TRƯỜNG ĐẠI HỌC KHOA HỌC HUẾ
KHOA CÔNG NGHỆ THÔNG TIN

PGS.TSKH NGUYỄN CÁT HỒ
TS. NGUYỄN CÔNG HÀO

Giáo trình

LOGIC MỜ VÀ ỨNG DỤNG

(*Dành cho học viên cao học*)

Huế, 2009

Chương 1

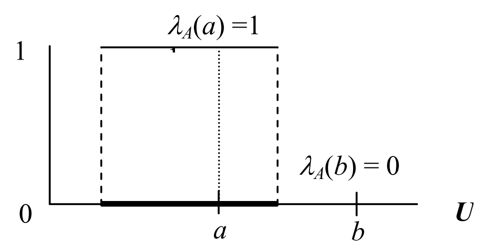
LÝ THUYẾT TẬP MỜ

1.1. Tập mờ và thông tin không chắc chắn

L.A. Zadeh là người sáng lập ra lý thuyết tập mờ với hàng loạt bài báo mở đường cho sự phát triển và ứng dụng của lý thuyết này, khởi đầu là bài báo "Fuzzy Sets" trên Tạp chí Information and Control, 8, 1965. Ý tưởng nổi bật của khái niệm tập mờ của Zadeh là từ những khái niệm trừu tượng về ngữ nghĩa của thông tin mờ, không chắc chắn như *trẻ*, *nhanh*, *cao-thấp*, *xinh đẹp*.., ông đã tìm ra cách biểu diễn nó bằng một khái niệm toán học, được gọi là tập mờ, như là một sự khái quát trực tiếp của khái niệm tập hợp kinh điển.

Để dễ hiểu chúng ta hãy nhớ lại cách nhìn khái niệm tập hợp kinh điển như là khái niệm các hàm số.

Cho một tập vũ trụ $\boldsymbol{U}$. Tập tất cả các tập con của $\boldsymbol{U}$ ký hiệu là $\mathscr{P}(\boldsymbol{U})$ và nó trở thành một đại số tập hợp với các phép tính hợp $\cup$, giao $\cap$, hiệu \ và lấy phần bù $-$, $(\mathscr{P}(\boldsymbol{U}), \cup, \cap, \backslash, -)$. Bây giờ mỗi tập hợp $A \in \mathscr{P}(\boldsymbol{U})$ có thể được xem như là một hàm số $\lambda_A : \boldsymbol{U} \to \{0, 1\}$ được xác định như sau:



$$\lambda_A(x) = \begin{cases} 1 & khi \quad x \in A \\ 0 & khi \quad x \notin A \end{cases}$$

Mặc dù λ_A và A là hai đối tượng toán học hoàn toàn khác nhau, nhưng chúng đều biểu diễn cùng một khái niệm về tập hợp: $x \in A$ khi và chỉ khi $\lambda_A(x) = 1$, hay x thuộc vào tập A với "độ thuộc vào" bằng 1. Vì vậy, hàm λ_A được gọi là hàm đặc trưng của tập A. Như vậy tập hợp A có thể được biểu thị bằng một hàm mà giá trị của nó là *độ thuộc về* hay đơn giản là *độ thuộc* của phần tử trong **U** vào tập hợp A: Nếu $\lambda_A(x) = 1$ thì $x \in A$ với độ thuộc là 1 hay 100% thuộc vào A, còn nếu $\lambda_A(x) = 0$ thì $x \notin A$ hay $x \in A$ với độ thuộc là 0 tức là độ thuộc 0%.

Trên cách nhìn như vậy, chúng ta hãy chuyển sang việc tìm kiếm cách thức biểu diễn ngữ nghĩa của khái niệm mờ, chẳng hạn, về lứa tuổi "*trẻ*". Giả sử tuổi của con người nằm trong khoảng $\boldsymbol{U}$ = [0, 120] tính theo năm. Theo ý tưởng của Zadeh, khái niệm *trẻ* có thể biểu thị bằng một tập hợp như sau: Xét một tập hợp $A_{trẻ}$ những người được xem là *trẻ*. Vậy, một câu hỏi là "Một người *x* có tuổi là *n* được hiểu là thuộc tập $A_{trẻ}$ như thế nào?" Một cách chủ quan, chúng ta có thể hiểu những người có tuổi từ 1 – 25 chắc chắn sẽ thuộc vào tập hợp $A_{trẻ}$, tức là với độ thuộc bằng 1; Nhưng một người có tuổi 30 có lẽ chỉ thuộc vào tập $A_{trẻ}$ với độ thuộc 0,6 còn người có tuổi 50 sẽ thuộc vào tập này với độ thuộc 0,0 … Với ý tưởng đó, ngữ nghĩa của khái niệm *trẻ* sẽ được biểu diễn bằng một hàm số $\mu_{trẻ} : \boldsymbol{U} \rightarrow [0, 1]$, một dạng khái quát trực tiếp từ khái niệm hàm đặc trưng λ_A của một tập hợp kinh điển *A* đã đề cập ở trên.

Một câu hỏi tự nhiên xuất hiện là tại sao người có tuổi 30 có lẽ chỉ thuộc vào tập $A_{trẻ}$ với độ thuộc 0,6 mà không phải là 0,65? Trong lý thuyết tập mờ chúng ta không có ý định trả lời câu hỏi kiểu như vậy mà ghi nhận rằng tập mờ của một khái niệm mờ phụ thuộc mạnh mẽ vào chủ quan của người dùng hay, một cách đúng đắn hơn, của một cộng đồng, hay của một ứng dụng cụ thể. Khía cạch này cũng thể hiện tính không chính xác về ngữ nghĩa của các khái niệm mờ. Tuy nhiên, thực tế này không ảnh hưởng đến khả năng ứng dụng của lý thuyết tập mờ vì mỗi giải pháp dựa trên lý thuyết tập mờ cũng chỉ nhằm vào một miền ứng dụng cụ thể trong đó các khái niệm mờ trong ứng dụng (hay trong cộng đồng sử dụng ứng dụng đó) sẽ có ý nghĩa chung thống nhất.

1.1.1. Khái niệm tập hợp mờ

Định nghĩa 1.1. Cho một tập vũ trụ $\boldsymbol{U}$. Tập hợp $A^{\sim}$ được xác định bởi đẳng thức: $A^{\sim} = \{ \mu_{A^{\sim}}(u)/u : u \in \boldsymbol{U}, \mu_{A^{\sim}}(u) \in [0, 1]\}$ được gọi là một tập hợp mờ trên tập $\boldsymbol{U}$.

Biến *u* lấy giá trị trong $\boldsymbol{U}$ được gọi là *biến cơ sở* và vì vậy tập $\boldsymbol{U}$ còn được gọi là tập tham chiếu hay *miền cơ sở*. Hàm $\mu_{A^{\sim}} : \boldsymbol{U} \rightarrow [0, 1]$ được gọi là hàm thuộc (membership function) và giá trị $\mu_{A^{\sim}}(u)$ tại *u* được gọi là độ thuộc của phần tử u thuộc về tập hợp mờ $A^{\sim}$. Nếu không gây nhầm lẫn, hàm thuộc $\mu_{A^{\sim}}$ cũng được ký hiệu là $A^{\sim}(.)$, nếu biến cơ sở u không biểu thị hiển, hay $A^{\sim}(u)$, nếu biến u xuất hiện hiển.

Lưu ý rằng vế phải của định nghĩa $A^{\sim}$ là một tập kinh điển và do đó định nghĩa trên là hoàn chỉnh.

Họ tất cả các tập mờ trên miền cơ sở $\boldsymbol{U}$ được ký hiệu là $\boldsymbol{F}(\boldsymbol{U})$,

$$\boldsymbol{F}(\boldsymbol{U}) = \{\mu_{A^{\sim}} : \boldsymbol{U} \to [0, 1]\} = [0, 1]^{\boldsymbol{U}}$$

Có nhiều cách biểu diễn hình thức một tập mờ. Trong trường hợp $\boldsymbol{U}$ là một tập hữu hạn, đếm được hay vô hạn liên tục, tập mờ $A^{\sim}$ có thể được biểu diễn bằng các biểu thức hình thức như sau:

Trong trường hợp $\boldsymbol{U}$ hữu hạn, $\boldsymbol{U} = \{u_i : 1 \le i \le n\}$, ta có thể viết:

$A^{\sim} = \mu_{A^{\sim}}(u_1)/u_1 + \mu_{A^{\sim}}(u_2)/u_2 + \ldots + \mu_{A^{\sim}}(u_n)/u_n$ hay $A^{\sim} = \sum_{1 \le i \le n} \mu_{A^{\sim}}(u_i)/u_i$

Trong trường hợp này tập mờ được gọi là tập mờ rời rạc (discrete fuzzy set).

Trong trường hợp $\boldsymbol{U}$ là vô hạn đếm được, $\boldsymbol{U} = \{u_i : i = 1, 2, \ldots\}$, ta có thể viết: $A^{\sim} = \sum_{1 \le i < \infty} \mu_{A^{\sim}}(u_i)/u_i$

Trong trường hợp $\boldsymbol{U}$ là vô hạn liên tục, $\boldsymbol{U} = [a, b]$, ta có thể viết

$$A^{\sim} = \int_a^b \mu_{A^{\sim}}(u)/u$$

Lưu ý rằng các biểu thức trên chỉ có tính hình thức, các phép cộng +, phép tổng Σ và phép lấy tích phân đều không có nghĩa theo quy ước thông thường. Tuy nhiên cách biểu diễn như vậy sẽ rất tiện dụng khi định nghĩa và thao tác các phép tính trên các tập mờ sau này.

Ví dụ 1.1. Xét tập $\boldsymbol{U}$ gồm 5 người là $x_1, x_2, \ldots x_5$ tương ứng có tuổi là 10, 15, 50, 55, 70 và $A^{\sim}$ là tập hợp các người *"Trẻ"*. Khi đó ta có thể xây dựng hàm thuộc như sau:

$\mu_{Trẻ}(10) = 0.95$, $\mu_{Trẻ}(15) = 0.75$, $\mu_{Trẻ}(50) = 0.35$, $\mu_{Trẻ}(55) = 0.30$, $\mu_{Trẻ}(70) = 0.05$ và tập mờ $A^{\sim} = \frac{0.95}{x_1} + \frac{0.75}{x_2} + \frac{0.35}{x_3} + \frac{0.30}{x_4} + \frac{0.05}{x_5}$

Định nghĩa 1.2. Tập mờ $A^{\sim}$ có dạng hình thang xác định bởi bộ 4 giá trị (a, b, c, d), ký hiệu $A^{\sim}$ = (a, b, c, d) và được xác định:

$$\mu_{A^{\sim}}(x) = \begin{cases} 0 & \text{nếu } x \le a \\ \dfrac{x-a}{b-a} & \text{nếu } a < x < b \\ 1 & \text{nếu } b \le x \le c \\ \dfrac{d-x}{d-c} & \text{nếu } c < x < d \\ 0 & \text{nếu } x \ge d \end{cases}$$

1.1.2. Tập lát cắt của tập mờ

Ở trên chúng ta thấy khai niệm tập mờ là một sự khái quát trực tiếp, đẹp đẽ của khái niệm tập kinh điển. Điều này cho phép hy vọng nó sẽ đặt cơ sở cho mối liên hệ chặt chẽ giữa hai khái niệm tập hợp này. Để dẫn đến việc nghiên cứu đó, trước hết chúng ta đưa ra khái niệm *tập lát cắt* α của một tập mờ.

Định nghĩa 1.3. Cho một tập mờ $A^{\sim}$ trên tập vũ trụ $\boldsymbol{U}$ và $\alpha \in [0, 1]$. *Tập lát cắt* α (hoặc α+) của tập $A^{\sim}$ là một tập kinh điển, ký hiệu là $A^{\sim}_{\alpha}$ (hoặc $A^{\sim}_{\alpha+}$), được xác định bằng đẳng thức sau:

$$A^{\sim}_{\alpha} = \{u \in \boldsymbol{U} : \mu_{A^{\sim}}(u) \ge \alpha\} \text{ (hoặc } A^{\sim}_{\alpha+} = \{u \in \boldsymbol{U} : \mu_{A^{\sim}}(u) > \alpha\}).$$

Như vậy, mỗi tập mờ $A^{\sim}$ sẽ cảm sinh một họ các tập kinh điển, ta có ánh xạ $\quad \boldsymbol{h} : A^{\sim} \in \boldsymbol{F}(\boldsymbol{U}) \quad \rightarrow \{A^{\sim}_{\alpha} \in \mathcal{P}(\boldsymbol{U}) : 0 \le \alpha \le 1\} \quad$ (1*)

Để đơn giản ký hiệu, ta viết họ các tập kinh điển như vậy bằng $\boldsymbol{h}(A^{\sim}) = \{A^{\sim}_{\alpha} : 0 \le \alpha \le 1\}$, $A^{\sim} \in \boldsymbol{F}(\boldsymbol{U})$. Họ các tập hợp như vậy có các tính chất sau:

Định lý 1.1. Cho $A^{\sim}, B^{\sim} \in \boldsymbol{F}(\boldsymbol{U})$, $\boldsymbol{h}$ là ánh xạ được cho trong (1*) và $\boldsymbol{h}(A^{\sim}) = \{A^{\sim}_{\alpha} : 0 \le \alpha \le 1\}$, $\boldsymbol{h}(B^{\sim}) = \{B^{\sim}_{\alpha} : 0 \le \alpha \le 1\}$. Khi đó,

(i) Mỗi họ $\boldsymbol{h}(A^{\sim})$ như vậy là dãy đơn điệu giảm, nếu $\alpha < \beta$, thì $A^{\sim}_{\alpha} \supseteq A^{\sim}_{\beta}$;

(ii) Nếu $A^{\sim} \neq B^{\sim}$ thì $\{A^{\sim}_{\alpha} : 0 \leq \alpha \leq 1\} \neq \{B^{\sim}_{\alpha} : 0 \leq \alpha \leq 1\}$.

Nghĩa là tồn tại một song ánh từ họ các tập mờ $\boldsymbol{F}(\boldsymbol{U})$ vào họ của những họ tập kinh điển $\mathscr{P}(\boldsymbol{U})$ ở dạng (1*).

Chứng minh: Tính chất (i) dễ dàng rút ra từ tính chất ($A^{\sim}(u) \geq \beta \Rightarrow A^{\sim}(u) \geq \alpha$).

Để chứng minh (ii), giả sử $A^{\sim} \neq B^{\sim}$, $\exists u \in \boldsymbol{U}(A^{\sim}(u) \neq B^{\sim}(u))$. Để định ý, ta giả sử rằng có $u_0 \in \boldsymbol{U}$ sao cho $A^{\sim}(u_0) > B^{\sim}(u_0)$. Chọn $\alpha \in [0, 1]$ sao cho $A^{\sim}(u_0) > \alpha > B^{\sim}(u_0)$. Điều này khẳng định $u_0 \in A^{\sim}_{\alpha}$ nhưng $u_0 \notin B^{\sim}_{\alpha}$ hay $A^{\sim}_{\alpha} \neq B^{\sim}_{\alpha}$. Vậy, $\{A^{\sim}_{\alpha} : 0 \leq \alpha \leq 1\} \neq \{B^{\sim}_{\alpha} : 0 \leq \alpha \leq 1\}$.

Hiển nhiên là nếu $A^{\sim} = B^{\sim}$ thì $\{A^{\sim}_{\alpha} : 0 \leq \alpha \leq 1\} = \{B^{\sim}_{\alpha} : 0 \leq \alpha \leq 1\}$. Như vậy ta đã chứng tỏ rằng ánh xạ $\boldsymbol{h}$ là song ánh.

1.1.3. Một số khái niệm đặc trưng của tập mờ

Định nghĩa 1.4. (i) *Giá của tập mờ*: Giá của tập mờ $A^{\sim}$, ký hiệu là *Support*($A^{\sim}$), là tập con của $\boldsymbol{U}$ trên đó $\mu_{A^{\sim}}(u) \neq 0$, *Support*($A^{\sim}$) = $\{u: \mu_{A^{\sim}}(u) > 0\}$.

(ii) *Độ cao của tập mờ*: Độ cao của tập mờ $A^{\sim}$, ký hiệu là *hight*($A^{\sim}$), là cận trên đúng của hàm thuộc $\mu_{A^{\sim}}$ trên $\boldsymbol{U}$, *hight*($A^{\sim}$) = $\sup\{\mu_{A^{\sim}}(u) : u \in \boldsymbol{U}\}$.

(iii) *Tập mờ chuẩn* (normal): Tập mờ $A^{\sim}$ được gọi là *chuẩn* nếu *hight*($A^{\sim}$) = 1. Trái lại, tập mờ được gọi là *dưới chuẩn* (subnormal).

(iv) *Lõi của tập mờ*: Lõi của tập mờ $A^{\sim}$, ký hiệu là *Core*($A^{\sim}$), là một tập con của $\boldsymbol{U}$ được xác định như sau:

$$Core(A^{\sim}) = \{u \in \boldsymbol{U}: \ \mu_{A^{\sim}}(u) = hight(A^{\sim})\}.$$

Bây giờ chúng ta sẽ lấy một số ví dụ về việc biểu diễn ngữ nghĩa của các khái niệm mờ thuộc các lĩnh vực khác nhau bằng tập mờ.

Ví dụ 1.2. Giả sử $\boldsymbol{U}$ là tập vũ trụ về số đo nhiệt độ thời tiết, chẳng hạn $\boldsymbol{U}$ = [0, 50] tính theo thang độ C. Chúng ta sẽ xác định tập mờ biểu thị khái niệm mờ thời tiết *NÓNG* và *LẠNH*. Trong ví dụ này ta sử dụng một hàm số mẫu, gọi là *S*-hàm vì đồ thị của nó có hình chữ *S*. Chúng ta ký hiệu hàm này là *S*(*u*, *a*, *b*,

c), trong đó a, b và c là những tham số. Nó là hàm từng khúc bậc 2 và được định nghĩa như sau:

$$
\begin{aligned}
S(u, a, b, c) &= 0 && \text{đối với } u \le a \\
&= 2\left(\frac{u-a}{c-a}\right)^2 && \text{đối với } a \le u \le b \\
&= 1 - 2\left(\frac{u-c}{c-a}\right)^2 && \text{đối với } b \le u \le c \\
&= 1 && \text{đối với } c \le u
\end{aligned}
$$

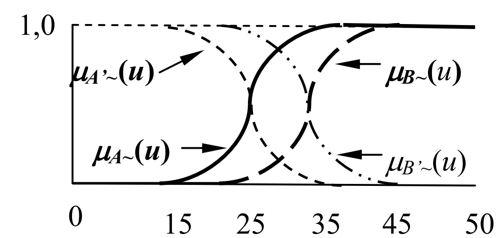
Hàm thuộc $\mu_{A\sim}(u) = S(u, 15, 25, 35)$ là khái niệm thời tiết *NÓNG* của người Lạng Sơn ở cực Bắc nước ta, còn hàm thuộc $\mu_{B\sim}(u) = S(u, 25, 35, 45)$ là khái niệm *NÓNG* của người Sài Gòn (xem Hình 1.1).

Với hai tập mờ này ta có: $Support(A^{\sim}) = [15, 50]$, $Support(B^{\sim}) = [25, 50]$, $Hight(A^{\sim}) = Hight(B^{\sim}) = 1$, $Core(A^{\sim}) = [35, 50]$ và $Core(B^{\sim}) = [45, 50]$.

Hàm thuộc biểu thị khái niệm mờ *LẠNH* được xác định qua hàm thuộc *NÓNG* bằng biểu thức sau:

$$\mu_{A'\sim}(u) = 1 - \mu_{A\sim}(u) \quad \text{và} \quad \mu_{B'\sim}(u) = 1 - \mu_{B\sim}(u)$$

Ví dụ này thể hiện tính chủ quan về ngữ nghĩa của khai niệm mờ và do đó thể hiện tính tự do trong việc xây dựng các hàm thuộc. Tình huống tương tự như vậy khi ta nói đến khái niệm cao của giới nữ và giới nam, hay khái niệm cao của người Việt Nam và người Châu Âu.

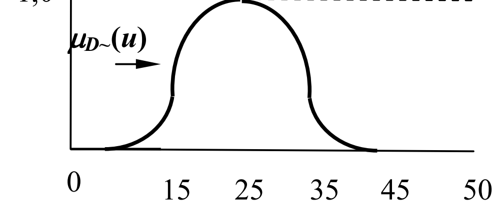


Hình 1.1: *Hàm thuộc của tập mờ NÓNG và LẠNH*

Ví dụ 1.3. *Tập mờ hình chuông*: Người ta có thể biểu diễn ngữ nghĩa của khái niệm mờ trời *mát mẻ* hay *dễ chịu* bằng hàm dạng hình chuông như sau:

$$\exp\left(-\left((u - u_0)/b\right)^2\right)$$

Chúng ta có thể chấp nhận *hàm chuông* trong Hình 1.2 là biểu thị ngữ nghĩa của khái niệm nhiệt độ *DỄ CHỊU* và khi đó tập mờ $D^{\sim}$ có dạng: $\mu_{D\sim}(u) = \exp\left(-\left((u - 24)/10\right)^2\right)$



Hình 1.2: *Hàm thuộc của tập mờ DỄ CHỊU*

Ví dụ 1.4. Ta sẽ đưa ra một ví dụ về tập mờ rời rạc (discrete fuzzy set). Xét $\boldsymbol{U}$ là tập các giá trị trong thang điểm 10 đánh giá kết quả học tập của học sinh về môn Toán, $\boldsymbol{U} = \{1, 2, \ldots, 10\}$. Khi đó khái niệm mờ về năng lực học môn toán giỏi có thể được biểu thị bằng tập mờ $G^{\sim}$ sau:

$$G^{\sim} = 0{,}1/4 + 0{,}2/5 + 0{,}4/6 + 0{,}7/7 + 0{,}9/8 + 1{,}0/9 + 1{,}0/10 \qquad (2^*)$$

ở đây các giá trị của miền $\boldsymbol{U}$ không có mặt trong biểu thức (2*) có nghĩa độ thuộc của chúng vào tập mờ $G^{\sim}$ là bằng 0,0.

Trong trường hợp tập mờ rời rạc ta có thể biểu diễn tập mờ bằng một bảng. Chẳng hạn, đối với tập mờ $G^{\sim}$ ở trên ta có bảng như sau:

Bảng 1.1: *Tập mờ* $G^{\sim}$

$\boldsymbol{U}$	1	2	3	4	5	6	7	8	9	10
$G^{\sim}$	0,0	0,0	0,0	0,1	0,3	0,5	0,7	0,9	1,0	1,0

Ví dụ 1.5. Trong ví dụ này chúng ta sẽ xây dựng tập mờ biểu thị ngữ nghĩa của khái niệm *GIÀ* và *TRẺ* của thuộc tính lứa tuổi.

Giả sử tập vũ trụ chỉ tuổi tính theo đơn vị năm là $\boldsymbol{U} = \{u : 0 \le u \le 120\}$, chẳng hạn tuổi của x là 8,37 năm. Khi đó khái niệm *GIÀ* có thể được biểu thị bằng tập mờ với hàm thuộc như sau:

$$\mu_{GI\grave{A}}(u) = \int_0^{120} \{1 + \left(\frac{u-60}{6}\right)^{-2}\}^{-1} / u$$

$$\mu_{TR\mathring{E}}(u) = 1 - \mu_{GI\grave{A}}(u) = \int_0^{120} \{1 - \{1 + \left(\frac{u-60}{6}\right)^{-2}\}^{-1}\} / u$$

Cần nhấn mạnh một lần nữa rằng đây là công thức hình thức biểu diễn các tập mờ. Dấu tích phân chỉ có nghĩa miền xác định $\boldsymbol{U}$ của hàm thuộc là vô hạn continuum, tập hợp có lực lượng tương đương với đoạn [0, 1].

Ví dụ 1.6. Tập rời rạc trên miền phi số: Trong thực tế ứng dụng người ta cũng hay sử dụng tập mờ trên miền phi số, chẳng hạn, miền giá trị ngôn ngữ. Ví dụ, ta xét biến ngôn ngữ *NHIỆT ĐỘ* có thể xem như xác định trên miền 3 giá trị ngôn ngữ $\boldsymbol{U}$ = {*Thấp*, *Trung-bình*, *Cao*}. Khi đó, một tập mờ rời rạc $T^{\sim}$ trên miền $\boldsymbol{U}$ có thể được biểu thị như sau:

$$T^{\sim} = \mu_1/\textit{Thấp} + \mu_2/\textit{Trung-bình} + \mu_3/\textit{Cao}$$

Chẳng hạn *Trời-mát* có thể biểu thị bằng tập mờ như sau:

$$Trời\text{-}mát = 0{,}7/Thấp + 0{,}8/Trung\text{-}bình + 0{,}2/Cao$$

Đối với tập hợp kinh điển *A* chúng ta có khái niệm số lượng các phần tử của một tập hợp, trong trường hợp *A* là hữu hạn, hay lực lượng của tập hợp, trong trường hợp *A* là vô hạn. Hai tập hợp *A* và *B* có lực lượng bằng nhau nếu có tồn tại một ánh xạ 1-1 từ *A lên B*.

Đối với tập mờ $A^{\sim}$, khái niệm lực lượng được khái quát hóa bằng định nghĩa sau:

Định nghĩa 1.5. **Lực lượng của tập mờ**

Cho $A^{\sim}$ là một tập mờ trên ***U***

(i) *Lực lượng vô hướng* (scalar cardinality): Lực lượng hay bản số thực của tập $A^{\sim}$, ký hiệu là $Count(A^{\sim})$, được tính theo công thức đếm sau (đôi khi được gọi là sigma count).

$$Count(A^{\sim}) = \sum\nolimits_{u\in U}^{arith} \mu_{A^{\sim}}(u)$$, nếu ***U*** là tập hữu hạn hay đếm được

$$= \int_{U}^{arith} \mu_{A^{\sim}}(u)du$$, nếu ***U*** là tập vô hạn continuum

ở đây $\sum^{arith}$ và $\int^{arith}$ là tổng và tích phân số học.

(ii) *Lực lượng mờ* (fuzzy cardinality): Lực lượng hay bản số mờ của tập *$A^{\sim}$ là một tập mờ* trên tập các số nguyên không âm ***N*** được định nghĩa như sau:

$$Card(A^{\sim}) = \int_{N} \mu_{Card(A^{\sim})}(n)dn$$

trong đó $\mu_{Card(A^{\sim})}(n)$ được xác định theo công thức sau, với $|A_t^{\sim}|$ là lực lượng của tập mức $A_t^{\sim}$, $\mu_{Card(A^{\sim})}(n) = \text{supremum } \{t \in [0, 1]: |A_t^{\sim}| = n\}$.

Có thể xem công thức tính $Count(A^{\sim})$ ở trên như là công thức "*đếm*" số phần tử trong ***U***. Thực vậy, nếu tập $A^{\sim}$ trở về tập kinh điển thì $\mu_{A\sim}(u) \equiv 1$ trên ***U*** và do đó công thức $Count(A^{\sim})$ trên chính là bộ đếm số phần tử. Khi $\mu_{A\sim}(u) \neq 1$, thì *u* chỉ thuộc về tập $A^{\sim}$ với tỷ lệ phần trăm bằng $\mu_{A\sim}(u)$ và do đó phần tử *u* chỉ được "*đếm*" vào số lượng các phần tử một đại lượng bằng $\mu_{A\sim}(u)$.

Lưu ý rằng, khác với trường hợp tập kinh điển, dù tập ***U*** là vô hạn đếm được hay vô hạn continuum, thì lực lượng của tập mờ $A^{\sim}$ vẫn có thể là hữu hạn, tùy theo dáng điệu của hàm $\mu_{A\sim}(u)$.

1.2. Biến ngôn ngữ

L.A.Zadeh viết *"khi thiếu hụt tính chính xác bề ngoài của những vấn đề phức tạp, một cách tự nhiên là tìm cách sử dụng các biến ngôn ngữ, đó là các biến mà giá trị của chúng không phải là số mà là các từ hoặc các câu trong ngôn ngữ tự nhiên hoặc nhân tạo. Động lực cho việc sử dụng các từ, các câu hơn các số là đặc trưng ngôn ngữ của các từ, các câu thường là ít xác định hơn của số"*.

Trong cơ sở dữ liệu quan hệ, các quan hệ hay các bảng dữ liệu chứa các thuộc tính hay các tên cột. Nó chỉ tính chất của đối tượng. Các thuộc tính này cũng thể hiện trong ngôn ngữ như để mô tả tính chất đối tượng là con người, trong ngôn ngữ tự nhiên chúng ta có những thuộc tính *TUỔI*, *CHIỀU CAO*, *LƯƠNG*, *NĂNG LỰC* … . Các thuộc tính này có thể được mô tả bằng giá trị ngôn ngữ như *trẻ*, *già*, *rất trẻ*, … Vì lý do như vậy, Zadeh gọi các thuộc tính kiểu như vậy là *biến ngôn ngữ* và miền giá trị của chúng là giá trị ngôn ngữ hay gọi là miền ngôn ngữ (linguistic domain hay term-domain). Tuy nhiên, như chúng ta đã đề cập trong Mục 1.1, vì bản thân giá trị ngôn ngữ không phải là đối tượng toán học, ngữ nghĩa của chúng được biểu thị bằng các tập mờ hay hàm thuộc. Để khái niệm biến ngôn ngữ trở thành một khái niệm toán học, Zadeh hình thức hóa khái niệm này như sau:

Định nghĩa 1.6. Biến ngôn ngữ là một bộ năm (X, T(X), U, R, M), trong đó X là tên biến, T(X) là tập các giá trị ngôn ngữ của biến X, U là không gian tham chiếu của biến cơ sở u, mỗi giá trị ngôn ngữ xem như là một biến mờ trên U kết hợp với biến cơ sở u, R là một qui tắc cú pháp sinh các giá trị ngôn ngữ của T(X), M là qui tắc ngữ nghĩa gán mỗi giá trị ngôn ngữ trong T(X) với một tập mờ trên U.

Ví dụ 1.7. Cho X là biến ngôn ngữ có tên là *AGE*, biến cơ sở u lấy theo số tuổi của con người có miền xác định là U = [0,100]. Tập các giá trị ngôn ngữ

T(AGE) = {*old*, *very old*, *more or less young*, *less young*, *very young*….}. R là một qui tắc sinh các giá trị này. M gán ngữ nghĩa mỗi tập mờ với một giá trị ngôn ngữ. Chẳng hạn, đối với giá trị nguyên thủy *old*, M (*old*) = $\{(u, \mu_{old}(u) \mid u \in [0,100]\}$, ở đây chọn

$$\mu_{old}(u) = \begin{cases} 0 & u \in [0,50] \\ (1+(\frac{u-50}{5})^{-2})^{-1} & u \in [50,100] \end{cases}$$

Các đặc trưng của biến ngôn ngữ

Trong thực tế có rất nhiều biến ngôn ngữ khác nhau về các giá trị nguyên thuỷ, chẳng hạn như biến ngôn ngữ *SỐ NGÀY LÀM VIỆC* có giá trị nguyên thuỷ là *ít*, *nhiều*, biến ngôn ngữ *LƯƠNG* có giá trị nguyên thuỷ là *thấp*, *cao*…..Tuy nhiên, những kết quả nghiên cứu đối với một miền trị của một biến ngôn ngữ cụ thể vẫn giữ được ý nghĩa về mặt cấu trúc đối với miền giá trị của các biến còn lại. Đặc trưng này được gọi là *tính phổ quát* của biến ngôn ngữ.

Ngữ nghĩa của các gia tử và các liên từ hoàn toàn độc lập với ngữ cảnh, điều này khác với giá trị nguyên thủy của các biến ngôn ngữ lại phụ thuộc vào ngữ cảnh. Ví dụ ta nói *LƯƠNG* của cán bộ An là *rất cao*, khi đó được hiểu rằng *LƯƠNG* khoảng trên 8.000.000 đồng, nhưng ta nói *CHIỀU CAO* của cán bộ An là *rất cao* thì được hiểu rằng *CHIỀU CAO* khoảng trên 1.8 m. Do đó khi tìm kiếm mô hình cho các gia tử và các liên từ chúng ta không quan tâm đến giá trị nguyên thuỷ của biến ngôn ngữ đang xét. Đặc trưng này được gọi là *tính độc lập ngữ cảnh của gia tử và liên từ*.

Các đặc trưng trên cho phép chúng ta sử dụng cùng một tập các gia tử và xây dựng một cấu trúc toán học duy nhất cho miền giá trị của các biến ngôn ngữ khác nhau.

1.3. Các phép tính trên trên tập mờ

Xét một biến ngôn ngữ X như đã được định nghĩa ở trên. Trước hết, chúng ta có nhận xét rằng, nhìn chung, tập ảnh của tập T(X) qua ánh xạ M(X) không có cấu trúc đại số, trên đó chúng ta không định nghĩa được các phép tính trên tập mờ. Một lý do nữa làm cho chúng ta không quan tâm đến điều này là cấu trúc đại số của tập gốc T(X) cũng chưa được phát hiện. Trong khi chúng ta chưa phát hiện ra cấu trúc đại số của miền T(X), trong mục này chúng ta sẽ định nghĩa trên tập $\boldsymbol{F}(\boldsymbol{U}, [0, 1])$ một cấu trúc đại số.

Cũng cần nhấn mạnh rằng mục tiêu của lý thuyết tập mờ là mô hình hóa toán học ngữ nghĩa của các khái niệm mờ và, quan trọng nhất, là mô hình hóa phương pháp lập luận của con người. Đây là một vấn đề cực kỳ khó và phức tạp vì những vấn đề này thuộc loại có cấu trúc yếu, hay khó có thể có một cấu trúc toán duy nhất mô hình hóa trọn vẹn những vấn đề nêu trên. Như là một hệ quả, khó lòng chúng ta tìm được một cấu trúc toán học chặt chẽ, đẹp của tập $\boldsymbol{F}(\boldsymbol{U}, [0, 1])$. Chính vì vậy chúng ta không có một ràng buộc chặt chẽ, minh bạch trong định nghĩa các phép toán trong $\mathrm{F}(\boldsymbol{U}, [0, 1])$. Như chúng ta sẽ thấy dưới đây, chúng ta có nhiều cách khác nhau để định nghĩa các phép tính và do đó nó tạo ra tính mềm dẻo, đa dạng trong tiếp cận, thích nghi với các bài toán ứng dụng khác nhau, miễn là nó cho phép giải quyết được các bài toán ứng dụng, đặc biệt các bài toán thuộc lĩnh vực trí tuệ nhân tạo.

Trước khi định nghĩa các phép tính trong $\boldsymbol{F}(\boldsymbol{U}, [0, 1])$, chúng ta hãy xem đoạn [0, 1] như là một cấu trúc dàn $\boldsymbol{L}_{[0,1]} = ([0, 1], \cup, \cap, -)$ với thứ tự tự nhiên trên đoạn [0, 1]. Khi đó, với mọi $a, b \in [0, 1]$, ta có:

$$a \cup b = \max\{a, b\}, \quad a \cap b = \min\{a, b\} \quad \text{và} \quad -a = 1 - b.$$

Chúng ta có thể kiểm chứng rằng $\boldsymbol{L}_{[0,1]} = ([0, 1], \cup, \cap, -)$ là một đại số De Morgan, hơn nữa nó có các tính chất sau:

- Các phép tính hợp $\cup$ và giao $\cap$ có tính giao hoán

$$a \cup b = b \cup a \quad \text{và} \quad a \cap b = b \cap a$$

- Các phép tính hợp $\cup$ và giao $\cap$ có tính chất phân phối lẫn nhau

$$a \cup (b \cap c) = (a \cup b) \cap (a \cup c) \quad \text{và} \quad a \cap (b \cup c) = (a \cap b) \cup (a \cap c)$$

- Tính chất nuốt (absorption) và nuốt đối ngẫu (dual absorption):
- Tính chất nuốt : $a \cap (a \cup b) = a$,
- Tính chất nuốt đối ngẫu : $a \cup (a \cap b) = a$.
- Tính lũy đẳng : $a \cup a = a$ và $a \cap a = a$
- Tính chất phủ phủ định : $-(-a) = a$
- Tính đơn điệu giảm : $a \leq b \Rightarrow -a \geq -b$

- Tính chất De Morgan : $-(a \cup b) = -a \cap -b$; $-(a \cap b) = -a \cup -b$.

Dựa trên cấu trúc $\boldsymbol{L}_{[0,1]}$ chúng ta sẽ định nghĩa các phép tính trên tập mờ thông qua các phép tính của dàn $\boldsymbol{L}_{[0,1]}$.

1.3.1. Phép hợp $\tilde{\cup}$

Cho hai tập mờ $A^{\sim}$ và $B^{\sim}$ trên tập vũ trụ $\boldsymbol{U}$. Hợp của hai tập mờ này là một tập mờ ký hiệu là $A^{\sim} \tilde{\cup} B^{\sim}$, mà hàm thuộc của nó được định nghĩa theo điểm (pointwise) như sau: $\mu_{A^{\sim} \tilde{\cup} B^{\sim}}(u) = \mu_{A^{\sim}}(u) \cup \mu_{B^{\sim}}(u)$

hay, trong trường hợp $\boldsymbol{U}$ là hữu hạn hay đếm được,

$$A^{\sim} \tilde{\cup} B^{\sim} = \sum_{1 \le i < \infty} \mu_{A^{\sim}}(u_i)/u_i \ \tilde{\cup} \ \sum_{1 \le i < \infty} \mu_{B^{\sim}}(u_i)/u_i = \sum_{1 \le i < \infty} [\mu_{A^{\sim}}(u_i) \cup \mu_{B^{\sim}}(u_i)]/u_i$$

hay, trong trường hợp $\boldsymbol{U}$ là tập continuum,

$$A^{\sim} \tilde{\cup} B^{\sim} = \int_{u \in U} \mu_{A^{\sim}}(u)du \ \tilde{\cup} \ \int_{u \in U} \mu_{B^{\sim}}(u)du = \int_{u \in U} [\mu_{A^{\sim}}(u) \cup \mu_{B^{\sim}}(u)]du\,.$$

Một cách tổng quát, cho $A_i^{\sim} \in \boldsymbol{F}(\boldsymbol{U})$, $i \in \boldsymbol{I}$, với $\boldsymbol{I}$ là tập chỉ số hữu hạn hay vô hạn nào đó. Khi đó, hợp của các tập mờ như vậy, ký hiệu là $\bigcup_{i \in I} A_i^{\sim}$, được định nghĩa bằng hàm thuộc như sau

$$\left(\bigcup_{i \in I} A_i^{\sim}\right)(u) = \mathrm{Sup}_{i \in \boldsymbol{I}}\ A_i^{\sim}(u) \qquad (3^*)$$

Chúng ta sẽ cho một số ví dụ về phép tính này.

Xét tập vũ trụ $\boldsymbol{U}$ như trong Ví dụ 1.3 và hai tập mờ $G^{\sim}$ và $K^{\sim}$ được cho như trong bảng dưới đây.

Bảng 1.2: *Tập mờ trên* $\boldsymbol{U}$

$\boldsymbol{U}$	1	2	3	4	5	6	7	8	9	10
$G^{\sim}$	0,0	0,0	0,0	0,1	0,3	0,5	0,7	0,9	1,0	1,0
$K^{\sim}$	1,0	0.9	0,8	0,6	0,4	0,2	0,0	0,0	0,0	0,0

Khi sử dụng cách biểu diễn tập mờ rời rạc, hợp của hai tập mờ $G^{\sim}$ và $K^{\sim}$ được thực hiện như sau:

$G^{\sim} \tilde{\cup} K^{\sim}$ = (0,0/1 + 0,0/2 + 0,0/3 + 0,1/4 + 0,3/5 + 0,5/6 + 0,7/7 + 0,9/8 +1,0/9 + 1,0/10)

$\tilde{\cup}$ (1,0/1 + 0,9/2 + 0,8/3 + 0,6/4 + 0,4/5 + 0,2/6 + 0,0/7 + 0,0/8 + 0,0/9 + 0,0/10)

= 1,0/1 + 0,9/2 + 0,8/3 + 0,6/4 + 0,4/5 + 0,5/6 + 0,7/7 + 0,9/8 + 1,0/9 + 1,0/10

Cách thực hiện phép tinh trong dàn $\boldsymbol{L}_{[0,1]}$ theo điểm như vậy gợi ý cho chúng ta thực hiện các phép tính như vậy ngay trên Bảng 1.3 như sau:

Bảng 1.3: *Hợp hai tập mờ trên **U***

U	1	2	3	4	5	6	7	8	9	10
$G^{\sim}$	0,0	0,0	0,0	0,1	0,3	0,5	0,7	0,9	1,0	1,0
$K^{\sim}$	1,0	0.9	0,8	0,6	0,4	0,2	0,0	0,0	0,0	0,0
$G^{\sim} \tilde{\cup} K^{\sim}$	1,0	0,9	0,8	0,6	0,4	0,5	0,7	0,9	1,0	1,0

Một cách tổng quát, nếu cho trước các tập mờ $A_i^{\sim}$, i = 1, …, m, thì hợp của các tập mờ này là tập mờ $A^{\sim}$ được định nghĩa mở rộng bằng quy nạp và được ký hiệu là

$$A^{\sim} = \tilde{\cup}_{i=1}^{n} A_i^{\sim}$$

Nhận xét 1.1: Các hạng thức dạng $\mu(u_i)/u_i$ có thể xem là một tập mờ mà giá của nó chỉ chứa duy nhất một phần tử u_i, hàm thuộc của nó bằng 0 tại mọi $u \neq u_i$ và bằng $\mu(u_i)$ tại phần tử u_i. Kí hiệu tập mờ này là $\mu(u_i)\{u_i\}$, tích của số vô hướng của $\mu(u_i)$ với tập kinh điển 1-phần tử $\{u_i\}$. Khi đó, với định nghĩa phép hợp như trên, các phép cộng hình thức "+" có thể được biểu thị bằng phép hợp, ta có, chẳng hạn với $\boldsymbol{U}$ là tập hữu hạn, $\boldsymbol{U} = \{u_1, \ldots, u_n\}$, tập mờ $A^{\sim}$ được biểu diễn qua phép hợp như sau:

$$A^{\sim} = \tilde{\cup}_{i=1}^{n} \mu(u_i)\{u_i\}$$

Tập $G^{\sim} \tilde{\cup} K^{\sim}$ thu được có những đặc điểm sau:

Support($G^{\sim} \tilde{\cup} K^{\sim}$) = $\boldsymbol{U}$

Nó là tập mờ chuẩn vì *Hight*($G^{\sim} \tilde{\cup} K^{\sim}$) = 1

$Core(G^{\sim} \tilde{\cup} K^{\sim}) = \{1, 9, 10\}$

$Count(G^{\sim} \tilde{\cup} K^{\sim}) = 1{,}0 + 0{,}9 + 0{,}8 + 0{,}6 + 0{,}4 + 0{,}5 + 0{,}7 + 0{,}9 + 1{,}0 + 1{,}0 = 7{,}8$.

1.3.2. Phép giao $\tilde{\cap}$

Cho hai tập mờ $A^{\sim}$ và $B^{\sim}$ trên tập vũ trụ ***U***. Hợp của hai tập mờ này là một tập mờ ký hiệu là $A^{\sim} \tilde{\cap} B^{\sim}$, mà hàm thuộc của nó được định nghĩa theo điểm (pointwise) như sau:

$$\mu_{A^{\sim} \tilde{\cap} B^{\sim}}(u) = \mu_{A^{\sim}}(u) \cap \mu_{B^{\sim}}(u)$$

hay, trong trường hợp ***U*** là hữu hạn hay đếm được,

$$A^{\sim} \tilde{\cap} B^{\sim} = \sum_{1 \le i < \infty} \mu_{A^{\sim}}(u_i)/u_i \ \tilde{\cap} \ \sum_{1 \le i < \infty} \mu_{B^{\sim}}(u_i)/u_i = \sum_{1 \le i < \infty} [\mu_{A^{\sim}}(u_i) \cap \mu_{B^{\sim}}(u_i)]/u_i$$

hay, trong trường hợp ***U*** là tập continuum,

$$A^{\sim} \tilde{\cap} B^{\sim} = \int_{u \in U} \mu_{A^{\sim}}(u) du \ \tilde{\cap} \ \int_{u \in U} \mu_{B^{\sim}}(u) du = \int_{u \in U} [\mu_{A^{\sim}}(u) \cap \mu_{B^{\sim}}(u)] du \, .$$

Một cách tổng quát, cho $A_i^{\sim} \in \boldsymbol{F}(\boldsymbol{U})$, $i \in \boldsymbol{I}$, với ***I*** là tập chỉ số hữu hạn hay vô hạn nào đó. Khi đó, hợp của các tập mờ như vậy, ký hiệu là $\bigcap_{i \in I} A_i^{\sim}$, được định nghĩa bằng hàm thuộc như sau

$$\left(\bigcap_{i \in I} A_i^{\sim}\right)(u) = \mathrm{Inf}_{i \in \boldsymbol{I}} \ A_i^{\sim}(u)$$

Chúng ta sẽ cho một số ví dụ về phép tính này.

Xét hai tập mờ $G^{\sim}$ và $K^{\sim}$ được cho trong Bảng 1.2. Khi sử dụng cách biểu diễn tập mờ rời rạc, giao của hai tập mờ $G^{\sim}$ và $K^{\sim}$ được thực hiện như sau:

$G^{\sim} \tilde{\cap} K^{\sim}$ = (0,0/1 + 0,0/2 + 0,0/3 + 0,1/4 + 0,3/5 + 0,5/6 + 0,7/7 + 0,9/8 +1,0/9 + 1,0/10)

$\tilde{\cap}$ (1,0/1 + 0,9/2 + 0,8/3 + 0,6/4 + 0,4/5 + 0,2/6 + 0,0/7 + 0,0/8 + 0,0/9 + 0,0/10)

= 0,0/1 + 0,0/2 + 0,0/3 + 0,1/4 + 0,3/5 + 0,2/6 + 0,0/7 + 0,0/8 + 0,0/9 + 0,0/10

Cách thực hiện phép tính trong dàn $\boldsymbol{L}_{[0,1]}$ theo từng điểm như vậy, tương tự như trên, chúng ta thực hiện các phép tính như vậy ngay trên Bảng 1.4 dưới đây:

Bảng 1.4: *Giao của hai tập mờ trên* ***U***

$\boldsymbol{U}$	1	2	3	4	5	6	7	8	9	10
$G^{\sim}$	0,0	0,0	0,0	0,1	0,3	0,5	0,7	0,9	1,0	1,0
$K^{\sim}$	1,0	0.9	0,8	0,6	0,4	0,2	0,0	0,0	0,0	0,0
$G^{\sim}\tilde{\cap}K^{\sim}$	0,0	0,0	0,0	0,1	0,3	0,2	0,0	0,0	0,0	0,0

Tập $G^{\sim}\tilde{\cap}K^{\sim}$ thu được có những đặc điểm sau:

Support$(G^{\sim}\tilde{\cap}K^{\sim}) = \boldsymbol{U}$

Nó là tập mờ dưới chuẩn vì *Hight*$(G^{\sim}\tilde{\cap}K^{\sim}) = 0,3 < 1$

Core$(G^{\sim}\tilde{\cap}K^{\sim}) = \{5\}$, tập một phần tử

Count$(G^{\sim}\tilde{\cap}K^{\sim}) = 0,1 + 0,3 + 0,2 = 0,6$

1.3.3. Phép lấy phần bù ~

Xét một tập mờ $A^{\sim}$ trên tập vũ trụ $\boldsymbol{U}$. Phép lấy bù của tập $A^{\sim}$, ký hiệu là $\sim A^{\sim}$, là tập mờ với hàm thuộc được xác định bằng đẳng thức sau:

$$\mu_{\sim A^{\sim}}(u) = 1 - \mu_{A^{\sim}}(u)$$

Tập mờ $\sim A^{\sim}$ biểu diễn ở dạng công thức hình thức có dạng sau:

Trường hợp ***U*** *là hữu hạn hay vô hạn đếm được*

$$\sim A^{\sim} = \sim \sum_{u\in U} \mu_{A^{\sim}}(u)/u = \sum_{u\in U} (1-\mu_{A^{\sim}}(u))/u =$$

Trường hợp ***U*** *là vô hạn continuum*

$$\sim A^{\sim} = \int_{u\in U} \mu_{\sim A^{\sim}}(u)du = \sim\int_{u\in U} \mu_{A^{\sim}}(u)du = \int_{u\in U} (1-\mu_{A^{\sim}}(u))du$$

Để lấy ví dụ. chúng ta xét hai tập mờ $G^{\sim}$ và $K^{\sim}$ được cho trong Bảng 1.2. Khi sử dụng cách biểu diễn tập mờ rời rạc, phép lấy phần bù của hai tập mờ $G^{\sim}$ và $K^{\sim}$ được thực hiện như sau:

$\sim G^{\sim}$ = ~ (0,0/1 + 0,0/2 + 0,0/3 + 0,1/4 + 0,3/5 + 0,5/6 + 0,7/7 + 0,9/8 +1,0/9 + 1,0/10)

= (1,0/1 + 1,0/2 + 1,0/3 + 0,9/4 + 0,7/5 + 0,5/6 + 0,3/7 + 0,1/8 +0,0/9 + 0,0/10)

còn

$\sim K^{\sim}$ = ~ (1,0/1 + 0,9/2 + 0,8/3 + 0,6/4 + 0,4/5 + 0,2/6 + 0,0/7 + 0,0/8 + 0,0/9 + 0,0/10)

= (0,0/1 + 0,1/2 + 0,2/3 + 0,4/4 + 0,6/5 + 0,8/6 + 1,0/7 + 1,0/8 + 1,0/9 + 1,0/10)

Tương tự như trên, phép lấy phần bù cũng có thể thực hiện trên bảng dữ liệu, cụ thể như sau:

Bảng 1.5: *Phần bù của tập mờ trên **U***

U	1	2	3	4	5	6	7	8	9	10
$G^{\sim}$	0,0	0,0	0,0	0,1	0,3	0,5	0,7	0,9	1,0	1,0
$\sim G^{\sim}$	1,0	1,0	1,0	0,9	0,7	0,5	0,3	0,1	0,0	0,0
$K^{\sim}$	1,0	0.9	0,8	0,6	0,4	0,2	0,0	0,0	0,0	0,0
$\sim K^{\sim}$	0,0	0,1	0,2	0,4	0,6	0,8	1,0	1,0	1,0	1,0

1.3.4. Phép tổng và tích đại số của các tập mờ

Phép cộng đại số hai tập mờ: Cho hai tập mờ $A^{\sim}$ và $B^{\sim}$ trên tập vũ trụ ***U***. Tổng đại số của hai tập mờ này là một tập mờ, ký hiệu là $A^{\sim} \oplus B^{\sim}$, được định nghĩa bởi đẳng thức sau:

Trong trường hợp ***U*** là hữu hạn hay vô hạn đếm được,

$$A^{\sim} \oplus B^{\sim} = \sum_{u\in U}[\mu_{A^{\sim}}(u)+\mu_{B^{\sim}}(u)-\mu_{A^{\sim}}(u).\mu_{B^{\sim}}(u)]/u\ ,$$

Trong trường hợp ***U*** là vô hạn continuum,

$$A^{\sim} \oplus B^{\sim} = \int_{u\in U}[\mu_{A^{\sim}}(u)+\mu_{B^{\sim}}(u)-\mu_{A^{\sim}}(u).\mu_{B^{\sim}}(u)]du\ .$$

Lưu ý rằng giá trị biểu thức $\mu_{A^{\sim}}(u) + \mu_{B^{\sim}}(u) - \mu_{A^{\sim}}(u).\mu_{B^{\sim}}(u)$ luôn luôn thuộc [0, 1] và do đó các định nghĩa của phép tính $\oplus$ trên là đúng đắn.

Phép nhân đại số hai tập mờ: Nhân đại số hai tập mờ $A^{\sim}$ và $B^{\sim}$ là một tập mờ, ký hiệu là $A^{\sim} \otimes B^{\sim}$, được xác định như sau:

Trong trường hợp ***U*** là hữu hạn hay vô hạn đếm được,

$$A^{\sim} \otimes B^{\sim} = \sum_{u\in U} \mu_{A^{\sim}}(u).\mu_{B^{\sim}}(u)/u\,,$$

Trong trường hợp ***U*** là vô hạn continuum,

$$A^{\sim} \otimes B^{\sim} = \int_{u\in U} \mu_{A^{\sim}}(u).\mu_{B^{\sim}}(u)du\,.$$

1.3.5. Phép tập trung hay phép co (concentration)

Cho tập mờ $A^{\sim}$ trên ***U***. Phép tập trung tập mờ $A^{\sim}$ là tập mờ, ký hiệu là CON($A^{\sim}$), được định nghĩa như sau:

$$\text{CON}(A^{\sim}) = \int_{u\in U} \mu^{\alpha}_{A^{\sim}}(u)du \;=\; (A^{\sim})^{\alpha},\ \text{với}\ \ \alpha > 1$$

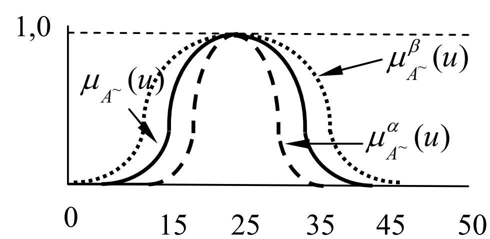
Vì $\alpha > 1$ nên $\mu^{\alpha}_{A^{\sim}}(u) < \mu_{A^{\sim}}(u)$ và do đó miền giới hạn bởi hàm $\mu^{\alpha}_{A^{\sim}}(u)$ sẽ nằm trọn trong miền giới hạn bởi hàm $\mu_{A^{\sim}}(u)$, hàm thuộc $\mu_{A^{\sim}}(u)$ của tập mờ bị co lại sau phép tập trung. Nói khác đi tập mờ CON($A^{\sim}$) biểu thị một khái niệm đặc tả hơn khái niệm gốc biểu thị bởi tập mờ $A^{\sim}$ (xem Hình 1.3). Về trực quan chúng ta thấy khái niệm mờ càng đặc tả thì nó càng chính xác hơn, ít mờ hơn và gần giá trị kinh điển hơn.

Thông thường người ta sử dụng phét tập trung để biểu thị ngữ nghĩa tác động của gia tử *rất* (very) vì ngữ nghĩa, chẳng hạn, của khái niệm *rất trẻ* là đặc tả hay ít mờ hơn so với khái niệm *trẻ*.

1.3.6. Phép dãn (Dilation)

Ngược với phép *tập trung* là *phép dãn*. Phép dãn khi tác động vào một tập mờ $A^{\sim}$, ký hiệu là DIL($A^{\sim}$), được xác định bởi đẳng thức sau:

$$\text{DIL}(A^{\sim}) = \int_{u\in U} \mu^{\beta}_{A^{\sim}}(u)du \;=\; (A^{\sim})^{\beta},\ \text{với}\ \ \beta < 1$$



Hình 1.3: *Phép tập trung*

Trong trường hợp này ta thấy $\mu_{A^{\sim}}^{\beta}(u) > \mu_{A^{\sim}}(u)$ và do đó phép dãn sẽ làm hàm thuộc của tập mờ đó *dãn nở* ra, hàm thuộc của tập mờ thu được sẽ xác định một miền thực sự bao hàm miền giới hạn bởi hàm thuộc của tập mờ gốc. Trên Hình 1.3, ta thấy đường cong nét chấm biểu thị hàm thuộc $\mu_{A^{\sim}}^{\beta}(u)$ còn đường cong nét liền biểu thị hàm thuộc $\mu_{A^{\sim}}(u)$. Ngữ nghĩa của khái niệm mờ biểu thị bởi tập mờ kết quả ít đặc tả hơn hay ngữ nghĩa của nó càng mờ hơn.

Ngược với hay đối ngẫu với việc sử dụng phép CON, phép DIL được sử dụng để biểu thị ngữ nghĩa của gia tử *có thể* hay *xấp xỉ* vì ngữ nghĩa của khái niệm *có thể trẻ* ít đặc tả hơn hay tính mờ của nó lớn hơn.

Ví dụ 1.8. Xét tập vũ trụ ***U*** = {1, 2, …, 8} và hai tập mờ $A^{\sim}$ và $B^{\sim}$ trên ***U*** được cho như sau:

$$A^{\sim} = 0{,}8/3 + 1{,}0/5 + 0{,}6/6 \quad \text{và} \quad B^{\sim} = 0{,}7/3 + 1{,}0/4 + 0{,}5/6$$

Khi đó ta có:

$$A^{\sim} \oplus B^{\sim} = 0{,}94/3 + 1{,}0/4 + 1{,}0/5 + 0{,}8/6$$

$$A^{\sim} \otimes B^{\sim} = 0{,}56/3 + 0{,}30/6$$

$$\mathrm{CON}(A^{\sim}) = 0{,}64/3 + 1{,}0/5 + 0{,}36/6 \text{, với } \alpha = 2.$$

$$\mathrm{DIL}(A^{\sim}) = \sqrt{0{,}8}/3 + 1{,}0/5 + \sqrt{0{,}6}/6 \text{, với } \beta = 1/2$$

1.3.7. Tích Đề-ca-tơ các tập mờ

Cho A_i là tập mờ của tập vũ trụ $\boldsymbol{U}_i$, i = 1, 2, …, n. Tích Đê-ca-tơ của các tập mờ $A_i^{\sim}$, i = 1, 2, …, n, ký hiệu là $A_1^{\sim} \times A_2^{\sim} \times \ldots \times A_n^{\sim}$ hay $\Pi_{i=1}^{n} A_i^{\sim}$, là một tập mờ trên tập vũ trụ $\boldsymbol{U}_1 \times \boldsymbol{U}_2 \times \ldots \times \boldsymbol{U}_n$ được định nghĩa như sau:

$$A_1^{\sim} \times A_2^{\sim} \times \ldots \times A_n^{\sim} = \int_{U_1 \times \ldots \times U_n} \mu_{A_1}(u_1) \cap \ldots \cap \mu_{A_n}(u_n) / (u_1, \ldots, u_n)$$

Ví dụ 1.9. Cho $\boldsymbol{U}_1 = \boldsymbol{U}_2$ = {1, 2, 3} và 2 tập mờ

$$A^{\sim} = 0{,}5/1 + 1{,}0/2 + 0{,}6/3 \quad \text{và} \quad B^{\sim} = 1{,}0/1 + 0{,}6/2$$

Khi đó,

$A^{\sim} \times B^{\sim} =$ 0,5/(1,1) + 1,0/(2,1) + 0,6/(3,1) + 0,5/(1,2) + 0,6/(2,2) + 0,6/(2,3).

Một ví dụ ứng dụng của tích Đê-ca-tơ là kết nhập (aggreegation) các thông tin mờ về các thuộc tính khác nhau của một đối tượng. Ví dụ, trong các hệ luật của các hệ trợ giúp quyết định hay hệ chuyên gia, hệ luật trong điều khiển thường có các luật dạng sau đây:

Nếu $X_1 := \tilde{A_1}$ and $X_2 := \tilde{A_2}$ and … and $X_n := \tilde{A_n}$ thì $Y := \tilde{B}$

trong đó các X_i là các biến ngôn ngữ (vì giá trị của nó là các ngôn ngữ được xem như là nhãn của các tập mờ) và A_i là các tập mờ trên miền cơ sở $\boldsymbol{U}_i$ của biến X_i. Hầu hết các phương pháp giải liên quan đến các luật nếu-thì trên đều đòi hỏi việc tích hợp các dữ liệu trong phần tiền tố "nếu" nhờ toán tử kết nhập, một trong những toán tử như vậy là lấy tích Đề-ca-tơ $\tilde{A_1} \times \tilde{A_2} \times \ldots \times \tilde{A_n}$.

1.3.8. Phép tổ hợp lồi (convex combination)

Cho $\tilde{A_i}$ là tập mờ của tập vũ trụ $\boldsymbol{U}_i$ tương ứng với biến ngôn ngữ X_i, i = 1, 2, …, n, và $w_i \in (0, 1]$, là các trọng số về mức độ quan trọng tương đối của biến X_i so với các biến khác, i = 1, 2, …, n, và thỏa ràng buộc $\sum_{i=1}^{n} w_i = 1$. Khi đó *tổ hợp lồi* của các tập mờ $\tilde{A_i}$, i = 1, 2, …, n, là một tập mờ $\tilde{A}$ xác định trên $\boldsymbol{U} = \boldsymbol{U}_1 \times \boldsymbol{U}_2 \times \ldots \times \boldsymbol{U}_n$, hàm thuộc của nó được định nghĩa như sau:

$$\mu_{\tilde{A}}(u_1,...,u_n) = \sum_{i=1}^{n} w_i \mu_{\tilde{A_i}}(u_i)$$

trong đó Σ là tổng số học (chứ không phải là tổng hình thức).

Phép tổ hợp lồi thường được sử dụng để biểu thị ngữ nghĩa của gia tử kiểu "cốt yếu" (essentially) hay "đặc trưng" hay "đặc tính tiêu biểu" (typically). Ví dụ, khái niệm mờ về người "*To lớn*" được biểu thị *một cách cốt yếu* từ ngữ nghĩa của các khái niệm người *Cao* và *Béo*. Như vậy ngữ nghĩa của "*To lớn*" có thể biểu thị qua ngữ nghĩa của "*Cao*" và của "*Béo*" thông qua phép tổ hợp lồi.

Cụ thể, giả sử ngữ nghĩa của các tập mờ *Béo* trên miền $\boldsymbol{U}_1$ = [40, 100] theo đơn vị kg và của *Cao* trên miền $\boldsymbol{U}_2$ = [50, 220] với đơn vị cm được biểu thị như sau:

$$Béo = \int_{40}^{100} \left\{ 1 + \left(\frac{u_1 - 40}{30} \right)^{-2} \right\}^{-1} du_1$$

$$Cao = \int_{40}^{100} \left\{ 1 + \left(\frac{u_2 - 140}{30} \right)^{-2} \right\}^{-1} du_2$$

Khi đó, tập mờ *To-lớn* được biểu thị nhờ phép tổ hợp lồi sau:

$$To\text{-}lớn = 0,6\ Béo + 0,4\ Cao = \int_{40}^{100}\int_{50}^{220} \{0,6\mu_{Béo}(u_1) + 0,4\mu_{Cao}(u_2)\} du_1 du_2$$

Chẳng hạn, ta có:

$$\mu_{To\text{-}lớn}(70,170) = 0,6\times0,5 + 0,4\times0,5 = 0,5$$
$$\mu_{To\text{-}lớn}(80,170) = 0,6\times0,64 + 0,4\times0,5 = 0,584$$
$$\mu_{To\text{-}lớn}(70,180) = 0,6\times0,5 + 0,4\times0,64 = 0,556$$

1.3.9. Phép mờ hóa (Fuzzification)

Việc mờ hóa có hai bài toán:

- Tìm tập mờ biểu thị một tập kinh điển hay, một cách tổng quát hơn, hãy mờ hóa một tập mờ đã cho $A^{\sim}$;
- Tìm độ thuộc của giá trị ngôn ngữ của một biến ngôn ngữ tương ứng với một dữ liệu đầu vào là thực hoặc mờ.

Theo nghĩa thứ nhất ta định nghĩa phép mờ hóa như sau:

Phép mờ hóa F của một tập mờ $A^{\sim}$ trên tập vũ trụ ***U*** sẽ cho ta một tập mờ F($A^{\sim}$, $K^{\sim}$) được xác định theo công thức sau:

$$F(A^{\sim}, K^{\sim}) = \int_U \mu_{A^{\sim}}(u) K^{\sim}(u) du$$

trong đó $K^{\sim}(u)$ là một tập mờ trên ***U***, $u \in$ ***U***, được gọi là nhân (kernel) của F.

Nếu $\mu_{A^{\sim}}(u)$ là hàm thuộc của tập kinh điển 1-phần tử $\{u\}$, $\mu_{A^{\sim}}(u)$ chỉ bằng 1 tại phần tử *u* còn lại là bằng 0 hay ta có tập "mờ" $\{1/u\}$, thì ta có

$$F(\{1/u\}, K^{\sim}) = K^{\sim}(u)$$

Nếu $A^{\sim}$ là tập kinh điển A, $\mu_A(u) = 1$ trên A và bằng 0 ngoài A, thì mờ hóa của A với nhân $K^{\sim}(u)$ sẽ là tập mờ sau:

$$\mathrm{F}(A, K^{\sim}) = \int_A K^{\sim}(u)du$$

Ví dụ 1.10. Cho hai tập mờ $A^{\sim}$ và $K^{\sim}$ trên $\boldsymbol{U}$ như sau:

$$\boldsymbol{U} = \{a, b, c, d\}, \quad A^{\sim} = 0{,}8/a + 0{,}6/b\ ,$$
$$K^{\sim}(a) = 1{,}0/a + 0{,}4/b \ \text{ và } \ K^{\sim}(b) = 1{,}0/b + 0{,}4/a + 0{,}4/c$$

Khi đó

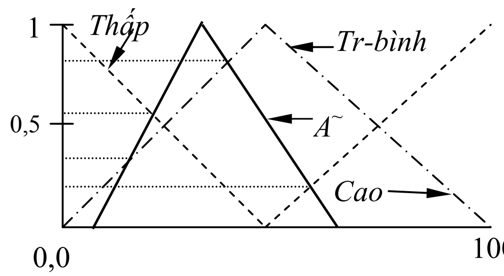
$$\begin{aligned} \mathrm{F}(A^{\sim}, K^{\sim}) &= 0{,}8(1{,}0/a + 0{,}4/b) + 0{,}6(1{,}0/b + 0{,}4/a + 0{,}4/c) \\ &= 0{,}8/a + 0{,}32/b + 0{,}6/b + 0{,}24/a + 0{,}24/c \\ &= (0{,}8 \cup 0{,}24)/a + (0{,}32 \cup 0{,}6)/b + 0{,}24/c \\ &= 0{,}8/a + 0{,}6/b + 0{,}24/c \end{aligned}$$

Người ta cho rằng phép mờ hóa như trên có vai trò quan trọng trong biểu diễn ngữ nghĩa của các gia tử như *ít nhiều* (more or less), *một chút* hay *hơi* (slightly), *nhiều* (much). Chẳng hạn, với khái niệm mờ *giỏi* chỉ về NĂNG LỰC của chuyên viên, thì khái niệm *hơi giỏi* có thể được biểu thị bằng phép mờ hóa tác động vào tập mờ biểu diễn khái niệm *giỏi*.

Bài toán mờ hóa thứ 2 được giới hạn trong trường hợp tập vũ trụ là tập hữu hạn các giá trị ngôn ngữ

Cụ thể bài toán mờ hóa trong trường hợp này như sau: Giả sử **T** là tập các giá trị ngôn ngữ của một biến ngôn ngữ X nào đó với miền cơ sở $\boldsymbol{U}$. Cho một tập kinh điển hoặc tập mờ $A^{\sim}$ trên $\boldsymbol{U}$. Hãy tìm tập mờ trên miền **T** biểu thị tập mờ $A^{\sim}$ hay, một cách tương đương, hãy tìm độ thuộc của giá trị τ trong **T** tương ứng với dữ liệu đầu vào $A^{\sim}$.

Chẳng hạn, ta xét biến NHIỆT ĐỘ thời tiết với **T** = {*Thấp*, *Trung-bình*, *Cao*} với không gian cơ sở là [0, 100] theo thang độ C. Vấn đề là cần xác định độ thuộc hay giá trị chân lý *TV* của mệnh đề $A^{\sim} := \tau$, $\tau \in \mathbf{T}$, với := được hiểu là "*xấp xỉ*



Hình 1.4: *Các hàm thuộc của biến NHIỆT ĐỘ*

bằng". Cụ thể chúng ta cần xác định giá trị chân lý như sau:

$\mu(Thấp) \quad = TV(A^{\sim} := Thấp)$
$\mu(Tr\text{-}bình) = TV(A^{\sim} := Tr\text{-}bình)$
$\mu(Cao) \quad = TV(A^{\sim} := Cao)$

Việc xác định giá trị chân lý này được tiến hành như sau (xem Hình 1.4): Chúng ta lần theo đồ thị của hàm thuộc của tập mờ đầu vào $A^{\sim}$ sẽ thấy nó cắt đồ thị của hàm thuộc *Thấp* ở giá trị 0,52. Giá trị này biểu thị độ phù hợp nhất của tập mờ $A^{\sim}$ biểu diễn qua tập mờ hay khái niệm mờ *Thấp* là 0,52. Tương tự, đồ thị của $A^{\sim}$ sẽ cắt đồ thị của tập mờ *Tr-bình* ở hai giá trị 0,34 và 0,82 và do đó độ phù hợp nhất của việc biểu diễn ngữ nghĩa của $A^{\sim}$ qua khái niệm mờ *Tr-bình* là giá trị 0,82 lớn hơn. Cũng như vậy, độ phù hợp của $A^{\sim}$ biểu thị qua khái niệm *Cao* là 0,18. Như vậy, việc mờ hóa sẽ đưa việc biểu diễn tập mờ $A^{\sim}$ trên $\boldsymbol{U}$ thành tập mờ trên tập các giá trị ngôn ngữ **T** sau:

$$NHIỆT\_ĐỘ(A^{\sim}) = 0{,}54/Thấp + 0{,}82/Tr\text{-}bình + 0{,}18/Cao \quad (4^*)$$

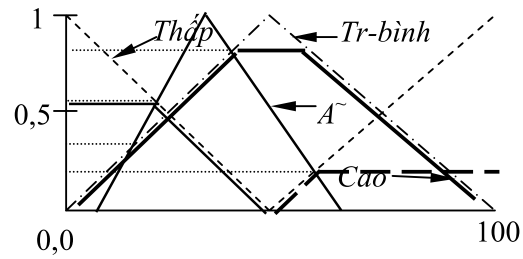
1.3.10. Phép khử mờ

Trong điều khiển mờ cũng như trong lập luận trong các hệ chuyên gia với các luật tri thức mờ, dữ liệu đầu ra nhìn chung đều là những tập mờ. Thực tế chúng ta cũng thường gặp nhu cầu chuyển đổi dữ liệu mờ đầu ra thành giá trị thực một cách phù hợp. Phương pháp chuyển đổi như vậy được gọi là phương pháp khử mờ (defuzzification). Nhu cầu này thường gặp nhất trong điều khiển mờ vì đầu ra đòi hỏi là giá trị thực để tác động vào một quá trình thực nào đó.

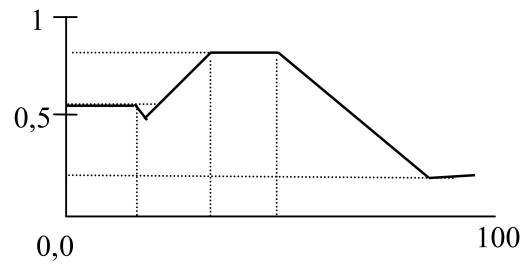
Giả sử dữ liệu đầu ra được biểu diễn ở dạng (4*) với các tập mờ của các giá trị ngôn ngữ được biểu thị trong Hình 1.4.

Trước khi trình bày một số phương pháp khử mờ, chúng ta hãy đưa ra phương pháp biến đổi để tính hàm thuộc của tập mờ được biểu diễn bằng biểu thức dạng (4*). Trước hết ta nhớ lại rằng tập mờ với hàm thuộc có dạng $\mu(u) \equiv a$, $a \in [0, 1]$, được ký hiệu là $a\boldsymbol{U}$, nó là tích của số vô hướng a và tập kinh điển $\boldsymbol{U}$. Khi đó, hạng thức trong (4*), chẳng hạn 0,54/*Thấp*, sẽ được hiểu là biểu thức 0,54$\boldsymbol{U}$ AND *Thấp*, trong đó *Thấp* là nhãn của tập mờ với hàm thuộc được cho trong Hình 1.4. Từ Nhận xét 1.1, chúng ta có thể hiểu các phép cộng hình thức "+" sẽ là phép OR mà ngữ nghĩa của nó là phép $\cup$ trong dàn L([0,1]).

Có nhiều cách biểu thị ngữ nghĩa phép AND và phép OR trên đoạn [0, 1]. Một cách tổng quát, ta có thể chọn một cặp đối ngẫu t-norm và t-conorm bất kỳ mà chúng sẽ được đề cập đến sau này khi nói về các đại số liên hợp tập hợp mờ để biểu thị ngữ nghĩa của hai phép AND và OR. Dưới đây ta sẽ chọn ngữ nghĩa của AND là phép Min, và OR là phép Max. Trong Hình 1.5 ta có các kết quả của việc thức hiện phép AND cho từng hạng tử trong công thức (4*): hạng tử thứ nhất được biểu thị bằng hình thang thứ nhất với chiều cao là 0,54; hạng tử thứ hai được biểu thị bằng hình thang thứ hai ở giữa, với chiều cao 0,82; hạng tử thứ ba được biểu thị bằng hình thang bên phải với chiều cao là 0,18.



Hình 1.5: *Các hàm thuộc của 3 hạng tử trong* (1.5)



Hình 1.6. *Hàm thuộc hợp của 3 hạng tử trong* (1.5)

Hình 1.6 biểu thị kết quả của phép OR của 3 hạng tử với ngữ nghĩa được biểu thị trong Hình 1.5.

Như vậy, bất kỳ một tập mờ nào được cho ở dạng công thức (4*) chúng ta đều có thể biến đổi về tập mờ có dạng ở Hình 1.6.

Bây giờ bài toán khử mờ được cụ thể hóa bằng bài toán cho trước một tập mờ với hàm thuộc được biểu thị bằng đồ thị, chẳng hạn như trong Hình 1.6. Hãy xác định phương pháp biến đổi tập mờ đó về một giá trị thực thuộc miền cơ sở $\boldsymbol{U}$. Với ví dụ đang xét, ta có biến *NHIỆT ĐỘ* với $\boldsymbol{U}$ = [0, 100] theo thang độ C.

Thường chúng ta có nhiều cách để giải bài toán khử mờ. Chúng ta không có những ràng buộc chặt chẽ nào về việc định nghĩa một phương pháp khử mờ. Bất kỳ nhà nghiên cứu ứng dụng nào cũng có thể đưa ra một định nghĩa về một phương pháp khử mờ, miễn là nó phù hợp với một ứng dụng nào đó hay nó phù hợp với một ý tưởng nào đó về ngữ nghĩa của phép khử mờ.

Tuy nhiên, về trực quan chúng ta có thể đưa ra những yêu cầu để một phương pháp khử mờ được xem là tốt. Hellendoorn, H. and C. Thomas năm 1993 đã đưa ra 5 tiêu chuẩn trực quan sau. (i) *Tính liên tục*, nghĩa là một sự thay đổi nhỏ của dữ liệu đầu vào của phương pháp nó cũng chỉ tạo ra những thay đổi nhỏ ở dữ liệu đầu ra; (ii) *Tính không nhập nhằng* (disambiguity), nghĩa là phương pháp chỉ sinh ra một giá trị đầu ra duy nhất; (iii) *Tính hợp lý* (plausibility) đòi hỏi rằng giá trị đầu ra phải nằm ở vùng trung tâm của tập mờ và độ thuộc hay giá trị hàm thuộc tại đó phải lớn (không nhất thiết lớn nhất); (iv) *Độ phức tạp tính đơn giản* (computational simplicity), một đòi hỏi tự nhiên và (v) *Tính trọng số của phương pháp* (weighting method) đòi hỏi phương pháp tính đến trọng số hay “sự ưu tiên” của các tập mờ kết quả đầu ra (đối với trường hợp bài toán cho nhiều kết quả đầu ra như đối với một số phương pháp lập luận mờ đa điều kiện).

Nói chung, chúng ta có thể hiểu các tiêu chuẩn cần bảo đảm giá trị khử mờ của tập mờ $A^{\sim}$ là *phần tử thực đại diện* một cách hợp lý của $A^{\sim}$.
Sau đây chúng ta nghiên cứu một vài phương pháp khử mờ.

1.3.10.1. Phương pháp cực đại trung bình (average maximum)

Cho tập mờ $A^{\sim}$ với hàm thuộc $\mu_{A^{\sim}}$. Gọi $u$min và $u$max tương ứng là hai giá trị nhỏ nhất và lớn nhất của miền cơ sở $\boldsymbol{U}$ mà tại đó hàm thuộc $\mu_{A^{\sim}}$ nhận giá trị lớn nhất (cực đại toàn phần). Ký hiệu giá trị khử ở của $A^{\sim}$ theo phương pháp cực đại trung bình là $D_{Av\text{-}max}(A^{\sim})$. Khi đó $D_{Av\text{-}max}(A^{\sim})$ được định nghĩa như sau:

$$D_{AveMax}(A^{\sim}) = \frac{u\min + u\max}{2}$$

Ý tưởng của phương pháp này là chúng ta chỉ quan tâm đến các giá trị của $\boldsymbol{U}$ mà tại đó nó phù hợp hay tương thích với ngữ nghĩa của tập mờ $A^{\sim}$ nhất, tại đó độ thuộc là cực đại toàn phần. Những giá trị khác của $\boldsymbol{U}$ mà tại đó độ thuộc nhỏ hơn 1 đều bị bỏ qua. Vì vậy, một khả năng lựa chọn giá trị khử mờ là giá trị trung bình của giá trị nhỏ nhất và giá trị lớn nhất tại đó độ thuộc vào tập mờ là lớn nhất. Đó chính là lý do người ta gọi phương pháp khử mờ này là *phương pháp cực đại trung bình*.

Ví dụ trên Hình 1.6, hàm thuộc $\mu_{A^{\sim}}$ đạt cực đại toàn phần trên đoạn [41, 59] và, do đó, chúng ta ta có:

$$D_{AveMax}(A^{\sim}) = \frac{41+59}{2} = 50 \,.$$

1.3.10.2. Phương pháp cực đại trung bình có trọng số

Ý tưởng của phương pháp này là tìm những đoạn tại đó hàm thuộc $\mu_{A^{\sim}}$ đạt cực đại địa phương. Nghĩa là tại các giá trị của miền cơ sở mà độ thuộc của chúng đạt cực đại địa phương. Nói khác đi các giá trị đó của ***U*** thuộc về tập mờ $A^{\sim}$ với độ tin cậy có độ trội nhất. Các giá trị như vậy cần được tham gia "đóng góp" vào việc xác định giá trị khử mờ của tập $A^{\sim}$ với trọng số đóng góp chính là độ thuộc của chúng vào tập $A^{\sim}$. Chúng ta chọn cách đóng góp như vậy bằng phương pháp lấy trung bình có trọng số (*weighted average maxima method*). Vì vậy cách tính giá trị khử mờ của tập mờ $A^{\sim}$ như sau:

Xác định các giá trị của ***U*** mà tại đó hàm thuộc $\mu_{A^{\sim}}$ đạt giá trị cực đại địa phương. Ký hiệu $u\min_i$ và $u\max_i$ là giá trị lớn nhất và nhỏ nhất trong các giá trị của ***U*** mà tại đó hàm thuộc đạt cực đại địa phương. Giá trị trung bình cộng của $u\min_i$ và $u\max_i$ sẽ được ký hiệu là $uave\max_i$, trong đó, chỉ số *i* chỉ nó là giá trị tương ứng với giá trị cực đại địa phương thứ *i*.

Giả sử hàm thuộc $\mu_{A^{\sim}}$ có *m* giá trị cực đại địa phương, *i* = 1, 2, …, *m*. Khi đó giá trị khử mờ của tập mờ $A^{\sim}$ được tính theo công thức trung bình cộng có trọng số như sau:

$$D_{w\text{-}AveMax} = \sum_{i=1}^{m} \frac{\mu(uave\max_i)uave\max_i}{\sum_{i=1}^{m}\mu(uave\max_i)}$$

Ví dụ, chúng ta xét tập mờ $A^{\sim}$ được cho trong Hình 1.6. Hàm thuộc $\mu_{A^{\sim}}$ đạt cực đại địa phương trên hai đoạn thẳng, đoạn [0, 23] và đoạn [41, 59]. Do đó, theo công thức ta có $uave\max_1 = (0 + 23)/2 = 11{,}5$ và $uave\max_2 = (41 + 59)/2 = 50$. Theo công thức chúng ta có:

$$D_{w\text{-}AveMax} = \frac{\mu(11{,}5)11{,}5 + \mu(50)50}{\mu(11{,}5) + \mu(50)} = \frac{0{,}54 \times 11{,}5 + 0{,}82 \times 50}{0{,}54 + 0{,}82} = \frac{47{,}21}{1{,}36} \approx 34{,}71$$

1.3.10.3. Phương pháp trọng tâm

Trong hai phương pháp trên, người ta chỉ quan tâm đến giá trị của miền ***U*** mà tại đó hàm thuộc đạt cực đại, còn các giá trị khác đều bị bỏ qua. Như vậy có vẻ "thiếu bình đẳng". Phương pháp trọng tâm (*centroid method* hay *centre of gravity*) xuất phát từ ý tưởng mọi giá trị của ***U*** đều được đóng góp với trong số vào việc xác định giá trị khử mờ của tập mờ $A^{\sim}$, ở đây trọng số của nó là độ thuộc của phần tử thuộc vào tập mờ $A^{\sim}$.

Theo nghĩa thông thường của trọng tâm, công thức tính giá trị khử mờ có dạng sau:

$$D_{Centroid}(A^{\sim}) = \frac{\int_a^b u\mu(u)du}{\int_a^b \mu(u)du}$$

Ví dụ, ta tính giá trị khử mờ theo phương pháp trong tâm của tập mờ trong Hình 1.6. Theo công thức trên ta tính:

$$\int_0^{100} u\mu(u)du = \int_0^{23} 0{,}54 * udu + \int_{23}^{25}(-\tfrac{1}{50}u+1)udu + \int_{25}^{41}(\tfrac{1}{50}u)udu + \int_{41}^{59} 0{,}82udu + \int_{59}^{91}(-\tfrac{1}{50}u+2)udu + \int_{91}^{100} 0{,}18udu$$

= 142,83 + 24,946 + 355,306 + 738,0 + 1145,386 + 154,71 = 2561,178

$$\int_0^{100} \mu(u)du = 12{,}42 + 1{,}04 + 10{,}56 + 14{,}76 + 10{,}56 + 10{,}88 + 1{,}44 = 61{,}66$$

Do đó, $$D_{Centroid}(A^{\sim}) = \frac{2561{,}178}{61{,}66} = 41{,}537.$$

1.3.11. Nguyên lý thác triển và số học các số mờ

1.3.11.1. Nguyên lý thác triển

Vấn đề được đặt ra là cho một tập mờ A trên không gian ***U*** và một quan hệ kinh điển ρ trên $\boldsymbol{U} \times \boldsymbol{V}$ (hay nó cũng là một ánh xạ đa trị từ ***U*** sang ***V***), liệu tập mờ A sẽ cảm sinh một tập mờ B nào trên ***V*** nhờ "thông tin" từ quan hệ ρ?

Nguyên lý thác triển (extension principle) cho ta một quy tắc xác định tập mờ B dựa trên các thông tin mà quan hệ ρ cung cấp. Nguyên lý này được phát biểu như sau:

Cho ρ là một quan hệ kinh điển trên $\boldsymbol{U} \times \boldsymbol{V}$. Với $v \in \boldsymbol{V}$, ta ký hiệu

$$\rho^{-1}(v) = \{u \in \boldsymbol{U}: \rho(u, v)\}$$

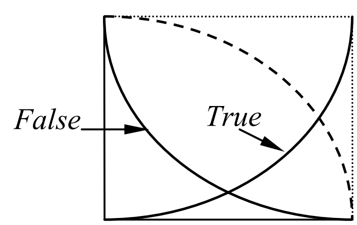
Khi đó, mỗi tập mờ A trên $\boldsymbol{U}$ sẽ cảm sinh một tập mờ B trên $\boldsymbol{V}$ nhờ quan hệ ρ với hàm thuộc $\mu_B(v)$ được tính theo công thức sau:

$$\mu_B(v) \ = \ \sup_{u\in\rho^{-1}(v)} \mu_A(u) \, .$$

Ta cho một vài ví dụ về ứng dụng của nguyên lý thác triển trên.

Ví dụ 1.11. Người ta thường biểu diễn khái niệm chân lý như là một tập mờ trên $\boldsymbol{U} = [0,1]$, chẳng hạn hàm thuộc μ_{True} của khái niệm *True* được cho trong Hình 1.7. Thông thường, phần bù của tập mờ *True* biểu thị phép phủ định và do đó đường cong gạch từng đoạn biểu thị khái niệm *False*. Về trực quan quan sát trên Hình 1.7 chúng ta thấy không hợp lý.

Bây giờ chúng ta định nghĩa khái niệm phủ định bằng việc áp dụng nguyên lý thác triển. Trong lôgic đa trị với miền giá trị chân lý trên đoạn [0,1], phép phủ định là 1-, $\neg\, t = 1 - $t. nó xác định một ánh xạ φ từ [0,1] vào [0,1]. Theo nguyên lý thác triển, tập mờ *True* sẽ cảm sinh tập mờ cũng trên [0,1], chính là tập mờ *False*, với hàm thuộc là



Hình 1.7

$$\mu_{False}(t) = \ \sup_{s\in\varphi^{-1}(t)} \mu_{True}(s) = \sup_{s\in\{1-t\}} \mu_{True}(s)$$
$$= \ \mu_{True}(1 - t)$$

Hàm thuộc này là đường cong đối xứng với μ_{True} qua đường thẳng s = 0,5 và nó biểu thị khái niệm *False* một cách hợp lý hơn.

Ví dụ 1.12. Bây giờ ta xét một ví dụ phức tạp hơn về việc áp dụng nguyên lý thác triển. Xét không gian $\boldsymbol{U} = \boldsymbol{R}$, tập tất cả các số thực và phép tính 2-ngôi $a * b$ trên các số thực. Phép tính này xác định một quan hệ hai ngôi, hơn nữa nó xác định một ánh xạ ψ: $\boldsymbol{R} \times \boldsymbol{R} \rightarrow \boldsymbol{R}$. Do đó, theo nguyên lý thác triển, mỗi cặp tập mờ A và B trên $\boldsymbol{R}$ sẽ cảm sinh một tập mờ C cũng trên $\boldsymbol{R}$ nhờ ánh xạ ψ với hàm thuộc được xác định như sau:

$$\mu_C(t) = \sup_{(a,b)\in\psi^{-1}(t)} \mu_A(a) \wedge \mu_B(b) = \sup_{a*b=t} \mu_A(a) \wedge \mu_B(b) \quad (5*)$$

Về hình thức hóa, công thức trên rõ ràng và dễ hiểu, nhưng về tính toán nó lại rất phức tạp: cho trước hai hàm thuộc $\mu_A(a)$ và $\mu_B(b)$ chúng ta khó có thể tính toán cụ thể được hàm thuộc $\mu_C(t)$ dựa theo công thức trên.

Để khắc phục khó khăn tính toán này, chúng ta ứng dụng cấu trúc số học trên các khoảng, cụ thể trên các tập mức hay lát cắt của tập mờ.

1.3.11.2. Số học các khoảng và ứng dụng đối với nguyên lý thác triển

Trước hết chúng ta khảo sát lại công thức (5*) dựa trên các tập mức. Chúng ta biết rằng có một tương ứng 1-1 giữa tập mờ A trên $\boldsymbol{U}$ và họ đơn điệu giảm các tập mức $\{A_\alpha : \alpha \in (0, 1]\}$, $\alpha < \beta \Rightarrow A_\alpha \supseteq A_\beta$. Vì vậy, thay vì tính trực tiếp hàm thuộc của một tập mờ, ta tính họ các tập mức. Đặc biệt trong trường hợp rời rạc hóa, số tập mức như vậy chỉ hữu hạn. Để cho gọn, ta ký hiệu giá của tập mờ A là $A(0)$, $A(0) = \{u \in \boldsymbol{U}: \mu_A(u) > 0\}$.

Giả thiết rằng ta chỉ xét các tập mờ mà hàm thuộc của chúng *liên tục*.

Để phân tích công thức (5*) trên quan điểm tập mức một cách cụ thể, ta giả thiết phép * là phép + số học trên $\boldsymbol{R}$. Ta sẽ chứng tỏ rằng

$$C_\alpha = A_\alpha + B_\alpha = \{a + b: a \in A_\alpha \;\&\; b \in B_\alpha\} \quad (6*)$$

Thực vậy, giả sử $t \in C_\alpha$, $\mu_C(t) \geq \alpha$. Từ (5*), ta suy ra $\mu_A(a) \geq \alpha$ và $\mu_B(b) \geq \alpha$ với ít nhất một cặp (a, b) sao cho $a + b = t$. Nghĩa là, ta có $C_\alpha \subseteq \{a + b: a \in A_\alpha \;\&\; b \in B_\alpha\}$. Ngược lại, dễ dàng thấy rằng với mọi cặp (a, b) sao cho $a \in A_\alpha \;\&\; b \in B_\alpha$, thì $\mu_A(a) \wedge \mu_B(b) \geq \alpha$ và do đó, theo (5*), với $t = a + b$, ta có $\mu_C(t) \geq \alpha$ hay $a + b \in C_\alpha$. Như vậy chúng ta đã chứng minh công thức (6*) là đúng.

Tương tự như vậy chúng ta có thể thiết lập các công thức tương tự như (6*) cho các phép tính số học khác.

Với giả thiết các hàm thuộc của các tập mờ được xét A là chuẩn, tức là $high(A) = 1$ hay $A_1 \neq \emptyset$, và liên tục, các tập mức đều là các đoạn thẳng. Khi đó, công thức (6*) có nghĩa đoạn C_α là tổng của 2 đoạn A_α và B_α. Như vậy, (6*) dẫn đến việc nghiên cứu số học các khoảng đóng.

Xét họ các khoảng đóng, giới nội trên tập số thực $\boldsymbol{R}$, ký hiệu là họ $Intvl(\boldsymbol{R})$.

Ta định nghĩa các phép tính số học trên các khoảng như vậy như sau. Gọi * là phép tính 2-ngôi bất kỳ trên số thực $\boldsymbol{R}$, * có thể là phép cộng (+), phép trừ (–), phép nhân (.) và phép chia (/) số học, thì nó sẽ cảm sinh một phép tính trên $Intvl(\boldsymbol{R})$ cũng được ký hiệu là phép * và được định nghĩa như sau:

$$[a, b] * [c, d] = \{u * v : u \in [a, b] \ \&\ v \in [c, d]\} \qquad (7^*)$$

với giả thiết rằng nếu $*$ là phép chia thì ta giả thiết đoạn $[c, d]$ không chứa phần tử 0 của $\boldsymbol{R}$.

Từ (7*) ta dễ dàng suy ra các công thức sau

$$[a, b] + [c, d] = [a + b, c + d]$$

$$[a, b] - [c, d] = [a - b, c - d]$$

$$[a, b] / [c, d] = [a, b] . [1/d, 1/c]$$

$[a, b] . [c, d] = [e, f]$, với $e = \min \{ac, ad, bc, bd\}$ còn $f = \max \{ac, ad, bc, bd\}$

Lưu ý rằng vì mỗi số a, b, c và d có thể âm hoặc dương nên ta phải tính min, max để xác định đầu mút của khoảng kết quả của phép nhân.

Trở lại với nguyên lý thác triển đối với ánh xạ xác định bởi phép tính số học $*$ trên số thực với tập mờ cảm sinh được tính trên các tập mức như ở dạng công thức (6*). Nếu các tập mờ A và B là chuẩn và liên tục, thì tất cả các tập mức A_α và B_α đều là các khoảng đóng giới nội, chúng là các phần tử của $Intvl(\boldsymbol{R})$ và do đó các công thức ở dạng (6*) đều được tính toán dựa trên số học các khoảng.

1.3.11.3. Số mờ và số học các số mờ

Xét tập mờ A trên tập các số thực $\boldsymbol{R}$. Về nguyên tắc, không có ràng buộc chặt đối với việc xây dựng các tập mờ để biểu thị ngữ nghĩa của các khái niệm ngôn ngữ. Tuy nhiên, để đơn giản trong xây dựng các tập mờ và trong tính toán trên các tập mờ, người ta đưa ra khái niệm tập mờ có dạng đặc biệt, gọi là số mờ để biểu thị các khái niệm mờ về số như *gần* 10, *khoảng 15*, *lớn hơn nhiều so với* 10, …

Số mờ là tập mờ có các đặc điểm sau:

Nó là tập mờ chuẩn, tức là $high(A) = 1$;

Mọi tập mức A_α, $\alpha \in (0,1]$, là các khoảng đóng;

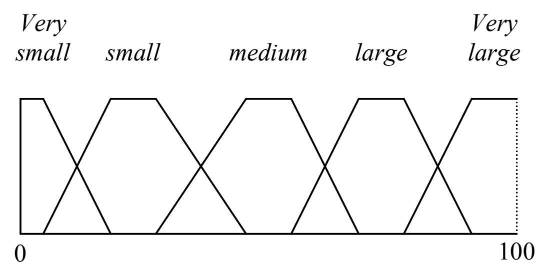
Support(A) là tập giới nội hay nó là một đoạn hữu hạn.

Trong nhiều tài liệu nghiên cứu và trong ứng dụng, người ta thường sử dụng các số mờ đặc biệt, gọi là các số mờ tam giác hay hình thang (xem Hình 1.8).

Các phép tính số học trên số mờ

Có hai cách định nghĩa các phép tính số học trên các số mờ.

Cách thứ nhất, dựa trên công thức (5*) khi sử dụng nguyên lý thác triển.



Hình 1.8: *Các số mờ của các giá trị ngôn ngữ*

Cách định nghĩa thứ hai, định nghĩa qua tập mức. Vì, như trong Mục 1.3.11.2, chúng ta đã thấy mỗi tập mờ A được xác định duy nhất bởi họ các tập mức $\{A_\alpha : \alpha \in (0,1]\}$. Khi đó ta có thể biểu diễn:

$$A = \bigcup_{\alpha \in (0,1]} A_\alpha$$

Giả sử $A = \bigcup_{\alpha \in (0,1]} A_\alpha$ và $B = \bigcup_{\alpha \in (0,1]} B_\alpha$ và $*$ là một phép tính số học hai ngôi nào đó trên số thực, $* \in \{+, -, ., /\}$. Theo định nghĩa số mờ, A_α và B_α là các khoảng đóng giới nội và $A_\alpha * B_\alpha$ là một phép tính số học trên các khoảng. Khi đó, ta định nghĩa:

$$A * B = \bigcup_{\alpha \in (0,1]} A_\alpha * B_\alpha$$

Ví dụ 1.13. Cho các tập mờ *A* và *B* với các hàm thuộc sau

$$\begin{aligned}
\mu_A(u) &= 0 && \text{với } u \le -1 \text{ và } u > 3 \\
&= (u+1)/2 && \text{với } -1 < u \le 1 \\
&= (3-u)/2 && \text{với } 1 < u \le 1, \\
\mu_B(u) &= 0 && \text{với } u \le 1 \text{ và } u > 5 \\
&= (u-1)/2 && \text{với } 1 < u \le 3 \\
&= (5-u)/2 && \text{với } 3 < u \le 5.
\end{aligned}$$

Khi đó, ta có thể kiểm chứng thấy rằng

$A_\alpha = [2\alpha - 1, 3 - 2\alpha]$, và $B_\alpha = [2\alpha + 1, 5 - 2\alpha]$.

và, do đó,

$$\begin{aligned}
(A+B)_\alpha &= [4\alpha, 8-4\alpha] \\
(A-B)_\alpha &= [4\alpha-6, 2-4\alpha] \\
(A\,.\,B)_\alpha &= [-4\alpha^2+12\alpha-5, 4\alpha^2-16\alpha+15] && \text{với } \alpha \in (0;0,5] \\
&= [4\alpha^2-1, 4\alpha^2-16\alpha+15] && \text{với } \alpha \in (0,5;1]
\end{aligned}$$

$$\begin{aligned}
(A/B)_\alpha &= [(2\alpha-1)/(2\alpha+1), (3-2\alpha)/(2\alpha+1)] \text{ với } \alpha \in (0;0,5] \\
&= [2\alpha-1)/(5-2\alpha), (3-2\alpha)/(2\alpha+1)] \text{ với } \alpha \in (0,5;1]
\end{aligned}$$

Trong trường hợp đơn giản này chúng ta có thể tính các hàm thuộc kết quả và thu được

$$\begin{aligned}
\mu_{A+B}(u) &= 0 && \text{với } u \le 0 \text{ và } u > 8 \\
&= u/4 && \text{với } 0 < u \le 4 \\
&= (8-u)/4 && \text{với } 4 < u \le 8 \\
\mu_{A-B}(u) &= 0 && \text{với } u \le 6 \text{ và } u > 2 \\
&= (u+6)/4 && \text{với } -6 < u \le -2 \\
&= (2-u)/4 && \text{với } -2 < u \le 2 \\
\mu_{A.B}(u) &= 0 && \text{với } u < -5 \text{ và } u > 15
\end{aligned}$$

$= [3 - (4 - u)^{1/2}]/2$ với $-5 \le u < 0$

$= (1 + u)^{1/2}/2$ với $0 \le u < 1/3$

$= [4 - (1 + u)^{1/2}]/2$ với $3 \le u < 15$

$\mu_{A/B}(u) = 0$ với $u < -1$ và $u > 3$

$= (u + 1)/(2 - 2u)$ với $-1 \le u < 0$

$= (5u + 1)/(2u + 2)$ với $0 \le u < 1/3$

$= (3 - u)/(2u + 2)$ với $1/3 \le u < 3$

1.3.11.4. Phương trình số học mờ

Cũng như trong số học, khi chúng ta có số học các số mờ thì chúng ta có thể giải các phương trình số học. Chúng ta sẽ thấy với biểu diễn tập mờ qua họ các tập mức, chứng ta có thể dễ dàng giải các phương trình số học mờ. Chúng ta hãy lấy một ví dụ.

Ví dụ 1.14. Xét phương trình mờ

$$A + X = B \qquad (8^*)$$

trong đó A và B là các tập mờ còn X là tập mờ ẩn. Ta hãy tìm nghiệm X và sẽ chứng tỏ rằng nghiệm $X = A - B$.

Xét tập mức mức $\alpha \in (0, 1]$ của 3 tập mờ trong (8*) và đặt $A_\alpha = [a_{1\alpha}, a_{2\alpha}]$, $B_\alpha = [b_{1\alpha}, b_{2\alpha}]$ và $X_\alpha = [x_{1\alpha}, x_{2\alpha}]$. Rõ ràng ta phải có các ràng buộc sau: $a_{1\alpha} \le a_{2\alpha}$, $b_{1\alpha} \le b_{2\alpha}$ và $x_{1\alpha} \le x_{2\alpha}$. Từ (8*) ta có:

$a_{1\alpha} + x_{1\alpha} = b_{1\alpha}$ và $a_{2\alpha} + x_{2\alpha} = b_{2\alpha}$ và do đó $x_{1\alpha} = b_{1\alpha} - a_{1\alpha}$, $x_{2\alpha} = b_{2\alpha} - a_{2\alpha}$.

Như vậy để cho phương trình mờ (8*) có nghiệm ta phải có giả thiết

(i) $b_{1\alpha} - a_{1\alpha} \le b_{2\alpha} - a_{2\alpha}$, với mọi $\alpha \in (0, 1]$;

Ngoài ra, từ điều kiện đơn điệu giảm của họ X_α, $\alpha < \beta \Rightarrow X_\alpha \supseteq X_\beta$ ta suy ra $x_{1\alpha} \le x_{1\beta} \le x_{2\beta} \le x_{2\alpha}$. Hay, một điều kiện tồn tại nghiệm nữa là

(ii) $\alpha < \beta \Rightarrow b_{1\alpha} - a_{1\alpha} \le b_{1\beta} - a_{1\beta} \le b_{2\beta} - a_{2\beta} \le b_{2\alpha} - a_{2\alpha}$.

Vậy, với điều kiện (i) và (ii), ta có

$$X = \bigcup_{\alpha \in (0,1]} X_\alpha .$$

Một cách tương tự, phương trình mờ

$A \,.\, X = B$ sẽ có nghiệm với hai điều kiện sau:

(i) $b_{1\alpha}/a_{1\alpha} \le b_{2\alpha}/a_{2\alpha}$, với mọi $\alpha \in (0;1]$;

(ii) $\alpha < \beta \Rightarrow b_{1\alpha}/a_{1\alpha} \le b_{1\beta}/a_{1\beta} \le b_{2\beta}/a_{2\beta} \le b_{2\alpha}/a_{2\alpha}$.

1.3.12. Phép toán kết nhập

Một phép toán trên tập mờ có ý nghĩa thực tiễn quan trọng là phép kết nhập (Aggregation Operator). Trong cuộc sống hàng ngày con người thường xuyên phải đánh giá các đối tượng trên cơ sở tổng hợp các đánh giá theo từng tiêu chí đánh giá nào đó. Ví dụ, đánh giá học lực của học sinh hay sinh viên trên cơ sở các điểm đánh giá của các môn học, hay đánh giá các phương án đầu tư một nhà máy trên cơ sở đánh giá hiệu quả kinh tế, xã hội, … đánh giá phương án phát triển sản phẩm của một xí nghiệp theo nhiều tiêu chí khác nhau …

Một cách hình thức, bài toán đặt ra là giả sử có n tiêu chí đánh giá C_i và mỗi tiêu chí được đánh giá bằng các từ ngôn ngữ với ngữ nghĩa biểu thị bằng các tập mờ $A_i^{\sim}$, $i = 1, \ldots, n$. Hãy xây dựng phép tính cho phép "kết nhập" các "điểm" đánh giá $A_i^{\sim}$, $i = 1, \ldots, n$.

Chúng ta sẽ xây dựng một lớp các toán tử như vậy, gọi là phép kết nhập, trên cơ sở các tính chất trực giác quan sát được từ bản chất của việc tích hợp các ý kiến và xem chúng là các tiên đề của phép kết nhập. Ta sẽ thấy sau này là phép kết nhập là một hàm g: $[0, 1]^n \to [0, 1]$. Khi đó việc kết nhập các điểm đánh giá $A_i^{\sim}$ trên không gian $\boldsymbol{U}_i$, $i = 1, \ldots, n$, sẽ là một tập mờ $A^{\sim}$ được xác định bằng phép kết nhập sau:

$$A^{\sim}(u) = g(A_1^{\sim}(u_1), A_2^{\sim}(u_2), \ldots, A_n^{\sim}(u_n)),\ \forall u = (u_1, \ldots, u_n) \in \boldsymbol{U}_1 \times \ldots \times \boldsymbol{U}_n$$

Như vậy, nếu chúng ta có thể phát triển một lý thuyết về các phép kết nhập, thì chúng ta có công cụ kết nhập các ý kiến hoặc các đánh giá theo các tiêu chuẩn khác nhau.

Trước hết, một cách hình thức hóa, phép kết nhập là một hàm 2-ngôi g : $[0;1]^2 \to [0;1]$ có các tính chất sau được coi là các tiên đề:

Tiên đề (Agg1). g có tính chất *kết hợp*, $g(a, g(b, c)) = g(g(a, b), c)$ và do đó ta có thể viết :

$$g(a, b, c) = g(a, g(b, c)).$$

Như vậy, một hàm kết nhập g có thể mở rộng thành hàm n-ngôi

$$g(a_1, a_2, \ldots, a_n):\ [0;1]^n \to [0;1]$$

Tiên đề (Agg2). g là phép toán lũy đẳng (idempotent).

$$g(a, a, \ldots, a) = a,\ \ a \in [0;1]$$

Ý nghĩa thực tiễn của tiên đề này là rõ ràng: nếu các ý kiến là giống nhau, thì kết quả kết nhập phải không thay đổi.

Tiên đề (Agg2*). g thỏa điều kiện biên sau:

$$g(0, 0, \ldots, 0) = 0 \quad \text{và} \quad g(1, 1, \ldots, 1) = 1.$$

Dĩ nhiên, tiên đề này là trường hợp riêng của Tiên đề (Agg2) và do đó nó là một ràng buộc nhẹ hơn khá nhiều Tiên đề (agg2).

Tiên đề (Agg3). g là hàm liên tục.

Đòi hỏi này là tự nhiên trên thực tế: các ý kiến xấp xỉ nhau thì kết quả kết nhập cũng xấp xỉ nhau.

Tiên đề (Agg4). $g(a_1, \ldots, a_n, g(a_1, \ldots, a_n)) = g(a_1, \ldots, a_n)$

Tiên đề (Agg4) mô tả một tính chất thực tế là nếu thêm một ý kiến mới trùng với giá trị kết nhập các ý kiến đã có không làm thay đổi giá trị kết nhập đã có.

Tiên đề (Agg5). *Tính đơn điệu tăng*: Với mọi cặp $(a_1, \ldots, a_n)$ và $(b_1, \ldots, b_n)$ các giá trị trong [0, 1], nếu $a_i \le b_i$, với $i = 1, 2, ..., n$, thì

$$g(a_1, \ldots, a_n) \le g(b_1, \ldots, b_n).$$

Tiên đề (Agg6). *Tính chất giao hoán*: Với bất kỳ một hoán vị vị trí, $\pi : \{1, 2, \ldots, n\} \to \{1, 2, \ldots, n\}$ của các toán hạng của phép kết nhập $g(a_1, \ldots, a_n)$, chúng ta có:

$$g(a_1, a_2, \ldots, a_n) \ = \ g(a_{\pi(1)}, a_{\pi(2)}, \ldots, a_{\pi(n)})$$

Ý nghĩa của Tiên đề (Agg6) là nó mô tả một kiểu tình huống thực tế trong đó thứ tự các ý kiến không quan trọng trong kết nhập. Điều này cũng hợp lý trong một lớp các bài toán ứng dụng. Tuy nhiên, từ một cách nhìn khác, một câu hỏi đặt ra là liệu ta có thể bỏ yêu cầu này không? Câu trả lời là được vì trong việc lấy quyết định tập thể trong thực tiễn nhiều khi ý kiến đầu tiên có ảnh hưởng mạnh đến kết quả của phép kết nhập, hoặc ngược lại, trong các tình huống khác ý kiến về cuối lại có ảnh hưởng mạnh hơn so với các ý kiến đầu. Đối với các loại bài toán này, chúng ta có lý thuyết các phép kết nhập không giao hoán. Tuy nhiên, trong giáo trình này chúng ta không đề cập đến lớp các phép tính này.

Bây giờ ta cho một số ví dụ về phép kết nhập. Do tính chất rất đa dạng của các bài toán ứng dụng, về nguyên tắc chúng ta không nhất thiết đòi hỏi một phép kết nhập nào đó phải thỏa mãn cả 6 tiên đề trên.

1) *Hàm min và max*: Giả sử $g(a_1, a_2) = \text{Min}\,\{a_1, a_2\}$ (hay $g(a_1, a_2) = a_1 \wedge a_2$)

Do tính kết hợp của phép Min, ta có thể dễ dàng mở rộng hàm này thành phép n-ngôi:

$$g(a_1, a_2, \ldots, a_n) = \text{Min}\,\{a_1, a_2, \ldots, a_n\}.$$

Bên cạnh hàm Min, xét hàm Max $h(a_1, a_2) = \text{Max}\,\{a_1, a_2\}$ (hay $h(a_1, a_2) = a_1 \vee a_2$).

Tương tự, do tính kết hợp của phép Max, ta có thể dễ dàng mở rộng hàm này thành hàm n-ngôi:

$$h(a_1, a_2, \ldots, a_n) = \text{Max}\,\{a_1, a_2, \ldots, a_n\}.$$

Dễ dàng kiểm tra hàm Min và Max thỏa tất cả các tiên đề từ (Agg1) – (Agg6).

2) *Trung bình có trọng số*:

$$WAvg(a_1, a_2, \ldots, a_n) = \sum\nolimits_{1\le i\le n} w_i a_i \text{, với } w_i \ge 0 \text{ và } \sum\nolimits_{1\le i\le n} w_i = 1.$$

Chúng ta khảo sát các tính thỏa các tiên đề của phép kết nhập $WAvg$.

(a) Rõ ràng rằng phép $WAvg$ thỏa các tiên đề lũy đẳng, liên tục và đơn điệu tăng.

(b) Bây giờ ta khảo sát tính thỏa Tiên đề (Agg4) của nó.

Định lý 1.2. Cho phép kết nhập có trọng số *wAvg*, nếu nó có *n*-đối số ta sẽ ký hiệu nó là $WAvg(a_1, a_2, \ldots, a_n) = \sum_{1\le i\le n} w_i^n a_i$, với $w_i^n \ge 0$ và $\sum_{1\le i\le n} w_i^n = 1$. Khi đó, *WAvg* thỏa Tiên đề (Agg4) nếu và chỉ nếu

$$w_i^n = \frac{w_i^{n+1}}{\sum_{1\le j\le n} w_j^{n+1}},\ i = 1, \ldots, n \qquad (9^*)$$

Chứng minh: Trước hết ta giả thiết rằng phép kết nhập *Wavg* thỏa Tiện đề (Agg4), ta có đẳng thức $WAvg(a_1, a_2, \ldots, a_n) = WAvg(a_1, a_2, \ldots, a_n, WAvg(a_1, \ldots, a_n))$, hay

$$\sum_{1\le i\le n} w_i^n a_i = \sum_{1\le i\le n} w_i^{n+1} a_i + w_{n+1}^{n+1} \sum_{1\le i\le n} w_i^n a_i = \sum_{1\le i\le n} (w_i^{n+1} + w_{n+1}^{n+1} w_i^n) a_i$$

Từ đẳng thức này ta suy ra $w_i^n = \dfrac{w_i^{n+1}}{1 - w_{n+1}^{n+1}} = \dfrac{w_i^{n+1}}{\sum_{1\le j\le n} w_j^{n+1}}$, ta thu được (9*).

Ngược lại, rất dễ dàng kiểm chứng rằng nếu các trọng số của phép kết nhập *WAvg* có mối liện hệ (9*) thì *WAvg* sẽ thỏa Tiên đề (Agg4).

(c) Có thể kiểm tra rằng *WAvg* không có tính chất kết hợp, không thỏa Tiên đề (Agg1).

(d) *WAvg* không có tính chất giao hoán, không thỏa Tiên đề (Agg6).

Định lý 1.3. Nếu tồn tại hai trong số w_i và w_j của phép kết nhập *WAvg* sao cho $w_i \ne w_j$, thì phép *WAvg* không giao hoán.

Chứng minh: Xét một bộ giá trị $(a_1, a_2, \ldots, a_n)$ sao cho $a_i = 1$, các giá trị khác đều bằng 0, $(0, \ldots, 0, a_i = 1, 0, \ldots, 0)$, và xét một phép hoán vị π hoán vị hai vị trí với chỉ số *i* và *j*, còn các vị trí khác giữ nguyên. Nếu phép *WAvg* có tính giao hoán ta phải có

$$WAvg(a_1, a_2, \ldots, a_n) = \sum_{1\le i\le n} w_i a_i = w_i = WAvg(a_{\pi(1)}, a_{\pi(2)}, \ldots, a_{\pi(n)}) =$$
$$= \sum_{1\le i\le n} w_i a_{\pi(i)} = w_j.$$

Điều này mâu thuân với giả thiết $w_i \ne w_j$.

3) *Phép trung bình cộng số học*

Phép kết nhập trung bình cộng được ký hiệu là *Avg* và được định nghĩa như sau

$$Avg(a_1, a_2, \ldots, a_n) = \sum_{1\le i\le n} \frac{1}{n} a_i .$$

Trong cuộc sống hàng ngày chúng ta thường hay gặp và sử dụng phép kết nhập này để tổng hợp các ý kiến đánh giá hay điểm đánh giá theo các thiêu chuẩn khác nhau. Bay giờ chúng ta khảo sát xem phép kết nhập quen thuộc này sẽ thỏa mãn các tiên đề nào về việc kết nhập.

(a) Rõ ràng là phép *Avg* thỏa các tiên đề về tính lũy đẳng, liên tục, đơn điệu tăng.

(b) Bây giờ ta xem xét tính thỏa của phép *Avg* đối với Tiên đề (Agg4). Ta tính biểu thức

$$Avg(a_1, a_2, \ldots, a_n, Avg(a_1, a_2, \ldots, a_n)) = [\sum_{1\le i\le n} \frac{1}{n+1} a_i] + \frac{1}{n+1} \sum_{1\le i\le n} \frac{1}{n} a_i$$

$$= \sum_{1\le i\le n} (\frac{1}{n+1} + \frac{1}{n(n+1)}) a_i = \sum_{1\le i\le n} \frac{1}{n} a_i = Avg(a_1, a_2, \ldots, a_n).$$

Biểu thức này chứng tỏ rằng phép kết nhập *Avg* thỏa Tiên đề (Agg4).

(c) Tính giao hoán của phép *Avg* có liên hệ chặt chẽ với Định lý 1.2. Cụ thể, ta có định lý sau:

Định lý 1.4.. Phép kết nhập trung bình có trọng số *WAvg* có tính giao hoán thì nó là phép lấy trung bình cộng số học *Avg*.

Chứng minh: Giả sử *WAvg* có tính chất giao hoán, với mọi phép hoán vị vị trí các hạng tử π, ta có

$$WAvg(a_1, a_2, \ldots, a_n) = WAvg(a_{\pi(1)}, a_{\pi(2)}, \ldots, a_{\pi(\mathrm{n})}) \quad (10^*)$$

Xét bộ giá trị (1, 0, …, 0) và phép hoán vị π chỉ đối với hai vị trí thư nhất và vị trí thứ i, các vị trí còn lại giữ nguyên. Thay vào (10*) ta thu được

$$WAvg(a_1, a_2, \ldots, a_n) = w_1 = w_i = WAvg(a_{\pi(1)}, a_{\pi(2)}, \ldots, a_{\pi(\mathrm{n})}).$$

Điều này đúng với mọi chỉ số $i = 1, \ldots, n$. Do $\sum_{1\le i\le n} w_i = 1$, ta suy ra $w_i = 1/n$, với mọi i.

4) *Phép trung bình cộng tổng quát hóa*

Phép kết nhập trung bình cộng tổng quát hóa được xác định bởi công thức sau

$$g_\alpha(a_1, a_2, \ldots, a_n) = \left(\frac{a_1^\alpha + \ldots + a_n^\alpha}{n}\right)^{1/\alpha}$$

với $\alpha \in \boldsymbol{R}$, tập số thực, và $a_i \neq 0$, $i = 1, \ldots, n$, khi $\alpha < 0$.

Dễ dàng kiểm tra thấy rằng phép kết nhập này thỏa các tiên đề lũy đẳng, liên tục, đơn điệu tăng và giao hoán. Tuy nhiên, bằng việc kiểm chứng với các bộ giá trị $(a_1, a_2, \ldots, a_n)$ cụ thể ta có thể chỉ ra rằng, nói chung, nó không thỏa tính chất kết hợp và Tiên đề (Agg4).

Với các giá trị $\alpha \in (-\infty, +\infty)$ khác nhau nó sẽ xác định các phép kết nhập trung bình cộng khác nhau.

- Khi $\alpha \to 0$ thì g_α xác định phép kết nhập trung bình hình học $g_0 = (a_1 . a_2 \ldots a_n)^{1/n}$

Thực vậy, ta tính giới hạn :

$$\lim_{\alpha\to 0} \ln g_\alpha = \lim_{\alpha\to 0} \frac{\ln(a_1^\alpha + \ldots + a_n^\alpha) - \ln n}{\alpha}.$$

Áp dụng quy tắc l'Hospital chuyển về lấy giới hạn theo thương của đạo hàm ta thu được

$$\lim_{\alpha\to 0} \ln g_\alpha = \lim_{\alpha\to 0} \frac{a_1^\alpha \ln a_1 + \ldots + a_n^\alpha \ln a_n}{a_1^\alpha + \ldots + a_n^\alpha} = \frac{\ln a_1 + \ldots + \ln a_n}{n} = \ln(a_1 .a_2 \ldots a_n)^{1/n}.$$

Đó là điều ta cần chứng minh.

- $g_{-\infty}(a_1, a_2, \ldots, a_n) = \text{Min } \{a_1, a_2, \ldots, a_n\}$ và $g_{+\infty}(a_1, a_2, \ldots, a_n) = \text{Max } \{a_1, a_2, \ldots, a_n\}$

Thực vậy, khi $\alpha \to -\infty$, với $a_{\min}$ = Min $\{a_1, a_2, \ldots, a_n\}$ và lưu ý rằng α là số âm, ta có

$$\lim_{\alpha\to-\infty} \ln g_\alpha = \lim_{\alpha\to-\infty} \frac{\ln(a_1^\alpha + \ldots + a_n^\alpha) - \ln n}{\alpha} = \lim_{\alpha\to-\infty} \frac{\ln a_{\min}^\alpha + \ln\left(\left(\frac{a_1}{a_{\min}}\right)^\alpha + \ldots + \left(\frac{a_n}{a_{\min}}\right)^\alpha\right) - \ln n}{\alpha}$$

$$= \ln a_{\min}$$

và ta thu được điều cần chứng minh.

Đối với trường hợp $g_{+\infty}$ việc chứng minh hoàn toàn tương tự.

- Với $\alpha = -1$, ta có $g_{-1}(a_1, a_2, \ldots, a_n) = \frac{n}{a_1^{-1} + \ldots + a_n^{-1}}$, ta có hàm lấy trung bình điều hòa; với $\alpha = +1$, ta có $g_{+1}(a_1, a_2, \ldots, a_n) = \frac{a_1 + \ldots + a_n}{n}$, ta có phép trung bình số học.

5) *Phép trung bình cộng trọng số theo thứ tự* (Phép toán *OWA*)

Trong nhiều công trình nghiên cứu và ứng dụng, người ta thường sử dụng phép kết nhập được gọi là *phép lấy trung bình cộng trọng số theo* (*quan hệ*) *thứ tự* (*ordered weighted averaging operations* (*OWA*)) và ký hiệu là $g_{\boldsymbol{w}}$. Nó được định nghĩa như sau. Cho một vectơ trong số $(w_1, w_2, \ldots, w_n)$, $w_i \in (0, 1]$ và $w_1 + \ldots + w_n = 1$. Khác với phép trung bình cộng có trọng số, ở đây, đối với mỗi bộ giá trị của đối số, $(a_1, a_2, \ldots, a_n)$, trước hết nó được sắp xếp theo quan hệ thứ tự giảm dần, ta thực hiện một hoán vị $(a_{\pi(1)}, a_{\pi(2)}, \ldots, a_{\pi(n)})$ sao cho $a_{\pi(i)}$ là số lớn nhất thứ i trong các giá trị của đối số đã cho. Nói khác đi, $a_{\pi(i)} \geq a_{\pi(j)}$, nếu $i < j$. Khi đó,

$$g_{\boldsymbol{w}} = w_1 a_{\pi(1)} + w_2 a_{\pi(2)} + \ldots + w_n a_{\pi(n)}$$

Ví dụ, cho vectơ trọng số (0,3, 0,1, 0,2, 0,4) và bộ giá trị (0,6, 0,9, 0,2, 0,6). Sắp xếp bộ các giá trị này theo thứ tự giảm, ta thu được 0,9, 0,6, 0,6, 0,2 và do đó, theo định nghĩa trên,

$g_w(0{,}6,\ 0{,}9,\ 0{,}2,\ 0{,}6) = 0{,}3 \times 0{,}9 + 0{,}1 \times 0{,}6 + 0{,}2 \times 0{,}6 + 0{,}4 \times 0{,}2 = 0{,}53$

Dễ dàng kiểm chứng rằng phép *OWA* thỏa các tiên đề lũy đẳng, liên tục, đơn điệu tăng và giao hoán.

Bây giờ ta khảo sát một số trường hợp đặc biệt:

- Với vectơ trọng số $\mathbf{w}_{\min} = (0, 0, \ldots, 0, 1)$ phép *OWA* sẽ trở thành phép Min và với $\mathbf{w}_{\max} = (1, 0, \ldots, 0)$, phép *OWA* sẽ trở thành phép Max.

- $g_{w_{\min}}(a_1,...,a_n) = \text{Min}\ \{a_1, a_2, \ldots, a_n\}$ và $g_{w_{\max}}(a_1,...,a_n) = \text{Max}\ \{a_1, a_2, \ldots, a_n\}$

- Với vectơ trọng số $\mathbf{w} = (1/n, \ldots, 1/n)$, rõ ràng phép *OWA* trở thành phép trung bình số học.

Định lý 1.5. Với mọi phép kết nhập g thỏa tiên đề lũy đẳng và đơn điệu tăng, ta có

$$\text{Min}\ \{a_1, a_2, \ldots, a_n\} \le g(a_1, a_2, \ldots, a_n) \le \text{Max}\ \{a_1, a_2, \ldots, a_n\} \quad (11^*)$$

Ngược lại, nếu hàm g thỏa công thức (11*) thì nó có tính chất lũy đẳng.

Chứng minh: Đặt $a_{min} = \text{Min}\ \{a_1, a_2, \ldots, a_n\}$ và $a_{max} = \text{Max}\ \{a_1, a_2, \ldots, a_n\}$. Khi đó, áp dụng tính lũy đẳng và đơn điệu tăng, ta thu được

$$a_{min} = g(a_{min}, \ldots, a_{min}) \le g(a_1, a_2, \ldots, a_n) \le g(a_{max}, \ldots, a_{max}) = a_{max}$$

Ngược lại, giả sử g thỏa công thức (11*). Khi đó,

$a = \text{Min}\ \{a, a, \ldots, a\} \le g(a, a, \ldots, a) \le \text{Max}\ \{a, a, \ldots, a\} = a$, nghĩa là g có thính chất lũy đẳng.

Bây giờ ta khảo sát một số tính chất hay của các phép kết nhập. Trước hết, ta nhắc lại bài toán về phương trình hàm Cauchy: Tìm họ các hàm số thực thỏa mãn phương trình hàm

$$f(x + y) = f(x) + f(y) \qquad (12^*)$$

Cauchy năm 1821 đã chứng tỏ rằng nếu với giả thiết f liên tục, chỉ có họ hàm có dạng $f(x) = cx$, $c \in \boldsymbol{R}$, là nghiệm của bài toán (12*). Sau đó người ta chứng minh rằng khẳng định này vẫn còn đúng với một trong các điều kiện ràng buộc sau:

- f liên tục tại một điểm nào đó;

- f đơn điệu trên một khoảng nào đó;

- f giới nội trên một khoảng nào đó.

Việc giải bài toán này trong trường hợp biến thực khá phức tạp. Để tham khảo, ta giải bài toán trong trường hợp f là liên tục.

Trước hết ta chứng minh khẳng định khi các biến nhận các giá trị trong tập các số hữu tỷ Q. Thực vậy, với điều kiện (12*) ta sẽ chứng tỏ rằng:

- $f(0) = 0$: Đặt $y = 0$, ta có $f(x + 0) = f(x) + f(0)$ và do đó $f(0) = 0$.

- $f(-x) = -f(x)$: Vì, khi đặt $y = -x$, ta có $f(x + (-x)) = f(x) + f(-x) = f(0) = 0$.

- $f(nx) = nf(x)$: Vì, $f(nx) = f(x + x + \ldots + x)$. Áp dụng hệ thức (12*) $n - 1$ lần ta thu được đẳng thức cần chứng minh.

- $f\left(\dfrac{x}{n}\right) = \dfrac{1}{n} f(x)$: Đặt $y = nx$ ta thu được $f(y) = nf(x) = nf(x) = nf\left(\dfrac{y}{n}\right)$, điều cần chứng minh.

Từ hai khẳng định cuối ta suy ra

$$f\left(m\frac{x}{n}\right) = \frac{m}{n} f(x) \text{ hay } f(xq) = f(x)q \qquad (12^*)$$

Xem q là biến nhận giá trị hữu tỷ, $q \in$ Q, và x là một hằng được chọn là 1, ta có : $f(q) = cq$, $c = f(1)$ và $q \in$ Q.

Vì Q trù mật trong tập các số thực $\boldsymbol{R}$, f là liên tục ta suy ra khẳng định đũng trên trường số thực $\boldsymbol{R}$.

Định lý 1.6. Giả sử hàm g: $[0, 1]^n \to \boldsymbol{R}^+$ thỏa điều kiện biên (Agg2*), tính đơn điệu (Agg5) và tính chất sau:

$$g(a_1 + b_1, \ldots, a_n + b_n) = g(a_1, \ldots, a_n) + g(b_1, \ldots, b_n) \quad (13^*)$$

trong đó $a_i, b_i \in [0, 1]$, $i = 1, \ldots, n$. Khi đó, ta có

$$g(a_1, \ldots, a_n) = \sum_{1\le i\le n} w_i a_i, \text{ với } w_i > 0, i = 1, \ldots, n.$$

Ngoài ra nếu nó thỏa tính lũy đẳng (Agg2) thì g là phép trung bình cộng có trọng số.

Chứng minh: Ta đặt $g_i(a_i) = g(0, ...,0, a_i, 0, ..., 0)$. Khi đó, g_i thỏa hệ thức (12*) và theo lời giải của bài toán phương trình hàm Cauchy với ràng buộc tính đơn điệu, nó phải có dạng $g_i(x) = w_i x$, trong đó $w_i = g_i(1) > 0$. Do vậy, từ giả thiết (13*), ta suy ra

$$g(a_1, \ldots, a_n) = g(a_1, 0, \ldots, 0) + g(0, a_2, \ldots, a_n)$$

và áp dụng tiếp tục như vậy ta thu được

$$g(a_1, \ldots, a_n) = g(a_1, 0, \ldots, 0) + g(0, a_2, 0, \ldots, 0) + \ldots + g(0, \ldots, 0, a_n)$$

$$= g_1(a_1) + g_2(a_2) + \ldots + g_n(a_n) = \sum_{1\le i\le n} w_i a_i$$

Nghĩa là ta thu được điều cần ta chứng minh.

Nếu g thỏa thêm tính chất lũy đẳng, ta có

$$a = g(a, \ldots, a) = a\sum_{1\le i\le n} w_i \text{, hay } \sum_{1\le i\le n} w_i = 1$$

Điều này chứng tỏ g là phép trung bình cộng có trọng số.

1.4. Quan hệ mờ

1.4.1. Khái niệm quan hệ mờ

Định nghĩa 1.7. Cho $\boldsymbol{U}$ là tích Đê-các của n miền cơ sở $\boldsymbol{U}_i$, $i = 1, \ldots, n$. Khi đó, mỗi một tập mờ trên $\boldsymbol{U}$ được gọi là một quan hệ mờ n-ngôi và được ký hiệu là R, gọi là tên của quan hệ đó, và nó được biểu thị bằng công thức sau:

$$R = \int_{U_1\times\ldots\times U_n} \mu(u_1,\ldots,u_n)/(u_1,\ldots,u_n)$$

trong đó $\mu(u_1, \ldots, u_n)$ là hàm thuộc của tập mờ R.

Độ thuộc $\mu(u_1, \ldots, u_n)$ có nghĩa các đối tượng $u_1, \ldots, u_n$ tương ứng của các miền cơ sở $\boldsymbol{U}_1, \ldots, \boldsymbol{U}_n$, có quan hệ R với nhau với độ tin cậy hay độ phù hợp chính là $\mu(u_1, \ldots, u_n)$.

Trong trường hợp R là quan hệ rời rạc thì nó có thể biểu thị bằng một bảng với tên hàng là tên các phần tử trong $\boldsymbol{U}$, còn tên cột là tên các phần tử trong $\boldsymbol{V}$. Trong trường hợp này ta còn nói R được biểu diễn bằng ma trận.

Ví dụ, xét hai miền cơ sở $\boldsymbol{U} = \boldsymbol{V} = \{1, 2, 3, 4\}$. Quan hệ mờ "lớn hơn rất nhiều" giữa các phần tử của $\boldsymbol{U}$ sẽ được biểu thị bằng bảng sau:

Bảng 1.6: Quan hệ mờ

"*Lớn hơn rất nhiều*"

R	1	2	3	4
1	0,0	0,0	0,0	0,0
2	0,3	0,0	0,0	0,0
3	0,8	0,0	0,0	0,0
4	1,0	0,8	0,3	0,0

Mỗi một giá trị độ thuộc trong bảng này, chẳng hạn giá trị 0,8 tại hàng 3 cột 1, có nghĩa cặp giá trị (3, 1) thỏa quan hệ "Lớn hơn rất nhiều" với độ phù hợp là 0,8, hay giá trị 3 *lớn hơn rất nhiều* giá trị 1 với độ phù hợp (với quan hệ *lớn hơn rất nhiều*) là 0,8.

Ta xét một ví dụ khác với $\boldsymbol{U} = \boldsymbol{V} = \boldsymbol{R}$, tập tất cả các số thực. Trên tập số thực này ta có khái niệm trực quan về sự *gần nhau* giữa các số thực. Quan hệ mờ *gần với*, ký hiệu là $R_{gần}$, có thể biểu thị bằng công thức sau:

$$R_{gần} = \int_{U \times V} e^{\frac{-|u-v|}{a}} / (u, v) \, .$$

1.4.2. Quan hệ mờ và tri thức dạng luật nếu-thì

Một dạng biểu diễn tri thức quan trọng trong trí tuệ nhân tạo là tri thức được phát biểu dưới dạng mệnh đề nếu-thì như sau:

"Nếu cường độ dòng điện I là *lớn*, thì vòng quay mô tơ điện N là *nhỏ*" (14*).

Mệnh đề (14*) biểu thị mối quan hệ "mờ" giữa đại lượng cường độ dòng điện và đại lượng số vòng quay của mô tơ điện. Theo nghĩa đó, phải có khả năng biểu diễn (14*) bằng một quan hệ theo Định nghĩa 1.7. Ý tưởng của phương pháp chuyển một mệnh đề *ngôn ngữ* như trên thành một quan hệ mờ được thực hiện như sau:

Giả sử miền cơ sở của biến ngôn ngữ I là $\boldsymbol{U_I} = [0, 10]$ theo đơn vị Ampe và miền cơ sở của biến ngôn ngữ N là $\boldsymbol{V_N} = [400, 2000]$ theo đơn vị vòng/phút. Khái niệm mờ *lớn* của I được biểu thị qua tập mờ với hàm thuộc $\mu_{I\text{-}lớn}$: $[0, 10] \rightarrow [0, 1]$, khái niệm *nhỏ* của N được biểu thị bằng tập mờ với hàm thuộc $\mu_{N\text{-}nhỏ}$: $[400, 2000] \rightarrow [0, 1]$. Khi đó, một ý tưởng trực quan biểu diễn theo từng điểm (u, v) (pointwise) mang tính định lượng của mệnh đề (14*) là:

Nếu $I := \mu_{I\text{-}lớn}(u)$ thì $N := \mu_{N\text{-}nhỏ}(v)$

Hay, một cách hình thức hơn, ta có thể viết

$$I := \mu_{I\text{-}lớn}(u) \Rightarrow N := \mu_{N\text{-}nhỏ}(v) \qquad (15^*)$$

Công thức (15*) cho phép ta nhìn nhận rõ ràng hơn mối quan hệ giữ 2 phần tử $u \in \boldsymbol{U_I}$ và $v \in \boldsymbol{V_N}$. Vấn đề còn lại là từ (15*) ta có thể tính giá trị độ thuộc của cặp phần tử (u, v). Vì hai giá trị $\mu_{I\text{-}lớn}(u)$, $\mu_{N\text{-}nhỏ}(v) \in [0, 1]$ cũng phản ảnh tính đúng đắn của đẳng thức u = *lớn* và v = *nhỏ*, ta có thể xem chúng như giá trị chân lý của một lôgic đa trị trên đoạn [0, 1]. Do đó ngữ nghĩa của (15*) có thể biểu thị bằng

$$\mu_{I\text{-}lớn}(u) \overset{*}{\rightarrow} \mu_{N\text{-}nhỏ}(v) \qquad (16^*)$$

trong đó $\overset{*}{\rightarrow}$ là một phép kéo theo lôgic nào đó của lôgic đa trị và do đó giá trị

$$\mu_{I\text{-}lớn}(u) \overset{*}{\rightarrow} \mu_{N\text{-}nhỏ}(v) \in [0, 1].$$

Một vài ví dụ của phép kéo theo $\overset{*}{\rightarrow}$ thường được sử dụng như:

Kéo theo nhị phân (binary): $s \overset{b}{\rightarrow} t = (1 - s) \vee t$

Kéo theo chuẩn (Standar): $s \overset{S}{\rightarrow} t$ = 1 nếu $s \le t$
= 0 nếu $s > t$

Kéo theo Goedel: $s \overset{g}{\rightarrow} t$ = 1 nếu $s \le t$
= t nếu $s > t$

Kéo theo Madami: $s \overset{m}{\rightarrow} t = s\,.\,t$ (trong đó "." là tích số học) .

Kéo theo Lukasiewicz $s \overset{L}{\rightarrow} t = 1 \wedge (1 - s + t)$

1.4.3. Các phép tính trên quan hệ

Vì quan hệ cũng là tập mờ nên các phép tính trên tập mờ được trình bày trong Mục 1.3 cũng là các phép tính trên quan hệ. Tuy nhiên, trên quan hệ có những phép tính đặc thù riêng mà trên tập mờ nói chung không có, chẳng hạn phép tính hợp thành dưới đây:

Định nghĩa 1.8. Giả sử R là quan hệ mờ trên $\boldsymbol{U}\times\boldsymbol{V}$ và S là quan hệ mờ trên $\boldsymbol{V}\times\boldsymbol{W}$. Khi đó, phép hợp thành của hai quan hệ này là một quan hệ trên $\boldsymbol{U}\times\boldsymbol{W}$, được ký hiệu là $R°S$ và được định nghĩa như sau:

$$R \text{ o } S = \int \vee_{v\in V} [\mu_R(u,v)°\mu_S(v,w)]/(u,w) \qquad (17^*)$$

trong đó ° có thể là một phép tính 2-ngôi trong [0,1] có tính giao hoán, kết hợp và phân phối đối với phép *max* $\vee$. Nếu ° là phép *min* $\wedge$, thì ta có phép hợp thành *max-min*, nếu ° là phép nhân số học "." ta có phép hợp thành *max-product*.

Đối ngẫu với phép hợp thành (17*) là

$$R \text{ o } S = \int \wedge_{v\in V} [\mu_R(u,v) * \mu_S(v,w)]/(u,w) \qquad (18^*)$$

trong đó * là một phép tính đối ngẫu với °. Với * là *max* ta có phép hợp thành *min-max* đối ngấu với phép hợp thành *max-min*, với * là $\oplus$ ta có phép hợp thành *min-sum* đối ngẫu với phép hợp thành *max-product*.

Nếu R và S là các tập mờ rời rạc, tức $\boldsymbol{U}$, $\boldsymbol{V}$ và $\boldsymbol{W}$ là hữu hạn, chẳng hạn $\boldsymbol{U} = \{u_1, \ldots, u_m\}$, $\boldsymbol{V} = \{v_1, \ldots, v_p\}$ và $\boldsymbol{W} = \{w_1, \ldots, w_n\}$. Khi đó, hàm thuộc của tập mờ (18*) tại cặp phần tử (u_i,w_j) có dạng

$$\mu_{RoS}(u_i,w_j) = \vee_{k=1}^{p} [\mu_R(u_i,v_k)°\mu_S(v_k,w_j)] \qquad (19^*)$$

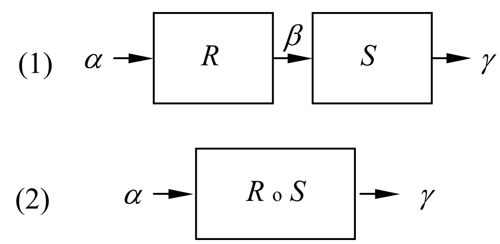
Quan sát công thức (19*) có thể nhận thấy sự "đồng dạng" của nó với biểu thức tính phần tử (i,j) của tích hai ma trận, với tổng ở đây được hiểu là phép max $\vee$, tích được hiểu là phép tính °. Nghĩa là, để tính giá trị $\mu_{RoS}(u_i,w_j)$, ta lấy các phần tử của hàng thứ i của bảng R nhân bằng phép ° với các phần tử tương ứng của cột thứ j của bảng S; lấy tổng bằng phép max $\vee$ các kết qua thu được.

Định lý 1.7. Phép hợp thành được định nghĩa như trong Định nghĩa 1.8 có tính chất kết hợp, nghĩa là, cho các quan hệ mờ R trên không gian $\boldsymbol{U}\times\boldsymbol{V}$, S trên không gian $\boldsymbol{U}\times\boldsymbol{W}$ và Q trên không gian $\boldsymbol{W}\times\boldsymbol{Z}$, chúng ta có đẳng thức sau:

$$(R \text{ o } S) \text{ o } Q = R \text{ o } (S \text{ o } Q) = R \text{ o } S \text{ o } Q \qquad (20^*)$$

Ví dụ 1.15. Xét một quá trình xử lý thông tin như trong hình (1) của Hình 1.9. R và S là các quan hệ mờ biểu diễn các tri thức, chẳng hạn, dưới dạng một tập luật if-then. Ngoài ra, giả sử R và S là các quan hệ mờ với giả thiết rời rạc như được xét ở trên. α là vectơ hàng m-chiều được xem như là dữ liệu đầu vào, β là vectơ hàng p-chiều được xem như kết quả xử lý trung gian, còn $\gamma = (g_1, \ldots, g_n)$ là vectơ output n-chiều biểu thị vectơ đầu ra. Ở đây chúng ta giả thiết rằng quá trình xử lý thông tin tương tác giữa các quá trình thực tế được mô phỏng bằng các phép hợp thành. Nghĩa là, dữ liệu đầu ra được tính theo các công thức sau:

$$\beta = \alpha \text{ o } R, \quad \gamma = \beta \text{ o } S \qquad (21^*)$$



Hình 1.9

Với tính chất của các phép tính nêu trong Định nghĩa 1.8, trong Hình 1.9 quá trình (1) là tương tương với quá trình (2).

Thực vậy, giả sử hàm thuộc của R là $\mu_R(u_i, v_k)$, hàm thuộc của S là $\mu_S(v_k, w_j)$ và $\alpha = (a_1, \ldots, a_m)$. Khi đó, do tính chất kết hợp, ta lần lượt tính theo công thức (21*) như sau:

$$\alpha \text{ o } R = (\vee_{i=1}^{m} [a_i \circ \mu_R(v_i, v_1)], \ldots, \vee_{i=1}^{m} [a_i \circ \mu_R(v_i, v_p)])$$

và thành phần thứ j của vectơ $\gamma = \beta \text{ o } S$ sẽ là:

$$g_j = \vee_{l=1}^{p} \{\vee_{i=1}^{m} [a_i \circ \mu_R(u_i, v_l)] \circ \mu_R(v_l, w_j)\}$$

Theo tính phân phối của phép tính ° đối với phép max, và do tính giao hoán và kết hợp của phép max, chúng ta có

$$g_j = \vee_{l=1}^{p} \vee_{i=1}^{m} [a_i \circ \mu_R(u_i, v_l) \circ \mu_R(v_l, w_j)] = \vee_{i=1}^{m} \vee_{l=1}^{p} [a_i \circ \mu_R(u_i, v_l) \circ \mu_R(v_l, w_j)]$$

$$= \vee_{i=1}^{m} \{a_i \circ \vee_{l=1}^{p} [\mu_R(u_i, v_l) \circ \mu_R(v_l, w_j)]\}$$

Có thể nhận thấy biểu thức $\vee_{l=1}^{p} [\mu_R(u_i, v_l) \circ \mu_R(v_l, w_j)]$ chính là phần tử c_{ij} của ma trận R o S, và do đó

$$\gamma = \alpha \text{ o } (R \text{ o } S) \qquad (22^*)$$

Như vậy, sự tương đương của các công thức (21*) và (22*) đem lại cho ta ý nghĩa thực tiễn của phép hợp thành.

1.4.4. Quan hệ mờ 2-ngôi

Một lớp các quan hệ mờ quan trọng là quan hệ mờ 2-ngôi, chẳng hạn quan hệ "bạn thân", "bạn hàng gần gũi", "học giỏi hơn", ...

Quan hệ 2-ngôi R trên $\boldsymbol{U} \times \boldsymbol{U}$, hay gọi là quan hệ trên không gian $\boldsymbol{U}$, có những tính chất đặc biệt mà các quan hệ khác không có. Trong những quan hệ này có quan hệ đơn vị E được định nghĩa bởi hàm thuộc sau:

$$\mu_E(u, u) = 1, \text{ với } \forall u \in \boldsymbol{U} \text{ và } \mu_R(u, v) = 0, \text{ với } \forall u, v \in \boldsymbol{U}, u \neq v.$$

Định nghĩa 1.9. Cho R là quan hệ mờ 2-ngôi trên $\boldsymbol{U} \times \boldsymbol{U}$. Khi đó ta nói R là:

Phản xạ : nếu và chỉ nếu $\mu_R(u, u) = 1$, với $\forall u \in \boldsymbol{U}$ hay $E \subseteq R$;

Phản phản xạ : nếu và chỉ nếu $\mu_R(u, u) = 0$, với $\forall u \in \boldsymbol{U}$;

Đối xứng : nếu và chỉ nếu $\mu_R(u, v) = \mu_R(v, u)$, với $\forall u, v \in \boldsymbol{U}$;

°-Bắc cầu : nếu và chỉ nếu $\mu_R(u, v) \geq \mu_R(u, w) \circ \mu_R(w, v)$, $\forall u, v, w \in \boldsymbol{U}$;

***-Bắc cầu đối ngẫu*:** nếu và chỉ nếu $\mu_R(u, v) \leq \mu_R(u, w) * \mu_R(w, v)$, $\forall u, v, w \in \boldsymbol{U}$, trong đó * là phép tính đối ngẫu đối với °.

Một quan hệ mờ có cả 3 tính chất phản xạ, đối xứng và °-bắc cầu được gọi là **quan hệ tương tự** (similarity) hay **quan hệ tương đương mờ**.

Một quan hệ mờ có 3 tính chất phản phản xạ, đối xứng, *-bắc cầu đối ngẫu được gọi là **quan hệ tương tự đối ngẫu**.

Ví dụ quan hệ "bạn thân", quan hệ "lớn hơn rất nhiều" là những quan hệ tương đương mờ vì chúng đều là các quan hệ phản xạ, đối xứng và bắc cầu.

Trong thực tế chúng ta dễ dàng xây dựng được quan hệ mờ phản xạ và đối xứng, nhưng khó có thể có ngay tính chất °-bắc cầu. Quan hệ mờ R có 2 tính chất phản xạ và đối xứng được gọi là quan hệ giống nhau (*resemblance*) hay quan hệ gần gũi (*proximity*). Để có được tính °- bắc cầu của quan hệ mờ R ta thực hiện phép lấy °-bắc cầu, ký hiệu là R^ được định nghĩa là nó là quan hệ °-bắc cầu nhỏ nhất chứa R, nghĩa là

$$R^\wedge = \bigcap \{S : S \text{ is } \circ - transitive \ \& \ R \subseteq S\}.$$

Ta sử dụng ký hiệu như sau: $R^2 = R \text{ o } R$; $R^{k+1} = R^k \text{ o } R$, với $k = 1, 2, \ldots$

Định lý 1.8. Giả sử R là quan hệ gần gũi. Khi đó, ta có

(i) $R^\wedge = R \cup R^2 \cup \ldots \cup R^k \cup \ldots = \bigcup_{k=1}^{\infty} R^k$

(ii) Nếu $\boldsymbol{U}$ hữu hạn và có n phần tử, thì ta có $R^\wedge = \bigcup_{k=1}^{n} R^k$

(iii) Nếu $\exists l$, $R^l = R^{l+1}$, thì ta có $R^\wedge = \bigcup_{k=1}^{l} R^k$

(iv) R là tương tự nếu $R^2 \subseteq R$ và khi đó $R^\wedge = R$.

Ví dụ 1.16. Để mô tả sừ gần gũi giữa các giá trị mô tả mầu mắt của con người ta có thể xây dựng một quan hệ mờ gần gũi như sau. Giả sử $\boldsymbol{U}$ = {*đen*, *nâu*, *xanh*, *xanh hơi sẫm*, *nâu đen*, *đen nâu*, *xanh nhạt*}. Có nhiều bài toán thực tế đòi hỏi so sánh, tìm kiếm và chúng ta có thể giải quyết bài toán này dựa trên việc xây dựng quan hệ tương tự, và trước hết xây dựng quan hệ gần gũi. Chẳng hạn quan hệ sau:

Bảng 1.7. Quan hệ giữa các màu

R	*đen*	*nâu đen*	*đen nâu*	*nâu*	*xanh*	*xanh hơi sẫm*	*xanh nhạt*
đen	1,0	0,7	0,85	0,6	0,0	0,3	0,0
nâu đen	0,7	1,0	0,92	0,86	0,0	0,4	0,0
đen nâu	0,85	0,92	1,0	0,82	0,0	0,2	0,0
nâu	0,6	0,86	0,82	1,0	0,0	0,25	0,0
xanh	0,0	0,0	0,0	0,0	1,0	0,9	0,84
xanh hơi sẫm	0,3	0,4	0,2	0,25	0,9	1,0	0,65
xanh nhạt	0,0	0,0	0,0	0,0	0,84	0,65	1,0

Dễ dàng kiểm chứng nó là quan hệ gần gũi: nó là phản xạ vì nó đồng nhất bằng 1 trên đường chéo chính; nó là đối xứng vì các giá trị đối xứng qua đường chéo chính.

1.5. Đại số các tập mờ

Lý thuyết tập mờ là cơ sở toán học cho việc phát triển các phương pháp mô phỏng lập luận của con người. Về nguyên tắc, vấn đề tư duy, lập luận của con người là vấn đề cực kỳ phức tạp và do đó không thể sử dụng một cấu trúc toán học duy nhất để mô phỏng. Vì vậy, mục tiêu của chúng ta là càng xây dựng được nhiều cấu trúc đại số các tập mờ thì càng tốt để chúng ta có thể linh hoạt trong tiếp cận các vấn đề ứng dụng.

1.5.1. T-norm và t-conorm

Trong định nghĩa các phép tính hợp và giao trên tập mờ trong Mục 1.3, chúng ta đã sử dụng hai cặp phép tính 2-ngôi trên [0;1] là cặp min ($\wedge$) và max ($\vee$) và cặp phép tính tích đại số $a.b$ (.) và tổng đại số ($\oplus$) $a \oplus b = a + b - a.b$. Dễ dàng kiểm chứng chúng là những cặp đối ngẫu De Morgan. Bây giờ chúng ta sẽ đưa ra một họ các cặp đối ngẫu t-norm và t-conorm.

Định nghĩa 1.10. Một hàm 2-biến $T : [0;1] \times [0;1] \to [0;1]$ được gọi là phép t-norm nếu nó thỏa các tính chất sau với $\forall\ a, a', b, c \in [0;1]$:

(T1) Tính chất điều kiện biên : $T(a, 1) = a$

(T2) Tính chất giao hoán : $T(a, b) = T(b, a)$

(T3) Tính chất đơn điệu : $a \le a' \Rightarrow T(a, b) \le T(a', b)$

(T4) Tính chất kết hợp : $T(T(a, b), c) = T(a, T(b, c))$

Chúng ta dễ dàng kiểm chứng rằng phép min ($\wedge$) và phép tích đại số (.) là các phép t-norm và chúng được ký hiệu tương ứng là T_m và T_p. Phép t-norm T_m ($\wedge$) được gọi là *phép giao mờ chuẩn* (fuzzy standard intersection).

Một tính chất khá hay của phép toán hai ngôi T nào đó là tính lũy đẳng (idempotency) nói rằng $T(a, a) = a$, với $\forall a \in [0;1]$. Tuy nhiên, sau đây chúng ta chỉ ra một tính chất "độc tôn" của phép giao tiêu chuẩn.

Định lý 1.9. Phép giao tiêu chuẩn là phép t-norm duy nhất có tính chất lũy đẳng.

Chứng minh: Tất nhiên ta thấy T_m có tính chất lũy đẳng, $\min\{a, a\} = a$, với $\forall a \in [0;1]$. Bây giờ ta xét bất kỳ phép t-norm nào mà $T(a, a) = a$, với $\forall a \in [0;1]$. Khi đó, với $\forall a, b \in [0;1]$ và không mất tính tổng quát ta giả sử $a \le b$, theo tính chất đơn điệu và tính chất về điều kiện biên ta có

$a = T(a, a) \le T(a, b) \le T(a, 1) = a$, Do vậy, $T(a, b) = a = \min\{a, b\}$.

Định lý này giải thích lý do tại sao chúng ta không xem tính chất này là tiên đề của phép t-norm.

Một số tính chất quan trọng của phép t-norm mà chúng ta cần đòi hỏi cần phải có trong nhiều ứng dụng, khi cần thiết cũng có thể được coi là các tiên đề, được phát biểu sau đây.

(T5) T là hàm hai biến liên tục (Tính liên tục);

(T6) $T(a, a) < a$ (Tính lũy đẳng dưới (subidempotency));

(T7) $a < a'$ và $b < b' \Rightarrow T(a, a') < T(b, b')$ (Tính đơn điệu chặt).

Ví dụ về những phép t-norm hay được sử dụng là các phép sau:

Phép giao mờ tiêu chuẩn: $T_m(a, b) = \min\{a, b\}$;

Phép tích đại số: $a.b$;

Phép hiệu giới nội: $T(a, b) = \max\{0, a + b - 1\}$;

Phép giao chặt[1]: $T(a, b) = \begin{cases} a & khi \quad b = 1 \\ b & khi \quad a = 1 \\ 0 & khi \quad a \neq 1 \,\&\, b \neq 1 \end{cases}$

Ngoài ra, các phép tính sau cũng là t-norm:

$$T_L(a, b) = \max\{0, a + b - 1\}$$

$$T^*(a, b) = \begin{cases} a & \text{nêu } b = 1 \\ b & \text{nêu } a = 1 \\ 0 & \text{nêu } a \neq 1 \,\&\, b \neq 1 \end{cases}$$

Chúng ta có các bất đẳng thức sau:

$$T^* \leq T_L \leq T_p \leq T_m \qquad (23^*)$$

và, với mọi T-norm T:

$$T^* \leq T \leq T_m \qquad (24^*)$$

Một phép tính "đối ngẫu" với phép t-norm được gọi là phép t-conorm và được định nghĩa như sau,

[1] Phép giao chặt tiếng Anh là *drastic intersection*. Nếu dịch theo nghĩa đen thì gọi là *phép giao mạnh*. Chúng tôi cho rằng chữ "chặt" trong tiếng Việt trong ngữ cảnh này có nghĩa phù hợp hơn.

Định nghĩa 1.11. Một hàm 2-biến $S : [0;1] \times [0;1] \rightarrow [0;1]$ được gọi là phép t-conorm, hay còn gọi là S-norm, nếu nó thỏa các tính chất sau với $\forall\ a, a', b, c \in [0;1]$:

(S1) Tính chất giới nội : $S(a, 0) = a$
(S2) Tính chất giao hoán: $S(a, b) = S(b, a)$
(S3) Tính chất đơn điệu: $a \le a' \ \Rightarrow\ S(a, b) \le S(a', b)$
(S4) Tính chất kết hợp: $S(S(a, b), c) = S(a, S(b, c))$

Như vậy, chỉ có tính chất (T1) và (S1) làm nên sự khác biệt giữa hai họ phép tính T-norm và S-norm.

Dưới đây là một vài S-norm:

$$S_m(a, b) = \max\{a, b\}$$
$$S_L(a, b) = \min\{1, a + b\}$$
$$S_p(a, b) = a + b - a.b$$
$$S^*(a, b) = \begin{cases} a & \text{nếu } b = 0 \\ b & \text{nếu } a = 0 \\ 1 & \text{nếu } a \neq 0 \ \&\ b \neq 0. \end{cases}$$

Về mặt ý nghĩa lôgic, phép T-norm được sử dụng để mở rộng ngữ nghĩa của phép lôgic AND, còn phép S-norm để mở rộng ngữ nghĩa của phép OR.

Bây giờ chúng ta mở rộng ngữ nghĩa của phép phủ định (negation). Giá trị chân lý trong đoạn [0, 1] chúng ta sử dụng phép "1–" để mô ta ngữ nghĩa phép phủ định. Dưới đây, chúng ta sẽ đưa ra một họ phép phủ định như sau:

Định nghĩa 1.12. Hàm $N : [0;1] \rightarrow [0;1]$ được gọi là phép phủ định nếu nó có các tính chất sau, với mọi $\forall\ a, a' \in [0;1]$:

(N1) Tính đơn điệu giảm : $a \le a' \ \Rightarrow\ N(a) \ge N(a')$
(N2) Tính lũy đẳng : $N(N(a)) = a$

Có thể suy ra rằng hàm N trong định nghĩa trên phải là ánh xạ 1-1.

Định nghĩa 1.13. Ba phép tính T-norm T, S-norm S và phép phủ định N được gọi là một hệ đối ngẫu (T, S, N) nếu chúng thỏa đều kiện sau:

$$N(S(a, b)) = T(N(a), N(b)) \qquad (25^*)$$

Chúng ta có thể kiểm chứng các hệ sau là hệ đối ngẫu:

$$(\wedge, \vee, 1\text{-}), (\otimes, \oplus, 1\text{-}), (T_L, S_L, 1\text{-}) \text{ và } (T^*, S^*, 1\text{-})$$

1.5.2. Đại số các tập mờ

Định nghĩa 1.14. Cho một hệ đối ngẫu bất kỳ $\Delta = (T, S, N)$ và không gian cơ sở $\boldsymbol{U}$. Gọi $\boldsymbol{F}(\boldsymbol{U})$ họ tất cả các tập mờ trên $\boldsymbol{U}$. Đại số các tập mờ trên $\boldsymbol{F}(\boldsymbol{U})$ dựa trên hệ đối ngẫu Δ là một cấu trúc $A_\Delta = (\boldsymbol{F}(\boldsymbol{U}), \cap_\Delta, \cup_\Delta, \sim_\Delta)$ với các phép tính được định nghĩa như sau:

- Phép giao $\cap_\Delta$: $A \cap_\Delta B = \int_U T(\mu_{A^\sim}(u), \mu_{B^\sim}(v)/(u,v)$

- Phép hợp $\cup_\Delta$: $A \cup_\Delta B = \int_U S(\mu_{A^\sim}(u), \mu_{B^\sim}(v))/(u,v)$

- Phép bù $\sim_\Delta$: $\sim_\Delta A = \int_U N(\mu_{A^\sim}(u))/u$

Đại số các tập mờ A_Δ có các tính chất sau:

1) Các phép tính $\cap_\Delta$ và $\cup_\Delta$ có tính giao hoán và kết hợp. Chẳng hạn chúng ta chứng tỏ chúng có tính kết hợp. Nhớ rằng phép S-norm S có tính chất kết hợp.

$$\begin{aligned}(A \cup_\Delta B) \cup_\Delta C &= \int_U S(S(\mu_{A^\sim}(u), \mu_{B^\sim}(u)), \mu_{C^\sim}(u))/u \\ &= \int_U S(\mu_{A^\sim}(u), S(\mu_{B^\sim}(u), \mu_{C^\sim}(u)))/u \\ &= A \cup_\Delta (B \cup_\Delta C)\end{aligned}$$

Tương tự, chúng ta có thể kiểm chứng các khẳng định còn lại phát biểu ở trên.

2) Nếu phép tính T-norm T có tính lũy đẳng, $T(a, a) = a$ với $\forall a \in \boldsymbol{U}$, thì phép giao cũng có tính lũy đẳng

$$A \cap_\Delta A = A, \text{ với } \forall A \in \boldsymbol{F}(\boldsymbol{U}, [0;1]).$$

Tương tự, nếu phép S-norm S là lũy đẳng, $S(a, a) = a$ với $\forall a \in \boldsymbol{U}$, thì phép hợp cũng có tính lũy đẳng

$$A \cup_\Delta A = A, \text{ với } \forall A \in \boldsymbol{F}(\boldsymbol{U}, [0;1]).$$

Thực vậy, ta kiểm chứng cho phép giao:

$$A \cap_\Delta A = \int_U T(\mu_{A^\sim}(u), \mu_{A^\sim}(u))/u = \int_U \mu_{A^\sim}(u)/u = A$$

Cũng kiểm chứng tương tự như vậy chúng ta có tính chất sau:

3) Nếu các phép T và S phân phối lẫn nhau thì các phép $\cap_\Delta$ và $\cup_\Delta$ cũng phân phối lẫn nhau

$$(A \cap_\Delta B) \cup_\Delta C = (A \cup_\Delta C) \cap_\Delta (A \cup_\Delta C) \text{ và}$$
$$(A \cup_\Delta B) \cap_\Delta C = (A \cap_\Delta C) \cup_\Delta (A \cap_\Delta C)$$

4) $$A \cap_\Delta \varnothing = \varnothing \quad \text{và} \quad A \cap_\Delta \boldsymbol{U} = A \text{ và}$$
$$A \cup_\Delta \varnothing = A \quad \text{và} \quad A \cup_\Delta \boldsymbol{U} = \boldsymbol{U}$$

5) Tính chất đối ngẫu De Morgan:

$$\sim_\Delta (A \cap_\Delta B) = (\sim_\Delta A) \cup_\Delta (\sim_\Delta B) \text{ và}$$
$$\sim_\Delta (A \cup_\Delta B) = (\sim_\Delta A) \cap_\Delta (\sim_\Delta B)$$

6) Tính chất lũy đẳng: $\sim_\Delta (\sim_\Delta A) = A$

7) Nhìn chung, chúng ta có tính chất sau mà nó rất khác biệt với tập mờ kinh điển:

$$A \cap_\Delta (\sim_\Delta A) \neq \varnothing; \quad A \cup_\Delta (\sim_\Delta A) \neq \boldsymbol{U}$$

1.5.3. Quan hệ giữa đại số tập mờ và đại số các tập kinh điển

Trong Mục 1.1 chúng ta biết rằng ánh xạ h : $A^{\sim} \in \boldsymbol{F}(\boldsymbol{U}) \to \{ A^{\sim}_{\alpha} \in \mathrm{P}(\boldsymbol{U})$: $0 \le \alpha \le 1\}$ thiết lập một song ánh từ tập tất cả các tập mờ $\boldsymbol{F}(\boldsymbol{U})$ vào tập tất cả các tập kinh điển $\mathrm{P}(\boldsymbol{U})$. Điều này gợi ý một hy vọng có một mối liên hệ chặt chẽ và đẹp đẽ giữa khái niệm tập mờ và khái niệm tập kinh điển.

Định lý 1.10. Cho $A^{\sim}_i \in \boldsymbol{F}(\boldsymbol{U})$ với $i \in \boldsymbol{I}$, $\boldsymbol{I}$ là tập chỉ số nào đó. Khi đó,

(i) $\bigcup_{i\in I} A^{\sim}_{i\alpha} \subseteq \left(\bigcup_{i\in I} A^{\sim}_i\right)_\alpha$ và $\bigcap_{i\in I} A^{\sim}_{i\alpha} = \left(\bigcap_{i\in I} A^{\sim}_i\right)_\alpha$

(ii) $\bigcup_{i\in I} A^{\sim}_{i\alpha+} = \left(\bigcup_{i\in I} A^{\sim}_i\right)_{\alpha+}$ và $\bigcap_{i\in I} A^{\sim}_{i\alpha+} \subseteq \left(\bigcap_{i\in I} A^{\sim}_i\right)_{\alpha+}$

Chứng minh: (i) Trước hết ta chứng minh đẳng thức trong (i). Ta thấy, $u \in \bigcap_{i\in I} A^{\sim}_{i\alpha}$ nếu và chỉ nếu với $\forall i \in \boldsymbol{I}$, $u \in A^{\sim}_{i\alpha}$ hay $A^{\sim}_{i\alpha}(u) \ge \alpha$. Điều này tương đương với khẳng định

$$\mathrm{Inf}_{i \in \boldsymbol{I}} A^{\sim}_{i\alpha}(u) \ge \alpha . \qquad (26^*)$$

Theo định ngĩa phép giao các tập mờ, (26*) tương đương với sự kiện $u \in \left(\bigcap_{i\in I} A_i^{\sim}\right)_\alpha$. Như vậy, đẳng thức trong (i) đã được chứng minh.

(ii) được chứng minh một cách tương tự.

Ta có thể chứng tỏ đẳng thức không thể xảy ra đói với các bao hàm thức trong định lý trên. Chẳng hạn, đối với bao hàm thức trong (i), ta xét ví dụ sau.

Giả sử các tập mờ $A_i^{\sim}$ có hàm thuộc cho bởi $A_i^{\sim}(u) = 1 - 1/i$, với mọi $u \in \boldsymbol{U}$ và $i \in \boldsymbol{N}$. Khi đó,

$$\left(\bigcup_{i\in I} A_i^{\sim}\right)(u) = \text{Sup}_{\,i \in \boldsymbol{N}}\, A_i^{\sim}(u) = \text{Sup}_{\,i \in \boldsymbol{N}}\, (1 - 1/i) = 1$$

Do đó, $$\left(\bigcup_{i\in I} A_i^{\sim}\right)_1 = \boldsymbol{U}.$$

Mặt khác, với mọi $i \in \boldsymbol{N}$ và mọi $u \in \boldsymbol{U}$, ta có $A_i^{\sim}(u) = 1 - 1/i < 1$ và do đó,

$$A_{i1}^{\sim} = \varnothing, \text{ với mọi } i \in \boldsymbol{N}$$

Vậy ta suy ra,

$$\bigcup_{i\in I} A_{i1}^{\sim} = \varnothing \neq \boldsymbol{U} = \left(\bigcup_{i\in I} A_i^{\sim}\right)_1.$$

Định lý 1.11. Xét $A^{\sim}, B^{\sim} \in \boldsymbol{F}(\boldsymbol{U})$. Khi đó, ta có

(i) $A^{\sim} \subseteq B^{\sim}$ nếu và chỉ nếu $A_\alpha^{\sim} \subseteq B_\alpha^{\sim}$, với bất kỳ $\alpha \in [0;1]$;

và $A^{\sim} \subseteq B^{\sim}$ nếu và chỉ nếu $A_{\alpha+}^{\sim} \subseteq B_{\alpha+}^{\sim}$, với bất kỳ $\alpha \in [0;1]$.

(ii) $A^{\sim} = B^{\sim}$ nếu và chỉ nếu $A_\alpha^{\sim} = B_\alpha^{\sim}$, với bất kỳ $\alpha \in [0;1]$;

và $A^{\sim} = B^{\sim}$ nếu và chỉ nếu $A_{\alpha+}^{\sim} = B_{\alpha+}^{\sim}$, với bất kỳ $\alpha \in [0;1]$.

Chứng minh: Để chứng minh (i), giả sử $A^{\sim} \subseteq B^{\sim}$, $A^{\sim}(u) \leq B^{\sim}(u)$, với $\forall u \in \boldsymbol{U}$. Điều này kéo theo khẳng định $A_\alpha^{\sim} \subseteq B_\alpha^{\sim}$, với bất kỳ $\alpha \in [0;1]$. Ngược lại, giả sử phản chứng là $A_\alpha^{\sim} \subseteq B_\alpha^{\sim}$, với bất kỳ $\alpha \in [0;1]$ nhưng $A^{\sim} \not\subset B^{\sim}$. Vậy, phải có $u_0 \in \boldsymbol{U}$ sao cho $A^{\sim}(u_0) > B^{\sim}(u_0)$. Lấy α sao cho $A^{\sim}(u) > \alpha > B^{\sim}(u)$. Với α như vậy, ta có $u_0 \in A_\alpha^{\sim}$ nhưng $u_0 \notin B_\alpha^{\sim}$, nghĩa là $A_\alpha^{\sim} \not\subset B_\alpha^{\sim}$, mà điều nàu mâu thuẫn với giả thiết là $A_\alpha^{\sim} \subseteq B_\alpha^{\sim}$, với bất kỳ $\alpha \in [0;1]$.

Một cách hoàn toàn tương tự, chúng ta dễ dàng chứng minh những khẳng định còn lại của định lý.

Định lý 1.12. Cho $A^{\sim} \in \boldsymbol{F}(\boldsymbol{U})$. Khi đó, với mọi $\alpha \in [0;1]$, ta có

(i) $A^{\sim}_{\alpha} = \bigcap_{\beta<\alpha} A^{\sim}_{\beta} = \bigcap_{\beta<\alpha} A^{\sim}_{\beta+}$;

(ii) $A^{\sim}_{\alpha+} = \bigcup_{\alpha<\beta} A^{\sim}_{\beta} = \bigcup_{\alpha<\beta} A^{\sim}_{\beta+}$.

Chưng minh: (i) Với $\beta < \alpha$, và với $\varepsilon > 0$ đủ nhỏ, ta có $\beta < \beta + \varepsilon < \alpha$ và rõ ràng, $A^{\sim}_{\alpha} \subseteq A^{\sim}_{\beta+\varepsilon} \subseteq A^{\sim}_{\beta+} \subseteq A^{\sim}_{\beta}$. Các bao hàm thức này kéo theo hệ thức sau

$$A^{\sim}_{\alpha} \subseteq \bigcap_{\beta<\alpha} A^{\sim}_{\beta} \subseteq \bigcap_{\beta<\alpha} A^{\sim}_{\beta+} .$$

Giả sử $u \in \bigcap_{\beta<\alpha} A^{\sim}_{\beta+}$ nhưng $u \notin A^{\sim}_{\alpha}$. Khi đó, ta có $A^{\sim}(u) < \alpha$ và với ε đủ nhỏ, ta cũng có $A^{\sim}(u) < \alpha - \varepsilon < \beta < \alpha$. Vậy, $u \notin A^{\sim}_{\alpha-\varepsilon}$ và do đó $u \notin A^{\sim}_{\beta+}$. Suy ra, $u \notin \bigcap_{\beta<\alpha} A^{\sim}_{\beta+}$ mâu thuẫn với giả thiết. Điều này chứng tỏ đẳng thức (i) phải xảy ra.

Bằng phương pháp tương tự, chúng ta có thể chứng minh (ii).

Định lý 1.13 (**Định lý phân tích thứ nhất**). Với mỗi $A^{\sim} \in \boldsymbol{F}(\boldsymbol{U})$ ta có công thức biểu diễn tập mờ qua các tập mức sau

$$A^{\sim} = \bigcup_{\alpha\in[0,1]} \alpha A^{\sim}_{\alpha} \qquad (27^*)$$

trong đó α là số vô hướng, $\alpha A^{\sim}_{\alpha}$ là tập mờ có dạng tích của số vô hướng và tập kinh điển và phép hợp vô hạn ở vế phải của (27*) là phép hợp các tập mờ.

Chứng minh: Để chứng minh đẳng thức giữa hai tập mờ trong (27*) ta sẽ chứng tỏ rằng hai hàm thuộc của chúng đồng nhất bằng nhau trên $\boldsymbol{U}$. Thật vậy, xét một phần tử $u \in \boldsymbol{U}$ bất kỳ và đặt $\beta = A^{\sim}(u)$. Theo định nghĩa, ta có đẳng thức sau

$$\left(\bigcup_{\alpha\in[0,1]} \alpha A^{\sim}_{\alpha}\right)(u) = \mathrm{Sup}_{\alpha \in [0,1]}\, \alpha A^{\sim}_{\alpha}(u)$$

$$= \max \{\mathrm{Sup}_{\alpha\in[0,\beta]}\, \alpha A^{\sim}_{\alpha}(u) , \mathrm{Sup}_{\alpha\in (\beta,1]}\, \alpha A^{\sim}_{\alpha}(u) \}. \quad (1)$$

Với $\alpha \in [0;\beta]$, $A^{\sim}(u) = \beta \geq \alpha$ hay $u \in A^{\sim}_{\alpha}$ và do đó $\alpha A^{\sim}_{\alpha}(u) = \alpha$. Với Với $\alpha \in (\beta;1]$, ta có $A^{\sim}(u) = \beta < \alpha$ và do vậy $u \notin A^{\sim}_{\alpha}$ và $\alpha A^{\sim}_{\alpha}(u) = 0$. Từ (1) ta suy ra

$$\left(\bigcup_{\alpha\in[0,1]} \alpha A^{\sim}_{\alpha}\right)(u) = \mathrm{Sup}_{\alpha\in[0,\beta]}\, \alpha A^{\sim}_{\alpha}(u) = \mathrm{Sup}_{\alpha\in[0,\beta]}\ \alpha\ = \beta\ = A^{\sim}(u).$$

Vậy đẳng thức (27*) đã được chứng minh.
Bằng cách chứng minh tương tự ta thu được định lý sau

Định lý 1.14 (**Định lý phân tích thứ hai**). Với mỗi $A^{\sim} \in \boldsymbol{F}(\boldsymbol{U})$ ta có công thức biểu diễn tập mờ qua các tập mức sau

$$A^{\sim}\ = \bigcup_{\alpha\in[0,1]} \alpha A^{\sim}_{\alpha+} \qquad (28^*)$$

trong đó α là số vô hướng và $\alpha A^{\sim}_{\alpha+}$ là tập mờ có dạng tích của số vô hướng và tập kinh điển.

Ký hiệu *Image*($A^{\sim}$) tập ảnh hay tập các giá trị của hàm thuộc $A^{\sim}(.)$. Trong thực tế chúng ta quan tâm đến trường hợp các tập mờ $A^{\sim}$ rời rạc, nghĩa là tập ảnh *Image*($A^{\sim}$) là hữu hạn. Ta chứng tỏ rằng tập chỉ số trong (27*) có thể thay bằng tập ảnh *Image*($A^{\sim}$) $\subseteq$ [0;1].

Định lý 1.15 (**Định lý phân tích thứ ba**). Với mỗi $A^{\sim} \in \boldsymbol{F}(\boldsymbol{U})$ ta có công thức biểu diễn tập mờ qua các tập mức sau

$$A^{\sim}\ = \bigcup_{\alpha\in \mathrm{Im}age(A^{\sim})} \alpha A^{\sim}_{\alpha} \qquad (29^*)$$

trong đó α là số vô hướng và $\alpha A^{\sim}_{\alpha}$ là tập mờ có dạng tích của số vô hướng và tập kinh điển.

Chứng minh: Trong chứng minh Định lý 1.13, ta giả thiết $\beta = A^{\sim}(u)$ hay $\beta \in$ *Image*($A^{\sim}$). Các công thức trong chứng minh đó vẫn đúng khi ta thay tập chỉ số [0;1] bằng tập ảnh *Image*($A^{\sim}$).

Chương 2

LÔGIC MỜ

2.1. Các mệnh đề mờ

Nhìn chung đối tượng nghiên cứu của lôgic là các mệnh đề cùng với giá trị chân lý của chúng. Trong chương này chúng ta nghiên cứu các mệnh đề mờ và việc định giá giá trị chân lý của chúng.

Mệnh đề mờ chứa những khái niệm không chính xác, không chắc chắn và do đó không có đủ thông tin để định giá giá trị chân lý là "tuyệt đối đúng" **I** hay "tuyệt đối sai" **O**, giá trị chân lý *đúng*, *sai* theo nghĩa kinh điển. Vì giá trị chân lý của các mệnh đề mờ có thể nằm trong đoạn [0;1].

Sau đây chúng ta sẽ khảo sát 4 loại mệnh đề mờ và việc định giá giá trị chân lý của chúng.

2.1.1. Mệnh đề mờ không điều kiện và không bị giới hạn

Trước hết ta làm sáng tỏ cụm từ "giới hạn" (qualified). Một mệnh đề bao giờ cũng có giá trị chân lý. Vấn đề là chúng ta có "tuyên bố" một cách rõ ràng giá trị chân lý của nó hay không. Nếu chúng ta tuyên bố rõ giá trị chân lý của nó, tức là chúng ta đã "giới hạn" giá trị chân lý của nó vào một giá trị cụ thể nào đấy, nếu không ta nói mệnh đề đó *không bị giới hạn*. Còn mệnh đề *điều kiện* là mệnh đề nếu-thì, nếu không như vậy mệnh đề đó được gọi là mệnh đề không điều kiện.

Mệnh đề mờ không điều kiện và không giới hạn là mệnh đề dạng sau:

$$p : X \text{ là } A, \qquad (1^*)$$

trong đó X là biến với miền tham chiếu $\boldsymbol{U}$, A là tập mờ trên $\boldsymbol{U}$ biểu thị ngữ nghĩa của giá trị ngôn ngữ như *trẻ*, *rất cao*, *nhanh nhẹn*, … Để đơn giản ta ký hiệu hàm thuộc của tập mờ A là $A(u)$.

Câu hỏi đặt ra là nếu X nhận giá trị cụ thể $u \in \boldsymbol{U}$ thì giá trị chân lý của mệnh đề p được cho bởi (1*) là bao nhiêu. Trong trường hợp cụ thể như vậy, (1*) trở thành

$$u \text{ là } A \qquad (2^*)$$

Như chúng ta đã biết (2*) được hiểu là u là phần tử của tập mờ A với độ thuộc $A(u)$ hay có thể hiểu $A(u)$ là giá trị chân lý của mệnh đề (2*) và ta ký hiệu $tv(p) = A(u), u \in \boldsymbol{U}$. (3*)

Chẳng hạn, ta xét $\boldsymbol{O}$ là một cộng đồng dân cư, biến X chỉ chiều cao của các cá thể trong cộng đồng nhận giá trị trong miền tham chiếu $\boldsymbol{U} = [0, 220]$ tính theo đơn vị *cm* và A là tập mờ biểu thị ngữ nghĩa của từ *cao*, mô tả chiều cao của các cá thể trong cộng đồng. Khi đó, mệnh đề (1*) được cụ thể hóa thành

p : *Chiều cao* (X) là *cao* (A) (4*)

Nếu X nhận giá trị 170 thì giá trị chân lý của mệnh đề (4*) là $tv(p) = 0{,}85 \in [0, 1]$, nếu X nhận giá trị $u \leq 150$ thì $tv(p) = 0{,}0$.

Trong thực tế người ta thường chỉ chiều cao của một đối tượng hay một cá thể cụ thể $o \in \boldsymbol{O}$, và (1*) khi đó được viết cụ thể như sau:

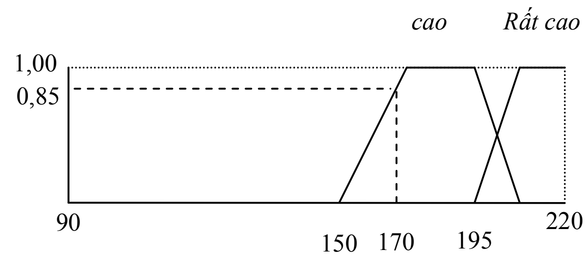
$p : X(o)$ là A

Chẳng hạn, X chỉ biến tuổi *Age* và A là tập mờ biểu thị khái niệm "trẻ" và o là một cá thể thì ta thường viết

p : *Age*(o) là *trẻ*

trong đó *Age*(o) chỉ tuổi tính theo năm của cá thể o. Giá trị chân lý của p khi đó là

$tv(p) = trẻ(Age(o))$.



Hình 2.1: *Tập mờ* cao *chỉ chiều cao của các cá thể*

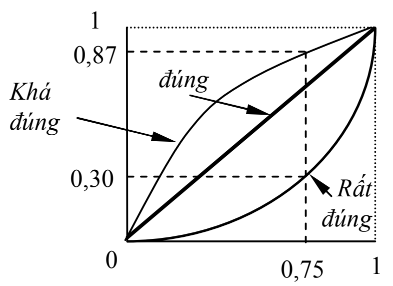
2.1.2. Mệnh đề mờ không điều kiện có giới hạn chân lý (qualified)

Thường một mệnh đề trong cuộc sống thực tiễn hàng ngày của chúng ta đều có một độ tin cậy hay một mức độ đúng hay sai nhất định. Chẳng hạn ta có mệnh đề khẳng định "Ngày mai chắc trời nắng" trong khi hôm nay trời đang u ám. Nếu khẳng định này được một dân làng nói thì độ tin cậy không bằng khẳng định như vậy của cơ quan dự báo thời tiết có uy tín. Một chuyên gia y tế khẳng định "Cháu bé đau ruột thừa" có độ tin cậy hay tính đúng chân lý cao hơn là khẳng định đó được nhận từ một sinh viên y khoa. Như vậy, một nhu cầu tự nhiên là chúng ta cần biểu thị một mệnh đề mờ cùng với giá trị chân lý của nó.

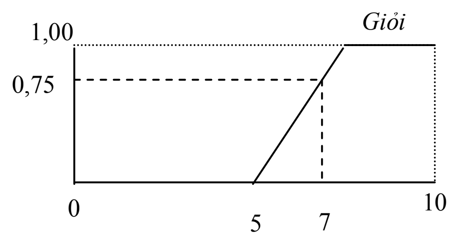
Một mệnh đề mờ không điều kiện và giới hạn được biểu thị ở dạng chuẩn sau p : “X là A” là τ (5*), trong đó X và A là các đại lượng giống như trường hợp trên, còn τ là phép *giới hạn chân lý mờ* (fuzzy truth qualifier) và nó là tập mờ trên tập $\boldsymbol{U}$ = [0;1].

Chẳng hạn, ta lấy ví dụ một mệnh đề dạng (i) “Kết quả học tập của sinh viên Nam là giỏi là *rất đúng*”, hay (ii) “Trình độ đội tuyển Olympic Toán của Việt Nam là giỏi là *khá đúng*”.

Câu hỏi được đặt ra là với một cá thể cụ thể o và một giá trị $u \in \boldsymbol{U}$ của biến cơ sở của A, giá trị chân lý của mệnh đề p ở dạng (5*) là bao nhiêu. Ý tưởng định giá giá trị chân lý này như sau:



Hình 2.2



Hình 2.3: *Tập mờ* giỏi

Xét mệnh đề (i) với khái niệm “giỏi” được biểu diễn bằng tập mờ trong Hình 2.3 và khái niệm chân lý “*rất đúng*” được cho trong Hình 2.2. Giả sử *Kết quả*(Nam) = 7. Khi đó, *tv*(*Kết quả*(Nam) là *giỏi*) = *giỏi*(7) = 0,75. Vì giá trị chân lý của mệnh đề “*Kết quả*(Nam) là *giỏi*” là *rất đúng* với hàm thuộc được cho trong Hình 2.2, nên giá trị chân lý của mệnh đề p sẽ bằng độ thuộc của giá trị 0,75 vào tập mờ biểu diễn khái niệm chân lý *rất đúng*.

$$tv(p) = \textit{rất đúng}(0{,}75) = 0{,}30.$$

Bây giờ chúng ta vẫn xét mệnh đề này với một sự thay đổi giá trị chân lý *rất đúng* của nó thành *khá đúng* ta có mệnh đề “Kết quả học tập của sinh viên Nam là giỏi là *khá đúng*” và ta kí hiệu là mệnh đề p'. Khi đó, ta có

$$tv(p') = \textit{khá đúng}(0{,}75) = 0{,}87.$$

hay $tv(p) < tv(p')$. Chúng ta hãy tự giải thích xem như vậy có hợp lý không?

Trong trường hợp tổng quát, với mệnh đề giới hạn chân lý p trong (5*) và với mỗi phần tử $u \in \boldsymbol{U}$, giá trị chân lý $tv(p)$ của mệnh đề p được định giá bằng công thức

$$tv(p) = \tau(A(u)) \qquad (6^*)$$

Dựa trên (6*), nếu τ là hàm đồng nhất, $\tau(t) = t$, với $t \in [0;1]$, ta sẽ có lại được công thức định giá chân lý (3*) của mệnh đề không giới hạn chân lý. Điều này chỉ ra rằng mệnh đề không giới hạn chân lý có thể xem như là mệnh đề giới hạn chân lý với $\tau = true$ mà hàm thuộc của nó là *hàm đồng nhất*. Lưu ý rằng không phải trong bất kỳ bài toán nào giá trị chân lý ngôn ngữ cũng được biểu thị ngữ nghĩa bằng hàm đồng nhất.

2.1.3. Mệnh đề điều kiện không giới hạn chân lý

Mệnh đề điều kiện không giới hạn chân lý (conditional and unqualified proposition) là mệnh đề có dạng sau

$$p: \text{ Nếu } X \text{ là } A, \text{ thì } Y \text{ là } B \qquad (7^*)$$

trong đó X và Y là các biến nhận các giá trị tương ứng trong miền cơ sở $\boldsymbol{U}$ và $\boldsymbol{V}$, còn A và B là các tập mờ tương ứng trên miền $\boldsymbol{U}$ và $\boldsymbol{V}$.

Như chúng ta đã dề cập trong Mục 1.4 chương 1 về quan hệ mờ và tri thực dạng luật nếu-thì, mệnh đề (7*) xác định một quan hệ mờ R giữa hai đại lượng X và Y . R là tập mờ trên tích Đề-các $\boldsymbol{U} \times \boldsymbol{V}$. Khi đó, (7*) có thể được hiểu là mệnh đề sau:

$$p: (X,Y) \text{ là } R, \qquad (8^*)$$

trong đó, như trong Mục 1.4 chương 1, quan hệ mờ R được xác định qua các tập mờ A và B và một phép kéo theo *Imp*, với $A(u)$ và $B(v)$ là các hàm thuộc tương ứng của A và B, ta có

$$R(u, v) = Imp(A(u), B(v)).$$

Nếu ký hiệu *Imp* là $\overset{*}{\rightarrow}$, thì biểu thức trên có dạng quen nhìn hơn là

$$R(u, v) = A(u) \overset{*}{\rightarrow} B(v).$$

2.1.4. Mệnh đề điều kiện và giới hạn chân lý

Mệnh đề điều kiện có giới hạn chân lý là mệnh đề có dạng sau

$$p: \text{"Nếu } X \text{ là } A, \text{ thì } Y \text{ là } B\text{" là } \tau \qquad (9^*)$$

với τ là giá trị chân lý ngôn ngữ biểu thị bằng hàm thuộc $\tau(t)$, $t \in [0;1]$. (9*) sẽ xác định một quan hệ mờ R^* với hàm thuộc $R^*(u, v)$ được định nghĩa như sau:

Như trên, mệnh đề "Nếu X là A, thì Y là B" sẽ xác định một quan hệ mờ R trên tích Đề-các $\boldsymbol{U} \times \boldsymbol{V}$, với độ thuộc $R(u, v) \in [0;1]$. Vì vậy, chúng ta có thể định nghĩa hàm thuộc

$$R^*(u, v) = \tau(R(u, v)).$$

2.2. Phép kéo theo mờ

Trong Mục 1.4 chương 1 khi đề cập về quan hệ mờ và việc biểu diễn tri thức dạng luật "*Nếu p, thì q*", chúng ta đã thấy mối quan hệ chặt chẽ giữa ngữ nghĩa của mệnh đề mờ dạng nếu-thì và các loại phép kéo theo lôgic $s \to t$, $s, t \in [0;1]$. Có thể với lý do đó, các phép kéo theo như vậy được gọi là các phép kéo theo mờ.

Vì tri thức dạng luật là một yếu tố quan trọng trong biểu diễn tri thức và trong lập luận xấp xỉ, nên việc nghiên cứu các phép kéo theo mờ có vị trí quan trọng. Sau đây chúng ta tìm hiểu một số kiến thức cơ bản về loại phép tính này.

Một cách tổng quát, phép kéo theo mờ là một hàm 2-ngôi $\mathrm{J} : [0;1]^2 \to [0;1]$ với ý nghĩa nói rằng với giá trị chân lý s và t tương ứng của hai mệnh đề p và q, $\mathrm{J}(s, t)$ sẽ cho ta giá trị chân lý của mệnh đề "*Nếu p, thì q*". Nó được xem như là một sự mở rộng của phép kéo theo kinh điển khi hạn chế giá trị chân lý vào tập $\{0, 1\}$. Trong Mục 1.4 chương 1 chúng ta đã đưa ra một số ví dụ về các phép kéo theo này, tuy không gọi là phép kéo theo mờ.

Có nhiều cách tiếp cận để xác định phép kéo theo mờ, mặc dù về nguyên tắc không có một ràng buộc cứng nhắc việc định nghĩa này. Ta sẽ đưa ra một số định nghĩa khác nhau sau để làm ví dụ:

- Định nghĩa dựa trên sự khái quát phép kéo theo 2-trị: Phép kéo theo Kleene-Dienes

$J_b(s, t) = - s \vee t,\ \forall s, t \in [0;1]$ và $-s = 1 - t$, ở đây chỉ số b có nghĩa là *binary*.

- Khi nghiên cứu phương pháp lập luận xấp xỉ để ứng dụng vào điều khiển quá trình tôi luyện thép, Mamdani đã đưa ra định nghĩa sau:

$$J(s, t) = \min(s, t).$$

- Một cách khái quát tương tự, giả sử N là hàm phủ định và S là phép hợp, chẳng hạn S là t-conorm, ta có thể định nghĩa

$$J(s, t) = S(N(s), t).$$

Mặt khác, trong lôgic kinh điển biểu thức Boole $(- s \vee t)$ tương đương với các biểu thức sau:

$$- s \vee (s \wedge t) \quad \text{và} \quad (- s \wedge - t) \vee t$$

và do đó ta có công thức khái quát sang miền trị chân lý [0;1], với T là phép t-norm, như sau:

$$J(s, t) = S(N(s), T(s, t)) \quad \text{và} \quad S(T(N(s), N(t)), t).$$

- Nếu xem tổng đại số như là phép hợp mờ, ta có phép kéo theo Reichenbach: $J_r(s, t) = 1 - s + s.t.$

- Nếu ta chọn phép hợp mờ là phép tổng giới nội $S(s, t) = \min \{1, s + t\}$, ta có phép kéo theo Lukasiewicz: $J_a(s, t) = \min \{1, 1 - s + t\}$.

- Phép kéo theo của Goguen đưa ra năm 1969:

$$J_{Goguen}(s, t) = \min\left\{1, \frac{t}{s}\right\}.$$

- Phép kéo theo Gaines-Rescher

$$J_{g\text{-}r}(s, t) = \begin{cases} 1 & nêu \quad s \le t \\ 0 & nêu \quad s > t \end{cases}$$

- Nếu chọn phép giao chuẩn, phép $\wedge$, ta có phép kéo theo Goedel

$$J_g(s, t) = \sup \{x : s \wedge x \le t\} = \begin{cases} 1 & nêu \quad s \le t \\ t & nêu \quad s > t \end{cases}$$

- Nếu chọn phép giao t-norm $T(s, t) = s.t$, ta có phép kéo theo Goguen

$$J_\Delta(s, t) = \sup \{x : s \,.\, x \le t\} = \begin{cases} 1 & nêu \quad s \le t \\ t/s & nêu \quad s > t \end{cases}$$

- Phép kéo theo Wu

$$J_{wu}(s,\ t) = \begin{cases} 1 & nêu \quad s \le t \\ \min\{1-s, t\} & nêu \quad s > t \end{cases}$$

Tuy nhiên, trong lôgic kinh điển và không kinh điển (chẳng hạn lôgic trực giác) một loại phép kéo theo được định nghĩa thông qua phép hội. Cụ thể, ta có : $J(s, t) = \max \{u \in \{0, 1\} : s \wedge u \le t\}$ (10*)

Có thể thấy phép kéo theo trong lôgic mệnh đề và lôgic vị từ thỏa biểu thức (10*).

Bây giờ ta trình bày phép kéo theo mờ được khái quát hóa từ định nghĩa (10*) và khảo sát các tính chất của loại phép kéo theo này.

2.2.1. Cách tiếp cận qua phép t-norm

Tổng quát hóa công thức (10*) ở trên bằng cách mở rộng giá trị chân lý trong miền 2-trị {0, 1} sang đoạn [0;1] và thay thế phép hội $\wedge$ bằng phép t-norm T, ta có công thức tính sau

$$J_T(s, t) = \sup \{u \in [0, 1] : T(s, u) \le t\} \qquad (11^*)$$

Định lý 2.1. Phép kéo theo J_T có các tính chất sau

(i) $T(s, u) \le t$ nếu và chỉ nếu $J_T(s, t) \ge u$;

(ii) $J_T(J_T(s, t), t) \ge s$;

(iii) $J_T(T(s, t), u) = J_T(s, J_T(t, u))$;

(iv) $s \le t \Rightarrow J_T(s, u) \ge J_T(t, u)$ và $J_T(u, s) \le J_T(u, t)$;

(v) $T(J_T(s, t), J_T(t, u)) \le J_T(s, u)$;

(vi) $J_T(\inf_{j\in J} s_j, t) \ge \sup_{j\in J} J_T(s_j, t)$;

(vii) $J_T(\sup_{j\in J} s_j, t) = \inf_{j\in J} J_T(s_j, t)$;

(viii) $J_T(t, \sup_{j\in J} s_j) \ge \sup_{j\in J} J_T(t, s_j)$;

(ix) $J_T(t, \inf_{j\in J} s_j) = \inf_{j\in J} J_T(t, s_j)$;

(x) $T(s, J_T(s, t)) \le t$

Chứng minh: Chúng ta sẽ chững minh một số tính chất phát biểu trong định lý trên.

(i) Nếu $T(s, u) \leq t$ thì $u \in \{x : T(s, x) \leq t\}$. Do vậy, theo định nghĩa, $u \leq \sup\{x : T(s, x) \leq t\} = J_{\sup}(s, t)$. Ngược lại, nếu $u \leq J_{\sup}(s, t)$ thì, theo tính đơn điệu của phép t-norm,

$$T(s, u) \leq T(s, J_{\sup}(s, t)) = T(s, \sup\{x : T(s, x) \leq t\}).$$

Do tính liên tục của phép t-norm T ta có,

$T(s, \sup\{x : T(s, x) \leq t\}) = \sup\{T(s, x) : x \in \{x : T(s, x) \leq t\}\} \leq t.$

(ii) Ta có, $J_{\sup}(s, t) = \sup\{x : T(s, x) \leq t\}$ và do đó, cùng với tính liên tục của T, ta có : $J_T(J_T(s, t), t) = \sup\{y : T(J_T(s, t), y) \leq t\} = \sup\{y : T(\sup\{x : T(s, x) \leq t\}, y) \leq t\} = \sup\{y : \sup\{T(x, y) : x \in \{x : T(s, x) \leq t\}\} \leq t\}$

(iii) Ta chứng minh $J_T(T(s, t), u) = J_T(s, J_T(t, u))$. Theo định nghĩa

$$J_T(s, J_T(t, u)) = \sup\{x : T(s, x) \leq J_T(t, u)\}$$

Ta thấy, $T(s, x) \leq J_T(t, u) \Leftrightarrow T(t, T(s, x)) \leq u \Leftrightarrow T(T(s, t), x) \leq u$ (do tính kết hợp của T) $\Leftrightarrow x \leq J_T(T(s, t), u)$. Do vậy,

$$J_T(s, J_T(t, u)) = \sup\{x \leq J_T(T(s, t), u)\} = J_T(T(s, t), u).$$

(iv) Giả sử $s \leq t$. Do tính đơn điệu của T, ta có $T(s, x) \leq T(t, x)$ và do đó

$$\sup\{x : T(s, x) \leq u\} \geq \sup\{x : T(t, x) \leq u\} \text{ hay } J_T(s, u) \geq J_T(t, u).$$

Một cách tương tự ta chứng minh bất đẳng thức còn lại.

(v) Theo định nghĩa

$$T(J_T(s, t), J_T(t, u)) = T(\sup\{x : T(s, x) \leq t\}, \sup\{y : T(t, y) \leq u\})$$

Theo tính liên tục của T và do $T(s, x) \leq t$ & $T(t, y) \leq u \Rightarrow T(s, T(x, y)) \leq u$. ta suy ra

$$\begin{aligned} T(J_T(s, t), J_T(t, u)) &= \sup_x\{T(x, \sup\{y : T(t, y) \leq u\}) : T(s, x) \leq t\} \\ &= \sup_x \sup_y\{T(x, y) : T(s, x) \leq t \ \&\ T(t, y) \leq u\} \\ &\leq \sup_x \sup_y\{T(x, y) : T(s, T(x, y)) \leq u\} \\ &= \sup_x \sup_y\{T(x, y) : T(x, y) \leq J_T(s, u)\} \\ &= J_T(s, u). \end{aligned}$$

(vi) Lưu ý rằng $\{x: \inf_{j\in J} T(s_j, x) \leq t\} \supseteq \{x: \sup_{j\in J} T(s_j, x) \leq t\}$. Do đó, ta có $J_T(\inf_{j\in J} s_j, t) = \sup_x\{x: T(\inf_{j\in J} s_j, x) \leq t\} = \sup_x\{x: \inf_{j\in J} T(s_j, x) \leq t\}$ $\geq \sup_x\{x: \sup_{j\in J} T(s_j, x) \leq t\} = \sup_{j\in J} \sup_x\{x: T(s_j, x) \leq t\} = \sup_{j\in J} J_T(s_j, t)$.

(vii) Đặt $s_0 = \sup_{j\in J} a_j$. Khi đó, $s_0 \geq a_j$ và, do tính chất (iv) đã chứng minh, ta có $J_T(s_0, t) \leq J_T(s_j, t)$, với mọi $j\in J$. Do vậy, $J_T(s_0, t) \leq \inf_{j\in J} J_T(s_j, t)$. Mặt khác, do $u = \inf_{j\in J} J_T(s_j, t) \leq J_T(s_i, t)$ với mọi $i\in J$, từ tính chất (i) ta suy ra $T(s_i, \inf_{j\in J} J_T(s_j, t)) \leq t$, với mọi $i\in J$. Vậy, từ tính liên tục của T, ta suy ra $T(s_0, \inf_{j\in J} J_T(s_j, t)) = \sup_{i\in J} T(s_i, \inf_{j\in J} J_T(s_j, t)) \leq t$. Nhờ tính chất (i) ta suy tiếp ra $J_T(s_0, t) \geq \inf_{j\in J} J_T(s_j, t)$. Kết hợp các kết quả lại ta có

$$J_T(\sup_{j\in J} a_j, t) = J_T(s_0, t) = \inf_{j\in J} J_T(s_j, t).$$

(viii) Với lưu ý rằng $\{x : T(t, x) \leq \sup_{j\in J} s_j\} \supseteq \{x : T(t, x) \leq s_j\}$, với mọi $j\in J$, ta có $J_T(t, \sup_{j\in J} s_j) = \sup_x \{x : T(t, x) \leq \sup_{j\in J} s_j\} \geq \sup_x \{x : T(t, x) \leq s_j\} = J_T(t, s_j)$, với mọi $j\in J$. Do vậy, ta rút ra tính chất (viii).

(x) Do tính liên tục của T, ta có

$T(s, J_T(s, t)) = T(s, \sup_x \{x: T(s, x) \leq t\}) = \sup_x \{T(s, x) : T(s, x) \leq t\} \leq t$, đó là điều ta cần chứng minh.

2.2.2 Cách tiếp cận tiên đề

Trong cách tiếp cận tiên đề chúng ta sẽ đưa ra các yêu cầu về tính chất của các phép kéo theo mờ và xem chúng là các tiên đề của phép kéo theo mờ. Bản chất ngư nghĩa kép theo mờ trong ngôn ngữ tự nhiên hay trong lập luận của con người rất phức tạp, khó có một hệ tiên đề chung cho mọi tình huống. Vì vậy, những tiên đề sau đây không nhất thiết bắt buộc mọi phép kéo theo mờ phải thỏa mãn. Chỉ có ứng dụng thực tiễn là tiêu chuẩn cuối cùng chứng minh tính phù hợp của một định nghĩa phép kéo theo mờ. Một số tiên đề là sự khái quát của phép kéo theo kinh điển.

Tiên đề 1. $s \leq s' \Rightarrow J(s, t) \leq J(s', t)$ (Tính đơn điệu tăng đối với biến thứ nhất).

Tiên đề 2. $t \leq t' \Rightarrow J(s, t) \leq J(s, t')$ (Tính đơn điệu tăng đối với biến thứ hai).

Tiên đề 3. $J(0, t) = 1$ (Tính chi phối của giá trị chân lý sai).
Tiên đề này có nghĩa nếu giá trị chân lý của phần tiền tố là sai thì nó chi phối giá trị chân lý của cả mệnh đề nếu-thì.

Tiên đề 4. $J(1, t) = t$ (Tính trung tính của giá trị chân lý đúng).

Điều này nói rằng giá trị chân lý đúng của phần tiền tố không đóng góp được bất kỳ sự thay đổi giá trị chân lý của phần hậu tố còn lại.

Tiên đề 5. J(s, s) = 1 (Tính đồng nhất).

Tiên đề 6. J(s, J(t, u)) = J(t, J(s, u)) (tính chất hoán đổi).

Tiên đề 7. J(s, t) = 1 nếu và chỉ nếu $s \le t$ (Tính chất về điều kiện giới nội).

Điều này nói rằng giá trị chân lý của mệnh đề nếu-thì là đúng nếu và chỉ nếu giá trị chân lý của phần tiền tố bị chặn bởi giá trị chân lý của phần hậu tố.

Tiên đề 8. J(s, t) = J($N(t)$, $N(s)$), trong đó N là hàm phủ định.

Tiên đề 8 là sự khái quát hóa tính chất của phép kéo theo kinh điển nói rằng $p \to q \cong \neg q \to \neg p$.

Tiên đề 9. J là hàm liên tục theo cả hai biến.

Mặc dù, trên một góc nhì nào đó, Tiên đề 9 là một đòi hỏi tự nhiên nhưng nhiều phép kéo theo trong các ví dụ được trình bày ở đầu tiết không thỏa tính liên tục, chẳng hạn phép kéo theo Goedel hay Goguen. Vì vậy, cần nhấn mạnh một lần nữa rằng không nhất thiết bắt buộc mỗi phép kéo theo mờ phải thỏa mãn mọi tiên đề nêu trên, đồng thời ta cũng có quyền đặt ra các yêu cầu về một tính chất nào đó khác mà một phéo kéo theo cần phải có.

Một câu hỏi nẩy sinh là liệu có tồn tại một phép kéo theo mờ thỏa mãn tất cả 9 đòi hỏi trên? Câu trả lời được phát biểu trong mệnh đề sau.

Định lý 2.2. Một hàm 2-biến J : $[0;1]^2 \to [0;1]$ thỏa các Tiên đề 1 – 9 về phép kéo theo mờ *nếu và chỉ nếu* có tồn tại một hàm liên tục đơn điệu tăng thực sự f : $[0;1] \to [0;+\infty)$ sao cho $f(0) = 0$ và J(s, t) $= f^{-1}(f(1) - f(s) + f(t))$, với $\forall s, t \in [0;1]$, và $N(s) = f^{-1}(f(1) - f(s))$, với $\forall s \in [0;1]$.

Trong các ví dụ đã đề cập ở trên, chỉ có phép kêó theo Lukasiewicz thỏa mãn cả 9 tiên đề, với hàm phủ định mờ chuẩn, t.l. $N(s) = 1 - s$, nghĩa là nó thỏa Định lý 2.2. với hàm f là hàm đồng nhất.

Một ví dụ khác về loại phép kéo theo mờ thỏa định lý trên với hàm f được cho như sau:

$f(s) = \ln(1 + s)$ với hàm tựa ngược là $f^{-1}(s) = \begin{cases} e^s - 1 & \text{nêu } 0 \le s \le \ln 2 \\ 1 & \text{nêu } \ln 2 < s \le 1 \end{cases}$

và hàm phủ định đi cùng với f là $N(s) = \dfrac{1-s}{1+s}, \quad s \in [0;1].$

Khi đó phép kéo theo mờ được xác định là

$$J(s, t) = \min\left\{1, \frac{1-s+2t}{1=s}\right\}, \text{ với mọi } s, t \in [0;1].$$

Ta mở rộng ví dụ này bằng việc thay hàm f ở trên bằng hàm $f' = \ln(1 + \lambda s)$, với λ là tham số dương. Khi đó,

$$f^{-1}(s) = \begin{cases} \dfrac{e^s - 1}{\lambda} & \text{nêu } 0 \le s \le \ln(1+\lambda) \\ 1 & \text{nêu } \ln 2 < s \le 1 \end{cases}$$

và hàm phủ định trong trường hợp này là hàm Sugeno

$$N_\lambda(s) = \frac{1-s}{1+\lambda s}, \quad s \in [0;1].$$

Phép kéo theo mờ khí đó được xác định là

$$J_\lambda(s, t) = \min\left\{1, \frac{1-s+t+\lambda t}{1+\lambda s}\right\}.$$

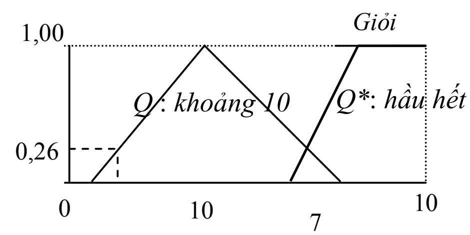
2.3. Lượng từ mờ

Trong lôgic vị từ chúng ta có khái niệm lượng hóa tồn tại và lượng hóa khái quát. Tương tự như vậy, trong lôgic mờ chúng ta cũng có những khái niệm mang hàm ý như vậy như *khoảng 10* học sinh thi tốt nghiệp giỏi; *nhiều hơn nhiều 100* có thể voọc mũi hếch đang sinh sống trong khu bảo tồn quốc gia; *ít nhất là 7* sinh viên đang làm thực tập tốt nghiệp ở Công ty Microsoft Việt Nam; *hầu hết* nữ sinh viên khóa 2005 đều có nguyện vọng theo học khoa CNTT, *khoảng một nửa* số sinh viên trong khóa 2006 là nữ, … Những từ in nghiêng trong các ví dụ trên đều thể hiện ngữ nghĩa không chính xác, mờ về số lượng và được gọi là các lượng hóa mờ.

Theo L.A. Zadeh, có hai loại lượng hóa mờ: (i) *Lượng hóa tuyệt đối* với ngữ nghĩa mờ được ấn định liên quan đến một giá trị (tuyệt đối) cụ thể trong tập các số thực không âm, chẳng hạn, như trong 3 ví dụ đầu nêu trên; (ii)

Lượng hóa tương đối, xác định trên tập [0;1], chỉ *tỷ lệ mờ* số phần tử thỏa một điều kiện hay mệnh đề nào đó, ví dụ như trong hai ví dụ sau cùng đề cập ở trên. Chẳng hạn *hầu hết* chỉ có *một số tỷ lệ mờ* số phận tử thỏa một mệnh đề *mờ* nào đó. Cũng tương tự như vậy ta hiểu cụm từ lượng hóa mờ *khoảng một nửa*.

Lượng hóa tuyệt đối Q được cho bởi một tập mờ trên tập các số thực dương $\boldsymbol{R}^{+}$. Chẳng hạn Q là lượng hóa mờ *khoảng 10* sẽ là tập mờ Q cho trong Hình 2.4. Lượng hóa tương đối được xác định dựa trên tính tỷ số giữa bản số của tập mờ và bản số của tập mờ giới hạn phạm vi các cá thể được đề cập. Ví dụ, trong mệnh đề "*hầu hết* nữ sinh viên khóa 2005 đều có nguyện vọng học khoa CNTT" phạm vi được đề cập là tập các nữ sinh viên khóa 2005. Phạm vi cũng có thể là tập mờ, chẳng hạn trong mệnh đề "*hầu hết* nữ sinh viên học *giỏi* khóa 2005 đều có nguyện vọng học khoa CNTT" phạm vi là tập mờ "các nữ sinh viên học *giỏi* khóa 2005". Khi đó, lượng hóa tương đối được hiểu là một tập mờ trên đoạn [0;1]. Chẳng hạn, tập mờ Q^* trong Hình 2.4 biểu thị phép lượng hóa *hầu hết*.



Hình 2.4: *Tập mờ Q và Q**

Một cách tổng quát, mệnh đề chứa phép lượng hóa mờ bất kỳ Q có dạng cơ bản sau:

$$p : \text{Có } Q \text{ cá thể } o \text{ trong } \boldsymbol{O} \text{ sao cho } X(o) \text{ là } A \quad (12^*),$$

trong đó Q là lượng hóa mờ bất kỳ, X là biến và mỗi cá thể $o \in \boldsymbol{O}$, $X(o)$ nhận giá trị trong miền tham chiếu của tập mờ A.

Ví dụ một mệnh đề như vậy là "Có khoảng 10 sinh viên nói tiếng Anh tốt", trong đó Q là "*khoảng 10*", $\boldsymbol{O}$ là một tập sinh viên trong một lớp học chẳng hạn, X là biến nhận giá trị chỉ trình độ *nói trôi chảy* tiếng Anh của sinh viên trong lớp còn A là tập mờ xác định trên tập các giá trị của biến X biểu thị khái niệm *tốt*.

Để đơn giản hóa cách viết của (12*), nếu ta ký hiệu giá trị chân lý của mệnh đề "cá thể o trong $\boldsymbol{O}$ sao cho $X(o)$ là A" là $P(o)$, $P(o) = A(X(o))$, thì P sẽ là một tập mờ trên $\boldsymbol{O}$ và mệnh đề (12*) trở thành mệnh đề:

$$p' : \text{Có } Q \text{ các cá thể } o \text{ có tính chất } P(o) \qquad (13^*)$$

Nếu P là tập kinh điển thì số lượng các cá thể o thỏa $P(o)$ chính là số lượng các phần tử của tập P. Trong trường hợp như trên P là tập mờ, số lượng của P chính là *bản số* của tập mờ P. Khi đó, giá trị chân lý của mệnh đề (13*) được xác định bởi độ tương thích của bản số C của tập mờ P với phép lượng hóa mờ Q, hay nó được xác định bởi mệnh đề sau, với biến C nhận giá trị trên $\boldsymbol{R}^+$,

$$p' : \text{C} \;\; \text{là} \;\; Q \qquad (14^*)$$

Nếu Q là lượng hóa tuyệt đối, nó là một tập mờ trên $\boldsymbol{R}^+$, thì giá trị C được tính theo bản số vô hướng, tức là $\text{C}(P) = count(P)$ và giá trị chân lý của (14*) hay cũng là của (13*) là $Q(\text{C})$.

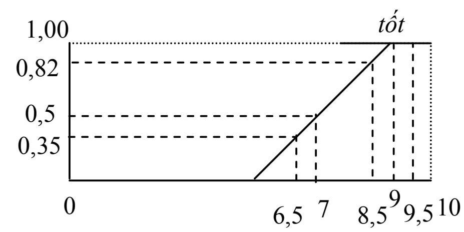
Ví dụ, chúng ta xét mệnh đề p : "Có *khoảng 10* sinh viên nói tiếng Anh *tốt*" với $\boldsymbol{O}$ = {Nam, Hoa, Chính, Hùng, Nga}. X là biến chỉ mức độ *nói tốt* tiếng Anh và giả sử điểm của các sinh viên này được cho như sau: X(Nam) = 6,5; X(Hoa) = 9,0; X(Chính) = 8,5; X(Hùng) = 7,0 và X(Nga) = 9,5. Như vậy, tập mờ P được xác định như sau (xem Hình 2.5)

$$P = \textstyle\sum_{o \in O} \; tốt(X(o))/o = 0{,}35/\text{Nam} + 1{,}0/\text{Hoa} + 0{,}5/\text{Hùng} + 0{,}82/\text{Chính} + 1{,}0/\text{Nga}$$

Khi đó,

$$\text{C}(P) = count(P) = 3{,}67$$

Do đó, gí trị chân lý của mệnh đề p, với Q được cho trong Hình 2.4 sẽ là



Hình 2.5: *Tập mờ "tốt"*

$$tv(p) = Q(3{,}67) = 0{,}26.$$

Một biến tướng của mệnh đề (14*) có dạng sau

$$p : \text{Có } Q \text{ cá thể } o \text{ trong } \boldsymbol{O} \text{ sao cho } X_1(o) \text{ là } A_1 \text{ và } X_2(o) \text{ là } A_2 \quad (15^*)$$

Một ví dụ về mệnh đề dàng này là "Có *khoảng 10* sinh viên nói tiếng Anh *tốt* có dáng người *khá cao*". Như vậy biến X_1 chỉ trình độ nói tiếng Anh nhận giá trị trong miền điểm [0;10], còn biến X_2 chỉ dáng người theo chiều cao trong miền [0;200] tính theo cm. A_1 là tập mờ biểu thị khái niệm *tốt*, A_2 là tập mờ biểu thị khái niệm *khá cao*.

Cũng như trong trường hợp mệnh đề dạng (12*), khi đặt

$$P_1(o) == A_1(X_1(o)) \text{ và } P_2(o) == A_2(X_2(o)) \qquad (16^*)$$

ta có thể chuyển mệnh đề *p* về mệnh đề *p'* sau

p' : Có *Q* các cá thể *o* có tính chất ($P_1(o)$ và $P_2(o)$),

hay, tương tự như mệnh đề (14*), ta thu được dạng mệnh đề của *p'* như sau

p' : C là *Q* (17*), trong đó C là bản số của tập mờ ($P_1(o)$ và $P_2(o)$).

Với *Q* là phép lượng hóa tuyệt đối, C sẽ là bản số vô hướng của tập mờ ($P_1(o)$ và $P_2(o)$) và được tính bằng công thức

$$\mathrm{C}(P_1 \cap P_2) = \sum_{o\in O} \min\{A_1(X_1(o)), A_2(X_2(o))\}$$

và giá trị chân lý của mệnh đề (15*) sẽ là $tv(p) = tv(p') = Q(\mathrm{C}(P_1 \cap P_2))$

Trong trường hợp *Q* là phép lượng hóa tương đối, ta tính tỷ số của các bản số các tập mờ. Ví dụ, đối với mệnh đề "*Hầu hết* sinh viên nói tiếng Anh *tốt* có dáng người *khá cao*" tỷ lệ này sẽ là (lưu ý là P_2 là tập con của tập P_1)

$$\begin{aligned} pr\mathrm{C}(P_1 \cap P_2) &= \mathrm{C}(P_1 \cap P_2)/\mathrm{C}(P_1) \\ &= \left(\sum_{o\in O} \min\{A_1(X_1(o)), A_2(X_2(o))\}\right) / \sum_{o\in O} A_1(X_1(o)) \end{aligned}$$

và giá trị chân lý của mệnh đề (15*) hay (17*) sẽ được tính bằng công thức $Q(pr\mathrm{C}(P_1 \cap P_2))$.

2.4. Lập luận xấp xỉ đơn điều kiện

Thuật ngữ lập luận xấp xỉ được L.A. Zadeh sử dụng lần đầu tiên và được nghiên cứu nhiều tác giả trong và ngoài nước nghiên cứu. Zadeh xuất phát từ ví dụ sau về phương pháp lập luận của con người:

Tiền đề 1: Nếu vỏ của quả cà chua là *đỏ*, thì quả cà chua là *chín*

Tiền đề 2: Vỏ của quả cà chua *c* là *rất đỏ* . (18*)

Kết luận: quả cà chua *c* là *rất chín*

Tiền đề thứ nhất thể hiện tri thức, sự hiểu biết của chúng ta, tiền đề thứ hai là dữ kiện hay sự kiện (fact) và kết luận được rút ra từ hai Tiền đề 1 và 2. (18*) được gọi là một lược đồ lập luận xấp xỉ đơn điều kiện, vì chỉ có một tiền đề có dạng luật nếu-thì.

Chúng ta thường hay gặp kiểu lập luận xấp xỉ như vậy trong suy luận của chúng ta bằng ngôn ngữ tự nhiên. Câu hỏi đặt ra là liệu chúng ta có thể có một cách tiếp cận tính toán để mô phỏng phương pháp lập luận nêu trên?

2.4.1. Quy tắc suy luận hợp thành

Một cách tổng quát, lược đồ lập luận (18*) được biểu thị như sau với *A*, *A'*, *B* và *B'* là các tập mờ tương ứng trên các không gian tham chiếu ***U*** của X và ***V*** của Y,

Tiền đề 1: Nếu *X* là *A*, thì *Y* là *B*
Tiền đề 2: *X* là *A'* . (19*)
Kết luận: *Y* là *B'*

Tiền đề 1 biểu thị mối quan hệ giữa hai đại lượng X và Y, với X nhận giá trị trong ***U*** và Y nhận giá trị trong ***V***. Lược đồ lập luận (19*) được gọi là ***quy tắc cắt đuôi tổng quát hóa*** (generalized modus ponens). Nó khác quy tắc cắt đuôi kinh điển ở chỗ sự kiện "X là *A'*" trong Tiền đề 2 không trùng với sự kiện trong phần "nếu" hay tiền tố của Tiền đề 1.

Chúng ta thiết lập quy tắc suy luận hợp thành để áp dụng vào lược đồ lập luận (19*) dựa trên quan sát các trường hợp sau.

1) Trường hợp X và Y có quan hệ hàm số, tức là $v = f(u)$, $v \in \boldsymbol{V}$ và $u \in \boldsymbol{U}$. Khi đó, nếu ta có sự kiện "X là *u'*" thì ta suy ra $v' = f(u')$, nhờ tri thức X xác định hàm Y. Nếu ta có sự kiện "X là *A'*", trong đó *A'* là tập con của ***U***, thì ta suy ra được tập $B' = \{v' \in \boldsymbol{V}: v' = f(u') \text{ và } u' \in \boldsymbol{U}\} \subseteq \boldsymbol{V}$.

2) Trường hợp X và Y có quan hệ được cho bởi quan hệ 2-ngôi kinh điển $R \subseteq \boldsymbol{U} \times \boldsymbol{V}$. Khi đó, nếu ta có sự kiện "X là *u'*" thì ta suy ra được tập $B = \{v' \in \boldsymbol{V}: (u', v') \in R\}$. Tương tự, nếu ta có sự kiện "X là *A'*", trong đó *A'* là tập con của ***U***, thì ta suy ra được tập

$$B' = \{v' \in \boldsymbol{V}: (u', v') \in R \text{ và } u' \in A'\} \subseteq \boldsymbol{V}$$

Sử dụng thuật ngữ hàm đặc trưng, với $\varphi_{A'}$, $\varphi_{B'}$ và φ_R là các hàm đặc trưng tương ứng của các tập *A'*, *B'* và *R*, công thức tính *B'* ở trên có thể viết dưới dạng sau

$$\varphi_{B'}(v') = \vee_{u' \in \boldsymbol{U}} [\varphi_{A'}(u') \wedge \varphi_R(u', v')], \forall v' \in \boldsymbol{V} \qquad (20^*)$$

3) Trường hợp X và Y có quan hệ được cho bởi *quan hệ mờ 2-ngôi R* trên $\boldsymbol{U} \times \boldsymbol{V}$. Như chúng ta biết, ngữ nghĩa của mệnh đề nếu-thì trong (19*) có thể được biểu thị bằng một quan hệ mờ *R* trên $\boldsymbol{U} \times \boldsymbol{V}$. Nó được xác định dựa trên tập mờ *A* trên $\boldsymbol{U}$ và tập mờ *B* trên $\boldsymbol{V}$, và dựa trên ngữ nghĩa của phép kéo theo mờ đã được nghiên cứu, tức là,

$$R = Impl(A, B) = A \overset{*}{\rightarrow} B, \text{ hay } \mu_R(u, v) = \mathrm{J}(\mu_A(u), \mu_{\boldsymbol{B}}(v)) \qquad (21^*)$$

Sự khác biệt của trường hợp này so với trường hợp đã đề cập trong 2) là thay vì các hàm đặc trưng chúng ta có các hàm thuộc $\mu_{A'}$, $\mu_{B'}$ và μ_R. Vì vậy, nếu ta có sự kiện "X là *A'*" với *A'* là tập mờ trên $\boldsymbol{U}$, thì chúng ta có thể suy luận ra tập mờ *B'* được tính bằng công thức được khái quát hóa từ (20*) như sau:

$$\mu_{B'}(v') = \vee_{u' \in \boldsymbol{U}} [\mu_{A'}(u') \wedge \mu_R(u', v')], \forall v' \in \boldsymbol{V} \qquad (22^*)$$

Như chúng ta đã nghiên cứu trong Mục trước, công thức (22*) có thể được biểu diễn ở dạng ma trận:

$$B' = A' \circ R \qquad (23^*)$$

trong đó o là phép hợp thành max-min (max-min composition). Chính vì *B'* được suy luận ra từ công thức (23*) nên phương pháp lập luận xấp xỉ này được gọi là *quy tắc suy luận hợp thành*.

Nếu ta thay phép min $\wedge$ bằng một phép t-norm *T* nào đó trong (22*) và (23*), ta có quy tắc suy luận hợp thành max-*T* được ký hiệu là $\circ_T$, cụ thể ta có

$$\mu_{B'}(v') = \vee_{u' \in \boldsymbol{U}} T(\mu_{A'}(u'), \mu_R(u', v')), \forall v' \in \boldsymbol{V} \qquad (22^*.1)$$

và

$$B' = A' \overset{T}{\circ} R \qquad (23^*.1)$$

Ví dụ, xét lược đồ suy luận (19*) với $\boldsymbol{U} = \{u_1, u_2, u_3\}$ và $\boldsymbol{V} = \{v_1, v_2\}$, $A = 0{,}5/u_1 + 1{,}0/u_2 + 0{,}6/u_3$ và $B = 1{,}0/v_1 + 0{,}4/v_2$. Cho sự kiện "X là *A'*" với $A' = 0{,}6/u_1 + 0{,}9/u_2 + 0{,}7/u_3$. Chúng ta sẽ suy luận dựa theo quy tắc suy luận cắt đuôi tổng quát và vì vậy trước hết chúng ta tính quan hệ mờ $R = A \overset{*}{\rightarrow} B$ dựa vào phép kéo mờ theo Lukasiewicz $s \overset{L}{\rightarrow} t = 1 \wedge (1 - s + t)$. Như vậy, $\mu_R(u, v) = \mu_A(u) \overset{L}{\rightarrow} \mu_B(v) = 1 \wedge (1 - \mu_A(u) + \mu_B(v))$, $u \in \boldsymbol{U}$ và $v \in \boldsymbol{V}$. Với các dữ liệu của bài toán, quan hệ mờ R có dạng ma trận sau :

$$R = \begin{pmatrix} 1{,}0 & 0{,}9 \\ 1{,}0 & 0{,}4 \\ 1{,}0 & 0{,}8 \end{pmatrix} \text{ và do đó } B' = A' \text{ o } R = (0{,}6\ 0{,}9\ 0{,}7) \circ \begin{pmatrix} 1{,}0 & 0{,}9 \\ 1{,}0 & 0{,}4 \\ 1{,}0 & 0{,}8 \end{pmatrix} = (0{,}9\ 0{,}7)$$

Như vậy, ta suy ra $B' = 0{,}9/v_1 + 0{,}7/v_2$.

Quy tắc suy luận hợp thành cũng có thể ứng dụng cho ***quy tắc modus tollens tổng quát hóa*** có dạng lược đồ lập luận sau:

Tiền đề 1: Nếu *X* là *A*, thì *Y* là *B*
Tiền đề 2: *Y* là *B'*. (24*)
Kết luận: *X* là *A'*

Lưu ý rằng nói chung $B' \neq B$. Khác với quan hệ hàm số, quan hệ mờ R có tính đối xứng giữa hai biến X và Y, cho nên sử dụng phép hợp thành trên các quan hệ mờ, việc suy luận ra A' có thể được tính theo công thức sau với B' là vectơ cột : $A' = R \text{ o } B'$ (25*)

Chúng ta xét một ví dụ với các dữ kiện giống như trong ví dụ vừa xem xét ở trên, trừ việc ta không có sự kiện "X là *A'*" mà ở đây ta lại so sự kiện "Y là *B'*" với B' được cho là $B' = 0{,}9/v_1 + 0{,}7/v_2$, nghĩa là nó chính là kết luận trong ví dụ trên. Khi đó, quan hệ mờ R vẫn như đã được tính trong ví dụ trên và kết luận A' được tính theo (25*) như sau:

$$A' = R \text{ o } B' = \begin{pmatrix} 1{,}0 & 0{,}9 \\ 1{,}0 & 0{,}4 \\ 1{,}0 & 0{,}8 \end{pmatrix} \circ \begin{pmatrix} 0{,}9 \\ 0{,}7 \end{pmatrix} = (0{,}9\ 0{,}9\ 0{,}9)$$

Như vậy, ta đa suy ra được kết luận $A' = 0{,}9/u_1 + 0{,}9/u_2 + 0{,}9/u_3$.

Như chúng ta thấy, phép kép theo có vị trí quan trọng trong lập luận. Trong môi trường thông tin không chắc chắn, chúng ta có nhiều cách biểu diễn ngữ nghĩa của phép kép theo. Trong Mục trước chúng ta đã nghiên cứu về phép kéo theo mờ để làm cơ sở định nghĩa ngữ nghĩa của các mệnh đề điều kiện nếu-thì hay các luật mờ và chúng ta đã liệt kê một danh sách các phép kéo theo. Để dễ đáp ứng với thực tiễn đa dạng và phức tạp, về nguyên tắc, danh sách như vậy càng dài càng tốt. Vì vậy, sau đây chúng ta trình bày một số ý tưởng định nghĩa các phép kéo theo mờ để "thâu tóm" ngữ nghĩa của luật.

Để cho gọn và dễ hiểu, trước hết chúng ta trình bày về việc biểu diễn ngữ nghĩa của luật mờ bằng biểu thức của đại số của quan hệ mờ nêu trong Bảng 2.1 dưới đây, trong đó một số phép kéo theo đã liệt kê trong Mục trước. Tuy nhiên, cần lưu ý là, do tính phong phú của các biểu thức giải tích, không phải biểu thức giải tích nào của phép kéo theo cũng viết được dưới dạng biểu thức đại số tập hợp.

Bảng 2.1. *Biểu thức đại số của quan hệ biểu thị ngữ nghĩa của luật*

Phép kéo theo	**Biểu thức giải tích** $J_b(A(u), B(v))$	**Biểu thức đại số của** R
Kleen-Dienes	$\max[1 - A(u), B(v)]$	$CA \times \boldsymbol{V} \cup \boldsymbol{U} \times B$
Mamdani	$\min(A(u), B(v))$	$(A \times \boldsymbol{V}) \cap (\boldsymbol{U} \times B)$
Zadeh 1973	$\max[\min(A(u), B(v)), 1 - A(u)]$	$((A \times \boldsymbol{V}) \cap (\boldsymbol{U} \times B)) \cup CA$
Reichenbach	$1 - A(u) + A(u)B(v)$	$CA \oplus B$

Bây giờ ta trình bày một số ý tưởng trực quan về sự "thâu tóm" ngữ nghĩa của phép kéo theo trong ngôn ngữ tự nhiên.

(1) Khi ta khẳng định A là đúng thì mệnh đề phủ định $\neg A$ cũng cung cấp một lượng thông tin nhất định.

Trong lôgic kinh điển điều này là hiển nhiên, hay từ giá trị chân lý của A ta suy ra giá trị chân lý của $\neg A$. Điều này không luôn luôn đúng trong lôgic mờ. Tuy nhiên, chúng ta có thể tận dụng ý nghĩa trực quan này để bổ sung vào định nghĩa của phép kéo theo hay quan hệ mờ. Chẳng hạn ta có thể định nghĩa quan hệ mờ R như sau:

- Quan hệ R dựa trên phép kéo theo Mamdani không có thông tin liên quan đến $\neg A$, khi thêm thông tin này ta có quan hệ được Zadeh định nghĩa năm 1973 cho trong Bảng 2.1.

- Quan hệ R xác định dựa trên phép kéo theo "tích" được cho như sau

$$R(u, v) = \mathrm{J}_{product}(A(u), B(v)) = A(u)\,.\,B(v).$$

Để bổ sung thông tin liên quan đến $\neg A$ ta có thể thiết lập quan hệ sau:

$$R(u, v) = \max\,[A(u)\,.\,B(v),\ 1 - A(u)]$$

(2) Một khẳng định $A \to B$ bao giờ cũng cho ta một thông tin nào đó về khẳng định $\neg A \to \neg B$. Chẳng hạn, khi ta khẳng định "Người *trẻ* thì chạy *nhanh*" thường kéo theo một khẳng định "Người *già* thì chạy *chậm*" ở mức độ tin cậy nào đó. Nhận xét trực quan này gợi ý cho ta một cách bổ sung thông tin vào định nghĩa phép kéo theo như sau: Nếu R được định ngĩa dựa trên một phép kéo theo $\mathrm{J}(A(u), B(v))$ nào đó, thì ta có thể sinh một định nghĩa khác như sau:

$$R(u, v) = \max\,[\mathrm{J}(A(u), B(v)),\ \mathrm{J}((\neg A)(u), (\neg B)(v))],$$

hay, ở dạng biểu thức đại số, trong đó $\xrightarrow{\Im}$ ký hiệu phép kéo theo J,

$$R = (A \times \boldsymbol{V} \xrightarrow{\Im} \boldsymbol{U} \times B) \cup (\neg A \times \boldsymbol{V}) \xrightarrow{\Im} (\boldsymbol{U} \times \neg B).$$

Chẳng hạn, nếu J là phép kéo theo Goedel, hay $R_g = A \times \boldsymbol{V} \xrightarrow{g} \boldsymbol{U} \times B$, thì ta có một định nghĩa khác là

$$R_{gg} = (A \times \boldsymbol{V} \xrightarrow{g} \boldsymbol{U} \times B) \cup (\neg A \times \boldsymbol{V}) \xrightarrow{g} (\boldsymbol{U} \times \neg B);$$

Nếu J là phép kéo theo Zadeh, hay $R_z = A \times \boldsymbol{V} \xrightarrow{z} \boldsymbol{U} \times B$, thì ta có một định nghĩa khác là

$$R_{zz} = (A \times \boldsymbol{V} \xrightarrow{z} \boldsymbol{U} \times B) \cup (\neg A \times \boldsymbol{V}) \xrightarrow{z} (\boldsymbol{U} \times \neg B).$$

Cũng với ý tưởng trực quan này nhưng không nhất thiết hai phép kéo theo là giống nhau, chẳng hạn ta có thể định nghĩa quan hệ sau để biểu diễn luật:

$$R_{zg} = (A \times \boldsymbol{V} \overset{z}{\rightarrow} \boldsymbol{U} \times B) \cup (\neg A \times \boldsymbol{V}) \overset{g}{\rightarrow} (\boldsymbol{U} \times \neg B),$$

hay

$$R_{gz} = (A \times \boldsymbol{V} \overset{g}{\rightarrow} \boldsymbol{U} \times B) \cup (\neg A \times \boldsymbol{V}) \overset{z}{\rightarrow} (\boldsymbol{U} \times \neg B).$$

2.4.2. Việc lựa chọn phép kéo theo mờ cho phương pháp lập luận xấp xỉ

2.4.2.1. Đối với quy tắc cắt đuối tổng quát hóa

Để thấy rõ vai trò của phép kéo theo mờ, dựa vào (22*.1) công thức (23*.1) có thể viết cụ thể như sau, trong đó $B'(v)$, $A(u)$ và $R(u, v)$ là các hàm thuộc tương ứng của các tập mờ B', A và R,

$$B'(v) = \vee_{u' \in \boldsymbol{U}} T[A'(u'), \mathrm{J}(A(u'), B(v))], \forall v \in \boldsymbol{V} \qquad (26^*)$$

Một câu hỏi đặt ra là một phương pháp lập luận khi nào được xem là tốt hay phù hợp. Một tiêu chuẩn đánh giá mức độ phù hợp là khi quay trở về lập luận kinh điển, tức là khi $A' = A$ thì ta cần có $B' = B$, hay ta cần có

$$B(v) = \vee_{u' \in \boldsymbol{U}} T[A(u'), \mathrm{J}(A(u'), B(v))], \forall v \in \boldsymbol{V} \quad (27^*)$$

Để trả lời cho câu hỏi này, ta có định lý sau

Định lý 2.3. Giả sử rằng phép t-norm T là hàm *liên tục*, phép kéo theo mờ được chọn là phép J_T, t.l. $\mathrm{J}_T(s, t) = \sup_u \{u : T(s, u) \le t\}$. Khi đó, nếu A là tập mờ chuẩn thì phương pháp lập luận xấp xỉ thỏa điều kiện (27*).

Chứng minh: Theo định nghĩa của phép $\mathrm{J}_T(s, t)$, ta có $T(s, \mathrm{J}_T(s, t)) \le t$. Với $s = A(u)$ và $t = B(v)$ ta thu được biểu thức $T(A(u), \mathrm{J}_T(A(u), B(v))) \le B(v)$, với mọi $u \in \boldsymbol{U}$ và $v \in \boldsymbol{V}$. Mặt khác, do A là tập mờ chuẩn, t.l. tồn tại $u_0 \in \boldsymbol{U}$ sao cho $A(u_0) = 1$, ta suy ra $T(A(u_0), \mathrm{J}_T(A(u_0), B(v))) = \mathrm{J}_T(1, B(v))) = B(v)$ và điều này chứng tỏ phép kéo theo mờ J_T thỏa (27*).

Định lý 2.4. Nếu tập mờ A có miền trị phủ toàn đoạn [0,1], thì các phép kéo theo mờ sau thỏa điều kiện (27*) đối với bất kỳ phép t-norm T nào:

(i) Phép kéo theo Gaines-Rescher J_{g-r};

(ii) Phép kéo theo Goedel J_g;

(iii) Phép kéo theo Wu J_{wu}.

Chứng minh: Trước hết chúng ta chứng minh trường hợp khó hơn trước.

(iii) Với mỗi $v \in \boldsymbol{V}$, ta tính biểu thức sau và nhớ rằng miền trị của A phủ toàn bộ đoạn [0;1]:

$$\sup_{u\in \boldsymbol{U}} T(A(u), J_{wu}(A(u), B(v))) = \sup_{s\in[0,1]} T(s, J_{wu}(s, B(v)))$$
$$= \max\{\sup_{s\le B(v)} T(s, J_{wu}(s, B(v))), \sup_{s>B(v)} T(s, J_{wu}(s, B(v)))\}$$
$$= \max\{\sup_{s\le B(v)} T(s, 1), \sup_{s>B(v)} T(s, \min(1-s, B(v))\}$$
$$= \max\{B(v), \sup_{s>B(v)} T(s, \min(1-s, B(v))\} = B(v),$$

vì, do T đơn điệu tăng theo từng biến, $T(s, \min(1-s, B(v)) \le T(1, B(v)) = B(v)$. Đều này nói rằng (27*) đúng đối với phép kéo theo mờ Wu.

Đối với trường hợp (i) và (ii) ta chứng minh hoàn toàn tương tự nhưng việc tính toán đơn giản hơn. Chẳng hạn, đối với trường hợp (i), ta thấy

$$\sup_{u\in \boldsymbol{U}} T(A(u), J_{g-r}(A(u), B(v))) = \sup_{s\in[0,1]} T(s, J_{g-r}(s, B(v)))$$
$$= \max\{\sup_{s\le B(v)} T(s, J_{g-r}(s, B(v))), \sup_{s>B(v)} T(s, J_{g-r}(s, B(v)))\}$$
$$= \max\{\sup_{s\le B(v)} T(s, 1), \sup_{s>B(v)} T(s, 0)\}$$
$$= \max\{B(v), 0\} = B(v), \text{ vì, } T(s, 0) \le T(1, 0) = 0.$$

Để dễ so sánh, chúng ta cho các kết quả nghiên cứu về vấn đề này đối với các phép kéo theo đã đề cập ở trên trong Bảng 2.1.

Bảng 2.2. *Quy tắc cắt đuôi tổng quát hóa*

Tên phép kéo theo	**t-norm min**	**t-norm tích đại số**	**t-norm hiệu giới nội**	**t-norm giao chặt**
Gaines-Rescher	B	B	B	B
Goedel (J_g)	B	B	B	B
Goguen (J_Δ)	$B^{1/2}$	B	B	B
Kleene-Dienes	$\max\{1/2, B\}$	$\max\{1/4, B\}$	B	B

Lukasiewicz (J_a)	$\frac{1}{2}(1+B)$	$\frac{1}{4}(1+B)^2$	B	B
Reichenbach (J_r)	$\frac{1}{2-B}$	$\max\{B, \frac{1}{4-4\min(B,1/2)}\}$	B	B
Wu (J_{wu})	B	B	B	B

2.4.2.2. Đối với quy tắc modus tollens tổng quát hóa

Tương tự như đối với trường hợp nghiên cứu về việc lựa chọn phép kéo theo mờ đối với phương pháp lập luận xấp xỉ dựa trên quy tắc suy luận cắt đuối tổng quát hóa ở trên, một tiêu chuẩn lựa chọn phép kéo theo mờ là khi sự kiện đầu vào $B' := \neg B$ thì kết luận hay đầu ra của quy tắc suy luận phải là $A' = \neg A$[1], hay chúng ta phải có:

$$N(A(u)) = \sup_{v\in V} T(N(B(v)), J(A(u), B(v))) \qquad (28^*)$$

Tương tự như việc nghiên cứu đối với quy tắc cắt đuôi tổng quát hóa ở trên, kết quả lập luận A' khi sử dụng quy tắc modus tollens tổng quát hóa với giá trị đầu vào $B' = B$ được cho trong Bảng 2.2.

2.4.2.3. Xây dựng phương pháp lập luận dựa trên phương trình quan hệ mờ

Trong hai Mục trước chúng ta đã trình tiêu chuẩn lựa chọn phép kéo theo mờ J để xác định quan hệ R sao cho nó thỏa biểu thức

$$B = A \circ R \qquad (29^*)$$

đối với quy tắc cắt đuôi tổng quát hóa, và thỏa biểu thức

$$N(A) = R \circ N(B) \qquad (30^*)$$

đối với quy tắc modus tollens tổng quát hóa.

Như vậy, bản chất của việc tìm một phương pháp lập luận xấp xỉ là việc xác định quan hệ mờ R một cách phù hợp. Tuy nhiên, khi quan sát hai biểu thức (29*) và (30*), chúng ta có thể coi chúng như là các phương trình quan

[1] Xem chú thích ngay trước.

hệ mờ và bài toán xây dựng một phương pháp lập luận xấp xỉ trở thành việc giải phương trình quan hệ mờ (29*) hay (30*) đến tìm lời giải R khi cho biết các "quan hệ mờ" A và B.

Bây giờ chúng ta đi nghiên cứu một số phương pháp giải các phương trình quan hệ ở hai dạng nếu trên.

Bảng 2.3. *Quy tắc modus tollens tổng quát hóa*

Tên phép kéo theo	**t-norm min**	**t-norm tích đại số**	**t-norm hiệu giới nội**	**t-norm giao chặt**
Gaines-Rescher	$\neg A$	$\neg A$	$\neg A$	$\neg A$
Goedel (J_g)	$\max\{1/2, \neg A\}$	$\max\{1/4, \neg A\}$	$\neg A$	$\neg A$
Goguen (J_Δ)	$\frac{1}{1-A}$	$\max\{1/(4A), \neg A\}$	$\neg A$	$\neg A$
Kleene-Dienes	$\max\{1/2, \neg A\}$	$\max\{1/4, \neg A\}$	$\neg A$	$\neg A$
Lukasiewicz (J_a)	$(1-\frac{1}{2}A)$	$\frac{1}{4}(A-2)^2$	$\neg A$	$\neg A$
Reichenbach (J_r)	$\frac{1}{1+A}$	$\begin{cases} \frac{1}{4A(u)} & A(u) \geq \frac{1}{2} \\ (\neg A)(u) & A(u) < \frac{1}{2} \end{cases}$	$\neg A$	$\neg A$
Wu (J_{wu})	$\neg A$	$\neg A$	$\neg A$	$\neg A$

1) *Phương trình quan hệ mờ*

Cho các quan hệ mờ 2-ngôi $P(u, v)$, $Q(v, w)$ và $R(u, w)$, với $u \in \boldsymbol{U}$, $v \in \boldsymbol{V}$ và $w \in \boldsymbol{W}$. Đối với việc nghiên cứu phương trình quan hệ, chúng ta giới hạn việc xét các quan hệ mờ rời rạc, $\boldsymbol{U}$, $\boldsymbol{V}$ và $\boldsymbol{W}$ là các tập hữu hạn

$\boldsymbol{U} = \{u_i : i = 1, \ldots, n\}$, $\boldsymbol{V} = \{v_j : j = 1, \ldots, m\}$ và $\boldsymbol{W} = \{w_k : k = 1, \ldots, l\}$

Khi đó các quan hệ mờ có thể biểu thị ở dạng ma trận.

Xét phương trình quan hệ mờ

$$R = P \overset{T}{\circ} Q \qquad (31^*)$$

Giả sử rằng các quan hệ R và Q là các dữ kiện cho trước. Bài toán đặt ra là tìm quan hệ mờ P sao cho nó thỏa phương trình quan hệ (31*). Vì phép $\overset{T}{\circ}$ không giao hoán, một bài toán tương tự là, cho trước R và P, tìm quan hệ Q sao cho nó thỏa phương trình (21*).

Công thức (31*) cũng có thể được xem như là một sự phân tích quan hệ R thành quan hệ Q khi cho trước P, hoặc một sự phân tích quan hệ R thành quan hệ P khi cho trước quan hệ Q.

Vì các quan hệ có thể biểu thị ở dạng ma trận, như chúng ta đã biết, phép hợp thành "$\overset{T}{\circ}$" ứng với phép t-norm T sẽ là phép tích ma trận tương tự như phép tích ma trận thông thường, với phép nhân là phép t-norm T và phép cộng là phép lấy max. Vì vậy chúng ta có thể sử dụng công cụ ma trận để giải phương trình (31*).

Một cách tổng quát, các quan hệ trong (31*) có thể suy biến thành các các ma trận một hàng hay một cột. Chẳng hạn, R và P có thể suy biên thành hai vectơ hàng, hoặc R và Q là hai vectơ cột.

Vấn đề phân hoạch bài toán

Trước hết ta xét bài toán cho trước ma trận R và Q, hãy xác định tập các ma trận $\mathbf{S}(Q, R)$ thỏa phương trình (31*), xác định tập lời giải của phương trình (31*)

$$\mathbf{S}(Q, R) = \{P : P \text{ o } Q = R\} \qquad (32^*)$$

trong đó, phép hợp thành o được giới hạn là phép hợp thành max-min.

Không mất tính chất tổng quát có thể thấy dễ dàng và tự nhiên rằng bài toán này sẽ được phân tách thành tập các bài toán biểu thị bằng phương trình ma trận sau:

$$\boldsymbol{p}_i \text{ o } Q = \boldsymbol{r}_i \qquad (33^*)$$

trong đó $i = 1, \ldots, n$, và các vectơ hàng $\boldsymbol{p}_i = (p_{ij} : j = 1, \ldots, m)$ và $\boldsymbol{r}_i = (r_{ik} : k = 1, \ldots, l)$. Công thức (33*) có nghĩa,

$$r_{ik} = \max_{1 \le j \le m} \min(p_{ij}, q_{jk}) \qquad (34^*)$$

Ký hiệu tập các lời giải của phương trình (33*) là

$$\mathbf{S}_i(Q, \boldsymbol{r}_i) = \{\boldsymbol{p}_i : \boldsymbol{p}_i \circ Q = \boldsymbol{r}_i\} \qquad (35^*)$$

với $i = 1, \ldots, n$. Khi đó lời giải của phương trình (31*) sẽ là vectơ cột

$$P = \begin{pmatrix} p_1 \\ p_2 \\ \cdot \\ \cdot \\ \cdot \\ p_n \end{pmatrix}, \text{ với } \boldsymbol{p}_i \in \mathbf{S}_i(Q, \boldsymbol{r}_i) \text{ với mọi } i = 1, \ldots, n.$$

Một câu hỏi đặt ra là khi nào phương trình ma trận (33*) có nghiệm hay không có nghiệm, hay khi nào $S_i(Q, \boldsymbol{r}_i) \neq \varnothing$?

Từ công thức (34*) co thể thấy ngay là nếu

$$\max_{1 \le j \le m} q_{jk} < \max_{1 \le i \le n} r_{ik} \qquad (36^*)$$

với một chỉ số k nào đó, thì $\mathbf{S}_i(Q, \boldsymbol{r}_i) = \varnothing$, nghĩa là không có một ma trận P nào thỏa mãn phương trình ma trận (31*).

Ví dụ 2.1. Xét phương trình ma trận dạng (31*) sau

$$\begin{pmatrix} p_{11} & p_{12} & p_{13} \\ p_{21} & p_{22} & p_{23} \end{pmatrix} \circ \begin{pmatrix} 0{,}9 & 0{,}5 \\ 0{,}7 & 0{,}8 \\ 1{,}0 & 0{,}4 \end{pmatrix} = \begin{pmatrix} 0{,}6 & 0{,}3 \\ 0{,}2 & 1{,}0 \end{pmatrix},$$

với ma trận thứ nhất P là ẩn số. Bài toán đặt ra là xác định tập nghiệm $\mathbf{S}(Q, R)$. Như chúng ta đã trình bày ở trên, bài toán này được phân hoạch thành một tập các bài toán con dạng (33*) sau

$$\begin{pmatrix} p_{11} & p_{12} & p_{13} \end{pmatrix} \circ \begin{pmatrix} 0{,}9 & 0{,}5 \\ 0{,}7 & 0{,}8 \\ 1{,}0 & 0{,}4 \end{pmatrix} = \begin{pmatrix} 0{,}6 & 0{,}3 \end{pmatrix}$$

và

$$\begin{pmatrix} p_{21} & p_{22} & p_{23} \end{pmatrix} \circ \begin{pmatrix} 0{,}9 & 0{,}5 \\ 0{,}7 & 0{,}8 \\ 1{,}0 & 0{,}4 \end{pmatrix} = \begin{pmatrix} 0{,}2 & 1{,}0 \end{pmatrix}.$$

Tuy nhiên, với $k = 2$, $i = 2$, ta có $r_{22} = 1{,}0$ và chúng ta kiểm chứng thấy

$\max_{1\le j\le m} q_{jk} = \max(0{,}5\ \ 0{,}8\ \ 0{,}4) < 1{,}0 = r_{22}$.

Vậy, phương trình ma trân đã cho không có nghiệm, $\mathbf{S}(Q, R) = \varnothing$.

Sau đây chúng ta nghiên cứu phương pháp giải phương trình (31*) hoặc (33*), kể cả phương pháp giải xấp xỉ trong trường hợp $\mathbf{S}(Q, R) = \varnothing$.

***Phương pháp giải phương trình ma trận với phép hợp thành* max-min**

Xét phương trình quan hệ $\boldsymbol{p} \circ Q = \boldsymbol{r}$ (37*)

của một phân hoạch nào đó, ta bỏ qua chỉ số phân hoạch i trong phương trình (33*). Trước hết, chúng ta khảo sát cấu trúc của tập lời giải của phương trình (37*), $\mathbf{S}(Q, \boldsymbol{r}) = \{\boldsymbol{p} : \boldsymbol{p} \circ Q = \boldsymbol{r}\}$.

Gọi P = $\{\boldsymbol{p} = (p_1, \ldots, p_m) : p_j \in [0, 1], j = 1, \ldots, m\}$, $\boldsymbol{p}$ là tập mờ trên không gian $\boldsymbol{V}$. Trên P ta định nghĩa quan hệ thứ tự bộ phận $\le$ trên các vectơ, $\boldsymbol{p} \le \boldsymbol{p}'$ nếu và chỉ nếu $p_j \le p'_j$, với mọi $j = 1, \ldots, m$. Với bây kỳ 2 phần tử $\boldsymbol{p}$ và $\boldsymbol{p}'$, với $\boldsymbol{p} \le \boldsymbol{p}'$, ta định nghĩa đoạn

$$[\boldsymbol{p}, \boldsymbol{p}'] = \{\boldsymbol{p}'' : \boldsymbol{p} \le \boldsymbol{p}'' \le \boldsymbol{p}'\}.$$

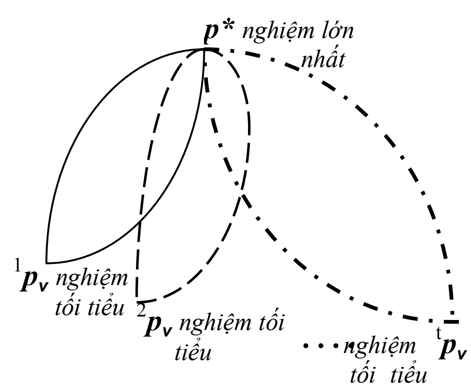
Chúng ta biết rằng tập $[\boldsymbol{p}, \boldsymbol{p}']$ sẽ tạo thành một dàn (lattice).
Dựa trên cấu trúc P ta định nghĩa một số khái niệm sau.

Xét tập lời giải hay tập nghiệm $\mathbf{S}(Q, \boldsymbol{r})$. Phần tử $\boldsymbol{p}^*$ của $\mathbf{S}(Q, \boldsymbol{r})$ được gọi là *nghiệm tối đại* nếu với mọi $\boldsymbol{p} \in \mathbf{S}(Q, \boldsymbol{r})$, ta có $\boldsymbol{p} \ge \boldsymbol{p}^* \Rightarrow \boldsymbol{p} = \boldsymbol{p}^*$, không tồn tại một nghiệm của (37*) nào thực sự lớn hơn $\boldsymbol{p}^*$. Nghiệm $\boldsymbol{p}^*$ được gọi là *lớn nhất* nếu $\boldsymbol{p}^* \ge \boldsymbol{p}$, với $\forall \boldsymbol{p} \in \mathbf{S}(Q, \boldsymbol{r})$.

Một cách tương tự, $\boldsymbol{p}_* \in \mathbf{S}(Q, \boldsymbol{r})$ được gọi là *nghiệm tối tiểu* nếu với mọi $\boldsymbol{p} \in \mathbf{S}(Q, \boldsymbol{r})$, ta có $\boldsymbol{p} \le \boldsymbol{p}_* \Rightarrow \boldsymbol{p} = \boldsymbol{p}_*$, t.l. không tồn tại một nghiệm của (37*) nào thực sự nhỏ hơn $\boldsymbol{p}_*$. Nghiệm $\boldsymbol{p}_v$ được gọi là *nhỏ nhất* nếu $\boldsymbol{p}_v \le \boldsymbol{p}$, với $\forall \boldsymbol{p} \in \mathbf{S}(Q, \boldsymbol{r})$.

Người ta đã xác định được cấu trúc của tập nghiệm $\mathbf{S}(Q, \boldsymbol{r})$ như sau:

- Tập $\mathbf{S}(Q, \boldsymbol{r})$ luôn tồn tại nghiệm lớn nhất $\boldsymbol{p}^*$;



Hình 2.6. Cấu trúc tập $S(Q, \boldsymbol{r})$

- Tập $\mathbf{S}(Q, \boldsymbol{r})$ chứa nhiều nghiệm tối tiểu, nhìn chung phương trình (37*) không có nghiệm nhỏ nhất;

- Với $\boldsymbol{p}'$ là một nghiệm tối tiểu và $\boldsymbol{p}^*$ là nghiệm lớn nhất của (37*), ta có $[\boldsymbol{p}', \boldsymbol{p}^*] \subseteq \mathbf{S}(Q, \boldsymbol{r})$. Nói khác đi, khi ký hiệu $\mathbf{S}_{min} = \mathbf{S}_{min}(Q, \boldsymbol{r})$ là tập các nghiệm tối tiểu của $\mathbf{S}(Q, \boldsymbol{r})$, ta có

$$\mathbf{S}(Q, \boldsymbol{r}) = \bigcup_{p_* \in S_*} [p_*, p^*]$$

Trên Hình 2.6 chúng ta thấy hình ảnh cấu trúc tập nghiệm $\mathbf{S}(Q, \boldsymbol{r})$ với chỉ một nghiệm lớn nhất và một số nghiệm tối tiểu còn tập $[{}^{l}\boldsymbol{p}_{v}, \boldsymbol{p}^*]$ biểu thị bằng hình chiếc lá.

Bây giờ ta xem xét phương pháp hay thủ tục xác định cấu trúc $\mathbf{S}(Q, \boldsymbol{r})$.

(i) *Xác định nghiệm lớn nhất*: Người ta cũng chứng tỏ rằng nếu $\mathbf{S}(Q, \boldsymbol{r}) \neq \varnothing$ thì nghiệm lớn nhất $\boldsymbol{p}^* = (p_j^* : j = 1, \ldots, m)$ được xác định như sau:

$$p_j^* = \min_{1 \leq k \leq l} \varphi(q_{jk}, r_k), \text{ với } \varphi(q_{jk}, r_k) = \begin{cases} r_k & \text{if } q_{jk} > r_k \\ 1 & \text{otherwise} \end{cases} \qquad (38^*)$$

và nếu $\boldsymbol{p}^*$ không thỏa mãn phương trình (37*) thì $\mathbf{S}(Q, \boldsymbol{r}) = \varnothing$, nghĩa là việc tồn tại nghiệm lớn nhất được xác định bởi (38*) là điều kiện cần và đủ để $\mathbf{S}(Q, \boldsymbol{r}) \neq \varnothing$.

(ii) *Xác định tập nghiệm tối tiểu* $\mathbf{S}_{min}(Q, \boldsymbol{r})$: Để xác định được cấu trúc của tập $\mathbf{S}(Q, \boldsymbol{r})$, tiếp theo ta chỉ cần xác định tập $\mathbf{S}_{min}(Q, \boldsymbol{r})$, ta giải bài toán tìm trong các phần tử $\boldsymbol{p} \leq \boldsymbol{p}^*$, tất cả các nghiệm tối tiểu của (37*). Không mất tính tổng quát giả thiết rằng $r_1 \geq \ldots \geq r_s > 0$, với $s \geq l$, nghĩa là giả thiết này kéo theo việc ta chỉ xét phương trình (37*) với việc rút gọn vectơ $\boldsymbol{r}$ xuống còn s thành phần. Thực vậy, nếu các thành phần của vectơ hàng $\boldsymbol{r}$ không phải là dãy số đơn điệu không tăng, ta chỉ cần thực hiện một phép hoán vị thích hợp các vị trí của chúng. Ta có quyền làm được điều mày vì tập chỉ số của vectơ $\boldsymbol{r}$ tương ứng với các phần tử của không gian $\boldsymbol{W}$ mà các phần tử của nó không bị buộc phải được xắp thứ tự. Sau đó, để không làm thay đổi phương trình ma trận (37*) chúng ta thực hiện chính phép hoán vị đó đối với các chỉ số cột của ma trận Q (lưu ý rằng chỉ số cột của Q trùng với chỉ số các thành phần của ***r***). Ngoài ra, với thành phần $r_k = 0$, $k > s$, ta có thể loại bỏ thành phần này của vectơ ***r*** và cột thứ k của Q, vì nếu ***p*** là nghiệm của phương trình (37*) đã được rút gọn thì ta cũng có $\max_{1\leq j\leq m} \min\{p_j, q_{jk}\} = r_k = 0$. Thực vậy, vì ***p**** là nghiệm nên ta phải có

$$\max_{1\leq j\leq m} \min\{ p_j^*, q_{jk}\} = r_k = 0. \qquad (39^*)$$

Từ đây ta suy ra nếu $q_{jk} \neq 0$ thì $p_j^* = 0$ và nếu $q_{jk} = 0$ thì p_j^* có thể nhận giá trị tùy ý trong [0, 1] mà ta vẫn có đẳng thức (39*). Vì ***p*** $\leq$ ***p****, nên ta có $p_j = 0$ đối với j mà $q_{jk} \neq 0$ và do đó ***p*** thỏa mãn (39*). Như vậy mọi nghiệm của phương trình (37*) rút gọn đều là nghiệm của phương trình gốc.

Bây giờ ta chỉ ra rằng ta có thể rút gọn tiếp phương trình (37*) bỏ các dữ liệu liên quan đến các chỉ số j mà $p_j^* = 0$. Cụ thể đối với những chỉ số j này ta loại bỏ thành phần p_j^* của vectơ ***p**** và hàng thứ j của ma trận Q. Khi đó, nếu ***p*** là nghiệm của pgương trình (37*) rút gọn, thì khi khôi phục lại thành phần thứ j đã loại với giá trị bằng 0 ta sẽ thu được nghiệm của phương trình gốc, việc ***p*** được khôi phục như vậy sẽ thỏa phương trình gốc.

Như vậy, chúng ta có thể giả thiết rằng mọi thành phần của vectơ nghiệm lớn nhất ***p**** và vectơ ***r*** đều khác 0, $p_j^* \neq 0$, với $j = 1, \ldots, m$, và $r_k \neq 0$, với $k = 1, \ldots, l$. Khi đó, cho trước Q, ***r*** và ***p**** thỏa mãn các điều kiện trên, tập nghiệm tối tiểu của phương trình rút gọn (37*) được xác định bằng một thủ tục.

Để tránh việc trình bày các kỹ thuật phức tạp chúng ta sẽ không chứng minh tính đúng đắn của thủ tục này. Nhưng để nắm được ý tưởng của thủ tục ta nêu ra một số nhận xét trực quan.

Ta viết lại công thức (34*) để xem xét, với lưu ý rằng ta bỏ qua chỉ số i trong công thức này vì ta đang xét bài toán của một phân hoạch với phương trình (37*):

$$r_k = \max_{1\leq j\leq m} \min(p_j, q_{jk}), \qquad (*)$$

trong đó p_j là thành phần của một vectơ nghiệm tối tiểu nào đó. Như vậy, phải có những chỉ số j để $\min(p_j, q_{jk}) = r_k$, với mọi k. Vì $\boldsymbol{p}$ là tối tiểu nên, đối với những chỉ số j như vậy, ta phải có $p_j = r_k$. Đối với những chỉ số j' khác, giá trị $p_{j'}$ của vectơ $\boldsymbol{p}$ không ảnh hưởng đến kết quả của công thức (*), vì $\min(p_{j'}, q_{jk}) < r_k$. Vì vậy, vì $\boldsymbol{p}$ là tối tiểu nên $p_{j'} = 0$. Vì vậy, các bước chính của thủ tục xác định tập $\mathbf{S}_{min}(Q, \boldsymbol{r})$ bao gồm:

1. Xác định các tập $J_k(\boldsymbol{p}^*) = \{j: 1 \le j \le m, \min(p_j^*, q_{jk}) = r_k\}$, $k = 1, \ldots, l$.

   Thiết lập tich Đề-các $J(\boldsymbol{p}^*) = J_1(\boldsymbol{p}^*) \times J_2(\boldsymbol{p}^*) \times \ldots \times J_l(\boldsymbol{p}^*)$.

   Ký hiệu các phần tử của $J(\boldsymbol{p}^*)$ là $\beta = (\beta_k : k = 1, \ldots, l)$.

2. Đối với mỗi $\beta \in J(\boldsymbol{p}^*)$ và mỗi chỉ số j, $1 \le j \le m$, ta xác định tập sau

$$K(\beta, j) = \{k : 1 \le k \le l, \beta_k = j\}.$$

3. Với mỗi phần tử $\beta \in J(\boldsymbol{p}^*)$, ta sinh các vectơ sau $g(\beta) = (g_j(\beta): j = 1, \ldots, m)$, trong đó :

$$g_j(\beta) = \begin{cases} \max_{k \in K(\beta, j)} r_k & if \quad K(\beta, j) \neq \varnothing, \\ 0 & otherwise \end{cases}.$$

4. Lựa chọn trong các vectơ m-chiều được sinh ra trong Bước 3 những phần tử tối tiểu theo quan hệ thứ tự một phần trong P. Các phần tử như vậy tồn tại vì số các phần tử $g(\beta)$ là hữu hạn và chúng là tập tất cả các nghiệm tối tiểu.

Ví dụ 2.2. Cho trước quan hệ Q và $\boldsymbol{r}$ như sau:

$$Q = \begin{pmatrix} 0{,}1 & 0{,}4 & 0{,}5 & 0{,}1 \\ 0{,}9 & 0{,}7 & 0{,}2 & 0{,}0 \\ 0{,}8 & 1{,}0 & 0{,}5 & 0{,}0 \\ 0{,}1 & 0{,}3 & 0{,}6 & 0{,}0 \end{pmatrix} \quad \text{và} \quad \boldsymbol{r} = \begin{pmatrix} 0{,}8 & 0{,}7 & 0{,}5 & 0{,}0 \end{pmatrix}$$

Hãy xác định tập tất cả các nghiệm $\mathbf{S}(Q, \boldsymbol{r})$ của (37*).

(i) Trước hết ta xác định nghiệm lớn nhất $\boldsymbol{p}^*$ dựa vào công thức (38*). Ta có,

$$p_1^* = \min(1{,}0\ \ 1{,}0\ \ 1{,}0\ \ 0{,}0) \qquad = 0{,}0$$
$$p_2^* = \min(0{,}8\ \ 1{,}0\ \ 1{,}0\ \ 1{,}0) \qquad = 0{,}8$$
$$p_3^* = \min(1{,}0\ \ 0{,}7\ \ 1{,}0\ \ 1{,}0) \qquad = 0{,}7$$
$$p_4^* = \min(1{,}0\ \ 1{,}0\ \ 0{,}5\ \ 1{,}0) \qquad = 0{,}5$$

và $\boldsymbol{p*} = (0{,}0\ \ 0{,}8\ \ 0{,}7\ \ 0{,}5)$. Chúng ta có thể kiểm chứng rằng $\boldsymbol{p*}$ là nghiệm và do đó $\mathbf{S}(Q, \boldsymbol{r}) \neq \varnothing$.

(ii) Xác định các nghiệm tối tiểu: Do $p_1^* = 0{,}0$ và $r_4 = 0{,}0$ ta có phương trình ma trận rút gọn sau:

$$\begin{pmatrix} p_2 & p_3 & p_4 \end{pmatrix} \circ \begin{pmatrix} 0{,}9 & 0{,}7 & 0{,}2 \\ 0{,}8 & 1{,}0 & 0{,}5 \\ 0{,}1 & 0{,}3 & 0{,}6 \end{pmatrix} = \begin{pmatrix} 0{,}8 & 0{,}7 & 0{,}5 \end{pmatrix}.$$

Bây giờ ta thức hiện thủ tục 4 bước đã trình bày ở trên.

1. Với $\boldsymbol{p*} = (0{,}8\ \ 0{,}7\ \ 0{,}5)$, ta có $J_1(\boldsymbol{p*}) = \{2\}$, $J_2(\boldsymbol{p*}) = \{2, 3\}$ và $J_3(\boldsymbol{p*}) = \{3, 4\}$. Vậy, $J(\boldsymbol{p*}) = \{2\} \times \{2, 3\} \times \{3, 4\}$. (Lưu ý rằng, sau khi rút gọn, $j = 2, 3, 4$ còn $k = 1, 2, 3$).

2. Tập $K(\beta, j)$ và các vectơ $g(\beta)$, $\beta \in J(\boldsymbol{p*}) = \{2\} \times \{2, 3\} \times \{3, 4\}$, được xác định và liệt kê trong Bảng 2.4 sau:

Bảng 2.4. Kết quả Bước 2 và 3 trong Ví dụ 2.2

$K(\beta, j)$	$j :=$ 2	3	4	$g(\beta)$ $j :=$ 2 3 4
$\beta =$ 2 2 3	{1, 2}	{3}	$\varnothing$	(0,8 0,5 0,0)
2 2 4	{1, 2}		$\varnothing$	(0,8 0,0 0,5)
2 3 3	{3}			(0,8 0,7 0,0)
2 3 4	{1}	{2, 3}	$\varnothing$	(0,8 0,7 0,5)
	{1}		{2}	
	{3}			

3. Đối với mỗi $\beta \in J(\boldsymbol{p^*})$, ta sinh các vectơ $g(\beta)$ được cho trong cột cuối của Bảng 2.4.

4. Dựa trên quan hệ thứ tự trên P, ta thấy có hai vectơ tối tiểu là (0,8 0,5 0,0) và (0,8 0,0 0,5) và chúng là lập thành tất cả các nghiệm tối tiểu của phương trình ma trận đã cho. Quay về phương trình gốc chưa rút gọn, nghiệm tối tiểu thu được bằng việc thêm thành phần $p_1^* = 0,0$ và, do đó, ta thu được $S_*(Q, \boldsymbol{r}) = \{(0,0\ \ 0,8\ \ 0,5\ \ 0,0), (0,0\ \ 0,8\ \ 0,0\ \ 0,5)\}$.

2.4.2.4. Lập luận với phương trình quan hệ dựa trên các phép hợp thành sup-T

Xét phương trình ma trận

$$P \overset{T}{\circ} Q = R \qquad (40^*)$$

trong đó, thay vì phép hợp thành max-min, o_T ở đây là phép hợp thành sup-T với phép t-norm T, còn các ký hiệu liên quan đến các quan hệ P, Q và R hoàn toàn giữ nguyên như trong mục trên. Tương tự như trong Mục trước, bài toán đặt ra là cho trước các ma trận Q và R, hãy tìm nghiệm ma trận P. Ta sẽ sử dụng cùng các ký pháp như trong mục trước, chẳng hạn $\mathbf{S}(Q, R)$ là tập tất cả các nghiệm của (40*), $\boldsymbol{p^*}$ là nghiệm lớn nhất, nó là phần tử lớn nhất của $\mathbf{S}(Q, R)$ trong tập sắp thứ tự một phần P.

Như ta biết, phương trình (40*) biểu thị một tập các phương trình có dạng

$$\sup_{1\le j\le m} T(p_{ij}, q_{jk}) = r_{ik} \qquad (41^*)$$

với mọi $i = 1, \ldots, n$ và $k = 1, \ldots, l$, và T là một phép t-norm cho trước.

Để giải bài toán này, chúng ta nghiên cứu hai loại phép tính hợp thành được gọi là phép hợp thành sup-T (hay max-T, trong trường hợp hữu hạn) và phép hợp thành $\inf_{IT}$.

1) *Phép hợp thành* sup-T *trên các quan hệ mờ*

Khái quát hóa của phép hợp thành sup-min, hay max-min trong trường hợp hữu hạn, là phép hợp thành sup-T, ký hiệu là $\overset{T}{\circ}$, được định nghĩa như sau:

$$(P \overset{T}{\circ} Q)(u, w) = \sup_{v \in V} T(P(u, v), Q(v, w)) \qquad (42^*)$$

với $\forall u \in \boldsymbol{U}$ và $\forall w \in \boldsymbol{W}$. Như vậy nó trở về phép hợp thành sup-min khi thay phép t-norm T bằng phép t-norm min ($\wedge$).

Giả sử P, P_j là các quan hệ mờ trên $\boldsymbol{U} \times \boldsymbol{V}$, Q và Q_j là các quan hệ mờ xác dịnh trên $\boldsymbol{U} \times \boldsymbol{W}$ và R là quan hệ mờ xác định trên $\boldsymbol{W} \times \boldsymbol{Z}$, trong đó chỉ số $j \in \boldsymbol{J}$. Khi đó, chúng ta có thể kiểm chứng tính đúng đắn của các tính chất sau:

(i) $(P \overset{T}{\circ} Q) \overset{T}{\circ} R = P \overset{T}{\circ} (Q \overset{T}{\circ} R)$, (tính chất kết hợp của phép o_T)

(ii) $P \overset{T}{\circ} (\bigcup_{j\in J} Q_j) = \bigcup_{j\in J} (P \overset{T}{\circ} Q_j)$,

(iii) $P \overset{T}{\circ} (\bigcap_{j\in J} Q_j) \subseteq \bigcap_{j\in J} (P \overset{T}{\circ} Q_j)$,

(iv) $(\bigcup_{j\in J} P_j) \overset{T}{\circ} Q = \bigcup_{j\in J} (P_j \overset{T}{\circ} Q)$,

(v) $(\bigcap_{j\in J} P_j) \overset{T}{\circ} Q \subseteq \bigcap_{j\in J} (P_j \overset{T}{\circ} Q)$,

(vi) $(P \overset{T}{\circ} Q)^{\mathrm{t}} = Q^{\mathrm{t}} \overset{T}{\circ} P^{\mathrm{t}}$, trong đó phép "t" là phép chuyển vị ma trận hàng thành cột hay, một cách tương đương, chuyển cột thành hàng;

(vii) $Q_1 \subseteq Q_2 \Rightarrow (P \overset{T}{\circ} Q_1 \subseteq P \overset{T}{\circ} Q_2)$ & $(Q_1 \overset{T}{\circ} P \subseteq Q_2 \overset{T}{\circ} P)$.

Bây giờ ta chỉ xét trường hợp mà tất cả các quan hệ mờ 2-ngôi đều xác định trên không gian $\boldsymbol{U} \times \boldsymbol{U}$, hay gọi đơn giản là các quan hệ 2-ngôi trên $\boldsymbol{U}$. Tập tất cả các quan hệ như vậy được kí hiệu là R($\boldsymbol{U}$). Tương tự như trong Mục trước, ở đây ta có khái niệm T-bắc cầu: Quan hệ 2-ngôi R trên $\boldsymbol{U}$ là ***T*-bắc cầu** nếu và chỉ nếu $R \overset{T}{\circ} R \subseteq R$.

Nếu R không phải là T-bắc cầu ta định nghĩa bao đóng T-bắc cầu của nó, ký hiệu là R_T, là quan hệ T-bắc cầu *nhỏ nhất* chứa R. Để nghiên cứu bao đóng T-bắc cầu của quan hệ R, ta đưa ra ký pháp sau: Ký hiệu $R^{2(\mathrm{T})} = R$, $R^{2(\mathrm{T})} = R \overset{T}{\circ} R$ và, bởi quy nạp, ta định nghĩa $R^{k(\mathrm{T})} = R^{k\text{-}1(\mathrm{T})} \overset{T}{\circ} R$. Nếu không có gì nhầm lẫn, để cho gọn, ta loại bỏ kí hiệu (T) ở số mũ, thay vì viết $R^{k(\mathrm{T})}$ ta viết R^k. Theo tính chất kết hợp của o_T, ta có $R^k \overset{T}{\circ} R^l = R^{k+l}$.

Theo định nghĩa của phép $\overset{T}{\circ}$, ta có thể thấy rằng

$$R^k(u, v) = \sup_{z_1,\ldots,z_{k-1}} T(R(u,z_1), R(z_1,z_2),\ldots, R(z_{k-1},v)) \qquad (43^*)$$

Định lý 2.5. Với mọi quan hệ 2-ngôi trên ***U***, bao đóng T-bắc cầu của R được tính theo công thức sau:

$$R_T = \bigcup_{1\le n<\infty} R^n \qquad (44^*)$$

Chứng minh: Trước hết ta chứng minh quan hệ $S = \bigcup_{1\le n<\infty} R^n$ thỏa tính chất bắc cầu. Thực vậy, theo tính chất (iv) và (ii), ta thấy

$$S \overset{T}{\circ} S = \bigcup_{1\le n<\infty} R^n \overset{T}{\circ} \bigcup_{1\le m<\infty} R^m = \bigcup_{1\le n<\infty} (R^n \overset{T}{\circ} \bigcup_{1\le m<\infty} R^m)$$

$$= \bigcup_{1\le n<\infty} \bigcup_{1\le m<\infty} (R^n \overset{T}{\circ} R^m) \quad = \bigcup_{1\le n,m<\infty} R^{n+m}$$

$$\subseteq \bigcup_{1\le n<\infty} R^n = S,$$

nghĩa là, theo định nghĩa, S là quan hệ T-bắc cầu.

Rõ ràng ta có $R \subseteq S$ và do đó ta chỉ còn cần chứng minh là S là nhỏ nhất trong các quan hệ T-bắc cầu chứa R. Giả sử Q là quan hệ T-bắc cầu bất kỳ chứa R. Khi đó, ta có $\quad R^2 = R \overset{T}{\circ} R \subseteq Q \overset{T}{\circ} Q \subseteq Q$.

Bằng quy nạp, giả sử $R^n \subseteq Q$, ta suy ra $R^{n+1} = R \overset{T}{\circ} R^n \subseteq Q \overset{T}{\circ} Q \subseteq Q$. Do vậy, với mọi n, ta thu được $R^n \subseteq Q$. Điều này kéo theo kết luận $S = \bigcup_{1\le n<\infty} R^n \subseteq Q$, S là nhỏ nhất trong những quan hệ mờ Q như vậy và do đó, theo định nghĩa, $S = R_T$.

Định lý 2.6. Giả sử R là quan hệ mờ phản xạ trên tập ***U*** hữu hạn n phần tử, $n \ge 2$. Khi đó, ta có $\quad R_T = R^{n-1}$.

Chứng minh: Vì R là phản xạ, ta có $E \subseteq R$, trong đó E là ma trận đơn vị. Khi đó, $R = E \overset{T}{\circ} R \subseteq R \overset{T}{\circ} R = R^2$. Từ đó suy ra rằng $R^n \subseteq R^{n+1}$, với mọi số nguyên n.

Trước khi chứng minh tiếp, để dễ hiểu ta nhắc lại một tính chất của các phép t-norm T. Do tính chất kết hợp của T ta có thể viết $T(a_1, T(a_2, a_3)) = T(a_1, a_2, a_3)$ và do đó, theo quy nạp, ta cũng có $T(a_1, a_2, \ldots, a_m) = T(a_1, T(a_2, a_3, \ldots, a_m))$, $a_i \in [0, 1]$. Dựa vào các tính chất của t-norm ta có thể thấy rằng

$$T(a_1, \ldots, a_{i-1}, a_i, a_{i+1}, \ldots, a_m) \le T(a_1, \ldots, a_{i-1}, a_{i+1}, \ldots, a_m), \quad (45^*)$$

tức là khi bỏ bớt một toán hạng thì hàm T không giảm. Thực vậy, do tính đơn điệu của T, tính chất $T(a, 1) = a$ và tính giao hoán, ta có

$$\begin{aligned} T(a_1, \ldots, a_{i-1}, a_i, a_{i+1}, \ldots, a_m) &\le T(a_1, \ldots, a_{i-1}, 1, a_{i+1}, \ldots, a_m) \\ &= T(a_1, \ldots, a_{i-1}, a_{i+1}, \ldots, a_m, 1) \\ &= T(T(a_1, \ldots, a_{i-1}, a_{i+1}, \ldots, a_m), 1) \\ &= T(a_1, \ldots, a_{i-1}, a_{i+1}, \ldots, a_m), \end{aligned}$$

ta thu được công thức (45*).

Bây giờ ta chứng tỏ rằng $R^n = R^{n-1}$ hay $R^n(u, v) = R^{n-1}(u, v)$, với $\forall u, v \in \boldsymbol{U}$, ở đây $n = |\boldsymbol{U}|$. Với $u = v$, ta có $1 \le R^{n-1}(u, u) \le R^n(u, u) \le 1$, hay $R^n(u, u) = R^{n-1}(u, u)$. Giả sử rằng $u \ne v$. Theo công thức (43*) ta có:

$$R^n(u, v) = \sup_{z_1,\ldots,z_{n-1}} T(R(u,z_1), R(z_1,z_2), \ldots, R(z_{n-1},v)).$$

Do $n = |\boldsymbol{U}|$, dãy các phần tử $u = z_0, z_1, z_2, \ldots, z_{n-1}, z_n = v$ phải có hai phần tử bằng nhau, chẳng hạn $z_i = z_j$, với $0 \le i < j \le n$. Khi đó, dựa vào (45*) và (43*), ta thu được:

$$\begin{aligned} &T(R(u,z_1), \ldots, R(z_{i-1},z_i), \ldots, R(z_j,z_{j+1}), \ldots, R(z_{n-1},v)) \\ &\le T(R(u,z_1), \ldots, R(z_{i-1},z_i), R(z_j,z_{j+1}), \ldots, R(z_{n-1},v)) \\ &\le R^k(u, v) \qquad (\text{với } k \le n-1) \\ &\le R^{n-1}(u, v). \end{aligned}$$

Từ bất đẳng thức này và (43*), ta suy ra $R^n(u, v) \le R^{n-1}(u, v)$, với $\forall u, v \in \boldsymbol{U}$, ta có $R^n \subseteq R^{n-1}$. Như vậy, ta đã chứng minh rằng $R^n = R^{n-1}$. Điều này kéo theo đẳng thức $R^m = R^{n-1}$, với mọi $m \ge n$. Vậy, (44*) dẫn đến đẳng thức $R_T = R^{n-1}$.

2) *Phép hợp thành* $\inf_{\Im_T}$ *trên các quan hệ mờ*

Cho phép t-norm T, phép kéo theo liên kết với T, J_T, được định nghĩa trong Mục trước là:

$$J_T(s, t) = \sup \{u \in [0, 1] : T(s, u) \le t\} \qquad (46^*)$$

Giả sử P và Q là hai quan hệ mờ xác định tương ứng trên $\boldsymbol{U} \times \boldsymbol{V}$ và $\boldsymbol{V} \times \boldsymbol{W}$. Phép hợp thành $\inf_{\Im_T}$, ký hiệu là $\overset{\Im_T}{\circ}$, trên các quan hệ mờ 2-ngôi được định nghĩa như sau

$$(P \overset{\Im_T}{\circ} Q)(u, w) = \inf_{v \in V} J_T(P(u, v), Q(v, w)) \qquad (47^*)$$

với mọi $(u, w) \in \boldsymbol{U} \times \boldsymbol{W}$.

Phép hợp thành $\overset{\Im_T}{\circ}$ có các tính chất sau:

Định lý 2.7. Cho các quan hệ mờ $P(\boldsymbol{U}, \boldsymbol{V})$, $Q(\boldsymbol{V}, \boldsymbol{W})$, $R(\boldsymbol{U}, \boldsymbol{W})$ và $S(\boldsymbol{W}, \boldsymbol{Z})$. Khi đó,

(i) Các khẳng định sau là tương đương

$$P \overset{T}{\circ} Q \subseteq R ; \qquad (48^*)$$

$$Q \subseteq P^{t} \overset{\Im_T}{\circ} R ; \qquad (49^*)$$

$$P \subseteq (Q \overset{\Im_T}{\circ} R^{t})^{t} ; \qquad (50^*)$$

(ii) Ta có: $\quad P \overset{\Im_T}{\circ} (Q \overset{\Im_T}{\circ} S) = (P \overset{T}{\circ} Q) \overset{\Im_T}{\circ} S. \qquad (51^*)$

Chứng minh: (i) Theo Định lý 2.6, ta có

$T(P(u, v), Q(v, w)) \le R(u, w)$ nếu và chỉ nếu $J_T(P(u, v), R(u, w)) \ge Q(v, w)$, với mọi $u \in \boldsymbol{U}$, $v \in \boldsymbol{V}$ và $w \in \boldsymbol{W}$. Từ đây ta suy ra

$\sup_{v \in V} T(P(u, v), Q(v, w)) \le R(u, w)$ nếu và chỉ nếu $\sup_{u \in U} T(P^{t}(v, u), R(u, w)) \ge Q(v, w)$. Nghĩa là, (48*) và (49*) là tương đương.

Để chứng minh (50*), ta hãy viết lại (48*) theo từng điểm (u, w), $u \in \boldsymbol{U}$, $w \in \boldsymbol{W}$, như sau: $\quad \sup_{v \in V} T(P(u, v), Q(v, w)) \le R(u, w)$.

Biểu thức này tương đương với $\sup_{v \in V} T(Q^{t}(w, v), P^{t}(v, u)) \leq R^{t}(w, u)$, ta có :

$$Q^{t} \overset{\Im_T}{\circ} P^{t} \subseteq R^{t}. \tag{52*}$$

Áp dụng sự tương đương giữa (48*) và (49*), ta suy ra (52*) tương đương với (50*) và do đó (50*) tương đương với (48*).

(ii) là hệ quả trực tiếp của khẳng định (iii), Định lý 2.6, nói rằng $J_T(T(s, t), u) = J_T(s, J_T(t, u))$, với mọi $s, t, u \in [0, 1]$.

Định lý 2.8. Cho các quan hệ mờ $P(\boldsymbol{U}, \boldsymbol{V})$, $P_j(\boldsymbol{U}, \boldsymbol{V})$, $Q(\boldsymbol{V}, \boldsymbol{W})$ và $Q_j(\boldsymbol{V}, \boldsymbol{W})$, với $j \in J$. Khi đó,

$$(\bigcup_{j \in J} P_j) \overset{\Im_T}{\circ} Q = \bigcap_{j \in J} (P_j \overset{\Im_T}{\circ} Q), \tag{53*}$$

$$(\bigcap_{j \in J} P_j) \overset{\Im_T}{\circ} Q \supseteq \bigcup_{j \in J} (P_j \overset{\Im_T}{\circ} Q), \tag{54*}$$

$$P \overset{\Im_T}{\circ} (\bigcap_{j \in J} Q_j) = \bigcap_{j \in J} (P \overset{\Im_T}{\circ} Q_j), \tag{55*}$$

$$P \overset{\Im_T}{\circ} (\bigcup_{j \in J} Q_j) \supseteq \bigcup_{j \in J} (P \overset{\Im_T}{\circ} Q_j). \tag{56*}$$

Chứng minh: Các khẳng định tương ứng được suy ra trực tiếp từ các khẳng định (vii), (vi), (ix) và (viii) trong Định lý 2.6.

Sau đây ta phát biểu định lý về tính đơn điệu của phép hợp thành $\overset{\Im_T}{\circ}$.

Định lý 2.9. Cho các quan hệ mờ $P(\boldsymbol{U}, \boldsymbol{V})$, $Q_1(\boldsymbol{V}, \boldsymbol{W})$, $Q_2(\boldsymbol{V}, \boldsymbol{W})$ và $R(\boldsymbol{U}, \boldsymbol{W})$. Khi đó, nếu $Q_1 \subseteq Q_2$, thì

$$P \overset{\Im_T}{\circ} Q_1 \subseteq P \overset{\Im_T}{\circ} Q_2 \text{ và } Q_1 \overset{\Im_T}{\circ} R \supseteq Q_2 \overset{\Im_T}{\circ} R.$$

Chứng minh: Do $Q_1 \subseteq Q_2$, ta có $Q_1 \cap Q_2 = Q_1$ và $Q_1 \cup Q_2 = Q_2$. Áp dụng (55*) ta thu được

$$(P \overset{\Im_T}{\circ} Q_1) \cap (P \overset{\Im_T}{\circ} Q_2) = P \overset{\Im_T}{\circ} (Q_1 \cap Q_2) = P \overset{\Im_T}{\circ} Q_1.$$

Đẳng thức này chứng minh tính đúng đắn của bao hàm thứ nhất phát biểu trong định lý. Bằng cách tương tự, ta có thể chứng minh bao hàm thức thứ hai dựa vào (53*).

Các định lý về phép hợp thành $\overset{\Im_T}{\circ}$ là cơ sở để chứng minh định lý sau mà sẽ có vai trò quan trọng trong việc giải các phương trình quan hệ.

Định lý 2.10. Cho các quan hệ mờ $P(\boldsymbol{U}, \boldsymbol{V})$, $Q(\boldsymbol{V}, \boldsymbol{W})$ và $R(\boldsymbol{U}, \boldsymbol{W})$. Khi đó, ta có:

$$P^{t} \overset{T}{\circ} (P \overset{\Im_T}{\circ} Q) \subseteq Q, \qquad (57^*)$$

$$R \subseteq P \overset{\Im_T}{\circ} (P^{t} \overset{T}{\circ} R), \qquad (58^*)$$

$$P \subseteq (P \overset{\Im_T}{\circ} Q) \overset{\Im_T}{\circ} Q^{t}, \qquad (59^*)$$

$$R \subseteq (R \overset{\Im_T}{\circ} Q^{t}) \overset{\Im_T}{\circ} Q. \qquad (60^*)$$

Chứng minh: Một cách hiển nhiên ta có $P \overset{\Im_T}{\circ} Q \subseteq (P^{t})^{t} \overset{\Im_T}{\circ} Q$. Thiết lập tương ứng các biểu thức con của biểu thức này với công thức (49*) và chuyển về dạng công thức tương đương (48*) ta thu được (57*). Một cách tương tự, từ công thức hiển nhiên đúng $P^{t} \overset{T}{\circ} R \subseteq P^{t} \overset{T}{\circ} R$, khi thiết lập tương ứng với công thức (48*) và áp dụng công thức (49*) tương đương với nó, ta thu được (58*).

Bây giờ ta lấy chuyển vị ma trận của (57*) ta thu được $[P^{t} \overset{T}{\circ} (P \overset{\Im_T}{\circ} Q)]^{t} \subseteq Q^{t}$. Do $(R \overset{T}{\circ} S)^{t} = S^{t} \overset{T}{\circ} R^{t}$, với mọi quan hệ R và S, ta có $(P \overset{\Im_T}{\circ} Q)^{t} \overset{T}{\circ} P \subseteq Q^{t}$. Lại thiết lập công thức con của công thức này tương ứng với công thức con của (48*) và thay thế vào cônt thức tương đương (49*) ta thu được (59*).

Công thức (60*) rõ ràng suy ra trực tiếp từ (59*).

Ta có nhận xét là nếu các dấu bao hàm trong (57*) và (58*) là dấu đẳng thức thì có thể xem hai phép hợp thành sup-*T* và inf-*T* là đối của nhau. Tuy nhiên, ở đây chúng ta chỉ có dấu bao hàm nhưng chúng cũng thể hiện mối liên hệ giữa hai phép hợp thành. Vì vậy, ta có thể xem chúng là đối của nhau theo nghĩa "yếu" như vậy.

Bây giờ ta trở lại việc giải phương trình (40*), $P \overset{T}{\circ} Q = R$, khi cho trước các quan hệ Q và R.

Định lý 2.11. Nếu đối với phương trình (40*) ta có $\mathbf{S}(Q, R) \neq \varnothing$, thì $P^* = (Q \overset{\Im_T}{\circ} R^{t})^{t}$ là lớn nhất trong $\mathbf{S}(Q, R)$.

Chứng minh: Lấy một nghiệm $P' \in \mathbf{S}(Q, R)$, $P' \overset{T}{\circ} Q = R$. Theo sự tương đương của công thức (48*) và (50*) ta có: $P' \subseteq (Q \overset{\Im_T}{\circ} R^t)^t = P^*$.

Vấn đề còn lại là chứng minh $P^* \in \mathbf{S}(Q, R)$. Đặt $S = (Q \overset{\Im_T}{\circ} R^t)^t \overset{T}{\circ} Q = P^* \overset{T}{\circ} Q$ và, do đó, $S^t = Q^t \overset{T}{\circ} (Q \overset{\Im_T}{\circ} R^t)$. Theo (57*) của Định lý 2.10, ta thu được $S^t \subseteq R^t$ hay $S \subseteq R$. Mặt khác, $S = P^* \overset{T}{\circ} Q \supseteq P' \overset{T}{\circ} Q = R$ và do đó $S = P^* \overset{T}{\circ} Q = R$, nghĩa là P^* là nghiệm của phương trình (40*).

Từ định lý trên ta thấy để phương trình (40*) có nghiệm hay điều kiện cần và đủ là: $(Q \overset{\Im_T}{\circ} R^t)^t \overset{T}{\circ} Q = R$.

Ví dụ 2.3. Xét phương trình (3.6-38) với phép t-norm T là phép nhân số học và Q và R được cho như sau:

$$Q = \begin{pmatrix} 0,1 \\ 0,2 \\ 0,3 \end{pmatrix} \text{ và } R = \begin{pmatrix} 0,12 \\ 0,18 \\ 0,27 \end{pmatrix}$$

Khi đó,

$$(P^*)^t = Q \overset{\Im_T}{\circ} R^t = \begin{pmatrix} 0,1 \\ 0,2 \\ 0,3 \end{pmatrix} \overset{\Im_T}{\circ} (0,12 \;\; 0,18 \;\; 0,27) = \begin{pmatrix} 1,0 & 1,0 & 1,0 \\ 0,6 & 0,9 & 1,0 \\ 0,4 & 0,6 & 0,9 \end{pmatrix}$$

Hay,
$$P^* = \begin{pmatrix} 1,0 & 0,6 & 0,4 \\ 1,0 & 0,9 & 0,6 \\ 1,0 & 1,0 & 0,9 \end{pmatrix}.$$

Kiểm chứng ta thấy $$P^* \overset{T}{\circ} Q = \begin{pmatrix} 1,0 & 0,6 & 0,4 \\ 1,0 & 0,9 & 0,6 \\ 1,0 & 1,0 & 0,9 \end{pmatrix} \overset{T}{\circ} \begin{pmatrix} 0,1 \\ 0,2 \\ 0,3 \end{pmatrix} = \begin{pmatrix} 0,12 \\ 0,18 \\ 0,27 \end{pmatrix},$$

nghĩa là $\mathbf{S}(Q, R) \neq \varnothing$ và P^* là nghiệm lớn nhất.

Định lý 2.12. Giả sử $P', P'' \in \mathbf{S}(Q, R)$. Khi đó,

(i) Điều kiện $P' \subseteq P \subseteq P''$ kéo theo $P \in \mathbf{S}(Q, R)$;

(ii) $P' \cup P'' \in \mathbf{S}(Q, R)$, trong đó $\cup$ là phép hợp ứng với t-conorm max.

Chứng minh: (i) Do $R = P' \overset{T}{\circ} Q \subseteq P \overset{T}{\circ} Q \subseteq P'' \overset{T}{\circ} Q = R$, ta suy ra $P'' \overset{T}{\circ} Q = R$, $P \in \mathbf{S}(Q, R)$.

(ii) Theo tính chất (iv) của phép hợp thành sup-T, $(\bigcup_{j\in J} P_j) \overset{T}{\circ} Q = \bigcup_{j\in J} (P_j \overset{T}{\circ} Q)$, ta có $(P' \cup P'') \overset{T}{\circ} Q = (P' \overset{T}{\circ} Q) \cup (P'' \overset{T}{\circ} Q) = R \cup R = R$. Điều này chứng tỏ $P' \cup P'' \in \mathbf{S}(Q, R)$.

3.6.4.5. Lập luận với phương trình quan hệ dựa trên các phép hợp thành $\inf_{\mathrm{I\text{-}T}}$

Xét phương trình quan hệ : $P \overset{\Im_T}{\circ} Q = R$ (61*)

trong đó phép hợp thành sup-T được thay bằng phép hợp thành $\inf_{\mathrm{I}\text{-}T}$ liên kết với phép t-norm T và P, Q và R là các quan hệ mờ xác định trên các không gian hữu hạn, nghĩa là chúng được biểu diễn dưới dạng ma trận.

Tương tự như đối với mục trên, cho trước P và R, ta ký hiệu $\mathbf{S}(Q, R)$ là tập tất cả các nghiệm P của phương trình (61*). Cho đến nay chưa có phương pháp nào xác định được tất cả các *nghiệm tối tiểu*, nhưng *nghiệm tối đại* được xác định bởi định lý sau:

Định lý 2.13. Nếu $\mathbf{S}(Q, R) \neq \emptyset$, thì $P^* = R \overset{\Im_T}{\circ} Q^{\mathrm{t}}$ là nghiệm lớn nhất trong $\mathbf{S}(Q, R)$.

Chứng minh: Lấy một nghiệm bất kỳ $P \in \mathbf{S}(Q, R)$, $P \overset{\Im_T}{\circ} Q = R$. Xét biểu thức $R \overset{\Im_T}{\circ} Q^{\mathrm{t}}$ và dựa vào công thức (59*), đối với mọi P và Q, ta có:

$$P^* = R \overset{\Im_T}{\circ} Q^{\mathrm{t}} = (P \overset{\Im_T}{\circ} Q) \overset{\Im_T}{\circ} Q^{\mathrm{t}} \supseteq P. \quad (62^*)$$

Theo công thức (60*) và kết hợp Định lý 2.9 về tính đơn điệu của $\overset{\Im_T}{\circ}$ với (62*), ta thu được

$$R \subseteq (R \overset{\Im_T}{\circ} Q^{\mathrm{t}}) \overset{\Im_T}{\circ} Q \subseteq P \overset{\Im_T}{\circ} Q = R.$$

Do vậy, $P^* \overset{\Im_T}{\circ} Q = (R \overset{\Im_T}{\circ} Q^{\mathrm{t}}) \overset{\Im_T}{\circ} Q = R$, nghĩa là P^* là nghiệm lớn nhất trong $\mathbf{S}(Q, R)$.

Nhận xét: Như là một hệ quả của định lý trên, phương trình (61*) có nghiệm nếu và chỉ nếu ta có $(R \overset{\Im_T}{\circ} Q^{t}) \overset{\Im_T}{\circ} Q = R$.

Ví dụ 2.4. Giả sử phép t-norm T trong (61*) là phép nhân số học và hai quan hệ mờ Q và R được cho như sau:

$$Q = \begin{pmatrix} 0{,}1 & 0{,}6 \\ 0{,}8 & 0{,}9 \end{pmatrix} \text{ và } R = \begin{pmatrix} 0{,}20 & 1{,}0 \\ 0{,}25 & 1{,}0 \end{pmatrix}.$$

Khi đó,

$$P^* = R \overset{\Im_T}{\circ} Q^{t} = \begin{pmatrix} 0{,}20 & 1{,}0 \\ 0{,}25 & 1{,}0 \end{pmatrix} \overset{\Im_T}{\circ} \begin{pmatrix} 0{,}1 & 0{,}8 \\ 0{,}6 & 0{,}9 \end{pmatrix} = \begin{pmatrix} 0{,}5 & 0{,}9 \\ 0{,}4 & 0{,}9 \end{pmatrix}$$

Kiểm chứng bằng việc thay P^* vào phương trình (61*) ta thấy

$$\begin{pmatrix} 0{,}5 & 0{,}9 \\ 0{,}4 & 0{,}9 \end{pmatrix} \overset{\Im_T}{\circ} \begin{pmatrix} 0{,}1 & 0{,}8 \\ 0{,}6 & 0{,}9 \end{pmatrix} = \begin{pmatrix} 0{,}20 & 1{,}0 \\ 0{,}25 & 1{,}0 \end{pmatrix},$$

nghĩa là P^* là nghiệm và là nghiệm lớn nhất của **S**(Q, R).

Do $(P \overset{\Im_T}{\circ} \mathrm{Q})^{t} \neq Q^{t} \overset{\Im_T}{\circ} P^{t}$, bài toán tìm nghiệm P của phương trình (61*) không thể chuyển trực tiếp về bài toán tìm nghiệm Q khi cho trước P và R. Ký hiệu **S**(P, R) là tập tất cả các nghiệm của bài toán sau, khi đó nghiệm tối tiểu trong **S**(P, R) được xác định bởi định lý sau. Tiếc là cho đến nay chúng ta chưa có phương pháp tìm các nghiệm tối đại trong **S**(Q, R).

Định lý 2.14. Nếu **S**(P, R) $\neq \varnothing$, thì $Q^* = P^{t} \overset{T}{\circ} R$ là nghiệm nhỏ nhất trong **S**(P, R).

Chứng minh: Lấy nghiệm tùy ý $Q \in$ **S**(P, R), t.l. $P \overset{\Im_T}{\circ} Q = R$. Xét biểu thức $P^{t} \overset{T}{\circ} R$ và dựa vào (3.6-55), đối với mọi quan hệ P và Q, ta có

$$Q^* = P^{t} \overset{T}{\circ} R = P^{t} \overset{T}{\circ} (P \overset{\Im_T}{\circ} Q) \subseteq Q \qquad (63^*)$$

Mặt khác, theo công thức (58*) và Định lý 2.9 về tính đơn điệu của $\overset{\Im_T}{\circ}$, từ (63*) ta thu được

$$R \subseteq P \overset{\Im_T}{\circ} (P^{t} \overset{T}{\circ} R) = P \overset{\Im_T}{\circ} Q^* \subseteq P \overset{\Im_T}{\circ} Q = R.$$

Do vậy, $P \overset{\Im_T}{\circ} Q^* = R$ và, do (53*) nghiệm Q^* là nhỏ nhất.

Tương tự như đối với tập nghiệm $\mathbf{S}(Q, R)$, ta cũng dễ dàng chứng minh các tính chất tương tự đối với tập $\mathbf{S}(P, R)$ như sau:

1) Nếu Q', $Q'' \in \mathbf{S}(P, R)$, thì điều kiện $Q' \subseteq Q \subseteq Q''$ kéo theo $Q \in \mathbf{S}(P, R)$;

2) Nếu Q', $Q'' \in \mathbf{S}(P, R)$, thì $Q' \cap Q'' \in \mathbf{S}(P, R)$.

Ví dụ 2.5. Giả sử phép t-norm T trong phương trình (61*) là phép nhân số học và hai quan hệ P và R được cho như sau

$$P = \begin{pmatrix} 0{,}5 & 0{,}9 \\ 0{,}4 & 0{,}9 \end{pmatrix} \text{ và } R = \begin{pmatrix} 0{,}20 & 1{,}0 \\ 0{,}25 & 1{,}0 \end{pmatrix}.$$

Nhớ rằng P ở đấy chính là nghiệm của bài toán trong ví dụ trước.
Khi đó,

$$Q^* = P^{\mathrm{t}} \overset{T}{\circ} R = \begin{pmatrix} 0{,}5 & 0{,}4 \\ 0{,}9 & 0{,}9 \end{pmatrix} \overset{T}{\circ} \begin{pmatrix} 0{,}20 & 1{,}0 \\ 0{,}25 & 1{,}0 \end{pmatrix} = \begin{pmatrix} 0{,}100 & 0{,}5 \\ 0{,}225 & 0{,}9 \end{pmatrix}.$$

Thay Q^* vào phương trình (61*), ta có

$$P \overset{\Im_T}{\circ} Q = \begin{pmatrix} 0{,}5 & 0{,}4 \\ 0{,}9 & 0{,}9 \end{pmatrix} \overset{\Im_T}{\circ} \begin{pmatrix} 0{,}100 & 0{,}5 \\ 0{,}225 & 0{,}9 \end{pmatrix} = \begin{pmatrix} 0{,}20 & 1{,}0 \\ 0{,}25 & 1{,}0 \end{pmatrix}.$$

Vậy, Q^* là nghiệm nhỏ nhất trong $\mathbf{S}(P, R)$ của phương trình (61*).

2.4.2.6. Nghiệm xấp xỉ của phương trình quan hệ

Xét phương trình quan hệ

$$P \overset{T}{\circ} Q = R \qquad (64^*)$$

Nhìn chung, với Q và R cho trước, (64*) không có nghiệm và, do đó, bài toán đặt ra là hãy nghiên cứu vấn đề nghiệm xấp xỉ của phương trình trên. Ý tưởng về nghiệm xấp xỉ như sau.

Giả sử phương trình (64*) không có nghiệm. Khi đó, ta biến đổi một chút các quan hệ mờ Q và R thành quan hệ Q' và R' sao cho phương trình

$$P \overset{T}{\circ} Q' = R' \qquad (65^*)$$

trở nên có nghiệm. Một cách tự nhiên, chúng ta xem các nghiệm của (65*) là các nghiệm xấp xỉ của (64*) nếu chúng thỏa mãn một số điều kiện hợp lý nào đó. Ý tưởng này gợi ý cho chúng ta đưa ra định nghĩa sau đây:

Định nghĩa 2.1. Một quan hệ mờ $P^{\sim}$ được gọi là nghiệm xấp xỉ của (64*) nếu các điều kiện sau thỏa mãn:

(i) Tồn tại các quan hệ $Q' \supseteq Q$ và $R' \subseteq R$ sao cho

$$P^{\sim} \overset{T}{\circ} Q' = R'. \qquad (66^*)$$

(ii) Nếu có những quan hệ P'', Q'' và R'' sao cho $Q \subseteq Q'' \subseteq Q'$, $R' \subseteq R'' \subseteq R$ và $P'' \overset{T}{\circ} Q'' = R''$, thì $Q'' = Q'$ và $R'' = R'$.

Để hiểu ý nghĩa của định nghĩa này, ta hãy viết phương trình quan hệ (64*) theo điểm như sau, $\sup_{v \in V} T(P(u, v), Q(v, w)) = R(u, w)$ (67*)

Khi đó, có thể thấy ý nghĩa của điều kiện (i) là: Vì phương trình (64*) không có nghiệm, nghĩa là có những điểm tại đó dấu đẳng thức trong (67*) phải thay bằng dấu nhỏ hơn '<', và do đó muốn phương trình quan hệ có nghiệm thì hoặc là phải thay thế Q bằng quan hệ lớn hơn, hoặc phải thay thế R bằng quan hệ nhỏ hơn, hoặc cả hai trường hợp. Ý nghĩa của (ii) rõ ràng hơn: R' và Q' tương ứng phải gần R và Q nhất nếu chúng tồn tại.Để làm sáng tỏ, ta đưa ra ví dụ sau.

Ví dụ 2.6. Xét phương trình

$$P \overset{T}{\circ} \begin{pmatrix} 0,1 & 0,3 \\ 0,2 & 0,4 \end{pmatrix} = \begin{pmatrix} 0,5 & 0,6 \end{pmatrix}. \qquad (68^*)$$

trong đó $\overset{T}{\circ}$ là phép hợp thành sup-tích đại số. Ta tính quan hệ

$$P^* = (Q \overset{\Im_T}{\circ} R^t)^t = \left(\begin{pmatrix} 0,1 & 0,3 \\ 0,2 & 0,4 \end{pmatrix} \overset{\Im_T}{\circ} \begin{pmatrix} 0,5 \\ 0,6 \end{pmatrix} \right)^t = \begin{pmatrix} 1,0 & 1,0 \end{pmatrix}.$$

Do

$$P^* \overset{T}{\circ} \begin{pmatrix} 0,1 & 0,3 \\ 0,2 & 0,4 \end{pmatrix} = \begin{pmatrix} 1,0 & 1,0 \end{pmatrix} \overset{T}{\circ} \begin{pmatrix} 0,1 & 0,3 \\ 0,2 & 0,4 \end{pmatrix} = \begin{pmatrix} 0,2 & 0,4 \end{pmatrix} \neq \begin{pmatrix} 1,0 & 1,0 \end{pmatrix},$$

theo Định lý 2.11, ta suy ra phương trình quan hệ (68*) đã cho không có nghiệm.

Bây giờ cúng ta đi tìm nghiệm xấp xỉ của phương trình (68*). Giả sử ta giảm R thành $R' = (0{,}2 \quad 0{,}4)$. Khi đó, nếu phương trình thu được có nghiệm thì nó phải có dạng $P^{\sim} = (Q \overset{\Im_T}{\circ} R'^{t})^{t}$ và có thể kiểm chứng thấy ta cũng có $P^{\sim} = (1{,}0 \quad 1{,}0)$, và

$$P^{\sim} \overset{T}{\circ} \begin{pmatrix} 0{,}1 & 0{,}3 \\ 0{,}2 & 0{,}4 \end{pmatrix} = \begin{pmatrix} 1{,}0 & 1{,}0 \end{pmatrix} \overset{T}{\circ} \begin{pmatrix} 0{,}1 & 0{,}3 \\ 0{,}2 & 0{,}4 \end{pmatrix} = \begin{pmatrix} 0{,}2 & 0{,}4 \end{pmatrix},$$

nghĩa là $P^{\sim}$ là nghiệm của (68*). Giả sử có $R'' = (r_1 \quad r_2)$ sao cho $R' \subseteq R'' \subseteq R$ và phương trình sau

$$P^{*} \overset{T}{\circ} \begin{pmatrix} 0{,}1 & 0{,}3 \\ 0{,}2 & 0{,}4 \end{pmatrix} = R''$$

có nghiệm $P'' = (p_1 \quad p_2)$. Khí đó, ta có hệ phương trình

$$\max\{0{,}1p_1,\ 0{,}2p_2\} = r_1,$$
$$\max\{0{,}3p_1,\ 0{,}4p_2\} = r_2.$$

Hệ này chỉ có nghiệm chỉ khi $r_1 \leq 0{,}2$ và $r_2 \leq 0{,}4$, nghĩa là chỉ khi $R'' \subseteq R'$. Vậy $R'' = R'$, và, theo định nghĩa, $P^{\sim}$ là nghiệm xấp xỉ của phương trình (68*). Hơn nữa, có thể kiểm chứng là bất kỳ quan hệ $(a \quad 1{,}0)$, với $a \in [0, 1]$, đều là nghiệm xấp xỉ của (68*) và do đó nghiệm xấp xỉ không duy nhất.

Bây giờ ta tìm nghiệm xấp xỉ khi tăng Q thành $Q' = \begin{pmatrix} 0{,}1 & 0{,}3 \\ 0{,}5 & 0{,}6 \end{pmatrix}$. Khi đó, dạng nghiệm của phương trình $P' \overset{T}{\circ} Q' = R$ là

$$P^{\sim} = (Q' \overset{\Im_T}{\circ} R^{t})^{t} = \left(\begin{pmatrix} 0{,}1 & 0{,}3 \\ 0{,}5 & 0{,}6 \end{pmatrix} \overset{\Im_T}{\circ} \begin{pmatrix} 0{,}5 \\ 0{,}6 \end{pmatrix} \right)^{t} = \begin{pmatrix} 1{,}0 & 1{,}0 \end{pmatrix}.$$

Kiểm chứng ta thấy

$$P^{\sim} \overset{T}{\circ} \begin{pmatrix} 0{,}1 & 0{,}3 \\ 0{,}5 & 0{,}6 \end{pmatrix} = \begin{pmatrix} 1{,}0 & 1{,}0 \end{pmatrix} \overset{T}{\circ} \begin{pmatrix} 0{,}1 & 0{,}3 \\ 0{,}5 & 0{,}6 \end{pmatrix} = \begin{pmatrix} 0{,}5 & 0{,}6 \end{pmatrix},$$

nghĩa là $P^{\sim}$ đúng là nghiệm của phương trình $P' \overset{T}{\circ} Q' = R$. Giả sử có quan hệ $Q'' = (q_{jk})$ sao cho $Q \subseteq Q'' \subseteq Q'$ và có tồn tại nghiệm $P'' = (p_1 \;\; p_2)$ của phương trình mới $P \overset{T}{\circ} Q'' = R$. Từ hệ thức $Q \subseteq Q'' \subseteq Q'$ ta suy ra $q_{11} = 0{,}1$, $q_{12} = 0{,}3$, $0{,}2 \leq q_{21} \leq 0{,}5$ và $0{,}4 \leq q_{22} \leq 0{,}6$. Thay vào phương trình ta có

$$(p_1 \;\; p_2) \overset{T}{\circ} \begin{pmatrix} 0{,}1 & 0{,}3 \\ q_{21} & q_{22} \end{pmatrix} = \begin{pmatrix} 0{,}5 & 0{,}6 \end{pmatrix},$$

hay chúng ta có

$$\max\{0{,}1p_1,\, q_{21}p_2\} = 0{,}5,$$

$$\max\{0{,}3p_1,\, q_{22}p_2\} = 0{,}5.$$

Suy ra, $q_{21}p_2 = 0{,}5$ và $q_{22}p_2 = 0{,}6$ và do đó ta phải có $q_{21} \geq 0{,}5$ và $q_{22} \geq 0{,}6$. Điều này chứng tỏ $Q'' \supseteq Q'$, và ta thu được $Q'' = Q'$. Theo Định nghĩa 2.1, $P^{\sim} = (1{,}0 \;\; 1{,}0)$ là nghiệm xấp xỉ của (68*).

Từ ví dụ này ta thấy không chỉ nghiệm xấp xỉ của một phương trình quan hệ là không duy nhất mà cả các quan hệ bị biến đổi R' và Q' cũng không duy nhất.

Một câu hỏi đặt ra là nghiệm xấp xỉ về sự tồn tại của một phương trình quan hệ bất kỳ?

Định lý 2.15. Phương trình quan hệ (64*) luôn luôn có nghiệm xấp xỉ và $P^{\sim} = (Q \overset{\Im_T}{\circ} R^t)^t$ là nghiệm xấp xỉ lớn nhất.

Chứng minh: Trước hết chúng ta chứng tỏ rằng $P^{\sim} = (Q \overset{\Im_T}{\circ} R^t)^t$ thỏa các điều kiện của Định nghĩa 2.1.. Chọn $Q' = Q$ và $R' = (Q \overset{\Im_T}{\circ} R^t)^t \overset{T}{\circ} Q$. Khi đó, hiển nhiên ta có $P^{\sim} \overset{T}{\circ} Q = R'$. Ta còn cần kiểm tra xem liệu $R \subseteq R'$. Từ hệ thức (57*) ta suy ra

$$R' = (Q \overset{\Im_T}{\circ} R^t)^t \overset{T}{\circ} Q = [Q^t \overset{T}{\circ} (Q \overset{\Im_T}{\circ} R^t)]^t \subseteq (R^t)^t = R, \quad (69^*)$$

nghĩa là điều kiện (i) thỏa mãn.

Bây giờ ta kiểm tra điều kiện (ii). Giả sử rằng có tồn tại Q'', R'' và P'' sao cho $Q \subseteq Q'' \subseteq Q'$, $R' \subseteq R'' \subseteq R$ và $P'' \overset{T}{\circ} Q'' = R''$. Vì $Q' = Q$, ta có $P'' \overset{T}{\circ} Q = R''$, nghĩa là phương trình $P \overset{T}{\circ} Q = R''$ có nghiệm. Theo Định lý 2.11, $P^* = (Q \overset{\Im_T}{\circ}$

$R''^{t})^{t}$ là nghiệm và là phần tử lớn nhất trong **S**(Q, R''). Vì nó là nghiệm nên ta có $(Q \overset{\Im_T}{\circ} R''^{t})^{t} \overset{T}{\circ} Q = R''$. Do $R' \subseteq R'' \subseteq R$, ta có hệ thức

$$Q \overset{\Im_T}{\circ} R'^{t} \subseteq Q \overset{\Im_T}{\circ} R''^{t} \subseteq Q \overset{\Im_T}{\circ} R^{t}. \tag{70*}$$

Mặt khác, để ý đến vế trái của bao hàm thức trong (69*) và hệ thức (58*), ta có :

$$Q \overset{\Im_T}{\circ} R'^{t} = Q \overset{\Im_T}{\circ} [Q^{t} \overset{T}{\circ} (Q \overset{\Im_T}{\circ} R^{t})] \supseteq Q \overset{\Im_T}{\circ} R^{t}.$$

Kết hợp với (70*) ta thu được $Q \overset{\Im_T}{\circ} R''^{t} = Q \overset{\Im_T}{\circ} R^{t}$ và, do đó, ta có

$$R'' = (Q \overset{\Im_T}{\circ} R''^{t})^{t} \overset{T}{\circ} Q = (Q \overset{\Im_T}{\circ} R^{t})^{t} \overset{T}{\circ} Q = R'.$$

Như vậy, ta đã chứng tỏ rằng điều kiện (ii) thỏa mãn và $P^{\sim}$ là nghiệm xấp xỉ của (64*).

Ta còn cần chứng tỏ rằng $P^{\sim}$ là nghiệm lớn nhất. Thực vậy, giả sử P' là một nghiệm xấp xỉ của (64*), nghĩa là có tồn tại Q', R' sao cho $Q \subseteq Q'$, $R' \subseteq R$ và $P' \overset{T}{\circ} Q = R'$. Khi đó, theo Định lý 2.11, $P' \subseteq (Q' \overset{\Im_T}{\circ} R'^{t})^{t}$ và, do tính đơn điệu (giảm theo biến thứ nhất, tăng theo biến thứ hai) của $\overset{\Im_T}{\circ}$, ta có $(Q' \overset{\Im_T}{\circ} R'^{t})^{t} \subseteq (Q \overset{\Im_T}{\circ} R^{t})^{t} = P^{\sim}$. Nghĩa là, $P' \subseteq P^{\sim}$.

2.5. Lập luận xấp xỉ đa điều kiện

Nhìn chung ý tưởng của phương pháp lập luận xấp xỉ là thiết lập cách tính kết luận từ một tập các tri thức dạng luật (mệnh đề nếu-thì) và các sự kiện, dựa trên lý thuyết tập mờ. Tri thức càng đầy đủ thì kết luận được tính càng phù hợp với thực tiễn hơn. Trong các quy tắc lập luận trình bày trong các phần trước, tiền đề chỉ chứa một luật và vì vậy đôi khi chúng ta gọi là phương pháp lập luận mờ đơn điều kiện (fuzzy single conditional reasoning method). Trong mục này chúng ta nghiên cứu phương pháp lập luận dựa vào nhiều luật và được gọi là phương pháp lập luận mờ đa điều liện (fuzzy multiple conditional reasoning method). Từ "mờ" trong thuật ngữ này đôi khi được bỏ qua cho gọn.

Phương pháp lập luận đa điều kiện được mô tả bằng lược đồ sau:

Tiền đề 1: Nếu X là A_1, thì Y là B_1
Tiền đề 1: Nếu X là A_2, thì Y là B_2
. . . . (71*)
Tiền đề n: Nếu Y là A_n, thì Y là B_n
Sự kiện: X là A', .
Kết luận: Y là B'

trong đó, X và Y là các biến ngôn ngữ với các không gian tham chiếu hay không gian cơ sở tương ứng là $\boldsymbol{U}$ và $\boldsymbol{V}$, còn A_i, B_i, A' và B', với $i = 1, 2, \ldots, n$, là những nhãn ngôn ngữ của các tập mờ xác định trên các không gian tham chiếu $\boldsymbol{U}$ hoặc $\boldsymbol{V}$. Tập n luật phát biểu trong các tiền đề trên được gọi là *mô hình mờ* vì nó mô tả hay mô hình hóa mối quan hệ giữa hai đại lượng được mô tả bằng các biến X và Y bằng các tập mờ.

Bất kỳ phương pháp nào cho phép tính kết luận B' từ các tiền đề và sự kiện trong (71*) được gọi là một phương pháp lập luận xấp xỉ đa điều kiện.

Vì chúng ta đang nằm trong môi trường thông tin không chắc chắn, mờ, nên sẽ không có một phương pháp lập luận chính xác và duy nhất. Mỗi phương pháp sẽ xuất phát từ một quan sát trực quan nào đó. Vì vậy, nhìn chung chúng ta sẽ có một số cách giải bài toán lập luận xấp xỉ.

Bây giờ chúng ta nghiên cứu một số phương pháp lập luận xấp xỉ đa điều kiện.

2.5.1. Phương pháp dựa trên quy tắc modus ponens

Phương pháp này dựa trên ý tưởng xem n luật trong mô hình mờ được liên kết với nhau bằng *phép tuyển* (disjuctive) hoặc *phép hội* (conjunctive). Như vậy ta có thể áp dụng quy tắc modus ponens cho từng luật sau đó kết nhập (aggregate) các kết luận thu được đối với từng luật.

2.5.1.1. Mô hình mờ được coi là tuyển của các luật

Phương pháp này bao gồm các bước sau:

Bước 1: Chọn một phương pháp thống nhất tính quan hệ mờ $R_j(u, v) = \mathrm{J}(A_j(u), B_j(v))$ để biểu thị ngữ nghĩa của các luật trong (71*). Khi đó, với dữ liệu đầu vào A' và với mỗi luật thứ j, $j = 1, 2, \ldots, n$, kết luận trung gian B'_j được tính theo quy tắc modus ponens tổng quát

$$B'_j = A' \overset{T}{\circ} R, \text{ hay } B'_j(v) = \sup_{u \in \boldsymbol{U}} T(A'(u), \mathrm{J}(A_j(u), B_j(v))).$$

Bước 2: Biểu thị phép hội liên kết các luật bằng phép t-norm chuẩn, phép hợp tập mờ, ta tính kết luận B' theo công thức

$$B' = \bigcup_{1 \le j \le n} B'_j \text{ , hay}$$

$$B'(v) = \max_{1 \le j \le n} \sup_{u \in \boldsymbol{U}} T(A'(u), \mathrm{J}(A_j(u), B_j(v))), v \in \boldsymbol{V}. \quad (72^*)$$

Trong trường hợp phép kéo theo được định nghĩa bởi Mamdani, $\mathrm{J}(A_j(u), B_j(v)) = \min\{A_j(u), B_j(v)\}$, và T là phép t-norm chuẩn, phép min, công thức trên sẽ trở thành

$$\begin{aligned} B'(v) &= \max_{1 \le j \le n} \sup_{u \in \boldsymbol{U}} \min\{A'(u), \min[A_j(u), B_j(v)]\} \\ &= \max_{1 \le j \le n} \sup_{u \in \boldsymbol{U}} \min\{\min[A'(u), A_j(u)], B_j(v))\} \\ &= \max_{1 \le j \le n} \min\{\sup_{u \in \boldsymbol{U}} \min[A'(u), A_j(u)], B_j(v))\} \\ &= \max_{1 \le j \le n} \min\{high(A' \cap A_j), B_j(v))\} \qquad (73^*) \end{aligned}$$

trong đó $high(.)$ là chiều cao của một tập mờ. Giá trị $high(A' \cap A_j)$ có thể được xem là độ tương hợp của dữ liệu đầu vào A' với tiền tố A_j của luật thứ j.

Với những giả thiết giới hạn như trên, từ công thức (73*) ta thu được một phương pháp lập luận đơn giản hơn như sau:

Bước 1: Vì các luật trong (71*) là các "điểm tựa" tri thức để chúng ta suy luận, nên với giá trị đầu vào A' ta hãy tính độ tương hợp giữa A' và các tiền tố A_j của luật thứ j, $j = 1, 2, \ldots, n$, bằng công thức:

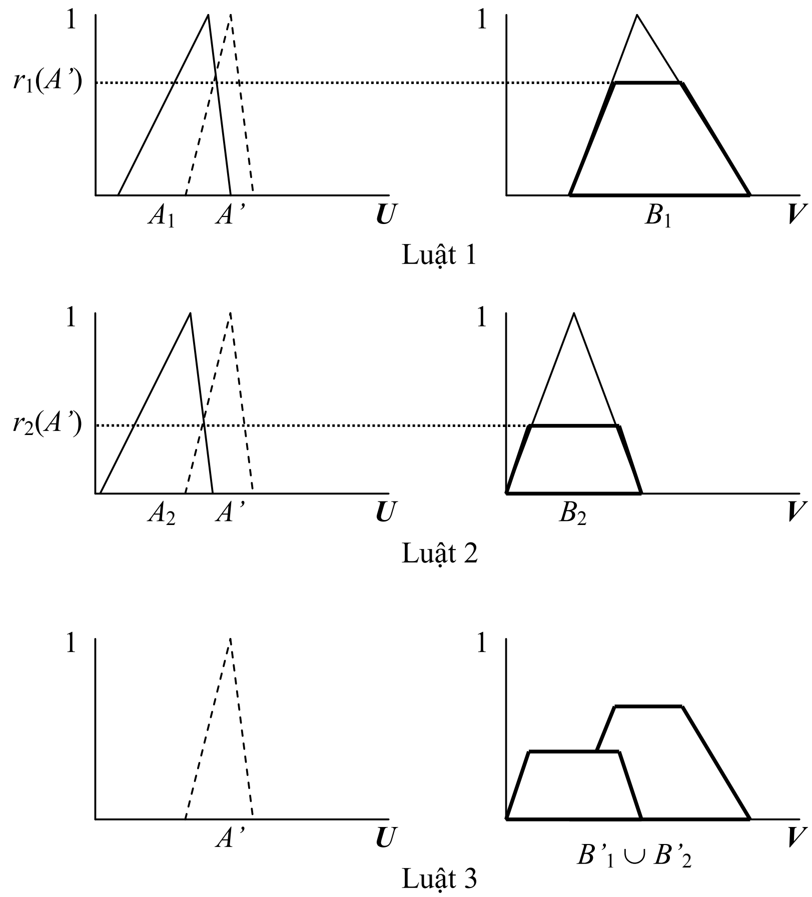
$$r_j(A') = \mathrm{high}(A' \cap A_j) = \sup_{u \in \boldsymbol{U}} \min\{A'(x), A_j(x)\}.$$

Bước 2: Vì độ tương hợp là $r_j(A')$, kết luận suy ra được dựa vào luật thứ j sẽ là $B'_j = \min\{r_j(A'), B_j\}$, B'_j là tập mờ B_j bị cắt ngọn sao cho chiều cao của phần còn lại là $r_j(A')$.

Bước 3: Vì sự liên kết các luật trong (71*) đã được xem như là phép tuyển, kết luận suy ra được từ n luật sẽ được tính bằng công thức

$$B' = \bigcup_{1 \le j \le n} B'_j$$

Giới hạn các tập mờ hình tam giác, phương pháp lập luận xấp xỉ như vậy có thể được biểu thị trong Hình 2.7, trong đó mô hình mờ chỉ chứa 2 luật.



Hình 2.7

2.5.1.2. Mô hình mờ được coi là hội của các luật

Phương pháp lập luận trong trường hợp này hoàn toàn tương tự như trên, chỉ khác biệt ở Bước 2 như sau:

Bước 2': Vì mô hình mờ được xem là hội của các luật nên kết luận B' được tính bằng giao của các kết luận trung gian B'_j như sau $B' = \bigcap_{1\le j\le n} B'_j$, hay

$$B'(v) = \min_{1\le j\le n} \sup_{u\in \boldsymbol{U}} T(A'(u), \mathrm{J}(A_j(u), B_j(v))),\ v \in \boldsymbol{V}. \quad (74^*)$$

Giới hạn phép kéo theo được xác định bởi Mamdani và T là phép t-norm chuẩn, phép min, ta có công thức tính B' như sau:

$$B'(v) \quad = \min_{1\le j\le n} \sup_{u\in \boldsymbol{U}} \min\{A'(u), \min[A_j(u), B_j(v)]\}$$

$= \min_{1\le j\le n} \sup_{u\in U} \min\{\min[A'(u), A_j(u)], B_j(v))\}$
$= \min_{1\le j\le n} \min\{\sup_{u\in U} \min[A'(u), A_j(u)], B_j(v))\}$
$= \min_{1\le j\le n} \min\{high(A' \cap A_j), B_j(v))\}$ (75*)

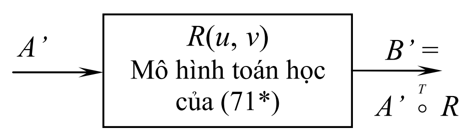
2.5.2. Phương pháp lập luận dựa vào việc mô hình hóa toán học của mô hình mờ

Mô hình mờ (71*) biểu thị tri thức chuyên gia trong một lĩnh vực ứng dụng nào đó. Khi xem nó như là một đối tượng chung, không tách rời, ta có nhu cầu mô hình hóa nó bằng một đối tượng toán học, cụ thể là bằng một quan hệ mờ. Với cách nhìn đó, ta xây dựng một phương pháp lập luận như sau:

Bước 1: Tương tự như Bước 1 trong Mục trước, mỗi luật trong mô hình mờ được biểu thị bằng một quan hệ $R_j(u, v) = J(A_j(u), B_j(v))$. Để xác định quan hệ mờ biểu diễn mô hình mờ (71*), chúng ta thực hiện việc kết nhập (aggregate) các quan hệ $R_j(u, v)$ bằng phép t-conorm chuẩn hay phép hợp các tập mờ:

$$R(u, v) = \bigcup_{1\le j\le n} R_j(u,v). \qquad (76^*)$$

Bước 2: Hình 2.8 thể hiện mô hình mờ (71*) được biểu thị bằng quan hệ mờ R và, tương tự như trường hợp quy tắc modus ponens, kết luận B' được tính theo qua tắc suy luận hợp thành:



Hình 2.8

$$B' = A' \overset{T}{\circ} R.$$

Nhìn chung có sự khác biệt lớn giữa phương pháp lập luận xấp xỉ ở đây với phương pháp được trình bày trong Mục 2.5.1. Tuy nhiên, trong những điều kiện hạn chế chúng lại đồng nhất với nhau. Thực vậy, cũng như trên, ta giả thiết phép kéo theo được xác định là Mamdani, phép $\overset{T}{\circ}$ là phép hợp thành sup-min, kết luận B' được tính như sau

$B'(v) = \sup_{u\in U} \min\{A'(u), R(u, v)\}$
$= \sup_{u\in U} \min\{A'(u), \max_{1\le j\le n} \min[A_j(u), B_j(v)]\}$
$= \max_{1\le j\le n} \sup_{u\in U} \min\{A'(u), \min[A_j(u), B_j(v)]\}$

$= \max_{1\le j\le n} \sup_{u\in \boldsymbol{U}} \min\{\min[A'(u), A_j(u)], B_j(v)]\}$

$= \max_{1\le j\le n} \min\{\sup_{u\in \boldsymbol{U}} \min[A'(u), A_j(u)], B_j(v))\}$

$= \max_{1\le j\le n} \min\{high(A' \cap A_j), B_j(v))\}$

$= \max_{1\le j\le n} B'_j(v)$

Vậy, $$B' = A' \text{ o } R = \bigcup_{1\le j\le n} B'_j$$

trong đó B'_j là phần tập mờ B_j bị cắt ngọn với chiều cao còn lại là $high(A' \cap A_j)$. Điều này chứng tỏ B' cũng tính được từ phương pháp lập luận dựa trên quy tắc modus ponens.

Chú ý: Trong trường hợp chúng ta xem các luật của mô hình mờ *được liên kết bằng phép hội*, quan hệ mờ R trong công thức (76*) sẽ được tính theo công thức sau

$$R(u, v) = \bigcap_{1\le j\le n} R_j(u,v). \qquad (77^*)$$

Nhìn chung, cho đến nay, cho một mô hình mờ (71*) và một cách biểu thị ngữ nghĩa $J(A_j(u), B_j(v))$ của các luật trong mô hình, chúng ta có 4 cách tính kết luận đầu ra B' như sau:

(1) $_1B' = A' \overset{T}{\circ} (\bigcup_{1\le j\le n} R_j)$ (2) $_2B' = A' \overset{T}{\circ} (\bigcap_{1\le j\le n} R_j)$

(3) $_3B' = \bigcup_{1\le j\le n} A' \overset{T}{\circ} R_j$ (4) $_4B' = \bigcap_{1\le j\le n} A' \overset{T}{\circ} R_j$

Trong trường hợp phép t-norm T của phép hợp thành $\overset{T}{\circ}$ là phép min, mối quan hệ giữa 4 phương pháp lập luận như vậy được thiết lập trong định lý sau:

Định lý 2.16. Nếu phép hợp thành $\overset{T}{\circ}$ là sup-min, ta có $_2B' \subseteq {}_4B' \subseteq {}_1B' = {}_3B'$.

Chứng minh: Xét công thức tính $_4B'$:

$$\begin{aligned}
{}_4B'(v) &= \min_{1\le j\le n} (A' \text{ o } R_j)(v) \\
&= \min_{1\le j\le n} \sup_{u\in \boldsymbol{U}} \min[A'(u), R_j(u, v)] \\
&\ge \sup_{u\in \boldsymbol{U}} \min_{1\le j\le n} \min[A'(u), R_j(u, v)] \\
&= \sup_{u\in \boldsymbol{U}} \min[A'(u), \min_{1\le j\le n} R_j(u, v)] \\
&= \sup_{u\in \boldsymbol{U}} \min[A'(u), (\bigcap_{1\le j\le n} R_j)(u, v)] \\
&= (A' \overset{T}{\circ} (\bigcap_{1\le j\le n} R_j))(v) \quad = {}_2B'(v).
\end{aligned}$$

Như vậy chúng ta đã chứng tỏ rằng ${}_2B' \subseteq {}_4B'$.

Tiếp theo, ta thấy rằng $A' \text{ o } R_j \subseteq A' \overset{T}{\circ} (\bigcup_{1\le j\le n} R_j)$, với mọi j: $1 \le j \le n$. Do vậy,

$${}_4B' = \bigcap_{1\le j\le n} A' \overset{T}{\circ} R_j \subseteq A' \overset{T}{\circ} (\bigcup_{1\le j\le n} R_j) = {}_1B'.$$

Cuối cùng, ta xét ${}_1B'$:

$$\begin{aligned} {}_1B'(v) &= \sup_{u\in \boldsymbol{U}} \min\,[A'(u), \bigcup_{1\le j\le n} R_j(u,v)] \\ &= \sup_{u\in \boldsymbol{U}} \max_{1\le j\le n} \min\,[A'(u), R_j(u, v)] \\ &= \max_{1\le j\le n} \sup_{u\in \boldsymbol{U}} \min\,[A'(u), R_j(u, v)] \\ &= (\bigcup_{1\le j\le n} A' \overset{T}{\circ} R_j)(v) = {}_3B'(v). \end{aligned}$$

Như vậy, định lý đã được hoàn toàn chứng minh.

Ví dụ 2.7. Xét mối quan hệ giữa hai biến ngôn ngữ áp suất *AS* và nhiệt độ *NĐ* trong hệ thống điều khiển phân phối chất lỏng trong một nhà máy. Giả sử nhiệt độ nằm trong giới hạn [400;1000] theo đơn vị psi và nhiệt độ trong giới hạn [130;140] độ F.

Giả thiết rằng quan hệ giữa hai đại lượng này tuân theo các luật của mô hình mờ sau:

Nếu *NĐ* := *rất cao* thì *AS* := *cao*

(78*)

Nếu *NĐ* := *thấp* thì *AS* := *khá thấp*

Tập mờ biểu thị ngữ nghĩa của các giá trị ngôn ngữ của hai biến ngôn ngữ được đặc trưng bởi các hàm thuộc sau:

Nhiệt độ *rất cao* = 0,0/134 + 0,0/135 + 0,2/136 + 0,4/137 + 0,7/138 + 1,0/139

Nhiệt độ *thấp* = 1,0/134 + 0,8/135 + 0,6/136 + 0,4/137 + 0,2/138 + 0,0/139

Áp suất *cao* = 0,0/400 + 0,2/600 + 0,4/700 + 0,6/800 + 0,8/900 + 1,0/1000

Áp suất *khá thấp* $= 1{,}0/400 + 0{,}9/600 + 0{,}8/700 + 0{,}6/800 + 0{,}4/900 + 0{,}0/1000$

1) Câu hỏi đặt ra là hãy tính áp suất của chất lỏng tương ứng với nhiệt độ là *cao* được đặc trưng bởi hàm thuộc sau:

Nhiệt độ *cao* $= 0{,}0/134 + 0{,}2/135 + 0{,}4/136 + 0{,}6/137 + 0{,}8/138 + 1{,}0/139$

Giải: Ta sẽ tính áp suất ứng với nhiệt độ đầu vào $A' = cao$ bằng 4 phương pháp trên:

Trước hết chúng ta tính các quan hệ mờ biểu thị hai luật trong (78*).

Giả sử rằng quan hệ mờ R_i, $i = 1, 2$, được tính dựa theo phép kéo theo Zadeh, $R(u, v) = \max\{\min[A(u), B(v)], 1 - A(u)\}$ và $\overset{T}{\circ}$ là phép hợp thành sup-min. Khi đó,

$$R_1(u,v) = \begin{pmatrix} 1{,}0 & 1{,}0 & 1{,}0 & 1{,}0 & 1{,}0 & 1{,}0 \\ 1{,}0 & 1{,}0 & 1{,}0 & 1{,}0 & 1{,}0 & 1{,}0 \\ 0{,}8 & 0{,}8 & 0{,}8 & 0{,}8 & 0{,}8 & 0{,}8 \\ 0{,}6 & 0{,}6 & 0{,}6 & 0{,}6 & 0{,}6 & 0{,}6 \\ 0{,}3 & 0{,}3 & 0{,}4 & 0{,}6 & 0{,}7 & 0{,}7 \\ 0{,}0 & 0{,}2 & 0{,}4 & 0{,}6 & 0{,}8 & 1{,}0 \end{pmatrix}; \quad R_2(u,v) = \begin{pmatrix} 1{,}0 & 0{,}9 & 0{,}8 & 0{,}6 & 0{,}4 & 0{,}0 \\ 0{,}8 & 0{,}8 & 0{,}8 & 0{,}6 & 0{,}4 & 0{,}2 \\ 0{,}6 & 0{,}6 & 0{,}6 & 0{,}6 & 0{,}4 & 0{,}4 \\ 0{,}6 & 0{,}6 & 0{,}6 & 0{,}6 & 0{,}6 & 0{,}6 \\ 0{,}8 & 0{,}8 & 0{,}8 & 0{,}8 & 0{,}8 & 0{,}8 \\ 1{,}0 & 1{,}0 & 1{,}0 & 1{,}0 & 1{,}0 & 1{,}0 \end{pmatrix}.$$

$${}_2B' = A' \overset{T}{\circ} \left(\bigcap\nolimits_{1\le j\le n} R_j\right) = (0{,}0 \quad 0{,}2 \quad 0{,}4 \quad 0{,}6 \quad 0{,}8 \quad 1{,}0) \circ \begin{pmatrix} 1{,}0 & 0{,}9 & 0{,}8 & 0{,}6 & 0{,}4 & 0{,}0 \\ 0{,}8 & 0{,}8 & 0{,}8 & 0{,}6 & 0{,}4 & 0{,}2 \\ 0{,}6 & 0{,}6 & 0{,}6 & 0{,}6 & 0{,}4 & 0{,}4 \\ 0{,}6 & 0{,}6 & 0{,}6 & 0{,}6 & 0{,}6 & 0{,}6 \\ 0{,}3 & 0{,}3 & 0{,}4 & 0{,}6 & 0{,}7 & 0{,}7 \\ 0{,}0 & 0{,}2 & 0{,}4 & 0{,}6 & 0{,}8 & 1{,}0 \end{pmatrix}$$

$$= (0{,}6 \ \ 0{,}6 \ \ 0{,}6 \ \ 0{,}6 \ \ 0{,}8 \ \ 1{,}0)$$

Đối với ${}_4B'$, ta tính

$$A' \circ R_1 = (0{,}0\ \ 0{,}2\ \ 0{,}4\ \ 0{,}6\ \ 0{,}8\ \ 1{,}0) \circ \begin{pmatrix} 1{,}0 & 1{,}0 & 1{,}0 & 1{,}0 & 1{,}0 & 1{,}0 \\ 1{,}0 & 1{,}0 & 1{,}0 & 1{,}0 & 1{,}0 & 1{,}0 \\ 0{,}8 & 0{,}8 & 0{,}8 & 0{,}8 & 0{,}8 & 0{,}8 \\ 0{,}6 & 0{,}6 & 0{,}6 & 0{,}6 & 0{,}6 & 0{,}6 \\ 0{,}3 & 0{,}3 & 0{,}4 & 0{,}6 & 0{,}7 & 0{,}7 \\ 0{,}0 & 0{,}2 & 0{,}4 & 0{,}6 & 0{,}8 & 1{,}0 \end{pmatrix}$$

$$= (0{,}6\ \ 0{,}6\ \ 0{,}6\ \ 0{,}6\ \ 0{,}8\ \ 1{,}0)$$

và
$$A' \circ R_2 = (0{,}0\ \ 0{,}2\ \ 0{,}4\ \ 0{,}6\ \ 0{,}8\ \ 1{,}0) \circ \begin{pmatrix} 1{,}0 & 0{,}9 & 0{,}8 & 0{,}6 & 0{,}4 & 0{,}0 \\ 0{,}8 & 0{,}8 & 0{,}8 & 0{,}6 & 0{,}4 & 0{,}2 \\ 0{,}6 & 0{,}6 & 0{,}6 & 0{,}6 & 0{,}4 & 0{,}4 \\ 0{,}6 & 0{,}6 & 0{,}6 & 0{,}6 & 0{,}6 & 0{,}6 \\ 0{,}8 & 0{,}8 & 0{,}8 & 0{,}8 & 0{,}8 & 0{,}8 \\ 1{,}0 & 1{,}0 & 1{,}0 & 1{,}0 & 1{,}0 & 1{,}0 \end{pmatrix}$$

$$= (1{,}0\ \ 1{,}0\ \ 1{,}0\ \ 1{,}0\ \ 1{,}0\ \ 1{,}0)$$

Vậy,
$${}_4B' = \bigcap_{1 \le j \le n} A' \overset{T}{\circ} R_j = (0{,}6\ \ 0{,}6\ \ 0{,}6\ \ 0{,}6\ \ 0{,}8\ \ 1{,}0)$$

Bây giờ ta tính ${}_1B' = {}_3B'$:

$${}_1B' = A' \overset{T}{\circ} \Big(\bigcup_{1 \le j \le n} R_j\Big) = (0{,}0\ \ 0{,}2\ \ 0{,}4\ \ 0{,}6\ \ 0{,}8\ \ 1{,}0) \circ \begin{pmatrix} 1{,}0 & 1{,}0 & 1{,}0 & 1{,}0 & 1{,}0 & 1{,}0 \\ 1{,}0 & 1{,}0 & 1{,}0 & 1{,}0 & 1{,}0 & 1{,}0 \\ 0{,}8 & 0{,}8 & 0{,}8 & 0{,}8 & 0{,}8 & 0{,}8 \\ 0{,}6 & 0{,}6 & 0{,}6 & 0{,}6 & 0{,}6 & 0{,}6 \\ 0{,}8 & 0{,}8 & 0{,}8 & 0{,}8 & 0{,}8 & 0{,}8 \\ 1{,}0 & 1{,}0 & 1{,}0 & 1{,}0 & 1{,}0 & 1{,}0 \end{pmatrix}$$

$$= (1{,}0\ \ 1{,}0\ \ 1{,}0\ \ 1{,}0\ \ 1{,}0\ \ 1{,}0)$$

2) Xét dữ liệu đầu vào nhiệt độ là *thấp* được đặc trưng bởi hàm thuộc sau:

Nhiệt độ *thấp* = 1,0/134 + 0,8/135 + 0,6/136 + 0,4/137 + 0,2/138 + 0,0/139

Tương tự như trên, ta tính

$${}_2B' = A' \overset{T}{\circ} \left(\bigcap_{1\le j\le n} R_j\right) = (1{,}0 \quad 0{,}8 \quad 0{,}6 \quad 0{,}4 \quad 0{,}2 \quad 0{,}0) \text{ o } \begin{pmatrix} 1{,}0 & 0{,}9 & 0{,}8 & 0{,}6 & 0{,}4 & 0{,}0 \\ 0{,}8 & 0{,}8 & 0{,}8 & 0{,}6 & 0{,}4 & 0{,}2 \\ 0{,}6 & 0{,}6 & 0{,}6 & 0{,}6 & 0{,}4 & 0{,}4 \\ 0{,}6 & 0{,}6 & 0{,}6 & 0{,}6 & 0{,}6 & 0{,}6 \\ 0{,}3 & 0{,}3 & 0{,}4 & 0{,}6 & 0{,}7 & 0{,}7 \\ 0{,}0 & 0{,}2 & 0{,}4 & 0{,}6 & 0{,}8 & 1{,}0 \end{pmatrix}$$

$$= (1{,}0 \;\; 0{,}9 \;\; 0{,}8 \;\; 0{,}6 \;\; 0{,}4 \;\; 0{,}4)$$

Đối với ${}_4B'$, ta tính:

$$A' \text{ o } R_1 = (1{,}0 \;\; 0{,}8 \;\; 0{,}6 \;\; 0{,}4 \;\; 0{,}2 \;\; 0{,}0) \text{ o } \begin{pmatrix} 1{,}0 & 1{,}0 & 1{,}0 & 1{,}0 & 1{,}0 & 1{,}0 \\ 1{,}0 & 1{,}0 & 1{,}0 & 1{,}0 & 1{,}0 & 1{,}0 \\ 0{,}8 & 0{,}8 & 0{,}8 & 0{,}8 & 0{,}8 & 0{,}8 \\ 0{,}6 & 0{,}6 & 0{,}6 & 0{,}6 & 0{,}6 & 0{,}6 \\ 0{,}3 & 0{,}3 & 0{,}4 & 0{,}6 & 0{,}7 & 0{,}7 \\ 0{,}0 & 0{,}2 & 0{,}4 & 0{,}6 & 0{,}8 & 1{,}0 \end{pmatrix}$$

$$= (1{,}0 \;\; 1{,}0 \;\; 1{,}0 \;\; 1{,}0 \;\; 1{,}0 \;\; 1{,}0)$$

và $$A' \text{ o } R_2 = (1{,}0 \;\; 0{,}8 \;\; 0{,}6 \;\; 0{,}4 \;\; 0{,}2 \;\; 0{,}0) \text{ o } \begin{pmatrix} 1{,}0 & 0{,}9 & 0{,}8 & 0{,}6 & 0{,}4 & 0{,}0 \\ 0{,}8 & 0{,}8 & 0{,}8 & 0{,}6 & 0{,}4 & 0{,}2 \\ 0{,}6 & 0{,}6 & 0{,}6 & 0{,}6 & 0{,}4 & 0{,}4 \\ 0{,}6 & 0{,}6 & 0{,}6 & 0{,}6 & 0{,}6 & 0{,}6 \\ 0{,}8 & 0{,}8 & 0{,}8 & 0{,}8 & 0{,}8 & 0{,}8 \\ 1{,}0 & 1{,}0 & 1{,}0 & 1{,}0 & 1{,}0 & 1{,}0 \end{pmatrix}$$

$$= (1{,}0 \;\; 0{,}9 \;\; 0{,}8 \;\; 0{,}6 \;\; 0{,}4 \;\; 0{,}4)$$

Vậy, $_4B' = \bigcap_{1\le j\le n} A' \overset{T}{\circ} R_j$ = (1,0 1,0 1,0 1,0 1,0 1,0) $\cap$ (1,0 0,9 0,8 0,6 0,4 0,4)

= (1,0 0,9 0,8 0,6 0,4 0,4).

Bây giờ ta tính $_1B' = {}_3B'$:

$$_1B' = A' \overset{T}{\circ} \left(\bigcup_{1\le j\le n} R_j\right) = (1{,}0\quad 0{,}8\quad 0{,}6\quad 0{,}4\quad 0{,}2\quad 0{,}0) \circ \begin{pmatrix} 1{,}0 & 1{,}0 & 1{,}0 & 1{,}0 & 1{,}0 & 1{,}0 \\ 1{,}0 & 1{,}0 & 1{,}0 & 1{,}0 & 1{,}0 & 1{,}0 \\ 0{,}8 & 0{,}8 & 0{,}8 & 0{,}8 & 0{,}8 & 0{,}8 \\ 0{,}6 & 0{,}6 & 0{,}6 & 0{,}6 & 0{,}6 & 0{,}6 \\ 0{,}8 & 0{,}8 & 0{,}8 & 0{,}8 & 0{,}8 & 0{,}8 \\ 1{,}0 & 1{,}0 & 1{,}0 & 1{,}0 & 1{,}0 & 1{,}0 \end{pmatrix}$$

= (1,0 1,0 1,0 1,0 1,0 1,0)

Lưu ý: Trong nhiều bài toán, chúng ta đòi hỏi dữ liệu đầu ra là giá trị thực nên chúng ta cần áp dụng một phương pháp khử mờ nào đó để chuyển dữ liệu đầu ra là tập mờ về giá trị thực. Trong ví dụ trên, nếu cần thiết ta có thể biến đổi tập mờ của dữ liệu đầu ra thành giá trị thực thuộc miền tham chiếu [134;140] của biến ngôn ngữ áp suất bằng một phương pháp khử mờ được trình bày trong Mục 3.3.10.

2.5.3. Phương pháp lập luận xấp xỉ đa điều kiện, nhiều biến

Trong mục các trên chúng ta đã nghiên cứu các phương pháp lập luận xấp xỉ trong đó phần tiền tố của luật (mệnh đề nếu-thì) chí có một biến ngôn ngữ. Trong mục này chúng ta sẽ đề cập đến phương pháp lập luận trên các luật mà phần tiền tố có nhiều biến ngôn ngữ tham gia và chúng được liên kết lôgic bằng các phép VÀ hay HOẶC. Như vậy, chúng ta có thể có các trường hợp sau:

Dạng tiền tố hội: NẾU X_1 là A_1 VÀ … VÀ X_m là A_m THÌ Y là B

Phương pháp lập luận đối với dạng này ta có thể được xây dựng bằng việc đưa về phương pháp đối với trường hợp tiền tố chỉ có một biến, nhờ thay tiền tố nhiều biến bằng mệnh đề X^* là A^*, với

$$A^* = A^*_1 \cap A^*_2 \cap \dots \cap A^*_m \qquad (79^*)$$

trong đó, A^*_i là mở rộng hình trụ của tập mờ A_i trong tích Đề-các $\boldsymbol{U}_1 \times \dots \times \boldsymbol{U}_m$, nghĩa là A^*_i là một tích Đề-các chỉ có riêng thành phần thứ i là A_i còn các thành phần còn lại là toàn không gian $\boldsymbol{U}_j$, $j \neq i$, và hàm thuộc của nó là

$$A^*(u_1, u_2, \dots, u_m) = \min\{A_1(u_1), A_2(u_2), \dots, A_m(u_m)\}.$$

Dạng tiền tố tuyển: NẾU X_1 là A_1 HOẶC … HOẶC X_m là A_m THÌ Y là B .

Tương tự như trên, nhưng A^* được tính theo công thức

$$A^* = A^*_1 \cup A^*_2 \cup \dots \cup A^*_m \qquad (80^*)$$

với hàm thuộc là $A^*(u_1, u_2, \dots, u_m) = \max\{A_1(u_1), A_2(u_2), \dots, A_m(u_m)\}$.

Mệnh đề với NẾU KHÔNG và TRỪ KHI

Những mệnh đề điều kiện có chứa NẾU KHÔNG hay TRỪ KHI có cấu trúc lôgic cho phép chuyển về dạng quen biết và do đó chúng ta có thể sử dụng các phương pháp lập luận xấp xỉ đã trình bày ở trên.

(1) Mệnh đề

NẾU X là A THÌ (Y là B_1 NẾU KHÔNG B_2)

có thể phân tách thành các mệnh đề điều kiện quen biết được liên kết với nhau bằng HOẶC như sau:

NẾU X là A THÌ Y là B_1

HOẶC

NẾU X là KHÔNG A THÌ Y là B_2

(2) Mệnh đề

NẾU X là A_1 THÌ Y là B TRỪ KHI X là A_2

cũng như trên, có thể phân tách thành các mệnh đề điều kiện quen biết được liên kết với nhau bằng HOẶC như sau:

NẾU X là A_1 THÌ Y là B

HOẶC

NẾU X là A_2 THÌ Y là KHÔNG B

(3) Mệnh đề

NẾU X_1 là A_1 THÌ Y là B NẾU KHÔNG (NẾU X_2 là A_2 THÌ Y là B_2)

có thể phân tách thành

NẾU X là A_1 THÌ Y là B

HOẶC

NẾU X là KHÔNG A_1 VÀ X_2 là A_2 THÌ Y là B_2

Dạng mệnh đề điều kiện kết tổ

Trong thực tế ta cũng thường gặp các mệnh đề dạng sau

NẾU X_1 là A_1 THÌ (NẾU X_2 là A_2 THÌ Y là B)

có thể viết thành mệnh đề dạng sau

NẾU X_1 là A_1 VÀ X_2 là A_2 THÌ Y là B

2.5.4. Phương pháp lập luận xấp xỉ bằng đồ thị

Phương pháp lập luận bằng đồ thị không có ý nghĩa trong tính toán máy tính nhưng nó có ý nghĩa cho việc chúng ta trực tiếp tính toán bằng tay trong việc thiết lập các ví dụ để kiểm tra tính đúng đắn của lôgic chương trình máy tính và cho chính việc trình bày và lĩnh hội nội dung của giáo trình này.

Xét một mô hình mờ nhiều biến hay còn gọi là một hệ luật sau (để đơn giản trong trình bày chúng ta giới hạn chỉ 2 biến đầu vào):

NẾU X_{1j} là A_{1j} VÀ X_{2j} là A_{2j} THÌ Y là B_j, $j = 1, 2, \ldots, n$. (81*)

Chúng ta giới hạn phương pháp lập luận đồ thị với những giả thiết sau:

- Các tập mờ đều ở dạng tam giác hay hình thang;

- Các luật trong (81*) liên kết bằng phép tuyển hoặc hội;
- Quan hệ mờ được định nghĩa dựa trên phép kéo theo Mamdani;
- Phép hợp thành là sup-min.

Với những giả thiết trên ta có thể xây dựng một phương pháp lập luận đồ thị khá đơn giản và dễ dàng thực hiện tính toán trực tiếp bằng tay.

Trước hết, chúng ta hãy thiết lập công thức tính tập mờ kết luận khi cho biết tập mờ đầu vào A_1' và A_2'.

Với phép hợp thành sup-min:

$B'(v) = (A_1' \times A_2') \text{ o } R(u_1, u_2, v)$

$= \sup_{u1 \in \boldsymbol{U}1,\, u2 \in \boldsymbol{U}2} \min\{\min[A_1'(u_1), A_2'(u_2)], \max_{1 \leq j \leq n} R_j(u_1, u_2, v)\}$

$= \sup_{u1 \in \boldsymbol{U}1,\, u2 \in \boldsymbol{U}2} \max_{1 \leq j \leq n} \min\{\min[A_1'(u_1), A_2'(u_2)], R_j(u_1, u_2, v)\}$

$= \max_{1 \leq j \leq n} \sup_{u1 \in \boldsymbol{U}1,\, u2 \in \boldsymbol{U}2} \min\{\min[A_1'(u_1), A_2'(u_2)], R_j(u_1, u_2, v)\}$

$= \max_{1 \leq j \leq n} (A_1' \times A_2') \text{ o } R_j(u_1, u_2, v) = \max_{1 \leq j \leq n} B'_j$ (82*)

Công thức (82*) chứng tỏ rằng, với những điều kiện phép hợp thành là sup-min và liên kết các luật là tuyển, thì kết luận B' có thể được tính theo kết quả lập luận đối với từng luật: $B'_j = (A_1' \times A_2') \text{ o } R_j(u_1, u_2, v)$, trong đó $R_j(u_1, u_2, v)$ là quan hệ mờ biểu diễn ngữ nghĩa của luật thứ j.

Ta hãy viết tường minh biểu thức giải tích của (82*) với R_j tính theo Mamdani:

$B'(v) = \max_{1 \leq j \leq n} \sup_{u1 \in \boldsymbol{U}1,\ u2 \in \boldsymbol{U}2} \min\{\min[A_1'(u_1), A_2'(u_2)], \min[A_{1j}(u_1), A_{2j}(u_2), B_j(v)]\}$

$= \max_{1 \leq j \leq n} \sup_{u1 \in \boldsymbol{U}1,\ u2 \in \boldsymbol{U}2} \min\{\min[A_1'(u_1), A_{1j}(u_1)], \min[A_2'(u_2), A_{2j}(u_2)], B_j(v)\}$

$= \max_{1 \leq j \leq n} \min\{\sup_{u1 \in \boldsymbol{U}1} \min[A_1'(u_1), A_{1j}(u_1)], \sup_{u2 \in \boldsymbol{U}2} \min[A_2'(u_2), A_{2j}(u_2)], B_j(v)\}$ (83*)

Phân tích biểu thức chứa sup trong vế phải của đẳng thức cuối cùng trong (82*) ta thấy, biểu thức $\sup_{u1 \in \boldsymbol{U}1} \min[A_1'(u_1), A_{1j}(u_1)]$ xác định chiều cao của tập mờ $A_1' \cap A_{1j}$, $high(A_1' \cap A_{1j})$. Một cách tương tự, $\sup_{u2 \in \boldsymbol{U}2} \min[A_2'(u_2),$

$A_{2j}(u_2)$] xác định chiều cao $high(A_2' \cap A_{2j})$ của tập mờ $A_2' \cap A_{2j}$. Do vậy, (83*) trở thành biểu thức sau:

$$B'(v) = \max_{1 \leq j \leq n} \min\{\min[high(A_1' \cap A_{1j}), high(A_2' \cap A_{2j})], B_j(v)\} \quad (84^*)$$

Ký hiệu $h_j = \min[high(A_1' \cap A_{1j}), high(A_2' \cap A_{2j})]$, biểu thức $\min[h_j, B_j(v)]$, với $\forall v \in \boldsymbol{V}$, xác định phần của tập mờ B_j, $\underline{B}_j$, bị cắt cụt ngọn, có chiều cao còn lại là h_j. Vậy, B' được tính theo công thức sau

$$B' = \underline{B}_1 \cup \underline{B}_2 \cup \ldots \cup \underline{B}_n \quad (85^*)$$

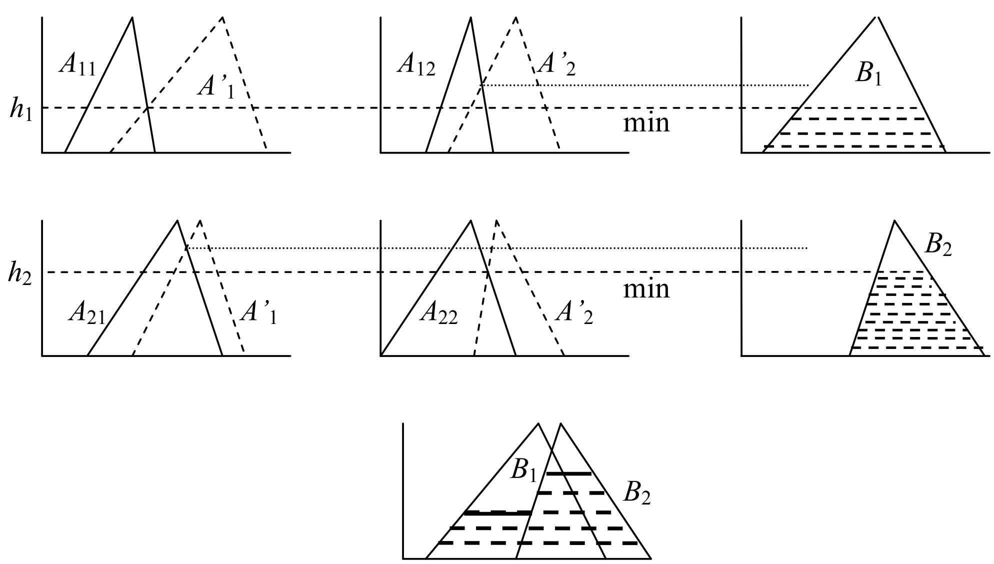
Vì các tập mờ ở dạng tam giác hay hình thang, biểu thức (85*) cho ta một phương pháp lập luận bằng đồ thị có thể tính trực tiếp bằng tay như sau:

Bước 1. Với mỗi $j = 1, \ldots, n$,

(i) Tính chiều cao của các hình tam giác hay hình thang $A_1' \cap A_{1j}$ và $A_2' \cap A_{2j}$. Lấy h_j là chiều cao thấp nhất trong các chiều cao đã tính.

(ii) Cắt phần ngọn của hình tam giác hay hình thang B_j sao cho phần còn lại $\underline{B}_j$ của nó có chiều cao là h_j.

Bước 2. Lấy hợp của các tập hợp $\underline{B}_j$, $j = 1, \ldots, n$, ta thu được tập mờ kết luận B'.



Hình 2.9

Hình 2.9 là một ví dụ giải tích cách tính tập mờ kết quả dựa trên phương pháp lập luận đồ thị. Ở ví dụ này, hệ luật (81*) có 2 luật và dữ liệu đầu vào của hệ là cặp tập mờ tam giác (A'_1, A'_2). Ứng với Bước 1, đối với $j = 1$, ta lấy giao của hai tam giác A_{11} và A'_1 và giao của A_{12} và A'_2, và min của chiều cao của hai tam giác thu được là h_1. Cắt ngọn tam giác B_1 ta thu được hình thang có chiều cao là h_1 được tô bằng các đường gạch song song với hai đáy. Một cách tương tự, đối với $j = 2$, ta thu được hình thang với chiều cao là h_2 bằng cách cắt ngọn tam giác B_2. Hợp của hai hình thang kết quả là hình được đánh dấu bằng các đường gạch song song với cạnh đáy. Nếu cần thiết biến đổi tập mờ kết quả về giá trị thức, ta có thể sử dụng một trong các phương pháp khử mờ được trình bày ở phần sau.

Chú ý: Trong trường hợp dữ liệu đầu vào là các tam giác (A'_1, A'_2) suy biến thành các giá trị thực, hàm thuộc của chúng là hàm đặc trưng chỉ khác không tại giá trị thực đó thì các bước của phương pháp lập luận đồ thị vẫn vận dụng được đúng đắn.

Chương 3

LẬP LUẬN NGÔN NGỮ VÀ THAO TÁC DỮ LIỆU MỜ

3.1. Đại số gia tử

Vấn đề sử dụng tập mờ để biểu diễn các giá trị ngôn ngữ và dùng các phép toán trên tập mờ để biểu thị các gia tử ngôn ngữ như $\mu_{\textit{rất trẻ}} = (\mu_{\textit{trẻ}})^2$, $\mu_{\textit{ít nhiều trẻ}} = (\mu_{\textit{trẻ}})^{1/2}$….đã cho phép thực hiện các thao tác dữ liệu mờ, đáp ứng nhu cầu thực tế của con người. Tuy nhiên, theo cách sử dụng tập mờ ta thấy có nhiều nhược điểm do việc xây dựng các hàm thuộc và xấp xỉ các giá trị ngôn ngữ bởi các tập mờ còn mang tính chủ quan, phụ thuộc nhiều vào ý kiến chuyên gia cho nên dễ mất mát thông tin. Mặc khác, bản thân các giá trị ngôn ngữ có một cấu trúc thứ tự nhưng ánh xạ gán nghĩa sang tập mờ, không bảo toàn cấu trúc đó nữa.

Do đó, vấn đề đặt ra là có một cấu trúc toán học mô phỏng chính xác hơn cấu trúc ngữ nghĩa của một khái niệm mờ. N.C.Ho và cộng sự đã đưa ra ĐSGT và ĐSGT mở rộng và ĐSGT tuyến tính đầy đủ đã giải đáp đầy đủ cho câu hỏi này.

3.1.1 Một số khái niệm

Chúng ta xét miền ngôn ngữ của biến chân lý *TRUTH* gồm các từ sau: *Dom*(*TRUTH*) = {*true, false, very true, very false, more-or-less true, more-or-less false, possibly true, possibly false, approximately true, approximately false, little true, little false, very possibly true, very possibly false.....*}, trong đó *true, false* là các từ nguyên thuỷ, các từ nhấn (modifier or intensifier) *very*, *more-or-less*, *possibly*, *approximately*, *little* gọi là các gia tử (hedges).

Khi đó, miền ngôn ngữ T = *dom*(*TRUTH*) có thể biểu thị như một đại số $\underline{X} = (\boldsymbol{X}, G, H, \leq)$, trong đó *G* là tập các từ nguyên thuỷ được xem là các phần tử sinh. $H = H^- \cup H^+$ với H^+ và H^- tương ứng là tập các gia tử dương, âm và được xem như là các phép toán một ngôi, quan hệ ≤ trên các từ (các khái niệm mờ) là quan hệ sắp thứ tự tuyến tính trên $\boldsymbol{X}$ cảm sinh từ ngữ nghĩa của ngôn ngữ. Ví dụ dựa trên ngữ nghĩa, các quan hệ thứ tự sau là đúng: *false*

$\leq$ *true, more true* $\leq$ *very true* nhưng *very false* $\leq$ *more false, possibly true* $\leq$ *true nhưng false* $\leq$ *possibly false* ... . Tập ***X*** được sinh ra từ *G* bởi các phép toán trong *H*. Như vậy mỗi phần tử của ***X*** sẽ có dạng biểu diễn $x = h_n h_{n-1} \ldots\ldots h_1 c$, $c \in G$. Tập tất cả các phần tử được sinh ra từ một phần tử *x* được ký hiệu là *H(x)*. Nếu *G* có đúng hai từ *nguyên thuỷ mờ*, thì một được gọi là *phần tử sinh dương* ký hiệu là c^+, một gọi là *phần tử sinh âm* ký hiệu là c^- và ta có $c^- < c^+$. Trong ví dụ trên *true* là *phần tử sinh dương* còn *false* là *phần tử sinh âm*.

Về mối quan hệ giữa các gia tử chúng ta có các khái niệm sau:

(1) : Mỗi gia tử hoặc là dương, hoặc là âm đối với bất kỳ một gia tử nào khác, kể cả chính nó.

(2) : Nếu hai khái niệm *u* và *v* độc lập, nghĩa là $u \notin H(v)$ và $v \notin H(u)$ thì $\forall x \in H(u)$ ta có $x \notin H(v)$. Ngoài ra nếu *u* và *v* là không sánh được thì bất kỳ $x \in H(u)$ cũng không sánh được với bất kỳ $y \in H(u)$.

(3) : Nếu $x \neq hx$ thì $x \notin H(hx)$ và nếu $h \neq k$ và $hx \leq kx$ thì $h'hx \leq k'kx$ với mọi gia tử *h*, *k*, *h'*, *k'*. Hơn nữa $hx \neq kx$ thì *hx* độc lập với *kx*.

(4) : Nếu $u \notin H(v)$ và $u \leq v$ ($u \geq v$) thì $u \leq hv$ ($u \geq hv$), đối với mọi gia tử *h*.

Định nghĩa trên mới chỉ dựa vào các tính chất ngữ nghĩa và di truyền ngữ nghĩa của ngôn ngữ nhưng đã tạo ra cấu trúc đủ giàu để xây dựng các quan hệ đối sánh trong mô hình CSDL mờ.

Tiếp theo là định lý thể hiện ý nghĩa trực quan trong ngôn ngữ về tính chất di truyền ngữ nghĩa của ngôn ngữ.

Định lý 3.1. Giả sử $x = h_n \ldots h_1 u$ và $y = k_m \ldots k_1 u$ là các biểu diễn chính tắc của *x* và *y* đối với *u*. Khi đó tồn tại một chỉ số $j \leq \min\{m,n\} + 1$ sao cho với mọi $i < j$ ta có $h_i = k_i$ và

(1) $x < y$ khi và chỉ khi $h_j x_j < k_j x_j$, trong đó $x_j = h_{j-1} \ldots h_1 u$;

(2) $x = y$ khi và chỉ khi $n = m = j$ và $h_j x_j = k_j x_j$;

(3) *x* và *y* là không sánh được khi và chỉ khi $h_j x_j$ và $k_j x_j$ là không sánh được.

Vì tất cả các thuộc tính có miền trị chứa giá trị số trong CSDL đều tuyến tính, nên một cách tự nhiên ta giả thiết trong chương này, ĐSGT được sử dụng là ĐSGT tuyến tính, do đó tập H^+ và H^- là tập sắp thứ tự tuyến tính. Như vậy, cho $\underline{X} = (\, \boldsymbol{X}, G, H, \leq \,)$ với $G = \{0, c^-, W, c^+, 1\}$, $H = H^- \cup H^+$ với giả thiết $H^- = \{h_1, h_2, .., h_p\}$, $H^+ = \{h_{-1}, ..., h_{-q}\}$, $h_1 > h_2 > ... > h_p$ và $h_{-1} < ... < h_{-q}$ là dãy các gia tử, ta có các định nghĩa liên quan như sau :

Định nghĩa 3.1. Cho $\underline{X} = (\, \boldsymbol{X}, G, H, \leq \,)$ là một ĐSGT, với mỗi $x \in \boldsymbol{X}$, độ dài của x được ký hiệu $|x|$ và xác định như sau:

(1) Nếu $x = c^+$ hoặc $x = c^-$ thì $|x| = 1$.

(2) Nếu $x = hx'$ thì $|x| = 1 + |x'|$, với mọi $h \in H$.

Định nghĩa 3.2. Hàm fm: $\boldsymbol{X} \rightarrow [0,1]$ được gọi là độ đo tính mờ trên $\boldsymbol{X}$ nếu thoả mãn các điều kiện sau:

(1) fm là độ đo mờ đầy đủ trên $\boldsymbol{X}$, tức là $\sum_{-q \leq i \leq p, i \neq 0} fm(h_i u) = fm(u)$ với mọi $u \in \boldsymbol{X}$.

(2) Nếu x là khái niệm rõ, tức là $H(x) = \{x\}$ thì $fm(x) = 0$. Do đó $fm(\boldsymbol{0}) = fm(\boldsymbol{W}) = fm(\boldsymbol{1}) = 0$.

(3) Với mọi $x, y \in \boldsymbol{X}$ và $h \in H$ ta có $\dfrac{fm(hx)}{fm(x)} = \dfrac{fm(hy)}{fm(y)}$, nghĩa là tỉ số này không phụ thuộc vào x và y, được kí hiệu là $\mu(h)$ gọi là độ đo tính mờ (*fuzziness measure*) của gia tử h.

Trong đại số gia tử, mỗi phần tử $x \in \boldsymbol{X}$ đều mang dấu âm hay dương, được gọi là *PN*-dấu và được định nghĩa đệ quy như sau:

Định nghĩa 3.3. Hàm *Sgn*: $\boldsymbol{X} \rightarrow \{-1, 0, 1\}$ là một ánh xạ được định nghĩa một cách đệ qui như sau, với $\forall h, h' \in H$, $c \in \{c^+, c^-\}$:

(1) $Sgn(c^-) = -1$ và $Sgn(c^+) = +1$.

(2) $Sgn(h'hx) = -Sgn(hx)$ nếu h' là *negative* với h và $h'hx \neq hx$.

(3) $Sgn(h'hx) = Sgn(hx)$ nếu h' là *positive* với h và $h'hx \neq hx$.

(4) $Sgn(h'hx) = 0$ nếu $h'hx = hx$.

Mệnh đề 3.1. Với $\forall x \in \boldsymbol{X}$, ta có: $\forall h \in H$, nếu $Sgn(hx) = +1$ thì $hx > x$, nếu $Sgn(hx) = -1$ thì $hx < x$ và nếu $Sgn(hx) = 0$ thì $hx = x$.

Để chuyển đổi một giá trị trong ĐSGT (giá trị ngôn ngữ) thành một số trong [0,1] ta sử dụng hàm định lượng ngữ nghĩa.

Định nghĩa 3.4. Cho *fm* là độ đo tính mờ trên $\boldsymbol{X}$, hàm định lượng ngữ nghĩa υ trên $\boldsymbol{X}$ được định nghĩa như sau :

(1) $\upsilon(\boldsymbol{W}) = \theta = fm(c^-)$, $\upsilon(c^-) = \theta - \alpha.fm(c^-)$ và $\upsilon(c^+) = \theta + \alpha.fm(c^+)$

(2) Nếu $1 \le j \le p$ thì $\upsilon(h_j x) = \upsilon(x) + Sgn(h_j x)\left[\sum_{i=1}^{j} fm(h_i x) - \omega(h_j x) fm(h_j x)\right]$

Nếu $-q \le j \le -1$ thì $\upsilon(h_j x) = \upsilon(x) + Sgn(h_j x)\left[\sum_{i=j}^{-1} fm(h_i x) - \omega(h_j x) fm(h_j x)\right]$

trong đó $\omega(h_j x) = \frac{1}{2}\left[1 + Sgn(h_j x) Sgn(h_q h_j x)(\beta - \alpha)\right] \in \{\alpha, \beta\}$.

3.1.2 Các tính chất của độ đo tính mờ trong ĐSGT

Dựa trên cấu trúc của ĐSGT, trong đó quan hệ giữa các phần tử là quan hệ thứ tự ngữ nghĩa, mô hình toán học của tính mờ và độ đó tính mờ của các khái niệm mờ đã được định nghĩa trong các công trình của N.C.Ho và cộng sự, ở đây chúng tôi chỉ trình bày một số mệnh đề và bổ đề liên quan đến hàm *fm* và hàm υ

Mệnh đề 3.2

(1) $fm(hx) = \mu(h) fm(x)$, với $\forall x \in \boldsymbol{X}$

(2) $fm(c^-) + fm(c^+) = 1$

(3) $\sum_{-q \le i \le p, i \ne 0} fm(h_i c) = fm(c)$, trong đó $c \in \{c^-, c^+\}$

(4) $\sum_{-q \le i \le p, i \ne 0} fm(h_i x) = fm(x)$, *với* $\forall x \in \mathbf{X}$

(5) $\sum_{i=-q}^{-1}\mu(h_i)=\alpha$ và $\sum_{i=1}^{p}\mu(h_i)=\beta$, với $\alpha, \beta > 0$ và $\alpha + \beta = 1$.

Bổ đề 3.1. Cho *fm* là hàm độ đo tính mờ trên ***X*** và hàm định lượng ngữ nghĩa υ trên ***X*** gắn với *fm*. Khi đó tồn tại một phân hoạch Δ gắn với *fm* sao cho phát biểu sau là đúng, với $\forall x \in \boldsymbol{X}$: $\upsilon(x) \in I(x)$ và $\upsilon(x)$ chia đoạn $I(x)$ thành hai đoạn con tỷ lệ $\alpha:\beta$. Và nếu $Sgn(h_1x) = 1$ thì đoạn con tương ứng với α lớn hơn đoạn con tương ứng với β và nếu $Sgn(h_1x) = -1$ thì đoạn con tương ứng với α nhỏ hơn đoạn con tương ứng với β.

Định lý 3.2. Cho $\underline{X} = (\boldsymbol{X}, G, H, \leq)$ là ĐSGT tuyến tính. Ta có các phát biểu sau:

(1) Với $\forall x \in \boldsymbol{X}$, $H(x)$ là tập sắp thứ tự tuyến tính.

(2) Nếu G là tập sắp thứ tự tuyến tính thì $H(G)$ cũng là tập sắp thứ tự tuyến tính.

Trong ĐSGT tuyến tính, bổ sung thêm vào hai phép tính Σ và Φ với ngữ nghĩa là cận trên đúng và cận dưới đúng của tập $H(x)$, khi đó ĐSGT tuyến tính được gọi là ĐSGT tuyến tính đầy đủ.

Cho một ĐSGT tuyến tính đầy đủ $\mathcal{AX} = (\boldsymbol{X}, G, H, \Sigma, \Phi, \leq)$, trong đó $Dom(\mathcal{X}) = \boldsymbol{X}$ là miền các giá trị ngôn ngữ của thuộc tính ngôn ngữ $\mathcal{X}$ được sinh từ tập các phần tử sinh $G = \{\boldsymbol{0}, c^-, \boldsymbol{W}, c^+, \boldsymbol{1}\}$ bằng việc tác động các gia tử trong tập H, Σ và Φ là hai phép tính với ngữ nghĩa là cận trên đúng và cận dưới đúng của tập $H(x)$, tức là $\Sigma x = supremum\ H(x)$ and $\Phi x = infimum\ H(x)$, quan hệ $\leq$ là quan hệ sắp thứ tự tuyến tính trên ***X*** cảm sinh từ ngữ nghĩa của ngôn ngữ.

3.2. Các phương pháp lập luận ngôn ngữ

3.2.1. Lập luận bằng các siêu luật

Trong phần này, chúng tôi giới thiệu phương pháp lập luận sử dụng các siêu luật đã được tác giả N.C.Ho giới thiệu. Trong các phần tiếp theo so sánh suy diễn mờ với lập luận ngôn ngữ để chứng tỏ có thể sử dụng cấu trúc đại số gia tử cho lập luận mà kết quả không khác nhiều với suy diễn mờ.

3.2.1.1. Mệnh đề mờ và cơ sở tri thức mờ

Trước hết dễ nhận thấy rằng con người sử dụng các câu trong ngôn ngữ kết hợp với một giá trị chân lý ngôn ngữ chỉ sự tin cậy về mức độ đúng đắn của các câu đó để biểu thị tri thức của mình. Ví dụ trong tri thức của chúng ta có câu "Trường Harvard *rất nỗi tiếng*" trong một khung cảnh nhất định có giá trị chân lý là "rất đúng". Như vậy, thành tố cơ bản của tri thức là một cặp gồm một câu nói chung có chứa các khái niệm mờ và một giá trị chân lý ngôn ngữ. Một cách hình thức mỗi thành tố cơ bản của tri thức sẽ được ký hiệu bằng một cặp A=(P(x,u),t), trong đó P(x,u) là mệnh đề mờ với biến đối tượng x, u là một khái niệm mờ và t là giá trị chân lý ngôn ngữ. Có thể giả thiết rằng t là một khái niệm sinh ra từ khái niệm *đúng* và ta gọi đó là một khẳng định.

Chúng ta xét câu sau đây : "Lan *rất già*", "Phần X của động cơ phụ thuộc *mạnh* vào phần Y" và "Nếu một sinh viên học ở trường đại học *có uy tín* và anh ta học *chăm chỉ* thì sẽ trở thành một nhân viên *tốt*".

Trong các câu trên, nếu loại bỏ các khái niệm mờ thì phần còn lại có thể xem là các mệnh đề theo nghĩa kinh điển. Vì vậy, một cách ký hiệu có thể viết cho các câu trên như sau: TUOI(Lan, *rất già*), PHUTHUOC(X, Y, *mạnh*), TRUONG(x, *có uy tín*) và HOC(x, *chăm chỉ*)→NHANVIEN(x, *tốt*).

Với cách hình thức hóa các mệnh đề mờ được xem là một khái quát trực tiếp từ các mệnh đề kinh điển. Vì vậy, cũng như đối với ngôn ngữ logic vị từ, ta kí hiệu các mệnh đề mờ bằng các chữ cái in hoa F, P, Q, R, S....Các khẳng định được kí hiệu là A với chỉ số khi cần. Vì tri thức thu nạp từ nhiều nguồn và trong những khung cảnh khác nhau, nên một cách tổng quát ta giả thiết một mệnh đề mờ có thể có nhiều giá trị chân lý ngôn ngữ. Một tập hữu hạn hay đếm được các khẳng định được gọi là *một cơ sở tri thức* và kí hiệu là K. Vì ở đây các mệnh đề chứa các khái niệm mờ nên có thể dùng từ *cơ sở tri thức mờ*. Một cơ sở tri thức K gọi là nhất quán nếu nó không chứa hai khẳng định dạng (P,t) và (¬P,s) với t>W và s>W.

Bài toán lập luận phát biểu như sau: Cho một cơ sở tri thức K, hãy tìm những khẳng định suy ra được từ K.

3.2.1.2. Lập luận ngôn ngữ bằng các siêu luật

Khác với phương pháp lập luận mờ, ở đây tác giả xây dựng các luật suy diễn và lập luận bằng dãy chứng minh, tương tự như logic kinh điển. Một cách tổng quát các luật suy diễn có dạng:

$$\frac{(P_1, t_1),........(Q_n, t_n)}{(Q_1, s_1),........(Q_m, t_m)}$$

trong đó (P_i,t_i), t_i>W, i=1..n là các tiên đề, (Q_i,s_i), s_j>W, j=1..m là các kết luận. Sau đây là các siêu luật suy diễn cho phép thao tác trực tiếp các giá trị ngôn ngữ.

Luật chuyển gia tử

$$(RT_1)\ \frac{((P, hu), \sigma T)}{((P, u), \sigma hT)} \qquad (RT_2)\ \frac{((P, u), \sigma hT)}{((P, hu), \sigma T)}$$

ở đây σ là xâu các gia tử và h là một gia tử nào đó, T∈{*đúng*, *sai*}

Luật chuyển gia tử trong mệnh đề

$$(RT_1)\ \frac{(hP, \sigma T)}{(P, \sigma hT)} \qquad (RT_2)\ \frac{(P, \sigma hT)}{(hP, \sigma T)}$$

Luật chuyển gia tử trong mệnh đề kéo theo

$$(RTI_1)\frac{(hP \rightarrow hQ, \alpha\ \text{đúng}), (hP, \sigma\ \text{đúng})}{(P \rightarrow Q, \alpha h\ \text{đúng})}$$

$$(RTI_2)\qquad \frac{(P \rightarrow Q, \alpha h\ \text{đúng}), (P, \sigma h\ \text{đúng})}{(hP \rightarrow hQ, \alpha\ \text{đúng})}$$

Luật modus ponens

$$(RMP)\qquad \frac{(P \rightarrow Q, \alpha\ \text{đúng}), (P, \text{đúng})}{(Q, \alpha\ \text{đúng})}$$

Luật modus tonens

$$(RMT)\qquad \frac{(P \rightarrow Q, \alpha\ \text{đúng}), (\neg Q, \text{đúng})}{(\neg P, \alpha\ \text{đúng})}$$

Luật kéo theo tỉ lệ

$$(RPI)\qquad \frac{(P(x, u) \rightarrow Q(x, v), \sigma\ \text{đúng})}{(\alpha P(x, u) \rightarrow \alpha Q(x, v), \sigma\ \text{đúng})}$$

trong đó α và σ là các xâu gia tử, P và Q là các công thức phân phối được đối với gia tử.

Luật phép thế

$$\text{(RSUB)} \qquad \frac{P(x, u)}{P(a, u)}$$

Trong đó x là biến cá thể, u là hằng cá thể

Luật thế công thức tương đương

$$\text{(RE)} \qquad \frac{P \leftrightarrow Q,\ (F(P), \alpha T)}{(F(Q), \alpha T)}$$

trong đó kí hiệu F(X) chỉ X là công thức con của F.

Cho K là một cơ sở tri thức, ta nói một khẳng định (P,t) suy được từ K nếu có tồn tại một dãy chứng minh (P_1,t_1)........(P_n,t_n) từ các khẳng định trong K nhờ việc sử dụng các luật (RT), (RTI_1), (RTI_2), (RMP), (RMT), (RPI), (RSUB) và (RE) sao cho (P_n,t_n)=(P,t). Khi đó ta có kí hiệu K |--- (P,t) và ta đặt C(K)={(P,t) : K|--- (P,t)}.

Sau đây chúng ta có định lý chỉ ra tính đúng đắn của phương pháp lập luận trên.

Định lý 3.3. Giả sử K là một cơ sở tri thức, khi đó :

(1) Nếu K |--- (P,t) thì t>W.

(2) Nếu K nhất quán thì C(K) cũng nhất quán.

(3) Nếu có tồn tại một phép gán 2-trị cho K thì K là nhất quán. Một phép gán 2-trị *val* chỉ có các giá trị 0 và 1, nghĩa là $\forall$(P,t), val(P)=1 (t>W).

Như vậy, có thể thấy rằng logic các giá trị ngôn ngữ có thể làm cơ sở cho phương pháp lập luận bằng các siêu luật suy diễn được đưa ra ở trên. Sau đây chúng ta đưa ra một ví dụ mô tả cho phương pháp này.

Ví dụ 3.1. Giả sử trong cơ sở tri thức có các công thức sau :

- Nếu một sinh viên học *chăm chỉ* và trường anh ta *có uy tín* thì sẽ là một nhân viên *tốt* là *đúng*.
- Trường của An *rất có uy tín* là *có thể đúng*.
- Lan học *tương đối chăm chỉ*.

Có thể rút được những thông tin gì từ những khẳng định trên trong cơ sở tri thức. Dựa vào phương pháp trên ta cần xây dựng một dãy các chứng minh sau, kí hiệu

p(x,chăm chỉ)	cho mệnh đề	*x học chăm chỉ*
q(U(x),có uy tín)	cho mệnh đề	*Trường của x có uy tín*
r(x,rất tốt)	cho mệnh đề	*x sẽ là một nhân viên tốt*

Khi đó ta hãy chứng minh :

(1) (q(U(An),rất có uy tín),có thể đúng) (theo giả thiết)

(2) (q(U(An),có thể rất có uy tín),đúng) (theo luật (RT))

(3) (p(U(An),tương đối chăm chỉ),đúng) (theo giả thiết)

(4) (p(x,chăm chỉ)$\wedge$q(U(x),có uy tín)$\rightarrow$r(x,tốt),đúng) (theo giả thiết)

(5) (q(U(x),có uy tín)$\rightarrow$(p(x,chăm chỉ)$\rightarrow$r(x,tốt),đúng) (theo luật (RE))

(6) (có thể rất q(U(An),có uy tín$\rightarrow$có thể rất(p(An),tương đối chăm chỉ)$\rightarrow$r(An,tốt)), đúng) (do (5) và các luật (RPI), (RSUB))

(7) (có thể rất(p(An,tương đối chăm chỉ)$\rightarrow$r(An,tốt)),đúng) (do (2),(6) và luật (RMP))

(8) ((p(An,tương đối chăm chỉ)$\rightarrow$r(An,tốt),có thể rất đúng) (do (7) và luật (RT_1))

(9) (p(An,tương đối chăm chỉ)$\rightarrow$r(An,tương đối tốt),có thể rất đúng) (do (8) và luật (PRI)).

(10) (r(An,tương đối tốt),có thể rất đúng) (do (3), (9) và luật (RMP))

(11) (r(An,có thể tương đối tốt),đúng) (do (10) và luật (RT))

(12) (r(An,tốt),có thể rất tương đối đúng) (do (11) và luật (RT))

3.2.1.3. So sánh suy diễn mờ và lập luận ngôn ngữ

Phương pháp lập luận ngôn ngữ có ưu điểm lớn là thao tác đơn giản, làm việc trực tiếp trên các giá trị ngôn ngữ mà không phải qua các bước trung gian như xây dựng hàm thuộc, quan hệ mờ, khử mờ.....độ tin cậy của kết quả chỉ phụ thuộc vào lựa chọn giá trị ngôn ngữ và các gia tử. Quan trọng hơn là sẽ cho một kết quả dưới dạng một ngôn ngữ tự nhiên mà không phải qua bước xấp xỉ ngôn ngữ. Đối với phương pháp suy diễn mờ phải thực hiện các bước trung gian nên có thể đem lại sai số lớn.

3.2.1.4. Nhận xét phương pháp lập luận ngôn ngữ bằng các siêu luật

Lập luận ngôn ngữ cho kết quả tương đương với suy diễn mờ và được coi là thỏa mãn các tiêu chuẩn suy diễn tốt của suy diễn mờ. Tuy nhiên, lập luận ngôn ngữ bằng các siêu luật vẫn còn một số hạn chế sau:

Giả sử ta có mệnh đề : "Quả chuối *chín* thì ăn *ngon*", dùng lập luận ngôn ngữ ta có "Quả chuối *không chín lắm* thì ăn *không ngon lắm*" là hợp lý, hoặc "Quả chuối *rất chín* thì ăn *rất ngon*" chấp nhận được, nhưng nếu "Quả chuối *rất rất chín* thì ăn *rất rất ngon*" sẽ không hợp lý cho lắm bới vì "chuối *rất rất chín*" đã chuyển sang dạng khác (chín nhũn). Rõ ràng khái niệm về *độ chín của quả chuối* có số phần tử sinh khác với *mức độ ngon của quả chuối* gây ra sự sai lệch.

Trong ví dụ 3.1 nếu đảo mệnh đề "Nếu một sinh viên học *chăm chỉ* và trường anh ta *có uy tín* thì sẽ là một nhân viên *tốt* là *đúng*" thành "Nếu trường của một sinh viên *có uy tín* và anh ta học *chăm chỉ* thì sẽ là một nhân viên *tốt* là *đúng* ", thì sẽ cho kết quả (r(An,tốt),tương đối có thể rất đúng) hay (r(An,tương đối có thể rất tốt),đúng) khác với kết quả trên. Do vậy, cần một phương pháp lập luận mới nhằm khắc phục những hạn chế này.

3.2.2. Phương pháp lập luận dựa trên đại số gia tử

Trong phần này, chúng tôi giới thiệu phương pháp lập luận xấp xỉ dựa trên đại số gia tử đối với bài toán đa điều kiện 1 biến và bài toán đa điều kiện nhiều biến.

3.2.2.1. Bài toán đa điều kiện, một biến

Cho mô hình mờ một biến gồm n mệnh đề IF-THEN như sau:

$$\begin{array}{l} \text{If } X=A_1 \text{ then } Y=B_1 \\ \text{If } X=A_2 \text{ then } Y=B_2 \\ \cdots\cdots\cdots \\ \text{If } X=A_n \text{ then } Y=B_n \end{array} \qquad (1^*)$$

Để giải quyết bài toán này, ta xem mỗi mệnh đề IF-THEN là một điểm và như vậy mô hình mờ diễn tả một đường cong C_f trong tích Decacs $\mathscr{X}$ và $\mathscr{Y}$, ở đây $\mathscr{X}$; $\mathscr{Y}$ là các miền ngôn ngữ và được xem như là đại số gia tử của X, Y. Như vậy, vấn đề lập luận mờ đối với "mô hình đã cho và một đầu vào A, tìm đầu ra B tương ứng của A" được xem như vấn đề nội suy đối với đường cong C_f trong $\mathscr{X}$ x $\mathscr{Y}$. Thuật toán sử dụng trong phương pháp này có thể mô tả như sau:

Thuật toán 3.1

Input **:** Mô hình mờ (1*), đầu vào A

Output **:** Giá trị của Y

(1) Xây dựng hàm định lượng ngữ nghĩa ν_X, ν_Y là các ánh xạ từ $\mathscr{X}$ và $\mathscr{Y}$ vào [0,1].

(2) Tác dụng của hàm ν_X và ν_Y là biến đổi giá trị ngữ nghĩa thành một giá trị thực thuộc đoạn [0,1] và sau đó biến đổi đường cong mờ C_f trong $\mathscr{X}$ x $\mathscr{Y}$ thành đường cong thực C_r trong [0,1]x[0,1].

(3) Sử dụng phương pháp nội suy tuyến tính để tính đầu ra tương ứng với mỗi đầu vào.

3.2.2.2. Bài toán đa điều kiện, nhiều biến

Cho mô hình mờ nhiều biến như sau:

MĐ1: If $X_1=A_{11}$ AND $X_2=A_{12}$ AND.....AND $X_n=A_{1n}$ then $Y=B_1$

MĐ2: If $X_1=A_{21}$ AND $X_2=A_{22}$ AND.....AND $X_n=A_{2n}$ then $Y=B_2$ (2*)

....

$MĐ_m$: If $X_1=A_{m1}$ AND $X_2=A_{m2}$ AND.....AND $X_n=A_{mn}$ then $Y=B_m$

$MĐ^*$: If $X_1=A_1$ AND $X_2=A_2$ AND.....AND $X_n=A_n$ Cần tính Y=B?

Để giải quyết bài toán này, chúng ta đặt

$\underline{A}_1$="$X_1=A_{11}$ AND $X_2=A_{12}$ AND.....AND $X_n=A_{1n}$" là tập mờ của biến $\underline{X}$, tương tự

$\underline{A}_2$="$X_1=A_{21}$ AND $X_2=A_{22}$ AND.....AND $X_n=A_{2n}$"

....

$\underline{A}_m$="$X_1=A_{m1}$ ND $X_2=A_{m2}$ AND.....AND $X_n=A_{mn}$"

$\underline{A}_0$="$X_1=A_1$ AND $X_2=A_2$ AND.....AND $X_n=A_n$"

Khi đó bài toán trở thành

$$\text{If } \underline{X}=\underline{A}_1 \text{ then } Y=B_1$$
$$\text{If } \underline{X}=\underline{A}_2 \text{ then } Y=B_2$$
$$\cdots\cdots\cdots\cdots\cdots\cdots$$
$$\text{If } \underline{X}=\underline{A}_n \text{ then } Y=B_m$$

$$\text{Cho } X=\underline{A}_0 \text{ tính } Y=B?$$

Như vậy bài toán trở thành dạng (1*), do đó có thể sử dụng phương pháp nội suy mờ. Muốn vậy phải tính độ "gần nhau giữa" $\underline{A}_0$ với các $\underline{A}_i$.

Định nghĩa 3.5. Giả sử $x,y \in X_1 \times X_2 \times \ldots\ldots \times X_n$. Ta định nghĩa $x < y$ khi và chỉ khi với mọi $x_i, y_i \in X_i$ thì $x_i < y_i$.

Định nghĩa 3.6. Khoảng cách giữa $x,y \in X_1 \times X_2 \times \ldots\ldots \times X_n$ được định nghĩa như sau:

$$\rho(x,y)=\sum_{i=1}^{n}\left|\nu(x_i)-\nu(y_i)\right|$$

Định nghĩa 3.7. Cho n bộ đại số gia tử $X_1, X_2, ..X_n$, khi đó ta gọi hàm tích hợp trên đại số gia tử là hàm có dạng $F : X_1 \times X_2 \times \ldots\ldots \times X_n \to [0,1]$ thỏa mãn các tính chất sau :

(1) $0 \le F(x) \le 1$ $\quad\quad \forall x \in X_1 \times X_2 \times \ldots\ldots \times X_n$.

(2) Nếu $x \le y$ thì $F(x) \le F(y)$ $\quad\quad \forall x,y \in X_1 \times X_2 \times \ldots\ldots \times X_n$.

Thuật toán 3.2

Input **:** Mô hình mờ (2*), đầu vào $A=(A_1, A_2, ...A_n)$

Output **:** Giá trị của Y

(1): Chuyển bài toán (2*) và dạng bài toán (1*)

(2): Với mỗi mệnh đề IF-THEN tính $\nu(X_i)$, $\nu(Y_i)$, tính $\nu^*(X_i)$

(3): Tính ρ_j của mệnh đề $MĐ^*$ với $MĐ^j$ với j=1..m

(4): Với mỗi mệnh đề j, xác định $F(X_j)$ tức là $F(X^j_1,X^j_2, ...X^j_n)$

Xác định F(A) tức là $F(A_1, A_2, ...A_n)$

(5): Xác định h,k sao cho: ρ_k và ρ_h là hai giá trị nhỏ nhất trong các ρ_j, $k \neq h$, j =1..m

Nếu $F(A) \in [F(X_h), F(X_k)]$ or $F(A) \in [F(X_k), F(X_h)]$ thì nội suy $[X_h, X_k]$ hoặc $[X_k, X_h]$, tương ứng

$$F(Y)=\nu(Y)=\nu_k(Y).\frac{F(A)-F(X_h)}{F(X_k)-F(X_h)}+\nu_h(Y).\frac{F(A)-F(X_k)}{F(X_h)-F(X_k)}$$

Nếu $F(A) \notin [F(X_h), F(X_k)]$ or $F(A) \notin [F(X_k), F(X_h)]$ thì

$F(Y)=\nu(Y)=\dfrac{F(X_k)+F(X_h)+F(X_l)}{3}$, với l là chỉ số thỏa mãn ρ_l là giá trị nhỏ nhất trong các giá trị ρ_j ($k \neq h \neq l$, $j = 1..m$).

3.2.3. Phương pháp lập luận trên đại số gia tử không thuần nhất

Theo tác giả L.X.Vinh, miền giá trị của biến ngôn ngữ mà các giá trị này có thể chứa *not so*, khi đó một cấu trúc được xây dựng gọi là đại số gia tử PN-không thuần nhất. Như vậy, đại số gia tử PN-không thuần nhất sẽ được xây dựng từ đại số gia tử $AH(X,C,H,\leq)$, ở đây H chứa *not so* (*not* hoặc N).

3.2.3.1. Hình thức hóa các mệnh đề

Mặc dù số từ ngữ của ngôn ngữ tự nhiên là hữu hạn nhưng khả năng biểu đạt của ngôn ngữ tự nhiên hầu như là vô hạn. Với một vài từ giàu thông tin chúng ta có thể mô tả nhiều trạng thái của sự vật. Chẳng hạn màu "xanh nhạt" của bầu trời ngày hôm nay và ngày hôm qua chắc chắn là không giống nhau. Do vậy, khi biểu đạt tri thức của mình bằng ngôn ngữ tự nhiên con người thường sử dụng chúng và các từ như thế gọi là các khái niệm mờ. Các câu chứa khái niệm mờ gọi là mệnh đề mờ. Ví dụ "An còn trẻ", "Sinh viên Lan học rất chăm"....là các mệnh đề mờ hay tổng quát là các vị từ mờ. Dưới dạng thể hiện của biến ngôn ngữ chúng có thể viết thành "Tuổi của An còn trẻ", "Việc học của Lan là rất chăm". Như vậy, một cách hình thức mỗi mệnh đề mờ cơ sở là một cặp (p,u) với p là một vị từ n-ngôi và u là một khái niệm mờ, chẳng hạn (Tuổi(An),trẻ), (Việc học(Sinh viên Lan),rất chăm).

Với mỗi vị từ p, tập các khái niệm mờ u của nó sẽ được nhúng vào một đại số gia tử PN-không thuần nhất hữu hạn đối xứng, kí hiệu $\{D_p, C_p, LH, \leq, -\}$, trong đó "-" là phép toán đối. Chẳng hạn đối với vị từ p=Tuổi(người) thì C_p={già, trẻ}, LH={rất, có thể, không, tương đối....}, khi đó các khái niệm mờ sẽ là {rất trẻ, có thể trẻ, tương đối già..}. Tập tất cả các khái niệm mờ tương ứng với vị từ p, kí hiệu là TER_p, được định nghĩa như sau:

Định nghĩa 3.8. TER_p là một bộ phận của D_p thỏa mãn các điều kiện sau:

(1) $C_p \subseteq TER_p$

(2) Nếu $u \in TER_p$ thì $hu \in TER_p, \forall h \in H$

(3) Nếu $u \in TER_p$ thì $-u \in TER_p$

Từ các mệnh đề cơ sở, bằng các phép toán logic như $\vee$, $\wedge$, $\neg$, $\rightarrow$ ta có thể xây dựng các mệnh đề phức tạp hơn. Kết quả thu được tập các công thức, kí hiệu FP và được định nghĩa như sau:

Định nghĩa 3.9. (1) Mệnh đề cơ sở $(p,u) \in FP$ với mọi $u \in TER_p$. Với $P = (p,u)$, $h \in H$ ta viết $hP = h(p,u)$ thay vì (p,hu).

(2) Với mọi $P,Q \in FP$ ta có $P \vee Q$, $P \wedge Q$, $\neg P$, $P \rightarrow Q$ thuộc FP.

Định nghĩa 3.10. Cho $T = (T, C, LH, \leq, \cup, \cap, \Rightarrow, -)$ là đại số gia tử PN-không thuần nhất hữu hạn đối xứng của biến ngôn ngữ Truth. Ánh xạ v : $FP \rightarrow T$ được gọi là hàm định giá trên T nếu các điều kiện sau đây thỏa mãn

(1) Nếu $P=(p,u)$ là mệnh đề cơ sở thì $v(P)$ luôn xác định, hơn nữa $v(\neg(p,u)) = v(p,-u)$.

(2) Nếu $P=(p,ku)$ thì $v(hP)=\delta l\tau$ khi và chỉ khi $v(P)=\delta l^* h^* \tau$ với mọi h, k, $l \in H$ và $\tau \in C$, trong đó :

$h^*=h^-$, $l^*=l$ nếu $k=N$

và $h^*=h$, $l^*=l^-$ nếu $k \neq N$ và $h=N$

và $h^*=h$, $l^*=l$ nếu $k \neq N$ và $h \neq N$

(3) Với mọi công thức P, Q mà $v(P)$ và $v(Q)$ xác định thì

$v(P \vee Q) = v(P \cup Q)$

$v(P \wedge Q) = v(P \cap Q)$

$v(P \rightarrow Q) = v(P) \Rightarrow v(Q)$

$v(\neg P) = \neg v(P)$

Ở đây trong vế trái là phép toán logic, vế phải là phép toán của T

Hai công thức P và Q được gọi là tương đương, kí hiệu $P \equiv Q$ nếu với mọi phép định giá v, khi $v(P)$ và $v(Q)$ xác định thì $v(P) = v(Q)$.

Định lý 3.4. Với mọi công thức P, Q, R, mọi $h \in H$ và với mọi vị từ p, ta có

(1) $\neg(p,u) \equiv (p,\neg u)$ và $(p,h\text{-}u) = \neg(p,hu)$

(2) $P \equiv P$ và $\neg\neg P = P$

(3) $P \vee P \equiv P$ và $P \wedge P \equiv P$

(4) $P \vee Q \equiv Q \vee P$ và $P \wedge Q \equiv Q \wedge P$

(5) $P \vee (Q \vee R) \equiv (P \vee) Q \vee R$ và $P \wedge (Q \wedge R) \equiv (P \wedge) Q \wedge R$

(6) $P \wedge (P \vee Q) \equiv P$ và $P \vee (P \wedge Q) \equiv P$

(7) $\neg(P \vee Q) \equiv \neg P \wedge \neg Q$ và $\neg(P \wedge Q) \equiv \neg P \vee \neg Q$

(8) $P \rightarrow Q \equiv \neg P \vee Q$

Tính chất phân phối giữa phép $\vee$ và $\wedge$ không thỏa mãn vì đại số gia tử PN-không thuần nhất hữu hạn đối xứng không phải là dàn phân phối.

3.2.3.2. Các qui tắc suy diễn

Một số tác tác giả đã xây dựng một số qui tắc suy diễn cho lập luận ngôn ngữ như qui tắc chuyển gia tử, qui tắc tỉ lệ.... Các qui tắc này giải quyết khá hiệu quả cho phần lớn các mệnh đề mờ thường gặp. Tuy nhiên, nếu sử dụng chúng thì trong quá trình lập luận có thể thu được một số kết quả không phù hợp. Chẳng hạn xét mệnh đề sau : "Một sinh viên học chăm thì kết quả tốt" và do đó "Nếu An học không chăm thì kết quả có thể là tốt". Điều này chấp nhận được kết quả có thể tốt được hiểu là không tốt lăm. Sử dụng qui tắc tỉ lệ cho câu thứ hai ta thu được "Nếu An học rất không chăm lắm thì kết quả rất có thể là tốt", kết luận này nói chung không phù hợp nữa. Điều này xảy ra do sự xuất hiện của "không chăm lăm" chứa gia tử "không" (*Not so*) và tính không thuần nhất của nó với gia tử "có thể" (Possibly) trong thành phần còn lại. Cũng vì lý do này mà xuất hiện những kết quả không phù hợp khi sử dụng qui tắc chuyển đổi gia tử. Vì vậy, trong tác giả L.X.Vinh đã mở rộng qui tắc chuyển đổi gia tử trước đây và đưa ra một số qui tắc suy diễn mới như thay thế gia tử đồng mức, phản tỉ lệ để giải quyết các tình huống nêu trên.

Chúng ta biết rằng qui tắc suy diễn là một sơ đồ mà dựa vào đó người ta có thể suy ra các kết luận từ một tập các khẳng định cho trước, nó có dạng :

$$\frac{(P_1, t_1),\ldots\ldots\ldots(P_n, t_n)}{(Q_1, s_1),\ldots\ldots\ldots(Q_m, s_m)}$$

trong đó (P_i, t_i) là các tiền đề và (Q_i, s_i) là các kết luận với các giá trị t_i, s_i >W. Một qui tắc được gọi là đúng đắn nếu $v(P_i) = t_i$, với mọi i=1..n thì $v(Q_j) = s_j$ với mọi j = 1..n với v là hàm định giá bất kỳ.

Qui tắc chuyển gia tử cho mệnh đề mờ đơn giản

Trong quá trình lập luận ngôn ngữ ở nhiều bước chúng ta cần chuyển một mệnh đề mờ sang dạng khác có ý nghĩa tương đương. Các qui tắc sau cho chúng ta cách xác định mức độ đúng của các mệnh đề thu được.

$$(RT_1) \qquad \frac{((P, hku), \delta l\tau)}{((P, ku), \delta l^{*}h^{*}\tau)}$$

$$(RT_2) \qquad \frac{((P, ku), \delta lh\tau)}{((P, h^{*}ku), \delta l^{*}\tau)}$$

trong đó δ là xâu gia tử bất kỳ, h, k, $l \in H$, h^*, l^* xác định theo công thức trong định nghĩa hàm định giá v và τ là khái niệm sinh nguyên thủy của biến ngôn ngữ Truth.

Qui tắc chuyển gia tử cho mệnh đề kéo theo

$$(RTI_1) \qquad \frac{(h(P, ku) \rightarrow h(Q, kv), \delta l True),((P, ku), \delta' True)}{((P, ku) \rightarrow (Q, kv), \delta l^{*}h^{*}True)}$$

$$(RTI_2) \qquad \frac{((P, ku) \rightarrow (Q, kv), \delta l True),((P, ku), \delta' True)}{(h^{*}(P, ku) \rightarrow h^{*}(Q, kv), \delta l^{*}True)}$$

trong đó δ và δ' là các xâu gia tử tùy ý, h, k, $l \in H$ và h^*, l^* xác định như trong hàm định giá v.

Hai qui tắc sau đây là mở rộng cho qui tắc Modus ponens và Modus tollens của logic kinh điển.

$$(RMP) \qquad \frac{(P \rightarrow Q, \delta True),(P, True)}{(Q, \delta True)}$$

$$(RMT) \qquad \frac{(P \rightarrow Q, \delta True),(\neg Q, True)}{(\neg P, \delta True)}$$

Các qui tắc phản tỉ lệ và tỉ lệ

Phân loại mệnh đề kéo theo

Trong thực tế nhiều mệnh đề kéo theo có tính tỉ lệ giữa hai thành phần của nó. Chẳng hạn "nếu sinh viên học càng chăm thì kết quả càng tốt" hoặc "trời càng nắng thì nhiệt độ càng cao"...Đối với các mệnh đề này, chúng ta có thể nói rằng "trời nắng thì nhiệt độ cao" là "tương đối đúng" dẫn đến " trời rất nắng thì nhiệt độ rất cao" hay "trời không nắng lắm thì nhiệt độ không cao lắm" cũng sẽ là "tương đối đúng"....Tuy nhiên, khi xuất hiện gia tử không (*Not so*) ở đúng một trong hai thành phần thì sẽ không còn tỉ lệ nữa, khi đó chúng ta gọi tính chất này là phản tỉ lệ. Ví dụ "nếu Lan học không chăm thì kết quả có thể tốt" là "tương đối đúng" không thể suy ra "nếu Lan học rất không chăm thì kết quả rất có thể tốt" là "tương đối đúng" mà phải là "nếu Lan học rất không chăm thì kết quả ít có thể tốt" mới là "tương đối đúng".

Các mệnh đề vừa đề cập trên đây có dạng $P(x^*,h_1u) \rightarrow Q(x^*,h_2v)$, trong đõ có thể là biến hoặc hằng u,v là các khái niệm mờ, h_1, h_2 là các gia tử. Ta chia các mệnh đề kéo theo thành hai loại khác nhau:

- *Loại tỉ lệ*: Khi h_1 và h_2 không phải là gia tử *Not so* ($h_1 \neq N$, $h_2 \neq N$) hoặc đồng thời là hai gia tử này ($h_1 = h_2 = N$)
- *Loại phản tỉ lệ*: Khi có đúng một gia tử h_1 hoặc h_2 là *Not so*.

Qui tắc tỉ lệ

(RPI) $$\frac{(P(x^*, h_1u) \rightarrow Q(x^*, h_2v), \delta True)}{(hP(x^*, h_1u) \rightarrow hQ(x^*, h_2v), \delta True)}$$

trong δ là các xâu gia tử, x^* có thể hằng hoặc biến, các công thức P, Q thuộc lớp có thể chuyển gia tử, h_1, h_2 là các gia tử tùy ý thỏa mãn điều kiện mệnh đề tỉ lệ.

Từ (RPI), (RMP), (RMT) với a là hằng ta suy ra :

Qui tắc tỉ lệ Modus ponens

(RPMP) $$\frac{(P(x^*, h_1u) \rightarrow Q(x^*, h_2v), \delta True), (hP(a, h_1u), True)}{(hQ(a, h_2v), \delta True)}$$

Qui tắc tỉ lệ Modus tollens

(RPMT) $$\frac{(P(x^*, h_1u) \rightarrow Q(x^*, h_2v), \delta True), (\neg hQ(a, h_2u), True)}{(\neg hP(a, h_1u), \delta True)}$$

Qui tắc phản tỉ lệ

(RNPI) $$\frac{(P(x^*, h_1u) \rightarrow Q(x^*, h_2v), \delta True)}{}$$

$$(hP(x^*, h_1u) \rightarrow h\text{-}Q(x^*, h_2v), \delta True)$$

trong δ là các xâu gia tử, x^* có thể hằng hoặc biến, h_1, h_2 là các gia tử tùy ý thỏa mãn điều kiện mệnh đề loại phản tỉ lệ và h- là gia tử đối xứng của h.

Từ các qui tắc (RNPI), (RMP), (RMT) ta suy ra

$$\text{(RNPMP)} \qquad \frac{(P(x^*, h_1u) \rightarrow Q(x^*, h_2v), \delta True), (hP(a, h_1u), True)}{(h\text{-}Q(a, h_2v), \delta True)}$$

và

$$\text{(RNPMT)} \qquad \frac{(P(x^*, h_1u) \rightarrow Q(x^*, h_2v), \delta True), (\neg hQ(a, h_2u), True)}{(\neg h\text{-}P(a, h_1u), \delta True)}$$

Các qui tắc tương đương và thay thế hằng cho biến

Việc thay thế các gia tử đồng mức h và k cho nhau ở vị trí tiền tố của một khái niệm mờ không làm thay đổi ý nghĩa của mệnh đề. Vì vậy, ta có qui tắc thay thế gia tử đồng mức sau đây :

$$\text{(REH)} \qquad \frac{P(x^*, hu)}{P(x^*, ku)}$$

Ngoài ra tương tự [5][7] cũng có qui tắc thay thế công thức tương đương :

$$\text{(REF)} \qquad \frac{P \equiv Q, (F(P), \delta\tau)}{(F(P/Q), \delta\tau)}$$

qui tắc thay thế hằng a bởi biến x^*

$$\text{(RSUB)} \qquad \frac{P(x^*, u)}{P(a, u)}$$

3.2.4. Phương pháp lập luận ngôn ngữ

Lập luận ngôn ngữ là tìm kiếm các kết luận không chắc chắn bằng phương pháp suy diễn theo nghĩa xấp xỉ từ các tiền đề không chắc chắn ở dạng ngôn ngữ. Tương tự như trong đại số gia tử, chúng ta đi tìm các khẳng định từ cơ sở tri thức dựa vào các qui tắc suy dẫn.

Chúng ta thừa nhận rằng, giá trị chân lý của mỗi khẳng định là không suy nhất, nó có thể nhận nhiều giá trị khác nhau miễn là các giá trị đó đều lớn hơn hay nhỏ hơn giá trị trung hòa W. Chẳng hạn cho P, Q là hai công thức mà $P \rightarrow Q$ thuộc loại tỉ lệ và có giá trị chân lý là $\sigma True$, nghĩa là $(P \rightarrow Q, \sigma True)$. Với tùy ý $h \in LH$, theo (RPI)

$$(hP \rightarrow hQ, \sigma True)$$

do đó, theo (RTI_1) nhiều trường hợp trở thành

$$(P \rightarrow Q, \sigma hTrue)$$

và trong đại số gia tử PN-không thuần nhất của biến ngôn ngữ Truth, ta có $\sigma True, \sigma hTrue \geq W$.

Do đó trong tác giả đã đưa ra hai khẳng định tương đương trong cơ sở tri thức mờ.

Ví dụ 3.2. Giả sử trong cơ sở tri thức có các khẳng định sau

-"Hôm nào trời càng nắng thì nhiệt độ càng cao" là "rất đúng".

-"Hôm nay trời rất là không nắng lắm".

Ta có thể rút ra kết luận gì từ các khẳng định trên.

Biểu diễn "Hôm nào trời nắng" bằng p(Hôm nào, nắng) và "nhiệt độ cao" là q(nhiệt độ, cao). Khi đó ta có :

(1) p((Hôm nay, rất không nắng lắm), đúng) (giả thiết)

(2) (p(Hôm nào, nắng) $\rightarrow$ q(Nhiệt độ, cao), rất đúng) (giả thiết)

(3) (p(Hôm nào, có thể nắng) $\rightarrow$ q(Nhiệt độ, có thể cao), rất đúng),từ (2) và (RPI)

(4) (p(Hôm nào, không nắng lắm) $\rightarrow$ q(Nhiệt độ, có thể cao), rất đúng),từ (3) và (REH).

(5) (p(Hôm nào, rất không nắng lắm) $\rightarrow$ q(Nhiệt độ, ít có thể cao), rất đúng), từ (4) và (RNPI).

(6) (p(Hôm nay, rất không nắng lắm) $\rightarrow$ q(Nhiệt độ, ít có thể cao), rất đúng), từ (5) và (RSUB).

(7) (q(nhiệt độ, ít có thể cao), rất đúng), từ (6), (1) và (RMP)

(8) (q(nhiệt độ, có thể cao), rất ít đúng), từ (7) và (RT_1)

(9) (q(nhiệt độ, cao), rất ít có thể đúng), từ (8) và (RT_1)

(10) (q(nhiệt độ, rất ít có thể cao), đúng), từ (9) và (RT_2)

Như vậy, ta có thể sử dụng kết luận (8) "Hôm nay nhiệt độ có thể cao" là "rất ít đúng" hoặc kết luận (10) "Hôm nay nhiệt độ rất ít có thể cao" là "đúng".

3.3. Thao tác dữ liệu mờ

3.3.1. Các mô hình Cơ sở dữ liệu mờ

Trước hết, một số khái niệm cơ bản ban đầu về cơ sở dữ liệu quan hệ được trình bày. Tiếp đến là một số cách tiếp cận để mở rộng mô hình cơ sở dữ liệu quan hệ để nhằm xử lý những thông tin mờ.

Định nghĩa 3.11. Cho D_1, D_2, …, D_n là n miền giá trị, r là một quan hệ trên các miền D_1, D_2, …, D_n nếu r là một tập con nào đó các n-bộ được sắp có dạng $(d_1, d_2, \ldots, d_n)$ sao cho $d_i \in D_i$ với i = 1, 2, .., n.

Rõ ràng, một quan hệ có thể biểu diễn dưới dạng một bảng, mỗi dòng biểu diễn một bộ của quan hệ, mỗi cột biểu diễn một thành phần của các bộ trong quan hệ. Số các bộ của một quan hệ gọi là lực lượng của quan hệ và số các thành phần được gọi là bậc của quan hệ. Để dễ dàng tham chiếu đến các thành phần của các bộ mà không cần biết thứ tự của thành phần trong bộ được sắp, người ta thường đặt tên cho các thành phần, tên của một thành phần là một thuộc tính. Như vậy, mỗi D_i trong định nghĩa 3.11 chính là miền giá trị của thuộc tính thứ i trong quan quan hệ r. Do đó, ta có thể định nghĩa một quan hệ xác định trên một tập thuộc tính như sau:

Định nghĩa 3.12. Cho $U = \{A_1, A_2, \ldots, A_n\}$ là một tập hữu hạn các thuộc tính. Với mỗi thuộc tính A_i, với i =1, 2,.., n có miền giá trị tương ứng là $Dom(A_i)$. Khi đó, r là quan hệ xác định trên tập thuộc tính U nếu : $r \subseteq Dom(A_1) \times Dom(A_2) \times \ldots\ldots \times Dom(A_n)$.

Đối với mô hình CSDL quan hệ, nếu người quản trị đang quản lý một cơ sở dữ liệu (CSDL) thực tế nào đó, khi xử lý các tình huống trong đó hoặc không có đầy đủ thông tin, hoặc có thông tin không chính xác, không chắc chắn mà gọi chung là thông tin mờ thì sẽ gặp những tình huống sau :

Tại thời điểm cần cập nhật một đối tượng nào đó vào CSDL nhưng chưa có đầy đủ thông tin về đối tượng đó. Chẳng hạn, đối tượng đó là một cán bộ khoa học có học vị TSKH nhưng giá trị thuộc tính học vị không có thông tin về năm bảo vệ (*at present unknown value*).

Biết đã hướng dẫn *nhiều* nghiên cứu sinh nhưng không biết cụ thể là bao nhiêu (*vague concept*).

Nếu giới hạn trong mô hình quan hệ thì phải đợi đủ thông tin về đối tượng đó ta mới nhập vào CSDL, hoặc nếu cứ nhập sẽ gây khó khăn, mất ngữ nghĩa và không nhất quán trong xử lý dữ liệu.

Một tình huống khác nữa : Giả sử trong CSDL quan hệ về các cán bộ, khi đó ta thường có những nhu cầu xử lý câu hỏi "*tìm các cán bộ khoa học trẻ và có nhiều công trình công bố trên các tạp chí quốc tế*". Rõ ràng trong phạm vi thao tác dữ liệu kinh điển chúng ta không thể xử lý câu hỏi chứa khái niệm mờ như vậy.

Do đó, một cách tự nhiên xuất hiện nhu cầu mở rộng CSDL. Có nhiều cách mở rộng nhưng dựa trên hai phạm trù chính:

Thứ nhất: cú pháp (các ký hiệu và các quy tắc kết hợp các ký hiệu).

Thứ hai : ngữ nghĩa (ý nghĩa các ký hiệu trong thế giới thực).

Một cách hình thức có hai cách mở rộng mô hình quan hệ là mở rộng ngữ nghĩa của dữ liệu để khai thác dữ liệu rõ với các yếu tố mờ và mở rộng miền giá trị của thuộc tính để biểu diễn dữ liệu mờ.

Với cách mở rộng ngữ nghĩa, dữ liệu tại mỗi bộ đối với thuộc tính là dữ liệu rõ. Tuy nhiên, cho phép chúng ta khai thác dữ liệu với nghĩa rộng hơn. Ví dụ chúng ta có thể "*tìm những cán bộ có kinh nghiệm và có lương tương đối cao trong năm 2005*". Cách mở rộng này có ưu điểm là có thể sử dụng các hệ quản trị CSDL mô hình quan hệ trong việc lưu trữ dữ liệu. Tuy nhiên, nó không cho phép biểu diễn dữ liệu mờ nên hạn chế nhiều đến khả năng quản lý dữ liệu thực tế.

Với cách mở rộng miền trị cho thuộc tính sẽ cho phép bổ sung thêm các cú pháp trong biểu diễn dữ liệu nhằm cho phép biểu diễn được dữ liệu mờ. Theo cách này, ngoài việc đưa vào hệ thống ký hiệu, việc quan trọng là giải quyết vấn đề ngữ nghĩa của các ký hiệu. Do đó, hướng nghiên cứu này sẽ mang tính tổng quát hơn.

Trong những năm qua, đã có nhiều tác giả trong và ngoài nước nghiên cứu và đề xuất các mô hình CSDL mờ theo hai cách trên, đặc biệt có ba mô hình sau đây được nhiều tác giả quan tâm nghiên cứu.

Mô hình CSDL mờ theo cách tiếp cận tập con mờ.

Mô hình CSDL mờ theo cách tiếp cận quan hệ tương tự.

Mô hình CSDL mờ theo cách tiếp cận lý thuyết khả năng.

Để xử lý, lưu trữ và biểu diễn các dữ liệu mờ đã có rất nhiều tác giả trong và ngoài nước quan tâm nghiên cứu nhằm mở rộng mô hình quan hệ do E.F.Codd đề xuất năm 1970. Một số tác giả nước ngoài tiêu biểu cho hướng nghiên cứu này có thể kể đến Buckles và Petry (1980), Baldwin và Zhou (1984), Raju K.V.S.N và Mazumdar (1988), Yager (1995), Bosc (1995), Cubero (1994), Mustafa ILKer Sozat, Adnan Yazici (2001)…..Các tác giả trong nước nghiên cứu đầu tiên phải kể đến PGS.TS Hồ Thuần, PGS.TS Lê Tiến Vương từ những năm 1985, 1989, tiếp đến là các tác giả Đinh Thị Ngọc Thanh, Trương Đức Hùng (1996), Hồ Cẩm Hà (2002), Trần Thiên Thành (2004), Nguyễn Công Hào (2003).

Tóm lại, các mô hình CSDL mờ của các tác giả tập trung nghiên cứu chủ yếu là ba mô hình dưới đây.

3.3.2. Mô hình CSDL mờ theo cách tiếp cận tập con mờ

Cách tiếp cận này do Baldwin và Zhou đưa ra năm 1984, Zvieli đưa ra năm 1986, với quan niệm rằng một quan hệ $r \subseteq D_1 \times D_2 \times \ldots \times D_n \times [0,1]$ được cho bởi một hàm thuộc $\mu_r : D_1 \times D_2 \times \ldots \times D_n \rightarrow [0,1]$. Như vậy, một bộ dữ liệu $t \in r$ có dạng: $(t_1, t_2, ..t_n, \mu_r(t_1, t_2, ..t_n))$, trong đó $t_i \in D_i$ và mỗi bộ dữ liệu phải thuộc về một quan hệ là khái niệm mờ, nhưng giá trị trên mỗi thuộc tính là giá trị rõ. Trên mỗi miền trị có yếu tố mờ, thay quan hệ đồng nhất trên miền trị thuộc tính bởi quan hệ xấp xỉ bằng nhau được xác định bởi hàm thuộc μ thoả mãn tính chất phản xạ và đối xứng.

Về mặt biểu diễn quan hệ mờ trong mô hình này giống trong mô hình quan hệ nhưng thêm cột μ để chỉ độ thuộc của một bộ vào quan hệ.

Ví dụ 3.3. Xét lược đồ quan hệ *COMPANY*($\mathscr{N}$, $\mathscr{E}$, $\mathscr{S}$, $\mathscr{P}$). Miền trị của các thuộc tính $\mathscr{E}$, $\mathscr{S}$, $\mathscr{P}$ là các tập mờ trên các vũ trụ tương ứng là $\mathscr{U}_{\mathscr{E}}$ = [0,30], $\mathscr{U}_{\mathscr{P}}$ = [500,3000], $\mathscr{U}_{\mathscr{S}}$ = [5000,30000]. Các hàm thuộc $\mu_{\mathscr{E}}$ tương ứng tập con mờ *"Số nhân viên ít"*, $\mu_{\mathscr{S}}$ tương ứng tập con mờ *"Lượng bán cao"*, μ_P tương ứng tập con mờ *"Lợi nhuận nhiều"* được xác định như sau:

$$\mu_{\mathscr{E}}(e) = \begin{cases} (1+|e-10|/4)^{-1} & \text{nếu } e \geq 10 \\ 1 & \text{nếu } e < 10 \end{cases}$$

$$\mu_{\mathcal{S}}(\mathrm{s}) = \begin{cases} (1+|s-12000|/4000)^{-1} & \text{nếu s} \le 12000 \\ 1 & \text{nếu s} > 120000 \end{cases}$$

$$\mu_{\mathcal{P}}(\mathrm{p}) = \begin{cases} (1+|p-1600|/400)^{-1} & \text{nếu p} \le 1600 \\ 1 & \text{nếu p} > 1600 \end{cases}$$

Một quan hệ *r* trên lược đồ *COMPANY* thể hiện "*Lượng bán hàng cao, lợi nhuận nhiều và ít nhân viên*" như sau:

TENCTY ($\mathcal{N}$)	SONVIEN ($\mathcal{E}$)	LUONGBAN ($\mathcal{S}$)	LOINHUAN ($\mathcal{P}$)	μ
CT Bách hoá 1	10	11000	1100	0.44
CT Bách hoá 2	11	10000	1600	0.67
CT Bách hoá 3	14	10000	1500	0.50
CT Bách hoá 4	12	13000	1800	0.67
CT Bách hoá 5	09	8000	1200	0.50

Bảng 3.1. *Quan hệ r trên lược đồ COMPANY*

Các phép tính quan hệ như: Phép chiếu, hợp, giao, tích Đề-Các được thực hiện như các phép toán tương ứng trên các tập mờ.

Quan hệ xấp xỉ bằng nhau *EQ* trên miền trị D_i là ánh xạ $\mu_{EQ} : D_i \to [0,1]$ thoả mãn các tính chất sau: với $\forall\ x, y \in D_i$:

(1) *Tính phản xạ* : $\mu_{EQ}(x,x) = 1$

(2) *Tính đối xứng* : $\mu_{EQ}(x,y) = \mu_{EQ}(y,x)$

Để so sánh giá trị của hai bộ dữ liệu trên một thuộc tính, trên một tập thuộc tính quan hệ *EQ* được sử dụng.

Giả sử $X = \{A_1, A_2, \ldots, A_m\}$ là tập con của *U*, *r* là quan hệ mờ trên *U* và t_1, t_2, t_3 là ba bộ dữ liệu thuộc *r*, ta định nghĩa:

Để biểu thị hai giá trị $t_1[A]$ và $t_2[A]$ giống nhau theo quan hệ *EQ* trên một thuộc tính *A* ta xác định $\mu_{EQ}(t_1[A],t_2[A])$.

Để biểu thị ba giá trị $t_1[A]$, $t_2[A]$ và $t_3[A]$ giống nhau theo quan hệ *EQ* trên một thuộc tính *A* ta có :

$\mu_{EQ}(t_1[A],t_2[A],t_3[A])= \min\{\mu_{EQ}(t_1[A],t_2[A]),\mu_{EQ}(t_2[A],t_3[A]),\mu_{EQ}(t_3[A],t_1[A])\}$.

Để biểu thị hai giá trị $t_1[X]$,$t_2[X]$ giống nhau trên tập thuộc tính X theo quan hệ EQ : $\mu_{EQ}(t_1[X],t_2[X]) = \min\{\mu_{EQ}(t_1[A_1],t_2[A_1]), \mu_{EQ}(t_1[A_2],t_2[A_2])\ldots, \mu_{EQ}(t_1[A_m],t_2[A_m])\}$.

3.3.3. Mô hình CSDL mờ theo cách tiếp cận quan hệ tương tự

Mô hình này đã được Buckles và Petry đề xuất năm 1980. Trong mô hình này, giá trị của mỗi bộ tại một thuộc tính có thể đa trị (một tập các giá trị có thể). Trên mỗi miền trị chứa dữ liệu mờ được trang bị một quan hệ tương tự để đánh giá độ *"gần nhau"* giữa các giá trị.

Đối với mô hình do hai tác giả Buckles và Petry đề xuất, giá trị tại mỗi thuộc tính của đối tượng có thể là đơn trị hoặc đa trị nhưng có một ràng buộc là các giá trị đòi hỏi phải *"đủ tương tự nhau"*, hay nói cách khác là độ tương tự của hai giá trị bất kỳ không nhỏ hơn ngưỡng cho trước.

Tuy nhiên, trong cuộc sống có thể gặp những thông tin không chắc chắn về một đối tượng, có thể gặp những thông tin cho biết một số khả năng mà điều xảy ra trong thực tế chỉ là một trong các khả năng này. Chẳng hạn, tại một thời điểm khi chẩn đoán bệnh cho một bệnh nhân An, các bác sĩ trong hội đồng đưa ra chẩn đoán bệnh như sau: Viêm amidan, viêm phế quản, sốt siêu vi trùng. Thực tế có thể bệnh nhân An bị một trong 3 bệnh nói trên hoặc bị cả 3 bệnh. Nếu xem một tập các khả năng có thể xảy ra cần được lưu trữ khi chúng ta chưa có cơ sở chắc chắn một khả năng nào trong đó là đúng thì tác giả Hồ Cẩm Hà đã mở rộng mô hình của P.Buckles và E.Petry theo ngữ nghĩa như vậy. Một điều đặc biệt trong mô hình này là các phần tử của mỗi giá trị thuộc tính không đòi hỏi *"đủ tương tự theo ngưỡng"*.

Các kết quả nghiên cứu và mở rộng theo cách tiếp cận mô hình này có thể kể đến S.K.De, R.Biswas và Mustafa ILKer Sozat, Adnan Yazici .

Để đối sánh giá trị của hai bộ dữ liệu trên một thuộc tính, trên một tập thuộc tính, sử dụng quan hệ tương tự (*similarity relation*) hay quan hệ tương đương mờ có ba tính chất phản xạ, giao hoán và bắc cầu max-min.

Quan hệ tương tự trên miền D là một ánh xạ s từ $D \times D \rightarrow [0,1]$ thoả mãn các tính chất sau: với $\forall x, y, z \in D$:

(1) *Tính phản xạ* : $s(x,x) = 1$

(2) *Tính đối xứng* : $s(x,y) = s(y,x)$

(3) *Tính bắc cầu max-min*: $s(x,z) \geq \max_{y \in D} \{\min(s(x,y), s(y,z))\}$

Cho $x, y \in D$, $\alpha \in [0,1]$, ta nói *x tương đương y* với ngưỡng α, ký hiệu $x \sim_{\alpha} y$, nếu $s(x,y) \geq \alpha$. Khi đó, quan hệ $\sim_{\alpha}$ là một quan hệ tương đương, do đó quan hệ này sẽ chia miền *D* thành các lớp tương đương $d_1, d_2, .., d_k$.

Một quan hệ mờ *r* trên tập các thuộc tính $A_1, A_2, \ldots A_n$ là một tập con tích Đề-Các: $2^{D1} \times 2^{D2} \times \ldots\ldots \times 2^{Dn}$.

Trên mỗi miền trị D_j xác định một quan hệ tương tự s_j và có một ngưỡng $\alpha_j \in [0,1]$. Một bộ $t_i \in r$ có dạng $t_i = (d_{i1}, d_{i2}, \ldots d_{in})$, với $d_{ij} \subseteq D_j$.

Để giải quyết vấn đề dư thừa dữ liệu của các bộ, trong mô hình này yêu cầu các giá trị tại mỗi thuộc tính của một bộ phải nằm trong một lớp tương đương với ngưỡng cho trước. Trong một quan hệ có các bộ dư thừa, khi đó ta trộn các thành phần tương ứng với nhau để tạo thành một bộ mới tương đương với hai bộ ban đầu.

Một hạn chế của mô hình này là sử dụng quan hệ tương tự, bởi vì nó là một quan hệ yêu cầu khá chặt do tính bắc cầu max-min làm hạn chế khả năng biểu diễn dữ liệu trong thực tế. Một số nghiên cứu thay quan hệ tương tự bởi quan hệ gần nhau (proximity relation) không cần phải thoả mãn tính bắc cầu max-min. Tuy nhiên, để đảm bảo các kết quả **trong mô hình quan** hệ, các tác giả đưa ra quan hệ tương đương α-gần nhau với mục đích phân hoạch miền trị mỗi thuộc tính thành các lớp tương đương. Do đó, các kết quả quan trọng của lý thuyết cơ sở dữ liệu quan hệ được mở rộng trên mô hình này vẫn đúng như : phụ thuộc dữ liệu, các dạng chuẩn, phép tách…..

3.3.4. Mô hình CSDL mờ theo cách tiếp cận lý thuyết khả năng

Trước hết, chúng ta xem xét mối quan hệ giữa tính mờ và khả năng thông qua một ví dụ. Sau đó, mô hình CSDL mờ nghiên cứu dựa trên lý thuyết khả năng được giới thiệu.

Ví dụ 3.4. Xét mệnh đề p = x là một số nguyên trong khoảng [0,5].
Khi đó, mệnh đề p khẳng định (i) : có thể bất kỳ số nguyên nào trong khoảng [0,5] là giá trị của x ; (ii) : không thể bất kỳ số nguyên nào ngoài khoảng [0,5] là giá trị của x.

Nói cách khác, $\mathscr{p}$ sinh ra một phân bố khả năng $\pi_{\mathscr{X}}$ gắn với mỗi số nguyên $u \in [0,5]$ khả năng u có thể là giá trị của $\mathscr{x}$. Do đó, $\pi_{\mathscr{X}} = \text{Poss}\{\mathscr{X} = u\} = 1$ với $0 \leq u \leq 5$ và $\pi_{\mathscr{X}} = \text{Poss}\{\mathscr{X} = u\} = 0$ với $u < 0$ hoặc $u > 5$. Ở đây, Poss $\{\mathscr{X} = u\}$ là khả năng $\mathscr{X}$ có thể nhận giá trị u.

Bây giờ, mệnh đề $\mathscr{p}$ được xem xét với nghĩa "*mờ*": $\mathscr{p}$ = $\mathscr{x}$ là một số nguyên nhỏ. Ở đây, *số nguyên nhỏ* là tập mờ được định nghĩa trong vũ trụ số nguyên dương như sau : *Số nguyên nhỏ* = 1/0 + 1/1 + 0.9/2 + 0.7/3 + 0.5/4 + 0.2/5. Trong đó 0.7/3 có nghĩa là mức độ thuộc của số nguyên 3 trong tập mờ *số nguyên nhỏ*. Vì vậy, mệnh đề $\mathscr{p}$ với nghĩa mờ khẳng định : có thể bất kỳ số nguyên nào là *số nguyên nhỏ* với khả năng của $\mathscr{X}$ nhận giá trị của u bằng mức độ thuộc của u trong tập mờ *số nguyên nhỏ*. Do đó, $\text{Poss}\{\mathscr{X} = 0\} = \text{Poss}\{\mathscr{X} = 1\}=1$, $\text{Poss}\{\mathscr{X} = 2\} = 0.9$, $\text{Poss}\{\mathscr{X} = 3\} = 0.7$, $\text{Poss}\{\mathscr{X} = 4\} = 0.5$, $\text{Poss}\{\mathscr{X} = 5\} = 0.2$, $\text{Poss}\{\mathscr{X} = u\} = 0$ với $u < 0$ hoặc $u > 5$.

Theo cách tiếp cận tập mờ, Zadeh xem phân bố khả năng Poss $\{\mathscr{X} = u\}$ như thu hẹp bởi tập mờ F trên miền trị $\mathscr{U}$, có hàm thuộc μ_F. Khi đó, Poss $\{\mathscr{X} = u\} = \mu_F(u)$, với mọi $u \in \mathscr{U}$.

Mô hình CSDL mờ dựa trên lý thuyết khả năng được đề xuất bởi Prade và Testemale vào năm 1983 bằng cách mở rộng miền trị thuộc tính, sử dụng phân bố khả năng để biểu diễn các dữ liệu mờ. Giá trị của một n-bộ t tại thuộc tính A được biểu diễn bởi phân bố khả năng chuẩn $\pi_{t[A]}$ (tồn tại $u \in \mathscr{U}$: $\pi(u) = 1$) trên miền trị mở rộng $D \cup \{e\}$. Trong đó e là phần tử bổ sung vào mỗi miền trị, được sử dụng trong trường hợp thuộc tính A không áp dụng (*inapplicable*) cho bộ t.

Một quan hệ mờ r trên tập thuộc tính $\{A_1, A_2, \ldots..A_n\}$ là một tập con của tích Đề-Các: $\prod(D_1) \times \prod(D_2)\ldots\times \prod(D_n)$, với $\prod(D_i)$ là tập các phân bố khả năng chuẩn trên miền trị D_i của thuộc tính A_i, i = 1. .n.

Một bộ dữ liệu t thuộc quan hệ r có dạng $t = (\pi_{t[A1]}, \pi_{t[A2]}, \ldots.\pi_{t[An]})$.

Độ gần nhau của hai giá trị $\pi_{t1[Ai]}$, $\pi_{t2[Ai]}$ của hai bộ t_1 và t_2 tại thuộc tính A_i, ký hiệu $\tau(t_1[A_i],t_2[A_i])$ được xác định: $\tau(t_1[A_i],t_2[A_i]) = \underset{\substack{x,y\in D_i \\ s_i(x,y)\geq\alpha_i}}{Sup} \min(\pi_{t1[A_i]}(x), \pi_{t2[A_i]}(y))$.

với s_i là quan hệ gần nhau (có hai tính chất phản xạ và đối xứng) trên thuộc tính A_i, α_i là ngưỡng kết hợp với quan hệ s_i trên miền trị D_i.

Nếu không có quan hệ gần nhau thì mặc định là quan hệ đồng nhất, khi đó độ gần nhau của hai giá trị $\pi_{t1[Ai]}$, $\pi_{t2[Ai]}$ được xác định :

$$\tau(t_1[A_i],t_2[A_i]) = \sup_{x\in D_i} \min(\pi_{t1[A_i]}(x), \pi_{t2[A_i]}(x)).$$

Độ gần nhau trên tập thuộc tính X của hai bộ t_1 và t_2 tại thuộc tính A_i, ký hiệu $\tau(t_1[X],t_2[X])$ được xác định: $\tau(t_1[A_i],t_2[A_i]) = \min_{1\le i\le m}\{\tau(t_1[A_i],t_2[A_i])\}$, với $X = \{A_1, A_2, ..A_m\}$.

Biểu diễn dữ liệu của phân bố khả năng

Dùng phân bố khả năng cho phép biểu diễn được nhiều loại dữ liệu như: dữ liệu rõ, dữ liệu không áp dụng được, dữ liệu chưa biết, dữ liệu không có thông tin, dữ liệu không chắc chắn. Chúng ta có thể xét một số khả năng biểu diễn dữ liệu mờ của phân bố khả năng tiêu biểu như sau:

3.3.4.1 Biểu diễn dữ liệu trong trường hợp kinh điển

(1) Biết chắc chắn lương của An là 500: $\pi_{t[\text{LUONG}]}(e) = 0$, $\pi_{t[\text{LUONG}]}(500) = 1$, $\pi_{t[\text{LUONG}]}(d) = 0$, $\forall$ d $\in D$-{500}.

(2) An là một người không có lương, hay nói cách khác thuộc tính Lương không áp dụng cho An : $\pi_{t[\text{LUONG}]}(e) = 1$, $\pi_{t[\text{LUONG}]}(d) = 0$, $\forall$ d $\in D$.

(3) Biết chắc chắn rằng An có lương nhưng không biết là bao nhiêu (*Unknown*). Trong trường hợp này tất cả các giá trị đều có khả năng bằng nhau và bằng 1 : $\pi_{t[\text{LUONG}]}(e) = 0$, $\pi_{t[\text{LUONG}]}(d) = 1$, $\forall$ d $\in D$.

(4) Không biết gì về thông tin Lương của An (Null): $\pi_{t[\text{LUONG}]}(e) = 1$, $\pi_{t[\text{LUONG}]}(d) = 1$, $\forall$ d $\in D$.

3.3.4.2 Biểu diễn dữ liệu trong trường hợp mờ

(1) Không biết chính xác lương bao nhiêu nhưng chắc chắn trong khoảng từ 200 đến 300 : $\pi_{t[\text{LUONG}]}(e) = 0$, $\pi_{t[\text{LUONG}]}(d) = 1$ nếu $200 \le d \le 300$, $\pi_{t[\text{LUONG}]}(d) = 0$ nếu $d < 200$ hoặc $d > 300$.

(2) Biết lương của An là *cao,* khi đó ta dùng tập mờ *cao* với hàm thuộc μ_{cao} để biểu diễn : $\pi_{t[\mathrm{LUONG}]}(e) = 0$, $\pi_{t[\mathrm{LUONG}]}(d) = \mu_{cao}(d)$, $\forall$ d $\in D$.

(3) Ta biết những thông tin rời rạc về Lương của An: $\pi_{t[\mathrm{LUONG}]}(e) = 0$, $\pi_{t[\mathrm{LUONG}]}(d_i) = a_i$, i = 1..m, $\pi_{t[\mathrm{LUONG}]}(d) = 0$, $\forall$ d $\in D$-$\{d_1, d_2, \ldots d_m\}$.

Ví dụ 3.5. Xem xét câu nói tuổi của An là *già*.
Giả sử rằng tập mờ *già* được định nghĩa trên $\mathcal{U} = \{u : 0 \le u \le 100\}$. Do đó, $\prod_{A(X)}(u)$ có thể nhận các giá trị sau :

$\mathcal{U}$	10	20	25	30	35	40	50	60	…
$\prod_{A(X)}(u)$	0	0.2	0.3	0.5	0.8	0.9	1	1	…

Bảng 3.2. *Bảng phân bố khả năng gắn với "Già"*

3.3.5. Phụ thuộc dữ liệu trong CSDL mờ

Trong mô hình quan hệ, hai dạng phụ thuộc dữ liệu quan trọng giúp cho việc chuẩn hoá tốt các CSDL là phụ thuộc hàm và phụ thuộc đa trị. Khi mở rộng mô hình quan hệ để có thể biểu diễn và xử lý được những thông tin không chắc chắn, không đầy đủ đã có rất nhiều công trình tập trung nghiên cứu mở rộng hai dạng phụ thuộc này trên mô hình mới. Như vậy, khái niệm phụ thuộc hàm mờ được nhiều tác giả nghiên cứu phát triển dựa trên ý nghĩa của khái niệm phụ thuộc hàm cổ điển với nhiều cách tiếp cận khác nhau. Tuy nhiên, các cách tiếp cận mở rộng phụ thuộc hàm kinh điển này dựa vào 2 nguyên tắc chính:

Nguyên tắc thứ nhất: Nguyên tắc mở rộng này thay cho quan hệ bằng nhau trên dữ liệu rõ bởi quan hệ gần nhau hoặc quan hệ tương tự trên dữ liệu mờ và đặt ngưỡng để xác định độ gần nhau.

Nguyên tắc thứ hai: Nguyên tắc này dựa vào ý nghĩa của các phụ thuộc dữ liệu để xây dựng định nghĩa tương ứng cho mô hình mới sao cho bảo toàn một số kết quả quan trọng đã được xây dựng trong mô hình quan hệ.
Trong phần này ta dùng các ký hiệu

$\tau(t_1[A], t_2[A])$: là một số thuộc [0,1] để chỉ độ gần nhau của hai giá trị bộ t_1 và t_2 trên thuộc tính A.

$\tau(t_1[X],t_2[X])$: Độ gần nhau của hai giá trị bộ t_1 và t_2 trên tập thuộc tính X.

$\vec{\tau}(t_1[X],t_2[X]) = \{(t_1[A_1],t_2[A_1]), (t_1[A_2],t_2[A_2], \ldots\ldots,(t_1[A_k],t_2[A_k])\}$: Véctơ độ gần nhau của hai giá trị bộ t_1 và t_2 trên tập thuộc tính X.

3.3.5.1 Phụ thuộc hàm mờ

Phụ thuộc hàm kinh điển $X \rightarrow Y$ có nghĩa là nếu có hai bộ dữ liệu thuộc r mà giá trị trên tập thuộc tính X bằng nhau thì kéo theo giá trị trên tập thuộc tính Y cũng bằng nhau. Vấn đề đặt ra ở đây, nếu với hai bộ dữ liệu bất kỳ mà giá trị trên tập thuộc tính X là *"xấp xỉ"* bằng nhau thì kéo theo giá trị trên tập thuộc tính Y cũng *"xấp xỉ"* bằng nhau. Vậy X và Y có ràng buộc gì không? Đây chính là câu trả lời chúng ta phải đi nghiên cứu mở rộng phụ thuộc hàm trong CSDL quan hệ.

Ví dụ 3.6. Cho lược đồ quan hệ U = {*TEN*, *HSLUONG*, *SONAMCTAC*, *THUNHAP*} và quan hệ *Sonam_thunhap* xác định trên U được cho ở bảng 3.3

TEN	HSLUONG	SONAMCTAC	THUNHAP
Thành	2.67	10	1000000
Thuỷ	3.0	10	1050000
Hiền	3.63	15	1400000
Lành	4.55	20	3000000
Mạnh	4.55	22	3100000

Bảng 3.3. *Quan hệ Sonam_thunhap*

Ta thấy, trong quan hệ *Sonam_thunhap* không tồn tại phụ thuộc hàm kinh điển nào. Tuy nhiên, với hai bộ bất kỳ thuộc *Sonam_thunhap*, nếu *SONAMCTAC* (Số năm công tác) "*xấp xỉ nhau*" thì *THUNHAP* (Thu nhập) cũng "*Xấp xỉ nhau*". Như vậy, có một ràng buộc không chính xác giữa hai thuộc tính *SONAMCTAC* và *THUNHAP*. Do đó, các phụ thuộc trong CSDL quan hệ cần được nghiên cứu mở rộng.

a. Khái niệm phụ thuộc hàm mờ của Raju

Khái niệm phụ thuộc hàm mờ được Raju xây dựng trên mô hình tập con mờ. Phụ thuộc hàm mờ $X \sim> Y$ đúng trên quan hệ r nếu và chỉ nếu với mọi $t_1, t_2 \in r$ ta có :

$$\tau(t_1[X], t_2[X]) \leq \tau(t_1[Y], t_2[Y]).$$

Đây được xem là mở rộng tiêu biểu của khái niệm phụ thuộc hàm mờ và được nhiều tác giả tiếp tục mở rộng và phát triển trên các mô hình khác.

b. Khái niệm phụ thuộc hàm mờ của Chen

Phụ thuộc hàm mờ $X \sim>_\phi Y$ đúng trên quan hệ r nếu và chỉ nếu :

$\underset{t_1,t_2\in r}{Min}\{I(t_1[X],t_2[X]),\tau(t_1[Y],t_2[Y])\} \geq \phi$, trong đó $\phi \in [0,1]$, I là phép kéo theo của Gödel.

Dễ thấy khái niệm phụ thuộc hàm này mở rộng hơn khái niệm Raju. Điểm đặc biệt của phụ thuộc hàm mờ của Chen là cho phép thay đổi ngưỡng ϕ.

c. Khái niệm phụ thuộc hàm mờ của Cubero

Xuất phát từ quan điểm độ mờ trên mỗi thuộc tính là khác nhau nên đặt ngưỡng độ gần nhau cho mỗi thuộc tính.

Phụ thuộc hàm mờ $X \overset{(\vec{\alpha},\vec{\beta})}{\sim>} Y$ đúng trên quan hệ r nếu và chỉ nếu với mọi t_1, $t_2 \in r$ nếu $\vec{\tau}(t_1[X],t_2[X]) \geq \vec{\alpha}$ thì $\vec{\tau}(t_1[Y],t_2[Y]) \geq \vec{\beta}$, trong đó $\vec{\alpha},\vec{\beta}$ tương ứng là véctơ ngưỡng của các tập thuộc tính X, Y.

Khái niệm phụ thuộc hàm mờ của Cubero được chứng minh là mở rộng khái niệm phụ thuộc hàm mờ của Raju và Chen, tuy nhiên véc tơ ngưỡng phải cố định.

Để mở rộng phụ thuộc hàm mờ, tác giả Trần Thiên Thành đưa lượng từ ngôn ngữ vào trong phụ thuộc hàm mờ nhằm mô tả các phụ thuộc dữ liệu gần với thực tế hơn.

Cho r là một quan hệ trên lược đồ U, X, $Y \subseteq U$, $\phi \in [0,1]$. Độ thỏa của phụ thuộc hàm mờ $X \sim> Y$ của một bộ t trong quan hệ r, ký hiệu $\sigma(t|X \sim> Y)$, được xác định :

$$\sigma(t|X \sim> Y) = \min_{t1\in r}\{I(\tau(t[X],t_1[X]),\tau(t[Y],t_1[Y]))\}.$$

trong đó I là phép kéo theo mờ của Gödel. Ký hiệu $r_{X\sim>\phi Y} = \{ t \in r : \sigma(t|X\sim>Y) \geq \phi\}$

d. Khái niệm phụ thuộc hàm mờ của Hồ Thuần vàTrần Thiên Thành

Cho r là một quan hệ trên lược đồ U, X, $Y \subseteq U$, Q là lượng từ ngôn ngữ được xác định bởi hàm thuộc μ_Q, $\phi \in [0,1]$. Quan hệ r được gọi là thỏa phụ thuộc hàm mờ X xác định Y với ngưỡng ϕ và lượng từ Q, ký hiệu $Q(X\sim>_\phi Y)$ khi và chỉ khi $\mu_Q(|r_{X\sim>\phi Y}|) = 1$ nếu Q là lượng từ tuyệt đối, hoặc $\mu_Q(|r_{X\sim>\phi Y}|/|r|) = 1$ nếu Q là lượng từ tỉ lệ.

3.3.5.2. Phụ thuộc đa trị hàm mờ

Tương tự như phụ thuộc hàm, phụ thuộc đa trị được nhiều tác giả nghiên cứu trên mô hình CSDL mờ. Một số kết quả tiêu biểu về phụ thuộc đa trị của các tác giả sẽ được trình bày.

a. Khái niệm phụ thuộc đa trị mờ của Jyothi và Babu

Dựa vào ý nghĩa của phụ thuộc đa trị trong CSDL quan hệ, các tác giả đưa ra khái niệm phụ thuộc đa trị mờ bằng cách thay thế quan hệ đồng nhất trên dữ liệu rõ bằng quan hệ gần nhau trên dữ liệu mờ, trong đó quan hệ gần nhau thỏa mãn 2 tính chất phản xạ và đối xứng.

Phụ thuộc hàm mờ $X \sim>\sim> Y$ đúng trên quan hệ r nếu và chỉ nếu với mọi t_1, $t_2 \in r$, tồn tại $t_3 \in r$ sao cho : $\tau(t_1[X],t_2[X],t_3[X]) \leq \max(\min(\tau(t_1[Y],t_3[Y]), \tau(t_2[Z],t_3[Z])), \min(\tau(t_2[Y],t_3[Y]), \tau(t_1[Z],t_3[Z])), \tau(t_1[Y],t_2[Y],t_3[Y]), \tau(t_1[Z],t_2[Z],t_3[Z]))$
trong đó $\tau(a,b,c) = \min(\tau(a,b), \tau(b,c), \tau(a,c))$.

b. Khái niệm phụ thuộc đa trị mờ của Bhattacharjee và Mazumdar

Phụ thuộc hàm mờ $X \sim>\sim>_\delta Y$ đúng trên quan hệ r nếu và chỉ nếu với mọi $t \in r$, đặt $x = t[X]$, $z = t[Z]$ ta có $Y_r(x) \approx_\delta Y_r(xz)$, với $Y_r(x) = \{y: \exists t \in r, t[X] = x, t[Y] = y\}$, $Y_r(x) \approx_\delta Y_r(xz)$ khi và chỉ khi $\forall y \in Y_r(x)$ thì $\exists y' \in Y_r(xz)$ sao cho $\tau(y,y') \geq \delta$ và ngược lại, $\delta \in [0,1]$.

c. Khái niệm phụ thuộc đa trị mờ của Hồ Thuần và Trần Thiên Thành

Phụ thuộc hàm mờ $X \underset{(\alpha X,\alpha Y)}{\sim>} Y$ đúng trên quan hệ *r* nếu và chỉ nếu với mọi t_1, t_2 $\in$ *r* nếu $\vec{\tau}(t_1[X], t_2[X]) \geq \alpha X$ thì tồn tại t_3 $\in$ *r* sao cho : $\vec{\tau}(t_1[X], t_3[X]) \geq \alpha X$, $\vec{\tau}(t_1[Y], t_3[Y]) \geq \alpha Y$, $\vec{\tau}(t_1[Z], t_3[Z]) \geq \alpha Z$. Điểm hạn chế của phụ thuộc đa trị này là phải cố định vectơ ngưỡng.

3.3.6. Ngôn ngữ truy vấn dữ liệu

3.3.6.1. Các điều kiện mờ

Trong ngôn ngữ truy vấn, thành phần quan trọng nhất đó là điều kiện dùng để chọn các bộ dữ liệu trong một CSDL nào đó. Trong CSDL mờ, các điều kiện được sử dụng khi truy vấn dữ liệu gọi là điều kiện mờ. Có thể phân tích các thành phần chính trong một điều kiện mờ được biểu diễn bởi các tập mờ như sau:

Vị từ nguyên tố (atomic predicate): là ánh xạ từ tập các miền trị thuộc tính vào [0,1]. Một vị từ nguyên tố thường tương ứng với một giá trị ngôn ngữ như "*già*", "*trẻ*"…

Toán tử sửa đổi : là ánh xạ từ [0,1] vào [0,1] ứng với một từ nhấn như "*rất*", "*có thể*"….

Các toán tử so sánh: là các phép đối sánh giữa các giá trị mờ trên những miền trị, chẳng hạn như phép so sánh "*xấp xỉ*", "*gần nhau*"….

Liên kết lôgc : thường dùng các phép toán hội, tuyển, phủ định.

3.3.6.2 Chọn các bộ thoả mãn điều kiện mờ

Giả sử $r = \{t_1, t_2, \ldots. t_n\}$ là một quan hệ xác định trên *U*, *P* là một vị từ mờ. Ký hiệu $\mu_P(t_i)$ là độ thoả vị từ *P* của bộ t_i ($t_i \in r$). Khi đó, kết quả phép chọn các bộ của quan hệ *r* thoả mãn vị từ P là một tập con mờ, ký hiệu r_P và được xác định:

$$r_P = \left\{ \frac{\mu_P(t_1)}{t_1}, \frac{\mu_P(t_2)}{t_2}, \ldots\ldots\ldots, \frac{\mu_n(t_n)}{t_n} \right\}$$

Trong trường hợp cụ thể, muốn chọn các bộ dữ liệu, thông thường ta đặt ngưỡng $\alpha \in [0,1]$ cho độ thoả vị từ mờ. Kết quả cuối cùng là một lát cắt trên tập mờ r_P mức α.

Chương 4

MÔ HÌNH CƠ SỞ DỮ LIỆU MỜ THEO CÁCH TIẾP CẬN ĐẠI SỐ GIA TỬ

4.1. Mô hình biểu diễn CSDL mờ theo cách tiếp cận Đại số gia tử

Xét một lược đồ CSDL trên miền vũ trụ $U = \{A_1, A_2, \ldots, A_n\}$. Mỗi thuộc tính A_i được gắn với một miền trị thuộc tính, ký hiệu là $Dom(A_i)$, trong đó một số thuộc tính cho phép nhận các giá trị ngôn ngữ trong lưu trữ hay trong các câu truy vấn và được gọi là thuộc tính mờ. Các thuộc tính còn lại được gọi là thuộc tính kinh điển. Thuộc tính kinh điển A_i được gắn với một miền giá trị kinh điển, ký hiệu là D_{A_i}. Thuộc tính mờ A_i sẽ được gắn một miền giá trị kinh điển D_{A_i} và một miền giá trị ngôn ngữ LD_{A_i} hay là tập các phần tử của một ĐSGT. Xem giá trị ngôn ngữ như là một phần tử của ĐSGT. Để bảo đảm tính nhất quán trong xử lý ngữ nghĩa dữ liệu trên cơ sở thống nhất kiểu dữ liệu của thuộc tính mờ, mỗi thuộc tính mờ sẽ được gắn với một ánh xạ định lượng ngữ nghĩa ĐSGT.

Theo cách tiếp cận này giá trị ngôn ngữ là dữ liệu, không phải là nhãn của các tập mờ biểu diễn ngữ nghĩa của giá trị ngôn ngữ và ưu điểm cơ bản của nó là việc cho phép tìm kiếm, đánh giá ngữ nghĩa của thông tin không chắc chắn chỉ bằng các thao tác dữ liệu kinh điển thường dùng và do đó bảo đảm tính thuần nhất của kiểu dữ liệu trong xử lý ngữ nghĩa của chúng.

Vì tất cả các thuộc tính có miền trị chứa giá trị số trong CSDL đều tuyến tính, nên một cách tự nhiên ta giả thiết ĐSGT được sử dụng là ĐSGT tuyến tính, do đó tập H^+ và H^- là tập sắp thứ tự tuyến tính. Như vậy, cho $\underline{X} = (\boldsymbol{X}, G, H, \leq)$ với $G = \{0, c^-, W, c^+, 1\}$, $H = H^- \cup H^+$ với giả thiết $H^- = \{h_1, h_2, \ldots, h_p\}$, $H^+ = \{h_{-1}, \ldots, h_{-q}\}$, $h_1 > h_2 > \ldots > h_p$ và $h_{-1} < \ldots < h_{-q}$ là dãy các gia tử.

Cho một ĐSGT tuyến tính đầy đủ $\mathcal{AX} = (\boldsymbol{X}, G, H, \Sigma, \Phi, \leq)$, trong đó $Dom(\mathcal{X}) = \boldsymbol{X}$ là miền các giá trị ngôn ngữ của thuộc tính ngôn ngữ $\mathcal{X}$ được sinh từ tập các phần tử sinh $G = \{\mathbf{0}, c^-, \mathbf{W}, c^+, \mathbf{1}\}$ bằng việc tác động các gia tử trong tập H, Σ và Φ là hai phép tính với ngữ nghĩa là cận trên đúng và cận dưới đúng của tập $H(x)$, tức là $\Sigma x = supremum\ H(x)$ and $\Phi x = infimum\ H(x)$, quan hệ $\leq$ là quan hệ sắp thứ tự tuyến tính trên $\mathbf{X}$ cảm sinh từ ngữ nghĩa của ngôn ngữ.

4.1.1. Ngữ nghĩa dữ liệu dựa trên việc định lượng Đại số gia tử

4.1.1.1. Đặt vấn đề

Cho một CSDL $DB = \{U ; R_1, R_2, \ldots, R_n ; Const\}$, với $U = \{A_1, A_2, \ldots, A_n\}$ là tập vũ trụ các thuộc tính, $R_1, R_2, \ldots, R_n$ là các lược đồ xác định trên U, Const là tập các ràng buộc trong CSDL. Mỗi thuộc tính A_i được gắn với một miền trị, ký hiệu là $Dom(A_i)$. Thuộc tính kinh điển A_i được gắn với một miền giá trị kinh điển, ký hiệu là D_{A_i}. Thuộc tính mờ A_i sẽ được gắn một miền giá trị kinh điển D_{A_i} và một miền giá trị ngôn ngữ LD_{A_i}. Như vậy, ta có $Dom(A_i) = D_{A_i} \cup LD_{A_i}$, với D_{A_i} là tập các giá trị kinh điển của A_i, LD_{A_i} là tập các giá trị ngôn ngữ của A_i. Tuy nhiên, để rút gọn khi trình bày, trong chương này nếu cho $U = \{A_1, A_2, \ldots, A_n\}$ thì ta cũng gọi U là một lược đồ quan hệ.

Ví dụ 4.1. Cho lược đồ quan hệ U = {*STT*, *TEN*, *SOCTRINH*, *SONCSINH, NAMSINH*} và quan hệ *Lylichkhoa hoc* được xác định như sau:

STT	TEN	SOCTRINH	SONCSINH	NAMSINH
1	Bình	6	1	1950
2	Nhanh	10	2	1953
3	Huế	*nhiều*	*rất nhiều*	1960
4	Hồng	2	3	1975
5	Hà	*khả năng ít*	5	1954
6	Thuỷ	2	*ít*	1950
7	Minh	5	6	1945

Bảng 4.1. Quan hệ Lylichkhoahoc

Trong quan hệ *Lylichkhoahoc*, các thuộc tính STT (Số thứ tự), TEN (Tên), NAMSINH (Năm sinh) được gọi là thuộc tính kinh điển và có miền trị tương ứng D_{STT}, D_{TEN}, D_{NAMSINH}. Các thuộc tính SOCTRINH (Số công trình), SONCSINH (Số nghiên cứu sinh) được gọi là thuộc tính mờ và có miền trị tương ứng $D_{\mathrm{SOCTRINH}} \cup LD_{\mathrm{SOCTRINH}}$, $D_{\mathrm{SONCSINH}} \cup LD_{\mathrm{SONCSINH}}$. Do đó, đối với mô hình CSDL mờ này, các khái niệm như lược đồ, quan hệ, bộ dữ liệu được hiểu tương tự như trong CSDL quan hệ. Tuy nhiên, miền trị của các thuộc tính mờ được xác định là một tập bao gồm miền trị kinh điển và miền giá trị ngôn ngữ được sinh ra khi tác động các gia tử vào các phần tử sinh. Chẳng hạn, trong quan hệ *Lylichkhoahoc*, miền trị thuộc tính LD_{SOCTRINH}, LD_{SONCSINH} chứa hai phần tử ít và nhiều. Vấn đề đặt ra ở đây, tìm một phương pháp đối sánh dữ liệu để ứng dụng thao tác dữ liệu trên miền trị của các thuộc tính mờ. Ví dụ tìm những cán bộ có nhiều công trình khoa học và hướng dẫn rất nhiều nghiên cứu sinh. Nếu chúng ta xem LD_{SOCTRINH}, LD_{SONCSINH} là hai ĐSGT và các giá trị nhiều, rất nhiều thuộc hai ĐSGT đó, thì việc đối sánh dữ liệu trên miền trị của thuộc tính mờ sẽ được dựa trên định lượng ngữ nghĩa của ĐSGT.

Để đề xuất các phép đối sánh dữ liệu trên mô hình CSDL mờ, một số định nghĩa được giới thiệu. Các định lý, hệ quả và bổ đề liên quan được chúng ta trình bày làm cơ sở cho phần tiếp theo.

4.1.1.2. Ngữ nghĩa dữ liệu dựa trên việc định lượng ĐSGT

Trong phần này, các khái niệm như: bằng nhau theo mức *k*, khác nhau theo mức *k* và bé hơn theo mức *k* được trình bày. Về nguyên tắc, chúng ta có thể định nghĩa với mức *k* là số nguyên dương bất kỳ. Tuy nhiên, trong ngôn ngữ tự nhiên, người ta thường chỉ sử dụng một số gia tử tác động liên tiếp, điều này dẫn đến trong CSDL chỉ có một số giới hạn các gia tử tác động liên tiếp vào phần tử sinh. Vì vậy, một cách hợp lý chúng ta giả thiết số gia tử tác động liên tiếp vào phần tử sinh không vượt quá *p* cho trước. Do đó, trong chương này, giá trị *k* được xét là $1 \le k \le p$, với k, p nguyên.

Vì tính mờ của các giá trị trong ĐSGT là một đoạn con của [0,1] cho nên họ các đoạn con như vậy của các giá trị có cùng độ dài sẽ tạo thành phân hoạch của [0,1]. Phân hoạch ứng với các giá trị có độ dài từ lớn hơn sẽ mịn hơn và khi độ dài lớn vô hạn thì độ dài của các đoạn trong phân hoạch giảm dần về 0. Do đó, các phân hoạch được xây dựng dựa trên tính mờ các giá trị trong ĐSGT hay là dựa trên tính mờ các giá trị trong $Dom(A_i)$.

Với A_i là thuộc tính mờ, để đối sánh hai giá trị trong $Dom(A_i)$ ta xây dựng phân hoạch của $Dom(A_i)$. Nếu đặt miền giá trị kinh điển D_{Ai} = [a,b], bằng một phép biến đổi tuyến tính hoặc sử dụng một hàm chuyển đổi nào đó thì ta có thể xem mỗi D_{Ai} = [0,1]. Do đó, xây dựng phân hoạch của $Dom(A_i)$ trở thành xây dựng phân hoạch của [0,1].

Định nghĩa 4.1. Cho $\boldsymbol{X_k} = \{x \in \boldsymbol{X}: |x| = k\}$, xét $P^k = \{I(x): x \in \boldsymbol{X_k}\}$ là một phân hoạch của [0,1]. Gọi υ là hàm định lượng ngữ nghĩa trên $\boldsymbol{X}$.

(1) *u* bằng *v* theo mức *k*, được ký hiệu $u =_k v$, khi và chỉ khi $I(u)$ và $I(v)$ cùng chứa trong một khoảng mờ mức *k*. Có nghĩa là với $\forall u, v \in \boldsymbol{X}$, $u =_k v \Leftrightarrow \exists \Delta^k \in P^k : I(u) \subseteq \Delta^k$ và $I(v) \subseteq \Delta^k$.

(2) *u* khác *v* theo mức *k*, được ký hiệu $u \neq_k v$, khi và chỉ khi $I(u)$ và $I(v)$ không cùng chứa trong một khoảng mờ mức *k*.

(3) *u* nhỏ hơn *v* theo mức *k*, được ký hiệu $u <_k v$, khi và chỉ khi $I(u)$ và $I(v)$ không cùng chứa trong một khoảng mờ mức *k* và $\upsilon(u) < \upsilon(v)$.

Ví dụ 4.2. Cho ĐSGT $\underline{X} = (\boldsymbol{X}, G, H, \leq)$, Trong đó $H = H^+ \cup H^-$, $H^+ = \{$*hơn*, *rất*$\}$, *hơn* < *rất*, $H^- = \{$*ít*, *khả năng*$\}$, *ít* > *khả năng*, $G = \{$ *trẻ*, *già*$\}$. Ta có $P^1 = \{I(\textit{trẻ}), I(\textit{già})\}$ là một phân hoạch của [0,1]. Tương tự, $P^2 = \{I(\textit{hơn trẻ}), I(\textit{rất trẻ}), I(\textit{ít trẻ}), I(\textit{khả năng trẻ}), I(\textit{hơn già}), I(\textit{rất già}), I(\textit{ít già}), I(\textit{khả năng già})\}$ là phân hoạch của [0,1].

(a) Ta có P^1 là phân hoạch của [0,1]. Do đó *hơn trẻ* $=_1$ *rất trẻ* vì $\exists \Delta^1 = I(\textit{trẻ}) \in P^1 : I(\textit{hơn trẻ}) \subseteq \Delta^1$ và $I(\textit{rất trẻ}) \subseteq \Delta^1$.

Ta có P^2 là phân hoạch của [0,1]. Do đó *ít già* $=_2$ *rất ít già* vì $\exists \Delta^2 = I(\textit{ít già}) \in P^2 : I(\textit{ít già}) \subseteq \Delta^2$ và $I(\textit{rất ít già}) \subseteq \Delta^2$.

(b) Ta có P^2 là phân hoạch của [0,1]. Chọn $\Delta^2 = I(\textit{rất trẻ}) \in P^2$, ta có $I(\textit{ít trẻ}) \not\subset \Delta^2$ và $I(\textit{rất trẻ}) \subseteq \Delta^2$ (1').

Mặc khác với mọi $\Delta^2 \neq I(\textit{ít trẻ}) \in P^2$, ta có $I(\textit{ít trẻ}) \not\subset \Delta^2$ và $I(\textit{rất trẻ}) \not\subset \Delta^2$ (2').

Từ (1') và (2') suy ra *ít trẻ* $\neq_2$ *rất trẻ*. Hơn nữa, vì *ít trẻ* $\neq_2$ *rất trẻ* và $\upsilon(\textit{ít trẻ}) > \upsilon(\textit{rất trẻ})$ nên *ít trẻ* $>_2$ *rất trẻ*.

Bổ đề 4.1. Quan hệ $=_k$ là một quan hệ tương đương trong P^k.

Chứng minh: *Tính phản xạ* : Ta chứng minh bằng quy nạp.

$\forall x \in Dom(A_i)$ nếu $|x| = 1$ thì $x = c^+$ hoặc $x = c^-$.

Ta có $\exists \Delta^1 = I(c^+) \in P^1$: $I(c^+) = I(x) \subseteq \Delta^1$ hoặc $\exists \Delta^1 = I(c^-) \in P^1$: $I(c^-) = I(x) \subseteq \Delta^1$.

Vậy $=_k$ đúng với $k = 1$, hay $x =_1 x$.

Giả sử $|x| = n$ đúng, có nghĩa $=_k$ đúng với $k = n$, hay $x =_n x$, ta cần chứng minh $=_k$ đúng với $k = n+1$. Đặt $x = h_1x'$, với $|x'| = n$. Vì $x =_n x$ nên theo định nghĩa ta có: $\exists \Delta^n \in P^n$: $I(x) \subseteq \Delta^n$. Mặc khác ta có $P^{n+1} = \{I(h_1x'), I(h_2x'),\ldots\ldots\}$, với $h_1 \neq h_2 \neq \ldots$là một phân hoạch của $I(x')$. Do đó $\exists \Delta^{(n+1)} = I(h_1x') \in P^{(n+1)}$: $I(h_1x') = I(x) \subseteq \Delta^{(n+1)}$. Vậy $=_k$ đúng với $k = n + 1$, hay $x =_{n+1} x$.

Tính đối xứng: $\forall x, y \in Dom(A_i)$, nếu $x =_k y$ thì theo định nghĩa $\exists \Delta^k \in P^k$: $I(x) \subseteq \Delta^k$ và $I(y) \subseteq \Delta^k$ hay $\exists \Delta^k \in P^k$: $I(y) \subseteq \Delta^k$ và $I(x) \subseteq \Delta^k$. Vậy $y =_k x$ thì $y =_k x$.

Tính bắc cầu: Ta chứng minh bằng phương pháp qui nạp

Trường hợp $k = 1$

Ta có $P^1 = \{I(c^+), I(c^-)\}$, nếu $x =_1 y$ và $y =_1 z$ thì $\exists \Delta^1 = I(c^+) \in P^1$: $I(x) \subseteq \Delta^1$ và $I(y) \subseteq \Delta^1$ và $I(z) \subseteq \Delta^1$ hoặc $\exists \Delta^1 = I(c^-) \in P^1$: $I(x) \subseteq \Delta^1$ và $I(y) \subseteq \Delta^1$ và $I(z) \subseteq \Delta^1$, có nghĩa là $\exists \Delta^1 \in P^1$: $I(x) \subseteq \Delta^1$ và $I(z) \subseteq \Delta^1$ hay $x =_1 z$. Vậy $=_k$ đúng với $k = 1$.

Giả sử quan hệ $=_k$ đúng với trường hợp $k = n$ có nghĩa là ta có $\forall x, y, z \in Dom(A_i)$ nếu $x =_n y$ và $y =_n z$ thì $x =_n z$.

Ta cần chứng minh quan hệ $=_k$ đúng với trường hợp $k = n+1$. Tức là $\forall x, y, z \in Dom(A_i)$ nếu $x =_{n+1} y$ và $y =_{n+1} z$ thì $x =_{n+1} z$.

Theo giả thiết nếu $x =_{n+1} y$ và $y =_{n+1} z$ thì $\exists\ \Delta^{(n+1)} \in P^{(n+1)}$: $I(x) \subseteq \Delta^{(n+1)}$ và $I(y) \subseteq \Delta^{(n+1)}$ và $I(z) \subseteq \Delta^{(n+1)}$, có nghĩa là $\exists\ \Delta^{(n+1)} \in P^{(n+1)}$: $I(x) \subseteq \Delta^{(n+1)}$ và $I(z) \subseteq \Delta^{(n+1)}$. Vậy $x =_{n+1} z$.

Bổ đề 4.2. Cho $u = h_n \ldots h_1 x$ và $v = h'_m \ldots h'_1 x$ là biểu diễn chính tắc của u và v đối với x.

(1): Nếu $u = v$ thì $u =_k v$ với mọi k.

(2): Nếu $h_1 \neq h'_1$ thì $u =_{|x|} v$.

Chứng minh:

(1) Ta có $u =_k u$ và $v =_k v$, vì $u = v$ nên theo bổ đề 2.2 ta có $u =_k v$, với mọi k.

(2) Nếu $|u| = |v| = 2$, tức là $u = h_1x$ và $v = h'_1x$, do $h_1 \neq h'_1$ nên $u \neq v$. Ta có $I(h_1x) \subseteq I(x)$, $I(h'_1x) \subseteq I(x)$ và $I(h_1x) \not\subset I(h'_1x)$ nên $\exists\Delta^1 = I(x) \in P^1$: $I(h_1x) \subseteq \Delta^1$ và $I(h'_1x) \subseteq \Delta^1$ hay $h_1x =_1 h'_1x$. Vậy $u =_{|x|} v$.

Nếu $|u| \neq |v|$, do $h_1 \neq h'_1$ nên $I(h_1x) \not\subset I(h'_1x)$ (1'). Giả sử $\exists\ k > 1$ sao cho $u =_k v$ thì $\exists\Delta^k \in P^k = \{ I(h_{k-1}\ldots h_1x), I(h'_{k-1}\ldots h'_1x)\}$, với P^k là một phân hoạch của $I(x)$: $I(u) \subseteq \Delta^k$ và $I(v) \subseteq \Delta^k$.

Nếu chọn $\Delta^k = I(h_{k-1}\ldots h_1x)$ thì $I(u) \subseteq I(h_{k-1}\ldots h_1x)$ và $I(v) \subseteq I(h_{k-1}\ldots h_1x)$ hay $I(h_n\ldots h_1x) \subseteq I(h_{k-1}\ldots h_1x)$ và $I(h'_m\ldots.h'_1x) \subseteq I(h_{k-1}\ldots h_1x)$ điều này mâu thuẩn vì $I(h'_m\ldots.h'_1x) \not\subset I(h_{k-1}\ldots h_1x)$ do (1').

Nếu chọn $\Delta^k = I(h'_{k-1}\ldots h'_1x)$ thì $I(h_n\ldots h_1x) \subseteq I(h'_{k-1}\ldots h'_1x)$ và $I(h'_m\ldots.h'_1x)$ $I(h'_m\ldots.h'_1x) \subseteq I(h'_{k-1}\ldots h'_1x)$, điều này mâu thuẩn vì $I(h_n\ldots h_1x) \not\subset I(h'_{k-1}\ldots h'_1x)$ do (1'). Vậy không tồn tại $k > 1$ sao cho $u =_k v$ hay $k = 1$. Vậy $u =_{|x|} v$.

Ví dụ 4.3. Cho u = *rất hơn trẻ* và v = *hơn rất trẻ*. Ta có h_1 = *hơn*, h'_1 = *rất*, x = *trẻ*. Vì $h_1 \neq h'_1$ nên theo tính chất (2) của bổ đề 4.2 ta có $u =_{|trẻ|} v$, hay $u =_1 v$.

Định lý 4.1. Cho $\boldsymbol{X}_k = \{x \in \boldsymbol{X}: |x| = k\}$, xét $P^k = \{I(x): x \in \boldsymbol{X}_k\}$ là một phân hoạch của [0,1], $u = h_n\ldots.h_1x$ và $v = h'_m\ldots.h'_1x$ là biểu diễn chính tắc của u và v đối với x.

(1) Nếu $u =_k v$ thì $u =_{k'} v$, $\forall\ 0 < k' < k$.

(2) Nếu tồn tại một chỉ số $j \leq \min(m,n)$ lớn nhất sao cho với mọi s = 1..j ta có $h_s = h'_s$ thì $u =_{j+|x|} v$.

Chứng minh: (1) Ta có $P^k = \{I(h_{k-1}\ldots h_1x), I(h'_{k-1}\ldots h_1x)\}$. Vì $u =_k v$ nên theo định nghĩa $\exists\Delta^k \in P^k$: $I(u) \subseteq \Delta^k$ và $I(v) \subseteq \Delta^k$ (1').

Ta có $P^1 = \{I(x)\}$, $P^2 = \{I(h_1x), I(h'_1x)\}, \ldots P^k = \{I(h_{k-1}\ldots h_1x), I(h'_{k-1}\ldots h_1x)\}$. Mặt khác, $I(h_{k-1}\ldots h_1x) \subseteq I(h_{k-2}\ldots h_1x) \subseteq \ldots. \subseteq I(h_1x) \subseteq I(x)$ và $I(h'_{k-1}\ldots h'_1x) \subseteq I(h'_{k-2}\ldots h'_1x) \subseteq \ldots \subseteq I(h'_1x) \subseteq I(x)$ nên $\exists\Delta^k = I(h_{k-1}\ldots h_1x) \in P^k$ hoặc $\exists\Delta^k = I(h'_{k-1}\ldots h'_1x) \in P^k$ và $\exists\Delta^{k-1} = I(h_{k-2}\ldots h_1x) \in P^{k-1}$ hoặc $\exists\Delta^{k-1} = I(h'_{k-2}\ldots h'_1x) \in P^{k-1}\ldots.$ và $\exists\Delta^2 = I(h_1x) \in P^2$ hoặc $\exists\Delta^2 = I(h'_1x) \in P^2$ và $\exists\Delta^1 = I(x) \in P^1$ sao cho: $\Delta^k \subseteq \Delta^{k-1} \subseteq \ldots \subseteq \Delta^2 \subseteq \Delta^1$ (2').

Từ (1') và (2') ta có $I(u) \subseteq \Delta^k \subseteq \Delta^{k-1} \subseteq \ldots \subseteq \Delta^2 \subseteq \Delta^1$ và $I(v) \subseteq \Delta^k \subseteq \Delta^{k-1} \subseteq \ldots \subseteq \Delta^2 \subseteq \Delta^1$, có nghĩa là $\forall\ 0 < k' < k$ luôn $\exists \Delta^{k'} \in P^{k'}$: $I(u) \subseteq \Delta^{k'}$ và $I(v) \subseteq \Delta^{k'}$. Vậy $\forall\ 0 < k' < k$ nếu $u =_k v$ thì $u =_{k'} v$.

(2): Nếu j =1 ta có $h_1 = h'_1$, khi đó $u = h_n \ldots h_2 h_1 x$ và $v = h'_m \ldots h'_2\, h'_1 x$ hay $u = h_n \ldots h_2 h_1 x$ và $v = h'_m \ldots h'_2 h_1 x$. Đặt $x' = h_1 x$ ta có $u = h_n \ldots h_2 x'$ và $v = h'_m \ldots h'_2 x'$. Vì $h_2 \neq h'_2$ nên theo bổ đề 2.3 ta có $u =_{|x'|} v$ (do $|x'| = 2$, $|x| = 1$) hay $u =_2 v$. Vậy $u =_{j+|x|} v$.

Nếu j $\neq$1, đặt k = j, ta cần chứng minh $u =_{k+|x|} v$. Vì $u =_k v$ nên theo giả thiết ta có $\forall$s =1..k ta có $h_s = h'_s$. Khi đó $u = h_n \ldots h_2 h_1 x$ và $v = h'_m \ldots h'_2\, h'_1 x$ hay $u = h_n . h_k h_{k-1} \ldots h_1 x$ và $v = h'_m \ldots h_k h_{k-1} \ldots h_1 x$.

Đặt $x' = h_k h_{k-1} \ldots h_1 x$ ta có $u = h_n \ldots h_{k+1} x'$ và $v = h'_m \ldots h'_{k+1} x'$. Vì $h_{k+1} \neq h'_{k+1}$ nên theo bổ đề 2.2 ta có $u =_{|x'|} v$ hay $u =_{k+|x|} v$ (do $|x'| = k$, $|x| = 1$).

Ví dụ 4.4. Cho u = *rất rất trẻ* và v = *hơn rất trẻ*. Ta có h_1 = *rất*, h_2 = *rất*, h'_1 = *rất*, h'_2 = *hơn*, x = *trẻ*. Ta thấy tồn tại chỉ số j =1 lớn nhất sao cho $h_1 = h'_1$, do đó theo tính chất (2) của định lý 4.1 ta có $u =_{\text{j}+|\textit{trẻ}|} v$, hay $u =_2 v$.

Hệ quả 4.1. Nếu u $\in$H(v) thì u $=_{|v|}$ v.

Ví dụ 4.5. Cho u = *rất rất trẻ* và v = *rất trẻ*. Vì $u \in H(v)$ nên theo hệ quả 4.1 ta có $u =_{|\textit{rất trẻ}|} v$, hay $u =_2 v$.

Bổ đề 4.3. Cho $X_k = \{x \in \boldsymbol{X}: |x| = k\}$, xét $P^k = \{I(x): x \in \boldsymbol{X}_k\}$ là một phân hoạch của [0,1], $u = h_n \ldots h_1 x$ và $v = h'_m \ldots h'_1 x$ là biểu diễn chính tắc của u và v đối với x.

(1) Nếu tồn tại chỉ số $k \leq$ min(m,n) lớn nhất sao cho $u =_k v$ thì $u \neq_{k+1} v$.

(2) Nếu $u <_k v$ hoặc $u >_k v$ thì với $\forall\ a \in H(u)$, với $\forall\ b \in H(v)$ ta có $a <_k b$ hoặc $a >_k b$.

Ví dụ 4.6. Cho u = *rất rất trẻ* và v = *hơn rất trẻ*. Theo ví dụ 4.4 ta có $u =_2 v$ nên theo bổ đề 4.3 ta có $u \neq_3 v$.

Hệ quả 2.2

(1) Nếu $u \in H(v)$ thì $u \neq_{|v|+1} v$.

(2) Nếu $u \neq_k v$ thì $u \neq_{k'} v\ \forall\ 0 < k < k'$.

Định nghĩa 4.2. Cho $Dom(A_i) = D_{A_i} \cup LD_{A_i}$, υ là hàm định lượng ngữ nghĩa của $Dom(A_i)$. Hàm $f: Dom(A_i) \to [0,1]$ được xác định như sau:

Nếu $LD_{A_i} = \varnothing$ và $D_{A_i} \neq \varnothing$ thì $\forall\omega \in Dom(A_i)$ ta có $f(\omega) = \dfrac{\omega - \psi_{\min}}{\psi_{\max} - \psi_{\min}}$

Nếu $D_{A_i} \neq \varnothing$, $LD_{A_i} \neq \varnothing$ thì $\forall\omega \in Dom(A_i)$ ta có $f(\omega) = \{\omega * \upsilon(\psi_{maxLV})\}/\psi_{\max}$

Với $D_{A_i} = [\psi_{\min}, \psi_{\max}]$ là miền trị kinh điển của A_i và $LD_{A_i} = [\psi_{minLV}, \psi_{maxLV}]$ là miền trị ngôn ngữ của A_i.

Ví dụ 4.7. Cho miền trị cơ sở $U(Tuoi)$ = {0…100, …*rất rất trẻ*,……, *rất rất già*}.

D_{TUOI} = {20, 25, 27, 30, 45, 60, 75, 66, 80}.

LD_{TUOI} = {*trẻ*, *rất trẻ*, *già*, *hơn trẻ*, *hơn già*, *ít già*, *rất già*, *rất rất trẻ*}.

$Dom(TUOI) = D_{TUOI} \cup LD_{TUOI}$. Nếu $LD_{TUOI} = \varnothing$ khi đó $Dom(TUOI) = D_{TUOI}$ = {20, 25, 27, 30, 45, 60, 75, 66, 80}. Do đó $\forall\omega \in Dom(Tuoi)$, chuyển đổi giá trị ω về một số trong [0,1] nhờ hàm $f(\omega)$, ta có $Dom(Tuoi)$ = {0.2, 0.25, 0.27, 0.3, 0.45, 0.6, 0.75, 0.66, 0.8}.

Nếu $D_{A_i} \neq \varnothing$ và $LD_{A_i} \neq \varnothing$ ta có $Dom(Tuoi) = D_{A_i} \cup LD_{Tuoi}$ = {*trẻ*, *rất trẻ*, *già*, *hơn trẻ*, *hơn già*, *ít giá*, *rất già*, *rất rất trẻ*, 20, 25, 27, 30, 45, 60, 75, 66, 80}. Giả sử tính được $\upsilon(\psi_{maxLV}) = \upsilon(\textit{rất rất già}) = 0.98$.

Khi đó $\forall\omega \in D_{A_i}$ ta có $f(\omega) = \{\omega * \upsilon(\psi_{maxLV})\}/\psi_{\max} = (\omega * 0.98)/100$, hay $\forall\omega \in D_{A_i}$ chuyển đổi giá trị ω về một số trong [0,1] nhờ hàm $f(\omega)$, ta có D_{A_i} = {0.196, 0.245, 0.264, 0.294, 0.441, 0.588, 0.735, 0.646, 0.784}.

Nếu chúng ta chọn các tham số W và độ đo tính mờ cho các gia tử sao cho $\upsilon(\psi_{maxLV}) = 1.0$ thì $(\{\omega * \upsilon(\psi_{maxLV})\}/\psi_{\max}) = \dfrac{\omega - \psi_{\min}}{\psi_{\max} - \psi_{\min}}$.

Tiếp theo, chúng ta đi xây dựng một hàm Φ_k để chuyển một giá trị trong [0,1] thành một giá trị ngôn ngữ tương ứng trong ĐSGT $\boldsymbol{X}$.

Định nghĩa 4.3. Cho ĐSGT $\underline{X} = (\boldsymbol{X}, G, H, \leq)$, υ là hàm định lượng ngữ nghĩa của $\boldsymbol{X}$. $\Phi_k: [0,1] \to \boldsymbol{X}$ gọi là hàm ngược của hàm υ theo mức k được xác định:

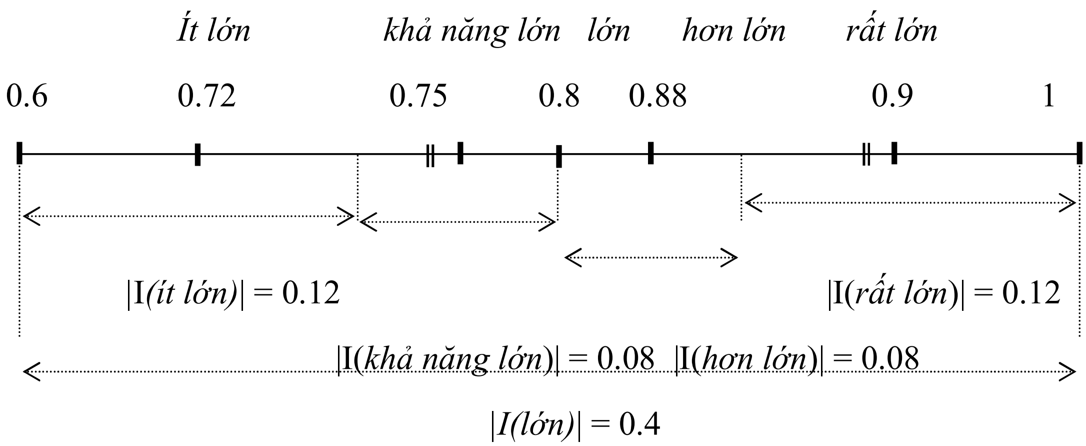
$$\forall a \in [0,1],\ \Phi_k(a) = x^k \text{ khi và chỉ khi } a \in I(x^k), \text{ với } x^k \in \boldsymbol{X}_k.$$

Ví dụ 4.8. Cho ĐSGT $\underline{X} = (\boldsymbol{X}, C, H, \le)$, Trong đó $H = H^+ \cup H^-$.
Trong đó H^+ = {*hơn*, *rất*} với *hơn* < *rất* và H^- = {*ít*, *khả năng*} với *ít* > *khả năng*.
G = {*nhỏ*, *lớn*}. Giả sử cho $\boldsymbol{W} = 0.6$, *fm*(*hơn*) = 0.2, *fm*(*rất*) = 0.3, *fm*(*ít*) = 0.3, *fm*(*khả năng*) = 0.2, *fm*(*nhỏ*) = 0.6, *fm*(*lớn*) = 0.4.

Ta có P^2 = {*I*(*hơn lớn*), *I*(*rất lớn*), *I*(*ít lớn*), *I*(*khả năng lớn*), *I*(*hơn nhỏ*), *I*(*rất nhỏ*), *I*(*ít nhỏ*), *I*(*khả năng nhỏ*)} là phân hoạch của [0,1]. Ta có *fm*(*rất lớn*) = 0.12, *fm*(*khả năng lớn*) = 0.08. Ta có |*I*(*rất lớn*)| = *fm*(*rất lớn*) = 0.12, hay *I*(*rất lớn*) = [0.88,1]. Do đó theo định nghĩa $\Phi_2(0.9)$ = *rất lớn* vì $0.9 \in$ *I*(*rất lớn*).

Tương tự ta có |*I*(*khả năng lớn*)| = *fm*(*khả năng lớn*) = 0.08, hay *I*(*khả năng lớn*) = [0.72,0.8]. Do đó theo định nghĩa $\Phi_2(0.75)$ = *khả năng lớn* vì $0.75 \in$ *I*(*khả năng lớn*).

Trong phần này, giả sử chúng ta chỉ xét các phần tử được sinh từ phần tử *lớn*.



Hình 4.1. *Tính mờ của phần tử sinh lớn*

Định lý 4.2. Cho ĐSGT $\underline{X} = (\boldsymbol{X}, C, H, \le)$, υ là hàm định lượng ngữ nghĩa của $\boldsymbol{X}$, Φ_k là hàm ngược của υ, ta có:

(1) $\forall\, x^k \in X_k$, $\Phi_k(\upsilon(x^k)) = x^k$.

(2) $\forall a \in I(x^k), \forall\, b \in I(y^k)$, $x^k \neq_k y^k$, nếu $a < b$ thì $\Phi_k(a) <_k \Phi_k(b)$.

Chứng minh:

(1) Đặt $a = \upsilon(x^k) \in [0,1]$. Vì $\upsilon(x^k) \in I(x^k)$ nên $a \in I(x^k)$. Vậy, theo định nghĩa ta có $\Phi_k(a) = x^k$, hay $\Phi_k(\upsilon(x^k)) = x^k$.

(2) Vì $a \in I(x^k)$ và $b \in I(y^k)$ nên theo định nghĩa ta có $\Phi_k(a) = x^k$ và $\Phi_k(b) = y^k$.

Mặc khác theo giả thiết $x^k \neq_k y^k$ nên $I(x^k) \neq I(y^k)$. Vì $a < b$ nên $I(x^k) < I(y^k)$, hay $\Phi_k(a) <_k \Phi_k(b)$.

4.1.2. Phương pháp xử lý giá trị khoảng

Trong cơ sở dữ liệu mờ, có nhiều quan hệ mà miền trị của các thuộc tính không phải là giá trị ngôn ngữ, không phải giá trị số mà là giá trị khoảng, chẳng hạn như quan hệ lưu trữ nhiệt độ sốt một căn bệnh của các bệnh nhân trong một bệnh viện nào đó, quan hệ thu nhập cá nhân trong một cơ quan.... Đối với loại dữ liệu này, việc lưu trữ phức tạp nên việc xử lý dữ liệu loại này càng phức tạp hơn. Vì vậy, trong phần này, một phương pháp để xử lý giá trị khoảng giúp cho việc thao tác dữ liệu dễ dàng được trình bày. Trước hết, một ví dụ được xem xét để từ đó phân tích ngữ nghĩa của các giá trị khoảng trong một quan hệ.

Ví dụ 4.8. Cho lược đồ quan hệ $U = \{$ *STT*, *TEN*, *TUOI*, *THUNHAP* $\}$ và quan hệ *Thunhapcanhan* được xác định như sau:

STT	TEN	TUOI	THUNHAP
1	An	30	2.500.000
2	Hải	*Khoảng* 25	1.500.000
3	Hằng	[25,40]	*Khoảng* 3.500.000
4	Phương	[45,50]	[1.500.00,1.800.000]
5	Thúy	45	*Khoảng* 1.000.000

Bảng 4.2. Quan hệ Thunhapcanhan

Chúng ta thấy rằng các giá trị trên thuộc tính *TUOI* và *THUNHAP* rất đa dạng, tuy nhiên, chúng ta chỉ quan tâm đến vấn đề xử lý các giá trị khoảng. Vì vậy, tất cả các giá trị trên quan hệ *Thunhapcanhan* có thể chuyển về các giá trị khoảng tương ứng. Một cách tổng quát, nếu là giá trị a ta chuyển thành [a,a], nếu là giá trị *khoảng a* ta chuyển thành [a-ε, a+ε], với ε được xem là bán kính với tâm a. Nếu giá trị từ *a đến b*, thì được chuyển thành [a,b]. Do đó, quan hệ *Thunhapcanhan* có thể chuyển thành quan hệ sau:

STT	TEN	TUOI	THUNHAP
1	An	[30, 30]	[2.500.000, 2.500.000]
2	Hải	[23, 27]	[1.500.000, 1.500.000]
3	Hằng	[25, 40]	[3.400.000, 3.600.000]
4	Phương	[45, 50]	[1.500.00, 1.800.000]
5	Thúy	[45, 45]	[9.00.00, 1.100.000]

Bảng 4.3. *Quan hệ Thunhapcanhan (sau khi chuyển đổi)*, *với* $\varepsilon_{TUOI} = 2$ và $\varepsilon_{THUNHAP} =$ 100.000

4.1.2.1. Chuyển các giá trị khoảng về đoạn con [0,1] tương ứng

Gọi $Dom(A_i) = [min, max]$ là miền trị kinh điển của thuộc tính mờ A_i trong một quan hệ, trong đó *min*, *max* tương ứng là giá trị nhỏ nhất và giá trị lớn nhất của $Dom(A_i)$. Trước hết, ta sử dụng hàm f để chuyển đổi giá trị thuộc $Dom(A_i)$ thành giá trị thuộc [0,1]. Tiếp theo, khoảng [a,b] được biến đổi thành đoạn con [0,1] tương ứng khi sử dụng hàm f, hay $[f(a), f(b)] \subseteq [0,1]$.

Ví dụ 4.9. Sử dụng quan hệ *Thunhapcanhan* (sau khi chuyển đổi)
Chọn $D_{TUOI} = [0,100]$ và $D_{THUNHAP} = [500.000, 6.000.000]$, khi đó ta có các kết quả chuyển các giá trị khoảng tương ứng về đoạn con của [0,1] như sau:

STT	TEN	TUOI	THUNHAP
1	An	[0.3, 0.3]	[0.36, 0.36]
2	Hải	[0.23, 0.27]	[0.18, 0.18]
3	Hằng	[0.25, 0.4]	[0.52, 0.56]
4	Phương	[0.45, 0.50]	[0.18, 0.23]
5	Thúy	[0.45, 0.45]	[0.07, 0.11]

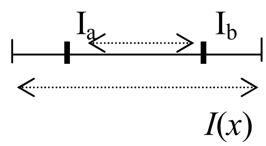
Bảng 4.4. Quan hệ Thunhapcanhan sau khi sử dụng hàm f

4.1.2.2. Đối sánh các giá trị khoảng

Cho ĐSGT $\underline{X} = (\boldsymbol{X}, G, H, \leq)$ và một giá trị khoảng [a,b]. Để so sánh một giá trị $x \in \boldsymbol{X}$ với [a,b], trước hết chuyển [a,b] về đoạn con của [0,1]. Vì tính mờ của x là một đoạn con của [0,1], do đó để so sánh $x \in \boldsymbol{X}$ và đoạn con [0,1], chúng ta chỉ cần dựa vào phần giao của hai đoạn con của [0,1] tương ứng.

Với $x \in \boldsymbol{X}$, ký hiệu $I(x) \subseteq [0,1]$ và $|I(x)| = fm(x)$, $[\mathrm{I_a}, \mathrm{I_b}] = [f(\mathrm{a}), f(\mathrm{b})] \subseteq [0,1]$ tương ứng với việc chuyển đổi giá trị khoảng [a,b] về đoạn con của [0,1].

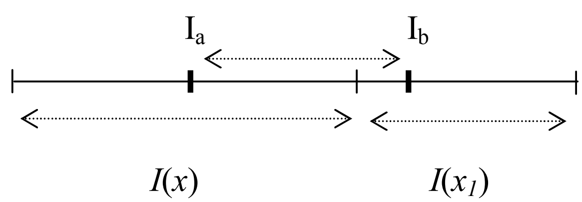
(1) Với mỗi $[\mathrm{I_a}, \mathrm{I_b}]$ nếu tồn tại $x \in \boldsymbol{X}$ sao cho $[\mathrm{I_a}, \mathrm{I_b}] \subseteq I(x)$ thì $[\mathrm{a,b}] =_{|x|} x$.



Hình vẽ 4.2. *Khi* $[\mathrm{I_a}, \mathrm{I_b}] \subseteq I(x)$

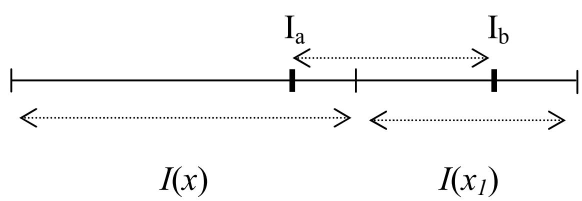
(2) Với mỗi $[\mathrm{I_a}, \mathrm{I_b}]$ sao cho $[\mathrm{I_a}, \mathrm{I_b}] \not\subset I(x)$ $\forall x, x_1 \in \boldsymbol{X}$ thì:

Khi đó với x và x_1, giả sử $x < x_1$ nếu $|[\mathrm{I_a}, \mathrm{I_b}] \cap I(x)| \geq |[\mathrm{I_a}, \mathrm{I_b}]| / £$ thì $[\mathrm{a,b}] =_{|x|} x$.



Hình vẽ 4.3. *Khi* $[\mathrm{I_a}, \mathrm{I_b}] \not\subset I(x)$ (*i*)

ngược lại nếu $|[\mathrm{I_a}, \mathrm{I_b}] \cap \mathrm{I}(\mathrm{x}_1)| \geq |[\mathrm{I_a}, \mathrm{I_b}]| / £$ thì $[\mathrm{a,b}] =_{|x1|} x_1$.

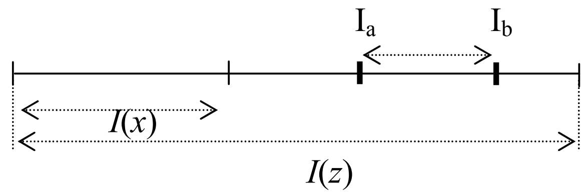


Hình vẽ 4.4. *Khi* $[\mathrm{I_a}, \mathrm{I_b}] \not\subset I(x)$ (*ii*)

với £ là số đoạn $I(x_i) \subseteq [0,1]$ sao cho $[\mathrm{I_a}, \mathrm{I_b}] \cap I(x_i) \neq \varnothing$.

(3) Với mỗi $[\mathrm{I_a}, \mathrm{I_b}]$ nếu tồn tại $x \in \boldsymbol{X}$ sao cho $[\mathrm{I_a}, \mathrm{I_b}] \cap I(x) = \varnothing$ thì:

Nếu tồn tại $z \in \boldsymbol{X}$ sao cho $[\mathrm{I_a}, \mathrm{I_b}] \subseteq I(z)$ và $I(x) \subseteq I(z)$ thì $[\mathrm{a,b}] =_{|z|} x$.



Hình vẽ 4.5. *Khi* $[I_a, I_b] \cap I(x) = \varnothing$

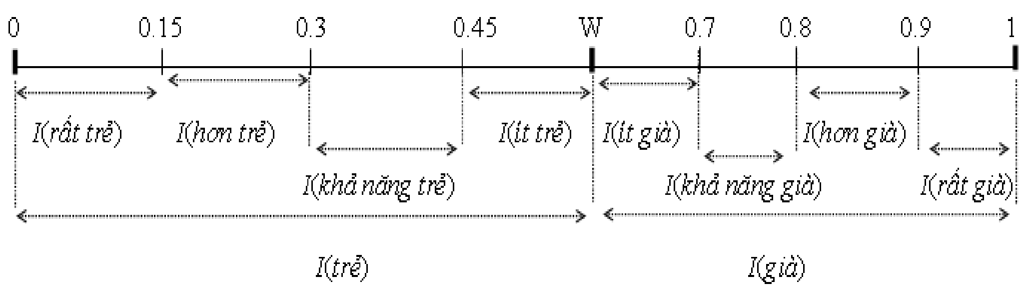
Ví dụ 4.10. Cho ĐSGT $\underline{X}_{Tuoi} = (\boldsymbol{X}_{Tuoi}, G_{Tuoi}, H_{Tuoi}, \leq)$, với G_{Tuoi} = {*trẻ*, *già*}, H^{+}_{Tuoi} = {*rất*, *hơn*}, H^{-}_{Tuoi} = {*khả năng*, *ít*}, *rất* > *hơn* và *ít* > *khả năng*. $\boldsymbol{W}_{\text{Tuoi}}$ = 0.6.

fm(*trẻ*) = 0.6, *fm*(*già*) = 0.4, *fm*(*rất*) = 0.25, *fm*(*hơn*) = 0.25, *fm*(*khả năng*) = 0.25, *fm*(*ít*) = 0.25. Ta có *fm*(*rất trẻ*) = 0.15, *fm*(*hơn trẻ*) = 0.15, *fm*(*ít trẻ*) = 0.15, *fm*(*khả năng trẻ*) = 0.15.

Vì *rất trẻ* < *hơn trẻ* < *trẻ* < *khả năng trẻ* < *ít trẻ* nên *I*(*rất trẻ*) = [0,0.15], *I*(*hơn trẻ*) = [0.15,0.3], *I*(*khả năng trẻ*) = [0.3,0.45], *I*(*ít trẻ*) = [0.45,0.6]

Ta có *fm*(*rất già*) = 0.1, *fm*(*hơn già*) = 0.1, *fm*(*ít già*) = 0.1, *fm*(*khả năng già*) = 0.1.

Vì *ít già* < *khả năng già* < *già* < *hơn già* < *rất già* nên *I*(*ít già*) = [0.6,0.7], *I*(*khả năng già*) = [0.7,0.8], *I*(*hơn già*) = [0.8,0.9], *I*(*rất già*) = [0.9,1].



Hình vẽ 4.6. *Tính mờ của trẻ* và *già*

Vì [0.23,0.27] $\subseteq$ *I*(*hơn trẻ*) mà [*f*(23),*f*(27)] = [0.23,0.27] nên [23,27] $=_2$ *hơn trẻ*. Hay *khoảng* 25 $=_2$ *hơn trẻ*. Tương tự ta có [0.25,0.4] $\cap$ *I*(*hơn trẻ*) = [0.25,0.3] và [0.25,0.4] $\cap$ I(*khả năng trẻ*) = [0.3,0.4].

Mặt khác ta có |[0.25,0.3]| = 0.05, |[0.3,0.4]| = 0.1, |[0.25,0.4]|/2 = 0.075. Vì [*f*(25),*f*(40)] = [0.25,0.4] và |[0.25,0.4] $\cap$ I(*khả năng trẻ*)| $\geq$ |[0.25,0.4]|/2 nên [25,40] $=_2$ *khả năng trẻ*.

4.1.3. Ngữ nghĩa dữ liệu dựa trên lân cận tôpô của ĐSGT

4.1.3.1. Độ tương tự mức k

Chúng ta có thể lấy các khoảng mờ của các phần tử độ dài k làm độ tương tự giữa các phần tử, nghĩa là các phần tử mà các giá trị đại diện của chúng thuộc cùng một khoảng mờ mức k là tương tự mức k. Tuy nhiên, theo cách xây dựng các khoảng mờ mức k, giá trị đại diện của các phần từ x có độ dài nhỏ hơn k luôn luôn là đầu mút của các khoảng mờ mức k. Một cách hợp lý, khi định nghĩa lân cận mức k chúng ta mong muốn các giá trị đại diện như vậy phải là điểm trong (theo nghĩa tôpô) của lân cận mức k. Vì vậy ta định nghĩa độ tương tự mức k như sau:

Chúng ta luôn luôn giả thiết rằng mỗi tập H^- và H^+ chứa ít nhất 2 gia tử. Xét $\boldsymbol{X}_k$ là tập tất cả các phần từ độ dài k. Dựa trên các khoảng mờ mức k và các khoảng mờ mức k+1 chúng ta mô tả không hình thức việc xây dựng một phân hoạch của miền [0,1] như sau :

Với $k = 1$, các khoảng mờ mức 1 gồm $I(c^-)$ và $I(c^+)$. Các khoảng mờ mức 2 trên khoảng $I(c^-)$ là $I(h_pc^-) \le I(h_{p\text{-}1}c^-) \le \dots \le I(h_2c^-) \le I(h_1c^-) \le \upsilon_A(c^-) \le I(h_{-1}c^-) \le I(h_{-2}c^-) \le \dots \le I(h_{-q+1}c^-) \le I(h_{-q}c^-)$. Khi đó, ta xây dựng *phân hoạch về độ tương tự mức* 1 gồm các lớp tương đương sau:

$S(\boldsymbol{0}) = I(h_pc^-)$; $S(c^-) = I(c^-) \setminus [I(h_{-q}c^-) \cup I(h_pc^-)]$; $S(\boldsymbol{W}) = I(h_{-q}c^-) \cup I(h_{-q}c^+)$; và một cách tương tự, $S(c^+) = I(c^+) \setminus [I(h_{-q}c^+) \cup I(h_pc^+)]$ và $S(\boldsymbol{1}) = I(h_pc^+)$.

Ta thấy, trừ hai điểm đầu mút $\upsilon_A(\boldsymbol{0}) = 0$ và $\upsilon_A(\boldsymbol{1}) = 1$, các giá trị đại diện $\upsilon_A(c^-)$, $\upsilon_A(\boldsymbol{W})$ và $\upsilon_A(c^+)$ đều là điểm trong tương ứng của các lớp tương tự mức 1 $S(c^-)$, $S(\boldsymbol{W})$ và $S(c^+)$.

Tương tự, với $k = 2$, ta có thể xây dựng phân hoạch các lớp tương tự mức 2. Chẳng hạn, trên một khoảng mờ mức 2, chẳng hạn, $I(h_ic^+) = (\upsilon_A(\Phi h_ic^+), \upsilon_A(\Sigma h_ic^+)]$ với hai khoảng mờ kề là $I(h_{i\text{-}1}c^+)$ và $I(h_{i+1}c^+)$ chúng ta sẽ có các lớp tương đương dạng sau: $S(h_ic^+) = I(h_ic^+) \setminus [I(h_ph_ic^+) \cup I(h_{-q}h_ic^+)]$, $S(\Phi h_ic^+) = I(h_{-q}h_{i\text{-}1}c^+) \cup I(h_{-q}h_ic^+)$ và $S(\Phi h_ic^+) = I(h_ph_ic^+) \cup I(h_ph_ic^+)$, với i sao cho $-q \le i \le p$ và $i \ne 0$.

Bằng cách tương tự như vậy ta có thể xây dựng các phân hoạch các lớp tương tự mức *k* bất kỳ.

Các giá trị kinh điển và các giá trị ngôn ngữ được gọi là *có độ tương tự mức k* nếu các giá trị đại diện của chúng (ở đây đại diện của giá trị thực là chính nó) cùng nằm trong một lớp tương tự mức *k*.

4.1.3.2. Lân cận mức k của khái niệm mờ

Giả sử phân hoạch các lớp tương tự mức *k* là các khoảng $S(x_1)$, $S(x_2)$, …, $S(x_m)$. Khi đó, mỗi giá trị ngôn ngữ *u* chỉ và chỉ thuộc về một lớp tương tự, chẳng hạn đó là $S(x_i)$ và nó gọi là lân cận mức *k* của *u* và ký hiệu là $\Omega_k(u)$. Dựa trên khái niệm độ tương tự, các quan hệ đối sánh được định nghĩa như sau :

Định nghĩa 4.4. Cho *U* là tập vũ trụ các thuộc tính, *r* là quan hệ xác định trên *U*, giả sử t_1 và t_2 là hai bộ dữ liệu thuộc quan hệ *r*. Ta ký hiệu $t_1[A_i] =_k t_2[A_i]$ và gọi là chúng bằng nhau mức *k*, nếu một trong các điều kiện sau xảy ra :

(1) Nếu $t_1[A_i], t_2[A_i] \in D_{A_i}$ thì $t_1[A_i] = t_2[A_i]$;

(2) Nếu một trong hai giá trị $t_1[A_i]$, $t_2[A_i]$ là khái niệm mờ, chẳng hạn đó là $t_1[A_i]$, thì ta phải có $t_2[A_i] \in \Omega_k(t_1[A_i])$;

(3) Nếu cả hai giá trị $t_1[A_i]$, $t_2[A_i]$ là khái niệm mờ, thì $\Omega_k(t_1[A_i]) = \Omega_k(t_2[A_i])$.

Như thông thường, nếu điều kiện $t_1[A_i] =_k t_2[A_i]$ không xảy ra ta có $t_1[A_i] \neq_k t_2[A_i]$.

Do quan hệ tương tự mức *k* được xây dựng bằng một phân hoạch của đoạn [0,1], nên có thể thấy quan hệ $=_k$ là tương đương trên [0,1]. Ngoài ra, ta cần nhấn mạnh rằng đẳng thức $t_1[A_i] =_k t_2[A_i]$ có nghĩa $L_k \leq t_1[A_i], t_2[A_i] \leq R_k$, trong đó L_k và R_k là hai điểm mút của khoảng $\Omega_k(t_1[A_i])$ hay $\Omega_k(t_2[A_i])$. Nghĩa là, việc kiểm chứng $t_1[A_i] =_k t_2[A_i]$ được đưa về việc kiểm chứng các quan hệ đối sánh kinh điển. Hơn nữa, tính mềm dẻo trong thích nghi với các ứng dụng cụ thể có thể đạt được bằng việc điều chỉnh các tham số của ánh xạ định lượng υ_{A_i}. Đây chính là ưu điểm nổi bật của cách tiếp cận đại số đến thông tin mờ. Dựa trên quan hệ tương đương này ta có thể dễ dàng định nghĩa các quan hệ đối sánh khác. Trước hết, để đơn giản ta quy ước là ký pháp $\Omega_k(t[A_i])$ có nghĩa cả khi $t[A_i] \in D_{Ai}$. Khi đó $\Omega_k(t[A_i])$ được hiểu là tập bao gồm chỉ đúng một giá trị thực $t[A_i]$. Với quy ước đó, với mọi cặp lân cận mức k, $\Omega_k(x)$ and $\Omega_k(y)$, ta sẽ viết $\Omega_k(x) < \Omega_k(y)$ khi $u < v$, với mọi $u \in \Omega_k(x)$ và mọi $v \in \Omega_k(y)$.

Định nghĩa 4.5. Cho U là tập vũ trụ các thuộc tính, r là quan hệ xác định trên U, giả sử t_1 và t_2 là hai bộ dữ liệu thuộc quan hệ r. Khi đó,

(1) Ta viết $t_1[A_i] \leq_k t_2[A_i]$, nếu $t_1[A_i] =_k t_2[A_i]$ hoặc $\Omega_k(t_1[A_i]) < \Omega_k(t_2[A_i])$;

(2) Ta viết $t_1[A_i] <_k t_2[A_i]$, nếu $\Omega_k(t_1[A_i]) < \Omega_k(t_2[A_i])$;

(3) Ta viết $t_1[A_i] >_k t_2[A_i]$, nếu $\Omega_k(t_1[A_i]) > \Omega_k(t_2[A_i])$.

Sau đây là định lý khẳng định họ các khoảng $\Omega_k(x)$ là một phân hoạch của $Dom(A_i)$ và giá trị định lượng của $x \in \boldsymbol{X}$ luôn là điểm trong của lân cận mức k của x.

Định lý 4.3. Cho một ĐSGT tuyến tính đầy đủ, tập các gia tử H^- và H^+ có ít nhất hai phần tử. Khi đó, họ các khoảng $\{\Omega_k(x): x \in \boldsymbol{X}\}$ được gọi là lân cận mức k của miền trị ngôn ngữ của thuộc tính A_i và là một phân hoạch của $Dom(A_i)$. Hơn nữa, mỗi giá trị x của A_i có duy nhất một lân cận mức k, $\upsilon_{A_i}(x)$ là điểm trong của $\Omega_k(x)$ với mọi $x \in \boldsymbol{X}$.

Mệnh đề 4.1. Quan hệ $=_k$ là tương đương trên $Dom(A_i)$.

4.2. Phụ thuộc dữ liệu trong cơ sở dữ liệu mờ

4.2.1. Phụ thuộc hàm mờ

Như chúng ta đã biết, trong mô hình quan hệ, hai dạng phụ thuộc dữ liệu quan trọng giúp cho việc chuẩn hoá tốt các CSDL là phụ thuộc hàm và phụ thuộc đa trị. Khi mở rộng mô hình quan hệ để có thể biểu diễn và xử lý được những thông tin không chắc chắn, không đầy đủ gọi chung là *dữ liệu mờ* đã có rất nhiều công trình tập trung nghiên cứu mở rộng hai dạng phụ thuộc này trên mô hình mới. Đối với mô hình trong các công trình này là sự mở rộng mô hình quan hệ theo hai cách : *mở rộng ngữ nghĩa* và *mở rộng miền trị của thuộc tính*. Tuy nhiên, cách mở rộng miền trị của thuộc tính là tốt hơn mở rộng ngữ nghĩa, bởi vì, cách mở rộng này cho phép bổ sung thêm các cú pháp trong biểu diễn dữ liệu nhằm cho phép biểu diễn được dữ liệu mờ. Vì thế, vấn đề mở rộng miền trị của thuộc tính, ngoài việc đưa ký hiệu vào hệ thống, việc quan trong hơn là giải quyết vấn đề ngữ nghĩa của các ký hiệu.

Như vậy, khái niệm phụ thuộc hàm mờ (*fuzzy functional dependencies*) được nhiều tác giả nghiên cứu phát triển dựa trên ý nghĩa của khái niệm phụ thuộc hàm cổ điển với nhiều cách tiếp cận khác nhau. Tuy nhiên, các cách tiếp cận mở rộng phụ thuộc hàm kinh điển này dựa vào 2 nguyên tắc chính:

Nguyên tắc thứ nhất (*mở rộng ký hiệu*)*:* Nguyên tắc mở rộng này thay cho quan hệ bằng nhau trên dữ liệu rõ bởi quan hệ gần nhau hoặc quan hệ tương tự trên dữ liệu mờ và đặt ngưỡng để xác định độ gần nhau.

Nguyên tắc thứ hai (*mở rộng ngữ nghĩa*)*:* Nguyên tắc này dựa vào ý nghĩa của các phụ thuộc dữ liệu để xây dựng định nghĩa tương ứng cho mô hình mới sao cho bảo toàn một số kết quả quan trọng đã được xây dựng trong mô hình quan hệ.

Ví dụ 4.11. Với cách tiếp cận mở rộng ngữ nghĩa, một phụ thuộc hàm mờ $X \sim> Y$ thoả trên quan hệ r khi và chỉ khi độ gần nhau của dữ liệu của các bộ trên tập thuộc tính X kéo theo độ gần nhau của các bộ trên tập thuộc tính Y. Do đó, phép kéo theo mờ đóng vai trò quan trọng trong cách tiếp cận này.

Với mô hình CSDL mờ được xây dựng trong mục 4.1, một số dạng phụ thuộc dữ liệu mờ trong mô hình này sẽ được đề xuất.

Xét CSDL mờ $\{U; R_1, R_2 \ldots, R_n, Const\}$, trong đó $U = \{A_1, A_2, \ldots, A_n\}$ là tập vũ trụ các thuộc tính, *const* là tập các ràng buộc dữ liệu. Một khi ngữ nghĩa của CSDL được mở rộng, như cho phép lưu trữ trong CSDL các thông tin không chắc chắn hay cho phép các câu truy vấn chứa các thông tin như vậy, khi đó ngữ nghĩa của các phụ thuộc dữ liệu cũng thay đổi, nghĩa là phải mở rộng định nghĩa các dạng phụ thuộc dữ liệu.

Trong thực tế, chúng ta thường gặp các tri thức dạng như *Nếu một tập thể T_1 và T_2 lao động chăm chỉ như nhau và Tính kỷ luật lao động là tốt thì Thu nhập của tập thể T_1 và T_2 cao như nhau*. Ở đây ta không nhìn nhận mối quan hệ trên như là một luật của một cơ sở tri thức nào đó mà xem như là mối quan hệ giữa các thuộc tính trong CSDL với thuộc tính *Số ngày làm việc trong tháng*, *Tính kỷ luật lao động* và *Thu nhập*. Hoặc trong một trường hợp khác *Nếu một tập thể T_1 và T_2 lao động không chăm chỉ và Tính kỷ luật kém giống* *nhau thì Thu nhập của tập thể T_1 và T_2 thấp như nhau*. Trong cả hai trường hợp trên, mối quan hệ giữa các thuộc tính là không chính xác không giống như mối quan hệ của các phụ thuộc kinh điển, và những phụ thuộc như vậy được gọi là phụ thuộc hàm mờ. Ta giả thiết rằng, trong một trường hợp cụ thể, mức độ xấp xỉ mức *k* được xác định ứng với mỗi thuộc tính mờ phù hợp với các phụ thuộc dữ liệu mờ. Nghĩa là, trên cùng một CSDL mờ, có thể có những hệ phụ thuộc dữ liệu mờ, tùy theo quan điểm khai thác dữ liệu của người dùng.

Với $X \subseteq U$, *r* là quan hệ xác định trên *U*, t_1 và t_2 là hai bộ thuộc *r*. Ta nói rằng bộ t_1 và t_2 *bằng nhau mức k* trên tập *X*, ký hiệu $t_1[X] =_k t_2[X]$, nếu với mọi $A \in X$, ta có $t_1[A] =_k t_2[A]$.

Định nghĩa 4.6. Cho *U* là một lược đồ quan hệ, *r* là một quan hệ xác định trên *U*, xét X, $Y \subseteq U$. Ta nói rằng, quan hệ *r* thỏa mãn phụ thuộc hàm mờ *X* xác định *Y* với mức *k*, ký hiệu là $X \sim>_k Y$ nếu ta có: với $\forall\ t_1, t_2 \in r$, $t_1[X] =_k t_2[X] \Rightarrow t_1[Y] =_k t_2[Y]$.

Khi đó, ta cũng nói *r* đúng với phụ thuộc hàm mờ $X \sim>_k Y$, hay ta $X \sim>_k Y$ thỏa trong quan hệ *r*.

Ví dụ 4.12. Ta xét lược đồ quan hệ *U* = { *MASO, TENCN, SONLV, THUNHAP* } với ý nghĩa: Mã số công nhân (*MASO*), Tên công nhân (*TENCN*) là 2 thuộc tính kinh điển, Số ngày làm việc trong tháng (*SONLV*), Thu nhập (*THUNHAP*) là 2 thuộc tính mờ. Trong đó D_{SONLV} = [0, 30] và $D_{THUNHAP}$ = [0, 100]. LD_{SONLV} và $LD_{THUNHAP}$ có cùng tập các xâu giống nhau với tập các phần tử sinh là {***0***, *thấp*, ***W***, *cao*, ***1***} và tập các gia tử là {*ít*, *khả năng*, *hơn*, *rất*}. Mặc dù các thuộc tính ngôn ngữ đang xét có cùng tập các xâu, nhưng ngữ nghĩa định lượng của chúng khác nhau.

(**a**). ***Đối với thuộc tính*** *SONLV*: *fm*(*cao*) = 0.35, *fm*(*thấp*) = 0.65, μ(*khả năng*) = 0.25, μ(*ít*) = 0.20, μ(*hơn*) = 0.15 và μ(*rất*) = 0.40. Ta phân hoạch đoạn [0, 30] thành 5 khoảng tương tự mức 1 là: *fm*(*rất cao*) × 30 = 0.35 × 0.35 × 30 = 3.675. Vậy *S*(***1***) × 30 = (26.325, 30];

(*fm*(*khả năng cao*) + *fm*(*hơn cao*)) × 30 = (0.25 × 0.35 + 0.15 × 0.35) × 30 = 4.2 và *S*(*cao*) × 30 = (22.125, 26.325];

(*fm*(*ít thấp*) + *fm*(*ít cao*)) × 30 = (0.25 × 0.65 + 0.25 × 0.35) × 30 = 7.5 và *S*(***W***) × 30 = (14.625, 22.125];

(*fm*(*khả năng thấp*) + *fm*(*hơn thấp*)) × 30 = (0.25 × 0.65 + 0.15 × 0.65) × 30 = 7.8 và *S*(*thấp*) × 30 = (6.825, 14.625], *S*(***0***) × 30 = [0, 6.825].

(**b**). ***Đối với thuộc tính*** *THUNHAP*: *fm*(*cao*) = 0.6, *fm*(*thấp*) = 0.4, μ(*khả năng*) = 0.15, μ(*ít*) = 0.25, μ(*hơn*) = 0.25 và μ(*rất*) = 0.35. Ta phân hoạch đoạn [0, 100] thành 5 khoảng tương tự mức 1 là: *fm*(*rất cao*) × 100 = 0.35 × 0.6 × 100 = 21. Vậy *S*(***1***) × 100 = (79, 100];

(*fm*(*khả năng cao*) + *fm*(*hơn cao*)) × 100 = (0.25 × 0.6 + 0.15 × 0.6) × 100 = 24 và *S*(*cao*)) × 100 = (55, 79];

(*fm*(*ít thấp*) + *fm*(*ít cao*)) × 100 = (0.25 × 0.6 + 0.25 × 0.4) × 100 = 25 và *S*(***W***) × 100 = (30, 55];

(*fm*(*khả năng thấp*) + *fm*(*hơn thấp*)) × 100 = (0.25 × 0.4 + 0.15 × 0.4) × 100 = 16 và *S*(*thấp*) × 100 = (14, 30], *S*(***0***) × 100 = [0, 14].

Quan hệ *Chamcong* trong ví dụ này được cho ở bảng 4.5

MASO	TENCN	SONLV	THUNHAP
N1	An	27	90
N2	Cường	17	25
N3	Hà	28	94
N4	Hương	*cao*	*cao*
N5	Lan	24	72
N6	Kiên	*cao*	*cao*
N7	Thanh	20	*thấp*
N8	Thủy	21	29
N9	Yến	*thấp*	11

Bảng 4.5. *Quan hệ Chamcong*

Chúng ta có thể thấy rằng phụ thuộc hàm mờ $SONLV \sim>_1 THUNHAP$ đúng trong quan hệ *Chamcong*.

Gọi $\mathcal{F}_k$ là họ tất cả các phụ thuộc hàm mờ $X \sim>_k Y$ trên lược đồ quan hệ *U*. Ta ký hiệu $\mathcal{F}_k^+$ là tập tất cả các phụ thuộc hàm mờ $X \sim>_k Y$ mà được *suy diễn* từ $\mathcal{F}_k$, tức là với mọi quan hệ *r* trên *U*, nếu *r* thỏa các thụ thuộc dữ liệu trong $\mathcal{F}_k$ thì r cũng thỏa $X \sim>_k Y$. Ta có thể dễ dàng kiểm chứng họ các phụ thuộc hàm mờ $\mathcal{F}_k^+$ thoả các tiên đề sau:

Định lý 4.4. Trong CSDL mờ với tập vũ trụ các thuộc tính U, họ $\mathcal{F}_k^+$ thỏa mãn các tiên đề sau:

(1) *Phản xạ*: $X \sim>_k X \in \mathcal{F}_k^+$.

(2) *Gia tăng*: $X \sim>_k Y \in \mathcal{F}_k^+ \Rightarrow XZ \sim>_k YZ \in \mathcal{F}_k^+$.

(3) *Bắc cầu*: $X \sim>_k Y \in \mathcal{F}_k^+$, $Y \sim>_k Z \in \mathcal{F}_k^+ \Rightarrow X \sim>_k Z \in \mathcal{F}_k^+$.

(4) *Bao hàm mức*: $X \sim>_k Y \in \mathcal{F}_k^+ \Rightarrow X \sim>_{k'} Y \in \mathcal{F}_k^+$ với mọi $0 < k' \leq k$.

Các tiên đề (1)-(4) trong định lý 4.4 là đúng đắn và đầy đủ.

Chứng minh: Tính đúng đắn của hệ tiên đề (1)-(4)

(1): Hiển nhiên ta có với $\forall t_1, t_2 \in r$, $t_1[X] =_k t_2[X] \Rightarrow t_1[X] =_k t_2[X]$, hay $X \sim>_k X \in \mathcal{F}_k^+$.

(2): Vì $X \sim>_k Y \in \mathcal{F}_k^+$ (giả thiết) nên theo định nghĩa ta có với $\forall t_1, t_2 \in r$, $t_1[X] =_k t_2[X] \Rightarrow t_1[Y] =_k t_2[Y]$. Từ $t_1[XZ] =_k t_2[XZ]$ và $t_1[X] =_k t_2[X]$ suy ra $t_1[Z] =_k t_2[Z]$ (1').

Từ $t_1[Y] =_k t_2[Y]$ và $t_1[Z] =_k t_2[Z]$ suy ra $t_1[YZ] =_k t_2[YZ]$ (2'). Vậy, từ (1'), (2') ta có $\forall t_1, t_2 \in r$, $t_1[XZ] =_k t_2[XZ] \Rightarrow t_1[YZ] =_k t_2[YZ]$, hay $XZ \sim>_k YZ \in \mathcal{F}_k^+$.

(3): Vì $X \sim>_k Y \in \mathcal{F}_k^+$ (giả thiết) nên theo định nghĩa ta có với $\forall\ t_1, t_2 \in r$, $t_1[X] =_k t_2[X] \Rightarrow t_1[Y] =_k t_2[Y]$ (1') và $Y \sim>_k Z \in \mathcal{F}_k^+$ (giả thiết) nên theo định nghĩa ta có với $\forall t_1, t_2 \in r$, $t_1[Y] =_k t_2[Y] \Rightarrow t_1[Z] =_k t_2[Z]$ (2'). Từ (1'), (2') ta có với $\forall\ t_1, t_2 \in r$, $t_1[X] =_k t_2[X] \Rightarrow t_1[Z] =_k s[Z]$, hay $X \sim>_k Z \in \mathcal{F}_k^+$.

(4): Vì $X \sim>_k Y \in \mathcal{F}_k^+$ (giả thiết) nên theo định nghĩa ta có với $\forall t_1, t_2 \in r$, $t_1[X] =_k t_2[X] \Rightarrow t_1[Y] =_k t_2[Y]$. Ta có $\forall t_1, t_2 \in r$, $t_1[X] =_k t_2[X] \Rightarrow t_1[X] =_{k'} t_2[X]$ với mọi $0 < k' \leq k$ (1'). Ta có với $\forall t_1, t_2 \in r$, $t_1[Y] =_k t_2[Y] \Rightarrow t_1[Y] =_{k'} t_2[Y]$ với mọi $0 < k' \leq k$ (2'). Từ (1'), (2') ta có với $\forall t_1, t_2 \in r$, $t_1[X] =_{k'} t_2[X] \Rightarrow t_1[Y] =_{k'} t_2[Y]$ với mọi $0 < k' \leq k$, hay $X \sim>_{k'} Y \in \mathcal{F}_k^+$ với mọi $0 < k' \leq k$.

Tính đầy đủ của hệ tiên đề (1)-(4)

Chúng ta dễ dàng thấy rằng trong một quan hệ 2-bộ r với các bộ t_1 và t_2 chỉ chứa các giá trị lớn nhất và nhỏ nhất trong miền giá trị thuộc tính, điều kiện $t_1[X] =_k t_2[X]$ và $t_1[Y] =_k t_2[Y]$ là tương đương với điều kiện $t_1[X] = t_2[X]$ và $t_1[Y] = t_2[Y]$, vì khi đó các khoảng phân hoạch (do nó chứa ít nhất hai lớp tương đương) của các thuộc tính mờ chỉ chứa một trong hai giá trị này. Do đó, nếu quan hệ 2-bộ r như vậy thỏa một phụ thuộc hàm mờ nào đó ở mức k, thì nó cũng thỏa phụ thuộc đó khi được xem như là một phụ thuộc hàm kinh điển. Vì vậy, cũng như đối với họ phụ thuộc hàm kinh điển, nếu $X \sim>_k Y \notin \mathcal{F}_k^+$ thì tồn tại một quan hệ 2-bộ r sao cho r thỏa $\mathcal{F}_k$ nhưng không thỏa phụ thuộc hàm mờ $X \sim>_k Y$. Có nghĩa là, hệ tiên đề (1)-(4) trong định lý 4.4 là đầy đủ.

Như vậy, chúng ta thấy ở mức cú pháp, phụ thuộc hàm mờ trùng với phụ thuộc hàm kinh điển nhưng ngữ nghĩa khác nhau, đặc biệt một quan hệ r có thể thỏa mãn một phụ thuộc hàm mờ $X \sim>_k Y$ nào đó nhưng không nhất thiết phải thỏa mãn $X \sim>_k Y$ với tư cách là một phụ thuộc hàm kinh điển.

Vì phụ thuộc hàm kinh điển là trường hợp riêng của phụ thuộc hàm mờ, nên ta có mệnh đề thể hiện mối quan hệ giữa hai loại phụ thuộc này.

Mệnh đề 4.2

(1) Nếu $X \to Y$ và $Y \sim>_k Z \in \mathcal{F}_k^+$ thì $X \sim>_k Z \in \mathcal{F}_k^+$.

(2) Nếu $X \sim>_k Y \in \mathcal{F}_k^+$ và $Y \to Z$ thì $X \sim>_k Z \in \mathcal{F}_k^+$.

Chứng minh :

(1): Vì $X \to Y$ nên với $\forall t_1, t_2 \in r$, $t_1[X] = t_2[X] \Rightarrow t_1[Y] = t_2[Y]$. Mặc khác, ta có $t_1[X] = t_2[X] \Rightarrow t_1[X] =_k t_2[X]$ và $t_1[Y] = t_2[Y] \Rightarrow t_1[Y] =_k t_2[Y]$, do đó $X \sim>_k Y \in \mathcal{F}_k^+$ (1’).

Theo giả thiết, ta có $Y \sim>_k Z \in \mathcal{F}_k^+$ (2’). Từ (1’), (2’) ta suy ra $X \sim>_k Z \in \mathcal{F}_k^+$.

(2): Vì $Y \to Z$ nên với $\forall t_1, t_2 \in r$, $t_1[Y] = t_2[Y] \Rightarrow t_1[Z] = t_2[Z]$. Mặc khác, ta có $t_1[Y] = t_2[Y] \Rightarrow t_1[Y] =_k t_2[Y]$ và $t_1[Z] = t_2[Z] \Rightarrow t_1[Z] =_k t_2[Z]$, do đó $Y \sim>_k Z \in \mathcal{F}_k^+$ (1’).

Theo giả thiết, ta có $X \sim>_k Y \in \mathcal{F}_k^+$ (2’). Từ (1’), (2’) ta suy ra $X \sim>_k Z \in \mathcal{F}_k^+$.

Một phụ thuộc hàm mờ $X \sim>_k Y$ đúng trong quan hệ r có nghĩa là “*với mọi*” hai bộ dữ liệu bất kỳ thuộc quan hệ r, nếu giá trị trên tập thuộc tính X bằng nhau theo mức k thì giá trị trên tập thuộc tính Y bằng nhau theo mức k. Tuy nhiên, trong thực tế, khi xét một quan hệ nào đó, có thể “*tồn tại*” hai bộ dữ liệu mà giá trị trên tập thuộc tính X bằng nhau theo mức k nhưng giá trị trên tập thuộc tính Y khác nhau theo mức k. Như vậy, ở đây không tồn tại phụ thuộc hàm mờ, bởi vì nó không thoả mãn “*với mọi*” nhưng có thể thoả mãn "*hầu hết*" hoặc "*một ít*", các dạng phụ thuộc này được gọi là phụ thuộc hàm mờ với lượng từ ngôn ngữ.

4.2.2. Phụ thuộc hàm mờ với lượng từ ngôn ngữ

4.2.2.1. Đặt vấn đề

Chúng ta thường gặp những tri thức dạng: *trong cơ quan những cán bộ có kinh nghiệm làm việc xấp xỉ nhau thì có thu nhập xấp xỉ nhau*. Đối với dạng tri thức như vậy, trong phần 4.2.1 chúng ta đã nghiên cứu và gọi đó là phụ thuộc hàm mờ. Ở phụ thuộc hàm mờ này có ý nghĩa là *với mọi* hai cán bộ bất kỳ trong cơ quan nếu *có kinh nghiệm làm việc xấp xỉ nhau thì có thu nhập xấp xỉ nhau*. Tuy nhiên, trong thực tế có những cán bộ *có kinh nghiệm làm việc xấp xỉ nhau nhưng có thu nhập khác nhau* do nhiều yếu tố khác tác động như: chủ nhiệm đề tài nghiên cứu cơ bản, kiêm nhiệm các chức vụ chủ chốt trong cơ quan…..Do đó các tri thức thỏa mãn *với mọi* đòi hỏi khá chặt về ràng buộc dữ liệu trong CSDL.

Vì vậy, việc sử dụng các lượng từ ngôn ngữ như *một vài*, *hầu hết*… vào trong phụ thuộc hàm mờ làm cho việc mô tả các phụ thuộc dữ liệu được mềm dẽo và thực tế hơn, chẳng hạn như: *hầu hết trong cơ quan những cán bộ có kinh nghiệm làm việc xấp xỉ nhau thì có thu nhập xấp xỉ nhau*.

4.2.2.2. Phụ thuộc hàm mờ với lượng từ ngôn ngữ

Trước tiên, phương pháp định giá lượng từ ngôn ngữ được trình bày trước khi xây dựng dạng phụ thuộc dữ liệu.

a. Phương pháp định giá lượng từ ngôn ngữ

Zadeh chia lượng từ ngôn ngữ thành hai loại đó là: lượng từ tuyệt đối và lượng từ tỉ lệ. Lượng từ tuyệt đối thường dùng trong các mệnh đề có số lượng xác định như *"ít nhất 5"*, *"nhiều hơn 3"*.... Lượng từ tỉ lệ thể hiện những số lượng phụ thuộc vào số lượng tập các đối tượng đang xử lý, chẳng hạn như *"hầu hết"*, *"một vài"*....

Gọi $D_r = [0..\|r\|]$, trong đó $\|r\|$ là số bộ dữ liệu trong quan hệ r. Chúng ta có thể chia lượng từ thành hai trường hợp:

Trường hợp Q là lượng từ tuyệt đối: Ký hiệu $\|Q\|$ là số lượng xác định lượng từ Q.

Nếu Q đơn điệu tăng : Ta xây dựng một hàm f_Q^A: $D_r \rightarrow \{0, 1\}$ sao cho: $\forall x \in D_r, f_Q^A(x) = 1$ nếu $x \geq \|Q\|$ và $f_Q^A(x) = 0$ nếu ngược lại.

Nếu Q đơn điệu giảm : Ta xây dựng một hàm f_Q^D: $D_r \rightarrow \{0, 1\}$ sao cho: $\forall x \in D_r, f_Q^D(x) = 1$ nếu $x \leq \|Q\|$ và $f_Q^D(x) = 0$ nếu ngược lại.

Trường hợp Q là lượng từ tỷ lệ: Khi ta nói *hầu hết* các bộ dữ liệu t trong r thỏa mãn điều kiện ($fc_1, fc_2, ...fc_n$), có nghĩa là tổng số bộ dữ liệu t phải xấp xỉ $\|r\|$. Hoặc trong trường hợp khác, chỉ *một ít* các bộ dữ liệu t trong r thỏa mãn điều kiện ($fc_1, fc_2, ...fc_n$), có nghĩa là tổng số bộ dữ liệu t phải xấp xỉ $1/\|r\|$. Hay một giả thiết ta thường gặp đó là *khoảng một nửa* các bộ dữ liệu t trong r thỏa mãn điều kiện ($fc_1, fc_2, ...fc_n$), khi đó chắc chắn rằng tổng số bộ dữ liệu t phải là xấp xỉ của $\|r\|/2$.

Điều này gợi ý cho chúng ta có thể đánh giá lượng từ tỉ lệ dựa trên sự phân hoạch của $[0..\|r\|]$. Theo mục 2.1 để chuẩn hóa $[0..\|r\|]$, nhờ một phép biến đổi tuyến tính, ta giả thiết mọi miền $D_r = [0..\|r\|]$ như vậy đều là khoảng [0,1]. Khi đó ta xây dựng hai khoảng mờ của hai khái niệm nguyên thủy *nhỏ* và *lớn*, ký hiệu là I(*nhỏ*) và I(*lớn*) với độ dài tương ứng là fm(*nhỏ*) và fm(*lớn*) sao cho chúng tạo thành một phân hoạch của miền tham chiếu [0,1]. Tiếp đến, đi xây dựng các lớp tương đương $S(\boldsymbol{1})$, S(*lớn*), $S(\boldsymbol{W})$, S(*nhỏ*), $S(\boldsymbol{0})$ dựa vào độ đo tính mờ của các gia tử và các khái niệm nguyên thủy.

Do đó, nếu gọi $\|r_1\|$, $\|r_2\|$ tương ứng là tổng số bộ dữ liệu t trong r thỏa mãn điều kiện ($fc_1, fc_2, ...fc_n$) với lượng từ *hầu hết* và *một ít* thì $\|r_1\| \in S(\boldsymbol{1}) \times \|r\|$ và $\|r_2\| \in S(\boldsymbol{0}) \times \|r\|$.

Như vậy, ta có thể khẳng định rằng tổng số bộ dữ liệu t trong r thỏa mãn điều kiện ($fc_1, fc_2, ...fc_n$) áp dụng với lượng từ Q được ký hiệu $\|r_Q\|$, khi đó: $\|r_Q\| \in S(\boldsymbol{1}) \times \|r\|$ hoặc $\|r_Q\| \in S(lớn) \times \|r\|$, hoặc $\|r_Q\| \in S(\boldsymbol{W}) \times \|r\|$, hoặc $\|r_Q\| \in S(nhỏ) \times \|r\|$, hoặc $\|r_Q\| \in S(\boldsymbol{0}) \times \|r\|$ hay nói cách khác: $\|r_Q\|/\|r\|$ phải thuộc về 1 trong các khoảng: $S(\boldsymbol{1})$, S(*lớn*), $\boldsymbol{S}(\boldsymbol{W})$, S(*nhỏ*), $S(\mathbf{0})$.

b. Đưa lượng từ ngôn ngữ vào phụ thuộc hàm mờ

Để đưa lượng từ ngôn ngữ vào phụ thuộc hàm mờ, chúng ta đề xuất một số khái niệm liên quan đến các bộ trong quan hệ.

Định nghĩa 4.7. Cho U là một lược đồ quan hệ, r là một quan hệ trên U, xét hai tập X, $Y \subseteq U$. Ta nói rằng bộ t *thỏa mãn tập X* và *tập Y* trong quan hệ r với mức k, được xác định $t_k(XY) = 1$ nếu một trong các điều kiện sau xảy ra:

(1) : Tồn tại bộ $t' \neq t$: $t[X] =_k t'[X]$ và $t[Y] =_k t'[Y]$ hoặc là

(2) : Với mọi bộ $t' \neq t$: $t[X] \neq_k t'[X]$.

Khi đó ta cũng nói bộ t' *thỏa mãn tập X* và *tập Y* trong quan hệ r với mức k, được xác định $t'_k(XY) = 1$.

Định nghĩa 4.8. Cho U là một lược đồ quan hệ, r là một quan hệ xác định trên U, xét X, $Y \subseteq U$. Ta nói rằng bộ t *thỏa mãn tập X* nhưng *không thỏa mãn tập Y* trong quan hệ r với mức k, được xác định $t_k(XY)=0$ nếu tồn tại bộ $t' \neq t$: $t[X] =_k t'[X]$ và $t[Y] \neq_k t'[Y]$.

Khi đó ta cũng nói bộ t' *thỏa mãn tập X* và *không thỏa mãn tập Y* trong quan hệ r với mức k, được xác định $t'_k(XY) = 0$.

Ký hiệu $r_{thoa} = \{t \in r : t_k(XY) = 1\}$ và $r_{khong} = \{t \in r : t_k(XY) = 0\}$. Trong phụ thuộc hàm mờ chỉ xét các lượng từ tuyệt đối và lượng từ tỉ lệ "*Hầu hết*".

Để đảm bảo một số kết quả như trong CSDL kinh điển, chúng tôi chỉ xét lượng từ Q đơn điệu tăng trong trường hợp lượng từ tuyệt đối.

Định nghĩa 4.9. Cho U là một lược đồ quan hệ, r là một quan hệ xác định trên U, xét X, $Y \subseteq U$. Ta nói rằng quan hệ r thỏa mãn phụ thuộc hàm mờ X xác định Y mức k và lượng từ ngôn ngữ Q, ký hiệu là $X \sim>_{kQ} Y$ khi và chỉ khi $f_Q^A(\|r_{thoa}\|) = 1$, với Q là lượng từ tuyệt đối hoặc $\|r_{thoa}\|/(\|r_{thoa}\| + \|r_{khong}\|) \in S(\boldsymbol{1})$, với Q là lượng từ "*Hầu hết*".

Ví dụ 4.13. Ta xét lược đồ quan hệ U = {*STT, TEN, HESO, THAMNIEN, LUONG*} với ý nghĩa: Số thứ tự (*STT*), Tên cán bộ (*TEN*), Hệ số lương (*HESO*) là 3 thuộc tính kinh điển, Thâm niên (*THAMNIEN*), Lương (*LUONG*) là 2 thuộc tính mờ. Trong đó $D_{THAMNIEN} = [0, 40]$ và $D_{LUONG} = [0, 500]$. $LD_{THAMNIEN}$ và LD_{LUONG} có cùng tập các xâu giống nhau với tập các phần tử sinh là {***0***, *thấp*, ***W***, *cao*, ***1***} và tập các gia tử là {*ít*, *khả năng*, *hơn*, *rất*}.

(**a**). ***Đối với thuộc tính*** *THAMNIEN*: *fm*(*cao*) = 0.35, *fm*(*thấp*) = 0.65, μ(*khả năng*) = 0.25, μ(*ít*) = 0.20, μ(*hơn*) = 0.15 và μ(*rất*) = 0.40.

Ta phân hoạch đoạn [0, 40] thành 5 khoảng tương tự mức 1 là: *fm*(*rất cao*) × 40 = 0.35 × 0.35 × 40 = 4.9. Vậy *S*(***1***) × 40 = (35.1, 40].

(*fm*(*khả năng cao*) + *fm*(*hơn cao*)) × 40 = (0.25 × 0.35 + 0.15 × 0.35) × 40 = 5.6 và *S*(*cao*) × 40 = (29.5, 35.1];

(*fm*(*ít thấp*) + *fm*(*ít cao*)) × 40 = (0.25 × 0.65 + 0.25 × 0.35) × 40 = 10 và *S*(***W***) × 40 = (19.5, 29.5];

(*fm*(*khả năng thấp*) + *fm*(*hơn thấp*)) × 40 = (0.25 × 0.65 + 0.15 × 0.65) × 40 = 10.4 và *S*(*thấp*) × 40 = (9.1, 19.5]; *S*(***0***) × 40 = [0, 9.1].

(**b**). ***Đối với thuộc tính*** *LUONG*: *fm*(*cao*) = 0.6, *fm*(*thấp*) = 0.4, μ(*khả năng*) = 0.15, μ(*ít*) = 0.25, μ(*hơn*) = 0.25 và μ(*rất*) = 0.35.

Ta phân hoạch đoạn [0, 500] thành 5 khoảng tương tự mức 1 là: *fm*(*rất cao*) × 500 = 0.35 × 0.6 × 500 = 105. Vậy *S*(***1***) × 500 = (395, 500].

(*fm*(*khả năng cao*) + *fm*(*hơn cao*)) × 500 = (0.25 × 0.6 + 0.15 × 0.6) × 500 = 120 và *S*(*cao*)) × 500 = (275, 395];

(*fm*(*ít thấp*) + *fm*(*ít cao*)) × 500 = (0.25 × 0.6 + 0.25 × 0.4) × 500 = 125 và *S*(***W***) × 500 = (150, 275];

(*fm*(*khả năng thấp*) + *fm*(*hơn thấp*)) × 500 = (0.25 × 0.4 + 0.15 × 0.4) × 500 = 80 và *S*(*thấp*) × 500 = (70, 150]; *S*(***0***) × 500 = [0, 70].

Quan hệ *Luong* trong ví dụ này được cho ở bảng 4.6

STT	TEN	HESO	THAMNIEN	LUONG
1	Thanh	3.5	20	300
2	Loan	3.7	25	*cao*
3	Hàng	4.5	36	450
4	Hà	4.3	37	470
5	Thủy	2.5	*thấp*	*thấp*
6	Nhật	1.9	*thấp*	185
7	Cường	3.0	27	350
8	Thương	3.1	28	*cao*
9	Miên	2.8	22	310

Bảng 4.6. *Quan hệ Luong*

Chúng ta có thể thấy rằng phụ thuộc hàm mờ *THAMNIEN* $\sim>_{1\ Hầu\ hết}$ *LUONG* không đúng trong quan hệ *Luong*.

Thật vậy, vì Q = "*Hầu hết*" nên ta tính $S(\mathbf{1})$. Chọn $fm(lớn)$ = 0.35, $fm(nhỏ)$ = 0.65, $\mu(khả\ năng)$ = 0.25, $\mu(ít)$ = 0.2, $\mu(hơn)$ = 0.15 và $\mu(rất)$ = 0.4. Ta có $fm(rất\ lớn) = 0.35 \times 0.35 = 0.1225$, do đó $S(\mathbf{1}) = (0.8775, 1]$.

Mặc khác ta có $Luong_{thoa} = \{t_1, t_2, t_3, t_4, t_7, t_8, t_9\}$ và $Luong_{khong} = \{t_5, t_6\}$. Do đó theo định nghĩa $\|Luong_{thoa}\|/(\|Luong_{thoa}\|+\|Luong_{khong}\|)= 7/(7 + 2) = 0.77 \notin (0.8775, 1]$.

Mệnh đề 4.3. Quan hệ r thỏa mãn phụ thuộc hàm mờ với lượng từ $X \sim>_{k\ Với\ mọi} Y$ khi và chỉ khi $X \sim>_k Y$.

Chứng minh: Quan hệ r thỏa mãn $X \sim>_{k\ Với\ mọi} Y$ khi và chỉ khi $\|r_{Với\ mọi}\| = 1$.
$\Leftrightarrow \|r_{khong}\| = 0$. Do đó với $\forall t_1, t_2 \in r$ nếu $t_1[X] =_k t_2[X] \Rightarrow t_1[Y] =_k t_2[Y]$. Vậy r thỏa $X \sim>_k Y$. Như vậy, nếu $X \sim>_{k\ Với\ mọi} Y$ thì ta có thể viết $X \sim>_k Y$.
Gọi $\mathcal{F}_Q^{\,k}$ là họ tất cả các phụ thuộc hàm mờ mức k với lượng từ ngôn ngữ Q trên lược đồ quan hệ U. Ta ký hiệu $\mathcal{F}_{Q^*}^{\,k}$ là tập tất cả các phụ thuộc hàm mờ mức k với lượng từ ngôn ngữ Q mà được *suy diễn* từ $\mathcal{F}_Q^{\,k}$. Ta có thể dễ dàng kiểm chứng họ các phụ thuộc hàm mờ mức k với lượng từ ngôn ngữ Q $\mathcal{F}_{Q^*}^{\,k}$ thoả các tiên đề sau:

Định lý 4.5. Trong CSDL mờ với tập vũ trụ các thuộc tính U, họ $\mathcal{F}_{Q^*}^{\,k}$ thỏa mãn các tiên đề sau:

(1) *Phản xạ* : $X \sim>_{kQ} X \in \mathcal{F}_{Q^*}^{\,k}$

(2) *Gia tăng* : $X \sim>_{kQ} Y \in \mathcal{F}_{Q^*}^{\,k} \Rightarrow XZ \sim>_{kQ} YZ \in \mathcal{F}_{Q^*}^{\,k}$

(3) *Bắc cầu 1* : $X \sim>_{kQ} Y \in \mathcal{F}_{Q^*}^{\,k}$, $Y \sim>_k Z \in \mathcal{F}_k^+ \Rightarrow X \sim>_{kQ} Z \in \mathcal{F}_{Q^*}^{\,k}$

(4) *Bắc cầu 2* : $X \sim>_k Y \in \mathcal{F}_k^+$, $Y \sim>_{kQ} Z \in \mathcal{F}_{Q^*}^{\,k} \Rightarrow X \sim>_{kQ} Z \in \mathcal{F}_{Q^*}^{\,k}$

4.3. Phụ thuộc đơn điệu

Trong cuộc sống hàng ngày, chúng ta thường gặp các điều kiện dạng như *nếu một tập thể T_1 lao động chăm chỉ (số ngày làm việc trong một tháng là nhiều) hơn tập thể T_2 thì Năng suất của tập thể T_1 cao hơn tập thể T_2*. Như vậy, giữa các thuộc tính *Số ngày làm việc trong tháng*, *Tính kỷ luật lao động* và *Năng suất* có mối quan hệ khá chặt chẽ. Hoặc trong một trường hợp khác *nếu một tập thể T_1 lao động không chăm chỉ (số ngày làm việc trong một tháng là ít) hơn tập thể T_2 thì Năng suất của tập thể T_1 thấp hơn tập thể T_2.*

Trong cả hai trường hợp trên, nó tồn tại mối quan hệ không chính xác không như mối quan hệ trong phụ thuộc hàm mờ, và những phụ thuộc như vậy gọi là phụ thuộc đơn điệu trong CSDL mờ.

Ví dụ 4.14. (1) Nếu công nhân An có *tay nghề bậc cao hơn* công nhân Thanh và *số ngày làm việc* của hai công nhân này như nhau thì *Lương* của An *cao hơn* lương của Thanh.

(2) Nếu hai thửa ruộng R_A, R_B *cùng một đơn vị diện tích* nhưng *năng suất* của R_A *cao hơn* R_B thì *sản lượng thu hoạch* của thửa ruộng R_A nhiều hơn thửa ruộng R_B.

Phụ thuộc đơn điệu được chia làm hai trường hợp là phụ thuộc đơn điệu tăng và phụ thuộc đơn điệu giảm.Vì phụ thuộc dữ liệu kinh điển được xem là một trường hợp riêng của phụ thuộc dữ liệu mờ, do đó chúng tôi xem xét phụ thuộc đơn điệu trong cả hai trường hợp kinh điển và mờ.

4.3.1. Phụ thuộc đơn điệu trong CSDL kinh điển

4.3.1.1 Phụ thuộc đơn điệu tăng: Với $X \subseteq U$, r là quan hệ xác định trên U, t_1, t_2 là hai bộ thuộc r, ta viết $t_1[X] \leq t_2[X]$, nếu với mọi $A \in X$, ta có $t_1[A] \leq t_2[A]$.

Định nghĩa 4.10. Cho U là một lược đồ quan hệ, r là một quan hệ xác định trên U, xét X, $Y \subseteq U$. Ta nói rằng quan hệ r thỏa mãn phụ thuộc đơn điệu tăng X xác định Y, ký hiệu là $X \xrightarrow{+} Y$, trong quan hệ r, nếu ta có với $\forall t_1, t_2 \in r$, $t_1[X] \leq t_2[X] \Rightarrow t_1[Y] \leq t_2[Y]$.

Ví dụ 4.15. Ta xét lược đồ quan hệ U = {*MAGV, TENGV, SOTIETGIANG, VUOTGIO*} với ý nghĩa: Mã số giáo viên (*MAGV*), Tên giáo viên (*TENGV*), Số tiết giảng trong năm học (*SOTIET*), Tiền vượt giờ (*VUOTGIO*). Quan hệ *Giangday* xác định trên U cho ở bảng 4.7.

MAGV	TENCN	SOTIETGIANG	VUOTGIO
G1	Anh	350	10050000
G2	Hiếu	450	13500000
G3	Nhân	600	18000000
G4	Giang	300	9000000
G5	Hải	370	11200000
G6	Hà	360	10800000
G7	Thanh	500	15000000
G8	Thiện	650	19000000
G9	Nhân	700	20000000

Bảng 4.7. *Quan hệ Giangday*

Trong quan hệ *Giangday* ta thấy rằng nếu *SOTIETGIANG* của giáo viên càng lớn thì *VUOTGIO* càng lớn, hay phụ thuộc đơn điệu tăng $SOTIETGIANG \xrightarrow{+} VUOTGIO$ đúng trong quan hệ *Giangday*. Thật vậy, với $\forall t_1, t_2 \in Giangday$ ta có $t_1[SOTIETGIANG] \leq t_2[SOTIETGIANG] \Rightarrow t_1[VUOTGIO] \leq t_2[VUOTGIO]$.

Gọi $\mathcal{F}_A$ là họ các phụ thuộc đơn điệu tăng trên lược đồ quan hệ *U*. Ta ký hiệu $\mathcal{F}_A{*}$ là tập tất cả các phụ thuộc đơn điệu tăng $X \xrightarrow{+} Y$ mà được *suy diễn* từ $\mathcal{F}_A$.

Định lý 4.6. Trong CSDL với tập vũ trụ các thuộc tính *U*, họ $\mathcal{F}_A{*}$ thỏa mãn các tiên đề sau:

(1) *Phản xạ*: $X \xrightarrow{+} X \in \mathcal{F}_A{*}$

(2) *Gia tăng*: $X \xrightarrow{+} Y \in \mathcal{F}_A{*} \Rightarrow XZ \xrightarrow{+} YZ \in \mathcal{F}_A{*}$, với $Z \subseteq U$

(3) *Bắc cầu*: $X \xrightarrow{+} Y \in \mathcal{F}_A{*}$, $Y \xrightarrow{+} Z \in \mathcal{F}_A{*} \Rightarrow X \xrightarrow{+} Z \in \mathcal{F}_A{*}$.

Chứng minh:

(1) : Phản xạ: Hiển nhiên vì $\forall t_1, t_2 \in r$, $t_1[X] \leq t_2[X] \Rightarrow t_1[X] \leq t_2[X]$. Vậy $X \xrightarrow{+} X \in \mathcal{F}_A{*}$.

(2) : Theo giả thiết ta có $X \xrightarrow{+} Y \in \mathcal{F}_{A^*}$ nên theo định nghĩa với $\forall t_1, t_2 \in r$, $t_1[X] \leq t_2[X] \Rightarrow t_1[Y] \leq t_2[Y]$ (1'). Với $Z \subseteq U$, từ $t_1[X] \leq t_2[X]$ ta có $t_1[XZ] \leq t_2[XZ]$ (2'). Tương tự, từ $t_1[Y] \leq t_2[Y]$ ta có $t_1[YZ] \leq t_2[YZ]$ (3'). Từ (1'), (2'), (3') ta có với $\forall t_1, t_2 \in r$, $t_1[XZ] \leq t_2[XZ] \Rightarrow t_1[YZ] \leq t_2[YZ]$. Vậy $XZ \xrightarrow{+} YZ \in \mathcal{F}_{A^*}$.

(3) : Theo giả thiết $X \xrightarrow{+} Y \in \mathcal{F}_{A^*}$, $Y \xrightarrow{+} Z \in \mathcal{F}_{A^*}$ nên ta có với $\forall t_1, t_2 \in r$, $t_1[X] \leq t_2[X] \Rightarrow t_1[Y] \leq t_2[Y]$ và với $\forall t_1, t_2 \in r$, $t_1[Y] \leq t_2[Y] \Rightarrow t_1[Z] \leq t_2[Z]$ hay với $\forall t_1, t_2 \in r$, $t_1[X] \leq t_2[X] \Rightarrow t_1[Z] \leq t_2[Z]$. Vậy $X \xrightarrow{+} Z \in \mathcal{F}_{A^*}$.

4.3.1.2. Phụ thuộc đơn điệu giảm: Với $X \subseteq U$ và r là quan hệ xác định trên U, t_1, t_2 là hai bộ thuộc r, ta viết $t_1[X] \geq t_2[X]$, nếu với mọi $A \in X$, ta có $t_1[A] \geq t_2[A]$.

Định nghĩa 4.11. Cho U là một lược đồ quan hệ, r là một quan hệ xác định trên U, xét $X, Y \subseteq U$. Ta nói rằng quan hệ r thỏa mãn phụ thuộc đơn điệu giảm X xác định Y, ký hiệu là $X \xrightarrow{-} Y$ trong quan hệ r, nếu ta có với $\forall t_1, t_2 \in r$, $t_1[X] \leq t_2[X] \Rightarrow t_1[Y] \geq t_2[Y]$.

Ví dụ 4.16. Ta xét lược đồ quan hệ U = { *SOBD, TENHS, SOPHUT, DIEMTHI* } với ý nghĩa: Số báo danh học sinh (*SOBD*), Tên học sinh (*TENHS*), Số giây chạy trong 100m (*SOGIAY*), Điểm thi (*DIEMTHI*). Quan hệ *Thihocky* xác định trên U cho ở bảng 4.8

SOBD	TENHS	SOGIAY	DIEMTHI
001	Thủy	16	8
002	Bình	18	6
003	Minh	17	7
004	Nhật	15	9
005	Thương	19	5
006	Huyền	18	6
007	Thuận	14	10
008	Thành	16	8
009	Hằng	20	4

Bảng 4.8. *Quan hệ Thihocky*

Trong quan hệ *Thihocky* ta thấy rằng, nếu *SOGIAY* mà học sinh chạy trong 100m càng lớn thì *DIEMTHI* của học sinh càng kém. Như vậy, trong quan hệ *Thihocky* tồn tại phụ thuộc đơn điệu giảm $SOGIAY \xrightarrow{-} DIEMTHI$ đúng trong quan hệ *Thihocky*. Thật vậy, với $\forall t_1, t_2 \in Thihocky$ ta có $t_1[SOGIAY] \leq t_2[SOGIAY] \Rightarrow t_1[DIEMTHI] \geq t_2[DIEMTHI]$.

Gọi $\mathcal{F}_D$ là họ các phụ thuộc đơn điệu giảm trên lược đồ quan hệ U. Ta ký hiệu $\mathcal{F}_{D^*}$ là tập tất cả các phụ thuộc đơn điệu giảm $X \xrightarrow{-} Y$ mà được *suy diễn* từ $\mathcal{F}_D$.

Định lý 4.7. Trong CSDL với tập vũ trụ các thuộc tính U, họ $\mathcal{F}_{D^*}$ thỏa mãn các tiên đề sau:

(1) *Phản xạ*: $X \xrightarrow{-} X \in \mathcal{F}_{D^*}$

(2) *Gia tăng*: $X \xrightarrow{-} Y \in \mathcal{F}_{D^*} \Rightarrow XZ \xrightarrow{-} YZ \in \mathcal{F}_{D^*}$, $Z \subseteq U$

(3) *Hỗn hợp bắc cầu:* $X \xrightarrow{-} Y \in \mathcal{F}_{D^*}$, $Y \xrightarrow{+} Z \in \mathcal{F}_{A^*} \Rightarrow X \xrightarrow{-} Z \in \mathcal{F}_{D^*}$.

Chứng minh: (1) Phản xạ: Hiển nhiên vì với $\forall t_1, t_2 \in r$, $t_1[X] \leq t_2[X] \Rightarrow t_1[X] \geq t_2[X]$. Vậy $X \xrightarrow{-} X \in \mathcal{F}_{D^*}$.

(2): Theo giả thiết ta có $X \xrightarrow{-} Y \in \mathcal{F}_{D^*}$ nên theo định nghĩa với $\forall t_1, t_2 \in r$, $t_1[X] \leq t_2[X] \Rightarrow t_1[Y] \geq t_2[Y]$ (1'). Với $Z \subseteq U$, từ $t_1[X] \leq t_2[X]$ ta có $t_1[XZ] \leq t_2[XZ]$ (2'). Tương tự, từ $t_1[Y] \geq t_2[Y]$ ta có $t_1[YZ] \geq t_2[YZ]$ (3'). Từ (1'), (2'), (3') ta có với $\forall t_1, t_2 \in r$, $t_1[XZ] \leq t_2[XZ] \Rightarrow t_1[YZ] \geq t_2[YZ]$. Vậy $XZ \xrightarrow{-} YZ \in \mathcal{F}_{D^*}$.

(3): Theo giả thiết ta có $X \xrightarrow{-} Y \in \mathcal{F}_{D^*}$ và $Y \xrightarrow{+} Z \in \mathcal{F}_{A^*}$ nên theo định nghĩa: $X \xrightarrow{-} Y \in \mathcal{F}_{D^*}$ ta có với $\forall t_1, t_2 \in r$, $t_1[X] \leq t_2[X] \Rightarrow t_1[Y] \geq t_2[Y]$ (1'). $Y \xrightarrow{+} Z \in \mathcal{F}_{A^*}$ ta có với $\forall t_1, t_2 \in r$, $t_1[Y] \geq t_2[Y] \Rightarrow t_1[Z] \geq t_2[Z]$ (2'). Từ (1'), (2') ta có: với $\forall t_1, t_2 \in r$, $t_1[X] \leq t_2[X] \Rightarrow t_1[Z] \geq t_2[Z]$. Vậy $X \xrightarrow{-} Z \in \mathcal{F}_{D^*}$.

Hệ quả 4.3. Từ định lý 4.7 ta có các hệ quả sau:

(4) *Hỗn hợp bắc cầu 1:* $X \xrightarrow{+} Y \in \mathcal{F}_{A^*}$, $Y \xrightarrow{-} Z \in \mathcal{F}_{D^*} \Rightarrow X \xrightarrow{-} Z \in \mathcal{F}_{D^*}$

(5) *Hỗn hợp bắc cầu 2:* $X \xrightarrow{-} Y \in \mathcal{F}_{D^*}$, $Y \xrightarrow{-} Z \in \mathcal{F}_{D^*} \Rightarrow X \xrightarrow{+} Z \in \mathcal{F}_{A^*}$.

4.3.2. Phụ thuộc đơn điệu trong CSDL mờ

4.3.2.1. Phụ thuộc đơn điệu tăng mức k: Với $X \subseteq U$ và r là quan hệ xác định trên U và t_1, t_2 là hai bộ thuộc r, ta viết $t_1[X] \leq_k t_2[X]$, nếu với mọi $A \in X$, ta có $t_1[A] \leq_k t_2[A]$.

Định nghĩa 4.12. Cho U là một lược đồ quan hệ, r là một quan hệ xác định trên U, xét $X, Y \subseteq U$. Ta nói rằng quan hệ r thỏa mãn phụ thuộc đơn điệu tăng X xác định Y với mức k, ký hiệu là $X \overset{+}{\sim>}_k Y$, trong quan hệ r, nếu ta có với $\forall t_1, t_2 \in r$, $t_1[X] \leq_k t_2[X] \Rightarrow t_1[Y] \leq_k t_2[Y]$.

Ví dụ 4.17. Ta xét lược đồ quan hệ U = { *MASO,TENCN, SONLV, THUNHAP* } với ý nghĩa: Mã số công nhân (*MASO*), Tên công nhân (*TENCN*) là 2 thuộc tính kinh điển, Số ngày làm việc trong tháng (*SONLV*), Thu nhập (*THUNHAP*) là 2 thuộc tính mờ. Trong đó D_{SONLV} = [0, 30] và $D_{THUNHAP}$ = [0, 100]. LD_{SONLV} và $LD_{THUNHAP}$ có cùng tập các xâu giống nhau với tập các phần tử sinh là {***0***, *thấp*, ***W***, *cao*, ***1***} và tập các gia tử là {*ít*, *khả năng*, *hơn*, *rất*}. Quan hệ *thunhap* xác định trên U được cho ở bảng 4.9.

MASO	TENCN	SONLV	THUNHAP (USD)
N1	Anh	27	90
N2	Ánh	*Cao*	*cao*
N3	Lan	28	94
N4	Hương	20	52
N5	Hằng	22	53
N6	Hồng	22	54
N7	Thúy	*thấp*	10
N8	Thanh	21	52
N9	Hiền	12	*thấp*

Bảng 4.9. *Quan hệ thunhap*

(**a**). ***Đối với thuộc tính*** *SONLV*: *fm*(*cao*) = 0.35, *fm*(*thấp*) = 0.65, μ(*khả năng*) = 0.25, μ(*ít*) = 0.2, μ(*hơn*) = 0.15 và μ(*rất*) = 0.4. Ta phân hoạch đoạn [0, 30] thành 5 khoảng tương tự mức 1 là: *fm*(*rất cao*) × 30 = 0.35 × 0.35 × 30 = 3.675. Vậy *S*(***1***) × 30 = (26.325, 30];

(*fm*(*khả năng cao*) + *fm*(*hơn cao*)) × 30 = (0.25 × 0.35 + 0.15 × 0.35) × 30 = 4.2 và *S*(*cao*) × 30 = (22.125, 26.325];

(*fm*(*ít thấp*) + *fm*(*ít cao*)) × 30 = (0.25 × 0.65 + 0.25 × 0.35) × 30 = 7.5 và *S*(***W***) × 30 = (14.625, 22.125];

(*fm*(*khả năng thấp*) + *fm*(*hơn thấp*)) × 30 = (0.25 × 0.65 + 0.15 × 0.65) × 30 = 7.8 và *S*(*thấp*) × 30 = (6.825, 14.625], *S*(***0***) × 30 = [0, 6.825].

(**b**). ***Đối với thuộc tính*** *THUNHAP*: *fm*(*cao*) = 0.6, *fm*(*thấp*) = 0.4, μ(*khả năng*) = 0.15, μ(*ít*) = 0.25, μ(*hơn*) = 0.25 và μ(*rất*) = 0.35. Ta phân hoạch đoạn [0, 100] thành 5 khoảng tương tự mức 1 là: *fm*(*rất cao*) × 100 = 0.35 × 0.6 × 100 = 21. Vậy *S*(***1***) × 100 = (79, 100];

(*fm*(*khả năng cao*) + *fm*(*hơn cao*)) × 100 = (0.25 × 0.6 + 0.15 × 0.6) × 100 = 24 và *S*(*cao*) × 100 = (55, 79];

(*fm*(*ít thấp*) + *fm*(*ít cao*)) × 100 = (0.25 × 0.6 + 0.25 × 0.4) × 100 = 25 và *S*(***W***) × 100 = (30, 55];

(*fm*(*khả năng thấp*) + *fm*(*hơn thấp*)) × 100 = (0.25 × 0.4 + 0.15 × 0.4) × 100 = 16 và *S*(*thấp*) × 100 = (14, 30], *S*(***0***) × 100 = [0, 14].

Chúng ta có thể thấy phụ thuộc đơn điệu tăng *SONLV* $\overset{+}{\sim>}_1$ *THUNHAP* đúng trong quan hệ *thunhap*. Thật vậy, ta có với $\forall t_1, t_2 \in$ *thunhap*, $t_1[SONLV] \leq_1 t_2[SONLV] \Rightarrow t_1[THUNHAP] \leq_1 t_2[THUNHAP]$. Do đó theo định nghĩa ta có *SONLV* $\overset{+}{\sim>}_1$ *THUNHAP*.

Ý nghĩa của phụ thuộc đơn điệu tăng *SONLV* $\overset{+}{\sim>}_1$ *THUNHAP* nói lên rằng với hai bộ bất kỳ, nếu giá trị trên thuộc tính *SONLV* của bộ thứ nhất bé hơn bộ thứ hai theo mức *1* thì giá trị trên thuộc tính *THUNHAP* của bộ thứ nhất cũng bé hơn bộ thứ hai theo mức *1*.

Ý nghĩa thực tế của phụ thuộc đơn điệu tăng *SONLV* $\overset{+}{\sim>}_1$ *THUNHAP* là nếu bất kỳ hai cán bộ nào có số ngày làm việc cách nhau một khoảng nào đó thì thu nhập cũng cách nhau một khoảng tương ứng.

Ví dụ 4.18. Cho lược đồ quan hệ *U* = { *TENCTY*, *NAM*, *DOANHTHU*, *LOINHUAN* } với ý nghĩa: Tên công ty (*TENCTY*), Năm (*NAM*) là 2 thuộc tính kinh điển, Doanh thu trong năm (*DOANTHU*), Lợi nhuận (*LOINHUAN*) là 2 thuộc tính mờ.

Trong đó, $D_{DOANHTHU}$ = [500, 3000] và $D_{LOINHUAN}$ = [50, 500]. $LD_{DOANHTHU}$ và $LD_{LOINHUAN}$ có cùng tập các xâu giống nhau với tập các phần tử sinh là {***0***, *thấp*, ***W***, *cao*, ***1***} và tập các gia tử là {*ít*, *khả năng*, *hơn*, *rất*}. Quan hệ *Loinhuancty* xác định trên *U* cho ở bảng 4.10.

TENCTY	NAM	DOANHTHU	LOINHUAN
Thăng Long	2006	*rất thấp*	135
Thuận Thành	2006	872	140
Tràng Tiền	2007	1275	170
Cầu Giấy	2007	*khá thấp*	*ít thấp*
Đống Đa	2007	1990	260
Dầu khí	2006	*khá cao*	*khá cao*
Hạ Long	2007	2575	375
Cố đô	2005	*rất cao*	*rất cao*
Hương Giang	2005	2950	490

Bảng 4.10. *Quan hệ Loinhuancty*

(**a**). ***Đối với thuộc tính*** *DOANHTHU* : Chọn *fm*(*cao*) = 0.35, *fm*(*thấp*) = 0.65, μ(*khả năng*) = 0.25, μ(*ít*) = 0.2, μ(*hơn*) = 0.15 và μ(*rất*) = 0.4. Ta phân hoạch đoạn [500, 3000] thành các khoảng tương tự mức 2 là:

fm(*rất rất cao*) × 2500 = 0.4 × 0.4 × 0.35 × 2500 = 140. Vậy *S*(***1***) × 2500 = (2860, 3000];

(*fm*(*hơn rất cao*) + *fm*(*khả năng rất cao*)) × 2500 = (0.15 × 0.4 × 0.35 + 0.25 × 0.4 × 0.35) × 2500 = 140 và *S*(*rất cao*) × 2500 = (2720, 2860];

(fm(*ít rất cao*) + fm(*rất hơn cao*)) × 2500 = (0.2 × 0.4 × 0.35 + 0.4 × 0.15 × 0.35) × 2500 =122.5;

(fm(*khả năng hơn cao*) + fm(*hơn hơn cao*)) × 2500 = (0.25 × 0.15 × 0.35 + 0.15 × 0.15 × 0.35) × 2500 = 52.5 và *S*(*hơn cao*) × 2500 = (2545, 2597.5];

(fm(*ít hơn cao*) + fm(*rất khả năng cao*)) × 2500 = (0.2 × 0.15 × 0.35 + 0.4 × 0.25 x 0.35) × 2500 = 113.75;

(fm(*hơn khả năng cao*) + fm(*khả năng khả năng cao*)) × 2500 = (0.15 × 0.25 × 0.35 + 0.25 × 0.25 x 0.35) × 2500 = 87.5 và *S*(*khả năng cao*) × 2500 = (2343.75, 2431.25];

(fm(*ít khả năng cao*) + fm(*rất ít cao*)) × 2500 = (0.2 × 0.25 × 0.35 + 0.4 × 0.2 × 0.35) × 2500 = 113.75;

(fm(*hơn ít cao*) + fm(*khả năng ít cao*)) × 2500 = (0.15 × 0.2 × 0.35 + 0.25 × 0.2 × 0.35) × 2500 = 70 và *S*(*ít cao*) × 2500 = (2160, 2230];

(fm(*ít ít cao*) + fm(*ít ít thấp*)) × 2500 = (0.2 × 0.2 × 0.35 + 0.2 × 0.2 × 0.65) × 2500 = 100 và *S*(***W***) × 2500 = (2060, 2160];

(fm(*hơn ít thấp*) + fm(*khả năng ít thấp*)) × 2500 = (0.15 × 0.2 × 0.65 + 0.25 × 0.2 × 0.65) × 2500 = 130 và *S*(*ít thấp*) × 2500 = (1930, 2060];

(fm(*ít khả năng thấp*) + fm(*rất ít thấp*)) × 2500 = (0.2 × 0.25 × 0.65 + 0.4 × 0.2 × 0.65) × 2500 = 211.25;

(fm(*hơn khả năng thấp*) + fm(*khả năng khả năng thấp*)) × 2500 = (0.15 × 0.25 × 0.65 + 0.25 × 0.25 × 0.65) × 2500 = 162.5 và *S*(*khả năng thấp*) × 2500 = (1556.25, 1718.75];

(fm(*ít hơn thấp*) + fm(*rất khả năng thấp*)) × 2500 = (0.2 × 0.15 × 0.65 + 0.4 × 0.25 × 0.65) × 2500 = 211.25;

(fm(*hơn hơn thấp*) + fm(*khả năng hơn thấp*)) × 2500 = (0.15 × 0.15 × 0.65 + 0.25 × 0.15 × 0.65) × 2500 = 97.5 và *S*(*hơn thấp*) × 2500 = (1247.5, 1345];

(fm(*ít rất thấp*) + fm(*rất hơn thấp*)) × 2500 = (0.2 × 0.4 × 0.65 + 0.4 × 0.15 × 0.65) × 2500 = 227.5;

(fm(*hơn rất thấp*) + fm(*khả năng rất thấp*)) × 2500 = (0.15 × 0.4 x 0.65 + 0.25 × 0.4 × 0.65) × 2500 = 260 và S(*rất thấp*) × 2500 = (857.5, 1117.5];

fm(*rất rất thấp*) × 2500 = 0.4 × 0.4 × 0.65 × 2500 = 260. Vậy $S(\boldsymbol{0})$ × 2500 = [500, 857.5].

(**b**). ***Đối với thuộc tính*** *LOINHUAN* : Chọn fm(*cao*) = 0.6, fm(*thấp*) = 0.4, μ(*khả năng*) = 0.15, μ(*ít*) = 0.25, μ(*hơn*) = 0.25 và μ(*rất*) = 0.35.

Tương tự, phân hoạch đoạn [50, 500] thành các khoảng tương tự mức 2 ta có các kết quả tương ứng như sau : $S(\boldsymbol{1})$ × 450 = (466.925, 500]; S(*rất cao*) × 450 = (429.125, 466.925];

S(*hơn cao*) × 450 = (354.875, 381.875];

S(*khả năng cao*) × 450 = (307.625, 323.825];

S(*ít cao*) × 450 = (246.875, 273.875]; $S(\boldsymbol{W})$ × 450 = (218.75, 246.875];

S(*ít thấp*) × 450 = (200.75, 218.75];

S(*khả năng thấp*) × 450 = (167.45, 178.25];

S(*hơn thấp*) × 450 = (128.75, 146.75]; S(*rất thấp*) × 450 = (72.05, 97.25]; $S(\boldsymbol{0})$ × 450 = [50, 72.05].

Chúng ta có thể thấy phụ thuộc đơn điệu tăng *DOANHTHU* $\overset{+}{\sim>}_2$ *LOINHUAN* đúng trong quan hệ *Loinhuancty*. Thật vậy, ta có với $\forall t_1, t_2 \in$ *Loinhuancty*, $t_1[DOANHTHU] \leq_2 t_2[DOANHTHU] \Rightarrow t_1[LOINHUAN] \leq_2 t_2[LOINHUAN]$. Do đó theo định nghĩa ta có *DOANHTHU* $\overset{+}{\sim>}_2$ *LOINHUAN* .

Gọi $\mathcal{F}_A{}^k$ là họ các phụ thuộc đơn điệu tăng mức k trên lược đồ quan hệ U. Ta ký hiệu $\mathcal{F}_{A^*}{}^k$ là tập tất cả các phụ thuộc đơn điệu tăng $X \overset{+}{\sim>}_k Y$ mức k mà được *suy diễn* từ $\mathcal{F}_A{}^k$.

Định lý 4.8. Trong CSDL mờ với tập vũ trụ các thuộc tính U, họ $\mathcal{F}_{A^*}{}^k$ thỏa mãn các tiên đề sau:

(1) *Phản xạ*: $X \overset{+}{\sim>}_k X \in \mathcal{F}_{A^*}{}^k$.

(2) *Gia tăng*: $X \overset{+}{\sim>}_k Y \in \mathcal{F}_{A^*}{}^k \Rightarrow XZ \overset{+}{\sim>}_k YZ \in \mathcal{F}_{A^*}{}^k, Z \subseteq U$.

(3) *Bắc cầu*: $X \overset{+}{\sim>}_k Y \in \mathcal{F}_{A^*}^k$, $Y \overset{+}{\sim>}_k Z \in \mathcal{F}_{A^*}^k \Rightarrow X \overset{+}{\sim>}_k Z \in \mathcal{F}_{A^*}^k$.

Chứng minh:

(1) Phản xạ: Hiển nhiên vì với $\forall\ t_1, t_2 \in r$, $t_1[X] \leq_k t_2[X] \Rightarrow t_1[X] \leq_k t_2[X]$. Vậy $X \overset{+}{\sim>}_k X \in \mathcal{F}_{A^*}^k$.

(2): Theo giả thiết ta có $X \overset{+}{\sim>}_k Y \in \mathcal{F}_{A^*}^k$ nên theo định nghĩa với $\forall t_1, t_2 \in r$, $t_1[X] \leq_k t_2[X] \Rightarrow t_1[Y] \leq_k t_2[Y]$ (1'). Với $Z \subseteq U$, từ $t_1[X] \leq_k t_2[X]$ ta có $t_1[XZ] \leq_k t_2[XZ]$ (2'). Tương tự, từ $t_1[Y] \leq_k t_2[Y]$ ta có $t_1[YZ] \leq_k t_2[YZ]$ (3'). Từ (1'), (2'), (3') ta có $\forall t_1, t_2 \in r$, $t_1[XZ] \leq_k t_2[XZ] \Rightarrow t_1[YZ] \leq_k t_2[YZ]$. Vậy $XZ \overset{+}{\sim>}_k YZ \in \mathcal{F}_{A^*}^k$.

(3): Theo giả thiết $X \overset{+}{\sim>}_k Y \in \mathcal{F}_{A^*}^k$, $Y \overset{+}{\sim>}_k Z \in \mathcal{F}_{A^*}^k$ nên ta có với $\forall\ t_1, t_2 \in r$, $t_1[X] \leq_k t_2[X] \Rightarrow t_1[Y] \leq_k t_2[Y]$ và với $\forall\ t_1, t_2 \in r$, $t_1[Y] \leq_k t_2[Y] \Rightarrow t_1[Z] \leq_k t_2[Z]$ hay, với $\forall\ t_1, t_2 \in r$, $t_1[X] \leq_k t_2[X] \Rightarrow t_1[Z] \leq_k t_2[Z]$. Vậy $X \overset{+}{\sim>}_k Z \in \mathcal{F}_{A^*}^k$.

4.3.2.2. Phụ thuộc đơn điệu giảm mức k: Với $X \subseteq U$ và r là một quan hệ xác định trên U, t_1 và t_2 là hai bộ thuộc r, ta viết $t_1[X] \geq_k t_2[X]$, nếu với mọi $A \in X$, ta có $t_1[A] \geq_k t_2[A]$.

Định nghĩa 4.13. Cho U là một lược đồ quan hệ, r là một quan hệ xác định trên U, xét $X, Y \subseteq U$. Ta nói rằng quan hệ r thỏa mãn phụ thuộc đơn điệu giảm X xác định Y với mức k, ký hiệu là $X \overset{-}{\sim>}_k Y$, trong quan hệ r, nếu ta có : với $\forall\ t_1, t_2 \in r$, $t_1[X] \leq_k t_2[X] \Rightarrow t_1[Y] \geq_k t_2[Y]$.

Ví dụ 4.19. Ta xét lược đồ quan hệ U = { *MASO,TENCN, SN_NGHILV, THUNHAP* } với ý nghĩa: Mã số công nhân (*MASO*), Tên công nhân (*TENCN*) là 2 thuộc tính kinh điển, Số ngày nghỉ làm việc trong tháng (*SN_NGHILV*), Thu nhập (*THUNHAP*) là 2 thuộc tính mờ. Trong đó, D_{SN_NGHILV} = [0, 30] và $D_{THUNHAP}$ = [0, 100]. LD_{SN_NGHILV} và $LD_{THUNHAP}$ có cùng tập các xâu giống nhau với tập các phần tử sinh là {***0***, *thấp*, ***W***, *cao*, ***1***} và tập các gia tử là {*ít*, *khả năng*, *hơn*, *rất*}. Phân hoạch các miền trị được tính giống như trong ví dụ 4.17. Quan hệ *thunhapthang* của ví dụ được cho trong bảng 4.11.

MASO	TENCN	SN_NGHILV	THUNHAP
N1	Anh	3	78
N2	Ánh	5	75
N3	Lan	2	*Cao*
N4	Hương	*thấp*	52
N5	Hằng	8	53
N6	Hồng	8	54
N7	Thúy	*cao*	10
N8	Thanh	9	52
N9	Hiền	18	*thấp*

Bảng 4.11. *Quan hệ thunhapthang*

Chúng ta có thể thấy rằng phụ thuộc đơn điệu giảm $SN\_NGHILV \overset{-}{\sim>}_1 THUNHAP$ đúng trong quan hệ *thunhapthang*. Thật vậy, ta có với $\forall\ t_1, t_2 \in thunhapthang$, $t_1[SN\_NGHILV] \leq_1 t_2[SN\_NGHILV] \Rightarrow t_1[THUNHAP] \geq_1 t_2[THUNHAP]$. Do đó theo định nghĩa ta có $SN\_NGHILV \overset{-}{\sim>}_1 THUNHAP$.

Ý nghĩa của phụ thuộc đơn điệu giảm $SN\_NGHILV \overset{-}{\sim>}_1 THUNHAP$ nói lên rằng với hai bộ bất kỳ, nếu giá trị trên thuộc tính *SN_NGHILV* của bộ thứ nhất bé hơn bộ thứ hai theo mức *1* thì giá trị trên thuộc tính *THUNHAP* của bộ thứ nhất lớn hơn bộ thứ hai theo mức *1*.

Ý nghĩa thực tế của phụ thuộc đơn điệu giảm $SN\_NGHILV \overset{-}{\sim>}_1 THUNHAP$ là nếu cán bộ thứ nhất có số ngày nghỉ làm việc bé hơn mức *1* đối với cán bộ thứ hai thì thu nhập của cán bộ thứ nhất lớn hơn mức *1* đối với cán bộ thứ hai.

Gọi $\mathcal{F}_D^k$ là họ các phụ thuộc đơn điệu giảm mức *k* trên lược đồ quan hệ *U*. Ta ký hiệu $\mathcal{F}_{D*}^k$ là tập tất cả các phụ thuộc đơn điệu giảm $X \overset{-}{\sim>}_k Y$ mức *k* mà được *suy diễn* từ $\mathcal{F}_D^k$.

Định lý 4.9. Trong CSDL mờ với tập vũ trụ các thuộc tính U, họ $\mathcal{F}_{D^*}{}^k$ thỏa mãn các tiên đề sau:

(1) *Phản xạ*: $X \overset{-}{\sim>}_k X \in \mathcal{F}_{D^*}{}^k$

(2) *Gia tăng*: $X \overset{-}{\sim>}_k Y \in \mathcal{F}_{D^*}{}^k \Rightarrow XZ \overset{-}{\sim>}_k YZ \in \mathcal{F}_{D^*}{}^k, Z \subseteq U$

(3) *Hỗn hợp bắc cầu:* $X \overset{-}{\sim>}_k Y \in \mathcal{F}_{D^*}{}^k$, $Y \overset{+}{\sim>}_k Z \in \mathcal{F}_{A^*}{}^k \Rightarrow X \overset{-}{\sim>}_k Z \in \mathcal{F}_{D^*}{}^k$.

Chứng minh: (1): Phản xạ: Hiển nhiên vì với $\forall t_1, t_2 \in r$, $t_1[X] \leq_k t_2[X] \Rightarrow t_1[X] \geq_k t_2[X]$. Vậy $X \overset{-}{\sim>}_k X \in \mathcal{F}_{D^*}{}^k$.

(2): Theo giả thiết ta có $X \overset{-}{\sim>}_k Y \in \mathcal{F}_{D^*}{}^k$ nên theo định nghĩa với $\forall t_1, t_2 \in r$, $t_1[X] \leq_k t_2[X] \Rightarrow t_1[Y] \geq_k t_2[Y]$ (1'). Với $Z \subseteq U$, từ $t_1[X] \leq_k t_2[X]$ ta có $t_1[XZ] \leq_k t_2[XZ]$ (2'). Tương tự, từ $t_1[Y] \geq_k t_2[Y]$ ta có $t_1[YZ] \geq_k t_2[YZ]$ (3'). Từ (1'), (2'), (3') ta có $\forall t_1, t_2 \in r$, $t_1[XZ] \leq_k t_2[XZ] \Rightarrow t_1[YZ] \geq_k t_2[YZ]$. Vậy $XZ \overset{-}{\sim>}_k YZ \in \mathcal{F}_{D^*}{}^k$.

(3): Theo giả thiết ta có $X \overset{-}{\sim>}_k Y \in \mathcal{F}_{D^*}{}^k$ và $Y \overset{+}{\sim>}_k Z \in \mathcal{F}_{A^*}{}^k$ nên theo định nghĩa: $X \overset{-}{\sim>}_k Y \in \mathcal{F}_{D^*}{}^k$ ta có với $\forall t_1, t_2 \in r$, $t_1[X] \leq_k t_2[X] \Rightarrow t_1[Y] \geq_k t_2[Y]$ (1'); $Y \overset{+}{\sim>}_k Z \in \mathcal{F}_{A^*}{}^k$ ta có $\forall t_1, t_2 \in r$, $t_1[Y] \geq_k t_2[Y] \Rightarrow t_1[Z] \geq_k t_2[Z]$ (2').

Từ (1'), (2') ta có: với $\forall t_1, t_2 \in r$, $t_1[X] \leq_k t_2[X] \Rightarrow t_1[Z] \geq_k t_2[Z]$. Vậy $X \overset{-}{\sim>}_k Z \in \mathcal{F}_{D^*}{}^k$.

Hệ quả 4.4. Từ định lý 4.9 ta có các hệ quả sau:

(4) *Hỗn hợp bắc cầu 1:* $X \overset{+}{\sim>}_k Y \in \mathcal{F}_{A^*}{}^k$, $Y \overset{-}{\sim>}_k Z \in \mathcal{F}_{D^*}{}^k \Rightarrow X \overset{-}{\sim>}_k Z \in \mathcal{F}_{D^*}{}^k$

(5) *Hỗn hợp bắc cầu 2:* $X \overset{-}{\sim>}_k Y \in \mathcal{F}_{D^*}{}^k$, $Y \overset{-}{\sim>}_k Z \in \mathcal{F}_{D^*}{}^k \Rightarrow X \overset{+}{\sim>}_k Z \in \mathcal{F}_{A^*}{}^k$.

4.4. Ngôn ngữ truy vấn trong cơ sở dữ liệu mờ

4.4.1. Truy vấn dữ liệu mờ

Ngôn ngữ truy vấn trong CSDL mờ được các tác giả quan tâm nghiên cứu và đã có nhiều kết quả. Truy vấn theo cách tiếp cận lý thuyết tập mờ, truy vấn theo cách tiếp cận quan hệ tương tự. Hầu hết, các tác giả đều xây dựng ngôn ngữ truy vấn với mong muốn thao tác mềm dẽo, "*chính xác*" với dữ liệu mờ bằng cách tập trung xây dựng các hàm thuộc, từ đó tùy theo ngữ nghĩa của dữ liệu để chọn các ngưỡng phù hợp khi thao tác dữ liệu.

Tuy nhiên, mỗi ngôn ngữ truy vấn chỉ phù hợp với một mô hình CSDL mờ cụ thể mà không có ngôn ngữ nào tổng quát cho các mô hình. Chẳng hạn, đối với ngôn ngữ truy vấn trên mô hình CSDL mờ theo cách tiếp cận tập mờ, việc xác định giá trị chân lý của điều kiện mờ trong truy vấn thường là khó khăn và kết quả của truy vấn phụ thuộc nhiều vào việc xây dựng các hàm thuộc.

Nếu chúng ta xem miền trị của thuộc tính mờ là một ĐSGT thì việc xây dựng ngôn ngữ truy vấn trên mô hình CSDL mờ để phù hợp với ngữ nghĩa mới là cần thiết cho việc thao tác và tìm kiếm dữ liệu, giá trị chân lý của điều kiện mờ trong câu truy vấn được xác định đơn giản và hiệu quả.

Vấn đề quan tâm ở đây là một ngôn ngữ được đề xuất dùng để truy vấn dữ liệu trong mô hình CSDL mờ đã được xây dựng trong phần 4.1. Do đó, trước tiên phải đi xây dựng các thuật toán xác định giá trị chân lý của điều kiện mờ để làm cơ sở khi truy vấn dữ liệu, tiếp đến áp dụng các thuật toán vừa xây dựng để thực hiện truy vấn.

4.4.1.1. Câu lệnh Select

Tương tự như trong CSDL quan hệ, dạng tổng quát của câu lệnh truy vấn SQL sử dụng truy vấn trong CSDL mờ được biểu diễn như sau:

SELECT [DISTINCT]<danh sách cột>| <biểu thức số học>
FROM <danh sách các bảng>|<danh sách các View>
[WHERE <biểu thức điều kiện >]
[GROUP BY <danh sách các cột>[HAVING <biểu thức điều kiện >]]
[ORDER BY <danh sách các cột>|<biểu thức>[ASC|DESC]]

[UNION|INTERSECT|MINUS <câu truy vấn>]
Trong đó mệnh đề WHERE được biểu diễn với các dạng sau:
WHERE [NOT] <biểu thức ĐK>
WHERE [NOT] <biểu thức ĐK mờ>
WHERE [NOT] <biểu thức ĐK> {AND|OR}[NOT] <biểu thức ĐK>
WHERE [NOT] <biểu thức ĐK mờ> {AND|OR} [NOT] <biểu thức ĐK>
WHERE [NOT] <biểu thức ĐK> {AND|OR} [NOT] <biểu thức ĐK mờ>
WHERE [NOT] <biểu thức ĐK mờ> {AND|OR} [NOT] <biểu thức ĐK mờ>.
HAVING <biểu thức ĐK >
HAVING <biểu thức ĐK mờ >

4.4.1.2. Thuật toán xác định giá trị chân lý của điều kiện mờ

Thuật toán 4.1. *Xác định giá trị chân lý của đơn điều kiện mờ với phép toán* θ.

Vào : *Cho* r *là một quan hệ xác định trên vũ trụ các thuộc tính* $U = \{A_1, A_2, \ldots, A_n\}$.

Điều kiện A_i θ *fvalue*, với *fvalue* là giá trị mờ và A_i là thuộc tính mờ có tính đơn điệu, $\theta \in \{=_k, \neq_k, <_k, >_k\}$.

Ra : *Với mọi* $t \in r$ *thoả mãn điều kiện* $t[A_i]$ θ *fvalue.*

Phương pháp

(1) **Begin**
(2) **for** *each* $t \in r$ **do**
(3) **if** $t[A_i] \in D_{Ai}$ **then** $t[A_i] = \Phi_k(\mathcal{F}(t[A_i]))$
// Xây dựng các P^k dựa vào độ dài các từ.
(4) $k = 1$
(5) **while** $k \le p$ **do**
(6) **begin**
(7) $P^k = \varnothing$
(8) **for** *each* $t \in r$ **do**
(9) **if** $|t[A_i]| = k$ **then** $P^k = P^k \cup \{I(t[A_i])\}$
(10) $k = k + 1$
(11) **end**
// Xác định giá trị chân lý của điều kiện mờ (A_i θ *fvalue*).

(12) **if** θ là phép toán $=_k$ **then**

(13) **begin**

(14) $k = p$

(15) **while** $k > 0$ **do**

(16) **begin**

(17) **for** *each* $\Delta^k \in P^k$ **do**

(18) **begin**

(19) **if** $(I(t[A_i]) \subseteq \Delta^k$ *and* $I(fvalue) \subseteq \Delta^k)$ **then** $(t[A_i] =_k fvalue) = 1$

(20) *exit*

(21) **end**

(22) $k = k - 1$

(23) **end**

(24) **end**

(25) **else** // trường hợp $\theta \in \{\neq_k, <_k, >_k\}$

(26) **begin**

(27) $k = 1$

(28) **while** $k \leq p$ **do**

(29) **begin**

(30) **for** *each* $\Delta^k \in P^k$ **do**

(31) **begin**

(32) **Case** θ **of**

(33) $\neq_k$: **if** $I(t[A_i]) \not\subset \Delta^k$ *or* $I(fvalue) \not\subset \Delta^k$ **then** $(t[A_i]\ \theta\ fvalue) = 1$

(34) $<_k$: **if** $(t[A_i] \neq_k fvalue)$ *and* $\upsilon(t[A_i]) < \upsilon(fvalue)$ **then** $(t[A_i]\ \theta\ fvalue) = 1$

(35) $>_k$: **if** $(t[A_i] \neq_k fvalue)$ *and* $\upsilon(t[A_i]) > \upsilon(fvalue)$ **then** $(t[A_i]\ \theta\ fvalue) = 1$

(36) *exit*

(37) **end;**

(38) **end**

(39) $k = k + 1$

(40) **end**

(41) **end**

(42) End.

Trong trường hợp đa điều kiện mờ với phép toán θ, không mất tính tổng quát, giả sử điều kiện mờ có dạng $A_i\ \theta\ fvalue_i\ \xi\ A_j\ \ \theta_1\ fvalue_j$, với $fvalue_i$, $fvalue_j$ là giá trị mờ, A_i, A_j là thuộc tính mờ có tính đơn điệu; θ, $\theta_1 \in \{=_k, \neq_k, <_k, >_k\}$ và ξ là phép toán *and* hoặc *or*. Muốn xác định giá trị chân lý trong trường hợp này, trước hết các điều kiện $A_i\ \theta\ fvalue_i$ và $A_j\ \theta_1\ fvalue_j$ được xác định dựa vào Thuật toán 4.1, tiếp theo tuỳ thuộc vào ξ là phép toán *and* hay *or* để kết hợp các giá trị chân lý vừa xác định. Do đó, Thuật toán xác định giá trị chân lý của đa điều kiện mờ được trình bày chi tiết :

Thuật toán 4.2. *Xác định giá trị chân lý của đa điều kiện mờ với phép toán* θ

Vào : *Cho* r *là một quan hệ xác định trên vũ trụ các thuộc tính* U = $\{A_1, A_2, \ldots, A_n\}$.

Điều kiện $A_i\ \theta\ fvalue_i\ \xi\ A_j\ \ \theta_1\ fvalue_j$.

Ra : Với mọi $t \in r$ thoả mãn điều kiện ($t[A_i]\ \theta\ fvalue_i\ \xi\ \ t[A_j]\ \theta_1\ fvalue_j$).

Phương pháp

(1) **Begin**

(2) **for** *each* $t \in r$ **do**

(3) **begin**

(4) **if** $t[A_i] \in D_{A_i}$ **then** $t[A_i] = \Phi_k(\mathcal{F}(t[A_i]))$

(5) **if** $t[A_j] \in D_{A_j}$ **then** $t[A_j] = \Phi_k(\mathcal{F}(t[A_j]))$

(6) **end**

// Xây dựng các ${P_{A_i}}^k$ và ${P_{A_j}}^k$ dựa vào độ dài các từ.

(7) $k = 1$

(8) **while** $k \leq p$ **do**

(9) **begin**

(10) ${P_{A_i}}^k = \varnothing$; ${P_{A_j}}^k = \varnothing$

(11) **for** *each* $t \in r$ **do**

(12) **begin**

(13) **if** $|t[A_i]| = k$ **then** $P_{A_i}{}^{k} = P_{A_i}{}^{k} \cup \{I(t[A_i])\}$

(14) **if** $|t[A_j]| = k$ **then** $P_{A_j}{}^{k} = P_{A_j}{}^{k} \cup \{I(t[A_j])\}$

(15) **end**

(16) $k = k + 1$

(17) **end**

(18) **for** *each* $t \in r$ **do**

(19) **begin**

// Trường hợp ξ là phép toán *and*

(20) **if** $((t[A_i]\ \theta\ fvalue_i) = 1)$ *and* $((t[A_j]\ \theta_1\ fvalue_j) = 1)$ **then**

$((t[A_i]\ \theta\ fvalue_i)$ *and* $(t[A_j]\ \theta_1\ fvalue_j)) = 1$

// Trường hợp ξ là phép toán *or*

(21) **if** $((t[A_i]\ \theta\ fvalue_i) = 1)$ *or* $((t[A_j]\ \theta_1\ fvalue_j) = 1)$ **then**

$((t[A_i]\ \theta\ fvalue_i)$ *or* $(t[A_j]\ \theta_1\ fvalue_j)) = 1$

(22) **end**

(23) **End.**

Thuật toán 4.3. *Thực hiện truy vấn SQL mờ trong trường hợp đơn điều kiện.*

Vào **:** Quan hệ r xác định trên tập vũ trụ các thuộc tính $U = \{A_1, A_2, \ldots, A_n\}$.

Câu truy vấn dạng *select …from…. r where* $A_i\ \theta_{(Ai)}\ fvalue$.

Ra **:** Quan hệ r_{result} thỏa mãn với mọi $t \in r_{result}$ ta có $t[A_i]\ \theta_{(Ai)}$ *fvalue*.

Phương pháp

// Khởi tạo các giá trị.

(1) **Begin**

(2) Cho $G_{Ai} = \{\ \boldsymbol{0}, c^{-}_{Ai}, \boldsymbol{W}, c^{+}_{Ai}, \boldsymbol{1}\}$, $H_{Ai} = H_{Ai}{}^{+} \cup H_{Ai}{}^{-}$. Trong đó $H_{Ai}{}^{+} = \{h_1, h_2\}$, $H_{Ai}{}^{-} = \{h_3, h_4\}$, với $h_1 < h_2$ và $h_3 > h_4$. Chọn độ đo tính mờ cho các phần tử sinh và gia tử.

(3) $D_{A_i} = [min_{Ai}, max_{Ai}]$, với min_{Ai}, max_{Ai}: giá trị nhỏ nhất và lớn nhất miền trị A_i.

(4) $LD_{A_i} = H_{Ai}(c^{+}) \cup H_{Ai}(c^{-})$.

(5) $r_{result} = \varnothing$.

// Phân hoạch D_{A_i} thành các khoảng *tương tự mức k*.

(6) $k = 1$

(7) **While** $k \leq p$ **do** // *mức tương tự lớn nhất* $k = p$

(8) **begin**

(9) Xây dựng các khoảng *tương tự mức k* : $S_{Ai}(x_1)$, $S_{Ai}(x_2)$, …, $S_{Ai}(x_m)$

(10) $k = k + 1$

(11) **end**

// Xác định *lân cận mức k* của *fvalue*

(12) **if** $fvalue \in S(x_i)$ **then** $\Omega_{\mathrm{k}}(fvalue) = S(x_i)$

// Duyệt các bộ t trong r để tìm các bộ thỏa mãn điều kiện $t[A_i]\ \theta_{(Ai)}\ fvalue$

(13) **for** *each* $t \in r$ **do**

(14) **if** $t[A_i] \in LD_{A_i}$ **then** Xác định *lân cận mức k* của $t[A_i]$ là $\Omega_{\mathrm{k}}(t[A_i]) = S(x_j)$

(15) **for** *each* $t \in r$ **do**

(16) **if** θ là phép $=_{\mathrm{k}}$ **then**

(17) **begin**

(18) **if** $t[A_i] \in \Omega_{\mathrm{k}}(fvalue)$ **then** $r_{result} = r_{result} \cup t$

(19) **elseif** $\Omega_{\mathrm{k}}(t[A_i]) = \Omega_{\mathrm{k}}(fvalue)$ **then** $r_{result} = r_{result} \cup t$

(20) **end**

(21) **else** // trường hợp $\theta \in \{ \neq_{\mathrm{k}}, <_{\mathrm{k}}, >_{\mathrm{k}} \}$

(22) **begin**

(23) **Case** θ **of**

(24) $\neq_{\mathrm{k}}$: **if** $\Omega_{\mathrm{k}}(t[A_i]) \neq \Omega_{\mathrm{k}}(fvalue)$ **then** $r_{result} = r_{result} \cup t$

(25) $<_{\mathrm{k}}$: **if** $\Omega_{\mathrm{k}}(t[A_i]) < \Omega_{\mathrm{k}}(fvalue)$ **then** $r_{result} = r_{result} \cup t$

(26) $>_{\mathrm{k}}$: **if** $\Omega_{\mathrm{k}}(t[A_i]) > \Omega_{\mathrm{k}}(fvalue)$ **then** $r_{result} = r_{result} \cup t$

(27) **end**

(28) return r_{result}

(29) **End.**

Trong trường hợp đa điều kiện mờ, muốn xác định giá trị chân lý trước hết các điều kiện $A_i\ \theta_{(Ai)}\ fvalue_i$ và $A_j\ \theta_{1(Aj)}\ fvalue_j$ được xác định dựa vào Thuật toán

4.3, tiếp theo tuỳ thuộc vào ξ là phép toán *and* hay *or* để kết hợp các giá trị chân lý vừa xác định. Do đó, Thuật toán được xây dựng chi tiết như sau:

Thuật toán 4.4. *Thực hiện truy vấn SQL mờ trong trường hợp đa điều kiện.*

Vào : Quan hệ *r* xác định trên tập vũ trụ các thuộc tính $U = \{A_1, A_2, \ldots, A_n\}$. Câu truy vấn dạng *select … from r where* $A_i\ \theta_{(Ai)}\ fvalue_i\ \xi\ A_j\ \theta_{1\,(Aj)}\ fvalue_j$, trong đó ξ là phép toán *and* hoặc *or*.

Ra : Quan hệ r_{result} thỏa mãn với mọi $t \in r_{result}$ ta có $t[A_i]\ \theta_{(Ai)}\ fvalue_i\ \xi\ t[A_j]\ \theta_{1\,(Aj)}\ fvalue_j$.

Phương pháp

// Khởi tạo các giá trị.

(1) **Begin**

(2) Cho $G_{Ai} = \{ \mathbf{0}, c^-_{Ai}, \mathbf{W}, c^+_{Ai}, \mathbf{1}\}$, $H_{Ai} = H_{Ai}^+ \cup H_{Ai}^-$. Trong đó $H_{Ai}^+ = \{h_1, h_2\}$, $H_{Ai}^- = \{h_3, h_4\}$, với $h_1 < h_2$ và $h_3 > h_4$. Chọn độ đo tính mờ cho các phần tử sinh và gia tử.

(3) Cho $G_{Aj} = \{ \mathbf{0}, c^-_{Aj}, \mathbf{W}, c^+_{Aj}, \mathbf{1}\}$, $H_{Aj} = H_{Aj}^+ \cup H_{Aj}^-$. Trong đó $H_{Aj}^+ = \{h_1, h_2\}$, $H_{Aj}^- = \{h_3, h_4\}$, với $h_1 < h_2$ và $h_3 > h_4$. Chọn độ đo tính mờ cho các phần tử sinh và gia tử.

(4) $D_{A_i} = [min_{Ai}, max_{Ai}]$, với min_{Ai}, max_{Ai}: giá trị nhỏ nhất và lớn nhất miền trị A_i.

(5) $D_{A_j} = [min_{Aj}, max_{Aj}]$, với min_{Aj}, max_{Aj}: giá trị nhỏ nhất và lớn nhất miền trị A_j.

(6) $LD_{A_i} = H_{Ai}(c^+_{Ai}) \cup H_{Ai}(c^-_{Ai})$.

(7) $LD_{A_j} = H_{Aj}(c^+_{Aj}) \cup H_{Aj}(c^-_{Aj})$.

(8) $r_{result} = \varnothing$.

// Phân hoạch D_{A_i} và D_{A_j} thành các khoảng *tương tự mức k*.

(9) $k = 1$

(10) **While** $k \le p$ **do**

(11) **begin**

(12) Xây dựng các khoảng *tương tự mức k*: $S_{Ai}(x_1)$, $S_{Ai}(x_2)$, …, $S_{Ai}(x_m)$

(13) Xây dựng các khoảng *tương tự mức k*: $S_{Ai}(y_1)$, $S_{Ai}(y_2)$, …,

$S_{Ai}(y_m)$

(14) $k = k + 1$

(15) **end**

// Xác định *lân cận mức k* của $fvalue_i$ và $fvalue_j$

(16) **if** $fvalue_i \in S_{Ai}(x_i)$ **then** $\Omega_k(fvalue_i) = S_{Ai}(x_i)$

(17) **if** $fvalue_j \in S_{Aj}(y_i)$ **then** $\Omega_k(fvalue_j) = S_{Aj}(y_i)$

// Duyệt các bộ t trong r để tìm các bộ thỏa mãn điều kiện $t[A_i]\ \theta_{(Ai)}\ fvalue_i\ \xi\ t[A_j] = \theta_{1\,(Aj)}\ fvalue_j$

// Trường hợp ξ là phép toán *and*

(18) **for** *each* $t \in r$ **do**

(19) **if** $\{(t[A_i]\ \theta_{(Ai)}\ fvalue_i) = 1\}$ *and* $\{(t[A_j]\ \theta_{1(Aj)}\ fvalue_j) = 1\}$ **then**

$r_{result} = r_{result} \cup t$

// Trường hợp ξ là phép toán *or*

(20) **if** $\{(t[A_i]\ \theta_{(Ai)}\ fvalue_i) = 1\}$ *or* $\{(t[A_j]\ \theta_{1(Aj)}\ fvalue_j) = 1\}$ **then**

$r_{result} = r_{result} \cup t$

(21) Return $r_{result.}$

(22) **End.**

4.4.1.3. Phương pháp truy vấn dữ liệu mờ

Các ngôn ngữ truy vấn trong mô hình CSDL quan hệ được nhiều tác giả quan tâm nghiên cứu và mở rộng trong mô hình CSDL mờ như: đại số quan hệ mờ, ngôn ngữ SQL. Vì thế, tương tự trong CSDL quan hệ, cấu trúc của câu lệnh SQL trong CSDL mờ được xem xét như trong Mục 4.4.1.1. Như vậy, câu lệnh SQL trong CSDL mờ có thể được tổng quát hoá sau:

(1) : Thực hiện phép tích Descartes giữa các quan hệ tham gia truy vấn.

(2) : Xác định giá trị chân lý của các điều kiện mờ (Sử dụng thuật toán 4.1, 4.2, 4.3, 4.4) và liên kết các giá trị chân lý vừa xác định.

(3) : Chọn các bộ dữ liệu thoả mãn bước (2).

Do đó, vấn đề quan trọng của câu lệnh SQL trong CSDL mờ chính là xác định giá trị chân lý của điều kiện mờ và liên kết các giá trị chân lý đó.

Ví dụ 4.20. Cho quan hệ lược đồ quan hệ U = {*SOCM*, *HOTEN*, *SUCKHOE*, *TUOI*, *LUONG*} và quan hệ *Suckhoe_tuoi* được xác định như sau:

SOCM	HOTEN	SUCKHOE	TUOI	LUONG
11111	Phạm Trọng Cầu	*rất rất tốt*	31	2.800.000
22222	Nguyễn Văn Tý	*rất tốt*	85	*cao*
33333	Trần Tiến	*xấu*	32	2.000.000
44444	Vũ Hoàng	*khá xấu*	45	500.000
55555	An Thuyên	*rất xấu*	41	*rất cao*
66666	Thuận Yến	*có thể xấu*	61	*thấp*
77777	Văn Cao	*khá tốt*	59	*ít cao*
88888	Thanh Tùng	*có thể tốt*	75	2.500.000
99999	Nguyễn Cường	*ít tốt*	25	*khá thấp*

Bảng 4.12. *Quan hệ Suckhoe_tuoi*

Tìm những cán bộ có *TUOI* $=_1$ *khả năng trẻ* và *LUONG* $=_1$ cao, sử dụng thuật toán 4.3 ta có:

(a). Xét $\underline{X}_{Tuoi} = (\boldsymbol{X}_{Tuoi}, G_{Tuoi}, H_{Tuoi}, \leq)$ là ĐSGT, với G_{Tuoi} = {*trẻ*, *già*}, H^{+}_{Tuoi} = {*rất*, *hơn*}, H^{-}_{Tuoi} = {*khả năng*, *ít*}, với *rất* > *hơn* và *ít* > *khả năng*. $\boldsymbol{W}_{tuoi}$ = 0.4, *fm*(*trẻ*) = 0.4, *fm*(*già*) = 0.6, *fm*(*rất*) = 0.3, *fm*(*hơn*) = 0.15, *fm*(*khả năng*) = 0.25, *fm*(*ít*) = 0.3.

Ta có *fm*(*rất trẻ*) = 0.12, *fm*(*hơn trẻ*) = 0.06, *fm*(*ít trẻ*) = 0.12, *fm*(*khả năng trẻ*) = 0.1. Vì *rất trẻ* < *hơn trẻ* < *trẻ* < *khả năng trẻ* < *ít trẻ* nên *I*(*rất trẻ*) = [0,0.12], *I*(*hơn trẻ*) = [0.12,0.18], *I*(*khả năng trẻ*) = [0.18,0.3], *I*(*ít trẻ*) = [0.3,0.4].

Nếu chọn ψ_1= 100 $\in \boldsymbol{X}_{tuoi}$ khi đó $\mathcal{F}$(31) = 0.31. Vì 0.31$\in$*I*(*ít trẻ*) nên Φ_2(0.31) = *ít trẻ*.

Xác định điều kiện (*TUOI* $=_1$ *có thể trẻ*) = 1. Vì υ(*ít trẻ*) $\in$ *I*(*trẻ*) và υ(*có thể trẻ*) $\in$ *I*(*trẻ*) nên ta có (*ít trẻ* $=_1$ *có thể trẻ*) = 1, (*có thể trẻ* $=_2$ *có thể trẻ*) = 1, nên suy ra (*có thể trẻ* $=_1$ *có thể trẻ*) = 1.

(b). Xét $\underline{X}_{\text{Luong}} = (\boldsymbol{X}_{luong}, G_{luong}, H_{luong}, \leq)$ là ĐSGT, với G_{luong} = {*cao*, *thấp*}, H^{+}_{luong} = {*rất*, *hơn*}, H^{-}_{luong} = {*khả năng*, *ít*}, *rất* > *hơn* và *ít* > *khả năng*.

$\boldsymbol{W}_{luong}$ = 0.6, *fm*(*thấp*) = 0.6, *fm*(*cao*) = 0.4, *fm*(*rất*) = 0.25, *fm*(*khá*) = 0.25, *fm*(*khả năng*) = 0.25, *fm*(*ít*) = 0.25.

Ta có *fm*(*rất thấp*) = 0.15, *fm*(*hơn thấp*) = 0.15, *fm*(*ít thấp*) = 0.15, *fm*(*khả năng thấp*) = 0.15. Vì *rất thấp* < *hơn thấp* < *thấp* < *khả năng thấp* < *ít thấp* nên *I*(*rất thấp*) = [0,0.15], *I*(*hơn thấp*) = [0.15,0.3], *I*(*khả năng thấp*) = [0.3,0.45], *I*(*ít thấp*) = [0.45,0.6].

Ta có *fm*(*rất cao*) = 0.1, *fm*(*khá cao*) = 0.1, *fm*(*ít cao*) = 0.1, *fm*(*khả năng cao*) = 0.1. Vì *ít cao* < *khả năng cao* < *cao* < *hơn cao* < *rất cao* nên *I*(*ít cao*) = [0.6,0.7], *I*(*khả năng cao*) = [0.7,0.8], *I*(*hơn cao*) = [0.8,0.9], *I*(*rất cao*) = [0.9,1].

Nếu chọn ψ_2 = *rất rất cao* $\in \boldsymbol{X}_{luong}$ và ψ_1=3.000.000, ta có υ(*rất rất cao*) = 0.985, khi đó $\mathcal{F}$(2.800.000) = 0.92 và $\mathcal{F}$(2.000.000) = 0.65 nên Φ_2(0.92) = *rất cao* và Φ_2(0.65) = *ít cao*.

Xác định điều kiện (*LUONG* $=_1$ *cao*) = 1. Vì υ(*rất cao*) $\in$ *I*(*cao*) và υ(*ít cao*) $\in$ *I*(*cao*) nên ta có (*rất cao* $=_1$ *cao*) = 1 và (*ít cao* $=_1$ *cao*) = 1.

Như vậy, câu truy vấn SQL mờ: *select * from Suckhoe_tuoi where* (*TUOI* $=_1$ *có thể trẻ*) *and* (*LUONG* $=_1$ *cao*) sẽ cho kết quả sau:

SOCM	HOTEN	SUCKHOE	TUOI	LUONG
11111	Phạm Trọng Cầu	*rất rất tốt*	31	2.800.000
33333	Trần Tiến	*xấu*	32	2.000.000

Bảng 4.13. *Kết quả truy vấn trên quan hệ Suckhoe_tuoi*

Ví dụ 4.21. Cho lược đồ quan hệ *U* = {*STT*, *TENNV*, *NGHENGHIEP*, *TUOI*, *LUONG*} và quan hệ *Nhanvien* được xác định như bảng 4.14

STT	TENNV	NGHENGHIEP	TUOI	LUONG
1	An	Giáo viên	45	*rất cao*
2	Bình	Kỹ sư	33	1100
3	Hà	Bác sĩ	*rất khả năng trẻ*	500
4	Hương	Y sĩ	36	700
5	Nhân	Giáo viên	46	1500

6	Thuỷ	Kiến trúc sư	26	*khả năng cao*
7	Thành	Y tá	*trẻ*	750
8	Xuân	Thư ký	21	*thấp*
9	Yến	Kỹ thuật viên	*ít già*	1125

Bảng 4.14. Quan hệ *Nhanvien*

Giả sử từ quan hệ *Nhanvien* chúng ta muốn thực hiện câu truy vấn SQL mờ : *Cho biết những cán bộ có tuổi khả năng trẻ*, sử dụng thuật toán 4.3 ta có:

Bước (1)-(5) : Cho G_{TUOI} = {**0**, *trẻ*, **W**, *già*, **1**}, D_{TUOI} = [0, 100], H_{TUOI} =$H_{TUOI}^+ \cup H_{TUOI}^-$Trong đó H_{TUOI}^+= {*hơn*, *rất*}, với *hơn* < *rất* và H_{TUOI}^- = {*ít*, *khả năng*}, với *ít* > *khả năng*. Chọn *fm*(*già*) = *fm*(c^+) = 0.35, *fm*(*trẻ*) = *fm*(c^-) = 0.65, μ(*khả năng*) = 0.25, μ(*ít*) = 0.20, μ(*hơn*) = 0.15 và μ(*rất*) = 0.40. LD_{TUOI} = H_{TUOI}(*trẻ*) $\cup$ H_{TUOI}(*già*),

$r_{result} = \varnothing$.

Bước (6)-(11): Vì |*khả năng trẻ*| = 2 nên ta chỉ cần đi xây dựng các khoảng *tương tự mức 2*.

Ta phân hoạch đoạn [0, 100] thành các khoảng *tương tự mức 2*: *fm*(*rất rất trẻ*) × 100 = 0.40 × 0.40 × 0.65 × 100 = 10.4. Vậy *S*(**0**) × 100 = [0, 10.4].

(*fm*(*hơn rất trẻ*) + *fm*(*khả năng rất trẻ*)) × 100 = (0.15 × 0.40 × 0.65 + 0.25 × 0.40 × 0.65)× 100 = 10.4 và *S*(*rất trẻ*) × 100 = (10.4, 20.8].

(*fm*(*ít rất trẻ*) + *fm*(*rất hơn trẻ*)) × 100 = (0.20 × 0.40 × 0.65 + 0.40 × 0.15 × 0.65) × 100 = 9.1.

(*fm*(*hơn hơn trẻ*) + *fm*(*khả năng hơn trẻ*)) × 100 = (0.15 × 0.15 × 0.65 + 0.25 × 0.15 × 0.65) × 100 = 3.9 và *S*(*hơn trẻ*) × 100 = (29.9, 33.8].

(*fm*(*ít hơn trẻ*) + *fm*(*rất khả năng trẻ*)) × 100 = (0.20 × 0.15 × 0.65 + 0.40 × 0.25 × 0.65) × 100 = 8.45.

(*fm*(*hơn khả năng trẻ*) + *fm*(*khả năng khả năng trẻ*)) × 100 = (0.15 × 0.25 × 0.65 + 0.25 × 0.25 × 0.65) × 100 = 6.5 và *S*(*khả năng trẻ*) × 100 = (42.25, 48.75].

(*fm*(*ít khá trẻ*) + *fm*(*rất ít trẻ*)) × 100 = (0.20 × 0.25 × 0.65 + 0.40 × 0.20 × 0.65) × 100 = 8.45.

(*fm*(*hơn ít trẻ*) + *fm*(*khả năng ít trẻ*)) × 100 = (0.15 × 0.20 × 0.65 +

$0.25 \times 0.20 \times\ 0.65) \times 100 = 5.2$ và S(*ít trẻ*) $\times 100 = (57.2, 62.4]$. Tương tự, chúng ta tính được $S(\boldsymbol{W})$, S(*ít già*), S(*khả năng già*), S(*hơn già*), S(*rất già*), $S(1)$.

Như vậy, các khoảng tương tự mức 2 là: $S(\boldsymbol{0})$, S(*rất trẻ*), S(*hơn trẻ*), S(*khả năng trẻ*), S(*ít trẻ*), $S(\boldsymbol{W})$, S(*ít già*), S(*khả năng già*), S(*hơn già*), S(*rất già*), $S(\boldsymbol{1})$.

Bước (12): Xác định *lân cận mức 2* của *khả năng trẻ*. Ta có *khả năng trẻ* $\in$ S(*khả năng trẻ*) nên *lân cận mức 2* của *khả năng trẻ* là Ω_2(*khả năng trẻ*) = S(*khả năng trẻ*) = (42.25, 48.75].

Bước (13)-(27): Ta thấy trong quan hệ *Nhanvien*, $t_1[TUOI] \in \Omega_2$(*khả năng trẻ*), $t_5[TUOI] \in \Omega_2$(*khả năng trẻ*) và $\Omega_2(t_3[TUOI]) = \Omega_2$(*rất khả năng trẻ*) = Ω_2(*khả năng trẻ*).

Bước (28): Vậy $r_{result} = \{t_1, t_3, t_5\}$.

Hay câu truy vấn SQL mờ *select * from Nhanvien where TUOI* $=_{2(TUOI)}$ *khả năng trẻ* cho kết quả sau :

STT	TENNV	NGHENGHIEP	TUOI	LUONG
1	An	Giáo viên	45	*rất cao*
3	Hà	Bác sĩ	*rất khả năng trẻ*	500
5	Nhân	Giáo viên	46	1500

Bảng 4.15. *Kết quả truy vấn trên quan hệ Nhanvien sử dụng thuật toán 4.3*

Ví dụ 4.22. Sử dụng quan hệ *Nhanvien* trong ví dụ 4.21. *Hãy cho biết những cán bộ Tuổi khả năng trẻ và có Lương không rất cao*. Sử dụng thuật toán 4.4 ta có:

Bước (1)-(8): Cho $G_{TUOI} = \{\boldsymbol{0}$, *trẻ*, $\boldsymbol{W}$, *già*, $\boldsymbol{1}\}$, $D_{TUOI} = [0, 100]$, $H_{TUOI} = H_{TUOI}^+ \cup H_{TUOI}^-$ Trong đó $H_{TUOI}^+ = \{$*hơn*, *rất*$\}$, với *hơn* < *rất* và $H_{TUOI}^- = \{$*ít*, *khả năng*$\}$, với *ít* > *khả năng*. Chọn fm(*già*) = $fm(c^+) = 0.35$, fm(*trẻ*) = $fm(c^-) = 0.65$, μ(*khả năng*) = 0.25, μ(*ít*) = 0.20, μ(*hơn*) = 0.15 và μ(*rất*) = 0.40. $LD_{TUOI} = H_{TUOI}$(*trẻ*) $\cup$ H_{TUOI}(*già*).

Cho $G_{LUONG} = \{\boldsymbol{0}$, *thấp*, $\boldsymbol{W}$, *cao*, $\boldsymbol{1}\}$ và $D_{LUONG} = [400, 1600]$, $H_{LUONG} = H_{LUONG}^+ \cup H_{LUONG}^-$

Trong đó $H_{LUONG}^+ = \{hơn, rất\}$, với *hơn* < *rất* và $H_{LUONG}^- = \{ít, khả\ năng\}$, với *ít* > *khả năng*. Chọn $fm(cao) = fm(c^+) = 0.6$, $fm(thấp) = fm(c^-) = 0.4$, $\mu(khả\ năng) = 0.15$, $\mu(ít) = 0.25$, $\mu(hơn) = 0.25$ và $\mu(rất) = 0.35$. $LD_{LUONG} = H_{LUONG}(cao) \cup H_{LUONG}(thấp)$, $r_{result} = \varnothing$.

Bước (9)-(15): Vì |*khả năng trẻ*| = 2 và |*rất cao*| = 2 nên ta chỉ cần đi xây dựng các khoảng *tương tự mức 2*.

Đối với thuộc tính *TUOI*: Ta phân hoạch đoạn [0, 100] thành các khoảng *tương tự mức 2*, theo ví dụ 4.2 ta có: *S*(***0***), *S*(*rất trẻ*), *S*(*hơn trẻ*), *S*(*khả năng trẻ*), *S*(*ít trẻ*), *S*(*W*), *S*(*ít già*), *S*(*khả năng già*), *S*(*hơn già*), *S*(*rất già*), *S*(***1***).

Đối với thuộc tính *LUONG*: Ta phân hoạch đoạn [400, 1600] thành các khoảng *tương tự mức 2* : *fm*(*rất rất cao*) × 1200 = 0.35 × 0.35 × 0.6 × 1200 = 88.2. Vậy *S*(***1***)× 1200 = (1511.8, 1600].

(*fm*(*hơn rất cao*) + *fm*(*khả năng rất cao*)) × 1200 = (0.25 × 0.35 × 0.6 + 0.15 × 0.35 × 0.6) × 1200 = 100.8 và *S*(*rất cao*) × 1200 = (1411, 1511.8].

(*fm*(*ít rất cao*) + *fm*(*rất hơn cao*)) × 1200 = (0.25 × 0.35 × 0.6 + 0.35 × 0.25 x 0.6) × 1200 = 126.

(*fm*(*khả năng hơn cao*) + *fm*(*hơn hơn cao*)) × 1200 = (0.15 × 0.25 × 0.6 + 0.25 × 0.25 × 0.6) × 1200 = 72 và *S*(*hơn cao*) × 1200 = (1213, 1285].

(*fm*(*ít hơn cao*) + *fm*(*rất khả năng cao*)) × 1200 = (0.25 × 0.25 × 0.6 + 0.35 × 0.15 × 0.6) × 1200 = 82.8.

(*fm*(*hơn khả năng cao*) + *fm*(*khả năng khả năng cao*)) × 1200 = (0.25 × 0.15 × 0.6 + 0.15 × 0.15 × 0.6) × 1200 = 43.2 và *S*(*khả năng cao*) × 1200 = (1087, 1130.2].

(*fm*(*ít khả năng cao*) + *fm*(*rất ít cao*)) × 1200 = (0.25 × 0.15 × 0.6 + 0.35 × 0.25 × 0.6) × 1200 = 90.

(*fm*(*hơn ít cao*) + *fm*(*khả năng ít cao*)) × 1200 = (0.25 × 0.25 × 0.6 + 0.15 × 0.25 × 0.6) × 1200 = 72 và *S*(*ít cao*) × 1200 = (925, 997]. Tương tự, chúng ta tính được *S*(***W***), *S*(*ít thấp*), *S*(*khả năng thấp*), *S*(*hơn thấp*), *S*(*rất thấp*), *S*(***0***).

Như vậy, các khoảng *tương tự mức 2* là: *S*(***0***), *S*(*rất thấp*), *S*(*hơn thấp*), *S*(*khả năng thấp*), *S*(*ít thấp*), *S*(***W***), *S*(*ít cao*), *S*(*khả năng cao*), *S*(*hơn cao*), *S*(*rất cao*), *S*(***1***).

Bước (16)-(17): Xác định *lân cận mức 2* của *khả năng trẻ* và *rất cao*.

Ta có *lân cận mức 2* của *khả năng trẻ* là Ω_2(*khả năng trẻ*) = *S*(*khả năng trẻ*) = (42.25, 48.75] và *lân cận mức 2* của *rất cao* là Ω_2(*rất cao*) = *S*(*rất cao*) = (1411, 1511.8].

Theo ví dụ 4.2, điều kiện $TUOI =_{2(TUOI)}$ *khả năng trẻ* ta có $r_{result} = \{t_1, t_3, t_5\}$. Xét điều kiện $LUONG \neq_{2(LUONG)}$ *rất cao*, trong quan hệ *Nhanvien*, ta có *lân cận mức 2* của $t_3[LUONG] = \Omega_2(t_3[LUONG]) \neq \Omega_2$(*rất cao*). Do đó ta có $r_{result} = \{t_3\}$.

Bước (18)-(20): Vì ξ là phép toán *and* nên kết hợp điều kiện $TUOI =_{(TUOI)}$ *khả năng trẻ and* $LUONG \neq_{2(LUONG)}$ *rất cao* ta có $r_{result} = \{t_3\}$.

Bước (21): Kết quả $r_{result} = \{t_3\}$.

Vậy, câu truy vấn SQL mờ *select * from Nhanvien where* $TUOI =_{(TUOI)}$ *khả năng trẻ and* $LUONG \neq_{2(LUONG)}$ *rất cao* cho kết quả sau :

STT	TENNV	NGHENGHIEP	TUOI	LUONG
3	Hà	Bác sĩ	*rất khả năng trẻ*	500

Bảng 4.16. *Kết quả truy vấn trên quan hệ Nhanvien sử dụng thuật toán 4.4*

4.4.1.4. Thuộc tính suy dẫn trong truy vấn dữ liệu mờ

Khi thiết kế một CSDL, nếu miền trị của một thuộc tính nào đó sẽ nhận được từ việc tính toán hay kết hợp của hai hay nhiều miền trị của thuộc tính khác bằng phương pháp nào đó thì thông thường thuộc tính này sẽ không cần đưa vào CSDL để để tinh giảm trong việc thiết kế. Chẳng hạn như trong một CSDL có thuộc tính *SOLUONG* và *DONGIA*, khi đó nếu *THANHTIEN* = *SOLUONG*DONGIA* thì thuộc tính *THANHTIEN* sẽ không cần thiết đưa vào CSDL nếu không muốn lưu trữ giá trị *THANHTIEN* mà chỉ muốn giá trị *THANHTIEN* phục vụ cho việc thao tác hay truy vấn dữ liệu. Hoặc trong một trường hợp khác, trong quan hệ *Suckhoe_svien*, chúng ta muốn tìm những sinh viên nữ có sức khỏe *khả năng yếu*, nhưng thuộc tính *SUCKHOE* không có trong quan hệ mà có thể kết hợp từ hai thuộc tính *CHIEUCAO* và *TRONGLUONG* trong quan hệ *Suckhoe_svien*, như vậy thuộc tính *THANHTIEN* và *SUCKHOE* gọi là thuộc tính suy dẫn.

Do đó, một cách tổng quát nếu ta gọi P_a là thuộc tính suy dẫn từ các thuộc tính mờ A_1, A_2, …A_q có miền trị tương ứng $Dom(A_i)$, i = 1..q, F_a: $Dom(A_1) \times Dom(A_2) \times \ldots \times Dom(A_q) \rightarrow [0,1]$ là hàm kết nhập các ĐSGT $Dom(A_1)$, $Dom(A_2)$, …$Dom(A_q)$, x' là một giá trị mờ. Khi đó, điều kiện $P_a =_k x'$ trong câu truy vấn SQL mờ được xác định như sau: $P_a =_k x' \Leftrightarrow \Phi_k(F_a(x)) =_k x'$, với $x \in Dom(A_1) \times Dom(A_2) \times \ldots \times Dom(A_q)$.

Ví dụ 4.23. Cho lược đồ quan hệ U = {*TENHS*, *QUEQUAN*, *DIEMK1*, *DIEMK2*} và quan hệ *Hoctap* được xác định như sau:

TENHS	QUEQUAN	DIEMK1	DIEMK2
An	Huế	9	9
Bình	Hà Nội	8	10
Lan	QTrị	3	5
Nhân	Huế	8	7
Hùng	Đà Nẵng	4	6

Bảng 4.17. *Quan hệ Hoctap*

Hãy tìm những học sinh có kết quả học tập cả năm xếp loại *rất tốt* (với k =2).

Để thực hiện câu truy vấn này, chúng ta nhận thấy rằng, thuộc tính *CANAM* không có trong quan hệ *hoctap* nhưng có thể nhận được từ thuộc tính *DIEMHK1* và *DIEMHK2*. Vì vậy, trong câu truy vấn này chúng ta sử dụng thuộc tính suy dẫn *CANAM*.

Gọi $\underline{X}_{Diemhk1} = (\mathbf{X}_{Diemhk1}, G_{Diemhk1}, H_{Diemhk1}, \leq)$ là một ĐSGT với $G_{Diemhk1}$ = {*yếu*, *tốt*}, $H^+_{Diemhk1}$ = {*rất*, *hơn*}, $H^-_{Diemhk1}$ = {*khả năng*, *ít*}, *rất* > *hơn* và *ít* > *khả năng*.

Gọi $\underline{X}_{Diemhk2} = (\mathbf{X}_{Diemhk2}, G_{Diemhk2}, H_{Diemhk2}, \leq)$ là một ĐSGT với $G_{Diemhk2}$ = {*yếu*, *tốt*}, $H^+_{Diemhk2}$ = {*rất*, *hơn*}, $H^-_{Diemhk2}$ = {*khả năng*, *ít*}, *rất* > *hơn* và *ít* > *khả năng*.

Chọn hàm kết nhập ĐSGT $F = \sum_{i=1}^{2} \delta_i . v_i(x_i)$, với $\delta_1 = \delta_2 = 0.5$, khi đó ta có miền trị của thuộc tính $Dom(CANAM)$ = { 0.9, 0.9, 0.4, 0.8, 0.5 }.

Gọi $\underline{X}_{Canam}$= ($\boldsymbol{X}_{Canam}$, G_{Canam}, H_{Canam}, $\leq$) là một ĐSGT với G_{Canam} = {*yếu*, *tốt*}, H^{+}_{Canam} = {*rất*, *hơn*}, H^{-}_{Canam} = {*khả năng*, *ít*}, *rất* > *hơn* và *ít* > *khả năng*. Giả sử ta chọn $\boldsymbol{W}_{\text{Canam}}$ = 0.5, *fm*(*rất*) = 0.25, *fm*(*hơn*) = 0.25, *fm*(*khả năng*) = 0.25, *fm*(*ít*) = 0.25. Khi đó ta có *fm*(*tốt*) = 0.5, *fm*(*yếu*) = 0.4, *fm*(*rất tốt*) = 0.125, *fm*(*hơn tốt*) = 0.125, *fm*(*khả năng tốt*) = 0.125, *fm*(*ít tốt*) = 0.125.

Vì *ít tốt* < *khả năng tốt* < *tốt* < *hơn tốt* < *rất tốt* nên ta có *I*(*rất tốt*)=[0.875,1.0]. Mặc khác ta có υ(*rất tốt*) $\in$*I*(*rất tốt*), do đó *rất tốt* $=_2$ 0.9.

Vậy câu lệnh *select * from hoctap where CANAM* $=_2$ *rất tốt* cho kết quả:

TENHS	QUEQUAN	DIEMK1	DIEMK2
An	Huế	9	9
Bình	Hà Nội	8	10
Lan	QTrị	3	5
Nhân	Huế	8	7
Hùng	Đà Nẵng	4	6

Bảng 4.18. *Kết quả truy vấn trên quan hệ Hoctap*

Các câu truy vấn như dạng trên chỉ cần tìm trong quan hệ tham gia truy vấn những bộ dữ liệu nào "*thoả mãn*" các điều kiện mờ. Tuy nhiên, trong thực tế có những dạng câu truy vấn khá phức tạp mà chúng ta thường gặp, chẳng hạn: "*cho biết hầu hết những cán bộ trẻ có lương khá cao*" hoặc "*cho biết ít nhất 3 mặt hàng trong siêu thị bán ra với số lượng rất cao*" … thì vấn đề xử lý truy vấn là khá phức tạp. Do đó, cần nghiên cứu các dạng câu truy vấn có sử dụng lượng từ như vậy để đáp ứng việc tìm kiếm dữ liệu trong CSDL.

4.4.2. Truy vấn dữ liệu mờ với lượng từ ngôn ngữ

Để khai thác dữ liệu trong mô hình CSDL mờ nhiều tác giả mở rộng những ngôn ngữ hỏi đáp trên mô hình quan hệ như đại số quan hệ, ngôn ngữ SQL...Điểm mở rộng đó chính là sử dụng các điều kiện mờ, chẳng hạn như "*tìm những cán bộ trẻ có nhiều công trình khoa học công bố trên tạp chí quốc tế có uy tín*", "*cho biết các mặt hàng bán trong siêu thị thu được lợi nhuận khá lớn*"....Việc xử lý các câu hỏi dạng như vậy chúng ta chỉ cần tìm những bộ dữ liệu "*thỏa mãn*" điều kiện mờ. Tuy nhiên, khi gặp những yêu cầu như "*cho biết ít nhất 5 cán bộ trẻ có nhiều công trình khoa học công bố trên tạp chí quốc tế có uy tín*", "*cho biết một vài mặt hàng bán trong siêu thị thu được lợi nhuận khá lớn*"... thì vấn đề xử lý câu hỏi là phức tạp. Bởi vì, ngoài việc tìm những bộ dữ liệu "*thỏa mãn*" những điều kiện mờ còn phụ thuộc vào các lượng từ "*ít nhất 5*" và "*một vài*".

4.4.2.1. Tiếp cận ngữ nghĩa dữ liệu dựa trên việc định lượng ĐSGT

a. Phương pháp định giá lượng từ ngôn ngữ

Để xây dựng phương pháp định giá lượng từ trong truy vấn, trước hết chúng ta đi đánh giá các điều kiện mờ (fc_1, fc_2, ...fc_n) *with k* đối với quan hệ tham gia truy vấn. Có nghĩa là tìm những bộ dữ liệu γ thuộc quan hệ tham gia truy vấn thỏa mãn điều kiện mờ (fc_1, fc_2, ...fc_n) *with k*, đây chính là thực hiện truy vấn mờ trong mục 4.4.1. Tiếp theo, định giá lượng từ trong câu truy vấn dựa vào những bộ dữ liệu vừa tìm được so với số bộ dữ liệu của quan hệ ban đầu tham gia truy vấn. Chúng ta có thể chia lượng từ thành hai trường hợp:

Trường hợp Q là lượng từ tuyệt đối

Nếu Q đơn điệu tăng : khi đó gọi $\|Q\|$ là số lượng xác định của lượng từ Q. Ta xây dựng một hàm ABQ: N $\rightarrow$ {0,1} sao cho $\forall$ $x \in$ N, ABQ(x) = 1 nếu $x \geq \|Q\|$ và ABQ(x) = 0 nếu $x < \|Q\|$.

Nếu Q đơn điệu giảm: ta xây dựng một hàm ABQ: N $\rightarrow$ {0,1} sao cho $\forall$ $x \in$ N, ABQ(x)= 1 nếu $x \leq \|Q\|$ và ABQ(x) = 0 nếu x > $\|Q\|$.

Trường hợp Q là lượng từ tỷ lệ

Gọi n là số bộ dữ liệu của quan hệ ban đầu tham gia truy vấn. Khi đó tùy theo ngữ nghĩa của lượng từ Q để xây dựng hàm PRQ: N $\rightarrow$ {0,1} sao cho PRQ(x) = 1 hoặc PRQ(x) = 0, với $x \in$ N và $x \leq \beta$*n hoặc $x \geq \beta$*n, $\beta \in [0,1]$.

b. Đưa lượng từ ngôn ngữ vào câu truy vấn

(1) : Đi tìm những bộ dữ liệu thỏa mãn điều kiện (fc_1, fc_2, ...fc_n) *with k* và đếm tổng số bộ dữ liệu này.

(2) : Tùy thuộc vào lượng từ tuyệt đối hay lượng từ tỉ lệ để áp dụng hàm ABQ hay PRQ.

(3) : Nếu giá trị của hàm ABQ hoặc PRQ bằng 1 thì những bộ thỏa mãn điều kiện *Q*(fc_1, fc_2, ...fc_n) *with k* chính là kết quả của truy vấn với điều kiện (fc_1, fc_2, ...fc_n) *with k.*

Thuật toán 4.5. *Thực hiện truy vấn với lượng từ Q*.

Vào : *Quan hệ* r *xác định trên vũ trụ các thuộc tính* U = {A_1, A_2, ..., A_n}.

Câu truy vấn *select* <Các trường> *from r where Q*(fc_1, fc_2, .fc_n) *with k*.

Ra : Quan hệ r_{result} chứa những bộ dữ liệu thỏa mãn điều kiện *Q*(fc_1,fc_2,...fc_n) *with k*.

Phương pháp

(1) **Begin**

(2) Count = ||*r*||, trong đó ||*r*|| là số bộ dữ liệu của *r*.

(3) $r_{result} = \varnothing$

(4) **for** *each* γ *in* (*select* * *from* *r*) **do**

(5) **if** với mọi $i \in [1..n]$ sao cho γ thoả mãn fc_i **then** $r_{result} = r_{result} \cup \gamma$

(6) Count1 = $\|r_{result}\|$

(7) **if** *Q* là lượng tự tuyệt đối **then**

(8) **if** ABQ(count1) = 1 **then** *select* * *from* r_{result}

(9) **if** *Q* là lượng từ tỉ lệ **then**

(10) **if** **PRQ**(count1) = 1 **then** *select* * *from* r_{result}

(11) **End.**

Ví dụ 4.25. Sử dụng quan hệ *Luongtuoi* thực hiện câu truy vấn sử dụng lượng từ: Hãy cho biết *ít nhất 2* cá nhân có *sức khỏe tốt* và *lương cao* (trong ví dụ này chọn *k* =1). Câu lệnh truy vấn có dạng *select * from Luongtuoi where* "*ít nhất 2*" (*SUCKHOE* = *tốt*, *LUONG* = *cao*) *with 1*.

Xem miền trị của *SUCKHOE* và *LUONG* là ĐSGT như trong ví dụ 4.20 và áp dụng thuật toán 4.1 ta có quan hệ *Result* như sau:

SOCM	HOTEN	SUCKHOE	TUOI	LUONG
11111	Phạm Trọng Cầu	*rất rất tốt*	31	2.800.000
22222	Nguyễn Văn Tý	*rất tốt*	85	*cao*
88888	Thanh Tùng	*có thể tốt*	75	2.500.000

Bảng 4.19. *Quan hệ Result*

Vì lượng từ Q = *ít nhất 2* là lượng từ tuyệt đối và $\|Result\| = 3 \geq \|Q\|$ nên ABQ($\|Result\|$) = 1. Vậy quan hệ *Result* chính là kết quả của truy vấn.

Ví dụ 4.24. Sử dụng quan hệ *Luongtuoi* trong ví dụ 4.20 thực hiện câu truy vấn sử dụng lượng từ: Hãy cho biết *hầu hết* cá nhân có *sức khỏe xấu* là có *lương thấp* (trong ví dụ này chọn k =1). *select * from Luongtuoi where* "*hầu hết*" (*SUCKHOE = xấu*, *LUONG = thấp*) *with 1*. Áp dụng thuật toán 4.1 ta có quan hệ *Result1* như sau:

SOCM	HOTEN	SUCKHOE	TUOI	LUONG
44444	Vũ Hoàng	*Khá* xấu	45	500.000
66666	Thuận Yến	*Có thể* xấu	61	Thấp

Bảng 4.20. *Quan hệ Result1*

Vì lượng từ Q = *hầu hết* là lượng từ tỉ lệ nên ta xây dựng hàm PRQ : $N \rightarrow \{0,1\}$ sao cho: $\forall\ x \in N$, PRQ(x) = 1 nếu $x \geq 0.8*\|Luongtuoi\|$ và PRQ(x) = 0 nếu ngược lại.

Ta có $\|Result1\| = 2 < 0.8*\|Luongtuoi\|$ nên PRQ($\|Result1\|$)= 0. Vậy kết quả của truy vấn là một quan hệ rỗng.

4.4.2.2. Tiếp cận ngữ nghĩa lân cận tôpô của ĐSGT

Do truy vấn sử dụng lượng từ có thể xem là một sự mở rộng của truy vấn mờ cho nên một câu truy vấn SQL mờ sử dụng lượng từ có thể tổng quát dạng: *select* <các thuộc tính> *from* <các quan hệ> *where* $Q(fc_1, fc_2, ...fc_n)$, trong đó Q là lượng từ và $fc_1, fc_2, ...fc_n$ là các điều kiện mờ. Chẳng hạn như trong quan hệ *Nhanvien* ở ví dụ 4.21, tìm "*ít nhất một nửa*" *nhân viên Tuổi trẻ và có Lương thấp*. Khi đó câu truy vấn có dạng: *select * from Nhanvien where* "*ít nhất một nửa*" ($TUOI =_{1(TUOI)}$ *trẻ and* $LUONG =_{1(LUONG)}$ *cao*).

Không mất tính tổng quát, chúng ta giả sử các điều kiện mờ fc_i trong câu truy vấn có dạng $A_i =_{k(Ai)} fvalue_i$, trong đó A_i là thuộc tính mờ có tính đơn điệu và $fvalue_i$ là giá trị mờ, phép toán liên kết các điều kiện là phép *and* hoặc *or*. Do đó, điều kiện $Q(fc_1, fc_2, ...fc_n)$ có dạng: $Q(A_1 =_{k(A1)} fvalue_1\ \xi\ A_2 =_{k(A2)} fvalue_2\ .....\xi\ A_n =_{k(An)} fvalue_n)$, trong đó ξ là phép *and* hoặc *or*.

Thuật toán xử lý lượng từ trong câu truy vấn SQL mờ

Có thể xây dựng thuật toán xử lý lượng từ trong truy vấn SQL mờ như sau:

(1) : Đếm tổng số những bộ dữ liệu thỏa mãn điều kiện ($fc_1, fc_2, ...fc_n$).

(2) : Tùy thuộc vào phân loại lượng từ để chọn phương pháp đánh giá lượng từ phù hợp với yêu cầu.

Thuật toán 4.6. *Thực hiện truy vấn với lượng từ Q.*

Vào : Quan hệ r xác định trên tập vũ trụ các thuộc tính $U = \{A_1, A_2, \ldots, A_n\}$. Câu truy vấn *select * from r where* $Q(A_1 =_{k(A1)} fvalue_1 \ \xi \ A_2 =_{k(A2)} fvalue_2 \ .....\xi \ A_n =_{k(An)} fvalue_n)$, trong đó ξ là phép *and* hoặc *or*, $Q \in Q_{ABS} \cup Q_{PRO}$, với Q_{ABS} là tập các lượng từ tuyệt đối, Q_{PRO} là tập các lượng từ tỉ lệ.

Ra : Một quan hệ $r_{resultQ}$ chứa những bộ dữ liệu t thỏa mãn điều kiện $Q(A_1 =_{k(A1)} fvalue_1 \ \xi \ A_2 =_{k(A2)} fvalue_2 \ .....\xi \ A_n =_{k(An)} fvalue_n)$.

Phương pháp

(1) **Begin**

(2) $r_{resultQ} = \varnothing$

(3) Sử dụng thuật toán 4.3 trong trường hợp đơn điều kiện và thuật toán 4.4 trong trường hợp đa điều kiện ta có kết quả là quan hệ r_{result}.

(4) **if** $Q \in Q_{ABS}$ **then**

(5) **if** $f_Q^A(\|r_{result}\|) = 1$ *or* $f_Q^D(\|r_{result}\|) = 1$ **then** $r_{resultQ} = r_{result}$

(6) **elseif**

(7) **if** $Q \in Q_{PRO}$ **then**

(8) **begin**

(9) Xây dựng các khoảng: $S(\boldsymbol{1})$, $S(lớn)$, $S(\boldsymbol{W})$, $S(nhỏ)$, $S(\boldsymbol{0})$

(10) **Case** Q **of**

(11) "*Một ít*" : **if** $(\|r_{result}\| / \|r\|) \in S(\boldsymbol{0})$ **then** $r_{resultQ} = r_{result}$

(12) "*Khoảng một nữa*" : **if** $(\|r_{result}\| / \|r\|) \in S(\boldsymbol{W})$ **then** $r_{resultQ} = r_{result}$

(13) "*Hầu hết*" : **if** $(\|r_{result}\| / \|r\|) \in S(\boldsymbol{1})$ **then** $r_{resultQ} = r_{result}$

(14) "*Với mọi*" : **if** $(\|r_{result}\| / \|r\|) = 1$ **then** $r_{resultQ} = r_{result}$

(15) **end**

(16) Return $r_{resultQ}$

(17) **End.**

Ví dụ 4.25. Sử dụng quan hệ Nhanvien trong ví dụ 4.21

(i) Cho biết *ít nhất 5* nhân viên có *Tuổi khả năng trẻ*, khi đó câu truy vấn SQL mờ có dạng: *select * from Nhanvien where* " *ít nhất 5*" (*TUOI* $=_{2(LUONG)}$ *khả năng trẻ*)

Bước (1)-(2): $r_{resultQ} = \varnothing$.

Bước (3): Vì điều kiện trong câu truy vấn là đơn điều kiện nên áp dụng thuật toán 4.3 ta thực hiện câu truy vấn *select * from Nhanvien where TUOI* $=_{2(TUOI)}$ *khả năng trẻ*. Theo ví dụ 4.21 ta có kết quả.

STT	TENNV	NGHENGHIEP	TUOI	LUONG
1	An	Giáo viên	45	*rất cao*
3	Hà	Bác sĩ	*rất khả năng trẻ*	500
5	Nhân	Giáo viên	46	1500

Bảng 4.21. *Kết quả truy vấn (i) trên quan hệ Nhanvien chưa sử dụng lượng từ*

Bước (4)-(15): Tiếp theo chúng ta đi đánh giá lượng từ *ít nhất 5* theo thuật toán 4.3. Vì lượng từ *ít nhất 5* $\in Q_{ABS}$ và đơn điệu tăng, ta có $f_{\text{ít nhất 5}}{}^{A}(||r_{result}||) = f_{\text{ít nhất 5}}{}^{A}(3) = 0$ nên kết quả của câu truy vấn *select * from Nhanvien where ít nhất 5* (*TUOI* $=_{2(LUONG)}$ *khả năng trẻ*) không có bộ nào.

Bước (16): Vậy $r_{resultQ} = \varnothing$.

(ii): Cho biết *hầu hết những cán bộ Tuổi khả năng trẻ và có Lương rất cao*, khi đó câu truy vấn SQL mờ : *select * from Nhanvien where* "*Hầu hết*" (*TUOI* $=_{2(TUOI)}$ *khả năng trẻ and LUONG* $=_{2(LUONG)}$ *rất cao*).

Bước (1)-(2): $r_{resultQ} = \varnothing$.

Bước (3): Vì điều kiện trong câu truy vấn là đa điều kiện nên áp dụng thuật toán 4.4 ta thực hiện câu truy vấn *select * from Nhanvien where* (*TUOI* $=_{2(TUOI)}$ *khả năng trẻ*) *and* (*LUONG* $=_{2(LUONG)}$ *rất cao*). Theo ví dụ 4.22 ta có kết quả.

STT	TENNV	NGHENGHIEP	TUOI	LUONG
1	An	Giáo viên	45	*rất cao*
5	Nhân	Giáo viên	46	1500

Bảng 4.22. *Kết quả truy vấn (ii) trên quan hệ Nhanvien chưa sử dụng lượng từ*

Bước (4)-(15): Vì lượng từ *Hầu hết* $\in Q_{PRO}$ nên đi xây dựng các khoảng $S(\boldsymbol{1})$, $S(lớn)$, $S(\boldsymbol{W})$, $S(nhỏ)$, $S(\boldsymbol{0})$. Chọn $fm(lớn) = 0.35$, $fm(nhỏ) = 0.65$, $\mu(khả\ năng) = 0.25$, $\mu(ít) = 0.2$, $\mu(hơn) = 0.15$ và $\mu(rất) = 0.4$. Ta phân hoạch đoạn [0, 1] thành 5 khoảng *tương tự mức 1* là: $fm(rất\ lớn) = 0.35 \times 0.35 = 0.1225$. Vậy $S(\boldsymbol{1}) = (0.8775, 1]$.

$(fm(khả\ năng\ lớn) + fm(hơn\ lớn)) = (0.25 \times 0.35 + 0.15 \times 0.35) = 0.14$. Vậy $S(lớn) = (0.7375, 0.8775]$.

$(fm(ít\ nhỏ) + fm(ít\ lớn)) = (0.25 \times 0.65 + 0.25 \times 0.35) = 0.25$. Vậy $S(\boldsymbol{W}) = (0.4875, 0.8775]$, $(fm(khả\ năng\ nhỏ) + fm(hơn\ nhỏ)) = (0.25 \times 0.65 + 0.15 \times 0.65) = 0.26$. Vậy $\boldsymbol{S}(nhỏ) = (0.2275, 0.4875]$ và $S(\boldsymbol{0}) = [0, 0.2275]$.

Vì $(\|r_{result}\| / \|r\|) = (2/9) = 0.222 \notin S(\boldsymbol{1})$ nên kết quả câu truy vấn *select * from Nhanvien where "Hầu hết"* ($TUOI =_{2(TUOI)}$ *khả năng trẻ and* $LUONG =_{2(LUONG)}$ *rất cao*) không có bộ nào thoả mãn.

Bước (16): Vậy $r_{resultQ} = \varnothing$.