapping on a finite set U is a mapping satisfied reflexibility, monotonicity, and idempotence properties. This is one of mathematical tools supporting theoretical aspects in several of IT fields, such as database and knowledge-base systems, deductive systems, data mining etc. .. Each closure mapping can be specified by a deductive system, called generation system. An antibase of a closure mapping f on U is the subset P of U satified f(P) ≠ U, and ∀A ∈U \ P: f(PA) = U. It is shown that antibases can be used in database design and data mining to reduce computational complexity of algorithms for computing such objects as closures, keys, normal forms, sensitive itemsets and association rules, etc… This paper presents some new results concerning representing the antibases of a given generation system. We show that each antibase in a generation system can be represented as a union of the maximal right-hand side and an antibase of the reduced system.*

***Keywords**: closure mapping, generation system, deductive system, antibase*

I. CÁC KHÁI NIỆM MỞ ĐẦU

Nhiều kết quả trong tin học lý thuyết dựa trên khái niệm *ánh xạ đóng* (AXĐ) như một toán tử thiết lập tương ứng giữa các tập con của tập hữu hạn cho trước, thỏa mãn các tiên đề *phản xạ*, *đồng biến* và *lũy đẳng*. Việc nghiên cứu tổng quát về các ánh xạ đóng cho phép ta mở rộng khả năng vận dụng một công cụ toán học trợ giúp phát triển một số kết quả trong các lĩnh vực nói trên. Mỗi ánh xạ đóng có thể được đặc trưng thông qua một hệ suy dẫn gọi là hệ sinh AXĐ. Từ đầu những năm 2000, các nhà khoa học trong nhiều công trình nghiên cứu đã công bố các lý thuyết về ánh xạ đóng, hệ sinh AXĐ, biểu diễn cơ sở, phản cơ sở của hệ sinh AXĐ thông qua phép thu gọn hệ sinh AXĐ nhằm mục đích nâng cao hiệu quả tính toán trên các đối tượng của AXĐ nói chung [5, 6, 7, 8] và các đối tượng đặc thù trong hướng nghiên cứu về khai phá và ẩn các đối tượng nhạy cảm như các tập thường xuyên và luật kết hợp [4]. Kết quả mới của bài viết là đề xuất một dạng biểu diễn phản cơ sở hệ sinh AXĐ theo vế phải cực đại của tập luật sinh. Kết quả này có ý nghĩa như sau: Thứ nhất, ta có thể sử dụng phản cơ sở thay cho vai trò của cơ sở vì cơ sở và phản cơ sở là hai khái niệm đối ngẫu và thuật toán xây dựng phản cơ sở từ cơ sở và ngược lại, xây dựng cơ sở từ phản cơ sở có độ phức tạp tính toán là tuyến tính [1]. Thứ hai, xác định được phản cơ sở thì cơ sở cũng được xác định cho phép ta giảm được không gian lưu trữ do tập luật được thu gọn và tăng hiệu quả tính toán cho các hệ thống quản lý các đối tượng phức tạp như cơ sở dữ liệu hướng đối tượng làm việc với các dữ liệu lớn, các hệ thống suy dẫn như datalog, ontology và khai phá dữ liệu, v.v..

Bài viết có cấu trúc như sau: Phần thứ nhất trình bày các định nghĩa, khái niệm về ánh xạ đóng, hệ sinh ánh xạ đóng và một số tính chất quan trọng liên quan đến các khái niệm này. Các khái niệm về cơ sở, phản cơ sở và phép thu gọn hệ sinh AXĐ được trình bày trong phần thứ hai. Phần thứ ba của bài viết trình bày các khái niệm về vế phải cực đại của các luật sinh, trên cơ sở đó phát biểu và chứng minh định lý về một dạng biểu diễn phản cơ sở của hệ sinh ánh xạ đóng thông

qua phép thu gọn hệ sinh theo vế phải cực đại của tập luật sinh.

Các ký hiệu và quy ước sau đây được sử dụng như trong [1, 2, 8]. Các phần tử của tập hợp được ký hiệu bằng các chữ *Latin viết in* đầu bảng chữ *A, B, C,…* Các tập được ký hiệu bằng các chữ *LATIN HOA* cuối bảng chữ *X, Y, Z,...* Các phần tử trong một tập thường được liệt kê như một *xâu ký tự*, chẳng hạn ta viết $X = abc$ thay vì viết $X = \{a,b,c\}$. XY biểu diễn *hợp* của hai tập X và Y. *Phép trừ* hai tập X và Y được ký hiệu là $X \backslash Y$. *Tập vũ trụ* hay *tập nền U* được cho trước luôn luôn là *hữu hạn* và *khác trống*. $|M|$ cho biết *lực lượng* của tập M. Ký hiệu *SubSet(U) là tập các tập con của U với thứ tự bộ phận bao hàm* ($\subseteq$). Với mỗi họ các tập con Ψ của U ta ký hiệu $\cap\Psi$ là *giao* của các tập con trong họ Ψ, cụ thể là $\cap\Psi = \bigcap_{X \in \Psi} X$

Cho $\Re, \Im \subseteq Poset(U)$ và $M, P \in Poset(U)$. Ta định nghĩa phép toán $\oplus$ trên $Poset(U)$ như sau:

$M \oplus P = MP$ (hợp của hai tập con M và P)

$M \oplus \Re = \{MX \mid X \in \Re\}$ và

$\Re \oplus \Im = \{ XY \mid X \in \Re, Y \in \Im \}$

Các khái niệm cơ bản sau đây được trình bày đầy đủ trong các công trình [8, 10].

Cho tập hữu hạn U. Ánh xạ f: $Subset(U) \rightarrow Subset(U)$ được gọi là *ánh xạ đóng* (AXĐ) trên U nếu với mọi tập con $X, Y \subseteq U$, f thỏa mãn các tính chất sau:

(C1) *Tính phản xạ*: $f(X) \supseteq X$,

(C2) *Tính đồng biến*: Nếu $X \subseteq Y$ thì $f(X) \subseteq f(Y)$,

(C3) *Tính lũy đẳng*: $f(f(X)) = f(X)$.

Gọi f là một AXĐ trên U . Tập con $X \subseteq U$ được gọi là *điểm bất động* (hay *tập đóng*) của AXĐ f nếu $f(X) = X$. Ta ký hiệu $Fix(f)$ là tập toàn thể các *điểm bất động* của AXĐ f trên U. Mặt khác, theo tính lũy đẳng của AXĐ, ta còn có thể mô tả $Fix(f)$ như sau, $Fix(f) = \{ f(X) \mid X \subseteq U \}$.

Một luật sinh AXĐ f trên U là biểu thức có dạng f: $L \rightarrow R$; $L, R \subseteq U$. Các tập L, R được gọi tương ứng là vế trái và vế phải của luật sinh f và được ký hiệu tương ứng là $LS(f)$ và $RS(f)$. Ta gọi một hệ sinh AXĐ là cặp $\alpha = (U, F)$ trong đó U là một tập hữu hạn, F là tập các luật sinh trên U.

Với mỗi hệ sinh AXĐ $\alpha = (U,F)$, ta xác định ánh xạ f_α: $SubSet(U) \rightarrow SubSet(U)$ như sau:

(i) $f_\alpha(X) \supseteq X$,

(ii) $\forall\, L \rightarrow R \in F, L \subseteq f_\alpha(X) \Rightarrow R \subseteq f_\alpha(X)$.

f_α được gọi là *ánh xạ cảm sinh* của hệ sinh AXĐ α, X là *vật*, $f_\alpha(X)$ là *ảnh* của *ánh xạ cảm sinh* f_α.

Dễ dàng nhận thấy, $f_\alpha(X)$ là *tập bao* (*nhỏ nhất*) của X trong hệ sinh AXĐ α.

Giả sử G là một họ các tập con đóng với phép giao của tập hữu hạn U. Cụ thể là,

$$G: (\forall X, Y \in G \Rightarrow X \cap Y \in G)$$

G được gọi là *giàn giao* trên U. Khi đó, G chứa duy nhất một họ con S sao cho mọi phần tử của G đều được biểu diễn qua giao của các phần tử trong S.

Cụ thể là, S là tập nhỏ nhất thỏa tính chất,

$$G = \{ X_1 \cap \ldots \cap X_k \mid k \geq 0, X_1, \ldots, X_k \in S \}$$

S được gọi là *tập sinh* của giàn giao G và được ký hiệu là $Gen(G)$.

Chú ý rằng, theo quy ước, giao của một họ rỗng các tập trong S chính là U, do đó mọi giàn giao đều chứa U và U không thuộc về $Gen(G)$.

<u>Định nghĩa 1.1</u> [8] Cho G là một *giàn giao* trên tập hữu hạn U. Ký hiệu $Coatom(G) = MAX(G \backslash \{U\})$. Các phần tử trong tập $Coatom(G)$ được gọi là *đối nguyên tử* của *giàn giao G*.

<u>Định nghĩa 1.2</u> [8] Cho $(M, \leq)$ là một tập hữu hạn có *thứ tự bộ phận*. Phần tử m trong M được gọi là *cực đại* nếu từ $x \geq m$ và $x \in M$, ta luôn có $x = m$. Ta ký hiệu $MAX(M)$ là tập các phần tử *cực đại* của M.

Ta nhận xét rằng, với mỗi phần tử x trong M luôn tồn tại phần tử m trong $MAX(M)$ thỏa mãn $x \leq m$.

Với mỗi họ các tập con của một tập hữu hạn U cho trước, ta xét thứ tự bộ phận $\subseteq$.

II. THU GỌN HỆ SINH AXĐ

Mục đích của thu gọn hệ sinh là giảm thiểu số luật sinh và giảm kích thước của mỗi luật sinh. Khi số lượng, kích thước các luật sinh được giảm thì không gian lưu trữ hệ sinh sẽ giảm, đồng thời độ phức tạp tính toán của các thuật toán sẽ được giảm theo.

Khác với các phép thu gọn tương đương các luật dẫn, thu gọn hệ sinh không đưa tập luật sinh về dạng tương đương nhưng vẫn bảo lưu được ảnh của AXĐ. Các khái niệm sau đây đã được trình bày một cách đầy đủ trong [7,8 9,10].

Cho hai hệ sinh AXĐ $\alpha = (U,F)$, $\beta = (V,G)$ và tập $M \subseteq U$. Ta nói hệ sinh AXĐ β *nhận được từ* hệ sinh AXĐ α *qua phép thu gọn theo tập M*, và ký hiệu là $\beta = \alpha\backslash M$, nếu sau khi loại bỏ mọi xuất hiện của các phần tử của M trong α thì thu được β.

Thao tác loại bỏ M được thực hiện trên hệ sinh AXĐ $\alpha = (U, F)$ để thu được hệ sinh AXĐ $\beta = (V,G)$ như sau:

1.Tính $V = U\backslash M$ có độ phức tạp $O(n)$ với $n = |U|$.

2.Với mỗi luật sinh $X \to Y$ trong F ta tạo một luật sinh $X\backslash M \to Y\backslash M$ cho G. Thủ tục này xác định tập các luật sinh được ký hiệu là $F\backslash M$. Để đơn giản ta ký hiệu $G = F\backslash M$ với ý nghĩa tập luật sinh G nhận được từ tập luật sinh F sau khi thực hiện thủ tục thu gọn như trên. Tính $F\backslash M$ đòi hỏi độ phức tạp $O(mn)$, với $m = |F|$.

Như vậy $\beta = \alpha\backslash M = (U\backslash M, F\backslash M)$ được thực hiện với độ phức tạp $O(mn)$, tức là tuyến tính theo chiều dài dữ liệu vào (của hệ sinh AXĐ α).

Sau khi thực hiện thủ tục $G = F\backslash M$, nếu:

- G chứa các *luật sinh tầm thường* (dạng $X \to Y$, $X \supseteq Y$) thì ta loại các luật sinh này khỏi G,
- G chứa các luật sinh trùng lặp thì ta lược bớt các luật sinh này.

<u>Thí dụ 2.1</u> Cho hệ sinh AXĐ $\alpha = (U, F)$, tập nền $U = ABCDEH$, $F = \{AE \to D, A \to DH, BC \to E, E \to BC\}$. Với $M = ADH$.

Hãy xác định $\beta = (V,G) = \alpha\backslash M$.

Ta có, $V = U\backslash ADH = ABCDEH\backslash ADH = BCE$,

$G = \{E \to \varnothing$ (loại), $\varnothing \to \varnothing$ (loại), $BC \to E, E \to BC\}$ $\equiv \{BC \to E, E \to BC\}$.

<u>Nhận xét</u>

Phép thu gọn thỏa tính *kết hợp*, cụ thể là nếu α là hệ sinh AXĐ trên tập U và X, Y là hai tập con rời nhau của U ta có $\alpha\backslash XY = (\alpha\backslash X)\backslash Y = (\alpha\backslash Y)\backslash X$.

Công thức biểu diễn AXĐ cảm sinh theo phép thu gọn hệ sinh được các tác giả trong [7] trình bày như sau:

Cho hệ sinh AXĐ $\alpha = (U,F)$ và hai tập con không giao nhau X và Y trong U. Khi đó: $f_\alpha(XY) = Xf_{\alpha\backslash X}(Y)$.

Từ công thức trên, trong [7] các tác giả đã đưa ra công thức tính ảnh cho tập $X \subseteq U$ trong một hệ sinh AXĐ $\alpha = (U, F)$ như sau: $f_\alpha(X) = X\, f_{\alpha\backslash X}(\varnothing)$.

Công thức tính ảnh cho một tập con được dùng làm cơ sở cho thuật toán tính ảnh của ánh xạ cảm sinh với thời gian tuyến tính theo kích thước của hệ sinh. Tư tưởng thuật toán là, để tính $f_\alpha(X)$ trong hệ sinh $\alpha = (U,F)$, ta thực hiện phép thu gọn $\beta = \alpha\backslash X$. Trong β, ta cần tính $f_\beta(\varnothing)$. Ta có thể nhận thấy rằng, $f_\beta(\varnothing) \neq \varnothing$ khi và chỉ khi hệ sinh β có luật sinh dạng $\varnothing \to R$ với $R \neq \varnothing$. Nếu có những luật như vậy, ta gom các vế phải R của chúng, đưa vào kết quả và lại thực hiện tiếp các phép rút gọn trên hệ sinh β. Quá trình kết thúc khi trong β không còn luật dạng $\varnothing \to R$, $R \neq \varnothing$.

<u>Thí dụ 2.2</u> Cho hệ sinh AXĐ $\alpha = (U, F)$, tập nền $U = ABCDEH$, $F = \{AE \to D, A \to DH, BC \to E, E \to BC\}$.

Tính: 1. $f_\alpha(AHE)$ và 2. $f_\alpha(E)$?

Theo hệ quả về công thức xác định ảnh cho một tập thì:

1. $f_\alpha(AHE) = AHE\, f_{\alpha\backslash AHE}(\varnothing)$; $G = F\backslash AHE = \{\varnothing \to D, BC \to \varnothing$ (loại), $\varnothing \to BC\} \equiv \{\varnothing \to BCD\}$.

Từ đây, ta tính được $f_{\alpha\backslash AHE}(\varnothing) = BCD$. Do đó, $f_\alpha(AHE) = AHEBCD = U$.

2. $f_\alpha(E) = E\, f_{\alpha\backslash E}(\varnothing)$; $G = F\backslash E = \{A \to D, A \to DH, BC \to \varnothing$ (loại), $\varnothing \to BC\} \equiv \{A \to DH, \varnothing \to BC\}$.

Từ đây, ta được $f_{\alpha\backslash E}(\varnothing) = BC$. Do đó $f_\alpha(E) = EBC$.

Trong các công trình [9,10], các tác giả cũng đã trình bày đầy đủ về các khái niệm cơ sở và phản cơ sở

của một hệ sinh AXĐ. Cụ thể như sau,

Tập con K của U được gọi là cơ sở của AXĐ f nếu thỏa mãn các tính chất sau đây:

(*i*) $f(K) = U$, *và*

(*ii*) $\forall X \subset K$: $f(X) \neq U$.

Ta ký hiệu *Base*(f) là tập các *cơ sở* của AXĐ f.

Một khái niệm đối ngẫu với *cơ sở* AXĐ là *phản cơ sở* AXĐ. Khái niệm đối ngẫu ở đây được hiểu với ý nghĩa *cơ sở* là tập con nhỏ nhất có ảnh là U, *phản cơ sở* là tập con lớn nhất có ảnh khác U. Khái niệm *phản cơ sở* AXĐ được phát biểu như sau,

Cho AXĐ f trên U. Tập con P của U được gọi là *phản cơ sở* của AXĐ f nếu thỏa mãn các tính chất:

(*i*) $f(P) \neq U$, *và*

(*ii*) $\forall A \in U \setminus P$: $f(PA) = U$.

Trong đó, theo truyền thống của lý thuyết cơ sở dữ liệu thì PA được hiểu là hợp của tập P với phần tử A.

Ta ký hiệu *Antibase*(f) là tập các *phản cơ sở* của AXĐ f. *Fix*(α) cũng được ký hiệu là tập toàn thể các điểm bất động của ánh xạ cảm sinh hệ sinh α.

Định lý sau sẽ trình bày mối tương quan giữa tập phản cơ sở của ánh xạ đóng và tập đối nguyên tử của giàn giao.

<u>Định lý 2.1</u> [9] Với mọi AXĐ f trên tập hữu hạn U, *Antibase*(f) = *Coatom*(f).

Ta gọi *cơ sở* (*phản cơ sở*) của hệ sinh là *cơ sở* (*phản cơ sở*) của ánh xạ cảm sinh của hệ sinh đó.

Với mỗi hệ sinh $\alpha = (U,F)$, ta ký hiệu *Antibase*(α) là tập các phản cơ sở của hệ sinh α.

<u>Định lý 2.2</u> [9] Cho hệ sinh AXĐ $\alpha = (U, F)$ và $\beta = (V, G)$. Biết $\beta = \alpha \backslash X$, $X, M \subseteq U$, $X \cap M = \varnothing$.

Khi đó,

1. $XM \in Fix(\alpha)$ khi và chỉ khi $M \in Fix(\beta)$.
2. $XM \in Gen(\alpha)$ khi và chỉ khi $M \in Gen(\beta)$.
3. $XM \in Coatom(\alpha)$ khi và chỉ khi $M \in Coatom(\beta)$.
4. $XM \in Antibase(\alpha)$ khi và chỉ khi $M \in Antibase(\beta)$

III. BIỂU DIỄN PHẢN CƠ SỞ HỆ SINH AXĐ

Trong các công trình nghiên cứu về phản cơ sở, chẳng hạn như [1, 8, 9], các tác giả đã phát biểu về mối tương quan giữa tập phản cơ sở AXĐ và tập đối nguyên tử của giàn giao qua định lý 2.1 cũng như chỉ ra được các tính chất của họ các tập đóng khi thực hiện phép thu gọn ở định lý 2.2. Tuy nhiên, việc xác định tập các đối tượng qua phép thu gọn thì chưa được các tác giả đề cập đến. Đây là các kết quả quan trọng để bài viết phát triển tiếp một dạng biểu diễn phản cơ sở được trình bày sau đây.

Từ thời điểm này trở đi, ta luôn giả thiết rằng mọi hệ sinh $\alpha = (U, F)$ đều được biểu diễn dưới dạng *thu gọn tự nhiên* với ý nghĩa như sau: tập U hữu hạn và không rỗng, tập luật sinh F không rỗng và thỏa mãn các tính chất:

1. F không chứa các luật sinh tầm thường:

$$\forall X,Y \subseteq U: X \supseteq Y \Rightarrow (X \rightarrow Y \notin F)$$

2. Hai vế trái và phải của mọi luật sinh trong F rời nhau (không giao nhau):

$$\forall f \in F: LS(f) \cap RS(f) = \varnothing$$

3. Các vế trái của mọi luật sinh trong F khác nhau đôi một:

$$\forall f, g \in F: LS(f) = LS(g) \Leftrightarrow f = g$$

Ta ký hiệu $MR(F)$ là tập các vế phải cực đại của F, $MR(F) = MAX\ \{RS(f) \mid f \in F\}$.

<u>Thí dụ 3.1</u> Hệ sinh AXĐ $\alpha = (U, F)$ có tập nền $U=ABCDEH$, $F = \{AE \rightarrow D, A \rightarrow DH, BC \rightarrow E, E \rightarrow BC\}$ có dạng *thu gọn tự nhiên*.

<u>Bổ đề 3.1</u> Cho hệ sinh $\alpha = (U, F)$, nếu $R \in MR(F)$ thì R là tập con của phản cơ sở nào đó của α khi và chỉ khi $f(R) \neq U$.

Chứng minh

($\Rightarrow$) Giả sử $P \in Antibase(\alpha)$, $R \in MR(F)$ và $R \subseteq P$. Theo giả thiết, do $R \subseteq P$ nên theo tính chất đồng biến của AXĐ, ta có $f(R) \subseteq f(P)$. Mặt khác, do P là phản cơ sở nên $f(P) \neq U$. Từ đó, ta suy ra $f(R) \neq U$.

($\Leftarrow$) Ngược lại, nếu $R \in MR(F)$ và $f(R) \neq U$. Theo tính chất điểm bất động thì $f(R)$ là một điểm bất động nên $f(R) \in Fix(f) \backslash \{U\}$ với f là ánh xạ cảm sinh của hệ sinh α. Suy ra, tồn tại $P \in MAX(Fix(f) \backslash \{U\})$ thỏa $f(R) \subseteq P$. Do tính phản xạ AXĐ, $R \subseteq f(R) \subseteq P$.

Vậy $R \subseteq P$ (1)

Mặt khác, theo định nghĩa 1.1 về tập *Coatom* và định lý 2.1 phát biểu về sự tương quan giữa tập phản cơ sở và tập đối nguyên tử của giàn giao, ta có, $Coatom(f) = MAX(Fix(f)\backslash\{U\}) = Antibase(f)$. Suy ra, $P \in Antibase(f)$. Vậy P là một phản cơ sở (2)

Từ (1) & (2), ta kết luận $R \subseteq P$, nghĩa là vế phải cực đại R phải chứa trong một phản cơ sở nào đó của hệ sinh α.

<u>Định lý 3.1</u> Mọi phản cơ sở của hệ sinh $\alpha = (U, F)$ đều biểu diễn được dưới dạng RM với R là vế phải cực đại không chứa cơ sở và M là phản cơ sở của hệ sinh $\beta = \alpha\backslash R$.

Chứng minh

Theo giả thiết thì R là vế phải cực đại không chứa cơ sở. Theo bổ đề 3.1, ta suy ra, R phải thuộc về một phản cơ sở P nào đó của hệ sinh α. Như vậy, ta đã chứng minh được mọi vế phải cực đại của tập luật sinh không chứa cơ sở thì luôn luôn thuộc về một phản cơ sở P nào đó của hệ sinh.

Đặt $M = P\backslash R$, ta suy ra $P = MR$. Áp dụng mệnh đề (4) trong định lý 2.2 về tính đóng khi thực hiện phép thu gọn hệ sinh đối với phản cơ sở, vì $P = MR$ là phản cơ sở của hệ sinh α và $M \cap R = \varnothing$ nên M cũng là phản cơ sở của hệ sinh $\beta = \alpha\backslash R$.

<u>Thí dụ 3.2</u> Cho hệ sinh AXĐ $\alpha = (U, F)$ với $U = ABCDEH$ và tập luật sinh $F = \{AE\rightarrow D, A\rightarrow C, E\rightarrow BC, EH\rightarrow A, AC\rightarrow EH, BD\rightarrow C\}$. Ta có $MR(F) = \{D, BC, A, EH\}$. Ta thấy, $f_\alpha(A) = U$. Vậy A không thuộc về bất kỳ phản cơ sở nào của hệ sinh α. Ngoài ra, $f_\alpha(BC) = BC \neq U$. Ta thu gọn α theo BC. Đặt $\beta = \alpha\backslash BC = (V,G)$, ta có $V = ABCDEH\backslash BC = ADEH$, $F = \{AE \rightarrow D, A \rightarrow \varnothing$ (loại), $E \rightarrow \varnothing$ (loại), $EH \rightarrow A$, $A \rightarrow EH$, $D \rightarrow \varnothing$ (loại) $\} = \{AE \rightarrow D, EH \rightarrow A, A \rightarrow EH\}$. Ta nhận thấy DH là phản cơ sở của β. Như vậy $BCDH$ là phản cơ sở của α, trong đó BC là một vế phải cực đại, DH là phản cơ sở của $\alpha\backslash BC$.

Bài toán xác định phản cơ sở của hệ sinh có độ phức tạp NP (Non-deterministic Polynomial). Việc xác định phản cơ sở hệ sinh theo kỹ thuật thu gọn vế phải cực đại của tập luật sinh sẽ cải thiện được hiệu quả tính toán hơn so với các thuật toán mà các tác giả khác đã trình bày do ta có thể thu gọn liên tiếp nhiều lần hệ sinh ban đầu cho đến khi nhận được một hệ sinh đơn giản mà việc xác định phản cơ sở của hệ sinh này được tìm thấy rất tự nhiên. Cụ thể là để xác định phản cơ sở P của một hệ sinh ban đầu, ta có thể thu gọn nhiều lần để nhận được $P=R_1R_2\ldots R_{i\ldots}R_kP_k$ với R_i ($i=1, 2, \ldots, k$) là vế phải cực đại của các hệ sinh trung gian sau khi thu gọn, P_k là phản cơ sở của hệ sinh sau khi thu gọn k lần.

IV. KẾT LUẬN

Việc nghiên cứu và vận dụng các ánh xạ đóng như một công cụ toán học cho phép thiết lập mối tương quan trong nhiều lĩnh vực như cơ sở dữ liệu, các hệ suy dẫn, logic, lý thuyết tập thô và tập mờ… Bài viết đã tổng quát hóa một số vấn đề lý thuyết của các hệ suy dẫn theo ngôn ngữ của ánh xạ đóng. Trên cơ sở đó đề xuất một dạng biểu diễn phản cơ sở của hệ sinh AXĐ theo vế phải cực đại của tập luật sinh kết hợp với phép thu gọn hệ sinh. Đóng góp này có thể giúp cho việc biểu diễn phản cơ sở trong một hệ suy dẫn trở nên đơn giản với hiệu quả tính toán tốt hơn. Các kết quả được ứng dụng trước hết trong lĩnh vực cơ sở dữ liệu và khai phá dữ liệu. Như ta biết, trong cơ sở dữ liệu quan hệ, ràng buộc dữ liệu là các phụ thuộc hàm. Toán tử lấy bao đóng của tập thuộc tính trong lược đồ quan hệ chính là một ánh xạ đóng, do đó các kết quả của ánh xạ đóng có thể vận dụng trực tiếp cho việc thiết kế các cơ sở dữ liệu quan hệ. Với các lớp phụ thuộc rộng hơn như phụ thuộc Boole dương, phụ thuộc Boole dương tổng quát, phụ thuộc sai khác, v.v.. thì lý thuyết ánh xạ đóng cũng được vận dụng rộng rãi trong việc xác định các phụ thuộc được suy dẫn từ một họ các phụ thuộc cho trước. Mặt khác, lý thuyết ánh xạ đóng có thể trợ giúp cho việc thu gọn tập các phụ thuộc nhằm giảm không gian lưu trữ và tăng tốc độ tính toán các thành phần cơ bản trong qui trình thiết kế

các lược đồ chuẩn hóa, như bao đóng, cơ sở và phản cơ sở.

Từ các nhận xét này, bài toán ứng dụng các khái niệm AXĐ vào việc biểu diễn các đối tượng của các hệ suy dẫn trong nghiên cứu, phát triển lý thuyết một số lĩnh vực của công nghệ thông tin chẳng hạn như lý thuyết cơ sở dữ liệu cũng được xem là hướng phát triển tiếp theo của chúng tôi trong thời gian tới.

TÀI LIỆU THAM KHẢO

[1] THI V. D., *Minimal Keys and Antikeys*, Acta Cybernetica 7 , 361-371, 1986

[2] DEMETROVICS, HO THUAN, NGUYEN XUAN HUY, LE VAN BAO, *Translation of Relation Schemes, Balanced Relation Schemes and the Problem of Key Representation*, J. Inf. Process. Cybern. EIK, 23(1987) 2/3, 81-97. Math. Reviews 88e:68022 68P15.

[3] BUROSCH G., DEMETROVICS J., AND KATONA G.O.H., *The Subset of Closure* as a Model of Changing Databases, Order 4, 1987, 127-142.

[4] HAI QUOC LE, SOMJIT ARCH-INT, HUY XUAN NGUYEN, NGAMNIJ ARCH-INT, *Association rule hiding in risk management for retail supply chain collaboration*, Computers in Industry 64 (2013), Elsevier B.V. 776–784.

[5] THURAISINGHAM, B. M., AND W. FORD, *1991, Issues on Design and Implementation of an Intelligent Database Inference Controller*, IEEE International Conference on Systems, Man, and Cybernetics, Charlottesville, VA, 1991.

[6] DAVEY, B.A.; PRIESTLEY, H. A. (2002), Introduction to Lattices and Order, Cambridge University Press, ISBN 978-0-521-78451-1

[7] NGUYỄN XUÂN HUY, LÊ THỊ MỸ HẠNH, *Thu gọn hệ sinh ánh xạ đóng*, Chuyên san các công trình nghiên cứu – triển khai Viễn thông và Công Nghệ Thông Tin, Số 15, 53-58, 12-2005.

[8] NGUYỄN XUÂN HUY, *Các phụ thuộc logic trong cơ sở dữ liệu*, Viện KH&CN VN, NXB Thống kê, 2006.

[9] NGUYỄN XUÂN HUY, LÊ THỊ MỸ HẠNH, HUỲNH MINH TRÍ , *Phản khóa của ánh xạ đóng qua phép thu gọn hệ sinh*, Kỷ yếu Hội thảo Khoa học Quốc gia lần thứ 3 "Nghiên cứu, ứng dụng và phát triển Công nghệ Thông tin và truyền thông", ICT.rda'06, 20-21/05/2006, NXB KHKT Hà nội, 2006, 1-6.

[10] NGUYỄN XUÂN HUY, LÊ THỊ MỸ HẠNH, HUỲNH MINH TRÍ, NGUYỄN ĐỨC VŨ, *Biểu diễn khóa của ánh xạ đóng*, Kỷ yếu Hội thảo FAIR: Nghiên cứu cơ bản và ứng dụng CNTT, NXB KHKT Hà Nội, 2006, 23-28.

Nhận bài ngày: 17/06/2013

SƠ LƯỢC VỀ TÁC GIẢ

BÙI ĐỨC MINH

Sinh ngày 24/3/1966.

Tốt nghiệp Đại học và Thạc sỹ ngành CNTT tại Trường Đại học Khoa học Tự nhiên, Đại học Quốc Gia TP.HCM năm 1998 và 2005. Hiện đang là nghiên cứu sinh tại Viện CNTT.

Hiện công tác tại Khoa CNTT, Trường Cao đẳng Giao thông Vận tải TP. HCM.

Hướng nghiên cứu: Cơ sở dữ liệu

Email: buiducminh@gmail.com

Mobile: 0903687898