

퍼지규칙 기반 시스템에서 불필요한 속성 감축에 의한 패턴분류

Pattern classification on the basis of unnecessary attributes reduction in fuzzy rule-based systems

손 창 식*
Chang-Sik Son

김 두 완**
Doo-Ywan Kim

요 약

본 논문에서는 퍼지규칙 기반 시스템에서 규칙 내에 포함된 불완전한 속성을 제거하여 보다 간략화 된 규칙으로도 분류할 수 있는 방법을 제안하였다. 제안한 방법에서는 규칙 내에 포함된 불완전한 속성을 제거하기 위해 러프집합을 이용하였고 보다 명확한 분류를 위해 출력부 소속함수의 적합도가 최대인 속성들을 추출하였다. 또한 모의실험에서는 제안된 방법의 타당성을 검증하기 위해 rice taste data를 기반으로 규칙 감축 전 퍼지 max-product 결과와 규칙 감축 후 퍼지 max-product 결과를 비교하였다. 그 결과, 규칙 감축 전 max-product 결과와 규칙 감축 후 max-product 결과가 정확히 일치함을 볼 수 있었고, 보다 객관적인 검증을 위해 비퍼지화 된 실수 구간을 비교하였다.

Abstract

This paper proposed a method that can be simply analyzed instead of the basic general Fuzzy rule that its insufficient characters are cut out. Based on the proposed method, Rough sets are used to eliminate the incomplete attributes included in the rule and also for a classification more precise; the agreement of the membership function's output extracted the maximum attributes. Besides, the proposed method in the simulation shows that in order to verify the validity, compare the max-product result of fuzzy before and after reducing rule based on the rice taste data; then, we can see that both the max-product result of fuzzy before and after reducing rule are exactly the same; for a verification more objective, we compared the defuzzificated real number section.

↳ Keyword : Fuzzy Pattern Classification, Attribution Reduction, Rough Set

1. 서 론

퍼지추론 시스템은 기존의 이치 논리를 다치 논리로 확장한 프레임워크를 퍼지집합으로 본 개념으로 주로 선형 혹은 비선형 문제나 불완전하고 부정확한 지식을 효과적으로 처리할 수 있다는 장점을 가지고 있다[1].

가장 일반적인 퍼지 모델로는 Mamdani 타입의 모델과 Sugeno 타입의 모델로 분류할 수 있다. 이들 2가지 타입의 모델의 주된 차이점은 출력부의 출력형태에 따라 결정된다. 특히 Mamdani 타입의 퍼지 모델은 Sugeno 타입의 모델에 비해 복잡성을 가지는 선형 혹은 비선형 문제에서 규칙의 수가 기하급수적으로 증가하고 추론 및 패턴분류 시 계산시간이 많이 걸리며, 시스템 설계에서 복잡성이 증대된다는 단점을 가지고 있다. 기존의 연구에서는 이러한 단점들을 좀 더 개선시키기 위해 규칙 수를 감축하고 시스템의 복잡성을 줄이기 위한 연구가 지속되었다[2-7]. 수행된 연구방

* 정 회 원 : 대구가톨릭대학교 컴퓨터정보통신공학부
강의전담교수
fuzzyrisk@paran.com

** 정 회 원 : 동방대학원대학교 문화정보학과 교수
dooywan@nate.com

[2006/11/10 투고 - 2006/12/20 심사 - 2006/02/12 심사완료]

법들로 Moraga의 추론 방법, Eshera의 추론 네트워크 모델, Chantanapornchai의 추론 그래프 모델, Chandramohan의 규칙감축 방법 그리고 Ishibuchi의 승자규칙 선택방법이다.

Moraga은 단순히 2개의 입력과 1개의 출력을 가진 병렬 퍼지추론 방법을 제시하였고[2], Eshera와 Barash는 각 프로세서에서 규칙을 지정하고 규칙의 단계를 재 정렬하는 알고리즘을 바탕으로 다단계적으로 추론 할 수 있는 방법을 제시하였다[3]. 또한 Chantanapornchai는 프로세스의 추론 계산 시간을 줄이기 위해 퍼지 min, max 연산자를 이용하여 그래프의 노드 수를 최소화할 수 있는 방법을 제시하였다[4]. 그러나 제안된 3가지 방법들은 시스템의 계산시간과 복잡성을 줄이기 위하여 단순히 추론 규칙의 수를 감축하는데 중점을 두기 때문에 정확성이 떨어진다는 단점을 가진다.

Chandramohan은 퍼지 규칙을 줄임과 동시에 추론의 정확성을 향상시키기 위해 최적화와 신경망 등에서 주로 사용되는 4가지 기법 (harmonic mean, geometric mean, arithmetic mean, root mean square)에 기반을 둔 mean 연산을 통한 규칙 결합방법을 제안하였다[5]. 그러나 제안된 방법은 규칙의 수를 효율적으로 줄임과 동시에 원래의 추론 규칙과 유사한 결과를 얻을 수 있지만 다량의 입력데이터가 존재하는 경우 최적의 결과를 얻는데 다소 많이 시간이 소비된다는 제약점을 가진다.

Ishibuchi는 수치적인 입력패턴에 대응되는 규칙의 이행도가 최대가 되는 규칙 (승자규칙)을 추출함으로써 패턴분류에서 사용되는 규칙을 최소화한 방법이다[6,7]. 그러나 제안된 방법은 수치적인 패턴으로부터 최대 적합도를 가지는 승자규칙을 추출함으로써 최소의 규칙과 분류 정확도를 향상시킬 수 있으나 추론문제 적용 시 입력패턴에 대응된 모든 규칙을 고려하지 않기 때문에 추론 결과의 객관성을 보장할 수 가 없다.

따라서 이러한 문제점을 해결하기 위한 방법으로 분류문제에서 발생하는 여러 가지 유형의 불

완전한 정보의 의존성과 속성의 수를 감축하기 위한 선택 도구로서 러프집합을 이용한 방법들이 제안되었다.[8-12].

제안된 방법들 중 Slowinski[10]의 러프 분류 방법은 기존 정보시스템에서 복합 속성으로 기술된 오브젝트(objects)들을 분류하기 위해 확장한 방법이다. 즉 러프집합을 이용한 기존의 정보 시스템은 정보 시스템에 축적된 데이터가 완전하고 명확하다는 가정 하에서 오브젝트를 분류하고 있으며 기존의 방법들의 특징뿐만 아니라 데이터가 손실(missing) 되거나 불명확한 환경 하에서 오브젝트를 분류할 수 있도록 확장하였다. 또한 Jensen[11]에 의해서 제안된 방법에서는 패턴분류 문제에서 특징 선택(feature selection)의 주된 문제점인 최소 데이터 감축을 최적으로 결정할 수 있는 방법을 제안하였고, Tsai[12]은 패턴분류 규칙을 추출하기 위해 Minimization Entropy Principle Algorithm (MEPA)에 기반을 둔 퍼지-러프 패턴 분류 방법을 제안하였다. 지금까지 살펴본 방법들의 공통점은 불완전한 속성을 가진 데이터를 보다 정량적·정성적인 속성 값으로 분류하여 부정확하거나 알 수 없는 속성 값에 대체함으로써 분류에서 불완전성을 제거할 수 있는 방법과 규칙 감축 방법들을 제시하였다.

따라서 본 논문에서는 이러한 불완전성을 제거하기 위해 퍼지 규칙기반 시스템에서 규칙을 간소화하기 위한 방법으로 러프집합의 특성을 이용한 패턴분류 방법을 제안한다. 제안한 방법에서 러프집합은 규칙의 상관성을 고려하여 규칙간의 불필요한 속성을 제거하고 가장 중요한 속성인 코어가 되는 규칙을 추출하기 위한 도구로 사용되었다. 또한 모의실험에서는 제안한 방법의 타당성을 검증하기 위해 Matsuda[13], Ishibuchi[14], 그리고 Nozaki[15]에서 사용된 rice taste data를 기반으로 규칙 감축 전 퍼지 max-product 결과와 규칙 감축 후 퍼지 max-product 결과를 비교하였다.

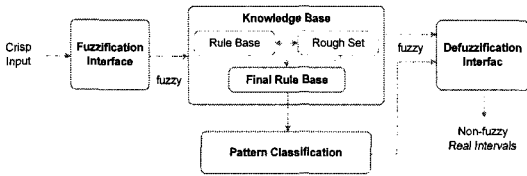
2장에서는 제안하는 패턴분류 방법을 설명하고, 3장에서는 패턴분류 방법의 타당성을 검증하

기 위해 rice taste data를 바탕으로 2가지 실험을 비교하였다. 그리고 4장에서는 결론에 대해서 논하였다.

2. 퍼지-러프집합을 기반으로 한 패턴분류 방법

2.1 패턴분류 시스템의 구조

퍼지-러프집합에 기반 한 패턴분류 시스템의 전반적인 구조는 다음과 같다.



(그림 2) 패턴분류를 위한 시스템의 전반적인 구조

(그림 1)에서와 같이 제안한 방법의 전반적인 구조는 4가지 모듈로 나뉜다. Fuzzification Interface 모듈은 사용자로부터 입력받은 수치적인 입력을 퍼지값으로 변환하는 역할을 하고, 퍼지값으로 변환된 입력은 다시 Knowledge Base 모듈로 들어간다. Knowledge Base 모듈은 기존의 Rule Base의 규칙을 Rough Set에 의해 불필요한 속성들을 감축한 후 Final Rule Base 갱신된다. Pattern Classification 모듈은 각 입력에 대해서 출력부 소속함수(규칙의 이행정도가 가장 큰 규칙)에 매칭된 결과 중에서 가장 큰 값을 가지는 속성을 추출한다. 마지막으로 Defuzzification Interface는 Final Rule Base와 Pattern Classification으로부터 얻은 정보로부터 비퍼지화된 실수값을 획득한다.

2.2 패턴분류 방법

패턴분류를 위해 사용된 IF-THEN 규칙의 구조와 분류 방법은 다음과 같다.

$$\text{Rule}^i : \text{IF } C_1 \text{ is } A \text{ and } C_2 \text{ is } B \dots C_n \text{ is } M \text{ THEN } E \text{ is } E \quad (1)$$

여기서, C_1, C_2, \dots, C_n 는 입력부 변수, E 는 출력부 변수이고, $A, B, M, \dots, E, (i=1, 2, \dots, n)$ 은 소속함수이다. 만약 규칙이 여러 개일 경우에는 각 규칙에 대한 관계를 계산하고 이것들을 하나로 결합하여 최종관계를 구하게 된다. 이렇게 구해진 관계에 입력 값을 입력하여 얻어진 출력을 비퍼지화하면 시스템에 대한 응답을 구할 수 있다.

[단계 1] 패턴분류를 위한 IF-THEN 규칙으로부터 식별가능 행렬을 계산한다.

$U = \{x_1, x_2, \dots, x_n\}$ 인 지식표현 시스템 $S = (U, A)$ 에서, $M(S)$ 로 표기되는 지식표현 시스템 S 의 식별가능 행렬(discernibility matrix)은 다음과 같이 $n \times n$ 행렬이다.

$$(c_{ij}) = \{a \in A : a(x_i) \neq a(x_j)\} \quad (2)$$

여기서 성분 $c_{ij}(i, j = 1, 2, \dots, n)$ 는 i 번째 입력 속성 x_i 와 j 번째 입력 속성 x_j 를 구별하는 모든 속성들의 집합이다.

[단계 2] 식별가능 행렬의 모든 단위 원소 엔트리의 집합으로부터 코어를 계산한다.

$$\text{CORE}(A) = \{a \in A : c_{ij} = (a)\}, \exists i, j \quad (3)$$

여기서 a 는 각각 i 와 j 번째 속성을 유일하게 구별할 수 있는 속성을 의미한다.

[단계 3] 모든 식별가능 행렬 $M(A)$ 로부터 식별가능 함수 $f(A)$ 를 계산한다.

$$f(A) = \prod_{(x,y) \in U^2} (\sum \delta(x,y) : (x,y) \in U^2 \text{ and } \delta(x,y) \neq 0) \quad (4)$$

여기서 $\sum \delta(x,y)$ 는 속성들의 집합 $\delta(x,y)$ 에 할당된 모든 부울 변수들의 집합과의 관계를 나타낸다.

[단계 4] 정의된 식별가능 행렬과 식별가능 함수 $f(A)$ 을 근거로 각 객체 x 에 대한 코어와 리덕트를 계산하기 위해서 식별가능 함수 $f^r(A)$ 을 계산한다.

$$f^r(A) = \prod_{y \in U} (\sum \delta(x,y) : y \in U \text{ and } \delta(x,y) \neq 0) \quad (5)$$

[단계 5] [단계 1]에서 [단계 4]를 통해 감축된 규칙을 바탕으로 주어진 입력에 대한 각 규칙의 입력부 적합도를 계산한다.

$$W_i = \mu_A(C_1) \wedge \mu_B(C_2) \wedge \dots \wedge \mu_m(C_m) \quad (6)$$

여기서 $W_i(i=1,2,\dots,m)$ 는 입력부 적합도를 나타낸다.

[단계 6] [단계 5]에서 계산한 적합도를 바탕으로 각 규칙의 이행정도(degree of fulfillment)를 계산한다.

$$\mu_{E_i}(E) = W_i \cdot \mu_E(E) \quad (7)$$

[단계 7] 여러 규칙의 이행정도 중에서 가장 큰 이행정도를 가지는 규칙의 속성을 패턴으로 분류한다.

$$S^i(E) = \begin{cases} \max_{i=1,2,\dots,m} \mu_{E_i}(E) & \mu_{E_j}(E) \neq \mu_{E_k}(E) \\ \emptyset & \mu_{E_j}(E) = \mu_{E_k}(E) \end{cases} \quad (8)$$

$j, k = 1, 2, \dots, m$

여기서 $S^i(E)$ 는 가장 큰 규칙의 이행정도를 가지는 규칙의 속성을 의미하고, $\mu_{E_j}(E) = \mu_{E_k}(E)$ 는 동일한 규칙에서 규칙의 이행정도의 같은 속성이 존재할 경우를 의미한다. 본 논문에서는 이러한 경우에 대해서 특정 패턴으로 분류할 수 없으므로 제외하였다.

[단계 8] 비퍼지화를 통해 확정치(실수 값)를 구한다. 비퍼지화 방법으로는 일반적으로 많이 사용되는 무게중심법을 사용하였다.

$$D = \frac{\sum_{i=1}^n (E_i \times \mu_{E_i}(E))}{\sum_{i=1}^n \mu_{E_i}(E)} \quad (9)$$

여기서 E_i 는 출력부 소속함수의 대집합을 의미한다.

3. 실험 및 결과분석

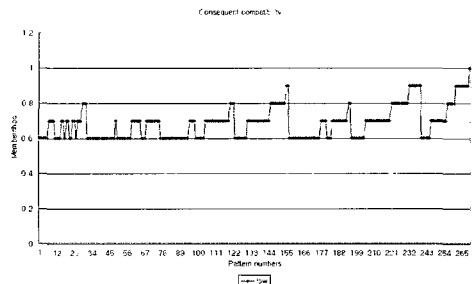
제안된 방법의 타당성을 검증하기 위하여

Matsuda[13], Ishibuchi[14], 그리고 Mozaki[15]에서 사용된 rice taste data를 기반으로 규칙감축 전 퍼지 max-product 결과와 규칙감축 후 퍼지 max-product 결과를 비교하였다. 규칙은 rice taste data를 분류하기 위한 IF-THEN 규칙과 각 입력 변수와 그에 대응하는 소속함수를 구성하였다.

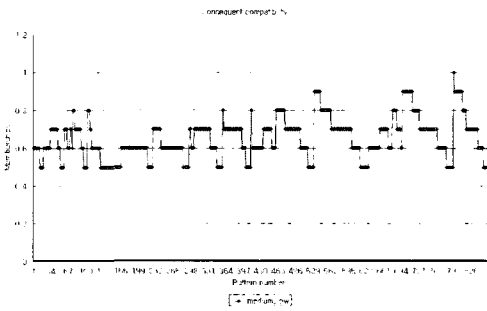
여기서 $x_1, x_2, x_3, x_4, x_5, y$ 는 5개의 입력변수와 1개의 출력변수를 나타내고, 각 속성의 의미는 Flavor, Appearance, Taste, Stickiness, Toughness, Evaluation을 나타낸다. 또한 5개의 각 입력변수 Flavor, Appearance, Taste는 각각 2가지 레벨(Good/Bad)로, Stickiness는 2가지 레벨(Sticky/Not Sticky)로, Toughness는 2가지 레벨(Tough /Tender)로 나누었고, 1개의 출력변수 Evaluation은 5가지 레벨(High/Medium High/Medium/Low/Low)로 나누었다.

3.1 속성감축 전 퍼지 max-product에 의한 패턴 분류

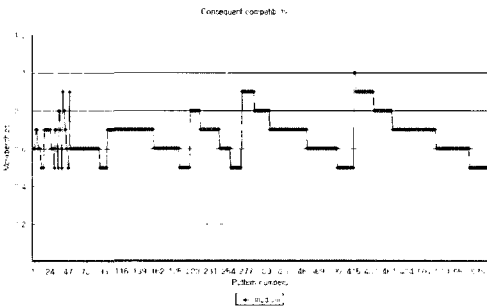
먼저 32개 규칙을 기반으로 속성감축 전 max-product에 의한 패턴 분류를 실험하였다. 실험에 사용된 패턴 분류 방법은 [단계 5]에서 [단계 7]의 식을 사용하였다. 다음 그림은 5개의 입력변수에서 2077개의 입력패턴에 대한 [단계 5]에서 [단계 7]을 적용하였을 때 입력 값에 따른 출력부 소속함수의 최대 적합도로 분류한 rice taste data의 분류 패턴(Low, Medium Low, Medium, Medium High, High)을 나타낸다.



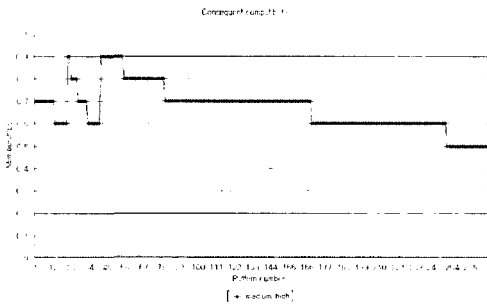
[그림 2] 속성감축 전 출력부 소속함수 Low의 적합도 패턴



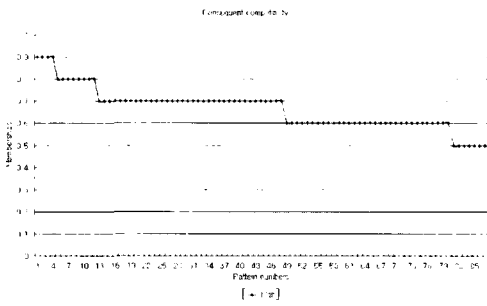
(그림 3) 속성감축 전 출력부 소속함수 Medium Low의 적합도 패턴



(그림 4) 속성감축 전 출력부 소속함수 Medium의 적합도 패턴



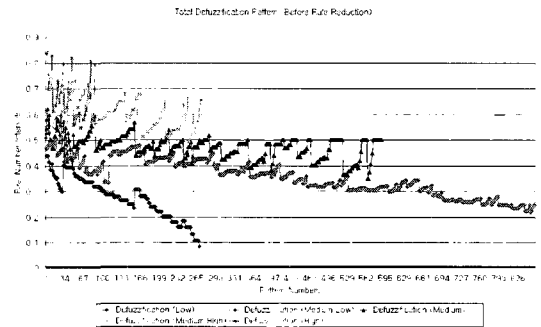
(그림 5) 속성감축 전 출력부 소속함수 Medium High의 적합도 패턴



(그림 6) 속성감축 전 출력부 소속함수 High의 적합도 패턴

(그림2-6)에서 Pattern numbers 항목은 각 출력 패턴으로 분류된 입력 패턴들의 수를 나타내고 소속함수는 각 입력 패턴에 대해서 사상된 출력부 소속함수의 적합도를 나타낸다. 즉 출력부 소속함수의 각 적합도는 식 (8)에 의해서 규칙의 이행정도 중에서 가장 큰 이행정도를 가지는 규칙의 속성으로 분류한 것을 의미한다.

다음은 2077개의 입력 패턴에 대해서 [단계 8]의 무게중심법을 이용하여 나타낸 전체 비퍼지화 패턴(실수 구간)을 보여준다.



(그림 8) 규칙 감축 전 전체 비퍼지화 패턴

(그림 7)에서 Pattern Numbers는 전체 입력 패턴 수를 나타내고, Real Number Intervals는 비퍼지화 된 실수 구간들을 의미한다. 예를 들어 Defuzzification(Low)의 비퍼지화 된 최소-최대의 실수구간은 대략 0.08에서 0.43정도임을 볼 수 있고 Low로 분류된 입력 패턴들의 수는 약 280개 이내임을 볼 수 있다. 따라서 Defuzzification(Low)의 경우와 동일한 방법으로 각 출력부 소속함수의 비퍼지화 된 최소-최대 실수 구간을 분석한 결과 (표 1)과 같음을 알 수 있었다.

(표 1) 속성감축 전 각 패턴의 비퍼지화 된 실수 구간

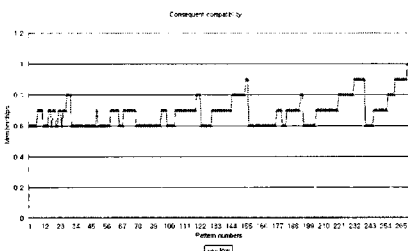
분류 패턴	실수 구간	패턴 수
Low	[0.082 - 0.439]	270개
Medium Low	[0.217 - 0.506]	856개
Medium	[0.351 - 0.622]	589개
Medium High	[0.494 - 0.741]	275개
High	[0.547 - 0.843]	87개

3.2 속성감축 후 퍼지 max-product에 의한 패턴 분류

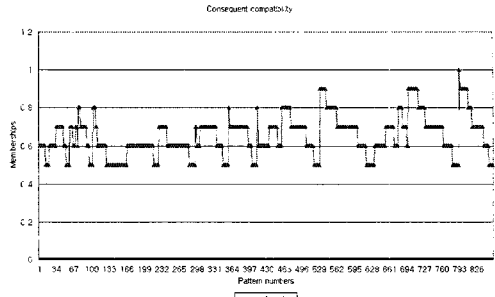
본 절에서는 3.1절의 실험과 동일하게 32개 규칙을 기반으로 [단계 2]에서 [단계 4]까지 적용하여 속성을 감축하였고 [단계 5]에서 [단계 8]을 이용하여 패턴 분류를 실험하였다. 실험에서 사용된 총 입력패턴의 수는 3.1절과 동일한 2077개를 사용하였다. (표 2)는 rice taste data의 분류 규칙에 [단계 2]에서 [단계 4]를 적용한 후 생성된 퍼지 규칙을 나타낸다. (표 2)에서 ‘.’은 [단계 2]에서 [단계 4]를 통해 감축된 속성, 즉 불필요한 속성이 제거된 것을 의미하고, 음영이 들어간 부분은 속성감축 후 생성된 둘 이상의 규칙들을 의미한다.

본 논문에서는 음영이 들어간 규칙 중 다음과 같은 형태로 규칙을 선택하였다 : Rule 5-2, Rule 6-2, Rule 9-1, Rule 11-1, Rule 16-2, Rule 18-2, Rule 19-3, Rule 21-3, Rule 22-1, Rule 24-2, Rule 25-2, Rule 26-2, Rule 28-2, Rule 29-2, Rule 30-1, Rule 31-1.

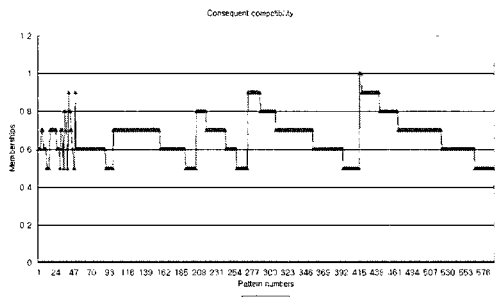
(표 3)은 중복된 규칙이 제거된 최종 rice taste data를 분류하기 위한 IF-THEN 규칙을 나타낸 것이고 이 규칙을 기초로 [단계 5]에서 [단계 7]을 수행하였을 때 출력부 소속함수의 적합도 패턴을 (그림 8)에서 (그림 12)로 나타냈다. (그림 2)에서 (그림 6)까지의 속성감축 전 출력부 소속함수에서 최대 적합도를 가지는 속성의 출력 패턴과 비교해 볼 때 속성을 감축하더라도 적합도의 출력 패턴이 정확히 일치하였다.



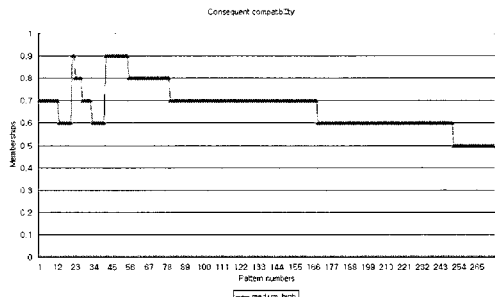
(그림 9) 속성감축 후 출력부 소속함수 Low의 적합도 패턴



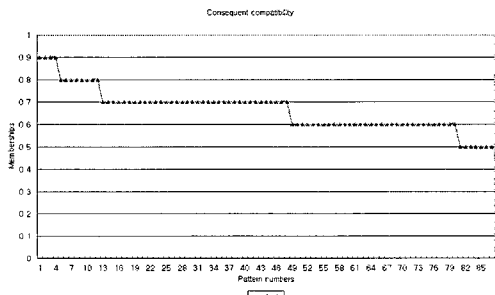
(그림 10) 속성감축 후 출력부 소속함수 Medium Low의 적합도 패턴



(그림 11) 속성감축 후 출력부 소속함수 Medium의 적합도 패턴



(그림 12) 속성감축 후 출력부 소속함수 Medium High의 적합도 패턴



(그림 13) 속성감축 후 출력부 소속함수 High의 적합도 패턴

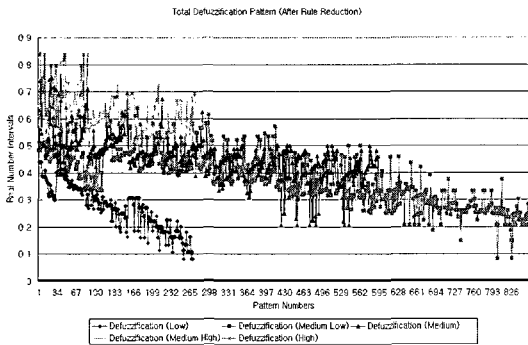
(표 2) 속성감축 후 생성된 IF - THEN 규칙

Rule	Input					Output	Rule	Input					Output
	x_1	x_2	x_3	x_4	x_5	y		x_1	x_2	x_3	x_4	x_5	y
1	G	G	G	S	-	H	19-2	B	G	G	NS	-	M
2	G	G	G	S	-	H	19-3	-	G	G	NS	To	M
3	G	B	G	S	To	MH	20	-	G	B	-	-	M
4	B	-	G	S	Te	MH	21-1	B	G	-	S	To	M
5-1	B	-	G	S	Te	MH	21-2	B	G	G	-	To	M
5-2	-	B	G	S	Te	MH	21-3	B	-	G	S	To	M
6-1	G	B	-	S	Te	MH	22-1	G	G	B	-	-	M
6-2	-	B	G	S	Te	MH	22-2	G	G	-	NS	To	M
7	G	G	G	NS	Te	MH	23	-	B	G	NS	-	ML
8	G	B	-	N	Te	MH	24-1	G	B	-	NS	To	ML
9-1	G	G	B	-	-	M	24-2	-	B	G	NS	TO	ML
9-2	-	G	B	S	-	M	25-1	G	B	B	NS	-	ML
9-3	G	-	B	S	To	M	25-3	-	B	B	NS	Te	ML
10	G	G	B	-	-	M	26-1	B	G	B	NS	-	ML
11-1	G	G	B	-	-	M	26-2	B	G	B	-	To	ML
11-2	-	G	B	S	-	M	27	B	B	B	S	To	ML
12	B	-	B	S	Te	M	28-1	B	G	B	-	-	ML
13	G	B	G	NS	Te	M	28-2	B	-	B	NS	Te	ML
14	G	-	B	S	To	M	29-1	B	B	-	NS	Te	ML
15	B	-	G	S	To	M	29-2	B	-	B	NS	Te	ML
16-1	G	G	-	NS	To	M	29-3	-	B	B	NS	Te	ML
16-2	-	G	G	NS	To	M	30-1	B	B	G	NS	-	ML
17	B	G	G	NS	-	M	30-2	B	B	-	NS	Te	ML
18-1	B	G	B	S	-	M	31-1	G	B	B	NS	-	ML
18-2	B	-	B	S	Te	M	31-2	G	B	-	NS	To	ML
18-3	-	G	B	-	Te	M	32	B	B	B	NS	To	L
19-1	B	G	G	NS	-	M							

(표 3) 최종 선택된 IF - THEN 규칙

Rule	Input					Output	Rule	Input					Output
	x_1	x_2	x_3	x_4	x_5	y		x_1	x_2	x_3	x_4	x_5	y
1	G	G	G	S	-	H	10	B	G	G	NS	-	M
2	G	B	G	S	To	MH	11	-	G	B	-	-	M
3	B	-	G	S	Te	MH	12	-	B	G	NS	-	ML
4	-	B	G	S	Te	MH	23	-	B	G	NS	To	ML
5	G	B	-	S	Te	MH	14	B	G	B	-	To	ML
6	G	G	B	-	-	M	15	B	B	B	S	To	ML
7	B	-	B	S	Te	M	16	B	-	B	NS	Te	ML
8	G	B	G	NS	Te	M	17	B	B	G	NS	-	ML
9	-	G	G	NS	To	M	18	B	B	B	NS	To	L

(그림 13)은 속성감축 후 2077개의 입력 패턴에 대해서 규칙감축 전 전체 비퍼지화 패턴과 동일한 방법으로 비퍼지화 한 전체 출력 패턴을 나타낸 것이다.



(그림 14) 규칙감축 후 전체 비퍼지화 패턴

(그림 13)의 비퍼지화 된 패턴들의 실수구간을 분석한 결과 각 패턴의 실수 구간은 (표 4)와 같다.

(표 4)의 속성감축 후 비퍼지화 된 실수 구간과 규칙의 속성감축 전 비퍼지화 된 실수 구간을 비교해 볼 때 속성감축 후 Medium Low와 Medium 구간에서 다소 차이를 보이지만 속성감축 전 실수 구간과 유사하였다. 또한 속성감축 전 패턴분류에서의 출력부 소속함수의 적합도 패턴과 속성감축 후 출력부 소속함수의 적합도 패턴이 정확히 일치하였다(그림 2-6과 그림 8-12).

(표 4) 속성감축 후 각 패턴의 비퍼지화 된 실수 구간

분류 패턴	실수 구간	패턴 수
Low	[0.082 - 0.439]	270개
Medium Low	[0.082 - 0.639]	856개
Medium	[0.207 - 0.752]	589개
Medium High	[0.494 - 0.741]	275개
High	[0.550 - 0.843]	87개

따라서 초기에 규칙 베이스를 러프집합을 이용

하여 속성을 감축함으로써 보다 적은 수의 규칙을 사용하여 패턴을 분류하더라도 정확하게 패턴을 분류할 수 있었다.

4. 결론

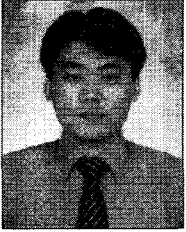
본 논문에서는 퍼지 규칙기반 시스템에서 규칙을 최적화하기 위한 방법으로 러프집합의 특성을 이용한 추론 방법을 제안하였다. 제안한 방법에서 러프집합은 초기의 규칙 집합에서 규칙의 상관성을 고려하여 규칙간의 불필요한 속성을 제거하고 코어가 되는 규칙을 추출하기 위해 사용되었다. 그리고 제안된 방법의 타당성을 보이기 위해 rice taste data를 기반으로 2077개의 입력 패턴에 대해서 출력부 소속함수의 최대 적합도를 가지는 속성들을 추출하여 속성감축 전과 후의 출력 패턴을 비교하였다. 비교 결과 속성감축 전과 후의 최대 적합도를 가지는 속성이 모두 동일한 패턴으로 출력됨을 볼 수 있었다. 또한 속성감축 전과 후의 비퍼지화 된 실수 구간을 비교함으로써 제안된 방법의 타당성을 재검증하였다.

참고 문헌

- [1] L.A.Zadeh. The concept of a linguistic variable and its application to approximate reasoning, Part I, II, III. pp. 199-249, pp. 301-357, pp. 43-80, 1975.
- [2] C.M.Moraga, J.Canas, R. Monge et al. Parallel Processing of Fuzzy Inference. In Proceeding of 24th International Symposium on Multiple-Valued Logic, pp. 134-139, 1994.
- [3] M.A.Eshera and S.C.Barash. Parallel Rule-Based Fuzzy Inference on Mesh-Connected Systolic Arrays. IEEE Expert, pp. 27-35, 1989.
- [4] C. Chantrapornchai et al. Minimization of Fuzzy Systems Based on Fuzzy Inference

- Graphs. Technical Report no. TR-96-06.
- [5] A. Chandramohan and M. V. C. Rao, A Novel Approach for Combining Fuzzy Rules Using Mean Operators for Effective Rule Reduction, *Soft Computing - A Fusion of Foundation, Methodologies and Applications*, Vol.10, No.11, pp. 1103-1108, 2006.
- [6] H. Ishibuchi, T. Murata and I.B. Turksen, Single-Objective and Two-Objective Genetic Algorithms for Selecting Linguistic Rules for Pattern Classification Problems, *Fuzzy Sets and Systems*, Vol.89, pp. 135-150, 1997.
- [7] H. Ishibuchi, T. Nakashima and T. Morisawa, Voting in Fuzzy Rule-Based Systems for Pattern Classification Problems, *Fuzzy Sets and Systems*, Vol.103, pp. 223-238, 1999.
- [8] E. Orłowska. *Incomplete Information : Rough Set Analysis*. A Springer-Verlag Company. pp. 23-57, pp. 95-108. 1997.
- [9] Z. Pawlak, *Rough Sets : Theoretical Aspects of Reasoning about Data*, Kluwer Academic Publishing, Dordrecht, 1991.
- [10] Slowinski, R. and Stefanowki, J. *Rough Classification in Incomplete Information Systems*, *Mathematical and Computer Modeling*, Vol.12, No.10/11, pp. 1347-1357, 1989.
- [11] R. Jensen and Q. Shen. *Fuzzy-Rough Data Reduction with Ant Colony Optimization*, *Fuzzy Sets and Systems*, Vol.149, pp. 5-20, 2005.
- [12] Y.C Tsai, C.H Cheng and J.R Chang, Entropy-Based Fuzzy Rough Classification Approach for Extracting Classification Rules, *Expert Systems with Applications*, Vol.31, pp. 436-443, 2006.
- [13] M. Matsuda and T. Kameoka, "Application of Fuzzy Measure, Fuzzy Integral and Neural Network To The System Which Estimate Taste By Using Industrial Analysis", *Proc. IIZUKA' 92*, pp. 601 - 606, 1992.
- [14] H. Ishibuchi, K. Nozaki, H. Tanaka, Y. Hosaka and M. Matsuda, "Empirical Study On Learning In Fuzzy Systems By Rice Taste Analysis", *Fuzzy Sets and Systems*, Vol. 64, pp. 129 - 144, 1994.
- [15] K. Nozaki, H. Ishibuchi and H. Tanaka, "A Simple But Powerful Heuristic Method For Generating Fuzzy Rules From Numerical Data", *Fuzzy Sets and Systems*, Vol. 86, pp. 251 - 270, 1997

◎ 저 자 소개 ◎



손 창 식

2000년 대구가톨릭대학교 전자정보공학부 졸업(학사)
2002년 대구가톨릭대학교 대학원 전산통계학과 졸업(석사)
2006년 대구가톨릭대학교 대학원 전산통계학과 졸업(박사)
2004년 ~ 현재 대구가톨릭대학교 컴퓨터정보통신공학부 강의전담교수
관심분야 : 패턴분류, 유전자알고리즘, 퍼지파라미터최적화, 감성공학
E-mail : fuzzyrisk@paran.com



김 두 완

1997년 경남대학교 수학과 졸업(학사)
1999년 대구가톨릭대학교 대학원 전산통계학과 졸업(석사)
2003년 대구가톨릭대학교 대학원 전산통계학과 졸업(박사)
동방대학원대학교 문화정보학과 교수
관심분야 : 퍼지, 신경망, 다치, 유비쿼터스시스템
E-mail : dooywan@nate.com