

관계형데이터베이스에서 FSQL 시스템의 설계 및 구현*

김 병 섭** · 박 순 철***

< 목 차 >

I. 서론	2.3. 튜플소속척도 구하는 방법
II. 퍼지 질의어의 정의	2.3.1. 평균치 방법
2.1. 퍼지 질의 모델	2.3.2. N-Root 방법
2.1.1. (모델 1)의 경우	III. FSQL 시스템의 구현
2.1.2. (모델 2)의 경우	3.1 시스템 전체 구조
2.2. 소속척도 (Membership Value)구하기	3.2. 동작설명
2.2.1. 계산 형태 (Computational metric)	3.2.1. Parse
2.2.2. 표 형태 (Tabular metric)	3.2.2. Fuzzy Query Processor
	3.2.3. MetaDatabase
	IV. 결과
	V. 결론
	참고문헌
	Abstract

I. 서론

최근 컴퓨터 산업의 발달과 멀티미디어의 발달 등으로 비속련된 사용자의 데이터베이스 시스템의 사용이 증가하고 있다. 기존의 관계형 데이터베이스 시스템(Date, 1985; Korth, 1991)에서의 질의는 부울조건에 근거하고 있기 때문에, 질의 조건에 정확히 일치

* 본 연구는 한국과학재단의 핵심 전문 연구 지원사업(과제번호: 971-0902-0202)으로 수행중에 있음

** 전북대학교 정보통신공학과

*** 전북대학교 정보통신공학과 조교수

하는 자료들만을 제공한다. 그러나, 사용자들이 요구하는 조건에 정확히 일치하는 자료가 존재하지 않을 경우 사용자는 데이터베이스에 대한 어떠한 정보도 얻지 못한다. 이러한 이유로 최근 퍼지 이론(Klir, 1988; Lee, 1991; Zadeh, 1965; Zadeh, 1973)을 적용하여 부울조건에 한정된 질의가 아닌 애매한 질의를 지원할 수 있는 FSQL(Fuzzy Standard Query Language)에 대한 연구가 계속되고 있다(이광규, 1996; 이도현, 1993; 이석균 1997; Buckles, 1982; Buckles, 1983; Motro, 1988; S.C. Park, 1995; Sheno, 1990; Zemankova, 1985). 이에 본 논문에서는 VAGUE 모델(이광형, 1991; Motro, 1988; S.C. Park, 1995)을 기반으로 확장된 질의어와 MetaDatabase를 사용한 FSQL 시스템을 설계 및 구현하였다.

본 논문에서는 비전문가의 사용에 중점을 두어 기존의 복잡한 FSQL 구문을 간략화 하였으며, GUI(Graphic User Interface)를 이용한 사용자의 입력을 최소화하였다. 또한, 관리자에 의해 수정이 가능한 MetaDatabase를 사용하여 퍼지 질의어에 대한 정보를 형성시켜 기존의 관계형 데이터베이스를 수정 없이 퍼지 질의를 할 수 있도록 하였다. 이렇게 함으로써 일반적으로 사용하는 관계형 데이터베이스를 사용하여 퍼지 질의어 시스템을 구축할 수 있도록 하였다. 데이터베이스의 연동은 마이크로소프트사의 DAO(Data Access Objects)(이상엽, 1997; Hipson, 1996; Michael Groh, 1996)를 사용하여 기존의 관계형 데이터베이스를 연동 하였다. DAO를 사용함으로써 향후 JET(Joint Engine Technology)엔진(Micheal Groh, 1996)을 사용하는 모든 데이터베이스와 연동 시킬 수 있도록 하였다.

본 연구에서는 퍼지 질의를 처리할 수 있는 퍼지 질의문을 정의하고 이러한 질의문을 토대로 기본적인 기능을 하는 퍼지 질의 시스템을 개발하였다. 본고의 2장에서는 퍼지 질의어의 정의에 대하여 살펴보고, 정의한 질의어를 중심으로 3장에서는 FSQL 시스템의 구현원리를 설명한 후, 4장에서는 구현 결과를 살펴보고, 끝으로 5장에서 결론을 맺는다.

II. 퍼지 질의어의 정의

퍼지 관계형 데이터베이스 시스템은 크게 FRDB(Fuzzy Relational Database) 모델과 VAGUE 모델로 나눌 수 있다(이광형, 1991). FRDB 모델은 실제 관계(Relation)들의 모임을 저장하는 VDB(value database)와 실제 관계에서 사용된 속성값들에 관련된 정보들을 저장하는 EDB(explanatory database)로 구성이 되어있으며, VDB에 속성값 자체를 저장하고, EDB에 저장되어 있는 유사관계를 참조하여 질의에 결과를 보여준다. 이처럼 FRDB 모델은 불확실한 데이터베이스와 질의어를 모두 수용하여 각 데이터의 불확실성을 처리하는 반면, VAGUE 모델은 기존 관계형 데이터베이스 시스템의 데이터를 그대로 사용하며, 질의어 형태만을 변경시켜 표 형태, 계산 형태 등의 거리계산법을 사용하여 소속척도를 구함으로써 사용자의 애매한 질의를 기존의 데이터베이스 변화 없이 처리할

수 있도록 지원해 준다(이광형, 1991). 본 논문에서는 기존의 관계형 데이터베이스의 수정없이 퍼지 질의어를 적용 시킬 수 있는 VAGUE 모델을 기반으로 FSQL 시스템을 구현하였다(이광형, 1991; S.C. Park, 1995).

2.1 퍼지 질의 모델

본 논문에서 사용하는 관계형 데이터베이스 테이블은 참고문헌 [17](S.C. Park, 1995)을 참조하였으며, <표 1>과 같다.

<표 1> 관계형 데이터베이스

이름	나이	최종학력	성별	결혼유무
김진아	30	대졸	여	무
채지숙	26	대중퇴	여	무
박형우	19	고졸	남	무
정희영	18	고졸	여	무
허연희	28	고졸	여	무
김기남	45	대졸	남	무
조아라	35	대중퇴	여	무
최창호	27	고졸	남	무
이하영	24	고졸	여	무
신서림	28	중졸	여	무

<표 1>과 같은 기존의 관계형 데이터베이스 시스템에서의 “고등학교 교육을 받은 24세 여성을 찾아라”(S.C. Park, 1995)와 같은 질의어는 아래의 (질의 1)과 같은 SQL로 표현이 가능하며, 해당 데이터베이스에 고졸 졸업자의 24세 여성의 자료가 존재하지 않는다면 아무런 정보를 얻을 수가 없다.

(질의 1)

```
Select *
From 개인신상
Where [최종학력]="고졸" and [나이]=24 and [성별]="여";
```

이러한 질의에 퍼지한 성격을 추가한 다음과 같은 두가지 질의 모델로 퍼지 질의 시스템을 구현한다.

(모델 1)

“고등학교 정도 교육을 받은 24세 정도의 여성을 찾아라”

(모델 2)

“고등학교 정도 교육을 받은 젊은 여성을 찾아라”(S.C. Park, 1995)

위에서와 같은 퍼지한 성질들은 Where문에 포함이 되었으며, 이러한 Where문에 포함된 질의를 본 논문에서는 다음과 같이 정의한다.

2.1.1 (모델 1)의 경우

Where [최종학력]=“고졸”/f and [나이]=24/f and [성별]=“여”; …………… (질의 2)

즉, “고졸”/f나 24/f처럼 “/f”를 포함시키면 정확한 속성이 아닌 퍼지한 속성을 요구하는 퍼지 질의를 나타낸다. 사용자가 질의를 하는데 정확히 찾고자하는 속성값을 모를 경우나, 또는, 찾고자 하는 질의에 가까운 정보들까지 보고 싶을 경우 사용할 수 있다. 또한, (질의 2)의 경우는 MetaDatabase의 거리정보를 이용한 세미퍼지(어느 정도 정확함을 요구하는 퍼지 질의어를 세미퍼지라 정의)를 도용하여 다음과 같이 사용자의 의도를 추가할 수 있다.

Where [최종학력]=“고졸”/2 and [나이]=24/3 and [성별]=“여”;……………(질의 3)

즉, 정확히 고등학교 졸업자와 24세 여성의 정보와 함께 [최종학력]에 대해서는 거리가 2(초등학교,중등학교,고등학교,대중퇴,대졸)이고, [나이]에 대해서는 거리가 3 (21, 22, 23, 24, 25, 26, 27)까지를 찾는다. 이렇게 함으로써 사용자가 찾고자 하는 정보가 없을 경우라도 사용자의 질의에 가까운 정보를 제공할 수 있다. 마치 영역 검색과 비슷하지만 소속척도를 제시하여 준다는 점에서 큰 차이가 있다. 결과에 각 속성의 소속척도와 튜플 소속척도를 제시하여 줌으로써 사용자가 찾고자 하는 정보와의 관련도를 보여준다. 사용자는 1차적으로 튜플 소속척도를 비교하고, 2차적으로 소속척도를 비교함으로써 최적의 정보를 얻을 수 있다.

2.1.2 (모델 2)의 경우

Where [최종학력]=“고졸”/f and [나이]=“young” and [성별]=“여”; …………… (질의 4)

즉, “young”과 같은 자연어에 가까운 퍼지 질의어는 숙련자에 의해 제공되는 MetaDatabase(본 논문의 3.2.3.절 참조)를 참조하여, 앞서 제시한 퍼지 질의어로 변환되어 처리된다. MetaDatabase를 사용함으로써 숙련자에 의해 퍼지 질의어를 자유로이 확장시킬 수 있으며, 독립적인 구성이 가능하다.

다음은 [나이]="young" 일 경우의 변환 예이다.

예)

"young" => [나이]가 20세에서 30세정도인 사람

=> Where [나이]=25/5; (With MetaDatabase)

결과적으로 다음과 같이 변환된다.

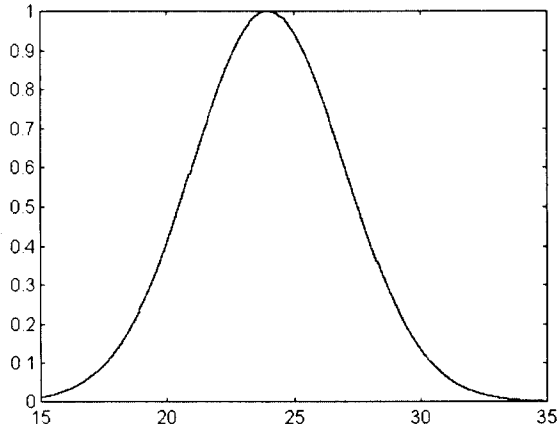
Where [최종학력]="고졸"/f and [나이]=25/5 and [성별]="여";

2.2 소속척도(Membership Value) 구하기

VAGUE 모델은 각 속성의 정의구역에 대하여 소속척도를 구하기 위해 거리계산법이 할당되는데, 거리계산법은 특정한 두 속성값의 서로 다른 정도를 알 수 있으며(이광형, 1991), 본 연구에서는 아래와 같이 VAGUE 모델(이광형, 1991)의 변형된 방법을 사용한다.

2.2.1 계산 형태(Computational metric)

소속척도를 결정하는 방법 중 속성값이 정수나 실수같은 연속적인 값을 갖을 경우에는



<그림 1> 데이터의 정규분포

함수를 사용하여 소속척도를 구할 수가 있으며, 본 논문에서는 식 (1)과 같은 범종형 함수인 정규분포(Normal distribution)함수를 이용하였다(이광형, 1991; Keller, 1988; S.C. Park, 1995).

$$\mu(x, a) = e^{-\frac{1}{2} \left[\frac{(x-a)}{a} \right]^2} \dots\dots\dots (1)$$

<표 2> MetaDatabase I

Label	Index
초졸	1
중졸	2
고졸	3
대중퇴	4
대졸	5
대졸이상	6
초등학교	1
초등졸	1
중학교	2
중학교졸업	2
고등학교	3
고등학교졸업	3
초등학교졸업	1
대학교중퇴	4
대학교졸업	5
학사	5
학사이상	6
국졸	1
박사	7
*	0

현재 레코드: 1

식 (1)에서 x 는 변수이고, a 는 평균 값, α 는 표준 편차이다. 즉, x 는 관계형 데이터 베이스로부터 가져온 비교 값, a 는 질의로부터 가져온 찾고자 하는 값, α 는 퍼지 정도로 α 만큼의 차이를 허용한다. 예로 24세 정도 되는 사람을 찾는데 위아래로 3살 정도의 차이를 허용한다면([나이]=24/3) $a=24$, $\alpha=3$ 으로 지정되며, μ 함수는 아래위로 3살까지는 $0.5 \leq \mu(x,a) \leq 1$ 의 값을, 나머지 영역은 $0 \leq \mu(x,a) \leq 0.5$ 로 결과를 보인다. 결과적으로, 질의자는 $\mu(x,a)$ 값이 0.5이상인 21, 22, 23, 24, 25, 26, 27세의 나이를 갖는 데이터들을 각각의 소속척도 0.80, 0.95, 1.00, 0.95, 0.80, 0.61과 함께 얻을 수 있다(S.C. Park, 1995). 다음은 정규분포 함수에 의해 얻어진 결과를 보이고 있다. <그림 1>에서 보는 바와 같이 찾고자 하는 24세에서 1값을 갖으며 가장 적합한 값이며, 24세를 중심으로 분포가 되어 있음을 알 수가 있다.

2.2.2 표 형태(Tabular metric)

퍼지 질의에 대한 해당 속성값이 숫자가 아니고 문자일 경우, 즉, [최종학력]이 고졸정도인 사람을 찾을 경우에는 기존의 퍼지 질의어에서는 소속척도를 표에 직접 저장 해놓는 표형태를 사용하였다(이광형, 1991; S.C. Park, 1995). 그러나, 이러한 방법을 사용했을 경우에는 주 데이터베이스에 의존도가 높으며 소속척도들의 삽입 삭제 및 관리가 매우 힘들었다. 이에 본 논문에서는 기존의 표를 이용한 방법처럼 소속척도를 직접 표에 저장하지 않고 유사도에 의해 정수화를 시켜 MetaDatabase화함으로써 주 데이터베이스와의 의존도를 줄이며 독립적인 관리가 가능하도록 하였다. 또한, <표 2>에서 처럼 “고졸”, “고등학교”, 그리고 “고등학교졸업”등과 같은 유사어를 같은 정수 “3”으로 정의함으로써 기존의 질의어가 정확한 속성 한가지의 단어만을 질의하는 제한을 극복할 수 있다는 장점이 있다.

만일, [최종학력]=“고졸”/2라는 질의를 받는다면 <표 2>의 MetaDatabase를 참조하여 [최종학력]=3/2와 같은 정수로 변환됨으로써 2.2.1.절과 같은 계산 방법을 이용하여 소속척도를 구할 수 있다. 이렇게 함으로써 2.2.1.절에서 소개한 계산형태와 동일한 방법으로 소속척도를 구하게 되는 장점을 갖는다.

2.3 튜플소속척도 구하는 방법

2.2.절에서 구한 각 칼럼의 소속척도를 가지고 퍼지한 질의에 맞는 최적의 튜플을 선택할 수 있도록 하는 튜플소속척도를 구하는 방법에는 여러 가지가 있지만 본 논문의 FSQL 시스템의 구현에는 모든 영역 소속척도를 튜플 소속척도에 반영해주는 참고문헌 [17]에서 제안한 2가지 방법(평균치 방법, N-Root 방법 : 각 방법에 대한 특징은 참고문헌[17]참조)을 사용자의 선택(기본은 평균치 방법)에 의해 각각의 방법을 사용할 수 있도록 구현하였다.

2.3.1 평균치 방법

본 논문의 2.2.절에서 제시한 방법으로 구한 각 칼럼의 소속척도들의 평균을 사용한 방법으로 식 (2)와 같다. 평균치 방법은 각 소속척도들을 균일한 중요도로 처리한다(S.C. Park, 1995).

$$\mu_t = \frac{1}{m} \cdot \left(\sum_{i=1}^m \mu_i \right), \quad (1 \leq m \leq n) \quad \dots\dots\dots (2)$$

(μ_t : 튜플소속척도, μ_i : 해당소속척도, m : 소속척도갯수, n : 총 컬럼수)

2.3.2 N-Root 방법

본 논문의 2.2.절에서 제시한 방법으로 구한 각 칼럼의 소속척도들을 곱한 값에 N-제곱근(N-square root)을 하는 방식으로, 식 (3)과 같다(S.C. Park, 1995).

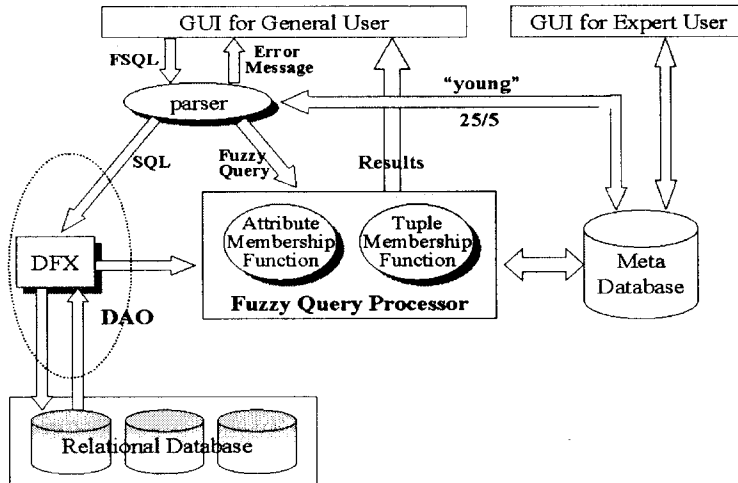
$$\mu_t = \sqrt[m]{\mu_1 \times \mu_2 \times \dots \times \mu_m} \quad \dots\dots\dots (3)$$

III. FSQL 시스템의 구현

FSQL의 구현은 Windows 95/NT 상에서 Visual C++ 5.0을 사용하였으며, 객체지향프로그래밍을 함으로써 기능을 모듈화하였다. 관계형 데이터베이스로는 마이크로소프트사의 ACCESS 97을 사용하여 구현하였다. 데이터베이스의 제어는 DAO(Data Access Objects)를 사용함으로써 DML(Data Manipulation Language)뿐만 아니라 DDL(Data Definition Language)(Hipson, 1996; Michael Groh,1996) 수준의 구현이 가능하므로 세부적으로 구현을 할 수 있으며, 향후 다른 데이터베이스와의 연동을 가능할 수 있게 하였다.

3.1 시스템 전체 구조

FSQL 시스템의 전체적인 구조는 <그림 2>와 같다.



<그림 2> FSQL 시스템의 전체구조

<그림 2>에서 보는 바와 같이 FSQ 시스템은 크게 parser, Fuzzy Query Processor와 MetaDatabase로 나눌 수가 있다.

3.2 동작설명

3.2.1 Parse

<표 3> MetaDatabase II

	word	query
▶	young	25/5
	젊음	25/5
	젊은	25/5
	old	50/10
	늙은	50/10
*		

현재 레코드: 1

parser는 사용자로부터 본고의 2장에서 정의된 형태의 퍼지 질의어를 넘겨 받는다. 넘겨 받은 퍼지 질의어는 <그림 2>와 같이 우선 “young”이나 “젊은”과 같은 자연어 성질의 질의를 분석한 후 자연어 질의가 있다면, MetaDatabase(<표 3> 참조)를 통하여 처리 가능한 형태로 변환시킨다. 이렇게 변형된 퍼지 성분이 포함된 Where문을 표준 SQL 부분과 퍼지한 성격의 조건으로 분리시킨다. 즉, 본 논문에서 사용한 퍼지 질의는 아래와 같이 분리된다.

2.1.절에서 정의한 (질의 3) :

Where [최종학력]=“고졸”/2 and [나이]=24/3 and [성별]=“여”;

=> (1) 표준 SQL : Where [성별]=“여”;

(2) 퍼지조건 : [최종학력]=“고졸”/2 and [나이]=24/3;

2.1.절에서 정의한 (질의 4) :

Where [최종학력]=“고졸”/f and [나이]=“young” and [성별]=“여”;

=> (1) 표준 SQL : Where [성별]=“여”;

(2) 퍼지 조건 : [최종학력]=“고졸”/f and [나이]=25/5

분리된 표준 SQL은 DAO객체를 통해서 해당 관계형 데이터베이스로 전달되어 일반적인 질의에 대한 결과를 얻는다. 표준 질의에 대한 결과는 DFX(DAO record field exchange)(Hipson, 1996; 이상엽, 1997)를 통해 접근할 수 있으며, 결과는 parser가 분리해 낸 또다른 퍼지한 성격의 조건과 함께 Fuzzy Query Processor로 전달된다.

3.2.2 Fuzzy Query Processor

Fuzzy Query Processor는 parser로부터 넘겨 받은 퍼지한 질의를 최소단위로 분리시킨다. 본 논문에서 사용한 퍼지 질의 모델은 다음과 같은 형태로 분리된다.

(질의 3) : [최종학력]="고졸"/2, [나이]=24/3;

(질의 4) : [최종학력]="고졸"/f, [나이]=25/5;

분리된 조건문은 2.2.절에서 기술한 퍼지조건의 성질에 맞는 소속척도를 구한다. 각각의 조건에 맞게 구해진 소속척도를 가지고 2.3.절에서 제시하는 방법으로 튜플소속척도를 구한다.

3.2.3 MetaDatabase

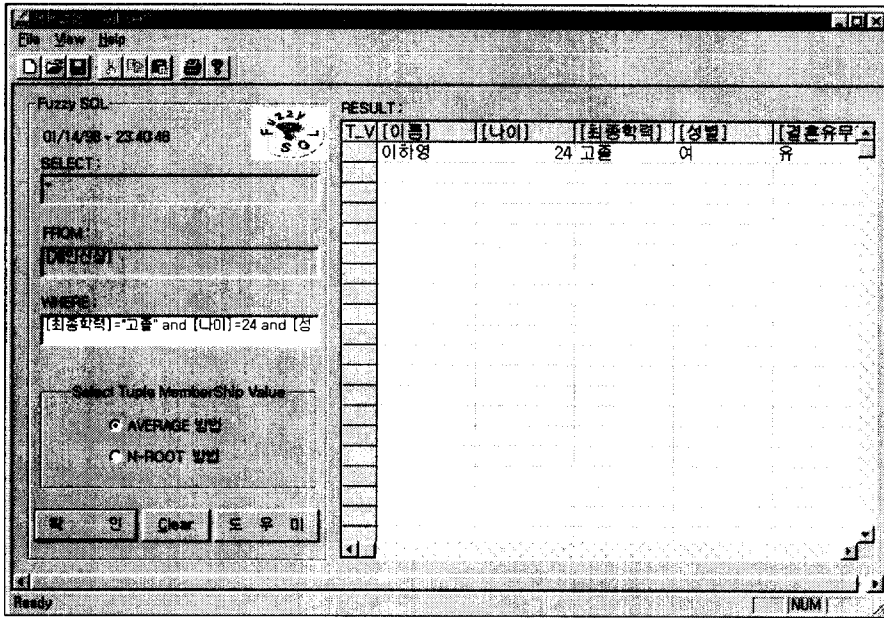
MetaDatabase는 퍼지 질의시에 참조하는 데이터베이스이다. MetaDatabase 역시 다루기 쉬운 관계형 데이터베이스로 구성하였으며, 숙련자에 의해 관리되며 특정 목적에 맞게 수정, 확장이 가능하다. 2.2.1절에서 소개한 <표 2>와 <표 3>은 본 논문에서 사용한 MetaDatabase를 보여준다.

FSQL 시스템에서 MetaDatabase의 사용은 본 논문의 2.1절과 2.2.1절을 참조하기 바란다.

IV. 결 과

본 논문의 2장에서 정의한 퍼지 질의어와 3장에서 설계한 FSQL 시스템의 구조를 기반으로 하여 FSQL 시스템을 윈도우즈 환경하에서 구현하였다. <그림 3>은 기존의 표준 SQL(질의 1)경우의 실행화면이다. <그림 3>에서 보는 것처럼 FSQL 시스템은 Where문 중심의 처리를 하며, 퍼지 질의어인 경우 Radio Button(이상엽, 1997; David J, 1996)을 사용하여 튜플 소속척도를 구하는 방법(평균치 방법, N-Root 방법)(S.C. Park, 1995)을

사용자로 하여금 선택할 수 있도록 하였다. 선택이 없을 시에는 기본적으로 평균치 방법을 사용한다. 또한, 소속척도 및 튜플 소속척도의 결과를 보여줌으로써 사용자의 질의에 대한 결과를 확인할 수 있도록 하였으며, 튜플 소속척도가 0.5이상인 튜플들만을 인정하였다.



<그림 3> 표준 질의어의 실행 결과

다음은 2장에서 정의한 퍼지 질의어를 사용한 실행 결과 화면들이다. 퍼지 질의어에 대한 사항은 본 논문의 2장을 참고하기 바란다.

- (질의 2)에 평균치 방법의 적용시 결과

Where [최종학력]="고졸"/f and [나이]=24/f and [성별]="여"; (그림 4참조)

T_V	[이름]	[나이]	[최종학력]	[성별]	[결혼유무]
0.95	채지숙	26/0.92	대중퇴/0.98	여	무
0.86	허연희	28/0.73	고졸/1.00	여	무
1.00	이하영	24/1.00	고졸/1.00	여	유
0.85	신서림	28/0.73	중졸/0.98	여	무

<그림 4> 평균치 방법을 이용한 (질의 2)의 실행

- (질의 2)에 N-Root 방법의 적용시 결과

Where [최종학력]="고졸"/f and [나이]=24/f and [성별]="여"; (그림 5참조)

RESULT :					
T.V	[이름]	[나이]	[최종학력]	[성별]	[결혼유무]
0.95	채지숙	26/0.92	대중퇴/0.98	여	무
0.85	허연희	28/0.73	고졸/1.00	여	무
1.00	이하영	24/1.00	고졸/1.00	여	유
0.84	신서림	28/0.73	중졸/0.98	여	무

<그림 5> N-Root 방식을 이용한 (질의 2)의 실행

- (질의 3)에 평균치 방법의 적용시 결과

Where [최종학력]="고졸"/2 and [나이]=24/3 and [성별]="여"; (그림 6참조)

RESULT :					
T.V	[이름]	[나이]	[최종학력]	[성별]	[결혼유무]
0.84	채지숙	26/0.80	대중퇴/0.88	여	무
1.00	이하영	24/1.00	고졸/1.00	여	유

<그림 6> 평균치 방법을 사용한 (질의 3)의 실행

- (질의 3)에 N-Root 방법의 적용시 결과

Where [최종학력]="고졸"/2 and [나이]=24/3 and [성별]="여"; (그림 7 참조)

RESULT :					
T.V	[이름]	[나이]	[최종학력]	[성별]	[결혼유무]
0.84	채지숙	26/0.80	대중퇴/0.88	여	무
1.00	이하영	24/1.00	고졸/1.00	여	유

<그림 7> N-Root 방법을 사용한 (질의 3)의 실행

- (질의 4)에 평균치 방법의 적용시 결과

Where [최종학력]="고졸"/f and [나이]="young" and [성별]="여"; (그림 8참조)

RESULT :					
T.V	[이름]	[나이]	[최종학력]	[성별]	[결혼유무]
0.76	김진아	30/0.61	대졸/0.92	여	유
0.98	채지숙	26/0.98	대중퇴/0.98	여	무
0.92	허연희	28/0.84	고졸/1.00	여	무
0.99	이하영	24/0.98	고졸/1.00	여	유
0.91	신서림	28/0.84	중졸/0.98	여	무

<그림 8> 평균치 방법을 사용한 (질의 4)의 실행

- (질의 4)에 N-Root 방법의 적용시 결과

Where [최종학력]="고졸"/f and [나이]="young" and [성별]="여"; (그림 9참조)

T.V	[이름]	[나이]	[최종학력]	[성별]	[결혼유무]
0.75	김진아	30/0.61	대졸/0.92	여	유
0.98	채지숙	26/0.98	대중퇴/0.98	여	무
0.91	허연희	28/0.84	고졸/1.00	여	무
0.99	이하영	24/0.98	고졸/1.00	여	유
0.90	신서림	28/0.84	중졸/0.98	여	무

<그림 9> N-Root 방법을 사용한 (질의 4)의 실행

개발한 FSQl 시스템의 실행 결과들에서 보여 주듯이 표준 질의어와 정의한 퍼지 질의어를 사용하여 정보를 검색할 수 있다. 검색 결과는 튜플 소속척도가 0.5이상인 튜플들만을 표시하여 주며 퍼지 질의어가 포함된 속성값의 결과에 대해서는 “/”로 구분된 소속척도를 함께 보여준다. 사용자는 1차적으로 튜플 소속척도를 기준으로 검색 결과를 확인할 수 있으며, 만족도를 높이기 위해서는 2차적으로 튜플내의 각각의 소속척도들을 확인함으로써 최적의 정보를 얻을 수 있다.

V. 결론

데이터베이스를 사용하는 사용자들의 보다 편리한 정보 검색에 대한 욕구는 끝임이 없다. 이러한 이유로 사용에 제약이 없는 질의어가 필요하게 되었으며, 이에 퍼지 이론을 도입한 퍼지 데이터베이스의 연구가 계속되고 있다.

본 논문에서는 참고문헌[17](S.C. Park, 1995)에서 사용했던 애매한 퍼지 질의를 기초로 하여 새롭게 정의한 퍼지 질의어(FSQL)를 사용하여 기존의 관계형 데이터베이스를 기반으로 한 FSQl 시스템의 프로토타입을 구현하였다. 기존의 방법들과는 달리 MetaDatabase를 사용하여 구현함으로써, 크게 두가지의 장점을 얻을 수 있었다. 첫째는, 사용자로부터 퍼지 질의어의 제한을 극복 할 수 있었다. 즉, “고졸”/f나 “고등학교졸업”/f 등을 같은 질의로 처리 할 수 있도록 해줌으로써, 일반 사용자들에게 단언기준이 아닌 의미에 기준한 검색이 가능하다. 둘째로, 퍼지 질의어 조건들을 수정 및 확장이 용이하다. 즉, 특정 목적에 FSQl 시스템을 적용 시킬 경우나 FSQl의 자연어 처리에 사용되는 어휘들을 확장시킬 경우 프로그램수준의 수정이 아닌 관계형 데이터베이스인 MetaDatabase의 수정 및 추가로 적용이 가능하다는 것이다. 또한, 마이크로소프트사의 JET엔진(Michael Groh, 1996) 내부의 DAO(Hipson, 1996; 이상엽 1997; Michael Groh, 1996)를 사용함으로써

일부 다른 데이터베이스들도 사용이 가능하다.

본 논문에서는 정의한 퍼지 질의어와 제안한 MetaDatabase를 사용한 처리 방법에 기준 최소한의 기능만을 지원하는 FSQL 시스템을 구현함으로써 제안한 방법의 퍼지 질의어와 자연어 처리가 가능한 데이터베이스 시스템의 실현 가능성만을 보였다. 따라서, 범용성을 포함한 몇가지의 문제점들을 내포하고 있으며, 아직 확실한 성능 분석이 되어 있지 않다. 향후에는 이러한 문제점들을 최소화시킬 것이며, 다양한 질의어를 수용하고, 자연어 처리 부분의 연구를 계속할 것이다.

참 고 문 헌

- [1] 이광규, 전근환, 김홍기, “퍼지 질의어를 이용한 FDBL 모델”, Journal of the Research Institute for Computer Science. Vol. 4, No. 1, 1996.2.
- [2] 이광형, 오길록, 퍼지이론 및 응용 II권, 홍릉과학출판사, 4장, 1991.
- [3] 이도현, 이광형, 김명호, “퍼지 질의를 위한 관계대수의 확장”, 정보과학회논문지 제20권, 제2호, 1993.2.
- [4] 이상엽, Visual C++ Programming Bible Ver 4.x, 영진, 1997.
- [5] 이석균, “신뢰도와 불확실성을 지원하는 확장 관계형 데이터 모델”, 정보과학회논문지 제24권 제4호, 1997.4.
- [6] Buckles, B. P. and Petry, F. E., A Fuzzy Representation of Data for Relational Database, Fuzzy Sets and Systems 7, pp.213-226, 1982.
- [7] Buckles, B. P. and Petry, F. E., “Information-Theoretical Characterization of Fuzzy Relational Databases”, *IEEE Trans. on Systems, Man, and Cybernetics SMC-13*, 1983, pp.74-77.
- [8] Date, C. J., An Introduction to Database Systems 1, Addison Wesley, 1985.
- [9] David J. Kruglinski, Inside Visual C++4, Microsoft, 1996.
- [10] Hipson & Jennings, Database Developer's Guide with Visual C++4, SAMS, 1996.
- [11] Keller, G., Warrack, B. and Bartel, H., Statistics for Management and Economics, Wadsworth, California, 1988.
- [12] Klir G. J. and Folger, T. A., Fuzzy Sets, Uncertainty, and Information, Prentice-Hall International(UK) Limited, London, 1988.
- [13] Korth H. F. and Silberschatz, A., Database System Concepts, McGRAWHILL, N. Y., 1991.
- [14] Lee K. H. and Oh, G. R., Fuzzy Sets and Applications 2, Hongreung Science Publishing Company, 1991.
- [15] Michael Groh, Microsoft Access 95 Power Toolkit, Ventana, 1996.
- [16] Motro, A. VAGUE, “A User Interface to Relational Databases that Permit Vague Queries”, *ACM Trans. on Office Information System 6*, pp.187-214, 1988.
- [17] S.C. Park, C.S. Kim and D.S. Kim, Fuzzy Logic and Its Applications to Engineering Information Sciences, and Intelligent Systems, KLUWER ACADEMIC PUBLISHERS, 1995, pp. 407-415.
- [18] Sheno, S. and Melton, A., “Proximity Relations in the Fuzzy Relational Database Model”, *Fuzzy Sets and Systems 31*, pp.285-296, 1989.

- [19] Shenoi, S., Melton, A. and Fan, L. T., "An Equivalence Classes Model of Fuzzy Relational Database", *Fuzzy Sets and Systems* 38, pp.153-170, 1990.
- [20] Zadeh, L. A., "Fuzzy Sets", *Information Control* 8, pp.338-353, 1965.
- [21] Zadeh, L. A., "Outline of a New Approach to the Analysis of Complex Systems and Decision Process", *IEEE Trans. on Systems, Man, and Cybernetics SMC-3*, 1973
- [22] Zemankova, M. and Kandel, A., "Implementing Imprecision in Information systems", *Information Science* 37, pp.107-141, 1985.

< Abstract >

Design and Implementation of FSQL System in Relational Databases

Kim Byoung-Seob · Park Soon-Cheol

The usages of the relational database systems have been increased nowadays. However there is a constraint that the standard query language for a database system requests users to know the exact condition for a query. In this paper a FSQL(Fuzzy SQL) system is proposed and implemented for the imprecise queries which do not include the exact condition. Moreover this system can supply the obvious results for the queries effectively. We define a simple fuzzy query language and use a method using the metadatabase to implement the FSQL system. This FSQL system can be combined with a relational database system without any painful effort.