Tạp chí Tin học và Điều khiển học, T.20, S.4 (2004), 329–342

CƠ SỞ ĐỐI TƯỢNG XÁC SUẤT MỜ VÀ PHÉP CHỌN

NGUYỄN HÒA[1], CAO HOÀNG TRỤ[2]

[1]*Khoa Tin học, Trường Đại học Mở - Bán công, Tp. Hồ Chí Minh*
[2]*Khoa Công nghệ thông tin, Trường Đại học Bách khoa Tp. Hồ Chí Minh*

Abstract. This article introduces a fuzzy and probabilistic object base model (FPOB) that develops the probabilistic object base model (POB) by Eiter and his fellows (2001). In order to develop POB into FPOB, a probabilistic interpretation of relations on fuzzy set values is proposed to integrate them into that probability-based framework. Then the definitions of fuzzy-probabilistic tuple values, fuzzy-probabilistic object base schemas, inheritance, instances, and the selection operation, which is a basic algebraic operation on data/object bases, are presented.

Tóm tắt. Bài báo giới thiệu một mô hình cơ sở đối tượng xác suất mờ (fuzzy and probabilistic object base - FPOB) là mở rộng mô hình cơ sở đối tượng xác suất (probabilistic object base - POB) của Eiter và các cộng sự (2001). Để mở rộng từ POB sang FPOB, một diễn dịch xác suất của các quan hệ hai ngôi trên các giá trị tập mờ được đề nghị nhằm liên kết chúng vào trong một khung làm việc dựa trên cơ sở xác suất. Sau đó các khái niệm giá trị bộ xác suất mờ, lược đồ cơ sở đối tượng xác suất mờ, thừa kế, thể hiện và phép chọn, một phép toán đại số cơ bản trên cơ sở dữ liệu/đối tượng được trình bày.

1. GIỚI THIỆU

Như chúng ta đã biết, mô hình hướng đối tượng đã chứng tỏ các ưu điểm của nó trong các vấn đề mô hình hóa, thiết kế và hiện thực các hệ thống lớn, từ phần mềm cho đến cơ sở dữ liệu. Để biểu diễn và suy luận trên các thông tin không chắc chắn và không rõ ràng phổ biến trong thực tế, nhiều nghiên cứu đã và đang được tiến hành để mở rộng mô hình hướng đối tượng cổ điển bằng cách áp dụng các kết quả của lý thuyết xác suất và lý thuyết tập mờ. Một số nghiên cứu đã mở rộng cho phép giá trị thuộc tính của đối tượng như là một tập mờ ([4, 9, 12]). Một số nghiên cứu dùng xác suất để biểu diễn tính không chắc chắn của sự có mặt của một thuộc tính trong một lớp đối tượng và áp dụng lí thuyết tập mờ để biểu diễn các giá trị mờ trong thuộc tính đó ([5, 6, 7, 11]). Một số nghiên cứu khác quan tâm đến cây phân cấp các lớp đối tượng bằng cách xem xét xác suất để một đối tượng thuộc lớp cha là thuộc về một lớp con, dẫn đến khái niệm cây phân cấp với xác suất ([10]). Trong bài báo này, chúng tôi giới thiệu một mô hình cơ sở đối tượng trong đó các giá trị thuộc tính mờ không chắc chắn và cây phân cấp xác suất đồng thời được xem xét.

Để có thể xây dựng được một mô hình hướng đối tượng trong đó sự phân cấp lớp có tính xác suất đồng thời với việc biểu diễn và xử lí được giá trị thuộc tính mờ không chắc chắn, chúng tôi đã tích hợp giá trị thuộc tính mờ vào trong mô hình cơ sở đối tượng xác suất của Eiter và các cộng sự [10]. Để có thể tính toán được xác suất xuất hiện các giá trị thuộc tính mờ, một diễn dịch xác suất của các quan hệ hai ngôi trên các tập mờ đã được đề nghị trong Mục 3. Sau đó các Mục 4, 5 và 6 lần lượt giới thiệu một sự mở rộng tổng quát của các định nghĩa lược đồ, thể hiện và phép chọn trên POB cho FPOB. Mục 7 là kết luận và đề nghị các nghiên cứu trong tương lai.

2. KIẾN TRÚC CỦA MỘT CƠ SỞ ĐỐI TƯỢNG XÁC SUẤT MỜ

Trong phần này, chúng tôi mô tả toàn bộ kiến trúc của một hệ thống FPOB. Hình 1 giới thiệu một kiến trúc cho quá trình xử lí truy vấn trong FPOB. Kiến trúc này bao gồm các thành phần sau:

1. Người sử dụng (NSD) biểu diễn truy vấn thông qua khối giao diện với hệ thống.

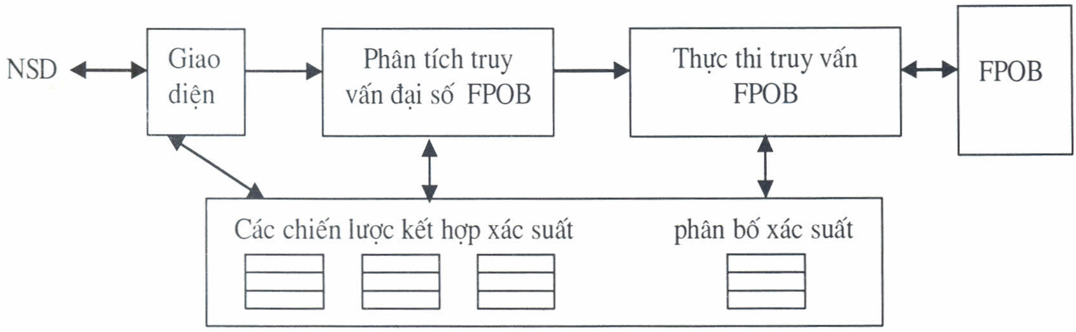
2. Các truy vấn được đưa vào khối phân tích truy vấn để chuyển thành các truy vấn đại số FPOB.

3. Các truy vấn đại số được chuyển vào khối thực thi trên FPOB.

4. Tất cả các bộ phận trên sử dụng các thư viện bao gồm:

(a) Một tập các chiến lược hội, tuyển và trừ xác suất ([10, 13]) cho phép người dùng biểu diễn thông tin về sự phụ thuộc của các sự kiện.

(b) Một tập các hàm phân bố xác suất cho phép người dùng chỉ ra cách thức mà xác suất được phân bố trên một không gian các giá trị thuộc tính.



Hình 1. Kiến trúc của hệ thống FPOB

3. DIỄN DỊCH XÁC SUẤT CỦA CÁC QUAN HỆ TRÊN CÁC TẬP MỜ

Để mở rộng mô hình POB với các giá trị tập mờ, chúng tôi áp dụng sự diễn dịch mô hình bầu cử của các tập mờ ([1, 2, 11]). Nghĩa là, đối với mỗi tập mờ A trên một miền U, mỗi cử tri đề cử một tập con của U như là một sự định nghĩa cổ điển của riêng họ về khái niệm mà A biểu diễn. Chẳng hạn, một cử tri có thể đề cử khoảng $[0, 35]$, 0 đến 35 tuổi, để biểu diễn cho khái niệm trẻ của con người, trong khi cử tri khác đề nghị khoảng $[0, 25]$ chứ không phải là $[0, 35]$.

Khi đó giá trị hàm thành viên $\mu_A(u)$ là phần trăm số cử tri mà định nghĩa cổ điển của họ cho khái niệm A có chứa u. Như vậy, A định nghĩa một phân bố xác suất trên các tập con của U và do đó, một mệnh đề mờ $x \in A$ xác định một họ các phân bố xác suất của biến x trên U.

Chúng ta xem ví dụ về trò chơi súc sắc trong [2]. Cho giá trị các mặt con súc sắc trong tập $\{1, 2, 3, 4, 5, 6\}$, giả sử rằng một điểm *cao* được định nghĩa bởi tập mờ rời rạc $\{3 : 0,2, 4 : 0,5, 5 : 0,9, 6 : 1\}$, nghĩa là, mức độ thành viên của giá trị 3 là 0,2 , của 4 là 0,5... Một sự bầu cử của một nhóm 10 người cho điểm *cao* này có thể như sau:

P_1	P_2	P_3	P_4	P_5	P_6	P_7	P_8	P_9	P_{10}
6	6	6	6	6	6	6	6	6	6
5	5	5	5	5	5	5	5	5	
4	4	4	4	4					
3	3								

Nghĩa là, tất cả các cử tri từ P_1 đến P_{10} đều bầu cho giá trị 6 như là một điểm cao, trong khi chỉ có hai cử tri P_1 và P_2, bầu cho 3 như là một điểm *cao*.... Nói cách khác, sự định nghĩa cổ điển của P_{10} cho điểm *cao* là $\{6\}$ trong khi của P_1 và P_2 là $\{3, 4, 5, 6\}$. Một giả định trong mô hình bầu cử này là bất kì một người nào chấp nhận một giá trị như là một điểm *cao* cũng sẽ chấp nhận tất cả các giá trị có mức độ thành viên lớn hơn trong tập mờ *cao*.

Mô hình này định nghĩa một phép gán khối, nghĩa là một phân bố xác suất, trên tập các tập con của $\{1, 2, 3, 4, 5, 6\}$ như sau:

$$\{6\} : 0{,}1, \{5,6\} : 0{,}4, \{4,5,6\} : 0{,}3, \{3,4,5,6\} : 0{,}2$$

trong đó khối (nghĩa là giá trị xác suất) được gán cho một tập con của $\{1, 2, 3, 4, 5, 6\}$ (chẳng hạn $m_{cao}(\{5, 6\}) = 0{,}4$) là tỉ lệ các cử tri đề cử tập con này như là một định nghĩa cổ điển cho khái niệm mờ điểm *cao*. Như đã chỉ ra trong [1], phép gán khối này tương ứng với một họ các phân bố xác suất trên $\{1, 2, 3, 4, 5, 6\}$.

Trên cơ sở mô hình bầu cử này, một sự diễn dịch xác suất của các quan hệ hai ngôi sau đây trên các tập mờ đã được đề nghị ([8]). Kí hiệu $Pr(E_1/E_2)$ được sử dụng để biểu diễn xác suất có điều kiện của E_1 khi đã có E_2.

Định nghĩa 3.1. Giả sử A, B là các tập mờ tương ứng trên các miền U và V, θ là một quan hệ hai ngôi từ $\{=, \leq, <, \subseteq, \in\}$. Diễn dịch xác suất của quan hệ $A\theta B$, kí hiệu $\mathrm{prob}(A\theta B)$ là một giá trị trong khoảng $[0, 1]$ được định nghĩa bởi:

$$\sum_{S\subseteq U, T\subseteq V} Pr(u\theta v \,|\, u \in S, v \in T).m_A(S).m_B(T).$$

Một cách trực quan, cho một mệnh đề mờ $x \in A$ và $y \in B$, $\mathrm{prob}(A\theta B)$ là xác suất để $x\theta y$ là đúng. Tính hợp lý của diễn dịch xác suất trên là, đối với mỗi định nghĩa cổ điển S của A và T của B, xác suất có điều kiện $u\theta v$ khi đã biết $u \in S$ và $v \in T$ được tính toán và nhân với hệ số là tích của các khối được kết hợp với S và T. Khi đó $\mathrm{prob}(A\theta B)$ là tổng của các giá trị xác suất có điều kiện được tích lũy như vậy. Ta cũng qui ước $\mathrm{prob}(A \geq B) = \mathrm{prob}(B \leq A), \mathrm{prob}(A > B) = \mathrm{prob}(B < A), \mathrm{prob}(A \supseteq B) = \mathrm{prob}(B \subseteq A)$ và $\mathrm{prob}(A \ni B) = \mathrm{prob}(B \in A)$.

Ví dụ 3.1. Trong ví dụ về súc sắc ở trên, giả sử *gần_5* được định nghĩa bởi tập mờ $\{6 : 0,3, 5 : 1, 4 : 0,3\}$, phép gán khối ứng với nó là: $\{5\}$: 0,7, $\{4, 5, 6\}$: 0,3.

Cho $x \in$ *gần_5* và $y \in$ *cao*, $\mathrm{prob}(\textit{gần\_5} = \textit{cao})$ đo xác suất $x = y$ được tính như sau:

$$
\begin{aligned}
\text{prob}(\textit{gần\_5} = \textit{cao}) &= Pr(u = v \mid u \in \{5\}, v \in \{6\}) \times m_{\textit{gần\_5}}(\{5\}) \times m_{cao}(\{6\}) \\
&+ Pr(u = v \mid u \in \{5\}, v \in \{5,6\}) \times m_{\textit{gần\_5}}(\{5\}) \times m_{cao}(\{5,6\}) \\
&+ Pr(u = v \mid u \in \{5\}, v \in \{4,5,6\}) \times m_{\textit{gần\_5}}(\{5\}) \times m_{cao}(\{4,5,6\}) \\
&+ Pr(u = v \mid u \in \{5\}, v \in \{3,4,5,6\}) \times m_{\textit{gần\_5}}(\{5\}) \times m_{cao}(\{3,4,5,6\}) \\
&+ Pr(u = v \mid u \in \{4,5,6\}, v \in \{6\}) \times m_{\textit{gần\_5}}(\{4,5,6\}) \times m_{cao}(\{6\}) \\
&+ Pr(u = v \mid u \in \{4,5,6\}, v \in \{5,6\}) \times m_{\textit{gần\_5}}(\{4,5,6\}) \times m_{cao}(\{5,6\}) \\
&+ Pr(u = v \mid u \in \{4,5,6\}, v \in \{4,5,6\}) \times m_{\textit{gần\_5}}(\{4,5,6\}) \times m_{cao}(\{4,5,6\}) \\
&+ Pr(u = v \mid u \in \{4,5,6\}, v \in \{3,4,5,6\}) \times m_{\textit{gần\_5}}(\{4,5,6\}) \times m_{cao}(\{3,4,5,6\}) \\
&= 0 \times 0{,}7 \times 0{,}1 + 1/2 \times 0{,}7 \times 0{,}4 + 1/3 \times 0{,}7 \times 0{,}3 + 1/4 \times 0{,}7 \times 0{,}2 \\
&+ 1/3 \times 0{,}3 \times 0{,}1 + 1/3 \times 0{,}3 \times 0{,}4 + 1/3 \times 0{,}3 \times 0{,}3 + 1/4 \times 0{,}3 \times 0{,}2 = 0{,}34.
\end{aligned}
$$

Chúng ta nhận xét rằng diễn dịch xác suất trên cũng có thể được làm tương thích cho các tập trên các miền liên tục, bằng cách sử dụng tích phân thay cho tổng như trong định nghĩa xác suất có điều kiện trong [3]. Nghĩa là $\text{prob}(A\theta B) = \int_0^1 \int_0^1 Pr(u\theta v | u \in^x A, v \in^y B)dxdy$, trong đó ${}^xA$ và ${}^yB$ là các lát cắt α ([13]) của tập mờ A và B tương ứng với $\alpha = x$ và $\alpha = y$.

Định nghĩa 3.2. Giả sử A và B là hai tập mờ trên một miền U. Diễn dịch xác suất của quan hệ $A \to B$, ký hiệu $\text{prob}(A \to B)$, là một giá trị trong [0, 1] được định nghĩa bởi:

$$\sum_{S,T \subseteq U} Pr(u \in T | u \in S).m_A(S).m_B(T).$$

Cho một mệnh đề mờ $x \in A$, ý nghĩa trực giác của $\text{prob}(A \to B)$ là giá trị xác suất để $x \in B$ là đúng. Nói cách khác, đó là xác suất có điều kiện để $x \in B$ khi đã biết $x \in A$ như được định nghĩa trong [2]. Chúng ta cũng qui ước $\text{prob}(A \leftarrow B)= \text{prob}(B \to A)$.

Ví dụ 3.2. Trong ví dụ xúc xắc, ta có:

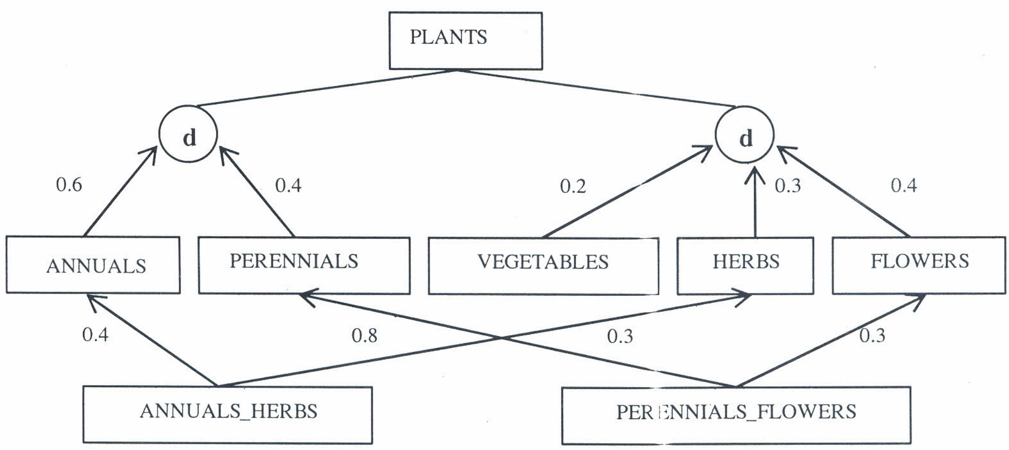
$$
\begin{aligned}
\text{prob}(\textit{cao} \to \textit{gần\_5}) &= Pr(u \in \{5\} \mid u \in \{6\}) \times m_{cao}(\{6\}) \times m_{\textit{gần\_5}}(\{5\}) \\
&+ Pr(u \in \{5\} \mid u \in \{5,6\}) \times m_{cao}(\{5,6\}) \times m_{\textit{gần\_5}}(\{5\}) \\
&+ Pr(u \in \{5\} \mid u \in \{4,5,6\}) \times m_{cao}(\{4,5,6\}) \times m_{\textit{gần\_5}}(\{5\}) \\
&+ Pr(u \in \{5\} \mid u \in \{3,4,5,6\}) \times m_{cao}(\{3,4,5,6\}) \times m_{\textit{gần\_5}}(\{5\}) \\
&+ Pr(u \in \{4,5,6\} \mid u \in \{6\}) \times m_{cao}(\{6\}) \times m_{\textit{gần\_5}}(\{4,5,6\}) \\
&+ Pr(u \in \{4,5,6\} \mid u \in \{5,6\}) \times m_{cao}(\{5,6\}) \times m_{\textit{gần\_5}}(\{4,5,6\}) \\
&+ Pr(u \in \{4,5,6\} \mid u \in \{4,5,6\}) \times m_{cao}(\{4,5,6\}) \times m_{\textit{gần\_5}}(\{4,5,6\}) \\
&+ Pr(u \in \{4,5,6\} \mid u \in \{3,4,5,6\}) \times m_{cao}(\{3,4,5,6\}) \times m_{\textit{gần\_5}}(\{4,5,6\}) \\
&= 0 \times 0{,}1 \times 0{,}7 + 1/2 \times 0{,}4 \times 0{,}7 + 1/3 \times 0{,}3 \times 0{,}7 + 1/4 \times 0{,}2 \times 0{,}7 \\
&1{,}0 \times 0{,}1 \times 0{,}3 + 1{,}0 \times 0{,}4 \times 0{,}3 + 1{,}0 \times 0{,}3 \times 0{,}3 + 3/4 \times 0{,}2 \times 0{,}3 = 0{,}53
\end{aligned}
$$

Chúng ta nhận xét rằng diễn dịch xác suất trên cũng có thể được làm tương thích cho các tập mờ trên các miền liên tục, bằng cách sử dụng tích phân thay cho tổng như trong định nghĩa xác suất có điều kiện trong [3]. Nghĩa là $\text{prob}(A \to B) = \int_0^1 \int_0^1 \frac{Pr({}^xA \cap {}^yB)}{Pr({}^xA)} dxdy$

$= \int_0^1 \int_0^1 \frac{|^xA \cap {}^yB|}{|^xA|} dxdy$, trong đó ^xA và ^yB là các lát cắt α của tập mờ A và B tương ứng với $\alpha = x$ và $\alpha = y$.

4. CÁC KIỂU VÀ LƯỢC ĐỒ FPOB

Đối với FPOB chúng tôi sử dụng cùng định nghĩa sự phân cấp lớp như trong POB. Hình 2 chỉ ra một ví dụ về sự phân cấp các thực vật (Plants) trong [10] mà được phân loại như là *Perennials* hoặc *Annuals* hay như là *Vegetables, Herbs* hoặc *Flowers*. Các lớp con như vậy của một lớp được liên kết với một nút ***d*** là loại trừ lẫn nhau (nghĩa là một đối tượng không thể thuộc về hai lớp tại cùng một thời điểm). Trong ví dụ này lớp PLANTS có hai nhóm lớp con là {ANNUALS, PERENNIALS} và {VEGETABLES, HERBS, FLOWERS}. Các giá trị số trong $[0, 1]$ trên các cung liên kết giữa một lớp với lớp con trực tiếp của nó biểu diễn xác suất có điều kiện để một đối tượng thuộc lớp cha là thuộc lớp con của nó. Chẳng hạn, sự phân cấp này chỉ ra rằng một đối tượng bất kì của PLANTS có 60% khả năng thuộc về ANNUALS trong khi chỉ có 40% khả năng còn lại thuộc về PERENNIALS. Ngoài ra chúng ta cũng để ý rằng một thực vật có thể vừa thuộc ANNUALS và HERBS như một đối tượng thuộc lớp ANNUALS_HERBS.



Hình 2. Một ví dụ về sự phân cấp lớp trong FPOB

Cũng như trong mô hình hướng đối tượng cổ điển, mỗi lớp trong POB được đặc trưng bởi một số thuộc tính mà các giá trị của chúng có các kiểu tương ứng nào đó. Đối với POB các kiểu có thể là các kiểu cơ sở, các kiểu tập hợp hoặc các kiểu bộ. Trong FPOB chúng tôi mở rộng các kiểu tập thành các kiểu tập hợp mờ và giá trị các kiểu này thành giá trị tập hợp mờ như trong các định nghĩa sau.

Định nghĩa 4.1. Giả sử A là một tập các thuộc tính và T là một tập các *kiểu cơ sở*. Các kiểu được định nghĩa một cách qui nạp như sau:

(1) Mọi kiểu cơ sở trong T là một kiểu.

(2) Nếu τ là một kiểu, thì $\{\tau\}$ là một *kiểu tập hợp mờ* của τ, được gọi là *kiểu tập hợp*.

(3) Nếu $A_1, ..., A_k$ là các thuộc tính đôi một khác nhau trong A và $\tau_1, ..., \tau_k$ là các kiểu, thì $\tau = [A_1 : \tau_1, ..., A_k : \tau_k]$ là một kiểu, được gọi là kiểu bộ trên tập các thuộc tính $\{A_1, ..., A_k\}$.

Với một kiểu $\tau = [A_1 : \tau_1, ..., A_k : \tau_k]$, chúng ta sử dụng $\tau.A_i$ để biểu thị τ_i. Ta gọi $A_1, ..., A_k$ là các thuộc tính ở mức cao nhất.

Ví dụ 4.1. Trong cơ sở đối tượng các thực vật đã giới thiệu ở trên, các thuộc tính có thể là soil, sun, water mô tả các điều kiện cho một thực vật phát triển và một số thuộc tính khác như name, size, width và height.... Một số kiểu cơ sở trong T có thể là integer, real, string, và soiltype. Một số kiểu tập mờ và kiểu bộ có thể là {real}, [soil: soiltype, sun: {real}, water: integer] và [name: string, soil: soiltype, sun: {real}, water: integer, size: [height: integer, width: integer]].

Định nghĩa 4.2. Mỗi kiểu cơ bản $\tau \in T$ có một miền xác định $\mathrm{dom}(\tau)$ kết hợp với nó. Giá trị được định nghĩa một cách qui nạp như sau:

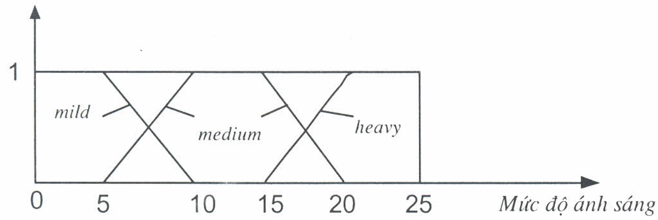
(1) Với mỗi kiểu cơ bản $\tau \in T$, thì mỗi $v \in \mathrm{dom}(\tau)$ là một giá trị kiểu τ.

(2) Với mỗi $\tau \in T$, mỗi tập mờ trên $\mathrm{dom}(\tau)$ là một giá trị kiểu $\{\tau\}$.

(3) Nếu $A_1, ..., A_k$ là các thuộc tính đôi một khác nhau trong A và $v_1, ..., v_k$ là các giá trị tương ứng của các kiểu $\tau_1, ..., \tau_k$ thì $[A_1 : v_1, ..., A_k : v_k]$ là một giá trị kiểu $[A_1 : \tau_1, ..., A_k : \tau_k]$.

Chúng ta cũng có thể coi một tập cổ điển A trên một miền U như là một tập mờ đặc biệt A_f trên U với hàm thành viên được định nghĩa bởi $\forall x \in U, A_f(x) = 1$ nếu $x \in A$ và $A_f(x) = 0$ nếu $x \notin A$. Tương tự như thế, mọi $v \in U$ cũng có thể coi như một tập mờ đặc biệt v_f trên U với hàm thành viên được định nghĩa bởi $\forall x \in U, v_f(x) = 1$ nếu $x = v$ và $v_f(x) = 0$ nếu $x \neq v$.

Ví dụ 4.2. Giả sử soiltype là kiểu liệt kê mà dom(soiltype) = {loamy, swampy, sandy} và mild, medium, heavy là các nhãn ngôn ngữ (linguistic labels) của các tập mờ trên dom(real) như chỉ ra trong hình 3. Thì, [soil: {loamy, swampy}, sun: mild, water: 3] là một giá trị kiểu [soil: {soiltype}, sun: {real}, water: integer].



Hình 3. Các giá trị tập mờ của thuộc tính sun

Trong POB, giá trị của mỗi thuộc tính có thể không chắc chắn và được ước lượng bởi các hàm phân bố xác suất cận dưới và cận trên trên một tập giá trị. Đối với FPOB, chúng tôi sẽ mở rộng định nghĩa giá trị bộ xác suất trong [10] để biểu diễn thông tin không chắc chắn cho các giá trị tập mờ.

Định nghĩa 4.3. Nếu $A_1, ..., A_k$ là các thuộc tính đôi một phân biệt trong A và $\langle V_1, \alpha_1, \beta_1\rangle, ..., \langle V_k, \alpha_k, \beta_k\rangle$ là các bộ ba xác suất, ở đó $V_1, ..., V_k$ là tập các giá trị của các kiểu $\tau_1, ..., \tau_k$ thì biểu thức $[A_1 : \langle V_1, \alpha_1, \beta_1\rangle, ..., A_k : \langle V_k, \alpha_k, \beta_k\rangle]$ là một giá trị bộ xác suất mờ của kiểu $[A_1 : \tau_1, ..., A_k : \tau_k]$ trên tập thuộc tính $\{A_1, ..., A_k\}$. Với mỗi giá trị bộ xác xuất $ptv = [A_1 : \langle V_1, \alpha_1, \beta_1\rangle, ..., A_k : \langle V_k, \alpha_k, \beta_k\rangle]$ ta sử dụng $ptv.A_i$ để biểu thị $\langle V_i, \alpha_i, \beta_i\rangle$.

Ý nghĩa trực giác của bộ ba xác suất $\langle V_i, \alpha_i, \beta_i \rangle$ là cho biết khoảng xác suất, biểu thị qua α_i và β_i, để một đối tượng có thuộc tính A_i nhận giá trị v_i trong V_i.

Ví dụ 4.3. Giả sử, chúng ta biết loại đất thích hợp của một thực vật là *loamy*. Hơn nữa chúng ta chắc chắn rằng thực vật này là Thyme, nhưng không chắc nó là French Thyme, Silver Thyme hay Wooly Thyme. Tuy nhiên, nếu chúng ta đảm bảo rằng 20-60% khả năng đó là một trong ba loại French Thyme, Silver Thyme hay Wooly Thyme thì chúng ta có thể biểu diễn thông tin này bởi một giá trị bộ xác suất mờ [**soil**: $\langle\{loamy\}, u, u\rangle$, **category**: $\langle\{french, silver, wooly\}, 0{,}6u, 1{,}8u\}\rangle$]. Trong đó u là hàm phân bố chuẩn và $0{,}6\,u$ và $1{,}8\,u$ biểu diễn hàm phân bố xác suất α và hàm β sao cho $\alpha(x) = 0{,}6\,(1/3) = 0{,}2$ và $\alpha(x) = 1{,}8\,(1/3) = 0{,}6 \forall x \in \{french, silver, wooly\}$.

Bây giờ lược đồ FPOB được định nghĩa bằng cách mở rộng khái niệm lược đồ POB trong [10] như sau:

Định nghĩa 4.4. Một lược đồ cơ sở đối tượng xác suất mờ là một bộ năm

$$S = (C, \tau, \Rightarrow, me, \wp),$$

trong đó:

C là một tập hữu hạn các lớp (đó là các lớp được kết hợp với FPOB),

τ là ánh xạ từ C đến tập các kiểu bộ $\tau(c)$ (cho biết kiểu dữ liệu của mỗi lớp),

$\Rightarrow$ là một quan hệ hai ngôi trên C sao cho $(C, \Rightarrow)$ là một đồ thị có hướng không có chu trình. Mỗi một node của $(C, \Rightarrow)$ là một lớp trong C, một cạnh $c_1 \Rightarrow c_2$ nghĩa là c_1 là lớp con trực tiếp của c_2,

me là ánh xạ đặt tương ứng mỗi lớp $c \in C$ với một phân hoạch của tập tất cả các lớp con trực tiếp của c sao cho các lớp trong mỗi nhóm của $me(c)$ là tách rời và loại trừ lẫn nhau,

$\wp$ là ánh xạ đặt tương ứng mỗi cạnh trong $(C, \Rightarrow)$ với một số trong khoảng $[0, 1]$ sao cho $\sum_{d \in P} \wp(d, c) \leq 1, \forall c \in C, \forall P \in me(c)$.

Kí hiệu $c_1 \Rightarrow {}^*c_k$ với $k \geq 1$ nếu có một đường đi $c_1 \Rightarrow c_2 \Rightarrow \ldots \Rightarrow c_k$ trong $(C, \Rightarrow)$. Đặc biệt $c \Rightarrow^* c\ \forall c \in C$.

Ví dụ 4.4. Một lược đồ FPOB cho cơ sở đối tượng các thực vật đã nêu trên có thể được định nghĩa như sau: $C=$ {PLANTS, ANNUALS, PERENNIALS, VEGETABLES, HERBS, FLOWERS, ANNUALS_HERBS, PERENNIALS_FLOWERS}.

τ được cho trong Bảng 1.

$(C, \Rightarrow), me$ và $\wp$ được cho như Hình 2.

Tuy nhiên, lược đồ được định nghĩa như trên có thể là không nhất quán, nghĩa là không phải luôn luôn tìm được một tập các đối tượng thỏa mãn một phép gán xác suất và phân loại các lớp được biểu diễn bởi đồ thị $(C, \Rightarrow)$ và phân hoạch các cạnh. Một lược đồ FPOB là nhất quán nếu và chỉ nếu nó có một mô hình như trong định nghĩa sau đây được mở rộng từ [10].

Định nghĩa 4.5. Giả sử $S = (C, \tau, \Rightarrow, me, \wp)$ là một lược đồ FPOB. Một *diễn dịch* của S là một ánh xạ ε từ C đến tập các tập con hữu hạn của một tập O. Một diễn dịch ε của S là một *mô hình* của S nếu và chỉ nếu:

1. $\varepsilon(c) \neq \emptyset, \forall c \in C$,
2. $\varepsilon(c) \subseteq \varepsilon(d), \forall c, d \in C$ với $c \Rightarrow d$,
3. $\varepsilon(c) \cap \varepsilon(d) = \emptyset, \forall c, d \in C$ với c và d phân biệt và $c, d \in P \in me(C)$,
4. $|\varepsilon(c)| = \wp(c, d).|\varepsilon(d)|\ \forall c, d \in C$ với $c \Rightarrow d$.

Bảng 1. Sự gán kiểu τ

c	$\tau(c)$
PLANTS	[name: **string**, soil: **soiltype**, water: **integer**]
ANNUALS	[name: **string**, soil: **soiltype**, water: **integer**, sun: {**real**}]
PERENNILAS	[name: **string**, soil: **soiltype**, water: **integer**, sun: {**real**}, expyears: **integer**]
VEGETABLES	[name: **string**, soil: **soiltype**, water: **integer**, sun: {**real**}, expyears: **integer**]
HERBS	[name: **string**, soil: **soiltype**, water: **integer**, sun: {**real**}, expyears: **integer**, category: **string**]
FLOWERS	[name: **string**, soil: **soiltype**, water: **integer**, sun: {**real**}, expyears: **integer**, category: **string**]
ANNUALS_ HERBS	[name: **string**, soil: **soiltype**, water: **integer**, sun: {**real**}, expyears: **integer**, category: **string**]
PERENNIALS_ FLOWERS	[name: **string**, soil: **soiltype**, water: **integer**, sun: {**real**}, expyears: **integer**, category: **string**]

5. THỪA KẾ VÀ CÁC THỂ HIỆN FPOB

Đối với FPOB, để giải quyết mối quan hệ thừa kế nói chung và đa thừa kế nói riêng của các lớp con và lớp cha, chúng tôi áp dụng cùng một chiến lược thừa kế như trong POB. Cho một lược đồ $S = (C, \tau, \Rightarrow, me, \wp)$ áp dụng chiến lược thừa kế này trên S sẽ dẫn đến một lược đồ mời $S^* = (C, \tau^*, \Rightarrow, me, \wp)$ chỉ khác S ở phép gán kiểu τ^*. Cụ thể với mỗi $c \in C$, $\tau^*(c) = [A_1 : \tau(d_1).A_1, ..., A_k : \tau(d_k).A_k]$, trong đó $A_1, ..., A_k$ là các thuộc tính mức cao nhất được thừa kế bởi c thông qua chiến lược thừa kế này từ các lớp cha $d_1, ..., d_k$ tương ứng. Chúng ta nói S^* là sự thừa kế hoàn toàn của S. Một lược đồ S được gọi là thừa kế đầy đủ nếu và chỉ nếu $S = S^*$. Kể từ bây giờ chúng tôi giả sử rằng tất cả lược đồ FPOB là nhất quán và được thừa kế đầy đủ.

Cho một lược đồ S, một thể hiện FPOB trên S được định nghĩa như một cơ sở đối tượng phù hợp với các kiểu dữ liệu và chiến lược thừa kế trên lược đồ này. Một thể hiện trên lược đồ FPOB cho ta thấy được tình trạng, quan hệ thừa kế của một tập đối tượng trong FPOB. Định nghĩa sau đây được mở rộng từ định nghĩa thể hiện POB trong [10].

Định nghĩa 5.1. Giả sử $S = (C, \tau, \Rightarrow, me, \wp)$ là một lược đồ FPOB và O là một tập danh hiệu đối tượng, một thể hiện FPOB trên S là một cặp (π, v), trong đó:

1. $\pi : C \rightarrow 2^O$ là ánh xạ đặt tương ứng mỗi lớp c thuộc C với một tập con hữu hạn của tập O, sao cho $\pi(c_1) \cap \pi(c_2) = \emptyset, \forall c_1 \neq c_2 \in C$.

Ngoài ra ánh xạ $\pi^* : C \rightarrow 2^O$ được định nghĩa bởi $\pi^*(c) = \cup\{\pi(c') \,|\, c' \in C, c' \Rightarrow^* c\}$ là tập các đối tượng thuộc lớp c (các đối tượng trong c hoặc trong lớp con, cháu của c).

2. v là ánh xạ đặt tương ứng mỗi $o \in \pi(C)$ với một giá trị bộ xác suất mờ kiểu $\tau(c)$ sao cho $o \in \pi(c)$.

Ví dụ 5.1. Một thể hiện $I = (\pi, v)$ trên lược đồ FPOB trong Ví dụ 4.4 được chỉ ra trong bảng 2 và bảng 3.

Bảng 2. Ánh xạ π và π^*

c	$\pi(c)$	$\pi^*(c)$
PLANTS	$\{o_1\}$	$\{o_1, o_2, o_3, o_4, o_5, o_6, o_7\}$
ANNUALS	$\{\}$	$\{o_2, o_3, o_5, o_6, o_7\}$
PERENNIALS	$\{\}$	$\{o_4\}$
VEGETABLES	$\{\}$	$\{\}$
HERBS	$\{\}$	$\{o_2, o_3, o_5, o_6, o_7\}$
FLOWERS	$\{\}$	$\{o_4\}$
ANNUALS_HERBS	$\{o_2, o_3, o_5, o_6, o_7\}$	$\{o_2, o_3, o_5, o_6, o_7\}$
PERENNIALS_FLOWERS	$\{o_4\}$	$\{o_4\}$

Bảng 3. Phép gán v

oid	v(oid)
o_1	[name: ⟨{*Lady-fern, Ostrich-fern*}, u, u⟩, soil: ⟨{*loamy*}, u, u⟩, water: ⟨{25,..., 30}, u, u⟩]
o_2	[name: ⟨{*Cuban-Basil, Lemon-Basil*}, u, u⟩, soil: ⟨{*loamy, sandy*},0.7u, 1.3u⟩, water: ⟨{20,...,30}, u, u⟩, sun: ⟨{*mild, medium*}, 0.8u,1.2u⟩, expyears: ⟨{2, 3, 4}, 0.6u, 1.8u⟩, category: ⟨{*french, silver, wooly*}, 0.6u, 1.8u⟩]
o_3	[name: ⟨{*Mint*}, u, u⟩, soil: ⟨{*loamy*}, u, u⟩, water: ⟨{20}, u, u⟩, sun: ⟨{*mild*}, u, u⟩, expyears: ⟨{2, 3, 4}, 0.6u, 1.8u⟩, category: ⟨{*french, silver, wooly*}, 0.6u, 1.8u⟩]
o_4	[name: ⟨{*Aster, Salvia*}, u, u⟩, soil: ⟨{*loamy, sandy*}, 0.6u, 1.4u⟩, water: ⟨{20,..., 25}, u, u⟩, sun: ⟨{*mild*}, u, u⟩, expyears: ⟨{2, 3, 4}, 0.6u, 1.8u⟩, category: ⟨ {*french, silver, wooly*}, 0.6u, 1.8u⟩]
o_5	[name: ⟨{*Thyme*}, u, u⟩, soil : ⟨{*loamy*}, u, u⟩, water: ⟨{20,...,25}, u, u⟩, sun: ⟨{*mild, medium*}, 0.8u, 1.2u⟩, expyears: ⟨{2, 3}, 0.8u, 1.2u⟩, category: ⟨{*french, silver, wooly*}, 0.6u, 1.8u⟩]
o_6	[name: ⟨{*Mint*}, u, u⟩, soil: ⟨{*loamy*}, u, u⟩, water: ⟨{20}, u, u⟩, sun: ⟨{*mild*}, u, u⟩, expyears: ⟨{2, 3, 4}, 0.6u, 1.8u⟩, category: ⟨{*french, silver, wooly*}, 0.6u, 1.4u⟩]
o_7	[name: ⟨{*Sage*}, u, u⟩, soil: ⟨{*sandy*}, u, u⟩, water: ⟨{20,..., 21}, u, u⟩, sun: ⟨{*mild*}, u, u⟩, expyears: ⟨{2, 3, 4}, 0.6u, 1.8u⟩, category: ⟨{*red, tricolor*}, 0.6u, 1.4u⟩]

Trong cơ sở đối tượng cổ điển, phạm vi của một lớp bao gồm tất cả các đối tượng thuộc về lớp đó. Trong FPOB cũng như trong POB, phạm vi xác suất của một lớp chỉ ra xác suất để cho mỗi đối tượng thuộc về lớp đó. Định nghĩa sau đây được mở rộng từ [10].

Định nghĩa 5.2. Giả sử $I = (\pi, v)$ là một thể hiện trên một lược đồ FPOB $S = (C, \tau, \Rightarrow, me, \wp)$. Với mỗi lớp $c \in C$, phạm vi xác suất của c, được kí hiệu ext(c), là một ánh xạ đặt tương ứng mỗi $o \in \pi(C)$ với một tập các số hữu tỉ trong khoảng $[0, 1]$ sao cho:

1. Nếu $o \in \pi^*(c)$ thì ext$(c)(o) = \{1\}$.
2. Nếu $o \in \pi^*(d)$ và $\varepsilon(d) \cap \varepsilon(c) = \emptyset$ với mọi mô hình ε của S thì ext$(c)(o) = \{0\}$.
3. Ngược lại, ext$(c)(o) = \{p | p$ là tích của các xác suất các cạnh trên một đường từ c đến một lớp $d \in C$, ở đó d là lớp nhỏ nhất sao cho $o \in \pi^*(d)$ và $c \Rightarrow^* d\}$.

Ví dụ 5.2. Giả sử I là thể hiện của lược đồ FPOB trong Ví dụ 5.1, phạm vi xác suất của các lớp ANNUALS_HERBS và PERENNIALS_FLOWERS đối với o_1 và o_2 như sau:

$$\text{ext(ANNUALS\_HERBS)}(o_1) = \{0{,}24\} \ \text{ext(PERENNIALS\_FLOWERS)}(o_1) = \{0{,}12\},$$
$$\text{ext(ANNUALS\_HERBS)}(o_2) = \{1\} \ \text{ext(PERENNIALS\_FLOWERS)}(o_2) = \{0\}.$$

6. PHÉP TOÁN CHỌN TRONG FPOB

Cũng như trong các cơ sở đối tượng cổ điển, phép chọn là một phép toán đại số cơ bản trên FPOB. Nói theo trực giác, kết quả của một truy vấn chọn trên một thể hiện I của một lược đồ S là một thể hiện I' trên S sao cho các đối tượng của các lớp trong I' và các giá trị thuộc tính của nó thỏa mãn điều kiện chọn của truy vấn này.

Trước khi định nghĩa phép chọn trên một thể hiện, chúng tôi giới thiệu cú pháp và ngữ nghĩa hình thức của các điều kiện chọn. Chúng tôi bắt đầu với cú pháp của các biểu thức đường đi và biểu thức chọn. Định nghĩa sau đây về các biểu thức đường đi là sự mở rộng định nghĩa tương ứng được đưa ra trong [10].

Định nghĩa 6.1. Giả sử $\tau = [A_1 : \tau_1, ..., A_k : \tau_k]$ là một kiểu bất kì. Biểu thức đường đi được định nghĩa một cách qui nạp cho mọ*i* i từ 1 đến k như sau:

1. A_i là một biểu thức đường đi cho τ.
2. Nếu P_i là một biểu thức đường đi cho τ_i thì $A_i.P_i$ là một biểu thức đường đi cho τ.

Ví dụ 6.1. Cho τ= [name: **string**, size: [height: **integer**, width: **integer**]], thế thì name và size.height là các biểu thức đường đi cho kiểu này.

Đối với các biểu thức chọn trên FPOB, chúng tôi mở rộng các quan hệ hai ngôi trong các biểu thức chọn trên POB thành các quan hệ hai ngôi trên các tập mờ và đưa thêm vào quan hệ kéo theo trên các giá trị tập mờ, và được định nghĩa như sau.

Định nghĩa 6.2. Giả sử $S = (C, \tau, \Rightarrow, me, \wp)$ là một lược đồ FPOB và X là một tập các biến đối tượng. Các biểu thức chọn mờ được định nghĩa một cách qui nạp và có một trong các dạng sau:

1. $x \in c$, trong đó $x \in X$ và $c \in C$.
2. $x.P\theta v$, trong đó $x \in X, P$ là một biểu thức đường đi, θ là một quan hệ hai ngôi từ trong $\{=, \neq, \leq, \geq, <, >, \subseteq, \supseteq, \in, \ni, \rightarrow, \leftarrow\}$ và v là một giá trị.
3. $x.P_1 =_{\otimes} x.P_2$, ở đó $x \in X, P_1$ và P_2 là hai biểu thức đường đi và $\otimes$ là một chiến lược hội xác suất kết hợp các xác suất để $x.P_1 = v_1$ và $x.P_2 = v_2$ sao cho $v_1 = v_2$.
4. $\phi \otimes \Psi$, trong đó ϕ và Ψ là các biểu thức chọn trên cùng một biến đối tượng và $\otimes$ là một chiến lược hội xác suất kết hợp các xác suất đối với ϕ và Ψ là đúng.
5. $\phi \oplus \Psi$, trong đó ϕ và Ψ là các biểu thức chọn trên cùng một biến đối tượng và $\oplus$ là một chiến lược tuyển xác suất kết hợp các xác suất đối với ϕ và Ψ là đúng.

Ba dạng đầu của biểu thức chọn được gọi là các biểu thức chọn mờ cơ sở. Các chiến lược xác suất đã được giới thiệu trong [10, 13].

Ví dụ 6.2. Giả sử S là lược đồ của thể hiện trong Ví dụ 5.1, thì yêu cầu tìm "tất cả các thực vật cần mild sun" có thể được biểu diễn bởi biểu thức chọn cơ sở $x.\text{sun} \rightarrow$ *mild*. Hay tìm "tất cả các thực vật cần mild sun hoặc cần hơn 21 đơn vị nước mỗi ngày" có thể được biểu diễn bởi biểu thức chọx.sun $\rightarrow$ $mild \oplus x.water > 21$. Trong đó mild là một giá trị mờ trên dom$(real)$.

Bây giờ điều kiện chọn được định nghĩa như biểu thức chọn thỏa mãn một xác suất trong một khoảng như trong [10].

Định nghĩa 6.3. Giả sử $S = (C, \tau, \Rightarrow, me, \wp)$ là một lược đồ FPOB. Các điều kiện chọn mờ được định nghĩa một cách qui nạp như sau:

1. Nếu ϕ là một biểu thức chọn mờ và L, U các số thực trong đoạn $[0,1]$, $L \leq U$ thì $(\phi)[L, U]$ là một điều kiện chọn mờ.

2. Nếu α và β là các điều kiện chọn mờ trên cùng một biến đối tượng thì $\neg\alpha, (\alpha \wedge \beta), (\alpha \vee \beta)$ là các điều kiện chọn mờ.

Ví dụ 6.3. Gọi S là lược đồ của thể hiện trong Ví dụ 5.1, thì "chọn tất cả các thực vật cần mild sun với xác suất ít nhất là 40% và trên 21 đơn vị nước với một xác suất ít nhất là 80%" có thể được thực hiện bởi điều kiện chọn: $(x.\text{sun} \rightarrow mild)[0{,}4, 1] \wedge (x.\text{water} > 21)[0{,}8, 1]$.

Để định nghĩa ngữ nghĩa cho các điều kiện chọn, sự diễn dịch xác suất các biểu thức và điều kiện chọn mờ cần phải được định nghĩa. Nhưng trước hết chúng tôi giới thiệu diễn dịch cho các đường đi như trong [10].

Định nghĩa 6.4. Giả sử $\tau = [A_1 : \tau_1, \ldots, A_k : \tau_k]$ là một kiểu bộ. Diễn dịch của một biểu thức đường đi P đối với τ theo giá trị $v = [A_1 : v_1, \ldots, A_k : v_k]$, biểu thị là $v.P$, được định nghĩa qui nạp như sau:

1. Nếu $P = A_i$ thì $v.P = v_i$.

2. Nếu $P = A_i.P_i$ trong đó P_i là một biểu thức đường đi cho τ_i thì $v.P = v_i.P_i$.

Ví dụ 6.4. Cho τ = [name: string, size : [height: integer, width: integer]], thì diễn dịch của các đường đi name và size.height theo giá trị v = [name: *Thyme*, size: [height: 4, width: 12]] đối với kiểu τ tương ứng là các giá trị *Thyme* và 4.

Định nghĩa 6.5. Giả sử $S = (C, \tau, \Rightarrow, me, \wp)$ là một lược đồ FPOB, $I = (\pi, v)$ là một thể hiện trên S và $o \in \pi(C)$. *Diễn dịch xác suất* theo S, I và o, được biểu thị bởi $\text{prob}_{S,I,o}$, là một ánh xạ từ tập tất cả các biểu thức chọn mờ đến tập tất cả các khoảng con đóng của khoảng $[0,1]$ và được định nghĩa qui nạp như sau:

1. $\text{prob}_{S,I,o}(x \in c) = [\min(\text{ext}(c)(o)), \max(\text{ext}(c)(o))]$.

2. $\text{prob}_{S,I,o}(x.P\theta v) = [\sum_{u \in V} \alpha(u).\text{prob}(u.P'\theta v), \min(1, \sum_{u \in V} \beta(u).\text{prob}(u.P'\theta v))]$, trong đó $P = A.P'$ và $v(o).A = \langle V, \alpha, \beta \rangle$.

3. $\text{prob}_{S,I,o}(x.P_1 =_{\otimes} x.P_2) = [\sum_{u \in V} \alpha(u).\text{prob}(u_1.P'_1 = u_2.P'_2), \min(1, \sum_{u \in V} \beta(u).\text{prob}(u_1.P'_1 = u_2.P'_2))]$, trong đó $P_1 = A_1 P'_1, v(o)A_1 = \langle V_1, \alpha_1, \beta_1 \rangle$, $P_2 = A_2 P'_2, v(o).A_2 = \langle V_2, \alpha_2, \beta_2 \rangle$, $[\alpha(u), \beta(u)] = [\alpha_1(u_1), \beta_1(u_1)] \otimes [\alpha_2(u_2), \beta_2(u_2)]$, $\forall u = (u_1, u_2) \in V = V_1 \times V_2$.

4. $\text{prob}_{S,I,o}(\phi \otimes \Psi) = \text{prob}_{S,I,o}(\phi) \otimes \text{prob}_{S,I,o}(\Psi)$.

5. $\text{prob}_{S,I,o}(\phi \oplus \Psi) = \text{prob}_{S,I,o}(\phi) \oplus \text{prob}_{S,I,o}(\Psi)$.

Về trực giác, $\text{prob}_{S,I,o}(x \in c)$ là khoảng xác suất để đối tượng o thuộc về lớp c, $\text{prob}_{S,I,o}(x.A.P'\theta\nu)$ là khoảng xác suất để thuộc tính A của đối tượng o có giá trị u sao cho $u.P'\theta\nu$, và $\text{prob}_{S,I,o}(x.A_1.P'_1 =_{\otimes} x.A_2.P'_2$ là khoảng xác suất để thuộc tính A_1 và A_2 của o tương ứng có giá trị u_1 và u_2 sao cho $u_1P'_1 = u_2P'_2$. Cần lưu ý là P', P'_1 và P'_2 có thể là rỗng.

Trong trường hợp của POB, chúng ta có $\text{prob}(u.P'\theta\nu) = 1$ nếu $(u.P'\theta\nu)$ được thỏa, hoặc $\text{prob}(u.P'\theta\nu) = 0$ nếu ngược lại. Tương tự như thế, diễn dịch xác suất $\text{prob}(u_1P_1^{'} = u_2P_2^{'})$ $= 1$ nếu $u_1P_1^{'} = u_2P_2^{'}$ hoặc $\text{prob}(u_1P_1^{'} = u_2P_2^{'}) = 0$ nếu ngược lại.

Trong trường hợp của $FPOB$, các giá trị $u.P'$, $\nu, u_1.P_1^{'}$ và $u_2.P_2^{'}$ có thể là các tập mờ, khi đó $\text{prob}(u.P'\theta\nu)$ và $\text{prob}(u_1P_1^{'} = u_2P_2^{'})$ được tính toán như đã trình bày ở Phần 3 (xem thêm [8]). Như vậy, diễn dịch xác suất của các biểu thức chọn trong FPOB thực sự là mở rộng diễn dịch xác suất của các biểu thức chọn trong POB. Trong định nghĩa diễn dịch xác suất của các biểu thức chọn cho POB, các diễn dịch xác suất $\text{prob}(u.P'\theta\nu)$ và $\text{prob}(u_1.P'_1 = u_2.P'_2)$ không tham gia vào $\text{prob}_{S,I,o}$ vì thực sự giá trị của chúng bằng 0 hoặc 1. Nói cách khác, diễn dịch xác suất của các biểu thức chọn trong POB chỉ là trường hợp riêng của khái niệm diễn dịch xác suất của các biểu thức chọn trong FPOB.

Ví dụ 6.5. Trong thể hiện trên lược đồ FPOB ở Ví dụ 5.1, chúng ta có $\text{prob}_{S,I,o_2}(x.\text{sun} \to mild) = [0{,}8 \times u(mild) \times prob(mild \to mild) + 0{,}8 \times u(medium) \times prob(medium \to mild), \min(1, 1{,}2 \times u(mild) \times \text{prob}(mild \to mild) + 1{,}2 \times u(medium) \times prob(medium \to mild))] = [0{,}8 \times 1/2 \times 0{,}903 + 0{,}8 \times 1/2 \times 0{,}068,\ 1{,}2 \times 1/2 \times 0{,}903 + 1{,}2 \times 1/2 \times 0{,}068] = [0{,}39,\ \min(1, 0{,}59)] = [0{,}39,\ 0{,}59]$.

Chúng ta lưu ý là các diễn dịch: 0,903 và 0,068 được tính theo Định nghĩa 3.2 với các giá trị mờ *mild* và *medium* cho trong Ví dụ 4.2. Bây giờ khái niệm ngữ nghĩa cho các điều kiện chọn sẽ được mở rộng từ định nghĩa tương ứng trong [10], bằng cách gán cho mỗi điều kiện chọn một giá trị chân lý gọi là sự thỏa mãn điều kiện chọn. Cũng lưu ý là, theo [2] và [3], $\text{prob}(A \to A)$ không nhất thiết phải bằng 1, vì A là một tập mờ.

Định nghĩa 6.6. Giả sử $\mathbf{S} = (C, \tau, \Rightarrow, me, \wp)$ là một lược đồ FPOB, $\mathbf{I} = (\pi, \nu)$ là một thể hiện trên $\mathbf{S}$ và $0 \in \pi(C)$, sự thỏa mãn các điều kiện chọn mờ theo diễn dịch xác suất $\text{prob}_{\mathbf{S},\mathbf{I},o}$ được định nghĩa như sau:

1. $\text{prob}_{\mathbf{S},\mathbf{I},o}| = (\phi)[L, U]$ nếu và chỉ nếu $\text{prob}_{\mathbf{S},\mathbf{I},o}(\phi) \subseteq [L, U]$.
2. $\text{prob}_{\mathbf{S},\mathbf{I},o}| = \neg\phi$ nếu và chỉ nếu $\text{prob}_{\mathbf{S},\mathbf{I},o}| = \phi$ không thỏa.
3. $\text{prob}_{\mathbf{S},\mathbf{I},o}| = \phi \wedge \psi$ nếu và chỉ nếu $\text{prob}_{\mathbf{S},\mathbf{I},o}| = \phi$ và $\text{prob}_{\mathbf{S},\mathbf{I},o}| = \psi$.
4. $\text{prob}_{\mathbf{S},\mathbf{I},o}| = \phi \vee \psi$ nếu và chỉ nếu $prob_{\mathbf{S},\mathbf{I},o}| = \phi$ hoặc $\text{prob}_{\mathbf{S},\mathbf{I},o}| = \psi$.

Ví dụ 6.6. Trong thể hiện ở Ví dụ 5.1, dễ dàng tính được $prob_{\mathbf{S},\mathbf{I},o2}(x.sun \to mild \otimes_{in} x.water > 21) = [0{,}32, 0{,}48] \subseteq [0{,}3,\ 0{,}5]$ nên ta có $prob_{\mathbf{S},\mathbf{I},o2}| = (x.sun \to mild \otimes_{in} x.water > 21)\ [0{,}3,\ 0{,}5]$. Trong đó $\otimes_{in}$ là chiến lược hội xác suất độc lập ([10, 13]).

Định nghĩa 6.7. Giả sử $\mathbf{S} = (C, \tau, \Rightarrow, me, \wp)$ là một lược đồ FPOB, $\mathbf{I} = (\pi, \nu)$ là một thể hiện trên $\mathbf{S}$ và ϕ là một điều kiện chọn mờ trên biến đối tượng x. Phép chọn trên $\mathbf{I}$ theo ϕ, được kí hiệu $\sigma_\phi(\mathbf{I})$ là một thể hiện $\mathbf{I}' = (\pi^{'}, \nu^{'})$ trên $\mathbf{S}$, trong đó:

1. $\pi'(c) = \{0 \in \pi(c) | \text{prob}_{\mathbf{S},\mathbf{I},0}| = \phi\}$ và
2. $\nu' = \nu | \pi'(C)$ (nghĩa là ánh xạ ν thu hẹp trên $\pi'(C)$).

Ví dụ 6.7. Giả sử $\mathbf{I} = (\pi, \nu)$ là thể hiện của lược đồ **S** trong Ví dụ 5.1. Khi đó $\sigma_\phi(\mathbf{I})$ với $\phi = (x.\text{sun} \rightarrow \mathit{mild})[0{,}39, 1{,}0] \wedge (x.\text{water} > 21)[0{,}8, 1{,}0]$ là một thể hiện $\mathbf{I}'$ trên S gồm chỉ một đối tượng o_2. Thật vậy $\text{prob}_{\mathbf{S},\mathbf{I},o_2}(x.\text{sun} \rightarrow \mathit{mild}) = [0{,}39,\ 0{,}59]] \subseteq [0{,}39, 1{,}00]$ và $\text{prob}_{\mathbf{S},\mathbf{I},o_2}(x.\text{water} > 21) = [0{,}82, 0{,}82] \subseteq [0{,}8, 1{,}0]$ nên $\text{prob}_{\mathbf{S},\mathbf{I},o_2} \models \phi$ thỏa mãn. Kiểm tra tương tự, và thấy rằng các đối tượng khác không thỏa điều kiện chọn này.

Đối với các phép toán còn lại như phép chiếu, phép đổi tên, phép tích Cartesian, phép kết, phép giao, phép hợp và phép trừ các thể hiện FPOB cũng có thể được mở rộng tương ứng từ POB. Ngoài ra các tính chất của các phép toán đại số trên POB vẫn còn đúng trong FPOB ([13]).

7. KẾT LUẬN

Trong bài báo này, chúng tôi đã giới thiệu một cơ sở đối tượng xác suất mờ là mở rộng của POB. Để tích hợp giá trị tập mờ vào khung dựa trên xác suất của POB, chúng tôi đã dùng một mô hình bầu cử dựa trên xác suất của các tập mờ và giới thiệu một diễn dịch các quan hệ trên chúng. Các định nghĩa lược đồ, thể hiện FPOB cũng như phép toán chọn trên FPOB đã được mở rộng tương ứng từ POB. Mặc dù có thể dễ dàng đưa ra một số ví dụ mới, chúng tôi vẫn dựa trên các ví dụ của [10] đã được sửa đổi bằng cách tích hợp các giá trị thuộc tính mờ vào trong đó, với mục đích làm rõ sự khác biệt giữa FPOB và POB. Chúng tôi cũng đã hiện thực một chương trình minh họa bằng ngôn ngữ C++ cho phép toán chọn trên FPOB ([13]). Ngoài ra, trong [13] chúng tôi cũng đã mở rộng các phép toán đại số đối tượng khác từ POB sang FPOB như phép chiếu, phép tích Cartesian, phép kết nối, phép giao, phép hợp và phép trừ. Các kết quả lý thuyết và thực nghiệm cho thấy mô hình FPOB có thể phát triển thành một hệ thống quản trị cơ sở đối tượng thực sự, với khả năng xử lý các thông tin không chính xác và không chắc chắn thường gặp trong thực tiễn. Để có thể mở rộng hơn nữa mô hình FPOB, chúng tôi đang nghiên cứu nhằm tích hợp thêm các phương pháp (method) vào trong các lớp đối tượng và hiện thực một hệ thống FPOB hoàn chỉnh.

TÀI LIỆU THAM KHẢO

[1] Baldwin J. F., Lawry J. M., and Martin T. P., A mass assignment theory of the probability of fuzzy events, *Fuzzy Sets and Systems* **83** (1996) 353–367.

[2] Baldwin J. F., Martin T. P., and Pilsworth B. W., *Fril - Fuzzy and Evidential Reasoning in Artificial Intelligence*, Research Studies Press, 1995.

[3] Baldwin J. F, Lawry J. M., and Martin T. P., A note on probability/possibility consistency for fuzzy events, *Proceedings of the 6th International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems*, IPMU'96, Granada, Spain, 521–525, 1996.

[4] Bordogna G., Pasi G., and Lucarella D., A fuzzy object-oriented data model managing vague and uncertain information, *International Journal of Intelligent Systems* **14** (1999) 623–651.

[5] Cao T. H. and Rossiter J. M., A deductive probabilistic and fuzzy object-oriented database language, *Fuzzy Sets and Systems*, Elsevier Science (2003).

[6] Cao T. H., Uncertain inheritance and recognition as probabilistic default reasoning, *International Journal of Intelligent Systems*, John Wiley & Sons **16** (2001) 781–803.

[7] Cao T. H., Rossiter J. M., Martin T. P., and Baldwin J. F., On the implementing of Fril++ for object-oriented logic programming with uncertainty and fuzziness. In Bouchon-Meunier B. et al. (Eds), *Technologies for Constructing Intelligent Systems, Studies in Fuzziness and Soft Computing*, Vol. 90, Physica-Verlag, 2002, 393–406.

[8] Cao T. H. and Nguyen H., Towards fuzzy and probabilistic object bases, *Proceedings of the 3rd International Conference on Intelligent Technologies and the 3rd Vietnam - Japan Symposium on Fuzzy Systems and Applications*, INTECH/VJFUZZY'2002, Hanoi - Vietnam, Dec. 3 - 5 (2002).

[9] Cross V. V., Fuzzy set theory + object-oriented model = fuzzy object model. http://www.sel.iit.nrc.ca/iitcolloq/Season1998-99/Cross.html, 1998.

[10] Eiter T., Lu J. J., Lukasiewicz T., and Subrahmanian V. S., Probabilistic object bases, *ACM Transactions on Database Systems* **26** (2001) 264–312 (or http://citeseer.nj.nec.com.).

[11] Gaines B. R., Fuzzy and probability uncertainty logics, *Journal of Information and Control* **38** (1978) 154–169.

[12] Na S. and Park S., Fuzzy object-oriented data model and fuzzy association algebra, http://www.wspc.com/books/compsci/3311.html, 2000.

[13] Nguyễn Hòa, "Phát triển một đại số để xử lý cơ sở đối tượng xác suất mờ". Luận văn Thạc sĩ, 2003, Trường Đại học Bách khoa Tp. Hồ Chí Minh.

Nhận bài ngày 2 - 12 - 2003
Nhận lại sau sửa ngày 15 - 8 - 2004