한국 퍼지 및 지능시스템학회 논문지 2000, Vol. 10, No. 1 (2000)12~15

소속함수에 사용자의 선호도를 반영한 퍼지 집합의 비교 알고리즘

Weighted Comparison of Fuzzy Sets using Preferences on Membership Function

김대원 · 신준범 · 정성원 · 이광형

Dae-Won Kim, Jun-Bum Shin, Sung-Won Jung and Kwang-Hyung Lee

한국과학기술원 전자전산학과 전산학전공 , AITrc, 인공지능 연구실

요 약

본 논문에서는 사용자의 선호도를 반영한 퍼지집합의 비교 알고리즘을 제안한다. 알고리즘에서 사용되는 선호도는 퍼지집합의 도메인에 대한 선호도와 만족도 함수에 대한 선호도이다. 제시된 두 가지의 선호도를 두 퍼지집합간의 유사도 비교에 사용하게 된다. 퍼지집합간의 유사도 함수는 각 선호도의 함수로 제시되며, 이에 따라 주어진 퍼지집합간의 유사도 결정에 일반적인 해를 줄 수 있을 것으로 보인다.

ABSTRACT

In this paper, a new fuzzy set comparison algorithm which exploits users' preferences is proposed. The preferences used in proposed algorithm are divided into two different kinds of ones: one is a preference on the domain of fuzzy sets and the other is a preference on the membership function. These two preferences are used in determining a similarity value between two fuzzy sets. The similarity function between fuzzy sets is formulated as the function of preferences. This proposed algorithm might give a general solution in determining a similarity value between given fuzzy sets.

1. 서 론

퍼지집합은 애매모호한 값을 표현하는데 적합하여 그 동안 많은 다양한 분야에서 응용되어져 왔으며, 지금까지 효과적으로 퍼지집합을 다루기 위한 여러 가지 연구가 진행되어 왔다. 퍼지집합은 애매모호한 값을 가능성 분포를 이용하여 표현한 것으로, 관심있는 모든 원소들을 포함한 전체집합에서 어떤 특정한 성질을 갖는 원소들과 그 원소들의 집합에 대한 소속 가능성을 만족도를 이용해서 나타낸 집합이다[1]. 이에 따라 퍼지집합은 각 원소들의 집합에 대한 만족도를 임의로 설정할 수 있어서, 전통적인 고전집합(crisp set)이 표현하지 못하는 불분명성과 애매성을 표현할 수 있다.

본 논문은 퍼지집합간의 새로운 비교 알고리즘을 제안한다. 본 연구의 동기는 주어진 표본 공간에 N개의 퍼지집합이 존재할 경우, 새로이 관측(observed)된 퍼지집합과 표본공간의 N개의 집합 중에서 가장 가까운 집합을 찾는 문제에서 제시되었다. 이러한 문제를 해결하는 전통적 방법으로는 비교하는 두 퍼지집한간의 유사도(similarity)를 결정하여, 가장 큰 유사도를 갖는 집합을 문제의 해로 갖게된다. 이에 본 논문에서는 두 퍼지집합간의 유사도 값을 비교하는 새로운 알고리즘을 제안한다. 두 퍼지집합간의 유사도 계산에 관한 기존 연구로는 퍼지집합간의 거리 척도를 이용한 것이 대표적이었다. 거리 척도로는 해밍 거리(Hamming distance), 유클리디언 거리(Euclidean distance) 등의 방법론을 비롯하여 많은 연구가 진행되어 왔다[2].

그러나 지금까지의 연구들은 그 방법론내에 여러 에러 요소를 지니고 있다. 또한 퍼지집합은 앞에서 언급한 바와 같이 보통집합처럼 명확한 값으로 표현된 것이 아니라, 각 원소들의 가능성 분포를 이용하여 표현된 애매모호한 집합이다. 따라서 두 퍼지집합의 유사도를 비교할 때 전체적인 집합의 가능성 분포를 고려하여야 한다. 더욱이 좀 더 일반적인 비교 척도로 사용되기 위해서는 비교하는 두 집합간에 사용자의 선호도(preference)를 반영할 수 있어야 한다. 이것은퍼지집합을 비교하는 관점, 즉 사용자의 선호도에 따라 다른 비교 결과가 나올 수 있기 때문이다. 최근들어 선호도를 반영하는 비교 방법이 연구되고 있으나, 아

본 연구는 첨단정보기술 연구센터를 통하여 과학재단의 지원을 받았음.

직 그 방법의 일관성이 부족하다.

이에 본 논문에서는 유사도를 이용한 퍼지집합의 비교 척도로서 보다 정확한 결과를 제시하며, 사용자의 선호도와 관심을 반영할 수 있는 새로운 알고리즘을 제시한다. 사용자의 선호도를 반영하기 위하여 비교하는 두 집합의 도메인에 대한 선호도(domain preference)와 집합의 만족도 대한 선호도(membership degree preference)의 개념을 도입하였다. 즉, 사용자의 의도를 퍼지집합 비교에 반영하여 비교하고자 하는 부분에 좀 더 가중치를 둘 수 있게 하였다. 여기서 유사도 비교값은 도메인 선호도와 만족도 선호도의 적분을 이용한 합성으로 계산된다. 본 논문의 접근은 기존의 유사도를 측정하는 방법의 일반화라는 점과 사용자의 선호도를 퍼지집합의 비교에 반영할 수 있다는 점에서 의미가 있다고 할 수 있다.

2. 사용자 선호도의 반영

퍼지집합의 유사도 비교가 일반적으로 확장되기 위해서는 사용자의 선호도나 의도를 고려해야 한다. 본 논문에서는 다음과 같은 두 가지의 선호도 요소를 정의하여, 사용자의 의도를 반영한다.

2.1 도메인 선호도 함수

$d\mathrm{Domain} = f_{Domain}(x)dx$

비교하는 두 퍼지집합의 도메인 축에 변화를 주는 선호도이다. 비교하는 도메인 상에서 좀 더 비중을 두고 싶은 영역에 가중치를 두게 된다. 그림 1의 (a)는 비교하는 두 퍼지집합의 예를 보인 것이다(여기서는 삼각꼴 퍼지집합을 가정하였다). 그림 (b)는 두 퍼지집합 A, B의 유사도 비교시에 적용한 도메인 선호도로, 이 경우에는 도메인의 오른쪽에 큰 비중을 두고있음을 나타낸다.

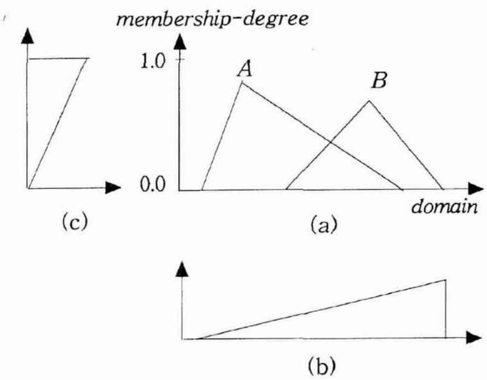


그림 1. (a) 비교하는 퍼지집합 A, B, (b) 도메인 선호도 함수, (c) 만족도 선호도 함수

2.2 만족도 선호도 함수

$d\mathrm{MV} = f_{MV}(y)dy$

비교하는 두 퍼지집합의 만족도 축에 변화를 주는 선호도이다. 비교하는 만족도 축([0.0, 1.0])상에서 좀 더 비중을 두고 싶은 영역에 가중치를 둘 수 있다. 그림 1의 (c)는 만족도 선호도의 관점을 제시한 것으로, 만족도 1.0에 가까운 영역에 사용자의 비교 가중치를 둔 것이다.

3. 집합간의 유사도 함수

본 절에서는 도메인 선호도와 만족도 선호도 함수를 이용하여 두 퍼지집합간의 유사도 비교를 계산하는 방법을 기술한다.

3.1 도메인 선호도의 적용

유사도 값의 계산은 먼저 특정 만족도 값에 대해서 도메인 선호도를 이용하여 두 집합의 유사도를 계산한다. 그 후 계산된 값은 다음 절에 설명될 전체 만족도 영역에 대해서 선호도의 가중치를 이용하여 전체 유사도가 결정된다.

도메인상의 한 점 x에 대해서 만족도 값 y에 해당하는 유사도 $\mu_{A,B}(x, y)$는 식 (1)과 같다.

$$\mu_{A,B}(x,y) = \begin{cases} 1 \text{ if } y \le (\mathrm{MIN}(\mu_A(x), \mu_B(x))) \\ 0 \ elsewhere \end{cases} \tag{1}$$

즉, 특정 레벨 수준 y에서 퍼지집합 A, B가 유사하다는 것은 두 집합 모두 y값보다 크다는 것을 의미한다. 이 계산값 식 (1)을 이용하여 도메인 선호도를 반영할 수 있으며, 그 식은 식(2)와 같이 적분의 형태

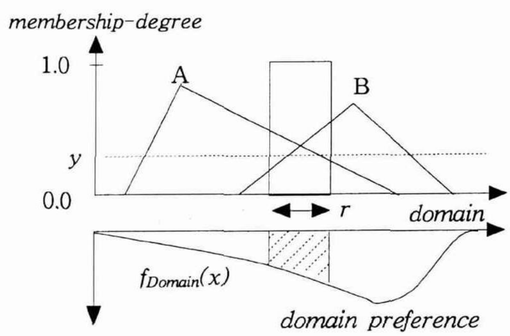


그림 2. 도메인 선호도 함수의 적용

를 가진다.

$$f(y) = \begin{cases} \int_{Domain} \mu_{A,B}(x,y) d\text{Domain} \\ \int_{Domain} \mu_{A,B}(x,y) f_{Domain}(x) dx \end{cases} \quad (2)$$

이 식이 의미하는 것을 그림으로 도시하면 그림 2와 같다. 도메인 선호도 함수 $f_{Domain}(x)$와 계산을 원하는 도메인 영역 r이 주어진 경우, 특정 만족도 y에서의 두 집합의 유사도 계산을 수행한다. $x \in r$에서 $\mu_{A,B}(x, y)$값은 식 (1)에 의해 모두 1이 되고, 식 (2)에 의해서 그림 2에서 보는 바와 같이 도메인 함수의 적분 값이 특정 만족도 y에서의 도메인 선호도를 반영한 유사도 값이 된다.

3.2 유사도 값의 계산

앞 절에서 계산된 유사도는 특정 만족도 y에서의 결과값이다. 따라서 전체 만족도 영역에 대해서 사용자의 만족도 선호도를 고려한 유사도 값을 계산하는 과정이 필요하며, 이를 식 (3)에 제시하였다.

$$Value_{A,B} = \int_{MV} f(y) d\text{MV} \quad (3)$$

$$Value_{A,B} = \int_{y=0}^{1} f(y) f_{MV}(y) dy \quad (4)$$

식에서 $Value_{A,B}$는 두 퍼지집합의 최종 유사도 값을 나타내며, 식 (2)에서 구한 $f(y)$를 만족도 영역(*MV, membership-value*)에 대해서 적분한 값이다. 이 식은 만족도 함수 $f_{MV}(y)$를 이용하여 식 (4)와 같이 나타낼 수 있으며, 그림으로 도시하면 그림 3과 같다.

특정 만족도 y에 대해 도메인 선호도를 이용하여 계산한 결과 $f(y)$를 사용자의 만족도 선호도 함수 $f_{MV}(y)$에 적분값을 취함으로써, 유사도를 결정한다. 이 계산을 전체 만족도 영역에 대해서 수행하면 두 퍼지집합 A, B의 유사도 값이 결정된다.

이제 서론에서 제기했던 동기의 문제를 해결 할 수 있다. 즉, 주어진 표본 공간에 N개의 퍼지집합이 존재할 경우, 새로이 관측된 입력 퍼지집합과 표본공간의 N개의 집합 중에서 가장 가까운 집합(S)을 찾는 문제는 식 (5)의 최대값 척도로서 구해진다.

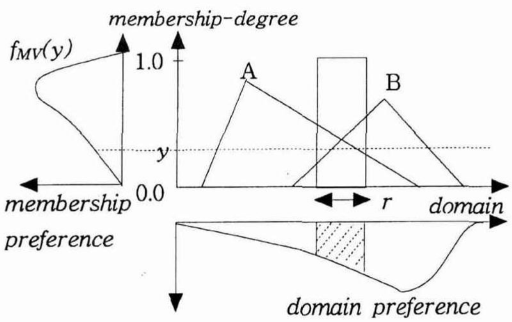


그림 3. 만족도 선호도 함수의 적용

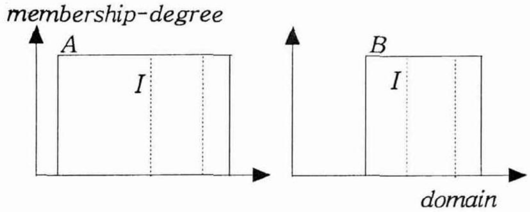


그림 4. 유사도 값이 동일한 두 경우

$$S = Max(Value_{A,I}, Value_{B,I}, \cdots, Value_{Z,I}) \quad (5)$$

, where A, B, ⋯, Z : 표본공간의 퍼지집합
I : 입력 퍼지집합

3.3 유사도 함수의 개선

기 제안된 유사도 함수는 사용자의 선호도를 반영할 수 있다는 점에서 의의가 있다. 하지만 그림 4에 보인 것과 같은 극단적인 경우, 유사도 계산은 에러를 포함한다.

위의 경우 두 표본 퍼지집합 A, B와 입력 퍼지집합 I의 유사도 비교값은 집합 B와 더 유사하다는 값을 주어야 하지만, 제안한 척도를 이용한 결과는 동일한 값을 가진다. 이것은 식 (1)에서 MIN연산의 정의에서 기인한 것으로서, 이를 보상해 주기 위해서 식 (6)과 같이 최종 유사도 값에 에러 보상 항목을 고려해 주어야 한다.

$$Value_{S,I} \times \frac{A_S \cap A_I}{A_S} \quad (6)$$

, where A_S: 집합 S의 면적, A_I: 집합 I의 면적

위 식에서 S는 샘플 퍼지집합을, I는 비교하는 입력 퍼지집합을 나타내며, A는 집합의 면적을 나타낸다. 즉, 식 (4)에서 구한 유사도 값에 비교하는 퍼지집합의 면적의 비를 곱으로 고려해줌으로써 위와 같은 경우를 해결할 수 있다.

4. 결 론

본 논문에서는 퍼지집합간의 유사도 비교에 사용될 새로운 척도를 제안하였다. 제안된 방법은 두 퍼지집합의 유사도를 비교할 때, 사용자의 선호도를 계산 과정에 반영할 수 있다. 반영되는 사용자의 의도는 퍼지

집합의 도메인에 대한 선호도와 만족도에 대한 선호도 두 가지이다. 각각 도메인 선호도 함수와 만족도 선호도 함수의 구간 적분값을 이용하였으며, 두 계산의 합성 결과를 최종 집합의 유사도로 정의하였다. 제안된 유사도 척도는 사용자의 비교 관점을 반영할 수 있다는 점에서 의의를 지닌다.

제안된 척도는 현재로서는 수학적인 개념적 단계에 있으며, 고려치 못한 에러 경우가 있을 수 있다. 따라서 척도에 대한 좀 더 다양한 분석이 필요하며 이 척도가 의의를 가질 수 있는 응용 분야에 대한 연구가 수행되어야 한다.

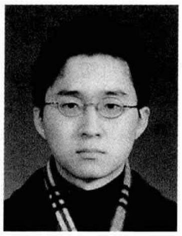
참고문헌

[1] D. Dubois and H. Prade. Fuzzy Sets and Systems: Theory and Applications. Academic Press, 1980.

[2] 이광형, 오길록, 퍼지 이론 및 응용: I권 이론. 홍릉과학출판사, 1991.

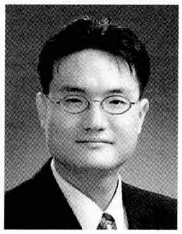
[3] 이지형, 이광형, 사용자 관심도를 반영하는 퍼지 숫자의 정렬 방법. 한국퍼지 및 지능시스템 학회 논문집, Vol. 8, pages 14-20, 1998.

[4] Jee-Hyung Lee. Comparison, Ranking and Determination of the Ranks of Fuzzy Numbers based on a Satisfaction Function. Ph.D Thesis, KAIST, 1999.



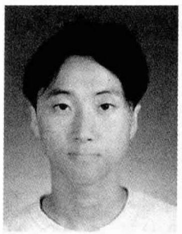
김 대 원 (Dae-Won Kim)

1997년 : 경북대학교 컴퓨터공학과 학사
1999년 : 한국과학기술원 전산학과 석사
1999년~현재 : 한국과학기술원 전산학과 박사과정
관심분야 : 퍼지이론 및 응용, 가상현실, 인공지능



신 준 범 (Jun-Bum Shin)

1995년 : 한국과학기술원 수학과 학사
1998년 : 한국과학기술원 수학과 석사
1998년~현재 : 한국과학기술원 전산학과 박사과정
관심분야 : 퍼지이론 및 응용, 보안, 전자상거래



정 성 원 (Sung-Won Jung)

1998년 : 한국과학기술원 전산학과 학사
1998년~현재 : 한국과학기술원 전산학과 석사과정
관심분야 : 퍼지이론 및 응용, 학술이론, 인공지능

이 광 형 (Kwang-Hyung Lee)

1978년 : 서울공대 산업공학학사
1980년 : 한국과학원 산업공학 석사
1982년 : 프랑스 INSA 전산학과 석사 (DEA)
1985년 : 프랑스 INSA 전산학과 공학박사
1988년 1월 : 프랑스 국가박사(전산학 INSA-LYONI대)
1985년~1995년 : 한국과학기술원 전산학과 조교수 및 부교수
1995년~현재 : 한국과학기술원 전산학과 교수
1985년 : 프랑스 INSA
1995년 : 미국 Stanford Research Institute
주관심분야 : 퍼지 이론 및 응용, 인공지능, 전문가 시스템 등