# PHÂN HOẠCH HỆ SUY DẪN ỨNG DỤNG TRONG THIẾT KẾ CƠ SỞ DỮ LIỆU

**Trần Minh Cảnh*, Phạm Ngọc Thịnh**
*Trường Đại học Phú Yên*
*Ngày nhận bài: 24/8/2020; Ngày nhận đăng: 08/01/2021*

**Tóm tắt**

*Hệ suy dẫn được xây dựng trên ánh xạ đóng với các luật dẫn dạng $X \rightarrow Y$. Hệ suy dẫn có thể coi là một tổng quát hoá cho một lớp công cụ đảm bảo ngữ nghĩa cho các hệ cơ sở dữ liệu và tri thức. Kết quả chủ yếu của báo cáo là thuật toán phân hoạch hệ suy dẫn và định lý về biểu diễn cơ sở của hệ suy dẫn thông qua các cơ sở của các hệ thành phần, được ứng dụng vào việc thiết kế và chuẩn hóa cơ sở dữ liệu.*

**Từ khóa:** *hệ suy dẫn, phân hoạch, cơ sở dữ liệu*

## 1. Mở đầu

Việc quan trọng nhất khi thiết kế cơ sở dữ liệu quan hệ là ta phải chọn ra tập các lược đồ quan hệ tốt nhất dựa trên một số tiêu chí nào đó. Và để có được lựa chọn tốt, thì chúng ta cần đặc biệt quan tâm đến mối ràng buộc giữa các dữ liệu trong quan hệ, đó chính là hệ suy dẫn. Để hiểu hơn về câu hỏi tại sao phải thiết kế một cơ sở dữ liệu tốt, chúng ta hãy cùng tìm hiểu ví dụ sau:

Cho lược đồ SV_DIEM với các thuộc tính sau MSV(Mã sinh viên), TENSV(Tên sinh viên), NS(Năm sinh sinh viên), MMH(Mã môn học), TENMH(Tên môn học), DIEM(Điểm). Với các luật dẫn {MSV → TENSV,NS; MMH → TMH; MSV,MMH → DIEM} và bảng dữ liệu cụ thể:

| MSV | TENSV | NS | MMH | TENMH | DIEM |
|---|---|---|---|---|---|
| 1 | Danh | 1998 | M1 | CSLD | 7 |
| 1 | Danh | 1998 | M2 | JAVA | 8 |
| 2 | Hoa | 1999 | M1 | CSDL | 9 |
| 3 | Dung | 1998 | M3 | WEB | 7 |
| 2 | Hoa | 1998 | M3 | WEB | 8 |

Ta thấy, quan hệ trên thiết kế chưa tốt vì: Dư thừa dữ liệu (Redundancy): Thông tin về tên sinh viên, năm sinh sinh viên và tên môn học bị lặp lại nhiều lần. Không nhất quán (Inconsistency), Dị thường (Anomalies),… dữ liệu xảy ra khi tiến hành cập nhật dữ liệu, giả sử xoá đi môn học có mã môn học là M3 thì thông tin về sinh viên Dung cũng mất. Chính vì vậy, việc phân hoạch Hệ suy dẫn và ứng dụng vào việc thiết kế cơ sở dữ liệu là vấn đề cấp thiết, quan trọng.

Hệ suy dẫn được xây dựng trên ánh xạ đóng với các luật dẫn dạng $X \rightarrow Y$ trong đó $X$ và $Y$ là các tích logic. Hệ suy dẫn có thể coi là một tổng quát hoá cho một lớp công cụ đảm bảo ngữ nghĩa cho các hệ cơ sở dữ liệu và tri thức. Ứng dụng đầu tiên và được triển khai khá rộng rãi là trong lĩnh vực tính toán các đối tượng của lược đồ quan hệ trên phụ thuộc hàm như bao đóng của tập thuộc tính, khoá được trình bày trong *"Các phụ thuộc logic trong cơ sở dữ liệu"* (Nguyễn Xuân Huy, 2006), sau đó là các phụ thuộc mở rộng được trình bày trong *"Stanford Encyclopedia of Philosophy"* (Hofweber, Rev. 2011), "*Computers and Artificial Intelligence"* (Vu Duc Thi, V7,2,165-184, 1988), tiếp đến là các hệ sinh cho các ánh xạ đóng *"Chuyên san các công trình*

* Email: canhdhpy@gmail.com

*nghiên cứu – triển khai Viễn thông và Công Nghệ Thông Tin*, Số 15, 53-58, 12-2005 (Nguyễn Xuân Huy, Lê Thị Mỹ Hạnh), "*Chuyên san các công trình nghiên cứu, phát triển và ứng dụng CNTT-TT, Tạp chí Công nghệ Thông tin & truyền thông,* Tập V-1, Số 5(25), 15-21 (Bùi Đức Minh, Lương Nguyễn Hoàng Hoa, tháng 6/2011) và bản thể học (ontology). Báo cáo trình bày các khái niệm cơ sở và các bài toán đặc trưng của hệ suy dẫn. Kết quả chủ yếu của báo cáo là khái niệm phân hoạch một hệ suy dẫn thành các hệ con, thuật toán phân hoạch hệ suy dẫn và định lý về biểu diễn cơ sở của hệ suy dẫn thông qua các cơ sở của các hệ con. Báo cáo có cấu trúc như sau. Phần mở đầu trình bày các khái niệm cơ sở về hệ suy dẫn và các kí hiệu qui ước. Khái niệm về phân hoạch hệ suy dẫn và thuật toán phân hoạch được trình bày trong phần thứ hai. Phần cuối cùng của báo cáo dành cho việc thảo luận về hướng phát triển và ứng dụng.

## 2. Hệ suy dẫn

**Ký hiệu và quy ước:** Các ký hiệu và quy ước sau đây được sử dụng như trong *"Các phụ thuộc logic trong cơ sở dữ liệu"* (Nguyễn Xuân Huy, 2006). Các phần tử của tập hợp được ký hiệu bằng các chữ *Latin viết thường* 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*, không có các ký hiệu biểu diễn tập, chẳng hạn ta viết $X = abc$ thay vì viết $X = \{a,b,c\}$; $XY$ biểu diễn *hợp* của hai tập $X$ và $Y$, $X\cup Y$; *Phép trừ* hai tập $X$ và $Y$ được ký hiệu là $X\backslash Y$. *Tập vũ trụ* hay *tập nền M* được cho trước luôn luôn là *hữu hạn* và *khác trống;* Kí hiệu *Poset(X) là tập các tập con của X với thứ tự bộ phận bao hàm* ($\subseteq$). Với mỗi họ các tập con $\Psi$ của $M$ ta kí hiệu $\cap\Psi$ là *giao* của các tập con trong họ $\Psi$.

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

$X \oplus Y = XY$ (hợp của hai tập con $X$ và $Y$)

$$X \oplus \Im = \{XY \mid Y \in \Im\} \text{ và}$$

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

Ví dụ, với $M = abc$ ta có $Poset(M)= \{\varnothing, a, b, c, ab, ac, bc, abc\}$. Cho $\Re = \{ab, bc\}$, $\Im = \{b, ac\}$, $X = c$, $Y = ab$. Ta có

$\cap\Re = ab\cap bc = b$, $X\oplus Y = abc$, $X\oplus\Re = c\oplus\{ab,bc\}=\{abc, bc\}$, $\Re\oplus\Im=\{ab, bc\}\oplus\{b, ac\}=\{ab, abc, bc\}$.

Để ý rằng toán tử $\oplus$ xây dựng từ phép hợp các tập con nên có tính chất giao hoán và kết hợp.

**Định nghĩa:** *Hệ suy dẫn* là cặp $p = (M,R)$ với $M$ là tập hữu hạn, không rỗng các phần tử được gọi là *tập nền*, $R$ là tập các *luật dẫn dạng* $r: X\to Y$; $X,Y \subseteq M$; $XY \neq \varnothing$, $X$ là *vế trái* và $Y$ là *vế phải* của luật $r$. Với mỗi luật dẫn $r: X\to Y$ ta kí hiệu $Set(r) = XY$ là tập các phần tử xuất hiện trong luật $r$, ta cũng kí hiệu $Set(R)$ là hợp của các $Set(r)$, $r \in R$.

**Ví dụ:** Cho $M = abcd$, $R = \{ab\to d, c\to b\}$. Ta có, $Set(ab\to d) = abd$, $Set(R) = abcd$.

**Cơ chế suy dẫn** *(Hệ tiên đề Armstrong)*

Cho hệ suy dẫn $p = (M,R)$, cơ chế suy dẫn của hệ dựa trên ba tiên đề sau đây:

$\forall X, Y, Z \subseteq M$:

(D1) Tiên đề phản xạ: $X \supseteq Y \Rightarrow X\to Y$,

(D2) Tiên đề gia tăng: $X\to Y \Rightarrow XZ\to YZ$,

(D3)Tiên đề bắc cầu: $X\to Y, Y\to Z \Rightarrow X\to Z$.

**Dạng thu gọn tự nhiên**

Tập luật $R$ trong các hệ suy dẫn thường được cho dưới *dạng thu gọn tự nhiên* thỏa các tính chất sau đây (Nguyễn Xuân Huy, 2006):

- Các vế trái $X$ và phải $Y$ của mỗi luật

$X \rightarrow Y \in R$ không giao nhau, $X \cap Y = \emptyset$ và không đồng thời cùng rỗng, $XY \neq \emptyset$,

- Mọi vế trái trong *R* khác nhau đôi một,

Ngoài ra, với mọi hệ suy dẫn $p = (M, R)$ ta luôn giả thiết $Set(R) = M$. Cụ thể là nếu có phần tử nào trong tập nền *M* không xuất hiện trong bất kỳ luật nào của *R* thì ta loại phần tử đó khỏi tập nền.

**Các tính chất mở rộng của luật dẫn**

Ngoài ba tiên đề D1 – D3, dưới đây là một số tính chất mở rộng của luật dẫn, chúng được chứng minh dễ dàng từ 3 hệ tiên đề D1, D2, D3 (Nguyễn Xuân Huy, 2006).

$\forall X, Y, Z, V \subseteq M, \forall a \in M$:

(D4) Tính tựa bắc cầu: $X \rightarrow Y, YZ \rightarrow V \Rightarrow XZ \rightarrow V$

(D5) Mở rộng vế trái và thu hẹp vế phải: $X \rightarrow Y \Rightarrow XZ \rightarrow Y \backslash V$

(D6) Cộng tính đầy đủ: $X \rightarrow Y, Z \rightarrow V \Rightarrow XZ \rightarrow YV$

(D7) Mở rộng vế trái: $X \rightarrow Y \Rightarrow XZ \rightarrow Y$

(D8) Bộ phận ở vế phải: $X \rightarrow YZ \Rightarrow X \rightarrow Y$

Khái niệm hệ suy dẫn được sử dụng khá rộng rãi trong nhiều lĩnh vực. Trong toán học, có thể coi *M* là tập các đối tượng nguyên thuỷ như điểm, đường thẳng, mặt phẳng, ... *R* là tập các tiên đề, tức là các mệnh đề thể hiện tương quan giữa các đối tượng nói trên. Ví dụ,

*r*: *Qua hai điểm phân biệt có thể xác định duy nhất một đường thẳng.*

Luật này có thể biểu diễn như sau:

$r: A, B \rightarrow [A;B]$

Hệ suy dẫn được sử dụng như một thành phần cơ sở cho chứng minh tự động *"ACM Transactions on Database Systems"* (Shaoxu Song and Lei Chen, Vol. 9, No. 4, Article 39, Publication date:March 2011).

Trong sinh học, *M* có thể là tập các kiểu gen, *R* là tập các luật gắn kết các gen trong *M*.

Trong bản thể học, luật $r: A,B \rightarrow C$ có thể cho biết "*khái niệm C được xây dựng từ hai khái niệm A và B*" (Shaoxu Song and Lei Chen, Vol. 9, No. 4, Article 39, Publication date:March 2011).

**Định nghĩa:** Hai luật dẫn *r* và *v* được gọi là *giao nhau* nếu $Set(r) \cap Set(v) \neq \emptyset$.

**Ví dụ:** Trong hệ $p = (M,R)$ với $M = abcdeghkm$, $R = \{1: a \rightarrow d, 2: c \rightarrow g, 3: h \rightarrow m, 4: km \rightarrow h, 5: d \rightarrow be\}$, các luật 1 và 5, 3 và 4 là giao nhau. Thật vậy, $Set(1) \cap Set(5) = ad \cap bde = d$; $Set(3) \cap Set(4) = hm \cap khm = hm$.

**Cơ sở của hệ suy dẫn**

Cho hệ suy dẫn $p = (M,R)$. Tập con $K \subseteq M$ được gọi là *cơ sở của hệ suy dẫn p* nếu *K* thoả hai tính chất sau:

(i) *Tính đủ mạnh*: $K \rightarrow M$,

(ii) *Tính đủ nhỏ* (*không dư*): $\forall a \in K: \neg((K-\{a\}) \rightarrow M)$ (*hiểu là* $K-\{a\}$ *không suy ra được M*).

Nếu *K* thoả tính chất (i) thì tập *K* được gọi là *siêu cơ sở của p*.

Về ngữ nghĩa, cơ sở của một hệ suy dẫn chính là các *phần tử nhân đủ nhỏ* phát triển toàn bộ tập nền.

Ta kí hiệu *Base*(*p*) là *tập các cơ sở* của hệ suy dẫn *p*.

**Ví dụ:** Cho $M = abcd$, $R = \{ab \rightarrow d, c \rightarrow b, d \rightarrow ac\}$. Ta có $Base(p) = \{ab, ac, d\}$.

**Nhận xét**

- Cho hệ suy dẫn $p = (M,R)$ và tập con $X \subseteq M$. Khi đó $X \rightarrow M$ khi và chỉ khi *X* là siêu cơ sở.

- Mỗi hệ suy dẫn đều có cơ sở. Thật vậy, đặt $K = M$, từ tính chất D1 ta có ngay $K \rightarrow M$. Vậy *K* chính là một siêu cơ sở. Ta loại bỏ khỏi *K* những phần tử không ảnh hưởng đến bất biến $K \rightarrow M$ sẽ nhận được tập đủ nhỏ. Tập đó chính là cơ sở của *p*.

## 3. Phân hoạch hệ suy dẫn

**Quan hệ tương đương ~ trên các tập luật:**

Cho hệ suy dẫn $p = (M,R)$. Ta xây dựng quan hệ ~ trên tập luật dẫn $R$ như sau:

$\forall\ r, v \in R$: $r \sim v$ khi và chỉ khi tồn tại một dãy luật dẫn $t_1 = r, t_2,..., t_k = v$ trong $R$ thoả tính chất: hai luật kề nhau trong dãy là giao nhau.

**Mệnh đề:** Quan hệ ~ là quan hệ tương đương.

*Chứng minh*

Ta chứng minh quan hệ ~ thoả các tính chất *phản xạ, đối xứng* và *bắc cầu*. Thật vậy, với mọi luật $r$, $u$ và $v$ trong $R$ ta có

- Nếu $r$: $X \rightarrow Y$ thì $Set(r) \cap Set(r) = XY \cap XY = XY \neq \varnothing$ vì ta giả thiết $XY \neq \varnothing$. Vậy $r \sim r$.

- Nếu $r \sim v$ thì tồn tại một dãy luật dẫn $t_1 = r, t_2,..., t_k = v$ trong $R$ thoả tính chất: hai luật kề nhau trong dãy là giao nhau. Lật ngược dãy trên ta thu được dãy $t_k = v, t_{k-1},..., t_1 = r$ và hai luật kề nhau trong dãy này giao nhau. Điều này chứng tỏ $v \sim r$.

- Nếu $r \sim v$ và $v \sim u$ thì tồn tại hai dãy

$t_1 = r, t_2, .., t_k = v$; $t_i$ , $t_{i+1}$ giao nhau, $i = 1,2,...,k-1$, và

$t_k = v, t_{k+1}, .., t_{k+m} = u$, $t_j$ , $t_{j+1}$ giao nhau, $j = k,...,k+m-1$,

Nối hai dãy này lại ta thu được $r \sim u$ ■

Do quan hệ ~ trên $R$ là *tương đương* nên $R$ được *phân hoạch* thành $m$ lớp tương đương $R_1, R_2, ..., R_m$. Đặt $M_i = Set(R_i)$ ta thu được các hệ suy dẫn con $p_i = (M_i, R_i)$, $i = 1, 2, ..., m$. Ta gọi $p_1, p_2, ..., p_m$ là *phân hoạch* của hệ suy dẫn $p$ theo quan hệ tương đương ~ và viết $p = p_1 \mid p_2 \mid \dots \mid p_m$.

**Nhận xét**

- Cho phân hoạch của hệ suy dẫn $p = (M, R)$

$p = p_1 \mid p_2 \mid \dots \mid p_m$; $\quad p_i = (M_i, R_i)$, $i = 1, 2,..., m$.

Khi đó ta có các phân hoạch tương ứng của $M$ và $R$ như sau:

$M = M_1 \mid M_2 \mid \dots \mid M_m$,

$R = R_1 \mid R_2 \mid \dots \mid R_m$.

- Từ nhận xét trên ta suy ra rằng cơ chế suy dẫn trong mỗi hệ thành phần của phân hoạch hoạt động độc lập với nhau. Mỗi phần tử $a$ trong $M$ thuộc và chỉ thuộc đúng một tập nền thành phần $M_i$, mỗi luật $r$ trong $R$ thuộc và chỉ thuộc đúng một tập luật thành phần $R_j$.

**Bổ đề**

Nếu $p = p_1 \mid p_2$ thì $Base(p) = Base(p_1) \oplus Base(p_2)$.

*Chứng minh*

Giả sử $K = K_1K_2$; $K_1 \in Base(p_1)$; $K_2 \in Base(p_2)$. Ta cần chứng minh $K$ là cơ sở của $p$. Thật vậy, theo định nghĩa cơ sở ta có $K_1 \rightarrow M_1$, $K_2 \rightarrow M_2$. Theo cộng tính ta suy ra $K = K_1K_2 \rightarrow M_1M_2 = M$. Giả sử $a \in K$. Vì $K$ là phân hoạch với các thành phần (rời nhau), $K = K_1|K_2$ nên tồn tại một cơ sở $K_i$ của $p_i$ để: $a \in K_i$, $i = 1$ hoặc $i = 2$. Giả sử $a \in K_2$. Khi đó ta có, theo định nghĩa cơ sở trong $p_2$, $\neg((K_2\text{-}\{a\}) \rightarrow M_2)$. Gọi $V$ là tập con lớn nhất trong $M_2$ thoả $(K_2\text{-}\{a\}) \rightarrow V$, ta có $V \subset M_2$. Theo tiên đề phản xạ, do $K_2 \supset K_2\text{-}\{a\}$ nên $K_2 \rightarrow K_2 – \{a\}$. Vận dụng tiên đề bắc cầu cho $K_2 \rightarrow K_2 – \{a\}$ và $(K_2\text{-}\{a\}) \rightarrow V$ ta suy ra $K_2 \rightarrow V$. Áp dụng luật cộng tính cho $K_1 \rightarrow M_1$ và $K_2 \rightarrow V$ ta thu được $K = K_1K_2 \rightarrow M_1V \neq M$, mâu thuẫn với $K \rightarrow M$. Trường hợp $a \in K_1$ được lập luận hoàn toàn tương tự. Vậy $K$ là cơ sở của $p$.

Đảo lại, giả sử $K = P|Q$ là một cơ sở của $p$, trong đó $P \subseteq M_1$ và $Q \subseteq M_2$. Gọi $V$ và $W$ là các tập con lớn nhất thoả $P \rightarrow V$, $Q \rightarrow W$. Do các hệ con hoạt động độc lập nhau nên theo cộng tính ta có $K = PQ \rightarrow VW = M = M_1M_2$. Từ đây suy ra $P$ là siêu cơ sở của $p_1$ và $Q$ là siêu cơ sở của $p_2$. Nếu $P$ hoặc $Q$ có phần tử dư thừa thì phần tử đó cũng là dư thừa trong $K$, điều này vi phạm định nghĩa siêu cơ sở. Vậy $P$ phải là cơ sở của $p_1$ và $Q$

là cơ sở của $p_2$ ■

**Định lý:** Nếu $p = p_1 \mid p_2 \mid \dots \mid p_m$ , $p_i = (V_i, R_i)$, $i = 1, 2, ..., m$, thì

$$Base(p) = Base(p_1) \oplus Base(p_2) \oplus ... \oplus Base(p_m).$$

*Chứng minh*

Ta chứng minh định lý bằng phương pháp qui nạp toán học theo $m$ là số thành phần của phân hoạch.

Cơ sở: $m = 1$ là hiển nhiên vì phân hoạch có duy nhất một thành phần là $p$.

Giả thiết qui nạp: Giả sử với phân hoạch

$p = p_1 \mid p_2 \mid \dots \mid p_m$ ; $p_i = (V_i, R_i)$, $i = 1, 2,..., m$ ta có

$$Base(p) = Base(p_1) \oplus Base(p_2) \oplus ... \oplus Base(p_m).$$

Ta cần chứng minh với phân hoạch

$p = p_1 \mid p_2 \mid \dots \mid p_m \mid p_{m+1}$ ; $p_i = (V_i, R_i)$, $i = 1, 2,..., m+1$ ta phải có

$$Base(p) = Base(p_1) \oplus Base(p_2) \oplus ... \oplus Base(p_m) \oplus Base(p_{m+1}).$$

Gộp hai hệ con cuối cùng là $p_m$ và $p_{m+1}$ ta thu được hệ con $p' = (M', R')$ trong đó $M' = M_mM_{m+1}$.và $R' = R_mR_{m+1}$. Theo bổ đề trên, $Base(p') = Base(p_m) \oplus Base(p_{m+1})$. Sau khi gộp ta thu được phân hoạch với $m$ thành phần là

$p = p_1 \mid p_2 \mid \dots \mid p_{m-1} \mid p'$ .

Theo giả thiết qui nạp ta có ngay

$$Base(p) = Base(p_1) \oplus Base(p_2) \oplus ... \oplus Base(p_{m-1}) \oplus Base(p').$$

Do toán tử $\oplus$ có tính kết hợp nên

$$Base(p) = Base(p_1) \oplus Base(p_2) \oplus ... \oplus Base(p_{m-1}) \oplus (Base(p_m) \oplus Base(p_{m+1})) =$$

$$= Base(p_1) \oplus Base(p_2) \oplus ... \oplus Base(p_{m-1}) \oplus Base(p_m) \oplus Base(p_{m+1})$$ ■

**Thuật toán phân hoạch hệ suy dẫn**

**Algorithm Partition**

**Input:** Hệ suy dẫn $p = (M, R)$ gồm $n$ phần tử trong tập nền $M$, $m$ luật trong $R$.

**Output:** Một phân hoạch của tập luật $R$.

**Method**

Bước 1. *Xây dựng đồ thị vô hướng* $G = (V, E)$ *với tập đỉnh là các số hiệu luật* {1, 2, ..., $m$}, cạnh $(i,j)$ thuộc *E khi và chỉ khi hai luật i và j giao nhau. Pha 1 có độ phức tạp tính toán* $O(n.m)$.

Bước 2. *Liệt kê các thành phần liên thông của đồ thị G. Mỗi thành phần liên thông sẽ là một lớp tương đương trong phân hoạch của tập luật R. Pha 2 có thể được thực hiện theo kỹ thuật Find-Union với độ phức tạp tính toán* $O(m^2)$.

**end Partition.**

**Ví dụ**

Ta minh hoạ hoạt động của thuật toán Partition qua thí dụ sau: Cho $p = (M,R)$ với $M = abcdeghkm$, $R = \{1: a \to d, 2: c \to g, 3: h \to m, 4: km \to h, 5: d \to be\}$.

Sau bước 1 ta thu được đồ thị $G = (V, E)$ với $V = \{1, 2, 3, 4, 5\}$ và $E = \{(1, 5), (3, 4)\}$.

Bước 2: Đồ thị $G$ có 3 thành phần liên thông là $R_1 = \{1, 5\}$, $R_2 = \{2\}$, $R_3 = \{3, 4\}$.

Tổng hợp lại, ta thu được kết quả sau:

$p_1 = (M_1, R_1)$, $M_1 = abde$, $R_1 = \{1: a \to d, 5: d \to be\}$, $Base(p_1) = \{a\}$.

$p_2 = (M_2, R_2)$, $M_2 = cg$, $R_2 = \{2: c \to g\}$, $Base(p_2) = \{c\}$.

$p_3 = (M_3, R_3)$, $M_3 = hkm$, $R_3 = \{3: h \to m, 4: km \to h\}$, $Base(p_3) = \{kh, km\}$.

Vậy $Base(p) = Base(p_1) \oplus Base(p_2) \oplus Base(p_3) = \{a\} \oplus \{c\} \oplus \{kh, km\} = \{ackh, ackm\}$.

## 4. Kết luận

Có thể vận dụng các kết quả trên để phân hoạch các lược đồ quan hệ ban đầu thành các lược đồ con trong việc thiết kế các cơ sở dữ liệu. Sau khi tìm khoá của mỗi lược đồ con ta tổng hợp lại để thu được các khoá của lược đồ ban đầu. Bài toán tìm các khoá được vận dụng rộng rãi trong thiết kế cơ sở dữ liệu chuẩn hoá.

Trong ontology cho giáo dục và đào tạo và các hệ quản trị khác khái niệm phân hoạch được sử dụng để xác định tập khái niệm và đối tượng cơ sở là những khái niệm và đối tượng mà nắm vững chúng sẽ dễ dàng phát triển sang các khái niệm và đối tượng khác❑

**TÀI LIỆU THAM KHẢO**

Hofweber, T. Logic and Ontology. In Zalta, Edward N (ed.) *Stanford Encyclopedia of Philosophy*. http://plato.stanford.edu/entries/logic-ontology, 2004, Rev. 2011.

Shaoxu Song and Lei Chen, Differential Dependencies: Reasoning and Discovery, *ACM Transactions on Database Systems*, Vol. 9, No. 4, Article 39, Publication date: March 2011.

Vu Duc Thi. (1988). Logical dependencies and irredundant relations, *Computers and Artificial Intelligence*, V 7, 2, 165-184, 1988.

Nguyễn Xuân Huy, Lê Thị Mỹ Hạnh. (2005). 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.

Nguyễn Xuân Huy. (2006). *Các phụ thuộc logic trong cơ sở dữ liệu*, Viện KH&CN VN, NXB Thống kê.

Bùi Đức Minh, Lương Nguyễn Hoàng Hoa. (2011). Hệ sinh cân bằng và bài toán biểu diễn sơ sở hệ sinh ánh xạ đóng, *Chuyên san các công trình nghiên cứu, phát triển và ứng dụng CNTT-TT, Tạp chí Công nghệ Thông tin & truyền thông,* Tập V-1, Số 5(25), 15-21.

---

## The partition of inference system applied in designing database

**Tran Minh Canh*, Pham Ngoc Thinh**

*Phu Yen University*

**Email: canhdhpy@gmail.com*

*Received: August 24, 2020; Accepted: January 08, 2021*

**Abstract**

*The inference system, built on a closed mapping with the rules of conduction $X \rightarrow Y$, can be considered a generalization for a tool class ensuring the semantics for database and knowledge systems. The main result of this report is the algorithm for partitions of the inference system and its theorem of the basic representation through the bases of the component systems, which can be applied to the design and standardization of the database.*

**Keywords**: *inference system, partition, database*