

유전자알고리즘에 의한 확장된 퍼지적분의 지원

이 건 명^o 이 광 형

한국과학기술원 전산학과 대전시 유성구 구성동 373-1

Genetic Algorithm-based Support for Extended Fuzzy Integral

Keon-Myung Lee Hyung Leekwang

Dept. Computer Science, KAIST(Korea Advanced Institute of Science and Technology), Taejon, 305-701

요 약

퍼지적분은 어떤 대상에 대한 여러 평가항목의 평가치를 평가항목의 중요도를 고려하여 합성하는 방법이다. 본 논문에서는 정보합성이나 증거결합 등에서 퍼지적분을 사용할 때 발생하는 문제점을 살펴보고, 이러한 문제점 해결을 위해 유전자알고리즘에 기반한 방법을 제안한다. 평가항목에 대한 퍼지척도의 성질을 만족하는 중요도를 결정하기 위해 유전자알고리즘을 사용한다. 한편 기존의 퍼지적분의 특성을 보완하기 위해 확장된 퍼지적분을 사용할 때 요구되는 파라미터 결정을 위해 유전자알고리즘을 사용하는 방법을 제안한다.

1. 서론

어떤 대상에 대한 평가를 할때 여러가지 평가항목에 대해서 평가를 하게 되고, 이들로 부터 전체적인 평가치를 결정해야 하는 상황이 발생한다. 이러한 목적으로 사용되는 방법중의 하나인 Sugeno의 퍼지적분은 여러가지 평가항목에 대한 중요도를 고려하여 평가치를 합성하는 Lebesgue 적분의 일종이다[6].

퍼지적분을 평가치들의 합성방법으로 사용할 때는 다음과 같은 점이 문제로 지적된다. 첫째, 퍼지적분을 평가치를 합성하는 목적으로 사용하기 위해서는 퍼지척도의 성질을 만족하는 평가항목에 대한 중요도를 결정해야 한다. 둘째, 확장된 퍼지적분에서 사용되는 파라미터를 응용응도에 적합하도록 결정해야 한다.

본 논문에서는 위에서 지적한 문제를 해결하는 방법에 대해서 다룬다. 먼저 퍼지척도의 성질을 만족하는 중요도를 결정하기 위해서는 유전자알고리즘[1]에 의한 방법을 사용한다. 평가항목에 대한 중요도는 일반적으로 사람이 직접 부여해야 하는데, 부여된 중요도가 퍼지척도의 성질에 부합하는 일관성있는 값이 되기 어렵다. 따라서 유전자알고리즘을 사용하여 이들 사람이 부여한 중요도 값들로 부터 퍼지척도의 성질에 부합되는 중요도값을 구해낸다. 본 연구에서는 퍼지척도중에서 λ -퍼지척도를 대상으로 한다. 한편, 확장된 퍼지적분[5]에는 보상정도, 상대적 중요도 등의 파라미터가 포함되어 있다. 그런데 이들 파라미터를 사람이 직접 결정하는 것은 쉽지 않다. 따라서 실제 응용분야에서 데이터를 수집해서 이들로 부터 파라미터를 결정하는 것이 바람직하다. 본 논문에서는 이들 파라미터 결정을 위해서 유전자알고리즘을 사용하는 방법을 제안한다.

본 논문은 다음과 같은 구성된다. 2절에서는 배경이론이 되는 퍼지척도와 퍼지적분에서 대해서 기술하고, 3절에서는 기존의 퍼지적분의 특성을 보완하기 위해 이미 제안했던 확장된 퍼

지적분에 대해 소개한다. 4절에서는 퍼지척도의 성질을 만족하는 평가항목의 중요도를 결정하기 위해 제안된 유전자알고리즘을 소개하고, 5절에서는 확장된 퍼지적분의 파라미터를 결정하는 유전자알고리즘을 제안한다. 6절에서는 유전자알고리즘에 의한 실험에 대해 기술한다.

2. 배경이론

퍼지척도는 퍼지적분을 사용할때 평가항목들에 부여된 중요도가 만족해야할 특성이다. 퍼지적분은 여러 평가항목에 대한 평가치를 평가항목의 중요도를 고려하여 합성하는 역할을 한다.

2.1 퍼지척도

퍼지척도 g 는 집합 X 의 멱집합(power set) $B(X)$ 에 대해 정의된 다음의 성질을 만족하는 함수이다.

- $g : B(X) \rightarrow [0, 1]$
- 1) $g(\phi) = 0, g(X) = 1$
- 2) $A, B \in B(X)$ 이고 $A \subset B$ 이면, $g(A) \leq g(B)$ 이다.
- 3) $F_n \in B(X) (1 \leq n < \infty)$ 이고 집합열 $\{F_n\}$ 이 포함관계에서 단조성(monotonicity)을 갖으면 $\lim_{n \rightarrow \infty} g(F_n) = g(\lim_{n \rightarrow \infty} F_n)$ 이다.

퍼지척도중에서 널리 사용되는 λ -퍼지척도 g_λ 는 다음의 성질을 만족하는 척도이다.

$$\forall A, B \in B, A \cap B = \phi \text{일때, } \lambda > -1 \text{에 대해}$$

$$g_\lambda(A \cup B) = g_\lambda(A) + g_\lambda(B) + \lambda g_\lambda(A)g_\lambda(B)$$

유한집합(finite set) $X = \{x_1, x_2, \dots, x_k\}$ 에 대한 λ -퍼지척도는 퍼지밀도(fuzzy density) $g_i = g_\lambda(\{x_i\})$ 의 값에 대해 다음과 같은 식이 성립한다.

$$g_\lambda(\{x_1, \dots, x_l\}) = \sum_{i=1}^{l-1} g_i + \lambda \sum_{i_1=1}^{l-1} \sum_{i_2=i_1+1}^l g_{i_1} g_{i_2} + \dots + \lambda^{l-1} g_1 g_2 \dots g_l = \frac{1}{\lambda} (\prod_{i=1}^l (1 + \lambda g_i) - 1)$$

퍼지척도는 의사결정 등의 문제에서 증거 또는 정보를 결합하는 목적으로 퍼지적분과 사용될 수 있다. 이때 퍼지척도의 값은 평가항목에 대한 중요도를 나타내게 된다. 그런데 이러한 퍼지척도의 값을 사람이 직접 주관적으로 결정해야 하기 때문에 일관성이 있는 값을 부여하기 힘들다. 따라서 사람이 부여한 퍼지척도의 값으로 부터 퍼지척도의 성질을 만족하는 퍼지척도 값을 구해내는 식별문제가 생긴다. 식별문제를 해결하기 위한 기존의 방법들이 몇가지 제안되어 있고, 저자가 제안한 유전자알고리즘을 사용한 식별방법이 있다[4].

2.2 퍼지적분

퍼지적분은 Lebesgue 적분으로서 몇가지 종류가 있지만 이중에서 Sugeno의 퍼지적분이 대표적이다. Sugeno의 퍼지적분은 다음과 같이 정의된다[8].

임의의 보통집합 X 에 대해 퍼지척도 $g: B(X) \rightarrow [0, 1]$ 가 정의되어 있고, X 를 정의구역으로 하고 구간 $[0, 1]$ 을 치역으로 하는 함수 $h: X \rightarrow [0, 1]$ 가 정의되어 있다고 하자. 이때 $A(A \subset X)$ 에서의 함수 h 의 퍼지척도 g 에 대한 Sugeno의 퍼지적분 $\int_A h(x) \circ g(\cdot)$ 은 다음과 같이 정의된다.

$$\begin{aligned} \int_A h(x) \circ g(\cdot) &= \sup_{E \subseteq X} \{ \min \{ \min_{x \in E} h(x), g(A \cap E) \} \} \\ &= \sup_{\alpha \in [0, 1]} \{ \min \{ \alpha, g(A \cap F_\alpha) \} \} \\ &\quad F_\alpha = \{ x \mid h(x) \geq \alpha \} \\ &= \sup_{E \subseteq A} \{ \min \{ \min_{x \in E} h(x), g(E) \} \} \end{aligned}$$

Sugeno의 퍼지적분은, 어떤 대상을 여러 항목(관점)에 대해서 평가할 때 이들 각 항목에 대한 평가치를 퍼지척도를 사용하여 종합하는 방법으로 사용될 수 있다. 이때 집합 X 는 어떤 대상에 대한 평가항목에 해당하고, X 의 멱집합의 원소 $E \in B(X)$ 에 대해 정의되는 퍼지척도 $g(E)$ 는, 대상의 전체적인 평가에 대해 항목 E 의 평가치가 기여하는 정도, 즉 평가항목의 중요도에 해당한다. 적분 영역이 되는 A 는 관심대상이 되는 평가항목에 해당한다.

Sugeno의 퍼지적분은 다음과 같이 해석될 수 있다. 적분식에서 $\min_{x \in E} h(x)$ 는 현재 평가항목 E 에서 가장 부정적인(보수적인) 평가치를 선택하는 것이고, $\min \{ \min_{x \in E} h(x), g(E) \}$ 는 평가항목중에서 가장 부정적인 평가치와 평가항목의 중요도중에서 작은 것을 선택하는 것이다. 이렇게 함으로써 가장 보수적인 평가치를 가짐과 동시에 평가치가 평가항목의 중요도보다 클 수 없게 된다. 한편, 적분 결과를 $\sup_{E \subseteq X} \min \{ \min_{x \in E} h(x), g(E) \}$ 로 함으로써 가장 유리한 항목들을 부각시켜 낙관적인 평가를 하는 측면이 있다.

3. 확장된 퍼지적분

Sugeno의 퍼지적분은 수행과정에서 관심 평가항목들의 평가치중에서 최소값을 취해서 계산함으로써 비관적인 면을 강조하는 경향으로 전체 평가치를 결정한다. 이러한 문제점을 극복하기 위해 다음과 같은 측면에서 확장될 수 있다[5].

퍼지적분식에서 특정 평가항목 집합 E 에 대한 평가치의 결합이 $\min \{ \min_{x \in E} h(x), g(E) \}$ 에 의해서 수행되고 있다. 그

런데 이식에서 평가항목들의 평가치에 대해 최소값을 취하는 연산을 함으로써 지나치게 부정적인 평가를 하고 있다. 일상의 의사결정에서는 하나의 평가항목에 대한 평가치가 열등하더라도 다른 평가항목의 평가치가 우수할 경우에는 열등한 평가치가 어느 정도 보상(compensation)되는 경우가 많다. 따라서 Sugeno의 퍼지적분에서는 $\min_{x \in E} h(x)$ 에서 \min 대신에 보상연산자를 사용함으로써 이러한 보상효과를 얻을 수 있다.

보상연산자에는 mean 연산자와 hybrid 연산자의 두가지 부류가 있다. mean 연산자로는 다음과 같은 연산자가 있다. 아래식에서 $\gamma \in [0, 1]$ 는 보상정도를 나타낸다.

weighted arithmetic mean

$$A \oplus_\gamma B = (1 - \gamma)(A \cap B) + \gamma(A \cup B) \quad 0 \leq \gamma \leq 1$$

geometric mean

$$A \otimes_\gamma B = (A \cap B)^{(1-\gamma)}(A \cup B)^\gamma \quad 0 \leq \gamma \leq 1$$

hybrid 연산자로는 다음과 같은 γ -모델 연산자가 대표적이다. 아래식에서 $x_i \in [0, 1]$ 는 결합되는 평가항목의 평가치를 나타내고, δ_i 는 i 번째 항목에 대한 중요도를 나타내고, $\gamma \in [0, 1]$ 는 보상정도를 나타낸다.

multiplicative γ -model

$$y = \left(\prod_{i=1}^n x_i^{\delta_i} \right)^{(1-\gamma)} (1 - \prod_{i=1}^n (1 - x_i)^{\delta_i})^\gamma$$

$$\sum_{i=1}^n \delta_i = n, \quad 0 \leq \gamma \leq 1$$

additive γ -model

$$y = (1 - \gamma) \prod_{i=1}^n x_i^{\delta_i} + \gamma \left(\prod_{i=1}^n (1 - x_i)^{\delta_i} \right)$$

Sugeno의 퍼지적분에서는 평가항목에 대한 중요도가 퍼지척도로 나타내어져서 적분할때 고려된다. 따라서 평가항목의 중요도를 고려하여 평가치를 합성하는 γ -model 등의 보상 연산자를 평가치의 합성에 사용하면 중요도가 두 측면에서 고려된다. 이에 반해 weighted arithmetic mean과 geometric mean 연산자는 보상정도만을 고려한다. 이들 연산자에서는 γ 가 0일때 \min 과 같은 효과가 있다. γ -model을 사용할 경우에는 δ_i 를 각 평가항목의 상대적인 중요도(relative importance degree)로 사용할 수 있다.

다음 식은 보상연산자 Φ 를 사용하여 확장한 퍼지적분을 나타낸다.

$$\int_A h(x) \circ g(\cdot) = \sup_{E \subseteq A} \{ \min \{ \Phi_{x \in E} h(x), g(E) \} \}$$

한편, 실제 응용에서는 평가치가 불확실성을 포함할 수 있다. 이러한 경우 퍼지적분에서 평가치의 신뢰도나 확신도를 고려하면 유용할 수 있다. 이러한 목적으로 관심대상 평가항목의 집합 A 를 퍼지집합으로 간주하고 소속정도를 확신도 또는 신뢰도로 사용할 수 있다. 이때 퍼지적분은 다음과 같은 식으로 확장된다.

$$\int_A h(x) \circ g(\cdot) = \sup_{E \subseteq \text{Supp}(A)} \{ \min \{ \min_{x \in E} Q(\mu_A(x), h(x)), g(E) \} \}$$

여기에서 $\text{Supp}(A)$ 은 퍼지집합 A 의 지지집합(support)을 나타내고, $Q(\mu_A(x), h(x))$ 는 확신도 $\mu_A(x)$ 를 평가치 $h(x)$ 에 적용하는 것을 나타낸다. 보상연산자가 퍼지적분에서 사용

될때 위의 식은 다음과 같이 된다.

$$\oint_A h(x) \circ g(\cdot) = \sup_{E \subseteq \text{Supp}(A)} \{ \min \{ \Phi_{x \in E} Q(\mu_A(x), h(x)), g(E) \} \}$$

4. 중요도 결정을 위한 유전자알고리즘

퍼지적분을 적용할때 평가항목에 대한 중요도는 퍼지척도의 성질을 만족하도록 부여되어야 한다. 일반적으로 평가항목에 대한 중요도는 사람에게 의해서 주관적으로 할당되어야 하기 때문에 퍼지척도의 성질을 만족하는 중요도를 부여하는 것이 쉽지 않다. 이러한 이유로 사람이 부여한 중요도(즉, 퍼지척도값)로부터 퍼지척도의 성질을 만족하는 퍼지척도의 값을 식별해 내는 퍼지척도 식별(fuzzy measure identification)문제가 발생한다[7].

본 연구에서는 퍼지적분에서 사용되는 퍼지척도로서 γ -퍼지척도를 사용한다. 따라서 중요도에 대한 퍼지척도 식별에서는 퍼지밀도값(fuzzy density value) $g_i = g_\lambda(\{x_i\})$ 와 파라미터 λ 를 결정하면 된다. 편의상 $\hat{g}_\lambda(A)$, $A \subset B(X)$ 와 \hat{g}_i 는 사람이 부여한 값을 나타내고, $g_\lambda(A)$ 와 g_i 는 식별된 결과값을 나타낸다.

λ -퍼지척도의 식별을 위한 유전자알고리즘은 두개의 부분 GA-I와 GA-II로 구성되어 있다[3]. GA-I은 퍼지밀도값을 결정하고, GA-II는 GA-I에서 얻어진 퍼지밀도값에 대해서 λ 값을 결정하는 역할을 한다. GA-I에서 염색체(chromosome)의 표현은 퍼지밀도값 g_j^i 와 λ 값을 나타내기 위해 벡터 $C_j = (g_1^j, g_2^j, \dots, g_k^j; \lambda_j)$ 로 표현한다. 염색체 C_j 에 대한 적합도함수 $f_1(C_j)$ 는 사람이 부여한 퍼지척도값 $\hat{g}_\lambda(A)$ 와, g_j^i 와 λ_j 에 의해 계산되는 퍼지척도값 사이의 제곱합의 근으로 한다.

$$f_1(C_j) = \sqrt{\sum_{A \in B(X)} (\hat{g}_\lambda(A) - \frac{1}{\lambda_j} [\prod_{x_i \in A} (1 + \lambda_j g_i^j) - 1])^2}$$

GA-I에 대한 유전연산자(genetic operator)로서 교차연산자(crossover)와 돌연변이연산자(mutation)가 있다. 교차연산자는 두개의 무작위로 선택된 염색체 C_1 과 C_2 로부터 임의의 위치 i 를 선택하여 다음과 같이 새로운 염색체 C' 을 생성한다.

$$\begin{aligned} C_1 &= (g_1^1, g_2^1, \dots, g_i^1, g_{i+1}^1, \dots, g_k^1; \lambda_1) \\ C_2 &= (g_1^2, g_2^2, \dots, g_i^2, g_{i+1}^2, \dots, g_k^2; \lambda_2) \\ C' &= (g_1^1, g_2^1, \dots, g_i^2, g_{i+1}^1, \dots, g_k^1; \lambda_1) \end{aligned}$$

교차연산자는 퍼지밀도값 g_j^i 에만 영향을 미치고 λ 값에는 무관하다. 생성된 염색체 C' 에 대한 λ' 값은 GA-II에서 결정된다. 돌연변이연산자는 무작위로 선택된 염색체 C_1 에 대해서 하나의 위치 i 를 선택한 다음, 지정된 구간에서 무작위로 선택한 값 r 을 사용하여 i 번째 위치를 다음과 같이 변경하여 새로운 염색체 C' 을 생성한다. 이때 λ' 의 값은 GA-II에 의해 결정된다.

$$\begin{aligned} C_1 &= (g_1^1, g_2^1, \dots, g_i^1, \dots, g_k^1; \lambda_1) \\ C' &= (g_1^1, g_2^1, \dots, g_i^1 + r, \dots, g_k^1; \lambda') \end{aligned}$$

GA-II에서 염색체는 λ 값을 나타내기 위해 하나의 실수값 $C_j = (\lambda_j)$ 으로 나타낸다. 염색체 C_j 에 대한 적합도함수

$f_2(C_j)$ 는 다음과 같이 정의된다.

$$f_2(C_j) = \sqrt{\sum_{A \in B(X)} (\hat{g}_\lambda(A) - \frac{1}{\lambda_j} [\prod_{x_i \in A} (1 + \lambda_j \hat{g}_i) - 1])^2}$$

GA-II에 대한 교차연산자는 두개의 염색체 $C_1 = (\lambda_1)$ 와 $C_2 = (\lambda_2)$ 를 선택한 다음, 구간 $[0, 1]$ 에서 무작위로 선택한 r 를 사용하여 새로운 염색체 $C' = (r\lambda_1 + (1-r)\lambda_2)$ 를 생성한다. 교차연산자는 염색체 $C_j = (\lambda_j)$ 를 선택하여 지정된 구간에서 무작위로 선택된 값 r 를 사용하여 $C'_j = (\lambda_j + r)$ 로 변경한다.

중요도에 대한 퍼지척도 식별을 위한 유전자알고리즘은 GA-I과 GA-II가 번갈아 수행되면서 퍼지밀도값 g_i 와 λ 값을 결정한다.

5. 확장된 퍼지적분의 파라미터 결정을 위한 유전자알고리즘

확장된 퍼지적분에서 결정되어야 할 파라미터로는 보상정도 γ 와 평가항목들에 대한 상대적 중요도 δ_i 가 있다. 이들 파라미터를 결정하기 위한 유전자알고리즘은 다음과 같은 후보해법에 대한 염색체표현, 모집단초기화, 유전연산자, 평가함수 등으로 구성된다.

5.1 염색체 표현

보상정도에 대한 파라미터만 결정하는 경우에는 염색체가 하나의 실수값 $C_i = (\gamma_i)$ 로 표현되고, 보상정도와 상대적 중요도를 모두 결정할 때는 $C_i = (\gamma_i, \delta_1^i, \delta_2^i, \dots, \delta_n^i)$ 으로 염색체를 표현한다.

5.2 모집단 초기화

보상정도만을 결정하는 경우에는 구간 $[0, 1]$ 에서 무작위로 값을 선택하여 모집단 내의 염색체를 초기화한다. 보상정도와 평가항목의 상대적인 중요도를 모두 고려하는 경우에는 다음과 같이 초기화를 수행한다. 보상정도에 대한 γ_i 는 구간 $[0, 1]$ 에서 무작위로 값을 선택하여 초기화한다. 상대적인 중요도를 초기화하기 위해서는 임의의 평가항목이 갖을 수 있는 최대 중요도값 Δ 를 결정한 다음, k 번째 항목의 중요도 δ_k^i 는 구간 $[0, \min\{n - \sum_{j=1}^{k-1} \delta_j, \Delta\}]$ 에서 무작위로 한 값을 선택하여 초기화한다. 단, n 번째 항목의 중요도는 $n - \sum_{j=1}^{n-1} \delta_j$ 로 초기화한다.

5.3 유전연산자

확장된 퍼지적분의 파라미터 결정을 위한 유전연산자로는 교차연산자와 돌연변이연산자를 사용한다. 보상정도 γ 만을 결정하는 유전자알고리즘의 경우에 대한 연산자는 다음과 같이 동작한다. 교차연산자는 두개의 염색체 $C_i = (\gamma_i)$ 과 $C_j = (\gamma_j)$ 를 선택한 다음, 구간 $[0, 1]$ 에서 무작위로 값 α 를 선택하여 새로운 염색체 $C' = (\alpha\gamma_i + (1-\alpha)\gamma_j)$ 를 생성한다. 돌연변이연산자는 염색체 $C_i = (\gamma_i)$ 에 대해 지정된 구간에서 무작위로 선택한 값 r 을 사용하여 염색체 $C' = (\gamma_i + r)$ 을 생성한다. 단, 이때 $\gamma_i + r \in [0, 1]$ 의 조건이 만족하도록 한다.

보상정도 γ 와 상대적 중요도 δ_i 를 모두 결정해야 하는 경우의 유전자알고리즘에서는 교차연산자가 다음과 같이 동작한다.

우선 두개의 염색체 C_i, C_j 를 선택한다.

$$C_i = (\gamma_i, \delta_1^i, \delta_2^i, \dots, \delta_k^i, \dots, \delta_n^i)$$

$$C_j = (\gamma_j, \delta_1^j, \delta_2^j, \dots, \delta_k^j, \dots, \delta_n^j)$$

이들 염색체에 대해 무작위로 하나의 교차점 k 를 선택한다. k 가 γ 값의 위치를 가리키는 경우에는 교차염산자에 의해서 아래와 같은 염색체 C_p 가 생성된다.

$$C_p = (\gamma_i, \delta_1^j, \delta_2^j, \dots, \delta_n^j)$$

k 가 δ_i 를 가리키는 경우에는 새로운 염색체 C_p 를 다음과 같이 생성한다.

$$C_p = (\gamma_p, \delta_1^p, \delta_2^p, \dots, \delta_k^p, \dots, \delta_n^p)$$

$$\gamma_p = \gamma_i$$

$$1 \leq s \leq k \text{ 일 때, } \delta_s^p = \delta_s^i$$

$$k < s \leq n \text{ 일 때, } \delta_s^p = \delta_s^j - \left[\frac{\sum_{i=1}^k \delta_i^i + \sum_{i=k+1}^n \delta_i^j - n}{n - k} \right]$$

돌연변이염산자는 하나의 염색체 C_i 를 선택하여 그 염색체에 대해서 두개의 위치 s, t 를 선택한다. s 나 t 가 γ 의 위치를 가리키는 경우에는 지정된 구간에서 무작위로 선택된 값 r ($\gamma_i + r \in [0, 1]$)을 사용하여 새로운 염색체 C_p 를 생성한다.

$$C_p = (\gamma_i + r, \delta_1^i, \delta_2^i, \dots, \delta_n^i)$$

s 와 t 가 δ 값들의 위치를 가리킬 경우에는 지정된 구간에서 무작위로 선택한 값 r 을 사용하여 s 위치의 값에는 r 을 더해주고, t 위치의 값에서는 r 을 빼주어서 새로운 염색체 C_p 를 생성한다.

$$C_p = (\gamma_i, \delta_1^i, \dots, \delta_s^i + r, \dots, \delta_t^i - r, \dots, \delta_n^i)$$

5.4 평가합수

확장된 퍼지적분을 사용할때 파라미터의 값을 결정하기 위해 현상에 대한 실제 데이터를 수집하여 이용한다. 이때 데이터는 대상들에 대한 각 평가항목별 평가치와 사람이 부여한 전체적인 평가치로 구성된다. 염색체 C_i 에 대한 평가합수 $f(C_i)$ 는 염색체에 표현된 파라미터를 이용하여 항목별 평가치 h_k 에 대해 수행한 퍼지적분의 결과 $fint(C_i, h_k)$ 와 사람이 평가치 h_k 를 고려하여 결정한 전체적인 평가치 d_k 의 차이의 제곱합의 근을 사용한다.

$$f(C_i) = \sqrt{\sum_k (fint(C_i, h_k) - d_k)^2}$$

6. 실험

제안된 방법을 실험하기 위해 확장된 퍼지적분을 사용하여 중고차의 선호도를 결정하는 실험을 했다. 실험에서는 평가항목으로 주행거리, 외양, 유지비, 차령, 사고이력 등을 사용했다.

퍼지적분을 적용하기 위해서는 먼저 평가항목의 중요도를 결정해야 한다. 이를 위해 평가항목에 대한 중요도를 사람이 직접 부여하는 방법으로 데이터를 수집한 다음, 이 데이터에 대해 γ -퍼지적도를 식별하는 유전자알고리즘을 적용하여 중요도를 선정했다.

중고차의 평가항목별 평가치에 대해 중요도를 고려하여 전체적인 평가치를 계산하는 모델로서 weighted arithmetic

mean 연산자를 사용하는 확장된 퍼지적분을 사용했다. 이때 연산자에서 보상정도에 대한 파라미터 γ 를 결정하기 위해, 실제 항목별 평가치에 대해 사람이 느끼는 전체적인 평가치를 수집하여, 이 데이터에 대해 제안된 유전자알고리즘을 적용하여 γ 값을 결정했다. 결정된 파라미터를 사용하여 확장된 퍼지적분을 적용한 결과 비교적 만족스러운 결과를 얻을 수 있었다.

7. 결론

정보합성, 증거결합 등에서 평가항목별로 대상을 평가하고, 평가항목별 중요도를 고려하여 평가치를 합성하는 방법으로 퍼지적분이 유용하다. 그런데 퍼지적분을 사용할 때는 퍼지적도의 성질에 부합되는 중요도를 결정하는 것과, 확장된 퍼지적분을 사용할 경우에 보상정도, 상대적 중요도에 대한 파라미터를 결정하는 것이 부담이 된다. 따라서 본 논문에서는 유전자알고리즘에 의해서 이들 문제를 해결하는 방법을 제안했다.

참고 문헌

- [1] L. Davis, *Handbook of Genetic Algorithms*(eds.), Van Nostrand Reinhold:New York, 1991.
- [2] K.De Jong, Learning with Genetic Algorithms: An Overview, *Machine Learning*, Vol.3, pp.121-138, 1988.
- [3] 이진명, 이광형, "유전자 알고리즘에 의한 퍼지적도의 식별", 한국퍼지시스템학회, '94춘계학술대회 논문집, Vol.4, No.1, pp.96-101, 1994.
- [4] K.-M. Lee, H. Leekwang, Genetic algorithms for Fuzzy Measure Identification, *the 3rd International Conference on Fuzzy Logic, Neural Networks, and Soft Computing*(Iizuka, Japan), pp.461-463, 1994.
- [5] K.-M. Lee, C.B. Chung, H. Leekwang, Information Aggregating Networks based on Fuzzy Integral, *Proc. of the 1994 IEEE/Nagoya University WWW on Fuzzy Logic and Neural Networks/Genetic Algorithms*(Nagoya, Japan), pp.41-45, 1994.
- [6] D. Ralescu, G. Adams, The Fuzzy Integral, *Journal of Mathematical Analysis and Applications*, Vol.75, pp.562-570, 1980.
- [7] S.T. Wierzchoń, An Algorithm for Identification of Fuzzy Measure, *Fuzzy Sets and Systems*, Vol.9, pp.69-78, 1983.
- [8] S.T. Wierzchoń, On Fuzzy Measure and Fuzzy Integral, *Fuzzy Information and Decision Processes*:M.M. Gupta and E. Sanchez(eds.), pp.79-86, 1982.
- [9] H.-J. Zimmermann, *Fuzzy Set Theory - and Its Applications*, Kluwer-Nijhoff Publishing: Boston, 364p, 1985.