

모호질의 처리에 관한 연구

이상준*, 김장형*

A Study on the Fuzzy Query Processing

*Sang-Joon Lee**, *Jang-Hyung Kim**

Summary

The queries involving fuzzy or imprecise conditions can not be processed in conventional DBMSs because their query languages are based on the boolean logic. Recently there have been many researches for processing fuzzy queries in the database systems, but their approaches have difficulties in defining the fuzziness embedded in queries.

To overcome such limitations, we propose a fuzzy query system model based on the multilevel definition of fuzziness. In the proposed system model, the fuzziness are interpreted based on the fuzzy set theory. Those fuzzy queries are expressed in the relational query language extended from QUEL.

서 론

근래에 들어 정보의 저장과 처리를 위한 데이터베이스의 중요성은 점점증하고 있다. 데이터베이스는 컴퓨터를 이용하는 거의 진분야의 자료관리 체제로 확고하게 자리를 잡았다. 그러나 데이터베이스의 응용영역이 확대됨에 따라 자연어에 많이 존재하는 모호한 표현을 DBMS가 처리할 수 있도록 하는 요구가 생겨나게 되었다. 부울논리에 기반을 둔 기존의 상용화된 DBMS에서는 사람들이 일상적으로 많이 쓰고 있는 표현인 '키가 큰' 또는 '매우 젊은' 등과 같은 모호한 표현을 처리하지 못하고 있다. 이러한 문제점

을 해결하기 위하여 ADB(Gala *et al.*, 1991)에서는 새로운 비교 연산자 'is'와 'close-to'를 도입하였으며, ARES(Ichikawa and Hirakawa, 1986)에서는 연산자 '≈'를 도입하여 모호성을 표현하고자 하였다. 그러나 이러한 연산자들은 매우 제한적인 의미의 모호성 밖에 표현할 수 없었다. 이러한 의미적 제한성을 극복하기 위해서 표현하기 위해 퍼지집합 이론(fuzzy set theory)(Zadeh, 1965)에 기반을 두고 모호성을 모델링하고자 하는 여러 연구들이 수행되었다(Leech, 1982; Bosc *et al.*, 1988; Tahani, 1977; 이 등, 1991a, 이 등, 1991b).

그러나 기존 연구들에서 제안된 방법들은 모호성에 대한 의미 표현력이 미약하거나 모호성의 해석을

* 공과대학 정보공학과(Dept. of Information Engineering, Cheju Univ., Cheju-do, 690-756, Korea)

위해 명백한 소속함수가 필요하였다. 그러나 이러한 소속함수를 정의하는 것은 결코 쉬운일이 아니다. 본 연구에서는 이러한 문제 해결을 위한 방법으로 복합적 의미를 내포한 합성값(compound value)의 도입과 다중 수준(multilevel)에서 모호성의 의미정의를 할 수 있는 기법을 제안한다.

ADB에서는 SQL에 두가지 형태의 연산자 'is'와 'close-to'를 추가하여 모호조건을 표현하였다. 'is'는 퍼지 연산자로서 언어값을 이용하며, 'close-to'는 명시된 값과 가까운 튜플들을 검색하기 위한 것이다. 이러한 연산자들을 이용하여 다음과 같은 질의를 할 수 있다.

- a) RETRIEVE(P. NAME)
WHERE P. AGE close-to 40 0.3
- b) RETRIEVE(P. NAME)
WHERE P. AGE is YOUNG

a)는 AGE가 40에 가까운 사람들을 검색하기 위한 것이고, b)는 나이가 'YOUNG'인 사람들을 검색하기 위한 것으로써 다음과 같이 시스템 내부에 정의되어 있는 조건으로 대치되어 평가된다.

YOUNG : AGE \Leftarrow 35

ADB에서의 모호조건은 숫자 애트리뷰트에 대해서만 제한적으로 적용이 가능하고, 언어값의 의미를 단지 부울조건으로 나타내어 단순히 참 아니면 거짓의 극단적인 진리값으로 평가되기 때문에 모호성을 다루는데 있어서 의미 표현력이 미약하다.

ARES에서는 새로운 비교연산자 ' \approx ' (approximately equal to)를 도입하여 근사적 동치조건을 표현한다. ARES는 근사적 동치조건을 처리를 위해서 특정 수준에서의 동치관계를 이용된다. 즉 시스템의 내부에 해당 도메인에 속하는 원소들간의 동치성에 관한 정보가 정의되어 있다. 예를 들어 '수준 0'에서 "job \approx professor"는 "job=professor"로 해석되고, '수준 1'에서는 "job \in {professor, reseacher, engineer}"로 해석된다. 따라서 부울조건만이 가능한 전통적인 시스템들에 비하여 모호성을 포함한 질의가 가능하지만 불확실한 의미 모델링을 위해서 일반적인 구조를 제공하지 못하고 있다.

ADB나 ARES등 지금까지의 대부분의 시스템들은 모호성을 다루는데 있어서 의미 표현력이 미약하거나 또는 그렇지 않을 경우라도 명백한 소속함수의 존재를 가정하고 있다. 따라서 기존의 방법들은 모호성을 정의하는 데 많은 어려움이 발생한다.

본 연구는 기존 시스템에 대한 두가지 문제 해결에 관심을 집중하고 있다. 즉, 모호성을 사용자가 쉽고 자연스럽게 정의할 수 있도록 새로운 기법을 도입하는 것과 사용자가 모호성을 표현하는 데에 보다 많은 융통성을 제공하고자 하는 것이다.

모호성 모델링

1. 모호성 정의

질의의 조건부에 존재하는 모호조건을 해석하기 위해서 필요한 지식은 사용자에 의해서 정의된다고 가정한다. 사용자가 정의하는 지식의 형태는 시스템이 실제로 이용할 수 있는 지식의 형태로 변환되어 질의 평가시에 이용된다.

질의에 내포될 수 있는 모호성의 종류는 '유사성(similarity)', '언어값(linguistic value)', '합성값(compound value)'이 있다. 유사성의 해석을 위하여 사용자는 유사관계를 직접 정의하지 않고 다중 수준들에서 동치관계를 나타내는 분할들을 정의하고, 마찬가지로 언어값(linguistic value)의 해석을 위해서 다중 수준들에서의 α -수준집합(α -cut set)들을 정의한다. 또한 합성값의 의미 해석을 위해서 합성규칙을 정의 한다.

- 임시 도메인(temporal domain)

아래와 같이 데이터베이스의 도메인을 임시 도메인으로 제한하여 분할이나 α -수준집합을 정의하며 데이터베이스 내의 릴레이션 R은 각 애트리뷰트 도메인들의 카티션 프로덕트의 부분집합이다.

$$R \subseteq D_1 \times D_2 \times \dots \times D_m \quad (1)$$

t가 R의 튜플이라고 할 때, t의 k번째 애트리뷰트의 값 t_k 는 해당 도메인 D_k 의 원소이다.

$$t_k \in D_k, t \in R \quad (2)$$

정의 1. D_k 의 임시도메인 \hat{D}_k 는 현재 데이터베이스 내에 존재하는 도메인의 원소들로 이루어진 집합이다.

$$\hat{D}_k = \bigcup_{i \in R} t_k C D_k$$

• 유사성 (similarity)

Zadeh는 (Zadeh, 1971)에서 동치관계를 유사관계로 일반화 하였으며, 유사관계는 해당 도메인에 속하는 원소들 간의 이진관계 (binary relation)로써 유사성의 정도 $[0, 1]$ 의 값으로 사상된다. 유사관계에 대한 자세한 정의는 다음과 같다.

정의 2. D_k 의 임시도메인 \hat{D}_k 상에 정의된 유사관계가 S 라고 할 때, $x, y, z \in \hat{D}_k$ 이면 $S(x, y) \in [0, 1]$ 이며 아래의 세가지 특성을 갖는다.

- ① 반사 : $S(x, x) = 1$
- ② 대칭 : $S(x, y) = S(y, x)$
- ③ 이행 : $S(x, z) \geq \max(\min(S(x, y), S(y, z)))$
 $y \in \hat{D}_k$

특정 도메인상에 정의된 동치관계 (equivalence relation)는 동치인 원소들로 분할되는 특성을 갖는다. 이처럼 동치관계를 갖는 원소들의 부분집합을 동치 클래스 (equivalence class)라 하며, 분할의 특성에 관한 정의는 다음과 같다.

정의 3. D_k 의 임시도메인 \hat{D}_k 상에 정의된 분할이 P 라고 할 때, $P = \{P_i | i \in I, P_i \subseteq P\}$, P_i 는 다음과 같은 세가지 조건을 만족한다.

- ① $P_i \neq \emptyset$
- ② $P_i \cap P_j = \emptyset, i \neq j, i, j \in I$
- ③ $\bigcup_{i \in I} P_i = P$

분할로 표현된 동치관계에 관한 정의는 다음과 같다.

정의 4. 임시도메인 \hat{D}_k 의 원소 x 와 y 가 동치관계 E 를 가지려면, \hat{D}_k 의 분할 P 의 같은 동치클래스 P_i 에 속해야 한다.

$$E(x, y) = 1, \text{ 단, } x, y \in P_i$$

유사관계 S 를 각 수준들에서의 α -cut들의 집합 $L(S)$ 로 다음과 같이 나타낼 수 있다.

$$L(S) = \{\alpha_1/S, \alpha_2/S, \dots, \alpha_n/S\} \quad (3)$$

이러한 α -cut된 동치관계로부터 아래의 공식으로 유사관계를 구할 수 있다.

$$S(x, y) = \max_{\alpha \in [0, 1]} \min(\alpha/S(x, y), \alpha) \quad (4)$$

유사관계 S 의 α -cut은 동치관계이므로 각각은 분할로 표현될 수 있다.

$$L(p) = \{P_1, P_2, \dots, P_n\} \quad (5)$$

이러한 분할들의 집합은 α 가 점차로 감소하는 순서를 갖는다. 따라서 사용자는 유사관계를 직접 정의하지 않고 각 수준들에서의 분할을 정의한다. 단일 각 분할의 수준값이 지정되지 않는다면 아래와 같은 공식으로 적절히 계산하여 할당하는 기법을 이용한다.

$$\alpha_i = 1 - (i-1) \times (1/n), \quad 1 \leq i \leq n \quad (6)$$

각 수준에서의 분할들로부터 유사관계를 생성하는 절차는 다음과 같다.

- 유사관계의 생성 절차

- 1) 수준값이 감소하는 순서로 해당 임시도메인에 대한 분할들을 입력
- 2) 수준값이 명시되지 않았으면 계산(식 6)
- 3) 각 분할들을 동치관계로 변화(정의 4)
- 4) 각 수준에서의 동치관계로부터 유사관계를 생성(식 4)

• 언어값 (linguistic value)

퍼지집합에서는 모호용어인 'TALL', 'YOUNG' 등으로 값을 줄 수가 있는데 이러한 값들을 언어값이라 한다. 언어값으로 표현된 퍼지집합 F 에서의 소속함수 μ_F 는 각 원소를 $[0, 1]$ 으로 대응시킨다.

$$\mu_F(x) \rightarrow [0, 1] \quad (7)$$

이러한 퍼지집합에서 일정한 소속정도 (mem-

bership grade) 이상을 가지는 원소들로만으로 보통집합들을 구성할 수 있다. 이것을 α -수준집합이라고 부르며 소속함수의 값이 α 이상인 원소들만으로 이루어진다.

퍼지집합 F를 α -수준집합들로 표현하면 다음과 같다.

$$F\alpha = \{x \in X \mid \mu_F(x) \geq \alpha\}, \quad \forall \alpha \in (0, 1] \quad (8)$$

이러한 보통집합들의 모임으로부터 퍼지집합을 구하는 공식은 다음과 같다.

$$\mu_F(x) = \max_{\alpha \in (0, 1]} \min(\mu_{F\alpha}(x), \alpha) \quad (9)$$

퍼지집합 F를 α -수준집합들의 집합 L(F)로 다음과 같이 나타낼 수 있다.

$$L(F) = \{\alpha_1/F, \alpha_2/F, \dots, \alpha_n/F\} \quad (10)$$

본 연구에서는 사용자가 직접 퍼지집합을 정의하지 않고 위와 같은 형태로 L(F)를 수준값 α 가 감소하는 순서로 명시하도록 하고, 만약 수준값들이 없다면 공식 6으로 적절한 값을 할당한다.

퍼지집합 F를 α 가 감소하는 순서를 갖는 α -수준집합들의 집합 L(F)로부터 구하는 절차는 다음과 같다.

- 퍼지집합 생성 절차

- 1) 수준값이 감소하는 순서로 퍼지집합의 α -수준 집합들을 입력
- 2) 수준값이 명시되지 않았으면 계산(식 6)
- 3) 각 수준에서의 α -수준집합들로부터 소속함수를 생성(식 9)

• 합성값(compound value)

보통의 모호조건들은 유사관계나 퍼지집합으로 의미를 평가할 수 있으나 합성값의 경우는 다른 모호조건들로부터 유도된다. 합성값은 합성규칙으로 의미가 정의되며 여러 애트리뷰트 값에 의존하는 특성을 갖는다. 사용자는 합성값을 질의에 이용하기 위해서 합성규칙을 정의해야 한다. 다른 개념들로부터 유도된 합성값 'BIG', 'OLD', 'SMALL'의 정의에 관한 규칙들의 예는 다음과 같다.

합성규칙의 예

BIG : AGE is TALL and WEIGHT is HEAVY
 OLD : -AGE is not YOUNG
 SMALL : -HEIGHT is not TALL

예에서 보인 규칙들에서 ':' = '의 왼쪽은 합성값의 이름을 나타내며 오른쪽은 합성값의 의미를 나타낸다. 오른쪽 부분에 존재하는 어휘에 대해서는 시스템 내부에 정의되어 있는 것으로 가정한다.

2. 모호 질의어

• 질의어의 형식

본 연구에서는 QUEL-like 질의어를 제공하며 조건절에 모호조건을 표현할 수 있도록 확장한 표현을 쓰는 것이 가능하다. 질의문은 명령어절과 WHERE절 두 부분으로 나뉘며, 'RETRIEVE' 명령은 다음과 같은 형식을 갖는다.

RETRIEVE (R, A1, R, A2, ..., R, An)
 WHERE <조건절>

R은 데이터베이스 내에 존재하는 릴레이션의 이름이고 A1, A2, ..., An은 R의 애트리뷰트 이름이다.

• 모호조건

조건절에는 부울조건(>, <, =)과 모호조건이 가능하며, 모호조건은 다음과 같이 세가지의 형태를 갖는다.

- Type1) A is {m} a {at t}
- Type2) A is {m} F {at t}
- Type3) A is {m} T {at t}

A는 해당 릴레이션의 애트리뷰트의 이름이며 "is"는 모호조건임을 나타내는 퍼지연산자이다. "{ }"로 표시된 부분은 생략이 가능한 부분이고, m은 모호조건을 만족하는 정도를 수정하기 위한 수정자(modifier)를 지시하며, t는 임계값(threshold)으로서 모호조건을 만족하는 정도를 t 이상으로 제한하기 위해서 이용된다.

Type1은 애트리뷰트 A의 값이 a와 유사한 값을 갖는 튜플들을 검색하기 위한 것이고, Type2는 F의

의미를 나타내는 퍼지집합에 포함된 값을 가진 튜플들을 검색하기 위한 것이고, Type3은 T의 의미를 나타내는 합성규칙을 만족하는 튜플들의 검색을 위한 모호조건이다.

Type1의 형태를 갖는 모호조건을 평가하기 위해서는 A의 도메인상에 정의된 유사관계 S가 이용되고, Type2에는 F의 소속함수 μ_F 가 이용되며, Type3에는 여러 조건들로 이루어진 T의 합성규칙이 필요하다.

결과 및 고찰

본 장에서는 표본 데이터베이스에 대한 모호질의 처리과정을 적용해보고 고찰한다.

데이터베이스에는 Fig. 1과 같은 릴레이션 PERSON이 존재한다고 가정하고, 다음과 같이 모호성을 내포한 세가지의 사용자 질의를 가정해보자.

- 질의 1. RETRIEVE (PERSON, NAME)
WHERE PERSON, AGE is 17
- 질의 2. RETRIEVE (PERSON, NAME)
WHERE PERSON, AGE is YOUNG
- 질의 3. RETRIEVE (PERSON, NAME)
WHERE PERSON is STRONG

PERSON

NAME	AGE	HEIGHT	WEIGHT
Park S. K	10	120	30
Cho K. D	12	130	32
Kim Y. K	14	147	40
Lee K. T	17	170	62
Lee H. J	19	174	66
Lee M. L	20	178	70
Sung B. K	30	175	75
Lee S. Y	33	170	65
Hwang S. C	35	165	60

Fig. 1. Personal data.

질의 1은 나이가 17과 유사한 사람들을 검색하기 위한 질의이며, 질의 2는 젊은 사람들을 검색하기 위한 질의이고, 질의 3은 체력인 강한 사람들을 검색하기 위한 것이다.

사용자의 질의에 포함된 모호성을 시스템이 해석할 수 있도록 평가 지식이 필요하며 Fig. 2와 같이 전문가 또는 사용자에게 의해서 정의된다고 가정하자.

①은 AGE의 도메인상에 정의된 각 수준에서의 동

<p>① partitions of AGE</p> $L(P(AGE)) = \{a_1/\{10, 12, 14\}, \{14, 19, 20\}, \{30, 33, 35\}, \{50\}\},$ $a_2/\{\{10, 12, 14, 17, 19, 20\}, \{30, 33, 35, 50\}\}$
<p>② α-cut sets of YOUNG, TALL, HEAVY</p> $L(YOUNG(AGE)) = \{a_1/\{10, 12\}, a_2/\{10, 12, 14, 17\},$ $a_3/\{10, 12, 14, 17, 19, 20\},$ $a_4/\{10, 12, 14, 17, 19, 20, 30\}\}$ $L(TALL(HEIGHT)) = \{a_1/\{178, 183\}, a_2/\{174, 175, 176, 178, 183\},$ $a_3/\{160, 165, 170, 174, 175, 176, 178, 183\}\}$ $L(HEAVY(WEIGHT)) = \{a_1/\{75, 80\}, a_2/\{68, 70, 75, 80\},$ $a_3/\{60, 62, 65, 66, 68, 70, 75, 80\}\}$
<p>③ Compound rule of STRONG</p> <p>STRONG : -AGE is YOUNG and WEIGHT is HEAVY and HEIGHT is TALL</p>

Fig. 2. Evaluation knowledges.

치관계를 나타내는 분할들의 집합으로, 첫번째의 분할은 정확도 수준이 나중의 것보다 명확하다는 순서관계를 갖고 있다. 즉, 첫번째의 분할에서는 14와 17이 다른 것으로 평가되고 두번째 분할에서는 같은 동치클래스에 속하여 같은 또는 유사한 값으로 해석할 수 있다. ②에는 언어값들 'YOUNG', 'TALL', 'HEAVY'에 대한 의미를 나타내는 각 수준에서의 집합들의 모임이며, 이들의 각 수준집합들도 정확도가

감소하는 순서관계를 내포하고 있다. ③은 합성값 'STRONG'의 해석을 위한 합성규칙으로 여러 애트리뷰트에 의존적인 모호조건으로 구성되는 몸체(body)를 갖는 규칙이다.

①과 ②로부터 Fig. 3과 같은 유사관계와 소속함수를 생성해서 질의 평가시에 모호성의 해석에 이용하며, 이러한 유사관계와 소속함수는 Fig. 4, 5와 같이 테이블의 형태로 유지된다.

$$\begin{aligned} \mu_S(\text{AGE}, \text{AGE}) &= \{1.00/(10, 10), 1.00/(19, 12), 1.00/(10, 14), \\ &\quad 1.00/(12, 14), \dots, 0.5/(35, 50)\} \\ \mu_{\text{YOUNG}}(\text{AGE}) &= \{1.00/10, 1.00/12, 0.75/14, 0.75/17, 0.50/19, \\ &\quad 0.50/20, 0.25/30\} \\ \mu_{\text{TALL}}(\text{HEIGHT}) &= \{1.00/183, 1.00/178, 0.67/176, 0.67/175, \\ &\quad 0.67/174, 0.34/170, 0.34/165, 0.34/160\} \\ \mu_{\text{HEAVY}}(\text{WEIGHT}) &= \{1.00/80, 1.00/75, 0.67/70, 0.67/68, \\ &\quad 0.34/66, 0.34/65, 0.34/62, 0.34/60\} \end{aligned}$$

Fig. 3. Similarity relation and membership grade table.

SIMILARITY (AGE, AGE)		
AGE	AGE	$\mu_S(\text{AGE}, \text{AGE})$
10	12	1.0
10	14	1.0
10	17	0.5
10	19	0.5
10	20	0.5
12	14	1.0
12	17	0.5
12	19	0.5
12	20	0.5
14	17	0.5
14	19	0.5
14	20	0.5
17	19	1.0
17	20	1.0

Fig. 4. Similarity relation table of AGE.

YOUNG (AGE)	
AGE	$\mu_{\text{YOUNG}}(\text{AGE})$
10	1.00
12	1.00
14	0.75
17	0.75
19	0.50
20	0.50
30	0.25

Fig. 5. Membership grade table of YOUNG.

Fig. 1은 각 개인들에 대한 자료를 가정한 것이다. 먼저 질의 1의 결과인 Fig. 6에서 보면 'Lee K. T'는 만족도가 1이다. 이러한 결과는 질의에 제시된 나이가 17이고 'Lee K. T'의 나이가 17이므로 유사성의 측정을 위해 저장된 유사 관계 테이블에 접근할 필요가 없이 1로 확정한다. 한편 'Lee H. J'의 나이는 19이고 유사관계 테이블에서 17과 19의 유사성이 1이므로 조건의 만족도는 1이 된다. 질의 2에서 'Park S.

K'의 나이는 10이므로 언어값에 대한 만족도를 구하기 위해서 Fig. 5의 소속테이블에서 해당 나이를 검색하면 만족도 1을 구할 수 있다. 질의 3의 'STRONG'을 만족하기 위해서는 각 튜플들을 'YOUNG'과 'TALL', 'HEAVY'를 모두 만족해야 한다. 따라서 질의 3의 결과를 나타내는 Fig. 8에서 'Kim S.K'는 나이가 20이므로 'YOUNG'의 소속테이블을 참조하여 만족도 0.5를 구하고 몸무게가 65이므로 'HEAVY'에 대해서는 같은 방법으로 만족도 1을 구한다. 여기에서 minimum연산자를 적용하면 'STRONG'에 대한 최종 만족도는 0.5가 된다.

이 름	적합도
Lee K. T	1.00
Lee H. J	1.00
Lee M. L	1.00
Kim S. K	1.00
Park S. K	0.50
Cho K. D	0.50
Kim Y. K	0.50

Fig. 6. Results of query 1.

이 름	적합도
Park S. L	1.00
Cho K. D	1.00
Kim Y. K	0.75
Lee K. T	0.75
Lee H. J	0.50
Kim S. K	0.50
Lee M. L	0.50

Fig. 7. Results of query 2.

이 름	적합도
Kim S. K	0.50
Lee M. L	0.50
Lee K. T	0.34
Lee H. J	0.34
Sung B. K	0.25

Fig. 8. Results of query 3.

본 연구에서 제공한 질의어는 WHERE-절에 모호 연산자 'is'가 추가된 형태를 가지며, 이를 이용하여 세가지 의미의 모호조건을 모델링할 수 있다. 질의 1과 같이 사용자가 제시한 실제값과의 유사성과 질의 2에서의 'YOUNG'과 같은 언어값 그리고 질의 3에서의 'STRONG'과 같은 합성값을 질의에 이용할 수 있다. 이러한 모호성의 의미 정의를 위해 사용자는 각 수준에서의 분할 또는 α -수준집합을 이용한다. 따라서 특정 수준의 정보의 이용이 가능하기 때문에 사용자의 지식이 완전하지 않은 상황에서도 쉽게 의미정의가 가능하다. 사용자는 이러한 모호성을 질의에 도입하여 적절한 결과를 얻을 수 있다. 질의의 결과 릴레이션을 구성하는 각각의 튜플들은 해당 모호조건에 대하여 (0, 1)의 값을 만족도로 갖게되어 인간의 사고에 어느정도 가까운 결과가 된다. 또한 여러 애트리뷰트에 의존하는 합성값의 도입으로 모호성의 의미 표현력을 향상시킬 수 있다. 예를 들어 사용자는 합성값 'STRONG'이 어느 애트리뷰트에 연관된 것인지 모른다고 하더라도 적절한 결과를 얻을 수 있다.

적 요

퍼지집합이론은 현재 공장자동화를 비롯한 자동화 분야 및 각종자동제어 분야, 소프트웨어 개발분야, 의사결정 분야등에 성공적으로 적용이 되고 있어 앞으로 계속되는 연구가 필요한 분야이다. 본 연구에서는 데이터베이스 분야에 퍼지집합 이론을 이용하여 모호한 질의를 처리할 수 있도록 하는 방안에 대하여 논하였다.

데이터베이스에 대한 퍼지집합론의 적용은 새롭고도 잠재적인 가치가 있는 분야이다. 본 연구에서는 현재 널리 이용되는 관계형 데이터모델에 기반을 두고 있으며 QUEL-like 질의어를 이용하였다. 사용자는 다양한 모호성을 질의에 내포시킬 수 있었으며 모호성의 해석은 퍼지집합론의 이론에 바탕을 두었다. 본 연구에서는 기존의 방식과 다른 방법의 다중 수준에서 모호성을 정의하도록 하는 기법을 도입하였다. 또한 여러 애트리뷰트에 종속적인 합성값은 규칙을 이용하여 의미 정의가 가능하도록 하였다.

추후 연구 과제로는 모호 조건의 평가를 위한 지식

을 용이하게 정의할 수 있도록 하기 위해 지식 습득을 지원하는 보다 효율적인 기법과 모호성의 편집을 지원하는 틀에 관한 연구가 필요하다.

참 고 문 헌

- Bosc, P., M. Galibourg, G. Hamon, 1988. Fuzzy Querying with SQL : Extension and Implementation Aspects, *Fuzzy Sets and Systems*, 28: 333-349.
- Gala, S., D. Chawla, C.M. Eastman, 1991. Combining Fuzzy and Nonfuzzy Approximate Reterival in a Database Management System, *North American Fuzzy Information Processing Society*, 57-61.
- Ichikawa, T., M. Hirakawa, 1986. ARES : A Relational Database with the Capability of Performing Flexible Interpretation of Queries, *IEEE Transactions on Software Engineering*, SE-12(5) : 80-85.
- Leech, M.Z., 1982. Uncertainty Propagation to Expert Systems, *Approximate Reasoning in Expert Systems*, North-Holland, 529-548.
- Tahani, V., A Conceptual Framework for Fuzzy Query Processing-A Step Toward very Intelligent Database Systems, *Information Proceeing & Management*, 289-303.
- Zedeh, L. A., 1965. Fuzzy Sets, *Inform. Contr.*, 8 : 338-353.
- Zadeh, L. A., 1971. Similarity Relations and Fuzzy Ordering, *Information Sciences*, 3(2) : 177-200.
- 이상준, 박석규, 김기태, 1991a. 관계 데이터베이스에서 퍼지한 질의를 처리하기 위한 지식 기반형 FQS (Fuzzy Query System)에 관한 연구, *한국정보과학회*, 18(2) : 289-292.
- 이상준, 조경달, 김기태, 1991b. 의사결정 지원을 위한 Fuzzy 질의에 관한 연구, *한국정보과학회*, 18(2) : 285-288.