Tạp chí Tin học và Điều khiển học, T.23, S.2 (2007), 110–121

MỘT CÁCH TIẾP CẬN ĐỂ XẤP XỈ DỮ LIỆU TRONG CƠ SỞ DỮ LIỆU MỜ

NGUYỄN CÁT HỒ[1], NGUYỄN CÔNG HÀO[2]

[1] *Viện Công nghệ thông tin, Viện Khoa học và Công nghệ Việt Nam*
[2] *Trường Đại học Khoa học Huế*

Abstract. In this paper, we introduced a method to approximate data on domain of fuzzy attributes in relation of fuzzy databases based hedge algebra. Because, domain of fuzzy attributes can except values are number, linguistic values, thus we have to effect and simply on method to approximate data.

Tóm tắt. Bài báo trình bày một phương pháp xấp xỉ dữ liệu trên miền trị thuộc tính mờ của một quan hệ trong cơ sở dữ liệu mờ dựa trên đại số gia tử. Bởi vì miền trị của thuộc tính mờ có thể là giá trị số, giá trị ngôn ngữ, do đó chúng ta cần có một phương pháp xấp xỉ dữ liệu một cách đơn giản và hiệu quả.

1. ĐẶT VẤN ĐỀ

Cơ sở dữ liệu mờ đã được nhiều tác giả trong và ngoài nước quan tâm nghiên cứu và đã có những kết quả đáng kể ([1–5, 10, 12]). Có nhiều cách tiếp cận khác nhau như cách tiếp cận theo lý thuyết tập mờ ([1]), theo lý thuyết khả năng ([4]) Prade và Testemale năm 1983, quan hệ tương đương ([2, 3, 5])... Tất cả các cách tiếp cận trên nhằm mục đích nắm bắt và xử lý một cách thỏa đáng trên một luận điểm nào đó các thông tin không chính xác (unexact), không chắc chắn (uncertainty) hay những thông tin không đầy đủ (incomplete). Do sự đa dạng của những loại thông tin này nên ta gặp rất khó khăn trong biểu thị ngữ nghĩa và thao tác với chúng.

Trong thời gian qua, đại số gia tử được nhiều tác giả nghiên cứu trong [6–8] và đã có những ứng dụng đáng kể, đặc biệt trong lập luận xấp xỉ và trong một số bài toán điều khiển. Vì vậy, việc nghiên cứu về cơ sở dữ liệu mờ theo cách tiếp cận đại số gia tử là một hướng mới cần quan tâm giải quyết.

2. ĐẠI SỐ GIA TỬ

Để xây dựng cách tiếp cận đại số gia tử, trong phần này sẽ trình bày tổng quan về một số nét cơ bản của đại số gia tử và khả năng biểu thị ngữ nghĩa dựa vào cấu trúc của đại số gia tử, hàm định lượng ngữ nghĩa và một số tính chất của đại số gia tử.

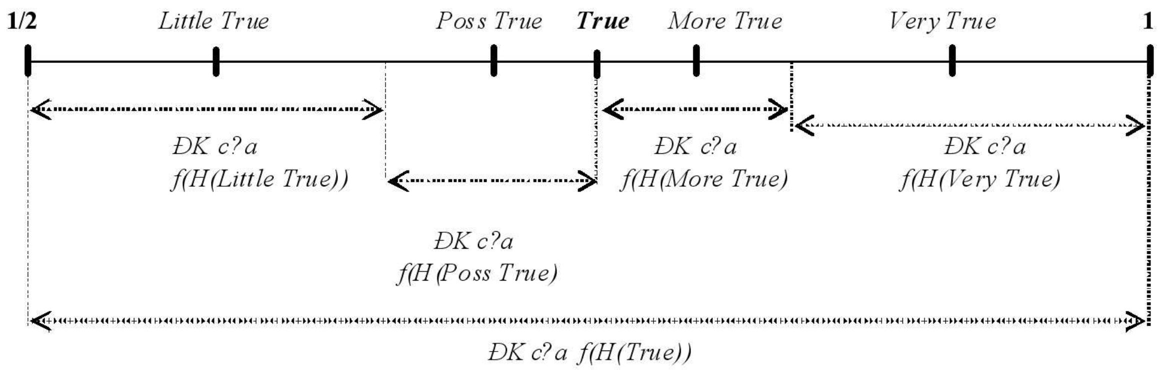
Ta xét miền ngôn ngữ của biến chân lý TRUTH gồm các từ sau:

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 (mordifier hay 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 $AH = (X, G, H, \leqslant)$, trong đó G là tập các từ nguyên thuỷ được xem là các phần tử sinh. H là tập các gia tử được xem như là các phép toán một ngôi, quan hệ (trên các từ (các khái niệm mờ) là quan hệ thứ tự được "cảm sinh" từ ngữ nghĩa tự nhiên. Ví dụ dựa trên ngữ nghĩa, các quan hệ thứ tự sau là đúng: false $\leqslant$ true, *more* true $\leqslant$ *very* true nhưng*very* false $\leqslant$*more* false, *possibly* true $\leqslant$ true nhưng false $\leqslant$ *possibly* false... Tập X được sinh ra từ G bởi các phép tính trong H. Như vậy mỗi phần tử của X sẽ có dạng biểu diễn $x = h_n h_{n-1} ... h_1 x,\ x \in G$. 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 dương còn *false* là âm. Cho đại số gia tử $X = (X, G, H, \leqslant)$, với $G = \{c^+, c^-\}$, trong đó c^+ và c^- tương ứng là phần tử sinh dương và âm, X là tập nền. $H = H^+ \cup H^-$ với $H^- = \{h_1, h_2, ..., h_p\}$ và $H^+ = \{h_{p+1}, ..., h_{p+q}\},\ h_1 > h_2 > ... > h_p$ và $h_{p+1} < ... < h_{p+q}$.

Định nghĩa 2.1. ([9]) $f : X \to [0, 1]$ gọi là hàm định lượng ngữ nghĩa của X nếu $\forall h, \in H^+$ hoặc $\forall h, k \in H^-$ và $\forall x, y \in X$, ta có:

$$\Big|\frac{f(hx) - f(x)}{f(kx) - f(x)}\Big| = \Big|\frac{f(hy) - f(y)}{f(ky) - f(y)}\Big|.$$

Với đại số gia tử và hàm định lượng ngữ nghĩa ta có thể định nghĩa tính mờ của một khái niệm mờ. Cho trước hàm định lượng ngữ nghĩa f của X. Xét bất kỳ $x \in X$. Tính mờ của x khi đó được đo bằng đường kính của tập $f(H(x)) \subseteq [0, 1]$.



Hình 1. Tính mờ của giá trị True

Định nghĩa 2.2. [9] Hàm $fm : X \to [0, 1]$ được gọi là *độ đo tính mờ* trên X nếu thoả mãn các điều kiện sau:

(1) $fm(c^-) = W > 0$ và $fm(c^+) = 1 - W > 0$

(2) Với $c \in \{c^-, c^+\}$ thì $\sum\limits_{i=1}^{p+q} fm(h_i c) = fm(c)$

(3) Với mọi $x, y \in X, \forall h \in H,\ \dfrac{fm(hx)}{fm(x)} = \dfrac{fm(hy)}{fm(y)} = \dfrac{fm(hc)}{fm(c)},$ với $c \in \{c^-, c^+\}$

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.

Mệnh đề 2.1. [9]

(1) $fm(hx) = \mu(h)fm(x)$, với mọi $x \in X$

(2) $\sum_{i=1}^{p+q} fm(h_i c) = fm(c)$, trong đó $c \in \{c^-, c^+\}$

(3) $\sum_{i=1}^{p+q} fm(h_i x) = fm(x)$, $\forall x \in X$

(4) $\sum_{i=1}^{p} \mu(h_i) = \alpha$ và $\sum_{i=p+1}^{p+q} \mu(h_i) = \beta$, với $\alpha, \beta > 0$ và $\alpha + \beta = 1$.

Định nghĩa 2.3. [9] Hàm $Sign : X \to \{-1, 0, 1\}$ là một ánh xạ được định nghĩa một cách đệ qui như sau, với mọi $h, h' \in H$:

(1) $Sign(c^-) = -1$ và $Sign(hc^-) = +Sign(c^-)$ nếu $hc^- < c^-$

$Sign(hc^-) = -Sign(c^-)$ nếu $hc^- > c^-$

$Sign(c^+) = +1$ và $Sign(hc^+) = +Sign(c^+)$ nếu $hc^+ > c^+$

$Sign(hc^+) = -Sign(c^+)$ nếu $hc^+ < c^+$

(2) $Sign(h'hx) = -Sign(hx)$ nếu h' là negative đối với h và $h'hx \neq hx$

(3) $Sign(h'hx) = +Sign(hx)$ nếu h' là positive đối với h và $h'hx \neq hx$

(4) $Sign(h'hx) = 0$ nếu $h'hx = hx$.

Định nghĩa 2.4. [9] Giả sử cho trước độ đo tính mờ của các gia tử $\mu(h)$, và các giá trị độ đo tính mờ của các phần tử sinh $fm(c^-), fm(c^+)$ và w là phần tử trung hòa. Hàm định lượng ngữ nghĩa (quantitatively semantic function) ν của X được xây dựng như sau với $x = h_{im}...h_{i2}h_{i1}c$:

(1) $\nu(c^-) = W - \alpha.fm(c^-)$ và $\nu(c^+) = W + \alpha.fm(c^+)$

(2) $\nu(h_j x) =$

$$\nu(x) + Sign(h_j x) \times \Big[\sum_{i=j}^{p} fm(h_i x) - \frac{1}{2}\Big(1 - Sign(h_j x) Sign(h_1 h_j x)(\beta - \alpha)\Big) fm(h_j x)\Big]$$

với $1 \leqslant j \leqslant p$, và

$$\nu(h_j x) = \nu(x) + Sign(h_j x) \times \Big[\sum_{i=p+1}^{j} fm(h_i x) - \frac{1}{2}\Big(1 - Sign(h_j x) Sign(h_1 h_j x)(\beta - \alpha)\Big) fm(h_j x)\Big]$$

với $j > p$.

3. MỘT CÁCH TIẾP CẬN ĐỂ XẤP XỈ DỮ LIỆU MỜ

Trong mục này, sẽ trình bày một phương pháp mới để xấp xỉ dữ liệu trên miền trị của thuộc tính mờ trong quan hệ của cơ sở dữ liệu mờ. Việc đánh giá dữ liệu trên miền trị thuộc tính mờ của quan hệ trong cơ sở dữ liệu mờ theo cách tiếp cận đại số gia tử được xây dựng dựa trên phân hoạch tính mờ của các giá trị trong đại số gia tử (giá trị ngôn ngữ). Như vậy, nếu gọi $Dom(A_i)$ là miền trị tương ứng với thuộc tính mờ A_i và xem như một đại số gia tử thì khi đó $Dom(A_i) = \mathrm{Num}(A_i) \cup LV(A_i)$, với $\mathrm{Num}(A_i)$ là tập các giá trị số của A_i và $LV(A_i)$ là tập các giá trị ngôn ngữ của A_i. Để xấp xỉ dữ liệu, ta xét hai trường hợp sau.

3.1. Miền trị của thuộc tính trong quan hệ là giá trị ngôn ngữ

Trong trường hợp này, chúng ta đi xây dựng các phân hoạch dựa vào tính mờ của các giá trị ngôn ngữ.

Vì tính mờ của các giá trị trong đại số gia tử 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.

Định nghĩa 3.1. Gọi fm là độ đo tính mờ trên ĐSGT X. Với mỗi $x \in X$, ta ký hiệu $I(x) \subseteq [0,1]$ và $|I(x)|$ là độ dài của $I(x)$.

Một họ các $\xi = \{I(x) : x \in X\}$ được gọi là phân hoạch của [0,1] gắn với x nếu:

(1) $\{I(c^+), I(c^-)\}$ là phân hoạch của [0,1] sao cho$|I(c)| = fm(c)$, với $c \in \{c^+, c^-\}$.

(2) Nếu đoạn $I(x)$ đã được định nghĩa và $|I(x)| = fm(x)$ thì $\{I(h_ix) : i = 1..p+q\}$ được định nghĩa là phân hoạch của $I(x)$ sao cho thoả mãn điều kiện $|I(h_ix)| = fm(h_ix)$ và $|I(h_ix)|$ là tập sắp thứ tự tuyến tính.

Tập $\{I(h_ix)\}$ được gọi là phân hoạch gắn với phần tử x. Ta có $\sum_{i=1}^{p+q} |I(h_ix)| = |I(x)| = fm(x)$.

Định nghĩa 3.2. Cho $P^k = \{I(x) : x \in \boldsymbol{X}^k\}$ với $\boldsymbol{X}^k = \{x \in \boldsymbol{X} : |x| = k\}$ là một phân hoạch. Ta nói rằng u xấp xỉ ν theo mức k trong P^k được ký hiệu $u \approx_k \nu$ khi và chỉ khi $I(u)$ và $I(v)$ cùng thuộc một khoảng trong P^k. Có nghĩa là $\forall u, v \in \boldsymbol{X},\ u \approx_k v \Leftrightarrow \exists \Delta^k \in P^k : I(u) \subseteq \Delta^k$ và $I(v) \subseteq \Delta^k$.

Ví dụ 3.1. Cho đại số gia tử $\underline{X} = (\boldsymbol{X}, G, H, \leqslant)$, trong đó $H = H^+ \cup H^-$, $H^+ = \{$*hơn, rất*$\}$, *hơn* $<$ *rất*, $H^- = \{$*ít, khả năng*$\}$, *ít* $>$ *khả năng*, $G = \{$ *trẻ, già*$\}$. Ta có $P^1 = \{I($trẻ$)$, $I($già$)\}$ là một phân hoạch của $[0,1]$. Tương tự, $P^2 = \{$*I(hơn trẻ), I(rất trẻ), I(ít trẻ), I(khả năng trẻ), I(hơn già), I(rất già), I(ít già), I(khả năng già)*$\}$ là phân hoạch của $[0,1]$.

Ví dụ 3.2. Theo Ví dụ 3.1, P^1 là phân hoạch của $[0,1]$. Ta có *hơn trẻ* $\approx_1$ *rất trẻ* vì $\exists \Delta^1 =$ *I(trẻ)* $\in P^1$ mà *I(hơn trẻ)* $\subseteq \Delta^1$ và *I(rất trẻ)* $\subseteq \Delta^1$.P^2 là phân hoạch của $[0,1]$, ta có *ít già* $\approx_2$ *rất ít già* vì $\exists \Delta^2 =$ *I(ít già)*$\in P^2$ mà *I(ít già)* $\subseteq \Delta^2$ và *I(rất ít già)* $\subseteq \Delta^2$.

Định nghĩa 3.3. Xét $P^k = \{I(x) : x \in \boldsymbol{X}^k\}$ với $\boldsymbol{X}^k = \{x \in \boldsymbol{X} : |x| = k\}$ là một phân hoạch. Ta nói rằng u không xấp xỉ v mức k trong P^k được ký hiệu $u \neq_k v$ khi và chỉ khi $I(u)$ và $I(v)$ không cùng thuộc một khoảng trong P^k. Có nghĩa là $\forall u, v \in \boldsymbol{X},\ u \neq_k v \Leftrightarrow \forall \Delta^k \in P^k$: $I(u) \not\subset \Delta^k$ hoặc $I(v) \not\subset \Delta^k$.

Ví dụ 3.3. Theo Ví dụ 3.1, $P^2 = \{$*I(hơn trẻ), I(rất trẻ), I(ít trẻ), I(khả năng trẻ), I(hơn già), I(rất già), I(ít già), I(khả năng già)*$\}$ là phân hoạch của $[0,1]$. Chọn $\Delta^2 =$ *I(rất trẻ)*$\in P^2$*, ta có I(ít trẻ)* $\not\subset \Delta^2$ và *I(rất trẻ)* $\subseteq \Delta^2$ *(1')*. Mặc khác với mọi $\Delta^2 \neq$ *I(ít trẻ)* $\in P^2$ *ta có I(ít trẻ)* $\not\subset \Delta^2$ và *I(rất trẻ)* $\not\subset \Delta^2$ *(2')* . Từ (1') và (2') ta suy ra *ít trẻ* $\neq_2$ *rất trẻ*.

Định nghĩa 3.4. Xét $P^k = \{I(x) : x \in \boldsymbol{X}^k\}$ với $\boldsymbol{X}^k = \{x \in \boldsymbol{X} : |x| = k\}$ là một phân hoạch. Gọi ν là hàm định lượng ngữ nghĩa trên $\boldsymbol{X}$. Ta nói rằng u nhỏ hơn v mức k trong P^k được ký hiệu $u <_k v$ khi và chỉ khi $I(u)$ và $I(v)$ không cùng thuộc một khoảng trong P^k và

$\nu(u) < u(v)$. Có nghĩa là $\forall u, v \in \boldsymbol{X},\ u <_k v \Leftrightarrow u \neq_k v$ và $\nu(u) < v(v)$.

Ví dụ 3.4. Theo Ví dụ 3.1 và 3.3 ta có $P^2 = \{$*I(hơn trẻ), I(rất trẻ), I(ít trẻ), I(khả năng trẻ), I(hơn già), I(rất già), I(ít già), I(khả năng già)*$\}$ là phân hoạch của $[0, 1]$. Vì *ít trẻ* $\neq_2$ *rất trẻ* và *v(rất trẻ)* $<$ *v (ít trẻ)* nên *rất trẻ* $<_2$ *ít trẻ.*

Các định lý, hệ quả và bổ đề liên quan đến những quan hệ được đề xuất trong Mục 3.1 như xấp xỉ, không xấp xỉ theo mức trong phân hoạch sẽ được trình bày và chứng minh đầy đủ làm cơ sở cho các phần tiếp theo.

Bổ đề 3.1. *Quan hệ $\approx_k$ là một quan hệ tương đương trên $Dom(A_i)$.*

Chứng minh: Ta chứng minh tính phản xạ 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 $\approx_k$ đúng với $k = 1$, hay $x \approx_1 x$.

Giả sử $|x| = n$ đúng, có nghĩa $\approx_k$ đúng với $k = n$, hay $x \approx_n x$, ta cần chứng minh $\approx_k$ đúng với $k = n + 1$. Đặt $x = h1x'$, với $|x'| = n$. Vì $x \approx_n x$ nên theo định nghĩa ta có $\exists \Delta^n \in P^n : I(x) \subseteq \Delta^n$. Mặc khác ta có $P^{n+1} = \{I(h_1x'), I(h_2x'), ...\}$, với $h_1 \neq h_2 \neq ...$ 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 $\approx_k$ đúng với $k = n + 1$, hay $x \approx_{n+1} x$.

Tính đối xứng: $\forall x, y \in Dom(A_i)$, nếu $x \approx_k y$ 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 $y \approx_k x$ thì $y \approx_k x$.

Tính bắt cầu: Ta chứng minh bằng phương pháp qui nạp.

Trường hợp $k = 1$:

Ta có $P^1 = \{I(c^+), I(c^-)\}$, nếu $x \approx_1 y$ và $y \approx_1 z$ 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 $x \approx_1 z$. Vậy $\approx_k$ đúng với $k = 1$.

Giả sử quan hệ $\approx_k$ đúng với trường hợp $k = n$ có nghĩa là ta có $\forall x, y, z \in Dom(A_i)$ nếu $x \approx_n y$ và $y \approx_n z$ thì $x \approx_n z$.

Ta cần chứng minh quan hệ $\approx_k$ đúng với trường hợp $k = n+1$. Tức là $\forall x, y, z \in Dom(A_i)$ nếu $x \approx_{n+1} y$ và $y \approx_{n+1} z$ thì $x \approx_{n+1} z$.

Theo giả thiết nếu $x \approx_{n+1} y$ và $y \approx_{n+1} z$ 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 $x \approx_{n+1} z$.

Bổ đề 3.2. *Cho $u = h_n..h_1x$ và $v = h'_m...h'_1x$ là biểu diễn chính tắc của u và v đối với x.*

(1) *Nếu $u = v$ thì $u \approx_k v$ với mọi k.*

(2) *Nếu $h_1 \neq h'_1$ thì $u \approx_{|x|} v$.*

Chứng minh:

(1) Theo Bổ đề 3.1, vì $u = v$ nên ta có $u \approx_k u$ hay $v \approx_k v$, 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 $h_1x \approx_1 h'_1x$. Vậy $u \approx_{|x|} v$.

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 $u \approx_k$ v thì

$\exists \Delta^k \in P^k = \{I(h_{k-1}...h_1x), I(h'_{k-1}...h'_1x)\}$, với P^k là một phân hoạch của $I(x) : I(u) \subseteq \Delta^k$ và $I(v) \subseteq \Delta^k$.

Nếu chọn $\Delta^k = I(h_{k-1}...h_1x)$ thì $I(u) \subseteq I(h_{k-1}...h_1x)$ và $I(v) \subseteq I(h_{k-1}...h_1x)$ hay $I(h_n...h_1x) \subseteq I(h_{k-1}...h_1x)$ và $I(h'_m...h'_1x) \subseteq I(h_{k-1}...h_1x)$ điều này mâu thuẫn vì $I(h'_m...h'_1x) \not\subset I(h_{k-1}...h_1x)$ do (1').

Nếu chọn $\Delta^k = I(h'_{k-1}...h'_1x)$ thì $I(h_n...h_1x) \subseteq I(h'_{k-1}...h'_1x)$ và $I(h'_m...h'_1x) \subseteq I(h'_{k-1}...h'_1x)$, điều này mâu thuẫn vì $I(h_n...h_1x) \not\subset I(h'_{k-1}...h'_1x)$ do (1'). Vậy không tồn tại $k > 1$ sao cho $u \approx_k v$ hay $k = 1$. Vậy $u \approx_{|x|} v$.

Định lý 3.1. *Xét $P^k = \{I(x) : x \in \boldsymbol{X}^k\}$ với $\boldsymbol{X}^k = \{x \in \boldsymbol{X} : |x| = k\}$ là một phân hoạch, $u = h_n...h_1x$ và $v = h'_m...h'_1x$ là biểu diễn chính tắc của u và v đối với x.*

(1) Nếu $u \approx_k v$ thì $u \approx_{k'} v$, $\forall 0 < k' < k$.

(2) Nếu tồn tại một chỉ số $j \leqslant \min(m, n)$ lớn nhất sao cho với mọi $s = 1...j$, ta có $h_s = h'_s$ thì $u \approx_{j+|x|} v$.

Chứng minh: (1) Ta có $P^k = \{I(h_{k-1}...h_1x), I(h'_{k-1}...h_1x)\}$. Vì $u \approx_k v$ nên theo định nghĩa $\exists \Delta^k \in P^k : I(u) \subseteq \Delta^k$ và $I(v) \subseteq \Delta^k$ (1').

Ta lại có $P^1 = \{I(x)\}$, $P^2 = \{I(h_1x), I(h'_1x)\}, ...,$ $P^k = \{I(h_{k-1}...h_1x), I(h'_{k-1}...h_1x)\}$. Mặt khác ta có $I(h_{k-1}...h_1x) \subseteq I(h_{k-2}...h_1x) \subseteq ... \subseteq I(h_1x) \subseteq I(x)$ và $I(h'_{k-1}...h'_1x) \subseteq I(h'_{k-2}...h'_1x) \subseteq ... \subseteq I(h'1x) \subseteq I(x)$ nên $\exists \Delta^k = I(h_{k-1}...h_1x) \in P^k$ hoặc $\exists \Delta^k = I(h'_{k-1}...h'_1x) \in P^k$ và $\exists \Delta^{k-1} = I(h_{k-2}...h_1x) \in P^{k-1}$ hoặc $\exists \Delta^{k-1} = I(h'_{k-2}...h'_1x) \in P^k - 1...$ 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: $\Delta^k \subseteq \Delta^{k-1} \subseteq ... \subseteq \Delta^2 \subseteq \Delta^1$ (2').

Từ (1') và (2') ta có $I(u) \subseteq \Delta^k \subseteq \Delta^{k-1} \subseteq ... \subseteq \Delta^2 \subseteq \Delta^1$ và $I(v) \subseteq \Delta^k \subseteq \Delta^{k-1} \subseteq ... \subseteq \Delta^2 \subseteq \Delta^1$, 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 $u \approx_k v$ thì $u \approx_{k'} v$.

(2): Nếu $j = 1$ ta có $h_1 = h'_1$, khi đó $u = h_n...h_2h_1x$ và $v = h'_m...h'_2h'_1x$ hay $u = h_n...h_2h_1x$ và $v = h'_m...h'_2h_1x$. Đặt $x' = h_1x$ ta có $u = h_n...h_2x'$ và $v = h'_m...h'_2x'$. Vì $h_2 \neq h'2$ nên theo Bổ đề 2.3 ta có $u \approx_{|x'|} v$ (do $|x'| = 2,\ |x| = 1$) hay $u \approx_2 v$. Vậy $u \approx_{j+|x|} v$.

Nếu $j \neq 1$, đặt $k = j$, ta cần chứng minh $u \approx_{k+|x|} v$. Vì $u \approx_k v$ nên theo giả thiết ta có $\forall s = 1...k$ ta có $h_s = h'_s$. Khi đó $u = h_n...h_2h_1x$ và $v = h'_m...h'_2h'_1x$ hay $u = h_n.h_kh_{k-1}...h_1x$ và $v = h'_m...h_kh_{k-1}...h_1x$.

Đặt $x' = h_kh_{k-1}...h_1x$ ta có $u = h_n...h_{k+1}x'$ và $v = h'_m...h'_{k+1}x'$. Vì $h_{k+1} \neq h'_{k+1}$ nên theo Bổ đề 2.2 ta có $u \approx_{|x'|} v$ hay $u \approx_{k+|x|} v$ (do $|x'| = k,\ |x| = 1$).

Hệ quả 3.1. *Nếu $u \in H(v)$ thì $u \approx_{|v|} v$.*

Định lý 3.2. *Xét $P^k = \{I(x) : x \in \boldsymbol{X}^k\}$ với $\boldsymbol{X}^k = \{x \in \boldsymbol{X} : |x| = k\}$, $u = h_n...h_1x$ và $v = h'_m...h'_1x$ là biểu diễn chính tắc của u và v đối với x. Nếu tồn tại chỉ số $k \leqslant \min(m, n)$ lớn nhất sao cho $u \approx_k v$ thì $u \neq_{k+1} v$.*

Hệ quả 3.2. (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 lý 3.3. *Xét $P^k = \{I(x) : x \in \boldsymbol{X}^k\}$ với $\boldsymbol{X}^k = \{x \in \boldsymbol{X} : |x| = k\}$, $u = h_n...h_1x$ và $v = h'_m...h'_1x$ là biểu diễn chính tắc của u và v đối với x. Nếu $u <_k v$ hoặc $u >_k v$ thì với* *mọi* $a \in H(u)$, *với mọi* $b \in H(v)$ *ta có* $a <_k b$ *hoặc* $a >_k b$.

3.2. Miền trị của thuộc tính trong quan hệ có chứa giá trị số

Trường hợp miền trị của thuộc tính có chứa giá trị số, chúng ta sẽ biến đổi các giá trị số thành các giá trị ngôn ngữ tương ứng theo một ngữ nghĩa xác định. Trước tiên, ta đi xây dựng một hàm IC chuyển một số về một giá trị thuộc $[0, 1]$ và hàm Φ_k để chuyển một giá trị trong $[0, 1]$ thành một giá trị ngôn ngữ x tương ứng trong đại số gia tử $\boldsymbol{X}$.

Định nghĩa 3.5. Cho $Dom(A_i) = Num(A_i) \cup LV(A_i)$, v là hàm định lượng ngữ nghĩa của A_i. Hàm $IC : Dom(A_i) \to [0, 1]$ được xác định như sau:

Nếu $LV(A_i) = \emptyset$ và $Num(A_i) \neq \emptyset$ thì $\forall \omega \in Dom(A_i)$ ta có $IC(\omega) = \dfrac{\omega - \psi_{\min}}{\psi_{\max} - \psi_{\min}}$ với $Dom(A_i) = [\psi_{\min}, \psi_{max}]$ là miền trị kinh điển của A_i.

Nếu $Num(A_i) \neq \emptyset$, $LV(A_i) \neq \emptyset$ thì $\forall \omega \in Dom(A_i)$ ta có $IC(\omega) = \{\omega^* v(\psi_{\max LV})\}/\psi_{\max}$, với $LV(A_i) = [\psi_{\min LV}, \psi_{\max LV}]$ là miền trị ngôn ngữ của A_i.

Ví dụ 3.5. Cho $Dom(Tuoi) = \{0...100, ...$ rất rất trẻ,......., rất rất già$\}$.

$Num(Tuoi) = \{20, 25, 27, 30, 45, 60, 75, 66, 80\}$.

$LV(Tuoi) = \{$trẻ, rất trẻ, già, khá trẻ, khá già, ít già, rất già, rất rất trẻ$\}$, $Dom(Tuoi) = Num(Tuoi) \cup LV(Tuoi)$.

Nếu $LV(Tuoi) = \emptyset$ khi đó $Dom(Tuoi) = Num(Tuoi) = \{20, 25, 27, 30, 45, 60, 75, 66, 80\}$.

Do đó $\forall \omega \in Dom(Tuoi)$, ta có $Dom(Tuoi) = \{$0,2, 0,25, 0,27, 0,3, 0,45, 0,6, 0,75, 0,66, 0,8$\}$.

Nếu $Num(A_i) \neq \emptyset$ và $LV(A_i) \neq \emptyset$ ta có $Dom(Tuoi) = Num(Tuoi) \cup LV(Tuoi) = \{$trẻ, rất trẻ, già, khá trẻ, khá già, ít giá, rất già, rất rất trẻ, 20, 25, 27, 30, 45, 60, 75, 66, 80$\}$. Giả sử tính được $v(\psi_{\max LV}) = v($rất rất già$) =$ 0,98. Khi đó $\forall \omega \in Num(A_i)$ ta có $IC(\omega) = \{\omega.v(\psi_{\max LV})\}/\psi_{\max} = (\omega \times 0,98)/100$, hay $\forall \omega \in Num(A_i)$ sử dụng $IC(\omega)$, ta có $Num(A_i) = \{$0,196, 0,245, 0,264, 0,294, 0,441, 0,588, 0,735, 0,646, 0,784$\}$.

Nếu ta chọn các tham số W và độ đo tính mờ cho các gia tử sao cho $v(\psi_{\max LV}) \approx 1,0$ thì $(\{\omega \times v(\psi_{\max LV})\}/\psi_{\max}) \approx 1 - \dfrac{\psi_{\max} - \omega}{\psi_{\max} - \psi_{\min}}$.

Định nghĩa 3.6. Cho đại số gia tử $\underline{X} = (\boldsymbol{X}, G, H, \leqslant)$, v là hàm định lượng ngữ nghĩa của $\boldsymbol{X}$. $\phi_k : [0, 1] \to \boldsymbol{X}$ gọi là hàm ngược của hàm v theo mức k được xác định:

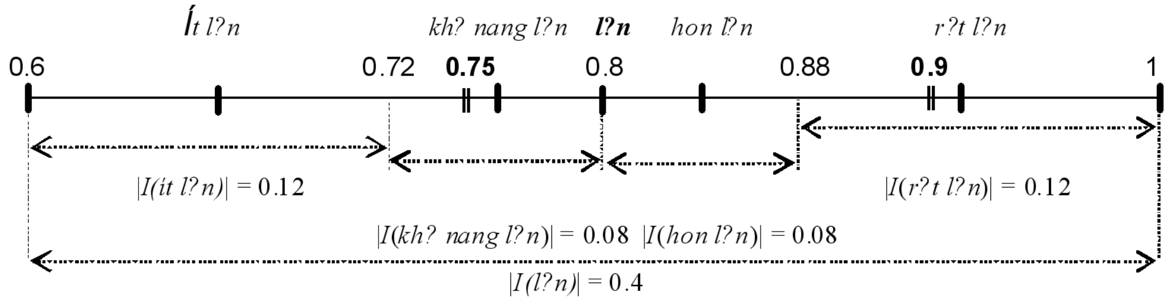
$$\forall a \in [0, 1],\ \Phi_k(a) = x^k \text{ khi và chỉ khi } a \in I(x^k), \quad \text{với } x^k \in \boldsymbol{X}^k.$$

Ví dụ 3.6. Cho đại số gia tử $\underline{X} = (\boldsymbol{X}, G, H, \leqslant)$, trong đó $H^+ = \{$*hơn, rất*$\}$ với *hơn* $<$ *rất* và $H^- = \{$*ít, khả năng*$\}$ với *ít* $>$ *khả năng*, $G = \{$*nhỏ, lớn*$\}$. Giả sử cho $W = 0,6$, *fm(hơn)* $= 0,2$, *fm(rất)* $= 0,3$, *fm(ít)* $= 0,3$, *fm(khả năng)* $= 0,2$.

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 $[0, 1]$. *fm(nhỏ)* $= 0,6$, *fm(lớn)* $=0,4$, *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 tôi chỉ xét các phần tử được sinh từ phần tử *lớn.*



Hình 3.1. Tính mờ của phần tử sinh lớn

Định lý 3.4. *Cho đại số gia tử* $\underline{X} = (\boldsymbol{X}, G, H, \leqslant)$, *v là hàm định lượng ngữ nghĩa của* $\boldsymbol{X}$, Φ_k *là hàm ngược của v, ta có*

(1) $\forall x^k \in \boldsymbol{X}^k,\ \Phi_k(v(x^k)) = x^k$

(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 $a = v(x^k) \in [0, 1]$. Vì $v(x^k) \in I(x^k)$ nên $a \in I(x^k)$. Theo định nghĩa ta có $\Phi_k(v(x^k)) = x^k$.

(2) Vì $x^k \neq_k y^k$ nên theo định nghĩa ta có $x^k <_k y^k$ hoặc $y^k <_k x^k$, suy ra $v(x^k) < v(y^k)$ hoặc $v(y^k) < v(x^k)$. Mặt khác ta có $v(x^k) \in I(x^k)$ và $v(y^k) \in I(y^k)$, theo giả thiết $a < b$ do đó $x^k <_k y^k$. Hay $\Phi_k(a) <_k \Phi_k(b)$.

3.3. Thuật toán xác định giá trị chân lý của điều kiện mờ

Như trong Mũc 3 đã trình bày, miền trị của thuộc tính mờ trong quan hệ của lược đồ cơ sở dữ liệu phức tạp và có thể nhận giá trị như số, giá trị ngôn ngữ hoặc vừa giá trị số vừa giá trị ngôn ngữ. Vì vậy, ta đi xây dựng thuật toán đánh giá điều kiện mờ để làm cơ sở cho việc thao tác và tìm kiếm dữ liệu sau này.

Gọi $Dom(A_i) = Num(A_i) \cup LV(A_i)$ là miền trị của thuộc tính mờ A_i trong một quan hệ của lược đồ cơ sở dữ liệu. Khi đó, thuật toán được xây dựng như sau.

Thuật toán 3.1

Vào: Cho r là một quan hệ xác định trên tập vũ trụ các thuộc tính U.

Điều kiện $t[A_i] \approx_k u$, với u là một giá trị số hoặc giá trị ngôn ngữ.

Ra: Với mọi $t \in r$ sao cho $(t[A_i] \approx_k u) = true$.

Phương pháp

// Đi xây dựng các $P^k = \{I(t[A_i]) : |t[A_i]| = k, \forall t \in r\}$, theo [2], một giới hạn hợp lý để phù hợp trong thực tế ta cho $k \leqslant 4$. Trước tiên, ta chuyển các giá trị số thành giá trị ngôn ngữ.

(1) for mỗi $t \in r$ do

(2) if $t[A_i] \in Num(A_i)$ then $t[A_i] = \Phi_k(IC(t[A_i]))$

//Xây dựng các P^k dựa vào độ dài các từ.

(3) $k = 1$

(4) While $k \leqslant 4$ do

(5) $P^k = \emptyset$

(6) for mỗi $t \in r$ do

(7) if $|t[A_i]| = k$ then $P^k = P^k \cup \{I(t[A_i])\}$

(8) $k = k + 1$

// Xác định giá trị chân lý của $(t[A_i] \approx_k u)$.

(9) if $u \in Num(A_i)$ then $u' = \Phi_k(IC(u))$

(10) $k = 4$ // Phân hoạch tương ứng với mức lớn nhất.

(11) While $k > 0$ do

(12) for mỗi $\Delta^k \in P^k$ do

(13) if $(I(t[A_i]) \subseteq \Delta^k$ and $I(u) \subseteq \Delta^k)$ or $(I(t[A_i]) \subseteq \Delta^k$ and $I(u') \subseteq \Delta^k)$ then $\{(t[A_i] \approx_k u) = true\}$ or $\{(t[A_i] \approx_k u') = true\}$

(14) exit

(15) $k = k - 1$

Thuật toán 3.2

Vào: Cho r là một quan hệ xác định trên tập vũ trụ các thuộc tính U.

Điều kiện $t[A_i]\theta u$, với u là một giá trị số hoặc giá trị ngôn ngữ, $\theta \in \{\neq_k, <_k, >_k\}$.

Ra: Với mọi $t \in r$ sao cho $(t[A_i]\theta u) = true$

Phương pháp

(1) Sử dụng các bước từ (1)-(8) trong Thuật toán 3.1

(2) if $u \in Num(A_i)$ then $u' = \Phi_k(IC(u))$

(3) $k = 1$

(4) While $k \leqslant 4$ do

(5) for với mọi $\Delta^k \in P^k$ do

(6) if $\{I(t[A_i]) \not\subset \Delta^k$ or $I(u) \not\subset \Delta^k\}$ then $(t[A_i] \neq_k u) = true$

(7) if $\{v(t[A_i]) > v(u)\}$ then $(t[A_i] >_k u) = true$

(8) else if $(t[A_i] <_k u) = true$

(7) if $\{I(t[A_i]) \not\subset \Delta^k$ or $I(u') \not\subset \Delta^k\}$ then $(t[A_i] \neq_k u') = true$

(9) if $\{v(t[A_i]) > v(u')\}$ then or $(t[A_i] >_k u') = true$

(10) else if $(t[A_i] <_k u') = true$

(11) $k = k + 1$

3.4. Ví dụ. Cho lược đồ quan hệ $U = \{SOCM, HOTEN, SUCKHOE, TUOI, LUONG\}$ và quan hệ *Luong_Tuoi* được xác định như sau:

Bảng 3.1. Quan hệ Lương tuổi

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	hơn xấu	45	500.000
55555	An Thuyên	rất xấu	41	rất cao
66666	Thuận Yến	khả năng xấu	61	thấp
77777	Văn Cao	hơn tốt	59	ít cao
88888	Thanh Tùng	khả năng tốt	75	1.500.000
99999	Nguyễn Cường	ít tốt	25	khá thấp

(a) Tìm những cán bộ có $TUOI \approx_2$ *hơn già* và $SUCKHOE \approx_2$ *khả năng tốt.*

(b) Tìm những cán bộ có $TUOI \approx_1$ *trẻ* hoặc có $LUONG \neq_1$ *cao.*

Trước hết chúng ta sẽ xem miền trị của SUCKHOE, TUOI và LUONG là ba đại số gia tử và được xác định như sau:

$\underline{X}_{Suckhoe} = \boldsymbol{X}_{suckhoe}, G_{suckhoe}, H_{suckhoe}, \leqslant)$, với $G_{suckhoe} = \{$*tốt, xấu*$\}, H^+_{suckhoe} = \{$*rất, hơn*$\}, H^-_{suckhoe} = \{$*khả năng, ít*$\}$, *rất* $>$ *hơn* và *ít* $>$ *khả năng.*

$W_{suckhoe} = 0,6$, fm*(xấu)* $= 0,6$, fm*(tốt)* $= 0,4$, fm*(rất)* $= 0,3$, fm*(khá)* $= 0,2$, fm*(khả năng)* $= 0,2$, fm*(ít)* $= 0,3$.

$\underline{X}_{Tuoi} = (\boldsymbol{X}_{tuoi}, G_{tuoi}, H_{tuoi}, \leqslant)$, với $G_{tuoi} = \{$*trẻ, già*$\}$, $H_t^+uoi = \{$*rất, hơn*$\}$, $H^-_{tuoi} = \{$*khả năng, ít*$\}$, *rất* $>$ *hơn* và *ít* $>$ *khả năng.* $W_{tuoi} = 0,4$, fm*(trẻ)* $= 0,.4$, fm*(già)* $= 0,6$, fm*(rất)* $= 0,3$, fm*(khá)* $=0,15$, fm*(khả năng)* $= 0,25$, fm*(ít)* $= 0,3$.

$\underline{X}_{Luong} = (\boldsymbol{X}_{luong}, G_{luong}, H_{luong}, \leqslant)$, với $G_{luong} = \{$*cao, thấp*$\}$, $H^+_{luong} = \{$*rất, hơn*$\}$,

$H^-_{luong} = \{$*khả năng, ít*$\}$, *rất* $>$ *hơn* và *ít* $>$ *khả năng.* $W_{luong} = 0,6$, fm*(thấp)* $= 0,6$, fm *(cao)* $= 0,4$, fm*(rất)* $= 0,25$, fm*(khá)* $= 0,25$, fm*(khả năng)* $= 0,25$, fm*(ít)* $= 0,25$.

Đối với thuộc tính TUOI: 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)* $= [0,3,\ 0,4]$.

Ta có fm*(rất già)* $= 0,18$, fm*(hơn già)* $= 0,09$, fm*(ít già)* $= 0,18$, fm*(khả năng già)* $= 0,15$.

Vì *ít già* $<$ *khả năng già* $<$ *già* $<$ *hơn già* $<$ *rất già* nên I*(ít già)* $= [0,4,\ 0,58]$, I*(khả năng già)* $= [0,58,\ 0,73]$, *I(hơn già)* $= [0,73,\ 0,82]$, I*(rất già)* $= [0,82,\ 1]$.

Nếu chọn $\psi_1 = 100 \in \boldsymbol{X}_{tuoi}$ khi đó $\forall \omega \in Num(TUOI)$, sử dụng $IC(\omega)$ ta có $Num(TUOI) = \{0,31,\ 0,85,\ 0,32,\ 0,45, 0,41,\ 0,61,\ 0,59,\ 0,75,\ 0,25\}$.

Do đó $\Phi_2(0,31) =$ *ít trẻ* vì $0,31 \in I$*(ít trẻ)*, tương tự $\Phi_2(0,85) =$ *rất già,* $\Phi_2(0,32) =$ *ít trẻ,* $\Phi_2(0,45) =$ *ít già,* $\Phi_2(0,41) =$ *ít già,* $\Phi_2(0,61) =$ *khả năng già,* $\Phi_2(0,59) =$ *khả năng già,* $\Phi_2(0,75) =$ *hơn già,* $\Phi_2(0,25) =$ *khả năng trẻ.*

Đối với thuộc tính LUONG: Ta cófm*(rất thấp)* $= 0,15$, fm*(khá 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 $fm(\textit{rất cao}) = 0,1,\ fm(\textit{hơn cao}) = 0,1,\ fm(\textit{ít cao}) = 0,1,\ fm(\textit{khả năng cao}) = 0,1$.

Vì *ít cao* $<$ *khả năng cao* $<$ *cao* $<$ *hơn cao* $<$ *rất cao* nên $I(\textit{ít cao}) = [0,6,\ 0,7]$, $I(\textit{khả năng cao}) = [0,7,\ 0,8]$, $I(\textit{hơn cao}) = [0,8,\ 0,9]$, $I(\textit{rất cao}) = [0,9,\ 1]$.

Nếu chọn ψ_2 =*rất rất cao* $\in \boldsymbol{X}_{luong}$ và $\psi_1 = 3.000.000$, ta có $v(\textit{rất rất cao}) = 0,985$ khi đó $\forall \omega \in Num(LUONG) = \{2.800.000, 2.000.000, 500.000, 1.500.000\}$, sử dụng $IC(\omega) = \{\omega \times v(\psi_2)\}/\psi_1$, ta có $Num(LUONG) = \{0,92,\ 0,65,\ 0,16,\ 0,49\}$.

Do đó $\Phi_2(0,92)=$ *rất cao,* $\Phi_2(0,65) =$ *ít cao,* $\Phi_2(0,16) =$ *hơn thấp,* $\Phi_2(0,49) =$ *ít cao.*

Vậy, những cán bộ có $TUOI \approx_2$ *hơn già* và $SUCKHOE \approx_2$ *khả năng tốt* là:

Bảng 3.2. Kết quả tìm kiếm của ví dụ (a)

Socm	Hoten	Suckhoe	Tuoi	Luong
88888	Thanh Tùng	khả năng tốt	75	1.500.000

và những cán bộ có $TUOI \approx_1$ trẻ hoặc có $LUONG \neq_1$cao.

Bảng 3.2. Kết quả tìm kiếm của ví dụ (b)

Socm	Hoten	Suckhoe	Tuoi	Luong
11111	Phạm Trọng Cầu	rất rất tốt	31	2.800.000
33333	Trần Tiến	xấu	32	2.000.000
44444	Vũ Hoàng	khá xấu	45	500.000
66666	Thuận Yến	khả năng xấu	61	thấp
99999	Nguyễn Cường	ít tốt	25	khá thấp

4. KẾT LUẬN

Bài báo xem xét một cách trọn vẹn việc đánh giá để đối sánh các giá trị khi miền trị thuộc tính của một quan hệ trong cơ sở dữ liệu mờ nhận giá trị đa dạng.Việc đánh giá này là phù hợp với thực tế, bởi vì giá trị của ngôn ngữ là tương đối phức tạp. Trên cơ sở này, bài báo đã phân tích các quan hệ đối sánh giữa hai giá trị theo ngữ nghĩa mới. Từ đó đưa ra một số ví dụ về các thao tác dữ liệu theo cách tiếp cận mới. Vấn đề xây dựng các phụ thuộc dữ liệu trên mô hình cơ sở dữ liệu mờ theo cách tiếp cận đại số gia tử sẽ được giới thiệu trong những bài báo tiếp theo.

TÀI LIỆU THAM KHẢO

[1] B. P. Buckles, F. E. Petry, A fuzzy representation of data for relational databases, *Fuzzy Sets and Systems* **7** (3) (1982) 213–226.

[2] Hồ Thuần, Hồ Cẩm Hà, An approach to extending the relational database model for handing incomplete information and data dependencies, *Tạp chí Tin học và Điều khiển học* **17** (3) (2001) 41–47.

[3] Hồ Thuần, Hồ Cẩm Hà, Đại số quan hệ và quan điểm sử dụng Null value trên một mô hình cơ sở dữ liệu mờ, *Tạp chí Tin học và Điều khiển học* **17** (4) (2001) 1–10.

[4] H. Thuan, T. T. Thanh, Fuzzy Functional Dependencies with Linguistic Quantifiers, *Tạp chí Tin học và Điều khiển học* **18** (2) (2002) 97–108.

[5] Mustafa LLKer Sozat, Adnan Yazici, A complete axiomatization for fuzzy functional and multivalued dependencies in fuzzy database relations, *Fuzzy Set and Systems* **117** (2001) 161–181.

[6] Nguyễn Cát Hồ, Trần Thái Sơn, Về khoảng cách giữa các giá trị của biến ngôn ngữ trong đại số gia tử, *Tạp chí Tin học và Điều khiển học* **11** (1) (1995) 10–20.

[7] Nguyễn Cát Hồ, Trần Thái Sơn, Trần Đình Khang, Lê Xuân Việt, Fuzziness measure, quantified semantic mapping and interpolative method of approximate reasoning in medical expert systems, *Tạp chí Tin học và Điều khiển học* **18** (3) (2002) 237–252.

[8] Nguyen Cat Ho, W. Wechler, Extended hedge algebras ans their application to fuzzy logic, *Fuzzy Set and Systems* **52** (1992) 259–282.

[9] Nguyễn Cát Hồ, Lý thuyết tập mờ và công nghệ tính toán mềm, *Hệ mờ, mạng nơron và ứng dụng*, Nhà xuất bản Khoa học và Kỹ thuật, năm 2001 (37–74).

[10] Le Tien Vuong, Ho Thuan, A relational database extended by application of fuzzy set theory and linguistic variables, *Computer and Artificial Intelligence* **8** (2) (1989) 153–168.

[11] E. Petry and P. Bosc, *Fuzzy Databases Principles and Applications*, Kluwer Academic Publishers, 1996.

[12] S. Shensoi, A. Melton, Proximity relations in the fuzzy relational databases, *Fuzzy Sets and Systems* **21** (1987) 19–34.

Nhận bài ngày 6 - 1 - 2006