DOI: 10.3745/KIPSTB.2010.17B.2.163

표준화된 매개변수 소속함수에 기반을 둔 언어적 케이스 기반 퍼지 추론

최 대 영†

요 약

표준화된 매개변수 소속함수에 기반을 둔 언어적 케이스 기반 퍼지 추론 방법을 제안한다. 제안된 방법은 선형 시간 복잡도를 갖는 퍼지 추론을 위한 효율적인 방법을 제공한다. 결과적으로 제안된 방법은 퍼지 추론의 속도를 개선하는데 사용될 수 있다.

언어적 케이스 기반 퍼지 추론 과정에서 표준화된 매개변수 소속함수에 기반을 둔 언어적 케이스 색인과 검색 방법을 제시한다. 이는 기존의 언어 근사 방법과 비교할 때 상대적으로 빠르게 계산될 수 있다. 공학적인 관점에서 이는 가치 있는 장점이 될 수 있다.

키워드 : 언어적 케이스 기반 퍼지 추론(LCBFR), 표준화된 매개변수 소속함수(SPMF), 언어적 케이스 색인과 검색

A Linguistic Case-based Fuzzy Reasoning based on SPMF

Dae-Young Choi†

ABSTRACT

A linguistic case-based fuzzy reasoning (LCBFR) based on standardized parametric membership functions (SPMF) is proposed. It provides an efficient mechanism for a fuzzy reasoning within linear time complexity. Thus, it can be used to improve the speed of fuzzy reasoning.

In the process of LCBFR, linguistic case indexing and retrieval based on SPMF is suggested. It can be processed relatively fast compared to the previous linguistic approximation methods. From the engineering viewpoint, it may be a valuable advantage.

Keywords : Linguistic Case-based Fuzzy Reasoning (LCBFR), Standardized Parametric Membership Functions (SPMF), Linguistic Case Indexing and Retrieval

1. 서 론

케이스 기반 추론(Case-Based Reasoning; CBR)은 문제를 해결하기 위해 과거의 경험을 활용하는 방법이다. 일반적으로 이는 복잡하고 비구조적인 의사결정 문제를 효율적으로 처리하는데 자주 사용된다. 또한 케이스 베이스가 실시간으로 변경될 수 있는데 이는 현실 문제에 적용할 때 매우 중요하다. 이러한 장점 때문에 CBR은 생산, 금융, 마케팅, 인터넷 쇼핑몰, 항공 운항 제어에서 충돌 해결, 반도체 설계, 의료 진단 등을 포함한 다양한 문제 해결에 적용되고 있다 [1, 4-5, 12]. 일반적으로 주요 인공지능 기법들이 문제와 결론 사이의 일반화된 관계에 의존하는 반면에 CBR은 과거의 경험과 주어진 문제 상황에 대한 특정 지식을 사용한다. CBR의 성능은 Case 색인과 Case 검색에 의존하기 때문에 이들의 과정이 가장 중요하다. CBR시스템은 Case-Base에서 가장 유사한 케이스들을 검색한다. 발견된 가장 유사한 케이스들은 문제 해결의 질에 영향을 준다. 결과적으로 효율적인 Case 색인과 Case 검색 방법을 설계하는 것이 매우 중요하다. CBR은 복잡하고 비구조적인 문제에 효과적이고 Case-Base는 변경하기도 쉽다[1]. 최근에 CBR 은 많은 관심을 받았고 다양한 분야에 성공적으로 사용되었다. 그러나 불확실성과 모호함(Fuzziness)을 내포하고 있는 실세계의 다양한 응용 분야에 적용하기 위해서 CBR은 불확실(Uncertain), 불완전(Incomplete), 모호한 정보를 다룰 수 있어야 한다. 이러한 관점에서 CBR과 퍼지 추론[15-16, 18, 23]을 결합하면 실세계의 다양한 응용 분야의 적용에 효과적으로 이용될 수 있다. 퍼지 추론은 다음과 같은 형태로 표현 된다.

$$\begin{array}{ll} \text{규칙 :} & \textit{IF}\ X \text{ is } A\ \textit{THEN}\ Y \text{ is } B \\ \text{관찰 :} & \underline{X \text{ is } A' \qquad\qquad\qquad} \\ \text{결론 :} & \qquad\qquad Y \text{ is } B' \end{array} \tag{1}$$

† 정 회 원 : 유한대학 경영정보과 교수
논문접수 : 2009년 10월 20일
수 정 일 : 1차 2009년 12월 18일
심사완료 : 2009년 12월 18일

이때 X와 Y는 언어 변수(Linguistic variable)이고 A, B, A'는 언어 값(Linguistic value)이다. 이러한 언어 값들은 퍼지 부분 집합으로 표현될 수 있다[22]. 프로그램 언어에서 변수에 값이 배정되는 것처럼 인간이 사용하는 언어의 언어 변수에는 언어 값이 배정될 수 있다. 예를 들어 언어 변수인 '나이'에는 언어 값으로 '매우 젊다', '젊다', '늙었다' 등이 배정될 수 있다. 일반적으로 규칙베이스에 있는 규칙에 관찰된 결과를 적용하여 추론 결과를 만들기 위해 Truth value restriction [21], Compositional rule of inference(CRI) [20, 23], Analogical reasoning[9-11], Approximate analogical reasoning schema (AARS) [18] 방법이 주로 사용된다. 이 중 CRI가 가장 많이 사용되고 있는데 이는 행렬 연산에 기반을 두고 있다. 그러나 소속함수 값 전파에 대한 행렬 연산의 효과는 개념적으로 명확하지 않다[3, 15-16]. 또한, Zadeh[21]는 퍼지 조건 명제('*IF X is A THEN Y is B*')를 퍼지 관계로 변환하기 위해 산술 규칙이라는 변환 규칙을 제안하였다. 산술 규칙을 사용하는 CRI의 주요한 문제점은 Modus ponens를 만족하지 못한다는 것이다. 즉, A'=A일 때, 추론된 B'가 다음과 같이 계산된다 : $\mu_{B'} = (1+\mu_B)/2 \neq \mu_B$. 이러한 추론 결과는 산술 규칙을 사용하는 CRI가 Modus ponens (즉, $(A\wedge(A\rightarrow B))\rightarrow B$)를 만족하지 못한다는 것을 나타낸다. 또한, 관찰(A')과 규칙베이스의 조건(A)간에 미결정 부분이 발생한다. 이는 규칙베이스의 조건(A)의 소속함수 값이 0인 범위가 관찰(A')의 소속함수 값이 0인 범위를 포함할 때 발생한다[3]. AARS [18]는 추론을 하기 위한 유사 측도(Similarity measure; SM)를 거리 측도(Distance measure; DM)를 사용해서 계산한다. 이때 두 퍼지 집합간의 유사 측도는 다음과 같이 정의 되었다 : $SM=(1+DM)^{-1}$ (단, $SM\in[0,1]$). 그러나 DM를 어떻게 계산하는가에 대해서는 정의되어 있지 않다. 또한, 퍼지 추론에서 중요한 역할을 하는 조정 함수(Modification function; MF)가 명확히 정의되어 있지 않고 주관적으로 처리된다. 이러한 문제점을 다루기 위해 표준화된 매개변수 소속함수(Standardized parametric membership functions; SPMF)에 기반을 둔 언어적 케이스 기반 퍼지 추론(LCBFR) 방법을 제안한다.

2. SPMF

소속함수는 퍼지 집합의 중요한 구성 요소이다. 퍼지 집합을 표현하기 위한 다양한 소속함수들이 제시되었다[6, 18-20]. 소속함수는 크게 매개변수 소속함수(Parametric membership functions; PMF)와 비매개변수 소속함수 (Non-parametric membership functions; NPMF)로 분류할 수 있다. 퍼지 시스템 개발에 관련된 많은 응용에서 매개변수 소속함수는 중요하고 이는 퍼지 시스템의 표준화를 가능하게 한다. 기술을 관리하는 핵심 요소중의 하나는 표준화 (예 : 인터넷 통신 프로토콜인 TCP/IP) 이다. 기술의 표준화 없이는 관리가 어렵게 된다. 기술 표준화는 패케지나 응용 서비스들이 단편적인 해결이 되지 않도록 한다. 또한, 표준화는 비표준화 보다 변화에 보다 빠르게 적응할 수 있게 한다 [17]. Mamdani [14]는 퍼지 집합의 표준화를 고려할 때라고 하였다. 그러나, 퍼지 집합의 표준화에 대한 연구는 미미한 실정이다. 주요한 이유는 퍼지 집합의 소속함수가 주로 개인의 주관성에서 만들어지기 때문이다. 언어 값을 사용하는 개인간에 그 언어 값에 대한 개념적인 차이가 명확히 존재하지만 개인간의 비교를 위한 표준을 만들기 위해서는 표준화된 소속 함수의 사용이 요구된다. 일반적으로 기존의 소속함수는 임의적인 (Ad-hoc) 방법을 사용해서 개인의 주관성에 기반을 두고 만들어 지므로 다양한 현상을 표현하기 위해 많은 소속함수가 필요하다. 만약 퍼지 시스템을 만들기 위해 적은 수의 소속함수를 사용하려면 소속함수가 적응적이 되어야 하는데 이는 매개변수를 사용하면 된다. 퍼지 집합을 표현하기 위해 조정 가능한 매개변수를 갖는 표준화된 소속함수를 사용하면 편리하다. 조정 가능한 매개변수를 갖는 표준화된 소속함수의 전형적인 형태로 삼각형 형태, 사다리꼴 형태, S 형태, Π 형태 등이 있다[20]. 언어 값을 표현하기 위한 퍼지 집합을 A라 하고 이는 전체 집합 X의 부분 집합이라 하자. $x\in X$에 대해 삼각형 형태 소속함수는 3점을 사용해서 표현될 수 있다. 즉, $\mu_A(x_L, x_M, x_H)$ (단 $x_L<x_M<x_H$). 만약 소속함수의 결과가 구간 [0, 1]로 정규화 된다면 $x\in[-\infty, x_L]\cup[x_H, \infty]$일 때 $\mu_A(x_L, x_M, x_H) = 0$ 이고 x_M 에서 $\mu_A(x_L, x_M, x_H) = 1$ 이 된다. 한편, 사다리꼴 형태 소속함수는 4점을 사용해서 표현될 수 있다. 즉, $\mu_A(x_L, x_{I1}, x_{I2}, x_H)$ (단, $x_L<x_{I1}<x_{I2}<x_H$). 만약 소속함수의 결과가 구간 [0, 1]로 정규화 된다면 $x\in[-\infty, x_L]\cup[x_H, \infty]$에서 $\mu_A(x_L, x_{I1}, x_{I2}, x_H) = 0$ 이고 $[x_{I1}, x_{I2}]$에서 $\mu_A(x_L, x_{I1}, x_{I2}, x_H) = 1$ 이 된다. 조정 가능한 매개변수를 갖는 표준화된 소속함수의 전형적인 형태에 대한 보다 자세한 내용은 [6, 20-21]에 기술되어 있다. 한편, 비매개변수 소속함수(NPMF)를 SPMF로 변환해 주기 위해 근사 변환 방법(Approximate transformation method; ATM) [6]이 제안되었다. 변환된 SPMF에서 매개변수는 퍼지 집합의 구조를 결정하는 특징점이라고 할 수 있다. 이러한 매개변수들을 사용해서 퍼지 집합간의 차이(Distance)를 효율적으로 계산할 수 있다.

3. SPMF에 기반을 둔 언어적 케이스 기반 퍼지 추론

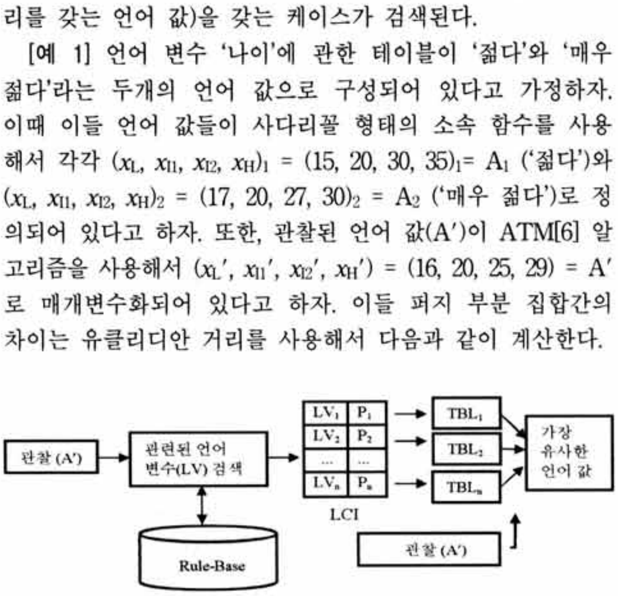
3.1 SPMF에 기반을 둔 언어적 케이스 색인과 검색

CBR의 성능은 Case 색인과 Case 검색에 의존하기 때문에 이들의 과정이 가장 중요하다[1]. 이 논문에서는 SPMF에 기반을 둔 언어적 Case 색인과 Case 검색 방법을 제시한다. Rule-Base 에 있는 언어 변수에 관련된 언어 값 중에서 관찰된 언어 값(A')에 가장 유사한 언어 값을 효율적으로 찾는 것을 언어 근사라 한다. 언어적 케이스 색인(Linguistic Case Index; LCI)에 언어 변수들이 정렬되어 있다고 가정한다. LCI에 있는 각 언어 변수는 관련된 언어 값들로 구성된 테이블에 관한 포인터를 가지고 있다. 즉, LCI는 관련된 언어 값들을 빠르게 발견하기 위한 색인 테이블이라고 할 수 있다. 이와 같이 언어 근사 과정에서 어떤 언

어 변수에 관련없는 언어 값들에 대한 탐색을 방지하기 위해 언어 변수들을 분리하는 분할 방법을 사용하였다. 결과적으로 각 테이블은 해당 언어 변수에 관련된 언어 값들로만 구성된다. 예를 들어, 언어 변수 '나이'에 관한 테이블은 '젊다', '매우 젊다' 등으로 구성될 수 있다. 이때 언어 값을 표현하는데 사용되는 퍼지 부분 집합[22]은 2장에서 제시된 SPMF인 삼각형 형태나 사다리꼴 형태 소속함수 등을 사용하여 정의된다고 가정한다.

제안하고자 하는 언어적 Case 색인과 Case 검색 방법에서는 (그림 1)에서 처럼 LCBFR 과정에서 사용된 언어 변수들을 분리하기 위해 언어 변수를 테이블 별로 분할한다. 이는 LCBFR 과정에서 관련없는 언어 값을 탐색하는 것을 피할 수 있게 한다. 결과적으로 식 (1)과 같이 관찰된 언어 값(A′)은 언어 변수(X)와 관련된 것이므로 Rule-Base에서 관련된 언어 변수를 참조하면 해당 언어 값들을 쉽게 찾을 수 있다. 관찰된 언어 값(A′)에 관련된 언어 변수(LV)가 LCI에서 결정된 후 그에 관련된 테이블(TBL)은 해당 언어 변수의 포인터(P_i, i = 1, 2, …, n)를 사용해서 얻어질 수 있다. 각 테이블(TBL_i, i = 1, 2, …, n)은 SPMF로 표현된 퍼지 부분 집합(즉, 언어 값)으로 구성되어 있다. 예를 들어 삼각형 형태나 사다리꼴 형태 소속함수는 (x_L, x_M, x_H)과 (x_L, x_{l1}, x_{l2}, x_H)으로 각각 표현될 수 있다. SPMF에서 매개변수는 퍼지 집합의 구조를 결정하는 특징점이라고 할 수 있다. 이러한 매개변수들을 사용해서 퍼지 집합간의 차이를 효율적으로 계산할 수 있다. 테이블에 있는 모든 퍼지 부분 집합과 관찰된 언어 값(A′)이 SPMF를 사용하여 정의된다고 가정하자. 만약 관찰된 언어 값(A′)이 비매개변수 소속함수로 되어 있다면 ATM[6]을 사용해서 SPMF로 표현된 퍼지 집합으로 변환될 수 있다. 본 논문에서는 두 퍼지집합 간의 차이를 SPMF의 매개변수를 사용하여 유클리디안 거리를 계산한다. 즉, 해당 언어 변수에 관련된 테이블에 있는 모든 퍼지 부분 집합과 관찰된 언어 값(A′)간의 유클리디안 거리가 계산된다. 결과적으로 가장 유사한 언어 값 (즉, 최소 거리를 갖는 언어 값)을 갖는 케이스가 검색된다.

[예 1] 언어 변수 '나이'에 관한 테이블이 '젊다'와 '매우 젊다'라는 두개의 언어 값으로 구성되어 있다고 가정하자. 이때 이들 언어 값들이 사다리꼴 형태의 소속 함수를 사용해서 각각 $(x_L, x_{l1}, x_{l2}, x_H)_1$ = $(15, 20, 30, 35)_1$= A_1 ('젊다')와 $(x_L, x_{l1}, x_{l2}, x_H)_2$ = $(17, 20, 27, 30)_2$ = A_2 ('매우 젊다')로 정의되어 있다고 하자. 또한, 관찰된 언어 값(A′)이 ATM[6] 알고리즘을 사용해서 $(x_L', x_{l1}', x_{l2}', x_H')$ = (16, 20, 25, 29) = A′로 매개변수화되어 있다고 하자. 이들 퍼지 부분 집합간의 차이는 유클리디안 거리를 사용해서 다음과 같이 계산한다.

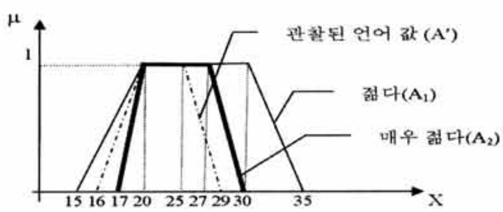


(그림 1) 언어적 케이스 색인과 검색

$d_1(A_1, A')$
$$= \sqrt{(X_L - X_L')^2 + (X_{l1} - X_{l1}')^2 + (X_{l2} - X_{l2}')^2 + (X_H - X_H')^2}$$
$$= \sqrt{(15-16)^2 + (20-20)^2 + (30-25)^2 + (35-29)^2} = \sqrt{62}$$
$d_2(A_2, A')$
$$= \sqrt{(17-16)^2 + (20-20)^2 + (27-25)^2 + (30-29)^2} = \sqrt{6}$$

A′는 $d_1 > d_2$이기 때문에 언어 근사로 언어 값 '매우 젊다'가 선택된다. 이 경우 규칙베이스에서 규칙 (*'IF '나이' is '매우 젊다' THEN Y is B'*)가 선택된다. 비슷한 방법으로 m개의 언어 값들로 구성된 테이블인 경우 유클리디안 거리를 반복 적용함에 의해 (d_1, d_2, …, d_m)를 계산하고 min(d_1, d_2, …, d_m)인 언어 값이 언어 근사로 선택된다.

제안된 방법은 SPMF의 매개변수를 특징점으로 사용하여 계산된다. 또한, 제안된 언어적 Case 색인과 검색 방법은 언어 근사시에 관련 없는 언어 값들에 대한 탐색을 방지하기 위해 (그림 1)처럼 언어 변수들을 분리하는 분할 방법을 사용하였다. 이는 LCBFR 과정에서 관련 없는 언어 값을 탐색하는 것을 피할 수 있게 한다. 결과적으로 제안된 언어 근사 방법은 이전의 연구[2, 7-8, 13, 19]와 비교할 때 처리 속도 측면에서 장점을 갖는다. 한편 [예 1]처럼 언어근사에 의해 어떤 규칙(Rule)이 선택되면 퍼지추론을 하기 위해 두 퍼지 집합 A(규칙의 조건부분)와 A′(관찰) 간의 거리 측도를 계산하고 거리 측도는 조정함수를 계산하는데 이용된다. 조정함수는 결론을 추론할 때 이용된다.



(그림 2) 언어 변수 '나이'에 관한 퍼지 부분 집합의 예

3.2 거리 측도 (DM)

인간 행동 실험의 결과로 Zwick [24] 등은 두 개의 퍼지 집합 A와 B간의 거리 측도로 사용하기 좋은 5 가지(즉, S_4, q_∞, q_*, Δ_∞, Δ_*)를 권고 하였다. 이들이 권고한 거리 측도는 평균 등의 방법을 사용한다기 보다는 하나의 값에 초점을 두고 있다. 예를 들어 S_4의 경우에는 A∩B의 소속함수가 최대값인 상태에서의 특정 값에 초점을 두고 있고, q_∞와 Δ_∞의 경우 두 퍼지 집합간 거리가 최대인 상태에서의 α-level set [20, 21]에 초점을 두고 있다. 그들의 실험 결과에 의하면 복잡한 소속함수를 어떤 특정 값으로 단순화 하는 것이 인간이 퍼지 집합을 처리할 때의 자연스런 방법이라는 것을 알 수 있다. 또한, 두 개의 퍼지 집합 A와 B간의 거리 측도는 몇 개의 특정 값에 의해 효율적으로 표현될 수 있음을 알 수 있다. 이러한 사실로부터 SPMF에 기반을 둔 거리 측도를 정의한다.

(1) 삼각형 형태 소속함수의 경우

삼각형 형태 소속함수는 3점 (즉, x_L, x_M, x_H)을 사용해서

표현될 수 있다 (단, $x_L < x_M < x_H$). 만약 규칙의 조건이 A = (x_L, x_M, x_H) 이고 관찰 A′ = (x_L', x_M', x_H')라고 하면 DM은 각각의 3점을 사용해서 계산된다.

$$\begin{aligned} DM_L &= x_L' - x_L \\ DM_M &= x_M' - x_M \\ DM_H &= x_H' - x_H \end{aligned} \tag{2}$$

(2) 사다리꼴 형태 소속함수의 경우

사다리꼴 형태 소속함수는 4점(즉, x_L, x_{I1}, x_{I2}, x_H)을 사용해서 표현될 수 있다(단, $x_L < x_{I1} < x_{I2} < x_H$). 만약 규칙의 조건이 A = ($x_L$, x_{I1}, x_{I2}, x_H) 이고 관찰 A′ = (x_L', x_{I1}', x_{I2}', x_H') 라고 하면 DM은 각각의 4점을 사용해서 계산된다.

$$\begin{aligned} DM_L &= x_L' - x_L \\ DM_{I1} &= x_{I1}' - x_{I1} \\ DM_{I2} &= x_{I2}' - x_{I2} \\ DM_H &= x_H' - x_H \end{aligned} \tag{3}$$

제안된 방법에서 거리 측도는 SPMF의 매개변수를 사용하여 식 (2)와 (3)처럼 쉽게 계산될 수 있다.

3.3 조정함수(Modification functions; MF)

검색된 규칙 R_i : $A_i \rightarrow B_i$ 는 조정함수(MF)를 이용해서 규칙 R_i의 결론 부분인 B_i를 조정한다.

(1) 삼각형 형태 소속함수의 경우

두 퍼지 집합 (즉, A, A′) 간의 최대 지원 구간(Maximum support interval; MSI)를 [x_{LL}, x_{MH}] (단, x_{LL} = min {x_L, x_L'}, x_{MH} = max {x_H, x_H'})라 하고 두 퍼지 집합의 MSI 거리 (DMSI)를 $|x_{MH} - x_{LL}|$라 하면 각각의 MF는 다음과 같이 계산된다.

$$\begin{aligned} MF_L &= (1+ (DM_L/DMSI)) \\ MF_M &= (1+ (DM_M/DMSI)) \\ MF_H &= (1+ (DM_H/DMSI)) \end{aligned} \tag{4}$$

(단, DM들은 식 (2)에 의해 만들어진다).

(2) 사다리꼴 형태 소속함수의 경우

두 퍼지 집합 (즉, A, A′) 간의 최대 지원 구간(Maximum support interval; MSI)를 [x_{LL}, x_{MH}] (단, x_{LL} = min {x_L, x_L'}, x_{MH} = max {x_H, x_H'})라 하고 두 퍼지 집합 의 MSI 거리 (DMSI)를 $|x_{MH} - x_{LL}|$라 하면 각각의 MF는 다음과 같이 계산된다.

$$\begin{aligned} MF_L &= (1+ (DM_L/DMSI)) \\ MF_{I1} &= (1+ (DM_{I1}/DMSI)) \\ MF_{I2} &= (1+ (DM_{I2}/DMSI)) \\ MF_H &= (1+ (DM_H/DMSI)) \end{aligned} \tag{5}$$

(단, DM들은 식 (3)에 의해 만들어진다).

식 (4) 와 (5)의 소속함수 형태별로 MF 값들의 평균을 이용해서 OMF를 계산한다.

$$OMF = avg\{\text{모든 } MF(A, A')\} \tag{6}$$

(단, MF는 식 (4) 또는 (5)를 각각 이용한다)

3.4 추론된 결과

먼저 하나의 관찰(A′)과 단순한 규칙형태 (즉, '*IF X is A THEN Y is B*')를 고려한다. B를 언어 변수 '*Y*' 의 퍼지 부분 집합이고 SPMF로 표현되어 있고 $\forall y \in Y$ 에 대해 언어 값 B는 삼각형 형태 또는 사다리꼴 형태 소속함수로 표현되어 있다면 각각의 경우에 대해 (y_L, y_M, y_H) 또는 (y_L, y_{I1}, y_{I2}, y_H)로 나타낼 수 있다. 그러면 OMF를 사용해서 추론된 결과 (B′)가 다음처럼 계산된다.

(1) 삼각형 형태 소속함수의 경우

$$\begin{aligned} y_L' &= OMF \times y_L \\ y_M' &= OMF \times y_M \\ y_H' &= OMF \times y_H \end{aligned} \tag{7}$$

(단, OMF는 식 (6)에 의해 만들어진다).

(2) 사다리꼴 형태 소속함수의 경우

$$\begin{aligned} y_L' &= OMF \times y_L \\ y_{I1}' &= OMF \times y_{I1} \\ y_{I2}' &= OMF \times y_{I2} \\ y_H' &= OMF \times y_H \end{aligned} \tag{8}$$

(단, OMF는 식 (6)에 의해 만들어진다).

[정의 1] 식 (6)에 따라 추론 단계에서의 조정의 방향성이 다음과 같이 결정된다:

Case 1: 만약 OMF<1 이면 y_L, y_H 같은 점들에 대해 OMF 만큼 좌측으로 이동한다.

Case 2: 만약 OMF=1 이면 이동이 발생하지 않는다. 특별한 경우로 규칙의 조건 부분 (A)과 관찰(A′) 간에 만약 모든 DM (식 (2) 또는 (3) 참조)이 0이면 A와 A′ 간에 정합이 발생한 것이다.

Case 3: 만약 OMF>1 이면 y_L, y_H같은 점들에 대해 OMF 만큼 우측으로 이동한다.

[정의 1]에서 규칙 '*IF X is A THEN Y is B*'에서 A 와 B의 관계가 긍정적인 의존관계인 것을 고려하였다. 예를 들어, 규칙 '*IF economic conditions are good THEN the earning is big*' 에서 '*good*' → '*big*' 처럼 긍정적인 의존관계에 있는 경우이다. 만약 규칙의 조건과 결과에 있는 언어값간에 부정적 의존관계인 경우, 예를 들어, '*high weight*' → '*low speed* '처럼 부정적 의존관계인 경우, 추론 단계에서의 조정의 방향성이 반대로 변화된다. 또한, [정의 1]의 Case 2의 특별한 경우가 발생했을 때(즉, A = A′) 제안된 방법의 추론 결과는 B′ = B이다. 이는 산술 규칙을 사용하는 CRI 에 대한 제안된 방법의 장점 중의 하나이다. 즉, 제안된 방법은 Modus ponens를 만족한다. 제안된 LCBFR의 절차를 예를 들어 설명한다.

[예 2] 하나의 관찰(A′) 과 단순한 규칙형태 (즉, '*IF X is A THEN Y is B*')를 고려한다. 관찰 (A′) = '*economic conditions are good*''이고, 3.1 장의 SPMF에 기반을 둔 언어적 케이스 색인과 검색을 사용해서 규칙베이스에서 규칙 '*IF economic conditions are good THEN the earning is* *big*' 이 검색되었다고 하자. 언어 변수 '*economic conditions*'에 관한 퍼지 부분집합 'Good'와 'Good′'의 소속함수가 구간 [0, 100]에서 사다리꼴 형태 소속함수를 사용해서 다음과 같이 정의되어 있다고 하자.

(그림 3)에서 검색된 규칙의 조건은 '*Good*'이고 관찰(A′)은 '*Good′*'이다. 이 경우 사다리꼴 형태 소속함수가 사용되었으므로 MF는 식 (5)를 사용해서 계산된다.

MF_L = (1+ ((88-80)/(100-80)) = (1 + (8/20)) = 1.4.

MF_{I1} = (1+ ((92-85)/(100-80)) = (1 + (7/20)) = 1.35.

MF_{I2} = (1+ ((96-90)/(100-80)) = (1 + (6/20)) = 1.3.

MF_H = (1+ ((100-95)/(100-80)) = (1 + (5/20)) = 1.25.

한편 식 (6)을 사용해서 OMF = (1.4+1.35+1.3+1.25)/4 = 1.33가 된다. 만약 (그림 4)처럼 '*big earning*'이 정의되어 있다면 B('*big earning*')에 OMF를 적용해서 B′를 계산할 수 있다.

y_L' = OMF × y_L = 1.33 × \$10 = \$13.3.

y_{I1}' = OMF × y_{I1} = 1.33 × \$12 = \$15.96.

y_{I2}' = OMF × y_{I2} = 1.33 × \$13 = \$17.29.

y_H' = OMF × y_H = 1.33 × \$15 = \$19.95.

결과적으로 추론된 결과 B′ = (y_L', y_{I1}', y_{I2}', y_H') = (13.3, 15.96, 17.29, 19.95)를 계산할 수 있다.

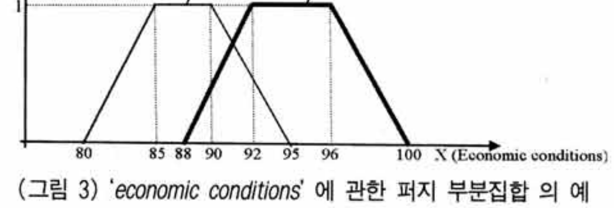
이제 규칙의 조건이 '*OR*' 또는 '*AND*'로 연결된 복합 규칙의 경우를 고려한다.

(1) '*OR*' 연결자의 경우

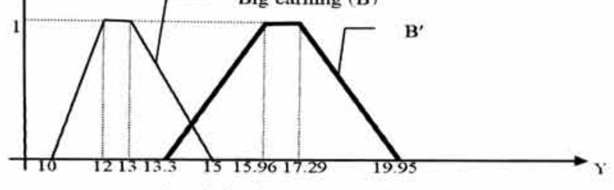
[A_{i1} *OR* A_{i2} *OR* ⋯*OR* A_{iK}] → B_i 형태의 규칙일 경우 $A_{i1} \rightarrow B_i$, $A_{i2} \rightarrow B_i$, ⋯, $A_{iK} \rightarrow B_i$ 과 같이 단순한 규칙들로 분할 될 수 있고 각각의 단순 규칙은 별개의 단순 규칙으로 처리될 수 있다 [18].

(2) '*AND*' 연결자의 경우

[A_{i1} *AND* A_{i2} *AND* ⋯ *AND* A_{iK}] → B_i 형태의 규칙일 경우 (A_{ij}, A_{ij}')에 대해 MF_{ij} 의를 평균을 이용해서 OMF_i를 계산한다 (단, i는 i 번째 규칙을 나타내고 j = 1, 2, ⋯, k).



(그림 3) '*economic conditions*' 에 관한 퍼지 부분집합 의 예



(그림 4) 추론된 결과(B′)의 예

이 경우 식 (6)은 다음과 같이 변경된다.

$$OMF_i = \text{avg}\ \{\text{avg}\ MF_{ij}(A_{ij}, A_{ij}')\} \quad (9)$$

(단, 각 $MF_{ij}(A_{ij}, A_{ij}')$는 식 (4) 또는 (5)로부터 만들어진다. 관찰은 [A_{i1}' *AND* A_{i2}' *AND* ⋯ *AND* A_{iK}']의 형태로 가정한다)

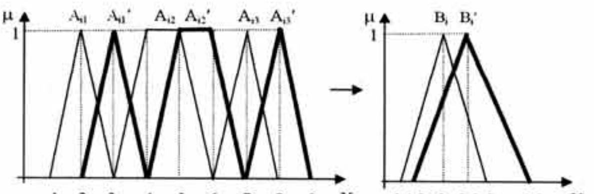
[예 3] i 번째 규칙이 '*IF* X_1 *is* A_{i1} *AND* X_2 *is* A_{i2} *AND* X_3 *is* A_{i3} *THEN Y is B*'와 같이 '*AND*' 연결자로 구성된 경우를 고려한다. A_{i1} = (1, 2, 3), A_{i2} = (3, 4, 5, 6), A_{i3} = (6, 7, 8) 그리고 A_{i1}' = (2, 3, 4), A_{i2}' = (4, 5, 6, 7), A_{i3}' = (7, 8, 9) 라고 하면 식 (4), (5), (9)를 이용해서 OMF_i를 계산할 수 있다.

$$OMF_i = (\sum_{j=1}^{3} \text{avg}\ (MF_{ij}))/3$$
$$= \{[(1+((2-1)/(4-1)))+(1+((3-2)/(4-1)))+(1+((4-3)/(4-1)))]/3$$
$$+[(1+((4-3)/(7-3)))+(1+((5-4)/(7-3)))+(1+((6-5)/(7-3)))$$
$$+(1+((7-6)/(7-3)))]/4+[(1+((7-6)/(9-6)))+(1+((8-7)/(9-6)))$$
$$+(1+((9-8)/(9-6)))]/3\}/3$$
$$= \{(1+(1/3))+(1+(1/4))+(1+(1/3))\}/3 = 1.3.$$

(그림 5)처럼 B_i = (y_L, y_M, y_H) = (1, 2, 3) 로 정의되어 있다면 OMF_i을 사용하여 추론된 결과 B_i'를 얻을 수 있다 : B_i' = (y_L', y_M', y_H') = (1.3, 2.6, 3.9).

이 논문에서는 삼각형과 사다리꼴 소속 함수를 사용해서 설명되었다. S 형태, Π 형태 등과 같은 SPMF에도 비슷한 방법을 적용할 수 있다.

실제 응용에서 중요한 문제 중의 하나는 적용 방법의 계산 속도이다. 제안된 방법은 선형 시간 복잡도를 갖는 퍼지 추론을 위한 효율적인 방법을 제공한다. CRI가 행렬 연산에 기반을 두고 있고 행렬의 크기가 퍼지 집합의 요소의 개수에 영향을 받는 반면 제안된 방법은 SPMF에 있는 적은 수의 매개변수를 사용해서 계산된다. 또한, SPMF의 매개변수를 사용해서 거리 측도나 조정 함수를 계산할 수 있어 추론 결과를 쉽게 만들 수 있다.



(그림 5) '*AND*' 연결자의 예

4. 기존의 방법과의 비교

기존의 퍼지 추론 방법에서 소속함수는 Ad-hoc [15, 16]하게 만들어지는데, 제안된 방법에서는 SPMF를 사용한다. 또한, 기존의 퍼지 추론 방법에서 추론 방법이 주로 CRI [20, 23] 등을 사용하여 복잡한 행렬 계산식을 사용하는데 비해 제안된 방법은 SPMF의 매개변수를 특징점으로 사용하여 계산되므로 퍼지 추론을 위한 효율적인 방법을 제공한다. 제안된 방법과 기존의 퍼지 추론 방법간의 차이점을 요

약하면 다음과 같다.

〈표 1〉 기존의 방법과의 비교

비교항목	기존의 방법	제안된 방법
소속함수	Ad-hoc [15, 16]	SPMF
추론 방법	CRI [20, 23] 등 사용	LCBFR
계산식	복잡 [15, 16]	단순

5. 결 론

제안된 방법은 SPMF의 매개변수를 특징점으로 사용하여 계산된다. 제안된 언어적 Case 색인과 검색 방법은 언어 근사 과정에서 관련 없는 언어 값들에 대한 탐색을 방지하기 위해 언어 변수들을 분리하는 분할 방법을 사용하였다. 이는 LCBFR 과정에서 관련 없는 언어 값을 탐색하는 것을 피할 수 있게 한다. 이러한 특징이 기존의 언어 근사 방법보다 상대적으로 빠른 언어 근사가 가능하여 관찰(A')과 규칙베이스의 조건(A)간의 패턴 매칭 속도를 개선할 수 있다. 산술 규칙을 사용하는 CRI와 비교할 때 제안된 방법은 퍼지 추론에서 요구되는 사항 중의 하나인 Modus ponens를 만족한다. 또한, 관찰(A')과 규칙베이스의 조건(A)간에 미결정 부분이 발생하지 않는다. 기존의 퍼지 추론 방법[15-16, 18, 23]과 비교할 때 SPMF의 매개변수를 사용해서 거리 측도나 조정함수를 계산할 수 있어 추론 결과를 쉽게 만들 수 있다. 제안된 방법은 선형 시간 복잡도를 갖는 퍼지 추론을 위한 효율적인 방법을 제공한다. 결과적으로 제안된 방법은 퍼지 추론의 속도를 개선하는데 사용될 수 있다.

참 고 문 헌

[1] H. Ahn, K. J. Kim, I. Han, A case-based reasoning system with the two-dimensional reduction technique for customer classification, Expert systems with application 32 (2007) 1011-1019.

[2] I. Batyrshin, M. Wagbnknbcht, Towards a linguistic description of dependencies in data, Int. journal of appl. math. compt. sci., 12(3) (2002) 391-401.

[3] T. C. Chang, K. Hasegawa, C. W. Ibbs, The effects of membership function on fuzzy reasoning, Fuzzy sets and systems 44 (1991) 169-186.

[4] C. Chiu, A case-based customer classification approach for direct marketing, Expert systems with application 22 (2002) 163-168.

[5] C. Chiu, P. C. Chang, N. H. Chiu, A case-based expert support system for due-date assignment in a water fabrication factory, Journal of intelligent manufacturing (14) (2003) 287-296.

[6] D. Y. Choi, ATM based on SPMF, Lecture notes in artificial intelligence 4251 (2006) 490-497.

[7] R. Degani, G. Bortolan, The problem of linguistic approximation in clinical decision making, International journal of approx. reasoning 2(2) (1988)143-162.

[8] F. Eshragh, E. H. Mamdani, A general approach to linguistic approximation, International journal of man-machine studies 11(1979) 501-519.

[9] D. Gentner, K. J. Holyoak, B. Kokinov, Analogy: perspectives from cognitive science, (MIT press. Cambridge: MA, 2000)

[10] K. D. Forbus, T. Mostek, R. Ferguson, An analogy ontology for integrating analogical processing and first-principles reasoning, Proc. IAAI-02 (2002) 878-885.

[11] T. Hamaguchi, H. Meng, K. Takeda, Y. Shimada, Y. Hashimoto, T. Itoh, A training system for maintenance personnel based on analogical reasoning, LNAI 4252 (2006) 587-594.

[12] Kolodner, J., Case-based Reasoning (Morgan Kaufmann, 1993).

[13] R. Kowalczyk, On linguistic approximation with genetic programming, Lecture notes in computer science 1415 (1998) 200-209.

[14] E. Mamdani, Soft knowledge as key enabler of future services, Proceedings of the 2001 BISC international workshop on fuzzy logic and the Internet (2001) 145-148.

[15] M. Mizumoto, Extended fuzzy reasoning, In : Gupta et al. (Eds.), Approximate reasoning in expert systems (North-Holland, 1985) 71-85.

[16] M. Mizumoto, H.-J. Zimmermann, Comparison of fuzzy reasoning methods, Fuzzy sets and systems 8 (1982) 253-283.

[17] T. Tanrikorur, Great expectations, Intelligent enterprise 4(12) (2001) 35-38.

[18] I. B. Turksen, Z. Zhong, An approximate analogical reasoning approach based on similarity measures, IEEE transactions on SMC 18(6) (1988) 1049-1056.

[19] F. Wenstop, Deductive verbal models of organization, International journal of man-machine studies 8(1976) 293-311.

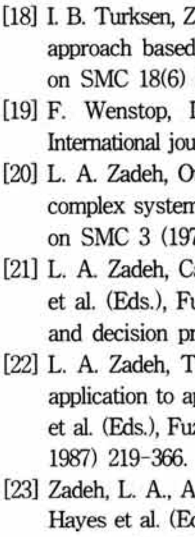
[20] L. A. Zadeh, Outline of a new approach to the analysis of complex systems and decision processes, IEEE transactions on SMC 3 (1973) 28-44.

[21] L. A. Zadeh, Calculus of fuzzy restriction, In L. A. Zadeh, et al. (Eds.), Fuzzy sets and their applications to cognitive and decision processes (Academic Press, 1975) 1-39

[22] L. A. Zadeh, The concept of a linguistic variables and its application to approximate reasoning 1, 2, 3, In R. R. Yager, et al. (Eds.), Fuzzy sets and applications (John Wiley & Sons, 1987) 219-366.

[23] Zadeh, L. A., A Theory of Approximate Reasoning, in J. E. Hayes et al. (Eds.), Machine Intelligence (Wiley, New York), pp.149-194, 1979.

[24] R. Zwick, E. Carlstein, D. V. Budescu, Measures of similarity among fuzzy concepts : A comparative analysis, International journal of approximate reasoning 1(1987) 221-242.



최 대 영

e-mail : dychoi@yuhan.ac.kr
1985년 서강대학교 컴퓨터학과(공학사)
1992년 서강대학교 컴퓨터학과(공학석사)
1996년 서강대학교 컴퓨터학과(공학박사)
1985년~1990년 한국국방연구원(연구원)
1994년 정보처리기술사(전자계산조직응용)
2000년~2001년 BISC Group, University of California, Berkeley (Visiting Scholar)
2004년~2006년 University of Colorado, Denver (Visiting Professor)
1997년~현 재 유한대학 경영정보과 교수
관심분야 : 비즈니스 인텔리전스, 웹 검색엔진, RFID 응용, 유비쿼터스 컴퓨팅