

PHỤ THUỘC BOOLE DƯƠNG THEO NHÓM BỘ

NGUYỄN XUÂN HUY, NGUYỄN XUÂN HOÀNG, VŨ MẠNH XUÂN, NGUYỄN VĂN TẢO

Abstract. This paper introduces a new type of dependencies in relational databases, named tuple grouping Boolean dependencies. This concept is a generalization of the Boolean dependencies with a set of tuples. We discuss in what conditions the following three types of consequences are equivalent for the tuple grouping Boolean dependencies:

- + consequences in the world of all relations
- + consequences in the world of k -tuple relations
- + consequences of propositional logic.

Tóm tắt. Bài báo giới thiệu khái niệm phụ thuộc Boole dương theo nhóm bộ như là một mở rộng tự nhiên của lớp phụ thuộc Boole dương tổng quát, trong đó thay vì xét các cặp bộ trong mỗi quan hệ ta xét một nhóm k bộ. Chứng minh rằng đối với phụ thuộc Boole dương theo nhóm bộ, ba loại suy dẫn sau đây là tương đương: suy dẫn theo các quan hệ, suy dẫn theo các quan hệ có không quá k bộ và suy dẫn theo logic mệnh đề.

1. MỞ ĐẦU

Năm 1970 Codd [3] giới thiệu mô hình cơ sở dữ liệu quan hệ và khái niệm phụ thuộc hàm để phản ánh ngữ nghĩa của dữ liệu trong thế giới thực. Một trong những kết quả chủ yếu của lý thuyết phụ thuộc hàm được thể hiện qua định lý tương đương khẳng định sự tương đương giữa ba loại suy diễn: *suy diễn logic, suy diễn theo quan hệ và suy diễn theo các quan hệ có không quá hai bộ*. Sử dụng định lý tương đương chúng ta có thể giải bài toán thành viên, xác định xem một ràng buộc có được suy dẫn ra từ một tập các ràng buộc cho trước hay không thông qua dãy biến đổi hình thức các công thức Boole. Các nghiên cứu của Nguyễn Xuân Huy và các cộng sự đã chỉ ra định lý tương đương có hiệu lực đối với một lớp khá rộng các phụ thuộc là phụ thuộc hàm và các biến thể của nó (phụ thuộc mạnh, phụ thuộc yếu, phụ thuộc đối ngẫu), phụ thuộc Boole dương, phụ thuộc Boole dương tổng quát, phụ thuộc Boole dương đa trị [1, 4, 6, 11].

Bài viết này tiếp tục phát triển lớp phụ thuộc Boole dương tổng quát thông qua việc xác định một ánh xạ từ một nhóm k bộ sang không gian Boole $\{0, 1\}^n$. Các phụ thuộc mới này được gọi là phụ thuộc Boole dương theo nhóm bộ.

Các khái niệm cơ bản liên quan đến cơ sở dữ liệu quan hệ được trình bày chi tiết trong [3, 5, 10].

1.1. Quan hệ

Cho tập thuộc tính $U = \{A_1, A_2, \dots, A_n\}$. Với mỗi thuộc tính A_i ta đặt tương ứng một tập trị d_i chứa ít nhất là hai phần tử, $1 \leq i \leq n$. Ký hiệu D là hợp của các d_i , $1 \leq i \leq n$. Quan hệ R trên tập thuộc tính U là tập hữu hạn các ánh xạ $t: U \rightarrow D$ thỏa tính chất

$$(\forall t \in R) (\forall A_i \in U) : t(A_i) \in d_i$$

Mỗi phần tử của quan hệ được gọi là một bộ của quan hệ đó.

Ta luôn luôn giả thiết là cho trước một tập hữu hạn gồm n thuộc tính U và mọi khái niệm đều xây dựng trên tập thuộc tính ban đầu này.

Ta ký hiệu $REL(U)$ là tập các quan hệ với tập thuộc tính U . $REL_k(U)$ là tập các quan hệ với tập thuộc tính U và chứa không quá k bộ, với k là một số tự nhiên lớn hơn 1.

Cho một bộ t trong quan hệ R và tập thuộc tính X . Ta ký hiệu $t(X)$ là hạn chế của ánh xạ t trên tập X .

1.2. Phụ thuộc hàm

Một phụ thuộc hàm là một phát biểu dạng $X \rightarrow Y$, trong đó $X, Y \subseteq U$. Ta nói quan hệ R thỏa phụ thuộc hàm $f : X \rightarrow Y$, và ký hiệu $R(f)$, nếu $(\forall u, v \in R, u(X) = v(X)) \Rightarrow u(Y) = v(Y)$.

Ta nói quan hệ R thỏa tập phụ thuộc hàm F , và ký hiệu $R(F)$, nếu $\forall f \in F : R(f)$

1.3. Các công thức Boole

Ta đặt tương ứng mỗi thuộc tính trong U với một biến logic biến thiên trong miền $B = \{0, 1\}$. Ký hiệu $L(U)$ là tập các công thức logic có thể thành lập từ các biến trong U liên kết với nhau qua các phép toán logic $\vee, \wedge, \neg, \rightarrow$ và các hằng logic true (1), false (0).

Mỗi vector $v = (v_1, \dots, v_n)$ trong không gian $B^n = B \times B \times \dots \times B$ được gọi là một phép gán trị.

Khi đó với mỗi công thức Boole $f \in L(U)$ ta có $f(v) = f(v_1, \dots, v_n)$ là trị của công thức f đối với phép gán trị v .

Ta gọi công thức $f : X \rightarrow Y$ là một công thức suy dẫn nếu X và Y có dạng tích (hội).

Ta quan tâm hai phép gán trị đặc biệt là phép gán trị đơn vị, $e = (1, 1, \dots, 1)$ và phép gán trị không, $z = (0, 0, \dots, 0)$.

Với mỗi tập hữu hạn các công thức Boole, $F = \{f_1, f_2, \dots, f_m\}$ trong $L(U)$ ta xem F như là một công thức dạng $F = f_1 \wedge f_2 \wedge \dots \wedge f_m$. Khi đó với mỗi phép gán trị v , giá trị chân lý của công thức F sẽ được tính là $F(v) = f_1(v) \wedge f_2(v) \wedge \dots \wedge f_m(v)$.

Với mỗi công thức f trên U , bảng chân lý của f , ký hiệu là T_f , là tập các phép gán trị v sao cho $f(v)$ nhận giá trị 1, $T_f = \{v \in B^n | f(v) = 1\}$. Khi đó *bảng chân lý* T_F của tập hữu hạn các công thức F trên U chính là giao của các bảng chân lý của mỗi công thức thành viên trong F . Ta có, $v \in T_F$ khi và chỉ khi $(\forall f \in F)(f(v) = 1)$.

1.4. Công thức Boole dương

Công thức $f \in L(U)$ được gọi là công thức Boole dương nếu $f(e) = 1$, trong đó e là phép gán trị đơn vị, $e = (1, 1, \dots, 1)$.

1.5. Suy dẫn logic

Cho f và g là hai công thức. Ta nói công thức f dẫn ra được công thức g và ký hiệu $f \vdash g$, nếu $T_f \subseteq T_g$. Ta nói f và g là hai công thức tương đương, ký hiệu $f \equiv g$, nếu $T_f = T_g$.

Với hai tập các công thức F, G trong $L(U)$, ta nói F dẫn ra được G , ký hiệu $F \vdash G$, nếu $T_F \subseteq T_G$. Ta nói F và G là tương đương, ký hiệu $F \equiv G$, nếu $T_F = T_G$.

1.6. Suy dẫn theo quan hệ

Cho tập phụ thuộc hàm F và một phụ thuộc hàm f . Ta nói F dẫn ra được f theo quan hệ, và ký hiệu $F \models f$ nếu $(\forall R \in \text{REL}(U), R(F)) \Rightarrow R(f)$.

1.7. Suy dẫn theo quan hệ có không quá k bộ

Cho tập phụ thuộc hàm F và một phụ thuộc hàm f . Ta nói F dẫn ra được f theo quan hệ có không quá k bộ, ký hiệu $F \models_k f$ nếu $(\forall R \in \text{REL}_k(U), R(F)) \Rightarrow R(f)$.

Mỗi phụ thuộc hàm $f : X \rightarrow Y$ có thể đặt tương ứng với một công thức của đại số mệnh đề theo cách xây dựng như trên.

1.8. Bảng chân lý của quan hệ ([2, 6])

Cho quan hệ R trên tập thuộc tính U . Ta xét ánh xạ $E : R^2 \rightarrow U^n$ như sau:

Giả sử $u, v \in R, u = (u_1, u_2, \dots, u_n)$ và $v = (v_1, v_2, \dots, v_n)$, ta đặt

$E(u, v) = t = (t_1, t_2, \dots, t_n)$, trong đó $t_i = 1$ nếu $u_i = v_i$ và $t_i = 0$ nếu $u_i \neq v_i$.

Ta lập bảng $T_R \subseteq B^n, T_R = \{E(u, v) | u, v \in R\}$ và gọi là bảng chân lý của quan hệ R .

Từ định nghĩa trên ta suy ra: Quan hệ R thỏa tập phụ thuộc hàm F khi và chỉ khi $T_R \subseteq T_F$.

1.9. Định lý tương đương ([9])

Ba mệnh đề sau là tương đương

- (i) $F \vdash f$ (suy dẫn logic).
- (ii) $F \models f$ (suy dẫn theo quan hệ).
- (iii) $F \models_2 f$ (suy dẫn theo quan hệ có không quá 2 bộ).

2. PHỤ THUỘC BOOLE DƯƠNG THEO NHÓM BỘ

2.1. Ý tưởng

Mở rộng phụ thuộc Boole dương tổng quát bằng cách xây dựng bảng chân lý của quan hệ theo nhóm k bộ.

Cho k là một số tự nhiên, $k > 1$. Với mỗi miền trị d_i của thuộc tính A_i ta xét ánh xạ $\beta_i : d_i^k \rightarrow \{0, 1\}$ thỏa ba tính chất sau:

- a. Tính phản xạ: $\forall x \in d_i^k : \beta_i(x) = 1$, nếu mọi thành phần trong x giống nhau.
- b. Tính đối xứng: $\forall x \in d_i^k : \beta_i(x) = \beta_i(x')$, trong đó x' là một hoán vị của x .
- c. Tính bộ phận: $\exists x \in d_i^k : \beta_i(x) = 0$.

2.2. Bảng chân lý của quan hệ

Giả sử các ánh xạ β_i đã được xác định trên mỗi miền trị d_i^k của các thuộc tính A_i trong tập U và R là một quan hệ trên U . Với k bộ u_1, u_2, \dots, u_k tùy ý trong quan hệ R ta định nghĩa $\beta(u_1, u_2, \dots, u_k)$ là phép gán trị: $\beta(u_1, u_2, \dots, u_k) =$

$$(\beta_1(u_1(A_1), u_2(A_1), \dots, u_k(A_1)), (\beta_2(u_1(A_2), u_2(A_2), \dots, u_k(A_2)), \dots, \beta_n(u_1(A_n), u_2(A_n), \dots, u_k(A_n))))$$

Khi đó bảng chân lý của quan hệ R được định nghĩa là:

$$T_R = \{\beta(u_1, u_2, \dots, u_k) \mid u_1, u_2, \dots, u_k \in R(U)\}$$

Ví dụ 1. Quan hệ R trên $U = \{A_1, A_2, A_3, A_4\}$ cho bởi bảng bên:

Với $k = 4$ và mỗi $i (1 \leq i \leq 4)$ ta định nghĩa $\beta_i : d_i^4 \rightarrow \{0, 1\}$,

$\beta_i(x) = 1$ nếu x có ít nhất hai thành phần giống nhau,

$\beta_i(x) = 0$ trong các trường hợp khác.

Khi đó $\beta(u_1, u_2, u_3, u_4) = (0, 1, 0, 1)$,

$\beta(u_2, u_3, u_4, u_5) = (0, 1, 1, 1)$.

	A_1	A_2	A_3	A_4
u_1	1	t	4	d
u_2	2	f	8	a
u_3	3	t	2	d
u_4	4	f	6	b
u_5	5	t	6	b
u_6	6	f	4	d

2.3. Phụ thuộc Boole dương theo nhóm bộ

Ta gọi một phụ thuộc Boole dương theo nhóm bộ là một phát biểu dưới dạng một công thức Boole dương. Ta nói quan hệ $R \in \text{REL}(U)$ thỏa phụ thuộc Boole dương theo nhóm bộ f , hay $R(f)$, nếu $T_R \subseteq T_f$.

2.4. Định lý tương đương

Cho tập phụ thuộc Boole dương theo nhóm bộ F và một phụ thuộc Boole dương theo nhóm bộ f trên tập thuộc tính U . Ba mệnh đề sau là tương đương:

- (i) $F \vdash f$ (suy dẫn logic);
- (ii) $F \models f$ (suy dẫn theo quan hệ);
- (iii) $F \models_k f$ (suy dẫn theo quan hệ có không quá k bộ).

Chứng minh.

(i) \Rightarrow (ii) Giả sử $F \vdash f$ và $R(F)$. Ta có, vì $F \vdash f$ nên $T_F \subseteq T_f$. Vì $R(F)$ nên $T_R \subseteq T_F$. Từ đó, theo tính chất bắc cầu ta suy ra $T_R \subseteq T_f$, tức là $R(f)$ (điều phải chứng minh).

(ii) \Rightarrow (iii) Hiển nhiên.

(iii) \Rightarrow (i) Giả sử $F \models_k f$ nghĩa là với mọi quan hệ không quá k bộ R ta có $R(F) \Rightarrow R(f)$. Ta cần chứng minh $F \vdash f$, nghĩa là $T_F \subseteq T_f$.

Với mỗi $t = (t_1, t_2, \dots, t_n) \in T_F$ ta xây dựng quan hệ R có k bộ như sau. Theo tính chất bộ phận (c) của các ánh xạ $\beta_i, \exists x^{(i)} \in d_i^k : \beta_i(x^{(i)}) = 0, i = 1, 2, \dots, n$. Với mỗi miền trị d_i của các thuộc tính trong $U, i = 1, 2, \dots, n$ ta lấy một trị a_i tùy ý trong vector $x^{(i)}$ nói trên.

Nếu $t_i = 1$ ta điền vào cột A_i, k giá trị bằng nhau và bằng a_i . Nếu $t_i = 0$ ta điền vào cột A_i chính vector $x^{(i)}$.

Theo cách xây dựng R ta có $T_R = \{t\} \subseteq T_F$. Quan hệ R có k bộ và thỏa tập phụ thuộc Boole dương theo nhóm bộ F , theo giả thiết R sẽ thỏa f , tức là $T_R = \{t\} \subseteq T_f$, hay $t \in T_f$. ■

2.5. Nhận xét

Nếu $k = 2$ ta xét các ánh xạ $\beta_i : d_i^2 \rightarrow \{0, 1\}$ như sau

$$(\forall a, b \in d_i) : \beta_i(a, b) = 1 \iff a = b,$$

ta thu được phụ thuộc Boole dương tổng quát [6].

Với $k = 2$ và các ánh xạ β_i là tổng quát (thỏa các tính chất a, b, c) ta sẽ thu được phụ thuộc Boole dương đa trị [4, 1].

2.6 Chính lý

Trong tài liệu [7], tiểu mục 5.1, do sơ xuất của nhóm tác giả nên tính chất (i) đã bị in sai. Nhóm tác giả rất mong được thứ lỗi và được thay bằng tính chất (a) của tiểu mục 2.1 trong bài này.

TÀI LIỆU THAM KHẢO

- [1] Berman J. and Blok W.J., Generalized Boolean dependencies, *Abstracts of AMS* **6** (1985) 163.
- [2] Berman J. and Blok W.J., Positive Boolean dependencies, *Inf. Processing Letters* **27** (1988) 147–150.
- [3] Codd E. F., A Relational Model of Data for Large Shared Data Banks, *CACM* **13** (6) (1970) 377–387.
- [4] Le Duc Minh, Vu Ngoc Loan, Nguyen Xuan Huy, Some results concerning with the class of multivalued positive Boolean dependencies in the relational data model in context of fuzzy semantics, *Proceedings of VJFUSSSY '98: Vietnam - Japan bilateral Symposium on Fussy Systems and Applications*, Ha Long Bay, Vietnam, 30th September - 2nd October, (1998) (378–382).
- [5] Maier D., *The Theory of Relational Databases*, Computer Science Press, 1983.
- [6] Nguyễn Xuân Huy, Lê Thị Thanh, Generalized Positive Boolean Dependencies, *J. Inform. Process Cybernet.*, EIK **28** (6) (1992) 363–370.
- [7] Nguyễn Xuân Huy, Đàm Gia Mạnh, Lê Kim Loan, Hòa Quang Dự, Tổng quan về các phụ thuộc logic trong cơ sở dữ liệu quan hệ, *Toàn văn báo cáo khoa học Hội nghị Kỷ niệm 25 năm thành lập Viện Công nghệ Thông tin*, Hà Nội, 24-25/12/2001, 277–284.
- [8] Post E. L., The two valued interative systems of mathematical logic, *Annals of Math. Studies* **5** Princeton University, (1991).
- [9] Sagiv Y., Delobel C., Parker D.S., Fagin R., An equivalence between relational database dependencies and a fragment of propositional logic, *J. ACM* **28**, 435–453; Corrigendum *J. ACM* **34** (1987) 1016–1018.
- [10] Ullman J., *Principles of Database and Knowledge-Base Systems* Vol.1&2, Computer Science Press, (1986).
- [11] Vũ Ngọc Loan, Lê Đức Minh, The class of extended functional dependencies in the relational data model, *VNU Journal of Science: Nat. Sci.*, Vol. XIII (2) (1997).

Nhận bài ngày 24 - 4 - 2002