

# PHỤ THUỘC HÀM ĐỐI TƯỢNG MỜ TRONG CƠ SỞ DỮ LIỆU HƯỚNG ĐỐI TƯỢNG MỜ\*

ĐOÀN VĂN BAN<sup>1</sup>, HỒ CẨM HÀ<sup>2</sup>, VŨ ĐỨC QUẢNG<sup>3</sup>

<sup>1</sup>*Viện Công nghệ Thông tin, Viện Khoa học và Công nghệ Việt Nam*

<sup>2</sup>*Trường Đại học Sư phạm Hà nội*

<sup>3</sup>*Trường Đại học Quảng Nam*

**Tóm tắt.** Bài báo này giới thiệu các phụ thuộc hàm đối tượng mờ cho phép biểu diễn các ràng buộc trên các thuộc tính của các kiểu đối tượng mờ có thể bao gồm cả kiểu của chính nó trong lược đồ CSDL mờ. Các ngữ nghĩa khác nhau của FOFD liên quan đến tính hợp lệ của trạng thái lược đồ CSDL hướng đối tượng mờ đối với phụ thuộc hàm đối tượng mờ cũng được đưa ra. Chúng được xem như một trong các công cụ nhận biết đối tượng trong CSDL hướng đối tượng mờ.

**Abstract.** In this article, we propose fuzzy object functional dependences which allow to express constraints on attributes of arbitrary object types including the types themselves. Different semantics of fuzzy object functional dependency (FOFD) relating to the validity of a fuzzy object-oriented database schema state for a FOFD are introduced and discussed. FOFDs are considered as one of the tools to identify fuzzy objects in fuzzy object-oriented databases.

## 1. GIỚI THIỆU

Trong những năm gần đây, việc nghiên cứu, ứng dụng cơ sở dữ liệu hướng đối tượng (CSDL HDT) mờ để đặt tả các đối tượng phức tạp khắc phục những hạn chế của cơ sở dữ liệu quan hệ/hướng đối tượng truyền thống trong việc biểu diễn và xử lý các thông tin không chắc chắn, không đầy đủ được nhiều người tập trung nghiên cứu và triển khai. Thông thường, các phụ thuộc dữ liệu là thông tin ngữ nghĩa về thế giới thực và được xem như các ràng buộc toàn vẹn đối với việc thiết kế CSDL và truy xuất dữ liệu.

Một trong số các khái niệm cơ bản của mô hình dữ liệu hướng đối tượng rõ hay mờ là định danh đối tượng (Object Identifier). Nó cho phép phân biệt các đối tượng, ngay cả khi chúng có cùng các giá trị thuộc tính hay chúng có giá trị được xem là giống nhau với một mức độ  $\alpha$  ( $\alpha \in [0, 1]$ ) nào đó. Tuy nhiên, các định danh đối tượng được xác lập bởi hệ thống hướng đối tượng, và ẩn đối với người sử dụng, tức là các ngôn ngữ truy vấn dữ liệu không được phép truy xuất trực tiếp đến các định danh này. Vấn đề đặt ra là làm thế nào để xác định được các đối tượng của một lớp trong mỗi trạng thái CSDL HDT mờ.

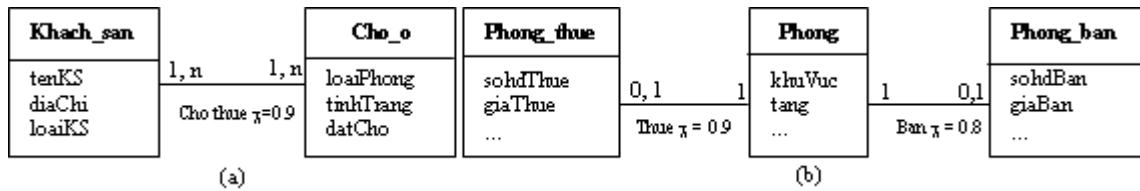
Trong thiết kế CSDL HDT mờ và việc thao tác trên tập các đối tượng mờ, thông tin về

---

\*Nghiên cứu này được hoàn thành dưới sự hỗ trợ từ Quỹ phát triển khoa học và Công nghệ quốc gia (NAFOSTED) mã số 102.01-2011.06

khả năng truy xuất các đối tượng được quan tâm đặc biệt. Kỹ thuật xác định đối tượng trong CSDL HDT đã được trình bày bởi một số tác giả [1, 8, 9]. Tuy nhiên, việc mô tả đặc trưng và nghiên cứu cách thức xác định đối tượng dựa vào giá trị của các thuộc tính cùng với các mối quan hệ của chúng chưa được đề cập đầy đủ. Trong [2], chúng tôi giới thiệu một dạng phụ thuộc hàm mờ cho CSDL HDT mờ làm cơ sở của việc chuẩn hóa lớp đối tượng mờ. Việc chuẩn hóa đối tượng rất quan trọng trong giai đoạn thiết kế CSDL HDT bởi vì nó hạn chế dư thừa dữ liệu khi thêm đối tượng vào trong lớp đối tượng hay hạn chế sự mất mát thông tin khi xóa đối tượng ra khỏi lớp. Trong bài báo này, dựa trên sự tương đương ngữ nghĩa giữa hai giá trị mờ, chúng tôi mở rộng khái niệm phụ thuộc hàm đối tượng của H.J. Klein và các cộng sự [5] cho CSDL HDT mờ, từ đó đề xuất kỹ thuật nhận biết các đối tượng trong mỗi trạng thái CSDL.

Trước tiên, chúng ta xét lược đồ CSDL HDT mờ cho trước như trong Hình 1. Hình 1(a) mô tả mối quan hệ giữa *Khach\_san* (khách sạn) cho thuê các phòng - Chỗ ở theo các loại khác nhau, chẳng hạn phòng đơn hay phòng đôi. Giá trị các thuộc tính của *Cho\_o* (Chỗ ở) không thể được sử dụng như các giá trị đầu vào để truy xuất một đối tượng *Cho\_o*, bởi vì các khách sạn khác nhau có thể cho thuê các phòng thuộc cùng một loại. Vì vậy, một đối tượng *Khach\_san* kết hợp với giá trị của thuộc tính *loaiPhong* giúp ta xác định được duy nhất một đối tượng *Cho\_o*. Trong Hình 1.b, chỉ ra một dạng nhận biết đối tượng khác: một tòa nhà gồm có hai loại phòng: phòng cho thuê của lớp *Phong\_thue* và phòng để bán của lớp *Phong\_ban*. Nếu thuộc tính *sohdThue* (số hợp đồng thuê) và *sohdBan* (số hợp đồng bán) lần lượt là khóa của lớp *Phong\_thue* và lớp *Phong\_ban*, khi đó, tổ hợp bộ phận giá trị của thuộc tính *sohdThue* và *sohdBan* có thể được dùng để xác định một đối tượng Phong – đây là một cách xác định đối tượng dựa vào giá trị của các thuộc tính, mối quan hệ ràng buộc giữa thuê và bán trong trường hợp này được xác định bởi phép toán *xor*.



Hình 1.1. Lược đồ CSDL hướng đối tượng mờ

Bài báo được trình bày như sau: mục 2 trình bày một số khái niệm cơ sở liên quan đến độ đo ngữ nghĩa của hai dữ liệu mờ theo phân bố khả năng và mô hình dữ liệu hướng đối tượng mờ; mục 3 đưa ra định nghĩa về phụ thuộc hàm đối tượng mờ; mục 4 giới thiệu thuật toán xây dựng một quan hệ mờ biểu diễn một phần trạng thái lược đồ CSDL được tham chiếu bởi phụ thuộc hàm đối tượng mờ, mục 5 trình bày các dạng phụ thuộc hàm đối tượng mờ và cuối cùng là phần kết luận.

## 2. CÁC KHÁI NIỆM LIÊN QUAN

### 2.1. Độ đo ngữ nghĩa giữa hai dữ liệu mờ

Theo cách tiếp cận của Zongmin Ma [10], giá trị mờ của thuộc tính  $X$  của lớp trong mô hình CSDL HDT được biểu diễn bởi phân bố khả năng như sau:

$$\pi_X = \{\pi_X(u_1)/u_1, \pi_X(u_2)/u_2, \pi_X(u_3)/u_3, \dots, \pi_X(u_n)/u_n\}$$

trong đó,  $U = \{u_1, u_2, u_3, \dots, u_n\}$  là một vũ trụ,  $\pi_X(u_i)$ ,  $u_i \in U$ , biểu thị khả năng  $X$  nhận giá trị  $u_i$ .

Với hai dữ liệu mờ  $\pi_A$  và  $\pi_B$  được định nghĩa trên miền  $U$  theo phân bố khả năng. Mức độ mà  $\pi_A$  bao hàm ngữ nghĩa  $\pi_B$ , ký hiệu  $SID(\pi_A, \pi_B)$  [11] được xác định như sau:

$$SID(\pi_A, \pi_B) = \sum_{i=1}^n \min(\pi_B(u_i), \pi_A(u_i)) / \sum_{i=1}^n \pi_B(u_i).$$

Mức độ tương đương ngữ nghĩa giữa hai dữ liệu mờ  $\pi_A$  và  $\pi_B$ , ký hiệu là  $SE(\pi_A, \pi_B)$ , được xác định như sau:

$$SE(\pi_A, \pi_B) = \min(SID(\pi_A, \pi_B), SID(\pi_B, \pi_A))$$

Với  $\beta \in [0, 1]$  là một ngưỡng cho trước, hai dữ liệu mờ  $\pi_A$  và  $\pi_B$  tương đương mức  $\beta$  nếu  $SE(\pi_A, \pi_B) \geq \beta$ .

## 2.2. Lược đồ đối tượng mờ

Một lược đồ CSDL HDT mờ  $S$  bao gồm các kiểu (lớp) đối tượng mờ và các mối quan hệ mờ nhị nguyên giữa các kiểu đối tượng mờ với mức độ kết hợp  $\chi \in [0, 1]$ , bao gồm cả quan hệ thừa kế [2, 3]. Ở đây, chúng ta xem một phân cấp theo thừa kế như là một tập các mối quan hệ mờ nhị nguyên với các ràng buộc về bội số được thêm vào ở đầu mỗi mối quan hệ mờ. Một kiểu đối tượng mờ  $O$  có một tập các thuộc tính  $attr(O)$  và thuộc tính  $\mu_O$  biểu thị độ thuộc thành viên của đối tượng thuộc về kiểu đối tượng  $O$ , mỗi thuộc tính  $A \in attr(O)$  có một miền giá trị  $dom(A)$  mờ (rõ). Các mối quan hệ mờ có thể có các bội số, được xem như các ràng buộc. Trong phạm vi bài báo này, chúng ta giả sử tên của các mối quan hệ mờ, tên các thuộc tính của các kiểu đối tượng mờ là duy nhất trong toàn bộ lược đồ CSDL HDT mờ.

Đặt  $I$  là tập hữu hạn các định danh đối tượng, mỗi đối tượng mờ  $o$  của kiểu đối tượng mờ  $O$  được biểu diễn bởi bộ ba  $(id, v, \mu_o)$ , trong đó  $id \in I$  và  $v$  là một bộ  $(a_1, a_2, \dots)$  với  $a_i \in dom(A_i)$ , được gọi là giá trị của đối tượng mờ  $o$ ,  $\mu_o$  là độ thuộc thành viên của đối tượng  $o$  thuộc vào kiểu đối tượng mờ  $O$ . Một thể hiện của kiểu đối tượng mờ  $O$ , ký hiệu  $ext(O)$ , là một tập các đối tượng mờ của kiểu đối tượng mờ  $O$ , ví dụ, với lược đồ CSDL HDT mờ trong Hình 1.a,  $ext(KhachSan) = \{(1, [Victoria, Cửa Đại, \{0.9/4\_sao, 0.5/5\_sao\}], \mu_1), (2, [Hội An, Trần Hưng Đạo, \{0.7/3\_sao, 0.5/4\_sao, 0.3/5\_sao\}], \mu_2)\}$ . Ở đây giá trị của thuộc tính *loaiKS* được biểu diễn bởi phân bố khả năng trên miền  $\{1\_sao, 2\_sao, 3\_sao, 4\_sao, 5\_sao\}$ .

Với mỗi mối quan hệ mờ  $r$  giữa hai kiểu đối tượng mờ  $O_1, O_2$ , một thể hiện của mối quan hệ mờ  $r$  giữa hai kiểu đối tượng mờ  $O_1, O_2$ , ký hiệu  $ext(r)$ , là tập các liên kết  $(id_1, id_2) \in I \times I$  với mức độ kết hợp  $\chi \in [0, 1]$ , trong đó,  $id_1 \in I(ext(O_1))$  và  $id_2 \in I(ext(O_2))$ . Nếu có một ràng buộc bội số được xác định trên mối quan hệ mờ  $r$  thì  $ext(r)$  cũng phải tuân theo ràng buộc đó.

Trạng thái  $s(S)$  của lược đồ CSDL HDT mờ  $S$  bao gồm tất cả các  $ext(O)$  và  $ext(r)$  của  $S$  sao cho  $I(ext(O_1)) \cap I(ext(O_2)) = \emptyset, \forall O_1, O_2 \in S, I(ext(O_1))$  và  $I(ext(O_2))$  lần lượt là tập

các định danh của các đối tượng mờ thuộc  $O_1$  và  $O_2$ . Để biểu diễn một CSDL HDT mờ, ta sử dụng đồ thị lược đồ mờ tương tự như [3]. Một đồ thị lược đồ mờ  $G_s = (V, E, l)$  của lược đồ CSDL HDT mờ  $S$  là đồ thị với cạnh được gán nhãn, trong đó, tập đỉnh  $V$  tương ứng với tập tất cả các kiểu đối tượng mờ; tập cạnh  $E$  tương ứng với tập các mối quan hệ mờ của lược đồ CSDL HDT mờ  $S$ ;  $l$  là hàm gán nhãn cạnh, được xác định như sau:  $\forall e \in E, l(e) = (r_n, \chi)$ , ở đây,  $r_n$  là tên của một mối quan hệ mờ  $r$  được biểu diễn bởi cạnh  $e$  và  $\chi$  là mức độ kết hợp giữa hai kiểu đối tượng mờ trong mối quan hệ mờ  $r$  của  $S$ . Một đường dẫn từ kiểu đối tượng mờ  $O_1$  đến kiểu đối tượng mờ  $O_n$  trong  $G_s$  là một chuỗi  $\Pi = O_1 e_1 O_2 e_2 O_3 \dots e_{n-1} O_n$ , trong đó,  $O_i \in V, e_j \in E, j \in \{1, 2, \dots, n-1\}$ ,  $e_j$  biểu diễn mối quan hệ giữa hai kiểu đối tượng mờ  $O_j$  và  $O_{j+1}$  với mức độ kết hợp  $\chi$ . Với  $s(S)$ , một chuỗi liên kết giữa đối tượng mờ  $o_1 \in ext(O_1)$  và đối tượng mờ  $o_n \in ext(O_n)$  là  $\pi = o_1 l_1 o_2 \dots l_{n-1} o_n$  sao cho  $o_i \in ext(O_i), l_j \in ext(r_j)$  với  $l_j$  là một cặp các định danh của  $(o_j, o_{j+1})$  và  $r_j$  là nhãn của cạnh  $e_j$ .

Đặt  $Sets(O) = 2^{attr(O)} \cup \{\{O\}\}$ ,  $O$  được viết ngắn gọn thay cho  $\{O\}$ , và đặt  $OT_s$  là tập tất cả các kiểu đối tượng mờ của  $S$ ,  $D_s = \bigcup_{O \in OT_s} Sets(O)$ . Một lược đồ CSDL HDT có nhất

một kiểu đối tượng.

### 2.3. Quan hệ mờ

Một  $ext(O)$  có thể được biểu diễn bởi một quan hệ mờ  $R_{ext(O)}$  trên tập thuộc tính  $\Omega_R = attr(O) \cup \{id_O\} \cup \mu_O$ , trong đó  $id_O$  được gọi là thuộc tính định danh có miền giá trị I,  $\mu_O$  là thuộc tính thành viên, và  $R_{ext(O)} = \{t \mid t \text{ là bộ giá trị xác định trên } \Omega_R \wedge (\exists(i, v, \mu_o) \in ext(O))(t[id_O] = i \wedge t[attr(O)] = v \wedge t[\mu_O] = \mu_o)\}$ .

Tương tự,  $ext(r)$  giữa hai đối tượng  $O_1, O_2$  có thể được biểu diễn bởi một quan hệ mờ  $R_{ext(r)}$  trên tập thuộc tính  $\Omega_R = \{id_{O_1}, id_{O_2}, \mu_O\}$ , và  $R_{ext(r)} = \{t = (id_1, id_2, \mu_o) \mid t \text{ là bộ giá trị xác định trên } \Omega_R \wedge (\exists(id_1, v, \mu_{o_1}) \in ext(O_1), \exists(id_2, w, \mu_{o_2}) \in ext(O_2))(t[id_{O_1}] = id_1 \wedge t[id_{O_2}] = id_2 \wedge (\mu_o = \min(\mu_{o_1} \cdot \chi, \mu_{o_2} \cdot \chi)), ((id_1, v, \mu_{o_1}) \text{ và } (id_2, w, \mu_{o_2}) \text{ có quan hệ với nhau với mức độ } \chi))\}$ . Miền giá trị của các thuộc tính của quan hệ mờ là miền giá trị của các thuộc tính của kiểu đối tượng mờ tương ứng. Ví dụ, một quan hệ mờ biểu diễn các đối tượng với các thuộc tính  $\{id_O, A, B, C, \mu_O\}$  (thuộc tính  $A$  có miền giá trị mờ) như trong Bảng 1 dưới đây.

Cho  $R_{fo}$  là một quan hệ mờ trên tập thuộc tính  $\Omega_R$  và tập thuộc tính  $X \subseteq \Omega_R$ . Một bộ  $t \in R_{fo}$  là đầy đủ trên  $X$  nếu  $t[C] \neq \perp'$  với mọi  $C \in X$ .  $NF(R_{fo}, X) = \{t \mid t \in R_{fo} \text{ và } t \text{ đầy đủ trên } X\}$ , được gọi là quan hệ mờ lọc null trên  $X$  (Null Filter - NF).  $WNF(R_{fo}, X) = \{t \mid t \in R_{fo} \text{ và } \forall t, \exists C \in X, t[C] \neq \perp'\}$ , được gọi là quan hệ mờ lọc null yếu trên  $X$  (Weak Null Filter - WNF).

Bảng 1. Một quan hệ mờ

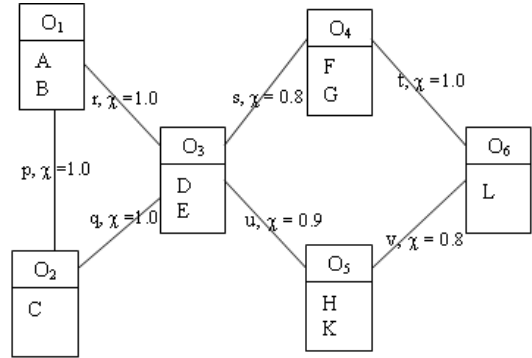
$id_O$	$A$	$B$	$C$	$\mu_O$
01	{0.9/1, 1.0/2, 0.9/3}	2	3	$\mu_1$
02	{0.8/1, 1.0/2, 0.8/3}	2	' $\perp'$	$\mu_2$
03	{0.8/1, 0.8/2, 1.0/3}	' $\perp'$	3	$\mu_3$
04	' $\perp'$	2	' $\perp'$	$\mu_4$

Một bộ  $t \in R_{fo}$  là không xác định trên  $X$  nếu  $t[C] = \perp'$  với mọi  $C \in X$ . Ở đây, chúng ta sử dụng ký hiệu ' $\perp'$ ' để chỉ những liên kết thiếu của một đối tượng hoặc để chỉ giá trị là null trong trường hợp không có giá trị,  $R_{fo}[X]$  là phép chiếu của quan hệ mờ  $R_{fo}$  lên tập thuộc tính  $X$ ,  $t[C]$  biểu thị giá trị của bộ  $t$  trên tập thuộc tính  $C$ .  $t_1, t_2$  là hai bộ của  $R_{fo}$ ,  $\alpha \in [0, 1]$  là một ngưỡng tương đương cho trước,  $t_1$  phủ  $t_2$  trên  $X$  nếu  $(\forall C \in X)(SE(t_1[C], t_2[C]) \geq \alpha \text{ hay } t_2[C] = \perp')$ . Ví dụ, với một quan hệ mờ trong Bảng 1 và  $\alpha = 0.8$ , ta có  $SE(t_1[A], t_2[A]) = \min(SID(t_1[A], t_2[A]), SID(t_2[A], t_1[A]))$  với  $SID(t_1[A], t_2[A]) = (0.8 + 1.0 + 0.8)/2.6 = 1$ ,  $SID(t_2[A], t_1[A]) = (0.8 + 1.0 + 0.8)/2.8 = 0.928$ , suy ra  $SE(t_1[A], t_2[A]) = \min(1, 0.928) = 0.928 > \alpha$ , và  $SE(t_1[B], t_2[B]) = 1 > \alpha$  và  $t_2[C] = \perp'$ . Vậy  $t_1$  phủ  $t_2$  trên  $\{A, B, C\}$ .

### 3. PHỤ THUỘC HÀM ĐỐI TƯỢNG MỜ

Mở rộng của khái niệm phụ thuộc hàm trong cơ sở dữ liệu quan hệ cho các lược đồ CSDL HDT mờ với các ràng buộc ở cả mức đối tượng và mức thuộc tính, tức là, về trái và về phải của một phụ thuộc hàm mờ  $f$  không chỉ chứa các thuộc tính của kiểu đối tượng mờ  $O$  mà còn chứa kiểu đối tượng của chính nó. Hướng tiếp cận này cho ta phụ thuộc hàm đối tượng mờ có dạng  $f : \Delta \xrightarrow{f} \Gamma$  với  $\Delta, \Gamma \in D_s$ , ở đây bất kỳ kiểu đối tượng của  $S$  có thể tham gia vào  $\Delta, \Gamma$ . Dạng phụ thuộc hàm mờ này có thể biểu diễn các phụ thuộc hàm mờ ở mức lược đồ, giống như các ràng buộc giữa các quan hệ trong lược đồ CSDL quan hệ. Các kiểu đối tượng trong lược đồ xác định một phụ thuộc hàm khi giữa chúng có một đường dẫn.

Sự nhập nhằng của một phụ thuộc hàm đối tượng mờ như đã đề cập ở trên có thể xuất hiện khi giữa hai kiểu đối tượng mờ bất kỳ xuất hiện trong  $\Delta$  hay giữa bất kỳ kiểu đối tượng mờ trong  $\Delta$  và kiểu đối tượng mờ trong  $\Gamma$  có thể có nhiều hơn một đường dẫn tồn tại trong  $G_s$ , ví dụ, với lược đồ CSDL HDT mờ trong Hình 3.2: phụ thuộc hàm đối tượng mờ có 2 đường dẫn kết nối  $O_3$  với  $O_6$  và hai đường dẫn (không chu trình) giữa  $O_1$  và  $O_3$ . Thường các đường dẫn khác nhau có ngữ nghĩa tương ứng khác nhau, và người thiết kế CSDL chỉ tập trung vào một trong số đường dẫn khi xác định phụ thuộc hàm mờ. Một điều hiển nhiên là, một phụ thuộc hàm đối tượng mờ có thể thỏa mãn đối với một đường dẫn và không thỏa mãn đối với các đường dẫn khác. Do đó, các đường dẫn cần phải được xác định cùng với các phụ thuộc hàm mờ. Vì vậy, việc biểu diễn phụ thuộc hàm đối tượng mờ cần gắn kết với một đồ thị biểu diễn đường dẫn kết nối giữa các kiểu đối tượng xuất hiện trong phụ thuộc hàm. Trong phạm vi bài báo này, ta chỉ giới hạn nghiên cứu các dạng phụ thuộc hàm mờ phi chu trình.



Hình 3.2. Sự nhập nhằng của một FOFD

thỏa mãn đối với một đường dẫn và không thỏa mãn đối với các đường dẫn khác. Do đó, các đường dẫn cần phải được xác định cùng với các phụ thuộc hàm mờ. Vì vậy, việc biểu diễn phụ thuộc hàm đối tượng mờ cần gắn kết với một đồ thị biểu diễn đường dẫn kết nối giữa các kiểu đối tượng xuất hiện trong phụ thuộc hàm. Trong phạm vi bài báo này, ta chỉ giới hạn nghiên cứu các dạng phụ thuộc hàm mờ phi chu trình.

**Định nghĩa 3.1.** Cho  $S$  là một lược đồ CSDL HDT mờ được biểu diễn bằng đồ thị lược đồ mờ  $G_s = (V, E, l)$ .  $f = (G_f, v_f)$  là một phụ thuộc hàm đối tượng mờ của  $S$ , trong đó:

- (i) Đồ thị phụ thuộc đối tượng mờ:  $G_f = (V_f, E_f, \eta_f)$  là một cây với tập cạnh  $E_f (E_f \subseteq E)$  nối các đỉnh thuộc  $V_f (V_f \subseteq V, V_f \neq \emptyset)$ , và  $\eta_f$  là hàm  $l$  thu hẹp xác định trên  $E_f$ .
- (ii)  $v_f : V_f \rightarrow D_s \times D_s$  là hàm bộ phận gán nhãn cho đỉnh của  $G_f$  sao cho với mỗi  $O \in V_f$ , nếu  $v_f$  xác định thì  $v_f(O) = (\delta, \gamma), \delta, \gamma \in Sets(O)$  và nếu  $O$  là nút lá thì  $\delta \neq \emptyset$  hoặc  $\gamma \neq \emptyset$ .

Với một kiểu đối tượng mờ  $O$ ,  $v_f(O) = (\delta, \gamma)$  và  $\delta \neq \emptyset (\gamma \neq \emptyset)$  được gọi là kiểu đối tượng mờ nguồn (trạm) của  $f$ . ( $\delta$  : các thành phần của một kiểu đối tượng mờ được dùng để xác định các kiểu đối tượng mờ khác,  $\gamma$  : các thành phần của một kiểu đối tượng mờ được xác định bởi các kiểu đối tượng mờ khác)

Các FOFD được biểu diễn bằng các cây nối các đối tượng mờ tương ứng với các đỉnh của đồ thị lược đồ mờ đảm bảo rằng không có sự nhập nhằng liên quan đến các kết nối giữa các kiểu đối tượng mờ nguồn (source) và kiểu đối tượng mờ trạm (sink). Với một kiểu đối tượng mờ có thể vừa là đối tượng mờ nguồn vừa là đối tượng mờ trung gian của một FOFD (chẳng

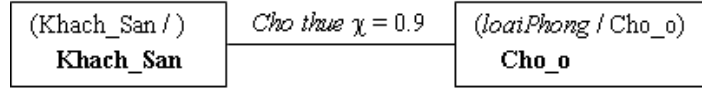
hạn, kiểu đối tượng *Cho\_o* trong Hình 3.3). Mỗi thành phần của nhãn bao gồm một tập các thuộc tính hay kiểu đối tượng của chính nó. Mỗi thành phần của nhãn được gán cho đỉnh trong đồ thị phụ thuộc đối tượng mờ không thể gồm cả các thuộc tính và kiểu đối tượng vì định danh của đối tượng xác định duy nhất một đối tượng nên nó cũng xác định duy nhất các thuộc tính của đối tượng.

Chúng ta sử dụng ký hiệu  $\Delta \xrightarrow{f} \Gamma$  thay cho  $v_f$ , trong đó  $\Delta = \bigcup_{O \in V_f, v_f(O) = (\delta, \gamma) \wedge \delta \neq \emptyset} \{\delta\}$ ,  $\Gamma = \bigcup_{O \in V_f, v_f(O) = (\delta, \gamma) \wedge \gamma \neq \emptyset} \{\gamma\}$ ,  $\Delta$  được gọi là vế phải và  $\Gamma$  được gọi là vế trái của FOFD. Khi đó,

một FOFD  $f = (G_f, v_f)$  có thể được biểu diễn bởi  $f : \Delta \xrightarrow{G_f} \Gamma$  với đồ thị phụ thuộc đối tượng

mờ  $G_f$  và  $\Delta \xrightarrow{f} \Gamma$ . Ví dụ, một FOFD biểu diễn ràng buộc đã được đề cập trong phần giới thiệu (Hình 1.a) và được chỉ ra trong Hình 3.3 với hàm gán nhãn cho đỉnh  $v_f(Khach\_san) = (\delta, \gamma)$ ,  $\delta = Khach\_san$ ,  $\gamma = \emptyset$  và  $v_f(Cho\_o) = (\delta, \gamma)$ ,  $\delta = loaiPhong$ ,  $\gamma = Cho\_o$ . Ta có FOFD:  $\{Khach\_San, loaiPhong\} \xrightarrow{G_f} \{Cho\_o\}$ .

Một kiểu đối tượng mờ  $O$  được tham chiếu bởi FOFD  $f$  (hay được bao hàm trong FOFD  $f$ ) nếu chính  $O$  hay bất kỳ tập con của tập thuộc tính của nó xuất hiện trong vế trái hay vế phải của  $f$ .  $f$  được gọi là phụ thuộc hàm đối tượng mờ chuẩn tắc nếu chỉ có một kiểu đối tượng mờ được bao hàm trong vế phải  $\Gamma$ .  $f$  được gọi là cục bộ (toàn cục) nếu chỉ có một kiểu đối tượng mờ (nhiều hơn một kiểu đối tượng mờ) được bao hàm trong  $f$ .



Hình 3.3. Phụ thuộc hàm đối tượng mờ

#### 4. QUAN HỆ MỜ BIỂU DIỄN MỘT BỘ PHẬN CỦA $s(S)$

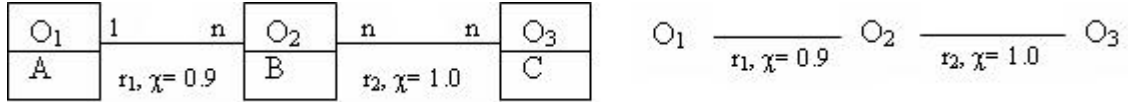
Làm thế nào để kiểm tra một trạng thái  $s(S)$  của lược đồ CSDL HDT mờ có thỏa FOFD  $f : \Delta \xrightarrow{G_f} \Gamma$  hay không. Tương tự như trong CSDL quan hệ, đơn giản chúng ta chỉ cần kiểm tra trên quan hệ mờ  $R_{ext(O)}$ , nếu  $f$  là FOFD cục bộ. Trường hợp,  $f$  là FOFD toàn cục, trước tiên, chúng ta tạo ra một quan hệ mờ bằng cách kết nối các quan hệ mờ tương ứng với các mối quan hệ mờ của lược đồ CSDL HDT mờ và các kiểu đối tượng mờ trong các đường dẫn kết nối đến các kiểu đối tượng mờ trong đồ thị phụ thuộc hàm đối tượng mờ  $G_f$ , sau đó kiểm tra quan hệ mờ thu được có thỏa  $f$  không.

Trước khi đề cập đến thuật toán tạo ra một quan hệ mờ như đã thảo luận ở trên, chúng ta phát triển phép toán liên kết ngoài đầy đủ trong [6] để kết nối các quan hệ mờ. Không mất tính tổng quát, những thuộc tính thành viên trong các quan hệ mờ sẽ không được xem xét khi thực thi phép toán.

**Định nghĩa 4.1.** Cho  $R$  và  $S$  là hai quan hệ mờ với tập các thuộc tính lần lượt là  $\Omega_R, \Omega_S, \Omega_R \cap \Omega_S \neq \emptyset$ . Phép liên kết ngoài mờ đầy đủ với ngưỡng tương đương  $\alpha$ , ( $\alpha \in [0, 1]$ ) của  $R$  và  $S$ , ký hiệu  $R \overset{f}{\bowtie}_{\alpha} S$  được xác định như sau:

$$\begin{aligned}
 R \stackrel{f}{\underset{\alpha}{\bowtie}} S &= \{t \mid t \text{ là bộ xác định trên } \Omega_R \cup \Omega_S \wedge \\
 &((t[C] \neq \perp', \forall C \in (\Omega_R \cap \Omega_S) \wedge (\exists t')(\exists t'')(t' \in R, t'' \in S)(SE(t'[\Omega_R \cap \Omega_S], t''[\Omega_R \cap \Omega_S]) \geq \\
 &\alpha \wedge t[\Omega_R - (\Omega_R \cap \Omega_S)] = t'[\Omega_R - (\Omega_R \cap \Omega_S)] \wedge t[\Omega_S - (\Omega_R \cap \Omega_S)] = t''[\Omega_S - (\Omega_R \cap \Omega_S)]) \wedge \\
 &t[\Omega_R \cap \Omega_S] = t'[\Omega_R \cap \Omega_S]) \\
 &\vee (t[\Omega_R] \in R \wedge t[C] \neq \perp', \forall C \in (\Omega_R \cap \Omega_S) \wedge \neg(\exists t' \in S)(SE(t[\Omega_R \cap \Omega_S], t'[\Omega_R \cap \Omega_S]) \geq \\
 &\alpha) \wedge t[C] = \perp', \forall C \in (\Omega_S - \Omega_R)) \\
 &\vee (t[\Omega_S] \in S \wedge t[C] \neq \perp', \forall C \in (\Omega_R \cap \Omega_S) \wedge \neg(\exists t' \in R)(SE(t[\Omega_R \cap \Omega_S], t'[\Omega_R \cap \Omega_S]) \geq \\
 &\alpha) \wedge t[C] = \perp', \forall C \in (\Omega_S - \Omega_R)) \\
 &\vee (t[\Omega_R] \in R \wedge t[C] = \perp', \forall C \in \Omega_S) \\
 &\vee (t[\Omega_S] \in S \wedge t[C] = \perp', \forall C \in \Omega_R))\}.
 \end{aligned}$$

trong đó, hai bộ  $t$  và  $t'$  được xem như giống nhau trên tập thuộc tính  $X$  nếu mức độ tương đương giữa hai bộ này trên giá trị của tập thuộc tính  $X$  lớn hơn hoặc bằng  $\alpha$ . Ba điều kiện đầu của phép liên kết ngoài mờ đầy đủ có được bằng cách mở rộng ba điều kiện của phép liên kết ngoài đầy đủ, các bộ thu được sau khi thực hiện phép toán này là đầy đủ trên các thuộc tính chung. Hai điều kiện sau được thêm vào cho phép các bộ có giá trị không xác định trên các thuộc tính chung được bổ sung vào quan hệ kết quả. Các điều kiện này là phù hợp với cách tiếp cận của chúng ta vì giá trị trên tập các thuộc tính liên kết là duy nhất.



Hình 4.4. Đồ thị lược đồ CSDL hướng đối tượng mờ

**Ví dụ 4.1:** Xét lược đồ trong Hình 4.1 bao gồm 3 kiểu đối tượng mờ  $O_1, O_2, O_3$  có tập thuộc tính lần lượt là  $A, B, C$  và hai mối quan hệ  $r_1, r_2$ , được biểu diễn như sau:

$$\begin{aligned}
 Ext(O_1) &= \{(1, [\{1.0/a, 0.7/b, 0.6/c\}]), (2, [\{1.0/a, 0.7/b, 0.7/c\}]), (3, [\{0.6/a, 1.0/b, 0.7/c\}])\}; \\
 Ext(O_2) &= \{(4, [\{0.6/a, 1.0/c\}]), (5, [\{0.6/a, 0.7/b, 1.0/c\}]), (6, [\{1.0/a, 0.6/b, 0.6/c\}]), \\
 &(7, [\{0.6/a, 0.6/c, 1.0/d\}]), (8, [\{0.7/a, 1.0/b, 0.6/c\}])\}; \\
 Ext(O_3) &= \{(9, [\{0.6/a, 0.6/b, 1.0/e\}]), (10, [\{0.7/a, 0.7/c, 1.0/f\}]), (11, [\{0.6/a, 0.7/b, 1.0/g\}])\}; \\
 Ext(r_1) &= \{(1, 5), (1, 6), (2, 7), (3, 8)\}; \\
 Ext(r_2) &= \{(6, 9), (7, 10)\}.
 \end{aligned}$$

Dựa trên khái niệm về liên kết bộ phận trong [5], ta có: các đối tượng  $(1, [\{1.0/a, 0.7/b, 0.6/c\}]), (2, [\{1.0/a, 0.7/b, 0.7/c\}]) \in O_1$  có liên kết bộ phận đến các đối tượng của  $O_3$ , tức là các đối tượng với  $id = 1$  và  $id = 2$  này có một số chuỗi liên kết đến các đối tượng của  $O_3$  (không cần phải có liên kết đến tất cả các đối tượng  $O_3$ ); đối tượng  $(3, [\{0.6/a, 1.0/b, 0.7/c\}]) \in O_1$  chỉ có một liên kết thiếu đến các đối tượng của  $O_3$ ; đối tượng  $(4, [0.6/a, 1.0/c]) \in O_2$  không có liên kết nào đến các đối tượng của  $O_3$ . Ví dụ, sử dụng phép toán liên kết ngoài mờ đầy đủ để nối các quan hệ mờ  $R_{ext(O_3)}$  và  $R_{ext(r_2)}$  ta thu được quan hệ mờ  $R$  (như trong Bảng 2).

Nếu quan hệ mờ  $R$  là một quan hệ trung gian cho việc thực hiện các phép toán liên kết ngoài mờ đầy đủ tiếp theo, điều kiện bốn và năm đảm bảo rằng bộ  $t = (\perp', 11, \{0.6/a, 0.7/b, 1.0/g\})$  không bị mất.

Với  $S$  là một lược đồ CSDL HDT mờ, dựa vào phép toán liên kết ngoài mờ đầy đủ được trình bày ở trên, thuật toán xác định quan hệ mờ  $R_{fo}$  biểu diễn một bộ

Bảng 2.  $R = R_{ext(O_3)} \stackrel{f}{\underset{0.9}{\bowtie}} R_{ext(r_2)}$

$id_{o_2}$	$id_{o_3}$	$C$
6	9	$\{0.6/a, 0.6/b, 1.0/e\}$
7	10	$\{0.7/a, 0.7/c, 1.0/f\}$
$\perp$	11	$\{0.6/a, 0.7/b, 1.0/g\}$

phận của  $s(S)$  được tham chiếu bởi  $f$  được trình bày như sau:

**Thuật toán xác định quan hệ mờ**

Vào:  $f : \{\delta_1, \delta_2, \dots, \delta_n\} \xrightarrow{G_f} \{\gamma_1, \gamma_2, \dots, \gamma_k\}$  với

$G_f = (V_f, E_f, \eta_f)$  và hàm gán nhãn  $v_f$ .

Ra:  $R_{fo}$  biểu diễn một bộ phận của  $s(S)$

**Phương thức:**

1. Đặt  $\delta'_i = \delta_i$  nếu  $\delta_i$  là tập thuộc tính,  $\delta'_i = \{id_o\}$  nếu  $\delta_i = O, O \in OT_s, \Delta' = \bigcup_{i=1}^n \delta'_i$ .

2. Đặt  $\gamma'_i = \gamma_i$  nếu  $\gamma_i$  là tập thuộc tính,  $\gamma'_i = \{id_o\}$  nếu  $\gamma_i = O, O \in OT_s, \Gamma' = \bigcup_{i=1}^k \gamma'_i$ .

(i) nếu  $f$  là một FOFD cục bộ tham chiếu đến kiểu đối tượng mờ,  $O, R_{fo} = R_{ext(O)}$ ;

(ii) nếu  $f$  là một FOFD toàn cục, với một ngưỡng tương đương  $\alpha \in [0, 1]$ ,  $R_{fo}$  được xác định như sau

Đặt  $\{O_1, \dots, O_k\}$  là tập các kiểu đối tượng mờ trạm thuộc  $V_f$  và  $\tau = \{id_{O_1}, id_{O_2}, \dots, id_{O_k}\}$  là tập các thuộc tính định danh của chúng. Với mỗi đỉnh  $O \in V_f$ , đặt  $\phi_O = \delta \cup \gamma \cup \{id_O\}$  nếu  $v_f(O) = (\delta, \gamma)$  và  $\phi_O = \{id_O\}$  nếu  $v_f(O)$  không xác định.

(a) Chọn một đỉnh bất đầu  $O \in V_f, R_{fo} = R_{ext(O)}[\phi_O]$

Với mỗi  $e \in E_f$

Nếu  $e = (O, O_i)$  and  $\eta_f(e) = r$  {

$$R_{fo} = R_{fo} \underset{\alpha}{\overset{f}{\bowtie}} R_{ext(r)}; E_f = E_f - \{e\}; }$$

$V_f = V_f - \{O\}$ ;

Với  $(V_f \neq \emptyset)$  {

Chọn một đỉnh  $O_j \in V_f$  sao cho  $id_{O_j} \in \Omega_{R_{fo}}$

$$R_{fo} = R_{fo} \underset{\alpha}{\overset{f}{\bowtie}} R_{ext(O_j)}[\phi_{O_j}];$$

Với mỗi  $e' \in E_f$

Nếu  $e' = (O_j, O_k)$  and  $\eta_f(e') = r'$

$$\{ R_{fo} = R_{fo} \underset{\alpha}{\overset{f}{\bowtie}} R_{ext(r')};$$

$$E_f = E_f - \{e'\}; }$$

$V_f = V_f - \{O_j\}$ ;

};

(b) Chuẩn hóa về phải: Nếu  $(\tau \neq \emptyset)$ , xóa tất cả các bộ không xác định trên  $\tau : R_{fo} = WNF(R_{fo}, \tau)$

(c) Chuẩn hóa phủ: Xóa tất cả các bộ được phủ bởi các bộ khác trên các thuộc tính thuộc về các kiểu của vế trái và vế phải của  $f : R_{fo} = \{t | \neg(\exists t' \in R_{fo}) \text{ mà } t' \text{ phủ } t \text{ với mức } \alpha \text{ trên tập thuộc tính } \Delta' \cup \Gamma'\}$ .

**end;**

**Ví dụ 2:** Giả sử ta có một FOFD  $f : \{A, O_3\} \xrightarrow{G_f} \{O_2\}$  của lược đồ CSDL HDT mờ (Hình

4.1). Với ngưỡng tương đương  $\alpha = 0.9$ ,  $R_{fo}$  được xác định bởi một chuỗi phép toán liên kết ngoài mờ đầy đủ bắt đầu từ  $O_3 : (((R_{ext(O_3)}[id_{O_3}] \underset{\alpha}{\overset{f}{\bowtie}} R_{ext(r_2)}[id_{O_2}]) \underset{\alpha}{\overset{f}{\bowtie}} R_{ext(O_2)}[id_{O_2}]) \underset{\alpha}{\overset{f}{\bowtie}}$



$$R_{ext(r_1)} \overset{f}{\underset{\alpha}{\bowtie}} R_{ext(O_1)}[A, id_{O_1}].$$

Với quan hệ mờ như ở trong Ví dụ 1 (bỏ đi thuộc tính  $C$ ) ta có:

$$R_{fo} = (R_{ext(O_3)}[id_{O_3}] \overset{f}{\underset{\alpha}{\bowtie}} R_{ext(r_2)}) \quad R_{fo} = (R_{fo} \overset{f}{\underset{\alpha}{\bowtie}} R_{ext(O_1)}[id_{O_2}])$$

$id_{O_2}$	$id_{O_3}$
6	9
7	10
$\perp$	11

$id_{O_2}$	$id_{O_3}$
6	9
7	10
5	$\perp$
8	$\perp$
4	$\perp$
$\perp$	11

$$R_{fo} = (R_{fo} \overset{f}{\underset{\alpha}{\bowtie}} R_{ext(r_1)})$$

$id_{O_2}$	$id_{O_3}$	$id_{O_1}$
6	9	1
7	10	2
5	$\perp$	1
8	$\perp$	3
4	$\perp$	$\perp$
$\perp$	11	$\perp$

$$R_{fo} = (R_{fo} \overset{f}{\underset{\alpha}{\bowtie}} R_{ext(O_1)}[A, id_{O_1}])$$

$id_{O_2}$	$id_{O_3}$	$id_{O_1}$	A
6	9	1	{1.0/a, 0.7/b, 0.6/c}
7	10	2	{1.0/a, 0.7/b, 0.7/c}
5	$\perp$	1	{1.0/a, 0.7/b, 0.6/c}
8	$\perp$	3	{0.6/a, 1.0/b, 0.7/c}
4	$\perp$	$\perp$	$\perp$
$\perp$	11	$\perp$	$\perp$

Bốn bộ đầu của  $R_{fo}$  có được từ điều kiện kết nối đầu tiên của phép liên kết ngoài mờ đầy đủ, bộ tiếp theo là bộ treo từ  $R_{ext(O_2)}$ , kết quả từ điều kiện 3. Bộ này biểu diễn một đối tượng mờ thuộc  $O_2$  không có các liên kết. Bộ thứ 5 được tạo ra bởi điều kiện cuối cùng của phép liên kết ngoài mờ đầy đủ. Quan hệ mờ  $R_{fo}$  sau khi thực hiện bước chuẩn hóa về phải, bộ thứ 5 sẽ được xóa khỏi  $R_{fo}$  do bộ này không được xác định trên  $id_{O_2}$ . Bước chuẩn hóa phủ không bộ nào bị xóa.

Việc xây dựng quan hệ mờ đảm bảo tất cả các đối tượng ở bên phải của trạng thái đã cho được biểu diễn trong  $R_{fo}$  nhưng loại bỏ đi các bộ thông qua việc chuẩn hóa về phải và chuẩn hóa phủ. Các bộ không xác định trên các thuộc tính của  $R_{fo}$  mà thuộc về các đối tượng mờ trạm, chúng mô tả các liên kết mà không có đối tượng của kiểu đối tượng về phải tham gia, việc loại bỏ chúng ra khỏi quan hệ  $R_{fo}$  không làm mất thông tin của các thể hiện của các kiểu đối tượng mờ ở về phải. Nếu hai bộ mô tả liên kết của một đối tượng về phải với các đối tượng về trái tồn tại, nếu bộ này được bao phủ bởi một bộ khác trong  $R_{fo}$  được loại bỏ bởi chuẩn hóa phủ.

### 5. CÁC DẠNG PHỤ THUỘC HÀM ĐỐI TƯỢNG MỜ

Theo cách tiếp cận của Raju [7], một phụ thuộc hàm mờ  $f : X \xrightarrow{f} Y$  đúng trên quan hệ  $r$  nếu và chỉ nếu với hai bộ bất kỳ thuộc quan hệ  $r$  độ gần nhau giữa hai giá trị của hai bộ trên  $Y$  luôn lớn hơn hoặc bằng độ gần nhau giữa hai giá trị của hai bộ trên  $X$ . Trong ngữ cảnh của CSDL HDT mờ, dựa trên quan hệ mờ, các khái niệm khác nhau liên quan đến tính đúng của các FOFD trong trạng thái của một lược đồ CSDL HDT mờ sẽ được trình bày trong mục này. Với một FOFD  $f : \Delta \xrightarrow{f} \Gamma$  chúng ta chỉ xem xét ngữ nghĩa của  $f$  trên tập các thuộc tính và các thuộc tính định danh trong quan hệ mờ  $R_{fo}$  mà thuộc về  $\Delta, \Gamma$ . Các thuộc tính khác liên quan đến các chuỗi liên kết giữa các đối tượng của  $\Delta$  và  $\Gamma$  không được xem xét mặc

dù chúng cũng được biểu diễn trong  $R_{f_o}$ . Với  $\Delta', \Gamma'$  và  $\tau = \{id_{O_1}, id_{O_2}, \dots, id_{O_k}\}$  được xác định như trong thuật toán xác định quan hệ mờ được trình bày ở trên, một FOFD được thỏa mãn hay yếu trong trạng thái lược đồ CSDL HDT được định nghĩa như sau:

**Định nghĩa 5.1.** Cho  $S$  là một lược đồ CSDL HDT mờ,  $s(S)$  là trạng thái của  $S$ ,  $\alpha$  là một ngưỡng tương đương, một FOFD  $f : \{\delta_1, \delta_2, \dots, \delta_n\} \xrightarrow{G_f} \{\gamma_1, \gamma_2, \dots, \gamma_k\}$  được thỏa mãn trong  $s(S)$  nếu thỏa mãn hai điều kiện sau đây:

- (i)  $(\forall t, t' \in NF(R_{f_o}, \Delta')) \Rightarrow SE(t[\Gamma'], t'[\Gamma']) \geq SE(t[\Delta'], t'[\Delta'])$
- (ii)  $R_{f_o}[\{id_{O_i}\}] = NF(R_{f_o}, \Gamma')[id_{O_i}]$  với  $id_{O_i} \in \tau, i = 1, 2, \dots, k$ .

Đối với CSDL HDT mờ, kiểu dữ liệu của các thuộc tính không đơn thuần là các kiểu cơ sở mà còn có các kiểu phức tạp khác như kiểu tập, kiểu bộ, ... Dựa trên thước đo độ tương đương ngữ nghĩa giữa hai giá trị mờ, và khái niệm về *hai đối tượng đồng nhất, hai đối tượng bằng nhau*, công thức xác định mức độ tương đương  $SE(\dots)$  của hai bộ trên tập thuộc tính  $X$  với quan hệ tương tự trên miền giá trị của các thuộc tính thuộc  $X$  trong cách tiếp cận này được tính tương tự như trong [2].

Ở điều kiện (i),  $f$  là hàm ánh xạ mỗi giá trị đầy đủ của  $\Delta'$  trong  $R_{f_o}$  vào đúng một giá trị của  $\Gamma'$ . Với các giá trị (tương ứng, các đối tượng) được đưa ra ở điểm đầu vào của mỗi  $\delta_i$ , nhiều nhất một tổ hợp các đối tượng về phải hay các giá trị được tìm thấy. Điều kiện (ii) đảm bảo tính chất toàn ánh khi tất cả các liên kết bộ phận và liên kết thiếu được loại bỏ. Do đó, với mỗi đối tượng mờ trạm, ít nhất một tổ hợp các đối tượng mờ nguồn (tương ứng, các giá trị của các đối tượng) phải tồn tại để thông qua đó các đối tượng trạm được xác định duy nhất.

**Ví dụ 3:** Cho  $f\{O_2\} \xrightarrow{G_f} \{A\}$  trên cùng lược đồ CSDL HDT mờ và trạng thái của lược đồ CSDL HDT mờ như Ví dụ 1. Theo thuật toán trên, quan hệ mờ  $R_{f_o}$  được xác định như sau:

$id_{O_2}$	$id_{O_1}$	$A$
6	1	{1.0/a, 0.7/b, 0.6/c}
7	2	{1.0/a, 0.7/b, 0.7/c}
5	1	{1.0/a, 0.7/b, 0.6/c}
8	3	{0.6/a, 1.0/b, 0.7/c}

Ta có, với hai bộ bất kỳ  $t, t' \in NF(R_{f_o}, \{id_{O_2}\})$  ta đều có  $SE(t[A], t'[A]) \geq SE(t[id_{O_2}], t'[id_{O_2}])$  và  $R_{f_o}[\{id_{O_1}\}] = NF(R_{f_o}, \{id_{O_2}\})[\{id_{O_1}\}]$ . Vậy  $f'$  là phụ thuộc hàm được thỏa mãn trong  $s(S)$ . Tương tự như phụ thuộc hàm trong CSDL quan hệ, điều này có nghĩa là với giá trị đầu vào là một đối tượng xác định thuộc kiểu đối tượng mờ  $O_2$  sẽ

xác định duy nhất một giá trị thuộc tính  $A$  của một đối tượng thuộc kiểu đối tượng  $O_1$  trên trạng thái của lược đồ CSDL HDT mờ.

**Định nghĩa 5.2.** Cho  $S$  là một lược đồ CSDL HDT mờ,  $s(S)$  là trạng thái của  $S$ ,  $\alpha$  là một ngưỡng tương đương, một FOFD  $f : \{\delta_1, \delta_2, \dots, \delta_n\} \xrightarrow{G_f} \{\gamma_1, \gamma_2, \dots, \gamma_k\}$  được thỏa yếu trong  $s(S)$  nếu:  $(\forall t, t' \in NF(R_{f_o}, \Delta')) \Rightarrow SE(t[\Gamma'], t'[\Gamma']) \geq SE(t[\Delta'], t'[\Delta'])$

**Ví dụ 4:** Với FOFD  $f'\{A, O_3\} \xrightarrow{G'_f} \{O_2\}$  như trong Ví dụ 2, ta có  $NF(R_{f_o}, \{id_{O_3}\})$  gồm 2 bộ  $(6, 9, 1, \{1.0/a, 0.7/b, 0.6/c\}), (7, 10, 2, \{1.0/a, 0.7/b, 0.7/c\})$ . Với hai bộ  $t, t' \in NF(R_{f_o}, \{id_{O_3}\})$  ta có  $SE(t[id_{O_2}], t'[id_{O_2}]) \geq SE(t[A, id_{O_3}], t'[A, id_{O_3}])$ . Theo định nghĩa,  $f'$  là được thỏa yếu trong trạng thái lược đồ đã cho và điều kiện (ii) trong định nghĩa 3 không được thỏa mãn.

Nếu một FOFD là được thỏa yếu, tất cả các đối tượng của các kiểu đối tượng mờ trạm được xác định duy nhất bằng việc sử dụng tất cả các  $\delta_i \in \Delta$  như các điểm đầu vào. Tuy nhiên, điều kiện đầy đủ trên  $\Delta$  là quá chặt nếu ta cần tìm kiếm khả năng khác để xác định

đối tượng như đã đề cập trong Hình 1.b. Vì vậy, các liên kết bộ phận từ đối tượng mờ trạm đến đối tượng mờ nguồn có thể được sử dụng để nhận biết các đối tượng.

**Định nghĩa 5.3.** Cho  $S$  là một lược đồ CSDL HDT mờ,  $s(S)$  là trạng thái của  $S$ ,  $\alpha$  là một ngưỡng tương đương, một FOFD  $f : \{\delta_1, \delta_2, \dots, \delta_n\} \xrightarrow{G_f} \{\gamma_1, \gamma_2, \dots, \gamma_k\}$  được thỏa mạnh trong  $s(S)$  theo liên kết bộ phận nếu thỏa mãn hai điều kiện sau đây:

- (i)  $(\forall t, t' \in NF(R_{fo}, \Delta')) \Rightarrow SE(t[\Gamma'], t'[\Gamma']) \geq SE(t[\Delta'], t'[\Delta'])$
- (ii)  $R_{fo}[\{id_{O_i}\}] = WNF(R_{fo}, \Gamma')[id_{O_i}]$  với  $id_{O_i} \in \tau, i = 1, 2, \dots, k$ .

Một FOFD được thỏa mạnh theo liên kết bộ phận biểu diễn một khía cạnh khác trong việc nhận biết đối tượng, ở đây, nhiều nhất một đối tượng ở vế phải có thể được tìm thấy với chỉ một số giá trị hay các đối tượng ở các điểm đầu vào  $\delta_1, \delta_2, \dots, \delta_n$ . Các đối tượng ở vế phải chỉ có liên kết thiếu hay không có liên kết đến các kiểu đối tượng mờ nguồn sẽ không được xem xét vì chúng không được tìm thấy từ bất kỳ các điểm đầu vào nào. Trong trường hợp này, nếu tồn tại các đối tượng như thế thì FOFD không được thỏa mạnh theo liên kết bộ phận. Một FOFD được thỏa yếu trong  $s(S)$  theo liên kết bộ phận nếu chỉ thỏa mãn điều kiện (i).

**Ví dụ 5:** Với  $f'' : \{A, O_3\} \xrightarrow{G''_f} \{O_2\}$  như trong Ví dụ 2,  $WNF(R_{fo}, \{id_{O_3}\})$  gồm các bộ sau:

Với hai bộ bất kỳ  $t, t' \in WNF(R_{fo}, \{A, id_{O_3}\})$  thì  $SE(t[id_{O_2}], t'[id_{O_2}]) \geq SE(t[A, id_{O_3}], t'[A, id_{O_3}])$  (thỏa mãn điều kiện i). Điều kiện (ii) không được thỏa mãn do đối tượng có  $id = 4$  thuộc  $O_2$  không có chuỗi liên kết nào đến các đối tượng của  $O_3$ . Vậy,  $f''$  không được thỏa mạnh trong  $s(S)$  theo liên kết bộ phận.

$id_{O_2}$	$id_{O_3}$	$id_{O_1}$	$A$
6	9	1	{1.0/a, 0.7/b, 0.6/c}
7	10	2	{1.0/a, 0.7/b, 0.7/c}
5	' $\perp$ '	1	{1.0/a, 0.7/b, 0.6/c}
8	' $\perp$ '	3	{0.6/a, 1.0/b, 0.7/c}

**Ví dụ 6:** Cho  $f''' : \{B, A\} \xrightarrow{G'''_f} \{O_1\}$  trên cùng lược đồ CSDL HDT mờ như trong Ví dụ 1.

Ta có  $WNF(R_{fo}, \{B, C\})$  gồm các bộ sau.

Với hai bộ  $t_1 = (6, \{1.0/a, 0.6/b, 0.6/c\}, 9, \{0.6/a, 0.6/b, 1.0/e\}, 1)$  và  $t_2 = (7, \{0.6/a, 0.6/c, 1.0/d\}, 10, \{0.7/a, 0.7/c, 1.0/f\}, 2)$ .

Ta có :  $SE(t_1[id_{O_1}], t_2[id_{O_1}]) = SE(\{1.0/a, 0.7/b, 0.6/c\}, \{1.0/a, 0.7/b, 0.7/c\}) = \min(\frac{2.3}{2.3}, \frac{2.3}{2.4}) =$

$0.96$ .  $SE(t_1[BC], t_2[BC]) = \min(SE(t_1[B], t_2[B]), SE(t_1[C], t_2[C])) = \min(\frac{1.2}{2.2}, \frac{0.6}{2.4}) = 0.25$ .

Suy ra,  $SE(t_1[id_{O_1}], t_2[id_{O_1}]) > SE(t_1[BC], t_2[BC])$ . Tương tự, với hai bộ khác của quan hệ mờ đều thỏa mãn điều kiện (i). Ngoài ra, ta còn có  $R_{fo}[\{id_{O_1}\}] = WNF(R_{fo}, BC)[\{id_{O_1}\}]$ . Vậy  $f'''$  được thỏa mạnh trong  $s(S)$  theo liên kết bộ phận.  $f'''$  có thể được xem như một cách thức xác định đối tượng  $O_1$  dựa vào giá trị.

Từ các định nghĩa về các dạng phụ thuộc hàm đối tượng mờ đã được trình bày ở trên, một phụ thuộc hàm mờ có dạng  $f : \Delta \xrightarrow{G_f} \{O\}$  có thể được xem như một công cụ nhận biết

đối tượng trong CSDL HDT mờ. Ở đây, một FOFD xác định một tập các điểm đầu vào có thể bao gồm cả các kiểu đối tượng mờ, nếu với một tập giá trị điểm đầu vào  $\delta_i (\delta_i \in \Delta)$  mà

$\delta_i \cap OT_s = \emptyset$  cho ta một cách thức nhận biết đối tượng chỉ dựa vào giá trị của các thuộc tính, đây là vấn đề được quan tâm nhiều bởi người sử dụng CSDL.

## 6. KẾT LUẬN

Tương tự như trong CSDL quan hệ, phụ thuộc hàm được xem như công cụ để biểu diễn các ràng buộc dữ liệu. Bài báo đã đề xuất một FOFD trên các thuộc tính của các kiểu đối tượng mờ có thể bao gồm cả các kiểu của chúng. Dựa trên tính đúng đắn của FOFD, với một giá trị đầu vào FOFD giúp ta xác định duy nhất một đối tượng trong mỗi trạng thái của lược đồ CSDL HDT mờ, nó được xem như cách thức xác định đối tượng trong mỗi trạng thái CSDL HDT mờ. Các luật suy dẫn của phụ thuộc hàm đối tượng mờ và nghiên cứu các ảnh hưởng của các phụ thuộc hàm đối tượng mờ trong quá trình thiết kế CSDL HDT mờ sẽ được nghiên cứu trong những công trình tiếp theo.

## TÀI LIỆU THAM KHẢO

- [1] B. S. Lee, Normalization in OODB Design, *ACM SIGMOD Record* **24**(3) (1995) 23–27.
- [2] Đ.V. Ban, H.C. Hà, V.Đ. Quảng, Chuẩn hóa các lớp đối tượng trong lược đồ CSDL hướng đối tượng mờ, *Tạp chí Tin học và Điều khiển học* **27**(2) (2011) 131–141.
- [3] Đ.V. Ban, H.C. Hà, V. D. Quang, Querying fuzzy object-oriented data based on fuzzy association algebra, *Proc. 3th Intl. Conf. on Knowledge and Systems Engineering (KSE2011), Hanoi-Vietnam, IEEE Computer Society Press*, 2011 (40–47).
- [4] K. D. Schewe and B. Thalheim, Fundamental concepts of object-oriented databases, *Acta Cybernetica*, **11** (1-2) (1993) 49–83.
- [5] H.J. Klein and J. Rasch, Value base identification and functional dependencies for object databases, *Proc. 3th Intl. Basque Workshop on Information Technology (BIWIT 97), Biarritz France, IEEE Computer Society Press*, 1997 (22–32).
- [6] M. Lacroix, A. Pirotte, Generalized Joins, *ACM SIGMOD Record* **29**(2) (1976) 5-16.
- [7] Raju et al, Fuzzy functional dependencies and lossless join decomposition of fuzzy relational database systems, *ACM TODS* **13**(2) (1998).
- [8] Vojtěch Merunka, Jiří Brožek, Martin Šebek, Martin Molhanec, Normalization rules of the object-oriented data model, *Proc. Intl. Workshop on Enterprises & Organizational Modeling and Simulation*, Amsterdam-Netherlands, 2009.
- [9] Zahir Tari, John Stokes, Stefano Spaccapietra, Object normal forms and dependency constraints for object-oriented schemata, *ACM Transactions on Database Systems* **22**(4) (1997) 513–569.
- [10] Zongmin Ma, *Fuzzy database modeling with XLM*, Springer, NewYork, 2005.
- [11] Z. M. Ma, W. J. Zhang, W. Y. Ma, Assessment of data redundancy in fuzzy relational databases based on semantic inclusion degree, *Information Processing Letters* **72** (1-2) (1999) 25–29.

Ngày nhận bài 28 - 12 - 2011

Nhận lại sau sửa 30 - 5 - 2012