

# 언어 값을 다루기 위한 알고리즘적인 접근법

최 대 영<sup>†</sup>

요 약

같은 언어 변수에서 정의된 인접 언어 값을 다루기 위한 알고리즘을 제안하였다. 제안된 방법을 사용해서 인접 언어 값에 대한 개인의 주관성의 차이를 명확히 발견할 수 있다. 제안된 방법은 같은 언어 변수에서 정의된 인접 언어 값들 사이의 숨겨진 관계를 발견하기 위한 유용한 도구로 이용될 수 있다. 결과적으로 제안된 방법은 퍼지 전문가시스템, 퍼지 의사결정 트리, 퍼지 인지 지도등과 같은 퍼지 시스템 개발 분야에서 지식 획득의 정확성을 개선하기 위한 기초를 제공해 줄 수 있다. 본 논문에서 제안된 방법을 다수의 전문가들 사이의 집단 언어 평가에 적용하였다.

## Algorithmic approach for handling linguistic values

Choi, Dae Young<sup>†</sup>

ABSTRACT

We propose an algorithmic approach for handling linguistic values defined in the same linguistic variable. Using the proposed approach, we can explicitly capture the differences of individuals' subjectivity with respect to linguistic values defined in the same linguistic variable. The proposed approach can be employed as a useful tool for discovering hidden relationship among linguistic values defined in the same linguistic variable. Consequently, it provides a basis for improving the precision of knowledge acquisition in the development of fuzzy systems including fuzzy expert systems, fuzzy decision tree, fuzzy cognitive map, etc. In this paper, we apply the proposed approach to a collective linguistic assessment among multiple experts.

키워드: 지식 획득(Knowledge acquisition), 언어 값(Linguistic values), 집단 언어 평가(Collective linguistic assessment)

### 1. 서 론

선택과 등급화를 위한 언어적인 평가(Linguistic assessment)는 인간의 의사결정 활동과 매우 긴밀한 관련이 있다. 대부분의 경우에 전문가들의 의견들은 주관적이고 모호한 언어적인 형태로 표현된다. 이러한 경우 발생하는 주요한 문제점은 정보 획득(Information acquisition)의 문제와 어떻게 그러한 언어적인 형태로 표현된 언어적인 평가를 적절히 모델링하는 것이다. 그러나, 주관적인 전문가의 의견을 정량화(Quantization) 하는 것은 중요성에 비해서 아직 많은 관심을 받지 못하고 있다. 몇몇 소수의 연구 결과들이 이러한 범주의 연구에 들어간다고 볼 수 있다. Zimmermann과 Zysno [18]는 퍼지 집합이론과 인간 의사결정의 경험적이고 실제적인 관련성을 보여주고 의사결정 이론 관점에서 신뢰 평가를 위한 평가의 틀(Evaluation framework)을 개발하였다. Biswas [2]는 학생들의 응답에 대한 퍼지 평가(Fuzzy evaluation)의

방법을 소개하였다. 그러나 그들의 방법은 퍼지 평가가 어떻게 효과적으로 생성되는가에 대해 고려하지 않았고 너무 단순한 절차를 사용하고 있다. 최근의 연구에서 Chakraborty [4]는 개인 평가자의 언어적인 반응이 퍼지 집합에 의해 어떻게 구조적으로 정량화 하는가를 보여주고 있다. 비록 퍼지 집합 이론[13]에서 소개된 소속 함수(Membership function)는 명확한 정의를 가지고 있지만 그에 대한 경험적이고 실험적인 측정을 위한 명확한 요구사항들에 대한 연구는 미미한 실정이다. 한편, 전문가 시스템의 개발에서 다수의 전문가들의 지식을 적절히 표현하는 것은 오랫동안 연구 목표의 하나였다. 전문가들은 서로 다른 주관적인 크기(Scales)에 바탕을 두고 언어 값(Linguistic value)을 사용하여 그들의 대부분의 판단을 표현한다. 이러한 관점에서 다수의 전문가들 사이의 언어적 평가에 내재된 불확실성(Uncertainty), 모호성(Vagueness), 부정확(Imprecision)을 다루기 위한 새로운 도구가 필요하다. 이러한 문제점을 다루기 위해 같은 언어 변수(Linguistic variable) [14-16]에서 정의된 인접 언어 값(Adjacent linguistic values)을 다루기 위한 알고리즘을 제안 하고자 한다. 프로그래밍언어에서 변수가 같은 자료형의 다양한 값들을 가질 수 있는 것처럼 인간이 사용하는 언

\* 이 논문은 2004년도 한국학술진흥재단의 지원에 의하여 연구되었음(KRF-2004-013-D00040)

† 경 회 원 : 유한대학 경영정보과 부교수

논문접수 : 2004년 3월 5일, 심사완료 : 2005년 4월 20일

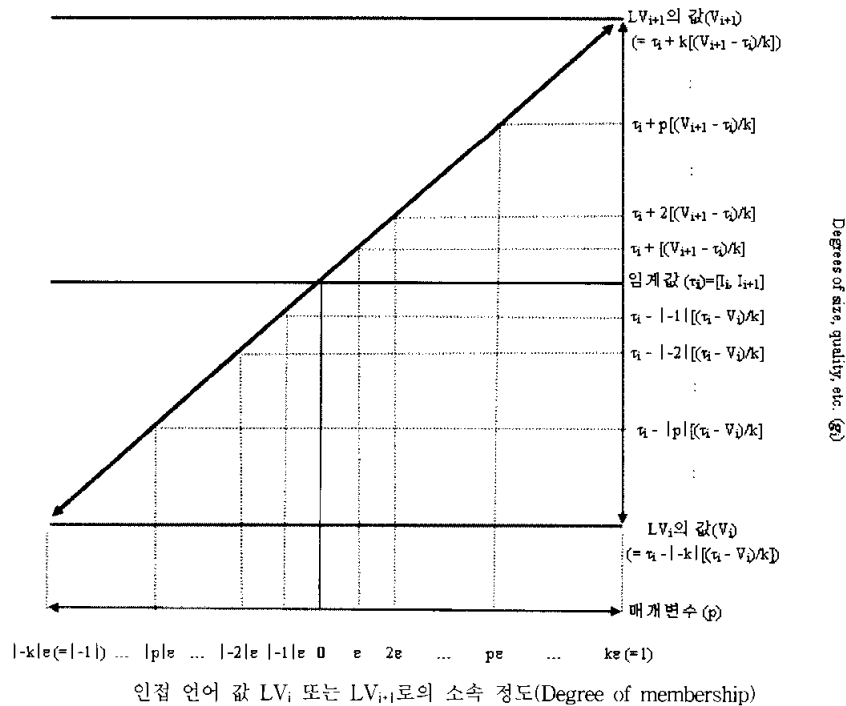
어에서도 언어 변수가 그에 관련된 다양한 언어 값들을 가질 수 있다. 예를 들어 '나이가 젊다'에서 '나이'는 언어 변수이고 '젊다'는 언어 값이 된다. 이 경우 '젊다', '매우 젊다' 등이 언어 변수 '나이'에 관련된 인접 언어 값들이라고 할 수 있다. 이 논문에서 제안하는 방법은 같은 언어 변수에서 정의된 인접 언어 값들 사이의 명확한 관계를 찾아 내는데 사용될 수 있다. 결과적으로 제안된 방법을 이용하여 다수 전문가들 사이의 언어적 평가에 내재된 차이를 명확히 발견할 수 있게 하여 다수 전문가들 사이의 언어적 평가에 내재된 모호성을 줄일 수 있다. 이러한 관점에서 다수 전문가들 사이의 보다 정교한 집단 언어 평가(Collective linguistic assessment)를 만들어 낼 수 있게 한다. 결과적으로 제안된 방법은 퍼지 시스템 개발에서 지식 획득(Knowledge acquisition)의 정확성을 개선하기 위한 기초를 제공해 줄 수 있다.

본 논문은 2장에서 같은 언어 변수에서 정의된 인접 언어 값을 다루기 위한 알고리즘을 제안한다. 3장에서는 제안된 방법에 기반을 둔 집단 언어 평가 방법을 제시한다. 4장에서 결론에 대해 논한다.

## 2. 인접 언어 값을 다루기 위한 알고리즘

퍼지 집합 이론의 주요한 등장 배경은 자연어로 표현된 인간 지식의 모호성을 표현하기 위한 정량적인 프레임워크를 만들기 위한 것이었다[6, 13]. 언어적인 평가(Linguistic assessment)는 인간의 의사결정 활동과 매우 긴밀한 관련이 있다. 예를 들어 어떤 평가에서 '좋다'와 '매우 좋다', 길

이에서 '길다'와 '매우 길다', 양의 크기에서 '많다', '적다' 등과 같이 숫자로 표현될 수 없는 많은 부분에서 언어적인 평가를 사용하고 있고 이러한 평가는 대부분 정성적(定性的)인 특성을 가지고 있다. 이러한 숫자로 표현될 수 없는 정성적인 언어적 평가에 사용되는 질의 정도(Degree of quality), 물체의 크기 정도(Degree of size) 등의 표현에는 불확실성, 모호성, 부정확 등이 내재되어 있다. 이러한 평가에 기반을 두고 많은 경우에 있어 인간의 추론(Human reasoning)과 개념 형성(Concept formation)이 만들어지기 때문에 이러한 평가 결과는 이분법적이라기 보다는 퍼지(Fuzzy)한 특성을 갖게 된다. 이러한 퍼지한 정보를 표현하기 위해 퍼지 집합 이론에서 소개된 소속 함수는 명확한 정의를 가지고 있지만 그에 대한 경험적이고 실험적인 측정을 위한 명확한 요구사항들에 대한 연구는 미미한 실정이다. 한편, 전문가 시스템의 개발에서 다수의 전문가들의 지식을 적절히 표현하는 것은 오랫동안 연구 목표의 하나였다. 전문가들은 서로 다른 주관적인 크기(Scales)에 바탕을 두고 언어 값(Linguistic value)을 사용하여 그들의 대부분의 판단을 표현한다. 이러한 관점에서 다수 전문가들 사이의 언어적 평가에 내재된 불확실성, 모호성, 부정확을 다루기 위한 새로운 도구가 필요하다. 이러한 문제점을 다루기 위해 같은 언어 변수(Linguistic variable)[14-16]에서 정의된 인접 언어 값(Adjacent linguistic values)을 다루기 위한 알고리즘을 제안하고자 한다. (그림 1)에서  $LV_{i+1}$ 과  $LV_i$ 를 같은 언어 변수에서 정의된 인접 언어 값들을 나타낸다고 하자. 예를 들어 어떤 평가에서 '좋다'와 '매우 좋다'와 같이 언어적 평가시에 인접되는 언어 값이 인접 언어 값이 될 수 있다.



(그림 1) 인접되는 언어 값들의 표현

(그림 1)에서 어떤 값  $V_i$ 와  $V_{i+1}$ 는 각각 인접 언어 값  $LV_i$ 과  $LV_{i+1}$ 를 위한 퍼지 부분집합(Fuzzy subset)의 커널(Kernel)에 포함된다. 즉, 어떤 값  $V_i$ 와  $V_{i+1}$ 는 각각 인접 언어 값  $LV_i$ 과  $LV_{i+1}$ 로의 소속 정도(Degree of membership)가 1이 된다. 점 또는 구간  $[I_i, I_{i+1}]$  (단,  $I_i \leq I_{i+1}$ ), 즉, 임계값 ( $\tau_i$ )에 서는 언어 값  $LV_i$ 와  $LV_{i+1}$ 사이에서 미결정 부분(Non-deterministic region)이 발생될 수 있다. 언어적인 평가에 대한 정보가 퍼지(Fuzzy)할수록 구간  $[I_i, I_{i+1}]$ 이 커지게 된다. (그림 1)에서 값  $V_i, V_{i+1}, I_i, I_{i+1}$  등은 인접 언어 값에 대한 개인의 주관성(Subjectivity)에 따라 다르게 결정된다.

**Remark 1.** (그림 1)에서  $\varepsilon$ 은 구간  $[-1, 1]$ 를  $[-k\varepsilon (= -1), -2\varepsilon, -\varepsilon, 0, \varepsilon, 2\varepsilon, \dots, k\varepsilon (=1)]$ , (단,  $0 < \varepsilon \leq 1, k = (1/\varepsilon)$ )로 나눈다. 예를 들어  $k$ 와  $\varepsilon$ 사이의 관계는 만약  $\varepsilon=10^{-n}$  (단,  $n$ 은 양의 정수) 라고 하면 식  $k=10^n$ 에 의해 결정될 수 있다.

**Remark 2.** (그림 1)에서 구간  $[V_i, V_{i+1}]$ 는 다음과 같이  $(2k+1)$  단계로 나뉘어진다:  $\tau_i - k[(\tau_i - V_i)k] (= V_i), \tau_i - (k-1)[(\tau_i - V_i)k], \dots, \tau_i - 2[(\tau_i - V_i)k], \tau_i - [(\tau_i - V_i)k], \tau_i, \tau_i + [(V_{i+1} - \tau_i)k], \tau_i + 2[(V_{i+1} - \tau_i)k], \dots, \tau_i + (k-1)[(V_{i+1} - \tau_i)k], \tau_i + k[(V_{i+1} - \tau_i)k] (= V_{i+1})$ . 이때 주어진 정성적인 언어적 평가에 사용되는 질의 정도(Degree of quality), 물체의 크기 정도(Degree of size)등을 나타내는  $g_i$ 는  $V_i$ 와  $V_{i+1}$ 사이에 있다고 가정한다(즉,  $V_i \leq g_i \leq V_{i+1}$ ). 언어 값은 크게 단조 증가 언어 값(Monotonically non-decreasing linguistic values)과 단조 감소 언어 값(Monotonically non-increasing linguistic values)의 두가지 형태로 나눌 수 있다. 단조 증가 언어 값은  $g_i$ 가 증가 함에 따라 (그림 1)에서 나타난 것처럼  $V_{i+1}$ 로 수렴하는 성질을 갖는 반면, 단조 감소 언어 값은  $g_i$ 가 증가 함에 따라  $V_i$ 로 수렴하는 성질을 갖는다. 예를 들어 단조 증가 언어 값의 경우 언어 변수를 어떤 웹 사이트의 '유명도(Popularity)' 라고 할 때 인기 정도는 웹 사이트의 방문자 수에 기반을 두고 언어적으로 평가할 수 있다. 이 경우  $LV_i$ 와  $LV_{i+1}$ 는 인접 언어 값 '유명함'과 '매우 유명함'을 각각 나타내고 그들의 값  $V_i$ 와  $V_{i+1}$ 는 각각의 방문자 수를 나타낸다고 하면,  $g_i$ 가 증가 함에 따라  $V_{i+1}$ 로 수렴한다. 이와는 반대로 단조 감소 언어 값의 경우에 만약  $LV_i$ 와  $LV_{i+1}$ 는 인접 언어 값 '유명하지 않은'과 '매우 유명하지 않은'을 각각 나타내고 그들의 값  $V_i$ 와  $V_{i+1}$ 는 각각의 방문자 수를 나타낸다고 하면,  $g_i$ 가 증가 함에 따라  $V_i$ 로 수렴한다. 이 논문에서는 (그림 1)에서와 같이 '유명함', '매우 유명함' 같은 단조 증가 언어 값을 가지고 설명한다. 한편, '유명하지 않은', '매우 유명하지 않은' 같은 단조 감소 언어 값은 단조 증가 언어 값의 역으로 생각하면 쉽게 다룰 수 있다.

**Remark 3.** (그림 1)에서 언어 값  $LV_i$ 로의 단계(Level)별 크기 ( $D_{del}$ )는  $[(\tau_i - V_i)k]$ 이고, 언어 값  $LV_{i+1}$ 로의 단계별 크기 ( $U_{del}$ )는  $[(V_{i+1} - \tau_i)k]$ 이다.

**Remark 4.** 주어진  $g_i$ 와 어떤 언어 값으로의 소속 정도(Degree of membership)의 관계를 연결시켜 주는 매개변수  $p$ 는 다음과 같이 계산된다: (i)  $g_i > \text{임계값}(\tau_i)$ 일 경우:  $p = (g_i - \tau_i) / (U_{del})$ , (ii)  $g_i < \text{임계값}(\tau_i)$ 일 경우:  $p = (g_i - \tau_i) / (D_{del})$ .

**Definition 1.** Remark 1, 2, 3, 4에 기반을 두고 정성적인 언어적 표현으로 크기, 속도, 색깔 정도(Degree)등과 같은 인지(Perception)의 정도( $g_i$ )를  $(2k+1)$  단계로 매개변수화 할 수 있다.

Case 1:  $g_i > \text{임계값}(\tau_i)$ 일때 언어 값  $LV_{i+1}$ 로 상향(Upward)으로 발생한다. 상향은  $k$  단계로 매개변수화 할 수 있다 (단,  $0 < p \leq k$ ). ( $k$  단계 발생)

Case 2:  $g_i = \text{임계값}(\tau_i)$ 일때 미결정 상황이 발생한다. 이 경우  $p = 0$  (즉, (그림 1)에서 임계값( $\tau_i$ )). (1단계 발생)

Case 3:  $g_i < \text{임계값}(\tau_i)$ 일때 언어 값  $LV_i$ 로 하향(Downward)으로 발생한다. 하향은  $k$  단계로 매개변수화 할 수 있다 (단,  $-k \leq p < 0$ ). ( $k$  단계 발생)

**Definition 2.** 언어 값  $LV_{i+1}$ 로 상향(Upward) (즉,  $\tau_i + p$  ( $U_{del}$ ))과 언어 값  $LV_i$ 로 하향(Downward) (즉,  $\tau_i - |p|$  ( $D_{del}$ ))는 주어진  $g_i$ 에 관한 인지(Perception)의 정도를 나타낸다. (그림 1)에서 매개변수  $p$ 는 주어진  $g_i$ 와 어떤 언어 값으로의 소속 정도(Degree of membership)의 관계를 연결시켜 주는 역할을 한다. 매개변수  $p$ 는 구간  $[-k, k]$ 에 있는 실수 값 (Real number) 이고, 이때  $k$ 는  $V_i$ 나  $V_{i+1}$ 로의 단계 수이다. 언어 값  $LV_i$ 나  $LV_{i+1}$ 로의 소속 정도는 매개변수  $p$ 값에 따라 결정된다. 즉, 언어 값  $LV_i$ 나  $LV_{i+1}$ 로의 소속 정도는 다음과 같이 결정된다: (i) 하향인 경우  $\tau_i - |p|$  ( $D_{del}$ )에서 언어 값  $LV_i$ 로의 소속 정도는  $|p| \varepsilon$ 이고 (ii) 상향인 경우  $\tau_i + p$  ( $U_{del}$ )에서 언어 값  $LV_{i+1}$ 로의 소속 정도는  $p \varepsilon$ 가 된다 (그림 1 참조).

Zadeh[17]는 실세계의 문제를 다루는데 있어 퍼지 논리(Fuzzy logic)의 성공과 효과성은 대부분 퍼지 정보(Fuzzy information)를 적절하게 표현하는데 있다고 하였으며, 이러한 관점에서 우리가 소홀히 하고 있는 점은 우리가 퍼지 정보에 대해 이야기할 때 하나의 퍼지 값만을 말하는 것이 아니고 여러 퍼지 값들의 관계에 대해 말하는 것이라고 하였다. 인간 인지(Human cognition)의 가장 기본적인 측면은 주어진 정보간의 관계(Relation)와 의존성(Dependency)에 관련되어 있다. 이러한 측면에서 같은 언어 변수에서 정의된 인접 언어 값의 관계를 다루기 위한 알고리즘을 제안한다. Remark 1, 2, 3, 4, Definition 1, 2에 기반을 두고 제안된 알고리즘은 다음과 같이 실행된다.

**알고리즘 1 :**

/\*  $k\varepsilon = 1$ 라 하자 (단,  $0 < \varepsilon \leq 1$ ). 매개변수  $p$ 는 주어진  $g_i$ 와 어떤 언어 값으로의 소속 정도(Degree of membership)의 관계를 연결시켜 주는데 사용된다. 이는 구간  $[-k, k]$ 에

있는 어떤 실수 값이다. \*/

Step 1: 주어진  $g_i$ 를 얻는다. /\* 이때  $g_i$ 는 정성적인 크기, 속도, 색깔 정도(Degree)등과 같은 인지(Perception)의 정도를 나타낸다. \*/

Step 2: 인접 언어 값을 결정한다 (즉,  $LV_i$ 와  $LV_{i+1}$ ) (단,  $V_i \leq g_i \leq V_{i+1}$ ).

Step 3: 분류 매개변수 (Classification parameter) 들을 결정한다 (즉,  $V_i, V_{i+1}, \tau_i$ ).

Step 4: 정밀도 매개변수(Preciseness parameter)를 결정한다 (즉,  $\epsilon$ ).

Step 5: 주어진  $g_i$ 와 임계값  $\tau_i$ 를 비교한다.

Case  $g_i > \tau_i$ : /\* 언어 값  $LV_{i+1}$ 로의 상향 발생 \*/

- 언어 값  $LV_{i+1}$ 로의 상향 단계별 크기 ( $U_{dgi}$ )를 계산한다:  $U_{dgi} = [(V_{i+1} - \tau_i)k]$  (Remark 3 참조).
- 매개변수  $p$ 를 계산한다:  $p = (g_i - \tau_i)(U_{dgi})$ .
- 언어 값  $LV_{i+1}$ 로의 상향 정도( $U_g$ )를 계산한다:  $U_g = \tau_i + p$  ( $U_{dgi}$ ).
- $U_g$ 에서의 언어 값  $LV_{i+1}$ 로의 소속 정도(Degree of membership)를 계산한다:  $\mu_{LV_{i+1}}(U_g) = p\epsilon$ .

Case  $g_i = \tau_i$ : 미결정 상황(Non-deterministic situation)이 발생한다 (즉, (그림1)에서 임계값( $\tau_i$ )일 때 발생). 이때는  $p = 0$ 가 되고 주어진  $g_i$ 에 대해 '언어 값  $LV_i$ 와  $LV_{i+1}$ 의 사이'로 평가 된다.

Case  $g_i < \tau_i$  : /\* 언어 값  $LV_i$ 로의 하향 발생 \*/

- 언어 값  $LV_i$ 로의 하향 단계별 크기 ( $D_{dgi}$ )를 계산한다:  $D_{dgi} = [(\tau_i - V_i)k]$  (Remark 3 참조).
- 매개변수  $p$ 를 계산한다:  $p = (g_i - \tau_i)(D_{dgi})$ .
- 언어 값  $LV_i$ 로의 하향 정도( $D_g$ )를 계산한다:  $D_g = \tau_i - |p| (D_{dgi})$ .
- $D_g$ 에서의 언어 값  $LV_i$ 로의 소속 정도(Degree of membership)를 계산한다:  $\mu_{LV_i}(D_g) = |p| \epsilon$ .

알고리즘 1에서 만약 매우 작은 값을 사용한다면 (즉,  $\epsilon \rightarrow 0$ ), 연속적인 퍼지 정보를 다룰 수 있다. 결과적으로 제안된 알고리즘은 이산적이거나 연속적인 퍼지 정보를 다룰 수 있다. 예를 '사과의 개수', '방문자의 수'등을 '많다', '적다'와 같은 언어 값으로 표현할 때 '사과의 개수', '방문자의 수'등은 이산적이라 할 수 있는데 비해서, '머리의 색깔', '음질'등을 '갈색이다', '좋다'등의 언어 값으로 표현할 때 '머리의 색깔', '음질'등은 연속적이라 할 수 있다.

**Remark 5.** 제안된 알고리즘에서 정밀도 매개변수(Preciseness parameter)인  $\epsilon$  (단,  $0 < \epsilon \leq 1$ ) 값을 조정해서 언어 값  $LV_i$ 나  $LV_{i+1}$ 로의 단계 수를 변화시킬 수 있다.

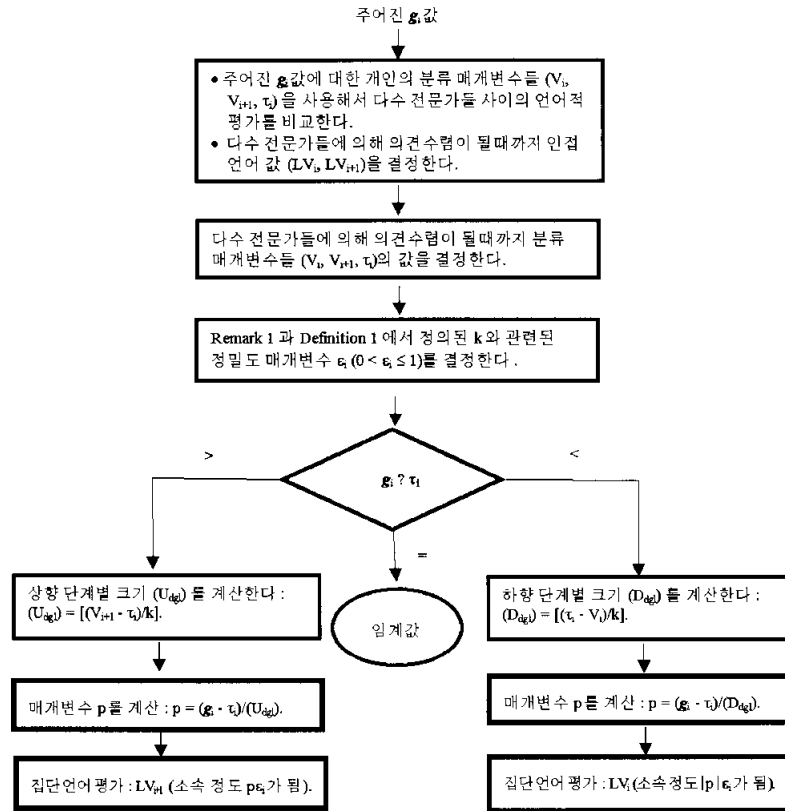
**Remark 6.** 일반적으로 단조성(Monotonicity)은 인간의 추론과 개념 형성(Concept formation)에 편리하다. 제안된 알고리즘은 (그림 1)에서 처럼 단조성에 기반을 두고 설계되었다.

**Remark 7.** 어떤 언어 값으로의 소속 정도는 주어진  $g_i$ 의 정도에 따라 단조 적용적으로(Monotonically and adaptively) 변한다 ((그림 1)과 [예 1] 참조). 또한, 방향성을 가지고 주어진  $g_i$ 의 감소와 증가에 따라 각각 언어 값  $LV_i$ 나  $LV_{i+1}$ 로의 소속 정도가 1에 수렴한다.

### 3. 다수 전문가에 의한 집단 언어적 평가 (Collective Linguistic assessment)

전문가 시스템의 개발에서 다수의 전문가들의 지식을 적절히 표현하는 것은 오랫동안 연구 목표의 하나였다[5]. 전문가들은 서로 다른 주관적인 크기(Scales)에 바탕을 두고 언어 값을 사용하여 그들의 대부분의 판단을 표현한다. 이러한 관점에서 다수 전문가들 사이의 언어적 평가에 내재된 불확실성, 모호성, 부정확을 다루기 위한 새로운 도구가 필요하다. 이러한 문제점을 다루기 위해 본 논문에서는 같은 언어 변수 [14-16]에서 정의된 인접 언어 값을 다루기 위한 알고리즘을 제안하였다. 제안된 방법은 분류 매개변수 (즉,  $V_i, V_{i+1}, \tau_i$ )와 정밀도 매개변수(즉,  $\epsilon$ )의 값들을 사용해서 같은 언어 변수에서 정의된 인접 언어 값들 사이의 관계를 명확하게 찾아 낼 수 있다 ((그림 1) 참조). 또한 제안된 방법을 사용해서 인접 언어 값에 대한 개인간의 주관성(Subjectivity)의 차이를 발견할 수 있다. 결과적으로 제안된 방법은 같은 언어 변수에서 정의된 인접 언어 값들 사이의 숨겨진 관계(Hidden relationship)를 발견하기 위한 유용한 도구로 이용될 수 있다. 이러한 관점에서 제안된 방법은 다수 전문가들 사이의 언어적 평가에 내재된 차이를 명확히 발견할 수 있는 도구로 이용될 수 있으므로 모든 전문가들이 같은 언어(The same language)를 말하도록 하는데 이용될 수 있다. 결과적으로 매개변수 (즉,  $V_i, V_{i+1}, \tau_i, \epsilon$ )의 값들을 조정해서 집단 언어 평가(Collective linguistic assessment)를 만들어 낼 수 있게 한다. (그림 2)에서 다수 전문가들 사이의 집단 언어 평가를 위한 절차를 나타내었다.

**[예 1]** 언어 변수가 어떤 웹 사이트의 '유명도(Popularity)'라고 할 때 인기 정도는 웹 사이트의 방문자 수에 기반을 두고 언어적으로 평가할 수 있다. 이 경우 주어진  $g_i$ 가 방문자 수가 된다. 이때  $g_i$ 가 (8630명/일)이고  $LV_i$ 와  $LV_{i+1}$ 는 인접 언어 값 '유명한' 과 '매우 유명한'을 각각 나타내고, 그들의 값  $V_i$ 와  $V_{i+1}$ 는 방문자 수를 (7,000명/일)과 (10,000명/일)로 각각 나타낸다고 하고, 임계값 ( $\tau_i$ )이 구간 [8400, 8600]이라고 하자. 만약  $= 10^2$ 를 제안된 알고리즘에 적용하면  $k = 10^2$ 이 된다. 이 경우  $g_i$ 가 상위 임계값( $\tau_{i-1}$ )인 8600보다 크므로 상향이 발생한다. 알고리즘 1에 의해 언어 값  $LV_{i+1}$ 로의 상향 단계별 크기는  $(V_{i+1} - \tau_i)/k = (10000 - 8600)/10^2 = 14$ 이 된다. 그리고 매개변수  $p$ 는  $[g_i - \tau_i]/[(V_{i+1} - \tau_i)/k] = [8630 - 8600]/[(10000 - 8600)/10^2] = 30/14 \approx 2.14$ 이 된다. 이 경우  $g_i$ 가 단계 2(즉, 8628)와 단계 3(즉, 8642) 사이에 놓이게 된다. 결과적으로 언어 값  $LV_{i+1}$  (즉, '매우 유명한') 로의 소속 정도는  $p \epsilon = 2.14 (10^{-2}) = 2.14/100$ 가 된다. 비슷한 방법으로  $g_i$ 가



\* 임계값 발생시에는 주어진  $g_i$  에 대해 '언어 값  $LV_i$ 와  $LV_{i+1}$ 의 사이'로 평가 된다.

(그림 2) 다수 전문가들 사이의 집단 언어 평가를 위한 절차

(8370명/일)이라면  $g_i$ 가 하위 임계값( $\tau_i$ )인 8400보다 작으므로 하향이 발생한다. 이때 매개변수  $p$ 는  $[g_i - \tau_i] / [(V_i - \tau_i) / k] = [8370 - 8400] / [(8400 - 7000) / 10^2] = -30 / 14 \approx -2.14$ 이 된다. 결과적으로 언어 값  $LV_i$  (즉, '유명한') 로의 소속 정도는  $|p| \epsilon = 2.14(10^{-2}) = 2.14/100$ 가 된다.

#### 4. 결 론

비록 퍼지 집합 이론에서 소개된 소속 함수는 명확한 정의를 가지고 있지만 그에 대한 경험적이고 실험적인 측정을 위한 명확한 요구사항들에 대한 연구는 미미한 실정이다[2, 4, 18]. 최근들어 Bagis[1]는 퍼지 논리 제어기의 성능을 개선하기 위해서 Tabu 검색에 기반을 두고 퍼지 규칙에 있는 입력과 출력 퍼지 집합의 소속함수를 최적화하는 방법을 제시했지만, 본 논문에서 제안한 같은 언어 변수에서 정의된 인접 언어 값들간의 관련성에 대해서는 해결 방법을 제시하고 있지 않다. 한편, Marchant[8]는 순서적이고 구간적인 크기를 갖는 퍼지 집합의 소속함수 측정을 위한 이론을 제시했지만 '어떤 값이 퍼지집합 B 보다 A에 속한다'는 인간의 평가를 위해서는 인간의 평가에 대한 경험적인 연구가 필요하다고 했다. 또한, Marchant[9]는 비율 측정에 기반을 둔 소속함수 측정 방법을 제시하고 측정의 수준을 결정하기 위한 방법을 제시하였다. 그는 두개의 퍼지집합 A, B에 대한

독립적인 소속함수  $\mu_A$ 와  $\mu_B$ 를 만드는 방법을 제시 하였지만 어떤 값  $X$ 에 대해  $\mu_A(x)$ 와  $\mu_B(x)$ 를 비교하는 두개의 퍼지 집합간의 관련성에 대해서는 해결 방법을 제시하고 있지 못하다. 본 논문에서는 같은 언어 변수에서 정의된 인접 언어 값을 다루기 위한 알고리즘을 제안하였는데 이는 퍼지 집합으로 표현된 인접 언어 값들간의 관련성에 대해 명확한 표현 방법을 제시하고 있다. 한편, 전문가 시스템의 개발에서 다수의 전문가들의 지식을 적절히 표현하는 것은 오랫동안 연구 목표의 하나였다[5]. 전문가들은 서로 다른 주관적인 크기(Scales)에 바탕을 두고 언어 값을 사용하여 그들의 대부분의 판단을 표현한다. 이러한 관점에서 다수 전문가들 사이의 언어적 평가에 내재된 불확실성, 모호성, 부정확을 다루기 위한 새로운 도구가 필요하다. 이러한 문제점을 다루기 위해 본 논문에서는 같은 언어 변수 [14-16]에서 정의된 인접 언어 값을 다루기 위한 알고리즘을 제안하였다. 제안된 방법은 같은 언어 변수에서 정의된 인접 언어 값들 사이의 명확한 관계를 매개변수들을 사용해서 찾아 낼 수 있다. 또한 제안된 방법을 사용해서 인접 언어 값에 대한 개인의 주관성의 차이를 발견할 수 있다. 결과적으로 제안된 방법은 같은 언어 변수에서 정의된 인접 언어 값들 사이의 숨겨진 관계를 발견하기 위한 유용한 도구로 이용될 수 있다. 이러한 관점에서 제안된 방법은 다수 전문가들 사이의 언어적 평가에 내재된 차이를 명확히 발견할 수 있는 도구

로 이용될 수 있으므로 모든 전문가들이 같은 언어(The same language)를 말하도록 하는데 이용될 수 있다. 결과적으로 매개변수 (즉,  $V_i, V_{i+1}, \tau_i, \epsilon$ )의 값들을 조정해서 집단 언어 평가를 만들어 낼 수 있게 한다. 제안된 방법을 사용해서 보다 정교한 집단 언어 평가를 만들어 낼 수 있으므로 제안된 방법은 퍼지 전문가시스템[3, 11], 퍼지 의사결정 트리[10, 12], 퍼지 인지 지도[5, 7]등과 같은 퍼지 응용 분야에서 지식 획득의 정확성을 개선하기 위한 기초를 제공해 줄 수 있다. 본 논문에서는 어떤 언어 값으로의 소속 정도는 주어진  $g_i$ 의 정도에 따라 단조 적응적으로 변하는 방법을 제시하였다. 또한, 방향성을 가지고 주어진  $g_i$ 의 감소와 증가에 따라 해당 언어 값으로의 소속 정도를 적응적으로 갖는 언어 평가를 만드는 방법을 제시하였다. 본 논문에서는 같은 언어 변수에서 정의된 인접 언어 값들 사이의 의미적 관련성(Semantic relationship)을 나타내기 위한 수학적인 모델을 제시하였다. 이는 인간의 언어적 평가에 내재된 불확실성, 모호성, 부정확을 다루기 위한 수단을 제공한다.

**참 고 문 헌**

[1] A. Bagis, "Determining fuzzy membership functions with Tabu search -an application to control", Fuzzy sets and systems, Vol. 139, pp. 209-225, 2003.

[2] R. Biswas, "An application of fuzzy sets in students' evaluation", Fuzzy sets and systems, Vol. 74, pp. 187-194, 1995.

[3] J. J. Buckley & W. Siler, "Fuzzy numbers for expert systems" : in M. M. Gupta and T. Yamakawa (Eds.), Fuzzy logic in knowledge-based systems, decision and control, Elsevier Science, pp. 153-172, 1988.

[4] D. Chakraborty, "Structural quantization of vagueness in linguistic expert opinions in an evaluation programme", Fuzzy Sets and Systems, Vol. 119, pp. 171-186, 2001.

[5] J. R. Cole & K. A. Persichitte, "Fuzzy cognitive mapping : Applications in education", Int. J. intelligent systems, Vol. 15, pp. 1-25, 2000.

[6] D. Dubois & H. Prade, "Fuzzy sets in approximate reasoning, part I : Inference with possibility distributions", Fuzzy sets and systems, Vol. 40, pp. 143-202, 1991.

[7] B. Kosko, "Fuzzy cognitive maps", Int. J. Man-machine studies, Vol. 24, pp. 65-75, 1986.

[8] T. Marchant, "The measurement of membership by comparisons", Fuzzy sets and systems, Vol. 148, pp. 157-177, 2004.

[9] T. Marchant, "The measurement of membership by subjective ratio estimation", Fuzzy sets and systems, Vol. 148, pp. 179-199, 2004.

[10] X. Wang, B. Chen, G. Qian & F. Ye, "On the optimization of fuzzy decision trees", Fuzzy sets and systems, Vol. 112, pp. 117-125, 2000.

[11] D. A. Waterman, A guide to expert systems, (Addison-Wesley, 1986).

[12] Y. Yuan & M. J. Shaw, "Induction of fuzzy decision trees", Fuzzy sets and systems, Vol. 69, pp. 125-139, 1995.

[13] L. A. Zadeh, "Fuzzy sets", Inform. and control, Vol. 8, pp. 338-353, 1965.

[14] L. A. Zadeh, "The concept of a linguistic variable and its applications to approximation reasoning", Inf. Sci., Vol. 8, part I, pp. 199-249, 1975.

[15] L. A. Zadeh, "The concept of a linguistic variable and its applications to approximation reasoning", Inf. Sci., Vol. 8, part II, pp. 301-357, 1975.

[16] L. A. Zadeh, "The concept of a linguistic variable and its applications to approximation reasoning", Inf. Sci., Vol. 9, part III, pp. 43-80, 1975.

[17] L. A. Zadeh, "Toward a theory of fuzzy information granulation and its centrality in human reasoning and fuzzy logic", Fuzzy sets and systems, Vol. 90, pp. 111-127, 1997.

[18] H. J. Zimmermann & P. Zysno, "Decisions and evaluations by hierarchical aggregation of information", Fuzzy sets and systems, Vol. 10, pp. 243-260, 1983.



**최 대 영**

e-mail : dychoi@yuhan.ac.kr  
 1985년 서강대학교 컴퓨터학과 졸업(공학사)  
 1985년~1990년 한국 국방 연구원(KIDA) 전산센터 연구원  
 1992년 서강대학교 이공대학원 컴퓨터학과 졸업(공학석사)

1994년 정보처리기술사(전자계산 조직응용 분야)  
 1996년 서강대학교 이공대학원 컴퓨터학과 졸업(공학박사)  
 2000년 한국과학재단 해외 Post-doc. 연수자로 선정됨  
 2001년~2002년 University of California, Berkeley (Post-doc.)  
 2001년 International Workshop on Fuzzy Logic and Internet, at UC, Berkeley (Local Chair)  
 2001년 BT-BISC (British Telecom-Berkeley Initiative in Soft Computing) 프로젝트 (Fellow)  
 2004년 한국학술진흥재단 대학교수 해외 방문 연구 연수자로 선정됨  
 2004년~2006년 University of Colorado, Denver(Visiting Professor)  
 1997년~현재 유한대학 경영정보과 부교수  
 관심분야 : Location-Based Services, Web Search Engines, Fuzzy Systems, Business Intelligence 등.