10(4):11-15 Tạp chí Tin học và Điều khiển học 12-1994

VỀ ĐỊNH LÝ TƯƠNG ĐƯƠNG TRONG LỚP CÁC PHỤ THUỘC BOOL DƯƠNG ĐA TRỊ

Vũ Ngọc Loãn và Nguyễn Xuân Huy

Viện Công nghệ thông tin

I. Giới thiệu

Trong lý thuyết cơ sở dữ liệu quan hệ, việc nghiên cứu các loại quan hệ phụ thuộc dữ liệu có ý nghĩa thực tiễn cho việc đảm bảo tính chất nhất quán dữ liệu. Các phụ thuộc Bool được giới thiệu trong [6] là một trong các loại phụ thuộc dữ liệu được nhiều người quan tâm phát triển. Tiếp đó các bài viết [1,2] đã đề cập đến một lớp con khá rộng của nó là lớp các phụ thuộc Boll dương, lớp các phụ thuộc Bool dương tổng quát,... Một trong những điều được quan tâm nhiều là định lý tương đương về tính dẫn được trong lớp các quan hệ có số bộ tùy ý cũng như trong lớp các quan hệ chỉ bao gồm hai bộ hay trong logic mệnh đề khi xét đến các lớp phụ thuộc Bool khác nhau. Trong [2], định lý này đã được chỉ ra cho các lớp phụ thuộc Bool dương dựa trên các ánh xạ bằng. Trong [5], định lý tương đương đã được phát biểu ở dạng mở rộng hơn cho lớp các phụ thuộc Bool dương tổng quát. Ở đây do việc đưa vào logic đa trị thay cho kiểu logic 2 trị và việc mở rộng ánh xạ α cho nên định lý tương đương sẽ được phát biểu trong dạng tổng quát hơn so với các trường hợp trên.

Bài viết gồm bốn phần. Phần đầu giới thiệu chung. Phần hai trình bầy một số định nghiã và các kết quả cơ sở chuẩn bị cho phần sau. Phần thứ ba phát biểu và chứng minh định lý tương đương cho lớp phụ thuộc Bool dương đa trị. Phần còn lại đề xuất một vài hướng nghiên cứu.

2. Các định nghiã cơ bản

Giả sử $U = \{A_1, ..., A_n\}$ là tập các thuộc tính. Với mỗi A_i, $1 \leq i \leq n$ có một tập d_i nào đó gồm ít nhất hai phần tử được gọi là miền trị của thuộc tính đó. Với $A \in U$ miền trị của A cũng được ký hiệu là dom(A).

Một tập con R hữu hạn của tích $d_1 \times d_2 \times ... \times d_n$ được gọi là một quan hệ trên U. Mỗi $t \in R$ được gọi là một bộ. Giả sử $t \in R$, $A \in U$ khi đó kí hiệu $t.A$ là giá trị t đối với thuộc tính A. Đặt $t.X$ là tập $\{t.A | A \in X\}$. với $X \subseteq U$.

Mỗi $s \in [0,1]$ được gọi là một hằng logic. Với $s_1, s_2 \in [0,1]$, ta xáx định các liên kết logic v, $\wedge, \rightarrow, \neg, \approx$ trên K như sau:
$s_1 v s_2 = \max\{s_1, s_2\}$, $s_1 \wedge s_2 = \min\{s_1, s_2\}$, $s_1 \rightarrow s_2 = \max\{1 - s_1, s_2\}$, $\neg s_1 = 1 - s_1$.

Nói rằng trên U xác định một logic đa trị nếu: Với mỗi phần tử $A_i \in U$, $1 \leq i \leq n$ có một tập hữu hạn B_i được gọi là miền đánh giá của A_i, thoả mãn các điều kiện sau:

1. $B_i \subseteq [0,1]$,
2. nếu $s \in B_i$ thì $1 - s \in B_i$,
3. $1 \in B_i$.

Ký hiệu $B = B_1 \times B_2 \times ... \times B_n$ và $K = B_1 \cup B_2 \cup ... \cup B_n$. Một ánh xạ $x : U \rightarrow K$ sao cho $x(A_i) \in B_i$ với $1 \leq i \leq n$ được gọi là một đánh giá trên U. Nếu $x(A_i) = x_i$, $1 \leq i \leq n$, khi đó ta ký hiệu x bởi $(x_1, x_2, ..., x_n) \in B$.

Các phần tử trong U được gọi là các biến logic hay là các biến sơ cấp. Các hằng logic trong [0,1] và các biến logic trong U được gọi là các công thức.

Giả sử g, h là các công thức, khi đó ta có thể xây dựng các công thức mới bằng cách sử dụng các liên kết logic v, $\wedge, \rightarrow, \neg, \approx$ và các công thức đã cho. Điều đó có nghĩa rằng $(g \wedge h)$. $gvh)$, $\neg(g).(g \rightarrow h)$ là các công thức mới. Ký hiệu tập tất cả các công thức được xây dựng từ U và các liên kết logic đã nêu ở trên bởi $\mathbf{F}$. Mỗi $f \in \mathbf{F}$ được gọi là một phụ thuộc Bool đa trị.

Cho $f \in \mathbf{F}$ và $x = (x_1, x_2, ..., x_n) \in B$. Khi đó ta gọi $f(x)$ là các giá trị chân lý của f đối với đánh giá x và được xác định như sau: Nếu f là một hằng $c \in [0,1]$ thì $f(x) = c$ với mọi $x \in B$. Nếu f là một biến A_i thì $f(x) = x_i$. Khi f được tạo nên từ các công thức g, h và các liên kết logic v, $\wedge, \rightarrow, \neg$ thì ta xác định $f(x)$ như sau

Nếu $f = (g \wedge h)$ thì $f(x) = g(x) \wedge h(x)$
Nếu $f = (gvh)$ thì $f(x) = g(x)^v h(x)$
Nếu $f = (g \rightarrow h)$ thì $f(x) = g(x) \rightarrow h(x)$
Nếu $f = (\neg g)$ thì $f(x) = g(x) \neg (g(x))$

Để ngắn gọn, thay cho các công thức $(g \wedge h)$, (gvh), $(g \rightarrow h)(g \neg h)$ ta viết $g \wedge h$, gvh, $g \rightarrow h$, $\neg g$ một cách tương ứng.

Với $m \in [0,1]$ $f \in \mathbf{F}$, $\sum \subseteq \mathbf{F}$, ta đặt $T^m f = \{x \in B | f(x) \geq m\}$ và $T^m \sum = \{x \in B | \ \forall f \in \sum,\ f(x) \geq m\}$. Nhận thấy $T^m \sum = \cap \{T^m f | \ f \in \sum\}$.

Hệ quả 2.1. *Cho $m, n \in [0,1]$, $m \leq n$. Giả sử có $\sum$, $\Gamma \subseteq \mathbf{F}$ và $\sum \subseteq \Gamma$ khi đó $T^m \Gamma \subseteq T^m \sum$. Nói cách khác toán tử $T^m \sum$ tỷ lệ nghịch với m và $\sum$.*

Chứng minh. Trước hết ta thấy rằng với mọi $m \in [0,1]$ và với $\sum$, $\Gamma \subseteq \mathbf{F}$, $\sum \subseteq \Gamma$ thì có

$$T^m\Gamma \subseteq T^m \sum. \tag{1}$$

Mặt khác với mọi m, $n \in [0,1]$, $m \leq n$ và $\Gamma \subseteq \mathbf{F}$ khi đó

$$T^n\Gamma \subseteq T^m\Gamma. \tag{2}$$

Từ (1) và (2) ta có điều phải chứng minh.

Định nghiã 2.1. Giả sử f, $g \in \mathbf{F}$, $m \in [0,1]$. Nói rằng g là m-suy dẫn f hoặc f là m-suy dẫn được từ g, được ký hiệu là $g|\underset{m}{---}f$, nếu với bất kỳ $x \in B$, thỏa mãn hòamãn $g(x) \geq m$ thì ta cũng có $f(x) \geq m$. Hai công thức f và g được gọi là m-tương đương nếu $f|\underset{m}{===}g$ và $g|\underset{m}{---}f$.

Với $m \in [0,1]$, $\sum \subseteq \mathbf{F}$, $f \in \mathbf{F}$, nói rằng $\sum$ là m-suy dẫn f hoặc f là m- dẫn được từ tập $\sum$, ký hiệu là $\sum|\underset{m}{---}f$, nếu $T^m \sum \subseteq T^m f$. Giả sử $\sum$, $\Gamma \subseteq \mathbf{F}$. Tập Γ được gọi là m-suy dẫn được từ tập $\sum$, ký hiệu là $\sum|\underset{m}{---}\Gamma$, nếu $\sum|\underset{m}{---}f$ cho mỗi $f \in \Gamma$. nói rằng, $\sum$ và Γ là tương đương, ký hiệu là $\sum|\underset{m}{---}\Gamma$, nếu $\sum|\underset{m}{---}\Gamma$ và $\Gamma|\underset{m}{---}\sum$.

Với các tập d_i, $1 \leq i \leq n$ ta xét các ánh xạ : $d_i \times d_i \to [0,1]$ thỏa mãn các điều kiện sau

1. $(\forall a \in d_i)(\alpha(a,a) = 1)$,
2. $(\forall a,\ b \in d_i)(\alpha(a,b) = \alpha(b,a))$,
3. $(\forall s \in B_i, \exists a, b \in d_i)(\alpha(a,b) = s)$.

Ví dụ 2.1. Giả sử $B_i = \{0,1\}$. d_i là một tập hợp mà trên đó có quan hệ so sánh bằng. Khi đó với $a, b \in d_i$ ta xác định $\alpha_i(a,b) = 1$ nếu $a = b$, $\alpha_i(a,b) = 0$ nếu $a \neq b$.

Ví dụ 2.2. Giả sử $B_i = \{0,1\}$, d_i là tập các từ trên một bảng chữ khác trống. Ta định nghiã: $\forall a, b \in d_i$, $\alpha_i(a,b) = 1$ nếu a, b có cùng độ dài $\alpha_i(a,b) = 0$ trong trường hợp ngược lại.

Ví dụ 2.3. Giả sử $B_i = \{0, 0.5, 1\}$, d_i là tập các số nguyên không âm với a, b ta xác định:

$\alpha_i(a,b) = 1$ nếu a, b có cùng tính nguyên tố và $|a - b| \leq 5000$,
$\alpha_i(a,b) = 0.5$ nếu a, b có cùng tính nguyên tố và $|a - b| > 5000$,
$\alpha_i(a,b) = 0$ trong các trường hợp còn lại.

Dễ thấy α_i trong các ví dụ trên thỏa mãn các yêu cầu cuả định nghiã.

Định nghiã 2.2. Một công thức $f \in \mathbf{F}$ được gọi là dương nếu $f(e) = 1$ với $e = (1,1,...,1) \in B$. Ký hiệu $\mathbf{F}p$ là tập tất cả các công thức dương trên U. Mỗi $f \in \mathbf{F}p$ được gọi là một phụ thuộc cuả Bool dương đa trị và được viết tắt PTBDĐT.

Giả sử có R là một mối quan hệ khác trống trên U và $u,.v \in R$ khi đó đánh giá $\{\alpha_1(u.A_1, v.A_1),\ (\alpha_2 9u.A_2, v.A_2), ..., \alpha_n(u.A_n, v.A_n)\}$ được ký hiệu bởi $\alpha(u,v)$. Đặt $T_R = \{\alpha(u,v) \mid u,v \in R\}$. Khi $R = \emptyset$ thì ta đặt $T_R = \emptyset$. Giả sử $m \in [0,1]$, R là một quan hệ trên U và $f \in \mathbf{F}p$. Nói rằng R là m-thoả lập PTBDĐT$\sum$, ký hiệu là $R^m(\sum)$ nếu với mọi $f \in \sum$ có $R^m(f)$. Điều đó tương đương với $T_R \subseteq T^m \sum$.

Cho $m \in [0,1]$, $\sum \subseteq \mathbf{F}p$ và $f \in \mathbf{F}p$. Khi đó ta ký hiệu $\sum \underset{m}{|{-}{-}{-}} f$ có nghiã với bất kỳ quan hệ R trên U nếu $R^m(\sum)$ thì có $R^m(f)$. $\sum \underset{m}{|{-}{-}{-}}_2 f$, có nghiã là với quan hệ R trên U và R chỉ có hai bộ, nếu $R^m(\sum)$ thì ta cũng có $R^m(f)$.

Hệ quả 2.2. Cho R là một quan hệ trên U và $m,n \in [0,1]$, $m \leq n$. Giả sử có $\sum$, $\Gamma \subseteq T^m_\Gamma(1)$. Từ (1) và hệ quả 2.1 ta nhận được $T_R \subseteq T^m \sum$, tức là $R^m(\sum)$.

Hệ quả 2.3. Với mọi $m \in [0,1]$, $\sum$, $\Gamma \subseteq \mathbf{F}$ và $\sum \subseteq \Gamma$. Khi đó nếu $\sum \underset{m}{|{=}{=}{=}} f$ thì $\Gamma \underset{m}{|{=}{=}{=}} f$.

Chứng minh. Giả sử R là quan hệ bất kỳ thoả mãn $R^m(\Gamma)(1)$. Khi đó $T_R \subseteq T^m_\Gamma(2)$. Do $\sum \subseteq \Gamma$ nên $T^m_\Gamma \subseteq T^m_\Gamma(3)$. Từ (20 và (3) suy ra $T_R \subseteq T^m \sum$ tức là $R^m(\sum)$. Theo giả thiết ta có $R^m(f)$.

Bổ đề 2.1. *Giả sử* $m \in [0,1]$, $\sum \subseteq \mathbf{F}p$, $f \in \mathbf{F}p$, *khi đó*

a. $\sum \underset{m}{|{-}{-}{-}} \Gamma$ *khi và chỉ khi* $T^m \sum \subseteq T^m_\Gamma$;

b. $\sum \underset{m}{|{-}{-}{-}} f$ *khi và chỉ khi mối quan hệ* R *thỏa* $T_R \subseteq T^m \sum$ *thì có* $T_R \subseteq T^m_f$.

Chứng minh. Nhận được từ chính định nghiã cuả các kiểu suy diễn.

3. Định lý tương đương

Định lý tương đương đã được chứng minh cho các lớp phụ thuộc Boole dương được xác định trên các ánh xạ bằng trong [2]. Đối với các phụ thuộc Boole dương tổng quát định lý tương đương cũng được đề cập đến trong [5]. Mục này sẽ phát biểu và chứng minh định lý tương đương trong dạng tổng quát hơn cho lớp các phụ thuộc Boole dương đa trị.

Định lý 3.1. (Định lý tương đương) *Giả sử* $m \in [0,1]$, $\sum \subseteq \mathbf{F}p$, $f \in \mathbf{F}p$, *khi đó các điểm sau là tương đương*

1. $\sum \underset{m}{|{-}{-}{-}} f$

2. $\sum \underset{m}{|{=}{=}{=}} f$

3. $\sum \underset{m}{|{=}{=}{=}} f$

Chứng minh. $1 \Rightarrow 2$. Giả sử rằng ta có $\sum \underset{m}{|{-}{-}{-}} f$. Từ giả thiết $\sum \underset{m}{|{-}{-}{-}} f$ và bổ đề 2.1 ta có $T^m \sum \subseteq T^m_f(1)$. Giả sử rằng R là một mối quan hệ trên U và $R^m(\sum)$. khi đó ta có $T_R \subseteq T^m \sum(2)$. Từ (1) và (2) suy ra $T_R \subseteq T^m_f$, tức là $R^m(f)$. Nói cách khác $\sum \underset{m}{|{=}{=}{=}} f$.

$2 \Rightarrow 3.$ là hiển nhiên.

$3 \Rightarrow 1.$ Theo giả thiết ta có $\sum \underset{m}{|===}_2 f$. Tức là với mọi quan hệ R gồm hai bộ mà $R^m(\sum)$ thì cũng có $R^m(f)$. Giả sử x là một phần tử bất kỳ và $x \in T^m \sum$. ta cần chỉ ra rằng $x \in T_f^m$. Thật vậy, vì $x \in B = B_1 \times B_2 \times ... \times B_n$, do đó ta có thể biểu diễn x trong dạng $x = (x_1, x_2, ..., x_n)$, ở đây $x_i \in B_i$, $1 \leq i \leq n$. Theo tính chất của ánh xạ α_i, khi đó sẽ tồn tại hai phần tử $a_i, b_i \in d_i$ sao cho $\alpha_i(a_i, b_i) = x_i$, với $1 \leq i \leq n$. Đặt $u = (a_1, a_2, ..., a_n)$ và $v = (b_1, b_2, ..., b_n)$. Gọi R là quan hệ chỉ gồm hai bộ u và v. Dễ dàng thấy rằng $T_R = \{e, x\} \subseteq T^m \sum$. Nói cách khác $R^m(\sum)$. Do phần 3, ta suy ra $R^m(f)$. Điều đó có nghĩa là $T_R \subseteq T_f^m$, suy ra $x \in T_f^m$. Vậy ta đã chứng minh được rằng với bất kỳ $x \in T^m \sum$ thì ta có $x \in T_f^m$, do đó ta có $\sum \underset{m}{|---} f$. Định lý đã được chứng minh.

4. Hướng nghiên cứu.

Trên cơ sở của định lý tương đương chúng tôi sẽ đề cập đến một số vấn đề liên quan đến quan hệ Armstrong, một số dạng suy diễn có dạng đặc biệt cho lớp các phụ thuộc Boole dương đa trị.

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. Kowalski R. A., *Logic for Data Description,* Logic and Data Bases, Gllaire H., Minker J.Eds, Plenum Press, New York, 1978.
4. Maier D., *The Theory of Relational Databeses,* Computer Science Press, Rockville, Md. 1983.
5. Nguyen Xuan Huy and Le Thi Thanh, *Generalized Positive Boolean Dependencies,* J. Inform. Process. Cybern. EIK 28 (1992), 363-370.
6. Sagiv Y., Delobel C, Parker D. S., and Fagin R., *An Equivalennce Between Relational Database Dependencies and a Fragment of Propositional Logic,* J. ACM, 28 (1981). 435-453. Also a correction to this paper in J. ACM, 34 (1978), 1016-1028.

Abstract

On the equivalent theorem in a class of multivalued positive Bollean dependencies

The main results of the paper is showing an equivalent theorem of consequences in the world of all relations. the world of 2-tuple relations and propositional logic for the class of multi-valued positive Boolean dependencies.