

# 자연어 질의 처리를 위한 Meta Data에 관한 연구\*

## Study of Meta Data for Natural Language Query Processing

신세영, 정은영, 김승권, 김수영, 박순철  
전북대학교 정보통신공학과

e-mail: gracia@internet.chonbuk.ac.kr

Se Young Shin, Eun Young Cheong, Seung Gwon Kim,  
Su Young Kim, Soon C. Park  
Dept. of Information and Communication Eng.,  
Chonbuk National University

### 요약

정보산업의 발달과 함께 일반 사용자들의 데이터베이스의 사용이 증가함에 따라 부정확한 자연어 질의 처리를 할 수 있는 인공 지능적인 질의시스템이 필요하게 되었다. 이러한 질의시스템이 자연어 질의를 처리하려면 불확실한 데이터들에 대한 정보를 제공하는 MetaData가 반드시 필요하고, 데이터베이스 분야와 인공지능 분야의 이론들을 바탕으로 MetaData의 정형화 및 분류가 필요하다. 본 연구에서는 퍼지이론, 확률이론을 기초로 하여 소속척도, 근접추론, 유사관계, 데이터마이닝 기법등을 이용하여 MetaData를 정형화하고 분류하였다.

### 1. 서론

데이터베이스 시스템은 대용량의 데이터를 저장하고 관리하는 시스템으로 컴퓨터의 가장 성공적인 응용분야 중의 하나이다. 데이터베이스 시스템은 컴퓨터 산업의 발달과 더불어 지속적인 발전을 하였으며, 더욱이 최근의 멀티미디어의 발달과 인터넷 사용의 확장은 데이터베이스 사용을 급증하게 하였다. 이러한 데이터베이스는 Codd가 제안한 관계형 데이터베이스 시스템을 근간으로 널리 사용되고 있다. 그러나 이러한 데이터베이스 시스템은 비숙련된 사용자라 할지라도 데이터베이스를 접근하기 위해서는 SQL에 익숙해 있어야 하며 사용자가 정확한 질의 조건을 생성시키

지 못한다면 해당 데이터베이스로부터 어떠한 정보도 얻지 못한다. 이러한 이유로 최근 국내외에서는 쉬운 질의와 지능적인 결과를 제공하기 위해서 부정확한 자연어 질의 처리에 대한 연구가 계속 진행되고 있다. 이러한 질의시스템이 자연어 질의를 처리하려면 불확실한 데이터들에 대한 정보를 제공하는 MetaData가 반드시 필요하고, MetaData를 구축하기 위하여 가능분포의 퍼지 데이터베이스에서 자연어의 불확실성의 유사관계와 거리개념을 정의하여 데이터를 질의하는 가능이론과 VAGUE모델과 퍼지이론의 근사추론을 수행한 요인공간이론 이외에 데이터마이닝 기법을 이용한 방법 등, 데이터베이스 분야와 인공지능 분야의 여러 이론을 바탕으로 MetaData의 정형화와 분류에 관한 연구가 계속 진행되고 있다

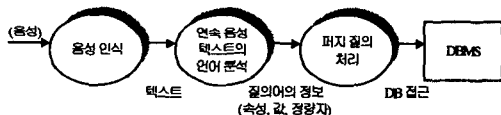
본 연구는 인적 사항 데이터베이스를 위한 음성 인터페이스 퍼지 질의처리 시스템을 구현하는 것을 최종목표로 하는 전체과제 가운데

11) 본 논문은 정보통신진흥연구원의 위탁으로 수행되고 있는 과제임( 과제명 : 인적사항 데이터베이스 조회를 위한 음성 인터페이스 퍼지질의 처리)

데 FSQL처리 단계에서 부정확한 질의가 처리 가능하도록 MetaData의 참조가 요구되는데, 이러한 MetaData의 정형화와 분류에 관한 연구이다. 현재까지는 자연어 질의처리를 위한 MetaData의 연구가 미비한 현인데, 본 연구는 MetaData를 구축하는 여러 이론과 방법들 중에서 퍼지이론, 확률이론을 기초로 하여 소속척도, 근접추론, 유사관계, 데이터마닝 기법등을 이용하여 MetaData를 정형화하고 분류하였다.

## 2. 전체 시스템의 구조

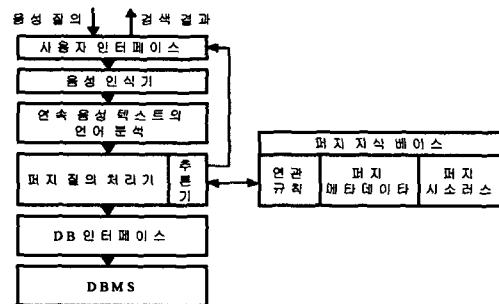
전체 시스템의 목표는 인적 사항 데이터베이스를 위한 음성 인터페이스 자연어 질의시스템을 구축하는 데에 있다. 전체 시스템에서 실질적으로 MetaData를 참조하여 FSQL 처리를 하려면 아래 그림과 같이 음성인식 단계와 자연어 처리를 통한 연속음성 텍스트의 언어 분석 단계를 거쳐야 한다. 예를 들어 음성을 통하여 “봉급이 많은 사람을 찾아라”[1,2,4]라는 질의가 입력되면, 음성인식 분야에서 음성을 인식하여 텍스트로 변환하고, 연속음성 텍스트 언어분석 단계에서는 Parser와 자연어 처리를 통하여 FSQL 처리부분에서 사용될 심볼 테이블의 키워드 즉, “봉급”, “많은”의 단어를 찾게 된다. 또한 키워드를 검색하여 퍼지 성분이 있으면 FSQL 처리를, 없으면 일반 SQL 처리를 하면 된다. 앞의 예제에서는 “많은”이라는 퍼지 성분이 포함되어 있으므로 FSQL 처리기에서 “봉급”에 관련된 MetaData를 통하여 “많은”의 정도를 나타내주는 데이터를 얻을 수 있다.



[그림 140] 음성 인터페이스 자연어 질의시스템

[그림 1]에서 본 전체 시스템의 단계별 과정은 [그림 2]와 같다. 먼저 input으로는 모든

컴퓨터 시스템에 대하여 가장 이상적인 사용자 인터페이스인 음성을 통한 자연어 질의가 들어온다. 데이터베이스에 대한 질의문으로 자연어 문장이 사용됨으로써, 질의 언어 문법에 대한 전문적 지식을 갖지 않은 일반 사용자들도 DBMS에 대한 용이한 접근이 가능하다. 이를 위해 음성 신호를 일단 텍스트로 변환한 후에 이 텍스트에 대하여 자연어 처리를 하여야 한다. 아울러 연속음성 텍스트는 단순히 신호처리를 거쳐서 나온 결과이기 때문에 언어처리기술을 이용하여 정련을 해 주어야 한다. 또한, 테스트 처리를 통한 결과를 데이터베이스가 인식할 수 있는 attribute, 정량자의 트리플로 변환시켜 주어야 한다. 마지막으로 이러한 트리플로 퍼지 질의를 위한 질의문을 구성하고 MetaData 처리를 수행하여 DB 인터페이스를 통해 사용자의 질의에 대한 검색 결과를 사용자 인터페이스를 통해 제공한다.



[그림141]음성 인터페이스 자연어 질의시스템의구조

본 연구에서는 전체 시스템 중 FSQL 처리 단계에서 사용되는 MetaData의 형태를 분류하고, 데이터베이스 분야와 인공지능 분야의 이용 가능한 이론과 방법을 바탕으로 MetaData에 적용한다.

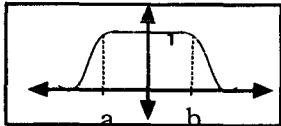
## 3. MetaData

MetaData는 자연어 질의 처리 시에 참조하는 자체 데이터베이스로써 퍼지 비교 연산자의 의미나 퍼지 상주 그리고 attribute 간의 유사한 정도를 정의해 둔 데이터베이스이다.

또한 데이터베이스 분야와 인공지능 분야의 여러 이론들에 의해 도출된 변환 데이터가 정의된 데이터베이스이다. 전자의 경우는 일반적인 MetaData의 형태이며, 그 형태에 따라 continuous form, index form, matrix form으로 분류할 수 있다. 후자의 경우는 확장된 MetaData의 형태이다.

### 3.1. continuous form

소속척도를 결정하는 방법 중 attribute가 정수나 실수 같은 연속적인 값을 가질 경우에는 함수를 사용하여 소속척도를 구할 수가 있다. 완전한 퍼지의 경우가 아닌 특정값에 퍼지 성분이 가미된 식(1)과 같은 범종형 함수를 사용된다. [4,5,6,7,8]

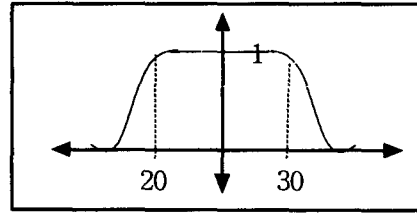


[그림 142] 데이터의 분포형태

$$\begin{aligned} \mu(x, a, b) &= e^{-\frac{1}{\alpha^2}[(x-a)]^2} \quad (x < a) \\ &= 1 \quad (a \leq x < b) \\ &= e^{-\frac{1}{\alpha^2}[(x-b)]^2} \quad (x > b) \dots \dots \dots (1) \end{aligned}$$

식 (1)에서 x는 변수 즉, 관계형 데이터베이스로부터 가져온 비교 값, a와 b는 질의로부터 가져온 찾고자 하는 값, α는 표준 편차이다. 따라서 a=b 즉, 소속척도가 1이 되는 구간이 한 점일 때에는 일반적인 범종형의 그래프로 나타내어진다.

예를 들어, “나이가 20에서 30세 정도인 사람을 찾아라”라는 질의가 들어오면, 식(1)의 값에 각각의 변수를 대입하여 소속척도를 구할 수 있다. threshold가 default로 0.8이고 α=10이라 가정하면, 소속척도 값이 0.8 이상인 18, 19, 20, 21, 22세의 나이를 갖는 데이터들을 소속척도 0.82, 0.94, 1, 0.94, 0.82와 함께 얻을 수 있다.



[그림 4] 데이터의 분포형태 예

또한 퍼지한 경우에 각 attribute들에 대하여 유사한 정도를 나타내는 소속함수를 구하여 MetaData로 저장한다 예를 들어 나이의 중간을 나타내는 소속함수는 [표 3]과 같이 여러 형태의 소속함수들이 존재할 수 있다. 다른 attribute들의 퍼지한 value의 소속함수들도 이와 같은 형태의 함수들을 따른다.

[표 1] MetaData I - Continuous Form

case	middle	분포형태
case 1		$\mu(x) = e^{-C(X-35)^2}$ (C=0.89*10 <sup>-2</sup> )
case 2		$\mu(x) = (x/10) - 2$ $= 1$ $= -(x/10) + 5$ ▶ 통계적 수치 표현의 어려움
case 3		$\mu(x) = (x/15) - (4/3)$ $= -(x/15) + (10/3)$ ▶ 비교적 소속척도의 값이 적게 나타남
case 4		$\mu(x) = e^{-dx-35}$ (c=0.045) ▶ 비교적 소속척도의 값이 적게 나타남

### 3.2. index form

퍼지 질의에 대한 attribute가 숫자가 아니고 문자일 경우, 즉, “최종학력”이 ‘고졸’ 정도인 사람을 찾을 경우에는 기존의 퍼지 질의어에서는 소속척도를 표에 직접 저장해 놓은 형태를 사용한다. 그러나, 이러한 방법을 사용

했을 경우에는 주 데이터베이스에 의존도가 높으며 소속척도들의 삽입 삭제 및 관리가 매우 힘들다. 이에 본 연구에서는 기존의 표를 이용한 방법처럼 소속척도를 직접 표에 저장하지 않고 유사도에 의해 정수화를 시켜 MetaData화 함으로써 주 데이터베이스와의 의존도를 줄이며 독립적인 관리가 가능하도록 하였다. 만일, “최종학력”=“고졸”라면 [표 4]의 MetaData를 참조하고 나온 값인 3을 소속함수에 대입하여서 디폴트 값으로 주어진  $\alpha = 0.8$  수준레벨을 이용하여 질의에 해당하는 값을 얻을 수 있다. [4,5,6,7,8]

[표 2] MetaDataII - Index Form

초등졸	1
중졸	2
고졸	3
대중퇴	4
대졸	5
대학원졸	6
박사	7

### 3.3. Matric form

자료간의 유사한 정도를 정략적으로 정의하는 형태이다. 대부분 2차 행렬로 표현하여, 자료간의 연관관계 정도를 특정 관점 기준으로 0과 1사이의 값으로 표시한다. 이러한 유사관계를 표현할 수 있는 MetaData가 반드시 필요하고, 다음과 같은 특징을 가진다.

- i) 원하는 자료와 의미상 가장 가까운 자료를 검색한다
- ii) 순서화된 결과를 얻을 수 있다

[표 4]는 전문분야에 대한 유사관계를 표현한 MetaData로 직업에 관련된 여러 카테고리들의 조합으로 유사도를 표현하였다. 유사도를 표현하는 방법에는 샤논의 엔트로피에 기초한 방법, 거리측정에 의한 방법 등이 있다.

[표 3] MetaDataIII-matric form

	DB	AI	IR	OS	KB
DB	1.0	0.7	0.8	0.0	0.9
AI	0.7	1.0	0.8	0.0	0.3
IR	0.8	0.8	1.0	0.0	0.6
OS	0.0	0.0	0.0	1.0	0.0
KB	0.9	0.3	0.6	0.0	1.0

또한 유사관계를 표현하는 대표적인 정의는 Shenoi와 Melton이 제한한 것으로 다음과 같다.

[정의 1] 근접관계  $s_j$ 는 도메인  $D_i$ 와 도메인  $D_j$  간의 곱이 0,1 사이의 실수 값과 사상관계(mapping)을 가진다.

(  $D_i * D_j \rightarrow [0,1]$  ) 이 때,  $x$ 와  $y$ 가 도메인  $D_i$  속하면,

(  $x, y \in D_i$  ) 다음의 식을 만족한다.

- i)  $s_j(x, x) = 1$  ( 반사성 )
- ii)  $s_i(x, y) = s_i(y, x)$  ( 대칭성 )

MetaData의 일반적인 분류는 위의 3가지 형태와 같이 분류할 수 있고, 또한 불확실한 질의를 처리하기 위하여 인공지능 분야의 여러 알고리즘을 이용할 수 있다. 그 중에 현재 널리 사용되고 있는 데이터마이닝은 KDD (Knowledge Discovery in Database) 라고 불리우는 복잡한 과정 중 ‘탐사단계’라는 중요한 단계이며, 대규모의 데이터베이스에서 숨겨진 지식, 예기치 않은 패턴, 새로운 규칙 등을 발견하여 것이다. 따라서 데이터마이닝은 부정확한 질의에 대해 발견된 지식을 사용하고 가공하여 인공지능적인 결과를 제시할 수 있다. 본 연구에서는 인공 지능적인 결과를 제시하기 위하여 데이터마이닝 기법 중의 하나인 의사결정트리를 참조하는 방법과 연관규칙을 이용하여 방법에 대하여 제시하였다.

### 3.4. reference form with dicision tree

의사결정트리는 의사결정규칙을 도표화하여 관심 대상이 되는 몇 개의 소집단으로 분류하거나 예측을 수행하는 분석방법으로 분석의 목적과 자료구조에 따라서 적절한 분리기준과 정지규칙을 지정하여 의사결정트리를 얻을 수 있다. 본 연구에서는 MetaData로써 사용되어지는 의사결정트리에 대하여 어떻게 기존의 데이터베이스로부터 분류, 분석하여 사용되는 지에 대해 설명한다.

예를 들어 경제적 계층, 성별, 카드정지 유무, 연령을 알 수 있는 [표 4]의 개인의 데이터와 함께 “신용상태가 좋은 사람을 찾아라”라는 질의가 입력되면 개인의 데이터를 사용하여

분석이 가능한 의사결정트리를 만들 수 있다. 의사결정트리의 학습 순서는 다음과 같다.

1. 문제 영역의 데이터 집합
2. 트리의 root node로 사용할 속성(attribute)을 선택
3. 선택된 속성의 값에 따라 데이터 집합을 분류 => 하위 집합 생성
4. 생성된 하위 집합 각각에 대하여 위 2, 3의 과정을 반복
5. 각각의 하위 집합에 속한 instance들이 모두 같은 클래스가 될 때까지 진행

[표 4] 개인 데이터의 예

사람	경제적 계층	성별	카드정지유무	연령	신용상태
P1	상	남	유	낮음	No
P2	상	남	유	높음	No
P3	중	남	유	낮음	Yes
P4	하	여	유	낮음	Yes
P5	하	여	무	낮음	Yes
...	...	...	...	...	...

의사결정트리를 생성하는데 있어서 가장 중요한 것은 ‘어떤 트리 구조를 갖느냐’의 문제이고, 그것은 어떤 ‘어떤 attribute를 선택하느냐’ 의해 결정된다. 일반적으로 최중적인 결정트리의 깊이가 작아질 수 있도록 attribute를 선택하여야 한다. 즉, 데이터를 attribute로 나눌 때 엔트로피가 감소하는 양인 information gain을 구하고 그 값이 최대가 되는 attribute를 선택하면 된다. 참고로 엔트로피와 information gain을 구하는 식은 다음과 같다.

$$\text{Entropy}(S) = -p^+ \log_2 p^+ - p^- \log_2 p^-$$

$p^+$ : 집합 S에서 결과가 YES인 인스턴스의 비율

$p^-$ : 집합 S에서 결과가 NO인 인스턴스의 비율

A: attribute

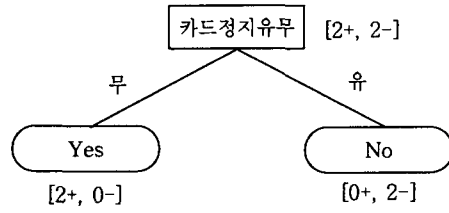
S: 데이터 집합

S1, S2, ..., Sn: 집합 S를 attribute A에 따라 분류한 하위 집합들

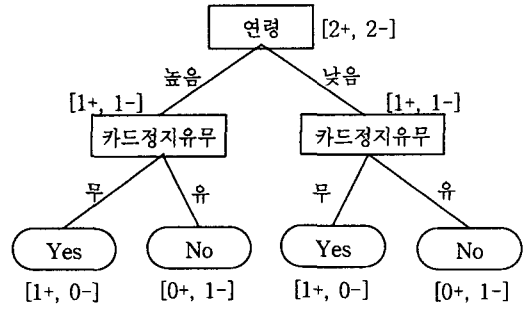
$$\text{Gain}(S, A) = \text{Entropy}(S) - \sum_{i=1}^n \left| \frac{S_i}{S} \right| \text{Entropy}(S_i)$$

예를 들어 [그림 5]와 [그림 6]의 트리의 경우, 위의 식을 통하여  $\text{Gain}(S, \text{카드정지유무}) = 1$ ,  $\text{Gain}(S, \text{연령}) = 0$ 의 값을 구할 수 있고 [그림 5]의

트리의 구조가 좋음을 알 수 있다.

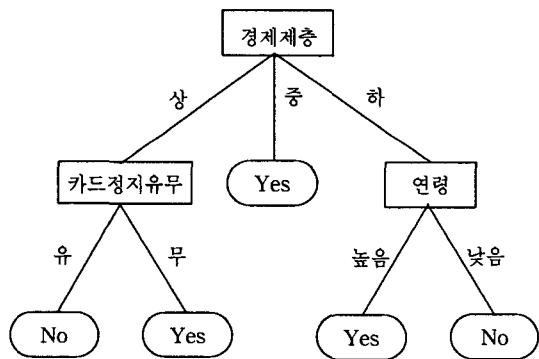


[그림 5] 트리 예(a)



[그림 6] 트리 예

따라서 [표 4]의 데이터를 사용하여 생성된 의사결정트리는 [그림 7]과 같고, 의사결정트리의 노드와 노드값을 일정한 형식으로 저장하여 MetaData로 사용하게 된다. 앞에서의 “신용상태가 좋은 사람을 찾아라”라는 질의는 사전의 학습에 의하여 저장된 MetaData를 참조하여 “경제적 계층이 상이고 카드정지유무가 유인 사람이나, 경제적 계층이 하인 사람으면서 연령이 높은 사람을 찾아라”라는 질의로 변환되어 데이터베이스에서 해당정보를 가져오게 된다.



[그림 7] 의사결정트리의 예

### 3.5. reference form with association rule

연관규칙도 데이터 마이닝의 한 기법으로 기존의

데이터베이스에서 각 attribute들에 대해서 나머지 attribute들에 대해 연관을 탐색하므로, 부정확한 질의에 대해 발견된 지식을 사용하고 가공하여 인공 지능적인 결과를 제시할 수 있다. 연관을 탐색하는 방법은, target attribute에 대하여 해당 데이터베이스의 attribute들의 연관을 조합하는 방식이다. 이 때, 단계별로 지지도와 신뢰도를 높이면서 만족하는 연관을 찾는다. 연관 규칙이 age=young, favorite=cartoon => buy\_Cartoon일 때 지지도와 신뢰도를 구하는 식은 다음과 같다.

$$\text{support} = P(\text{age}=\text{young} \cap \text{favorite}=\text{cartoon} \cap \text{buy\_Cartoon})$$

$$\text{confidence} = P(\text{buy\_Cartoon}) / P(\text{age}=\text{young} \cap \text{favorite}=\text{cartoon})$$

의사결정트리와 마찬가지로 탐색된 연관 규칙을 Metadata로 저장하여 부정확한 질의를 변환하여 데이터베이스에서 해당정보를 가져오게 된다.

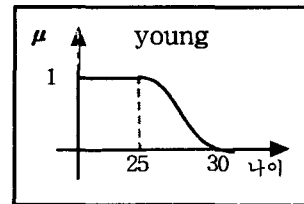
#### 4. Metadata의 응용

본 연구에서는 자연어 질의처리 시스템을 목표로 전체과제 중 그 전 단계인 FSQQL 처리기를 완성하였다.FSQQL 처리기에서 사용되어지는 Metadata의 응용과정은 다음과 같다.

예를 들어, [그림 8]과 같은 관계형 데이터베이스와 Metadata가 주어졌다고 가정하자. 이것을 이용하여 “나이가 젊고, 최종학력이 높으며 데이터베이스와 유사한 전문분야를 가진 사람을 찾아라”라는 자연어 질의가 처리되는 과정을 단계별로 나타내면 다음과 같다.

name	age	school_career	profession
Frank	25	석사	IR
Susan	26	박사	DB
Tom	40	중졸	OS
Anna	25	고졸	KB
Magaret	38	대졸	IR
John	20	고졸	IR
Steve	27	박사	KB
Smith	19	중졸	OS
Jane	28	대중퇴	DB
Maria	48	대졸	DB

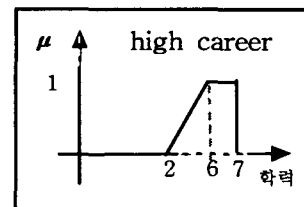
(a) personal\_info



(b) 소속함수 'young'

초등졸	1
중졸	2
고졸	3
대중퇴	4
대졸	5
대학원졸	6
박사	7

(c) index  
'school\_career'



(d) 소속함수 'high school-career'

	DB	AI	IR	OS	KB
DB	1.0	0.7	0.8	0.0	0.9
AI	0.7	1.0	0.8	0.0	0.3
IR	0.8	0.8	1.0	0.0	0.6
OS	0.0	0.0	0.0	1.0	0.0
KB	0.9	0.3	0.6	0.0	0.1

(e) matrix 'profession'

[그림 8] Metadata의 저장 예

**4.1. 퍼지질의 (FSQL 형태) 입력단계**  
 FSQL처리기로 입력되는 질의어의 형태는 다음과 같다.

```
SELECT name, age, school_career, profession
      bound 0.8
FROM   personal info
WHERE  age=~young bound 0.8 AND
      school_career=~high bound 0.8 AND
      profession=~'DB' bound 0.8;
```

**4.2. 퍼지 질의 변환 단계**

- (1) age=~young bound 0.8을 continuous form을 이용하여  $0 \leq \text{age} \leq 30$  로 바꾼다. 이때 0세부터 31세까지의 소속척도는 [그림 8]의 (a)의 MetaData에 따른다.
- (2) school\_career=~high bound 0.8을 index form을 이용하여  $6 \leq \text{school\_career} \leq 7$  로 바꾼다. 여기서도 최종학력의 소속척도는 [그림 8]의 (b)의 MetaData에 따른다.
- (3) profession=~'DB' bound 0.8는 matrix form의 유사관계를 이용하여 만족하는 근접한 전문분야로 대체한다. 즉, profession='DB' OR profession='IR' profession='KB'로 변환한다. 이러한 과정을 거쳐 변환된 입력 질의는 다음과 같다.

```
SELECT name, age, school_career, profession
FROM   personal info
WHERE  (0≤age≤30) AND (6≤school_career≤7)
      AND (profession='DB' OR profession='IR'
      profession='KB');
```

**4.3. 변환된 SQL을 튜플 소속척도에 의해 처리하는 단계**

변환된 질의를 관계형 데이터베이스 관리 시스템으로 처리하면 다음과 같은 튜플을 얻는다.

<Frank 31 석사 IR>  
 <Steve 29 박사 KB>

이것은 1차원 데이터만 가지므로 질의어에 대하여 유사도의 순위를 알기 위하여 구해진

MetaData와 매핑하여 튜플 소속척도를 구하면 [표 5]와 같은 결과를 얻을 수 있다.

[표 5] 질의처리 결과

name	age	school_career	profession	M.V.
Frank	25(1.0)	석사(1.0)	IR(0.9)	0.9
Steve	27(0.72)	박사(1.0)	KB(0.8)	0.8

결과적으로 튜플소속척도 값의 bound를 0.8로 정했으므로 Frank와 Steve가 해당되며, 질의에 가장 잘 만족하는 사람은 Frank이며, 두 번째는 Steve이다.

**5. 결론 및 향후 연구**

데이터베이스 사용의 증가와 데이터베이스 활용의 지능화

고도화에 대한 끊임없는 욕구는 계속되고 있다. 그러나 질의어가 필요하게 되고 데이터베이스의 사용의 증가와 연구가 계속 되고 있다. 데이터베이스 사용의 증가와 데이터베이스 활용의 지능화, 고도화에 대한 욕구는 끊임 없이 계속되고 있다. 이러한 이유로 사용에 제약이 없는 질의어가 필요하게 되었으며, 이에 인공지능 이론을 도입한 자연어 질의처리의 연구가 계속되고 있다.

본 연구에서는 보다 인공 지능적이고 유용한 정보를 제공 가능하게 하는 MetaData에 대하여, 퍼지이론, 확률이론을 기초로 하여 소속척도, 근접추론, 유사관계를 이용하여 정의하는 기본적인 형태와 데이터마닝 기법을 이용하여 정의하는 확장된 MetaData의 형태를 정형화하고 분류하였다. 또한 본 연구에서는 자연어 질의처리 시스템을 위한 그 전 단계인 FSQL 처리기를 완성하여 실제 MetaData를 구축하여 자연어 질의처리가 가능함을 보였다.

그러나 아직은 구현한 FSQL처리기는 제한적인 방법으로 자연어처리가 가능한 자연어 질의 시스템의 실현 가능성을 보였다. 아직까지는 제한적인 질의어 처리가 이루어지고 있으며, 방대한 량의 데이터베이스에 대해서도 MetaData를 구축하는데 있어서도 다양하고 인공 지능적인 방법이 필요하다. 구체적으로, 보다 정확한 통계량을 기준으로 MetaData를 설정

해야하고, 부정확한 자연어에 대하여 좀더 효과적인 소속함수의 개발이 필요하며, 새로운 이론을 도입하는 MetaData 확장에 대해서도 연구가 필요하다. 향후에는 이러한 제약점들을 최소화시킬 것이며, 자연어 질의의 확장과정형화에 관한 연구를 계속 할 것이다.

## 참 고 문 헌

- [1] Korth H. F. and Silberschatz, A., Database System Concepts, McGRAWHILL, N. Y., 1991.
- [2] Date, C. J., An Introduction to Database Systems 1, Addison Wesley, 1985.
- [3] Zadeh, L. A., "Fuzzy Sets", *Information Control* 8, pp.338-353, 1965.
- [4] S.C. Park, C.S. Kim and D.S, Kim, Fuzzy Logic and its applications to engineering information sciences, and intelligent systems, KLUWER ACADEMIC PUB LICSHERS, 1995, pp. 407-415.
- [5] 이광형, 오길록, 퍼지이론 및 응용 II권 홍릉과학출판사, 4장, 1991.
- [6] 이도현, 이광형, 김명호, "퍼지 질의를 위한 관계대수의 확장", 정보과학회논문지 제20권, 제2호, 1993.2.
- [7] Buckles, B. P. and Petry, F. E., A Fuzzy Representation of Data for Relational Database, *Fuzzy Sets and Systems* 7, pp.213-226, 1982.
- [8] Lee K. H. and Oh, G. R., Fuzzy Sets and Applications 2, Hongreung Science Publishing Company, 1991.
- [9] 이광규, 전근환, 김홍기, "퍼지 질의어를 이용한 FDBL 모델", *Journal of the Research Institute for Computer Science*. Vol. 4, No. 1, 1996.2.
- [10] 이석균, "신뢰도와 불확실성을 지원하는 확장 관계형 데이터 모델", 정보과학회논문지 제24권 제4호, 1997.4.
- [11] 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.
- [12] Motro, A. VAGUE, "A User Interface to Relational Databases that Permit Vague Queries", *ACM Trans. on Office Information System* 6, pp.187-214, 1988.
- [13] Sheno, S. and Melton, A., "Proximity Relations in the Fuzzy Relational Database Model", *Fuzzy Sets and Systems* 31, pp.285-296, 1989.
- [14] 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
- [15] Klir G. J. and Folger, T. A., Fuzzy Sets, Uncertainty, and Information, Prentice-Hall International(UK) Limited, London, 1988.
- [16] Sheno, S., Melton, A. and Fan, L. T., "An Equivalence Classes Model of Fuzzy Relational Database", *Fuzzy Sets and Systems* 38, pp.153-170, 1990.
- [17] Zemankova, M. and Kandel, A., "Implementing Imprecision in Information systems", *Information*



**신 세 영**



1995-1999 전북대학교 정보통신공학과  
1999-현재 전북대학교 대학원  
정보통신학과  
관심분야 : 데이터베이스, 네트워크 프로그래밍, Mobile Computing

**정 은 영**



1995-1999 전북대학교 정보통신공학과  
1999-현재 전북대학교 대학원  
정보통신학과  
관심분야 : 데이터베이스, 네트워크 프로그래밍, 인터넷 서비스

**김 승 권**



1991-1999 전북대학교 정보통신공학과  
1999-현재 전북대학교 대학원 정보통신학과  
관심분야 : Mobile Agents, 무선 인터넷(WAP), Fuzzy -SQL/데이터베이스

**김 수 영**



1993-2000 전북대학교 정보통신공학과  
2000-현재 전북대학교 대학원 정보통신학과  
관심분야 : XML, 네트워크 프로그래밍

**박 순 철**



1972-1979 인하대 응용물리학과  
1986-1991 미국 루이지애나 주립대 전자계산 박사  
1991-1993 한국 전자통신 연구소 데이터베이스 연구실 과제 책임자  
1993-현재 전자정보공학부 부교수  
관심분야 : 데이터베이스, 네트워크 컴퓨팅, SAN(Storage Area Network),