

**MỘT PHƯƠNG PHÁP XỬ LÝ TRUY VẤN CON TRONG CƠ SỞ DỮ LIỆU MỜ  
THEO CÁCH TIẾP CẬN ĐẠI SỐ GIA TỬ**

*Nguyễn Công Hòa  
Đại học Huế*

**TÓM TẮT**

*Trong bài báo này, chúng tôi giới thiệu ngôn ngữ truy vấn con để thao tác dữ liệu trong mô hình cơ sở dữ liệu mờ theo cách tiếp cận đại số gia tử. Ngôn ngữ thao tác dữ liệu được đề xuất phù hợp với mô hình cơ sở dữ liệu mờ theo cách tiếp cận mới. Các phương pháp biến đổi truy vấn con thành truy vấn tương ứng cũng được đề xuất trong bài báo này.*

**1. Mở đầu**

Với những ưu điểm của cấu trúc đại số gia tử (ĐSGT), các tác giả đã xây dựng mô hình cơ sở dữ liệu (CSDL) mờ dựa trên cách tiếp cận của đại số gia tử và ngôn ngữ để truy vấn dữ liệu trên mô hình đó [1-4]. Trong đó, ngữ nghĩa ngôn ngữ được lượng hóa bằng các ánh xạ định lượng của ĐSGT. Theo cách tiếp cận này, ngữ nghĩa ngôn ngữ có thể biểu thị bằng một lân cận các khoảng được xác định bởi độ đo tính mờ của các giá trị ngôn ngữ của một thuộc tính với vai trò là biến ngôn ngữ. Truy vấn con là một dạng truy vấn thường gặp trong việc xử lý, tìm kiếm dữ liệu trong mô hình CSDL và đã có một số công trình nghiên cứu vấn đề này theo cách tiếp cận lý thuyết tập mờ [6-8] nhưng còn nhiều hạn chế. Tuy nhiên, nghiên cứu dạng truy vấn này đối với cách tiếp cận ĐSGT là vấn đề mới. Vì vậy, nội dung bài báo tập trung nghiên cứu dạng truy vấn này và ứng dụng của nó.

Trước tiên, một số khái niệm cơ bản về ĐSGT và CSDL mờ sẽ được trình bày ngắn gọn mục 2. Trong mục 3, sẽ trình bày một cách xử lý truy vấn con trong CSDL mờ theo cách tiếp cận ĐSGT.

**2. Một số khái niệm cơ sở**

Cho một ĐSGT tuyến tính đầy đủ  $\mathcal{A}X = (X, G, H, \Sigma, \Phi, \leq)$ , trong đó  $Dom(X) = X$  là miền các giá trị ngôn ngữ của thuộc tính ngôn ngữ  $X$  được sinh tự do từ tập các phần tử sinh  $G = \{1, c^+, W, c^-, 0\}$  bằng việc tác động tự do các phép toán một ngôi trong tập  $H$ ,  $\Sigma$  và  $\Phi$  là hai phép tính với ngữ nghĩa là cận trên đúng và cận dưới đúng của tập  $H(x)$ , tức là  $\Sigma x = \sup H(x)$  and  $\Phi x = \inf H(x)$ , trong đó  $H(x)$  là tập các phần tử sinh ra từ  $x$ , còn quan hệ  $\leq$  là quan hệ sắp thứ tự tuyến tính trên  $X$  cảm sinh từ ngữ nghĩa của ngôn ngữ. Ví dụ, nếu ta có thuộc tính Lương là “Tổng lương của cán bộ trong

một tháng nào đó”, thì  $Dom(Luong) = \{high, low, very high, more high, possibly high, very low, possibly low, less low, \dots\}$ ,  $\mathbf{G} = \{1, high, W, low, 0\}$ ,  $\mathbf{H} = \{very, more, possibly, less\}$  và  $\leq$  một quan hệ thứ tự cảm sinh từ ngữ nghĩa của các từ trong  $Dom(Luong)$ , chẳng hạn ta có  $very high > high$ ,  $more high > high$ ,  $possibly high < high$ ,  $less high < high$ , ...

Cho tập các gia tử  $\mathbf{H} = \mathbf{H}^- \cup \mathbf{H}^+$ , trong đó  $\mathbf{H}^+ = \{h_1, \dots, h_p\}$  và  $\mathbf{H}^- = \{h_{-1}, \dots, h_{-q}\}$ , với  $h_1 < \dots < h_p$  và  $h_{-1} < \dots < h_{-q}$ , trong đó  $p, q > 1$ .

Ký hiệu  $fm: X \rightarrow [0, 1]$  là độ đo tính mờ của ĐSGT  $\mathcal{AX}$ . Khi đó,

**Định nghĩa 2.1.** Với mỗi  $x \in X$ , 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  $h \in H$ .

**Mệnh đề 2.1.** Độ đo tính mờ  $fm$  và độ đo tính mờ của gia tử  $\mu(h)$ ,  $\forall h \in \mathbf{H}$ , có các tính chất sau:

- (1)  $fm(hx) = \mu(h)fm(x)$ ,  $\forall x \in X$
- (2)  $fm(c^-) + fm(c^+) = 1$
- (3)  $\sum_{-q \leq i \leq p, i \neq 0} fm(h_i c) = fm(c)$ , trong đó  $c \in \{c^-, c^+\}$
- (4)  $\sum_{-q \leq i \leq p, i \neq 0} fm(h_i x) = fm(x)$ ,  $x \in X$
- (5)  $\sum \{\mu(h_i) : -q \leq i \leq -1\} = \alpha$  và  $\sum \{\mu(h_i) : 1 \leq i \leq p\} = \beta$ , trong đó  $\alpha, \beta > 0$  và  $\alpha + \beta = 1$ .

### 2.1. Khoảng mờ

Giả sử thuộc tính  $X$  có miền tham chiếu thực là khoảng  $[a, b]$ . Để chuẩn hóa, nhờ một phép biến đổi tuyến tính, ta giả thiết mọi miền như vậy đều là khoảng  $[0, 1]$ . Khi đó, tính chất (2) trong mệnh đề 2.1 cho phép ta xây dựng hai khoảng mờ của hai khái niệm nguyên thủy  $c^-$  và  $c^+$ , ký hiệu là  $I(c^-)$  và  $I(c^+)$  với độ dài tương ứng là  $fm(c^-)$  và  $fm(c^+)$  sao cho chúng tạo thành một phân hoạch của miền tham chiếu  $[0, 1]$  và  $I(c^-)$  và  $I(c^+)$  là đồng biến với  $c^-$  và  $c^+$ , tức là  $c^- \leq c^+$  kéo theo  $I(c^-) \leq I(c^+)$ .

**Định nghĩa 2.2.** (hàm PN-dấu  $Sgn$ ):  $Sgn : X \rightarrow \{-1, 0, 1\}$  là hàm dấu được xác định như sau, ở đây  $h, h' \in \mathbf{H}$ , và  $c \in \{c^-, c^+\}$ :

- (1)  $Sgn(c^-) = -1$ ,  $Sgn(c^+) = +1$
- (2)  $Sgn(h'hx) = 0$ , nếu  $h'hx = hx$ , còn ngược lại ta có

$Sgn(h'hx) = -Sgn(hx)$ , nếu  $h'hx \neq hx$  và  $h'$  là âm tính đối với  $h$  (hoặc  $c$ , nếu  $h = I$  và  $x = c$ )

$Sgn(h'hx) = +Sgn(hx)$ , nếu  $h'hx \neq hx$  và  $h'$  dương tính đối với  $h$  (hoặc  $c$ , nếu  $h = I$  và  $x = c$ ).

**Định nghĩa 2.3.** Giả sử  $\mathcal{AX} = (X, G, H, \Sigma, \Phi, \leq)$  là một ĐSGT đầy đủ, tuyến tính và tự do,  $fm(x)$  và  $\mu(h)$  tương ứng là các độ đo tính mờ của ngôn ngữ và của gia tử  $h$  thỏa mãn các tính chất trong mệnh đề 2.1. Khi đó, ta nói  $\nu$  là ánh xạ cảm sinh bởi độ đo tính mờ  $fm$  của ngôn ngữ nếu nó được xác định như sau:

$$(1) \quad \nu(W) = \kappa = fm(c^-), \nu(c^-) = \kappa - \alpha fm(c^-) = \beta fm(c^-), \nu(c^+) = \kappa + \alpha fm(c^+)$$

$$(2) \quad \nu(h_j x) = \nu(x) + Sgn(h_j x) \left\{ \sum_{i=Sgn(j)}^j \mu(h_i) fm(x) - \omega(h_j x) \mu(h_j) fm(x) \right\},$$

$$\text{trong đó} \quad \omega(h_j x) = \frac{1}{2} [1 + Sgn(h_j x) Sgn(h_p h_j x) (\beta - \alpha)] \in \{\alpha, \beta\}, \text{ với mọi } j, -q \leq j \leq$$

$p$  và  $j \neq 0$

(3)  $\nu(\Phi c^-) = 0, \nu(\Sigma c^-) = \kappa = \nu(\Phi c^+), \nu(\Sigma c^+) = 1$ , và với mọi  $j, -q \leq j \leq p$  và  $j \neq 0$ , chúng ta có:  $\nu(\Phi h_j x) = \nu(x) + Sgn(h_j x) \left\{ \sum_{i=sign(j)}^{j-1} \mu(h_i) fm(x) \right\}$  và

$$\nu(\Sigma h_j x) = \nu(x) + Sgn(h_j x) \left\{ \sum_{i=sign(j)}^j \mu(h_i) fm(x) \right\}.$$

## 2.2. Độ tương tự mức $k$

Xét  $X_k$  là tập tất cả các phần tử 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  $I(h_p c^-) \leq I(h_{p-1} c^-) \leq \dots \leq I(h_2 c^-) \leq I(h_1 c^-) \leq \nu_A(c^-) \leq I(h_{-1} c^-) \leq I(h_{-2} c^-) \leq \dots \leq I(h_{-q+1} c^-) \leq I(h_{-q} c^-)$ . Khi đó, ta xây dựng phân hoạch về độ tương tự mức 1 gồm các lớp tương đương sau:  $S(\theta) = I(h_p c^-); S(c^-) = I(c^-) \setminus [I(h_{-q} c^-) \cup I(h_p c^-)]; S(W) = I(h_{-q} c^-) \cup I(h_{-q} c^+)$ ; tương tự ta có  $S(c^+) = I(c^+) \setminus [I(h_{-q} c^+) \cup I(h_p c^+)]$  và  $S(I) = I(h_p c^+)$ .

Ta thấy, trừ hai điểm đầu mút  $\nu_A(\theta) = 0$  và  $\nu_A(I) = 1$ , các giá trị đại diện  $\nu_A(c^-)$ ,  $\nu_A(W)$  và  $\nu_A(c^+)$  đều là điểm trong tương ứng của các lớp tương tự mức 1  $S(c^-)$ ,  $S(W)$  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,  $I(h_i c^+) = (\nu_A(\Phi h_i c^+), \nu_A(\Sigma h_i c^+))$  với hai khoảng mờ kề là  $I(h_{i-1} c^+)$  và  $I(h_{i+1} c^+)$  chúng ta sẽ có các lớp tương đương dạng sau:  $S(h_i c^+) = I(h_i c^+) \setminus [I(h_p h_i c^+) \cup I(h_{-q} h_i c^+)]$ ,  $S(\Phi h_i c^+) = I(h_{-q} h_{i-1} c^+) \cup I(h_{-q} h_i c^+)$  và  $S(\Sigma h_i c^+) = I(h_p h_i c^+) \cup I(h_p h_i c^+)$ , với  $i$  sao cho  $-q \leq i \leq p$  và  $i \neq 0$ .

Giả sử phân hoạch các lớp tương tự mức  $k$  là các khoảng  $S(x_1), S(x_2), \dots, S(x_m)$ . Khi đó, mỗi giá trị mờ  $u$  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  $\Omega_k(u)$ .

### 2.3. Cơ sở dữ liệu mờ

Xét một lược đồ CSDL  $F_{DB} = \{U, R_1, R_2, \dots, R_m\}$ , trong đó  $U = \{A_1, A_2, \dots, A_n\}$  là tập vũ trụ các thuộc tính,  $R_i$  lược đồ quan hệ, tức là một tập con của  $U$ . Mỗi thuộc tính  $A$  được gắn với một miền giá trị thuộc tính, trong đó một số thuộc tính cho phép nhận các giá trị ngôn ngữ trong lưu trữ trong CSDL và được gọi là *thuộc tính mờ*. Những thuộc tính còn lại được gọi là *thuộc tính kinh điển*. Thuộc tính kinh điển  $A$  được gắn với một miền giá trị kinh điển, ký hiệu là  $D_A$ . Thuộc tính mờ  $A$  sẽ được gắn một miền giá trị kinh điển  $D_A$  và một miền giá trị ngôn ngữ  $LD_A$  hay là tập các phần tử của một ĐSGT. Một CSDL như vậy được gọi là CSDL mờ theo cách tiếp cận đại số gia tử.

### 2.4. Các quan hệ đối sánh trên miền trị thuộc tính

**Định nghĩa 2.4.** Giả sử  $t$  và  $s$  là hai bộ dữ liệu trên tập vũ trụ  $U$  các thuộc tính. Ta nói  $t[A_i] =_k s[A_i]$  và gọi là chúng bằng nhau mức  $k$ , nếu một trong các điều kiện sau xảy ra:

Nếu  $t[A_i], s[A_i] \in D_{A_i}$  thì  $t[A_i] = s[A_i]$  hoặc là

Nếu một trong hai giá trị  $t[A_i], s[A_i]$  là khái niệm mờ, chẳng hạn đó là  $t[A_i]$ , thì ta phải có  $s[A_i] \in \Omega_k(t[A_i])$  hoặc là

Nếu cả hai giá trị  $t[A_i], s[A_i]$  đều là giá trị mờ, thì  $\Omega_k(t[A_i]) = \Omega_k(s[A_i])$ .

**Định nghĩa 2.5.** Giả sử  $t$  và  $s$  là hai bộ dữ liệu trên tập vũ trụ  $U$  các thuộc tính. Khi đó:

Ta viết  $t[A_i] \leq_k s[A_i]$ , nếu hoặc là  $t[A_i] =_k s[A_i]$  hoặc là  $\Omega_k(t[A_i]) < \Omega_k(s[A_i])$

Ta viết  $t[A_i] <_k s[A_i]$ , nếu  $\Omega_k(t[A_i]) < \Omega_k(s[A_i])$

Ta viết  $t[A_i] >_k s[A_i]$ , nếu  $\Omega_k(t[A_i]) > \Omega_k(s[A_i])$ .

## 3. Truy vấn con trong truy vấn mờ

Một truy vấn con trong truy vấn mờ là một câu lệnh *select...from...where...* mà nó được lồng trong một truy vấn mờ khác. Thay vì cách biểu diễn truy vấn dưới dạng một khối thực hiện trên nhiều quan hệ, có thể biểu diễn truy vấn này dưới dạng nhiều khối lồng nhau, mỗi khối thực hiện đúng trên một quan hệ. Trong phần này, bài báo tập trung phân tích một số phép toán thường sử dụng trong truy vấn con và việc biến đổi tương đương một câu truy vấn con thành một câu truy vấn mờ tương ứng.

### 3.1. Toán tử In

Trong truy vấn mờ, toán tử “In” sử dụng để kiểm tra giá trị của thuộc tính trong bộ hiện hành thuộc về tập giá trị được trả lại bởi truy vấn con. Câu lệnh truy vấn con có dạng:

*Truy vấn Q3.1*

*select < danh sách các trường > from <P<sub>1</sub>>*

*where <fc<sub>1</sub>> and A<sub>1</sub> In ( select A<sub>2</sub> from <P<sub>2</sub>> where <fc<sub>2</sub>>)*

trong đó <fc<sub>1</sub>> và <fc<sub>2</sub>> là các điều kiện mờ

**Ví dụ 3.1.** Cho hai quan hệ *dssvien*(MA#, TENS<sub>V</sub>, QUEQUAN, HOCBONG, MAKHOA), *dskhoa*(#MAKHOA, TENKHOA, SOSVIEN).

Tìm những sinh viên (ma#, tensv, quequan) có *học bổng cao* và học ở khoa có *số sinh viên ít*. (giả sử ta chọn bằng nhau theo mức  $k = 1$ ).

(1) *select MA#, TENS<sub>V</sub>, QUEQUAN from dssvien where HOCBONG =<sub>1</sub> cao and MAKHOA In (select #MAKHOA from dskhoa where SOSVIEN =<sub>1</sub> ít).*

Từ truy vấn Q3.1 chúng ta có thể biến đổi thành câu truy vấn Q'3.1 như sau:

*Truy vấn Q'3.1*

*select < danh sách các trường > from <P<sub>1</sub>>, <P<sub>2</sub>>*

*where <fc<sub>1</sub>> and (A<sub>1</sub> = A<sub>2</sub>) and <fc<sub>2</sub>>*

(1') *select MA#, TENS<sub>V</sub>, QUEQUAN from dssvien, dskhoa where (HOCBONG =<sub>1</sub> cao and MAKHOA = #MAKHOA and SOSVIEN =<sub>1</sub> ít).*

**Bổ đề 3.1.** Kết quả câu truy vấn Q3.1 là tương đương với kết quả câu truy vấn Q'3.1 trong CSDL mờ.

Chứng minh:

*Trường hợp 1:* Nếu với  $\forall t \in P_1$  sao cho  $\{t[A_1] \text{ In } (select A_2 \text{ from } \langle P_2 \rangle \text{ where } \langle fc_2 \rangle)\}$  là sai, khi đó câu lệnh *select < danh sách các trường > from <P<sub>1</sub>> where <fc<sub>1</sub>> and A<sub>1</sub> In ( select A<sub>2</sub> from <P<sub>2</sub>> where <fc<sub>2</sub>>)* trong truy vấn **Q3.1** cho kết quả là quan hệ rỗng.

Đối với truy vấn **Q'3.1**, trước hết thực hiện phép tích Decac  $P_1$  và  $P_2$  ( $P_1 \times P_2$ ). Vì với  $\forall t_1 \in P_1$  và với  $\forall t_2 \in P_2$  mà  $t_1[A_1] \neq t_2[A_2]$  nên điều kiện <fc<sub>1</sub>> and (A<sub>1</sub> = A<sub>2</sub>) and <fc<sub>2</sub>> không thoả mãn. Hay, với  $\forall t_3 \in P_1 \times P_2$ , ta có  $t_3$  không thoả mãn điều kiện <fc<sub>1</sub>> and (A<sub>1</sub> = A<sub>2</sub>) and <fc<sub>2</sub>>. Do đó, truy vấn **Q'3.1** cho kết quả là quan hệ rỗng.

*Trường hợp 2:* Nếu  $\exists t \in P_1$  sao cho  $\{t[A_1] \text{ In } (select A_2 \text{ from } \langle P_2 \rangle \text{ where } \langle fc_2 \rangle)\}$  là đúng, khi đó câu lệnh *select < danh sách các trường > from <P<sub>1</sub>> where <fc<sub>1</sub>> and A<sub>1</sub> In ( select A<sub>2</sub> from <P<sub>2</sub>> where <fc<sub>2</sub>>)* trong truy vấn **Q3.1** cho kết quả là quan hệ gồm các bộ thoả mãn <fc<sub>1</sub>>. Đối với truy vấn **Q'3.1**, trước hết thực hiện phép nối  $P_1$  và  $P_2$  ( $P_1 \bowtie P_2$ ). Vì  $\exists t \in P_1$  sao cho  $\{t[A_1] \text{ In } (select A_2 \text{ from } \langle P_2 \rangle \text{ where } \langle fc_2 \rangle)\}$  là đúng nên  $t$  thoả mãn điều kiện (A<sub>1</sub> = A<sub>2</sub>) and <fc<sub>2</sub>>. Do đó câu lệnh *select < danh sách các trường > from <P<sub>1</sub>>, <P<sub>2</sub>> where <fc<sub>1</sub>> and (A<sub>1</sub> = A<sub>2</sub>) and <fc<sub>2</sub>>* trong truy vấn **Q'3.1** cho kết quả

là quan hệ gồm các bộ thỏa mãn  $\langle fc_1 \rangle$ . Vậy, câu truy vấn Q3.1 là tương đương với câu truy vấn Q'3.1 ■

### 3.2. Toán tử $In_f$

Đối với toán tử “ $In$ ” như trong 3.1, khi truy vấn dữ liệu thì sử dụng quan hệ bằng nhau trên thuộc tính rõ còn các điều kiện là mờ. Tuy nhiên, trong nhiều trường hợp truy vấn dữ liệu, người sử dụng mong muốn thay thế quan hệ bằng nhau bởi quan hệ “xấp xỉ” và gọi toán tử trong trường hợp này là “ $In_f$ ”. Khi đó, câu lệnh truy vấn con có dạng:

*Truy vấn Q3.2*

```
select < danh sách các trường > from < P1 >  
where < fc1 > and A1 Inf( select A2 from < P2 > where < fc2 >)
```

**Ví dụ 3.2.** Sử dụng hai quan hệ *dssvien* và *dskhoa* ở ví dụ 3.1.

Tìm những sinh viên (MA#, TENS<sub>V</sub>) có học bổng “xấp xỉ” học bổng của sinh viên có quê quán ở Huế và học ở khoa có số sinh viên rất nhiều.

(1) *select* MA#, TENS<sub>V</sub> *from* *dssvien* *where* HOCBONG  $In_f$ (*select* HOCBONG *from* *dssvien* *where* QUEQUAN = ”Huế” *and* MAKHOA  $In$  (*select* #MAKHOA *from* *dskhoa* *where* SOSVIEN =<sub>1</sub> rất nhiều)).

Từ truy vấn Q3.2 có thể biến đổi thành câu truy vấn Q'3.2 như sau:

*Truy vấn Q'3.2*

```
select < danh sách các trường > from < P1 >, < P2 >  
where < fc1 > and (A1 =k A2) and < fc2 >
```

(1') *select* MA#, TENS<sub>V</sub> *from* *dssvien*, *dskhoa* *where* (QUEQUAN = ”Huế”) *and* (*dssvien*.HOCBONG =<sub>1</sub> *dssvien*.HOCBONG ) *and* (SOSVIEN =<sub>1</sub> rất nhiều).

Trong đó, *dssvien*.HOCBONG là thuộc tính HOCBONG của quan hệ *dssvien*.

**Bổ đề 3.2.** Kết quả câu truy vấn Q3.2 là tương đương với kết quả câu truy vấn Q'3.2 trong CSDL mờ.

Chứng minh: Tương tự như bổ đề 3.1

### 3.3. Toán tử “any” và toán tử “all”

Trong câu lệnh SQL, khi truy vấn dữ liệu chúng ta mong muốn có được “một vài” hoặc “tất cả” bộ dữ liệu từ các quan hệ để so sánh với giá trị của một thuộc tính nào đó, nghĩa là câu lệnh có dạng:  $A_i \theta \lambda \langle SQ \rangle$ , trong đó  $A_i$  là thuộc tính,  $\theta$  là một phép so sánh,  $\lambda$  là toán tử “any” hoặc “all”,  $\langle SQ \rangle$  là truy vấn con.

### 3.3.1. Khi $\lambda$ là “any”

Câu lệnh truy vấn có dạng:

*Truy vấn Q3.3*

*select < danh sách các trường > from <P<sub>1</sub>>*

*where <fc<sub>1</sub>> and A<sub>1</sub>  $\theta$  any ( select A<sub>2</sub> from <P<sub>2</sub>> where <fc<sub>2</sub>>)*

Từ truy vấn Q3.3 chúng ta có thể biến đổi thành câu truy vấn Q’3.3 như sau:

*Truy vấn Q’3.3*

*select < danh sách các trường > from <P<sub>1</sub>>, <P<sub>2</sub>>*

*where <fc<sub>1</sub>> and {  $\exists t \in P_2: t$  thoả mãn điều kiện <fc<sub>2</sub>> and (A<sub>1</sub>  $\theta$  t[A<sub>2</sub>]) }*

**Ví dụ 3.3.** Cho hai quan hệ *nkban\_thang01*(MAHANG, TENHANG, SOLUONG, THANHTIEN) và *nkban\_thang02*(MAHANG, TENHANG, SOLUONG, THANHTIEN)

Tìm mặt hàng (MAHANG, TENHANG) bán trong tháng 01 có THANHTIEN “lớn hơn” một vài mặt hàng bán trong tháng 02 mà có *số lượng bán khả năng nhiều*.

*select MAHANG, TENHANG from nkban\_thang01 where THANHTIEN ><sub>1</sub> any (select THANHTIEN from nkban\_thang02 where nkban\_thang02.SOLUONG =<sub>1</sub> khả năng nhiều).*

Tương tự như ở câu truy vấn Q3.1, ta có bổ đề:

**Bổ đề 3.3.** Kết quả câu truy vấn Q3.3 là tương đương với kết quả câu truy vấn Q’3.3 trong CSDL mờ.

Chứng minh: Tương tự như bổ đề 3.1

### 3.3.2. Khi $\lambda$ là “all”

Câu lệnh truy vấn có dạng:

*Truy vấn Q4.4*

*select < danh sách các trường > from <P<sub>1</sub>>*

*where <fc<sub>1</sub>> and A<sub>1</sub>  $\theta$  all ( select A<sub>2</sub> from <P<sub>2</sub>> where <fc<sub>2</sub>>)*

Từ truy vấn Q3.4 chúng ta có thể biến đổi thành câu truy vấn Q’3.4 như sau:

*Truy vấn Q’3.4*

*select < danh sách các trường > from <P<sub>1</sub>>, <P<sub>2</sub>>*

*where <fc<sub>1</sub>> and {  $\forall t \in P_2: t$  thoả mãn điều kiện <fc<sub>2</sub>> and (A<sub>1</sub>  $\theta$  t[A<sub>2</sub>]) }*

**Ví dụ 3.4.** Sử dụng hai quan hệ trong ví dụ 3.2

Tìm mặt hàng (MAHANG, TENHANG) bán trong tháng 01 có THANHTIEN

“lớn hơn” tất cả các mặt hàng bán trong tháng 02 mà có số lượng bán rất nhiều.

*select MAHANG, TENHANG from nkban\_thang01 where THANHTIEN >\_1 all (select THANHTIEN from nkban\_thang02 where nkban\_thang02.SOLUONG =\_1 rất nhiều).*

Tương tự như ở câu truy vấn Q3.1, ta có bổ đề:

**Bổ đề 3.4.** Kết quả câu truy vấn Q3.4 là tương đương với kết quả câu truy vấn Q'3.4 trong CSDL mờ.

Chứng minh: Tương tự như bổ đề 3.1

**Bổ đề 3.5.** Trong CSDL mờ, mọi câu truy vấn con có thể biến đổi tương đương thành câu truy vấn mờ.

Chứng minh: Dựa vào các bổ đề từ 3.1 đến 3.4.

## 5. Kết luận

Phương pháp xử lý truy vấn con theo các cách tiếp cận truyền thống trước đây như: cách tiếp cận lý thuyết tập mờ, cách tiếp cận quan hệ tương tự có nhiều nhược điểm do việc xây dựng các quan hệ đối sánh mang nhiều ý kiến chuyên gia và không trực quan. Chẳng hạn xây dựng quan hệ đối sánh theo cách tiếp cận lý thuyết tập mờ thì chủ yếu chuyển các tập mờ về dạng hình thang, tam giác để đối sánh, từ đó chọn ngưỡng  $\alpha$  ( $\alpha \in [0,1]$ ) phù hợp khi truy vấn. Cách tiếp cận theo quan hệ tương tự (*quan hệ tương đương mờ*) thì tính bắc cầu của nó chỉ là bắc cầu max-min, min-max... Đây không phải là tính bắc cầu mạnh.

Vì vậy, trong bài báo này, với ưu điểm của ĐSGT, chúng tôi đã nghiên cứu và đề xuất một phương pháp mới để xử lý câu truy vấn con trong mô hình CSDL mờ theo cách tiếp cận ĐSGT. Các câu truy vấn con được biến đổi về dạng câu truy vấn thông thường làm cho việc xử lý dữ liệu được đơn giản hơn. Ngoài ra, việc thực hiện truy vấn con theo cách tiếp cận ĐSGT có ưu điểm hơn so với một số cách tiếp cận giải quyết trước đây như lý thuyết tập mờ, quan hệ tương tự đó là không sử dụng phương pháp khử mờ các giá trị kết quả truy vấn.

## TÀI LIỆU THAM KHẢO

1. Nguyễn Cát Hồ, Nguyễn Công Hào. *Một cách tiếp cận để xấp xỉ dữ liệu trong cơ sở dữ liệu mờ. Tạp chí tin học và điều khiển học*, T23, S2, 2007, 110-121.
2. Nguyễn Công Hào. *Một phương pháp xử lý giá trị khoảng trong cơ sở dữ liệu mờ*, Tạp chí Bưu chính Viễn thông và Công nghệ thông tin, “Chuyên san Các công trình nghiên cứu khoa học, nghiên cứu triển khai Công nghệ thông tin-Truyền thông”, Số 18, 2007, 68-74.



3. Nguyễn Cát Hồ, Nguyễn Công Hào. *Một phương pháp xử lý truy vấn trong cơ sở dữ liệu mờ theo cách tiếp cận đại số gia tử*. Tạp chí tin học và điều khiển học, T24, S4, 2008, 281-294.
4. Nguyễn Cát Hồ, Lê Xuân Vinh, Nguyễn Công Hào. *Thống nhất cơ sở dữ liệu và xây dựng quan hệ tương tự trong cơ sở dữ liệu thông tin ngôn ngữ bằng Đại số gia tử*. Tạp chí tin học và điều khiển học, T25, S4, 2009, 314-332
5. Hiroshi NAKAJIMA, Taiji SOGOH, Masaki ARAO, *Fuzzy Database Query Language and Library-Fuzzy Extention to SQL*, 1993 *IEEE*, 477-482.
6. Bosc P, Pivert O. *SQLf: a relational database language for fuzzy query*. *IEEE Transaction on Fuzzy Systems*, Vol. 3, No.1, 1995, 1-19.
7. Qi Yang, Chengwen Liu, Jing Wu, Clement Yu, Son Dao, Hiroshi Nakajima. *Efficient Processing of Nested fuzzy SQL Queries*. 1995 *IEEE*, 131-138.
8. Patrick Bosc. *Subqueries in SQLf, a Fuzzy Database Query Language*. 1995 *IEEE*, 3636-3641.

## **A METHOD FOR PROCESSING SUBQUERIES IN FUZZY DATABASES WITH HEDGE ALGEBRA BASED SEMANTICS**

*Nguyen Cong Hao  
Hue University*

### **SUMMARY**

*In this paper, we introduce an approach for processing fuzzy database subqueries to manipulate data in fuzzy database with hedge algebra based semantics. Data manipulation languages are proposed according to new models of fuzzy database. Finally, methods for transforming subqueries into corresponding queries are also introduced in this paper.*