

## DEFINITION MEMBERSHIP FUNCTION BASED ON APPROACH TO HEDGE ALGEBRAS

DOAN VAN THANG<sup>1</sup>, DOAN VAN BAN<sup>2</sup>

<sup>1</sup>*Ho Chi Minh City Industry and Trade College*  
<sup>2</sup>*Duy Tan University*



**Abstract.** In this paper, we describe a hedge algebras based approach to modelling uncertainty in fuzzy object-oriented databases. Membership value reflects the degree of fuzziness existing in the data values and uncertainty is extended to the class definition level and is the basis for the determination of the membership of an object in a class. On this basis, we recommend methods of determining the membership degree on characteristics of fuzzy attributes, object/class, class/superclass, and in addition, multiple inheritance was discussed and analyzed.

**Keywords.** Fuzzy object-oriented database, hedge algebra.

### 1. INTRODUCTION

The relational database model and fuzzy object-oriented database (FOODB) model and related problems have been extensively researched in recent years by many domestic and foreign authors [1–15]. To perform fuzzy information in the data model, there are several basic approaches: the model based on similarity relation and the model based on possibility distribution, etc... All above approaches aim to achieve and process the fuzzy values to build valuation and comparison methods among them to manipulate data more flexibly and accurately.

Based on the advantages of the structure of hedge algebra (HA) [7, 8], the authors studied the relational database model [9–15] and fuzzy object-oriented database model [2–6] based on the approach of *HA*, in which linguistic semantics be quantified by quantitative semantic mapping of hedge algebra. In this approach, language semantics can be expressed in a neighborhood of intervals determined by the fuzziness measure of linguistic values of an attribute as a linguistic variable.

As well as fuzzy database model, in the fuzzy object oriented databases also needs to a data query language really flexible and the "precision" high. In order to do that, we need to focus on building the membership functions to determine the dependencies between the components in the model. Based on approach to hedge algebras to performing fuzzy values and measure the semantic approximation of two fuzzy data, in this paper, we present the method to determine the degree of the relationships in the fuzzy object oriented database model.

This paper is organized as follows: Section 2 presents some fundamental concepts related to hedge algebraic as the basis for the next section. Section 3 proposes the method of determining the degree of membership in the model fuzzy OODB, and section 4 concludes this paper.

## 2. FUNDAMENTAL CONCEPTS

This section presents a general overview of the complete linear hedge algebra proposed by Nguyen Cat Ho and *et al* [7,8], [2,3] and some related concepts on quantifying mapping and how to determine the quantitative semantic neighboring systems according to *HA* approach.

### 2.1. Hegde algebra

Consider a complete hedge algebra (Comp-HA)  $\mathcal{AX} = (X, G, H, \Sigma, \Phi, \leq)$ , where  $G$  is a set of generators which are designed as primary terms denoted by  $c^-$  and  $c^+$ , and specific constants  $0$ ,  $W$  and  $1$  (zero, neutral and unit elements, respectively),  $H = H^- \cup H^+$  and two artificial hedges  $\Sigma, \Phi$ , the meaning of which is, respectively, taking in the poset  $X$  the supremum (sup, for short) or infimum (inf, for short) of the set  $H(x)$  - the set generated from  $x$  by using operations in  $H$ . The word complete means that certain elements added to usual hedge algebras for the operations  $\Sigma$  and  $\Phi$  will be defined for all  $x \in X$ . Set  $textLim(X) = X \setminus H(G)$ , the set of the so-called limit elements of  $\mathcal{AX}$ .

**Definition 1.** A Comp-HAs  $\mathcal{AX} = (X, G, H, \Sigma, \Phi, \leq)$  is said to be a linear hedge algebra (Lin-HA, for short) if the sets  $G = \{\mathbf{1}, c^-, \mathbf{W}, c^+, \mathbf{0}\}$ ,  $H^+ = \{h_1, \dots, h_p\}$  and  $H^- = \{h_{-1}, \dots, h_{-q}\}$  are linearly ordered with  $h_1 < \dots < h_p$  and  $h_{-1} < \dots < h_{-q}$ , where  $p, q > 1$ . Note that  $H = H^- \cup H^+$ .

**Proposition 1.** *Fuzziness measures  $fm$  and fuzziness measures of  $\mu(h)$ ,  $\forall h \in H$ , has the following properties:*

- (1)  $fm(hx) = \mu(h)fm(x), \forall x \in \mathbf{X}$ .
- (2)  $fm(c^-) + fm(c^+) = 1$ .
- (3)  $\sum_{-q \leq i \leq p, i \neq 0} fm(h_i c) = fm(c)$ , where  $c \in \{c^-, c^+\}$ .
- (4)  $\sum_{-q \leq i \leq p, i \neq 0} fm(h_i x) = fm(x), x \in \mathbf{X}$ .
- (5)  $\sum \{\mu(h_i) : -q \leq i \leq -1\} = \alpha$  and  $\sum \{\mu(h_i) : 1 \leq i \leq p\} = \beta$ , where  $\alpha, \beta > 0$  and  $\alpha + \beta = 1$ .

In *HA*, each term  $x \in \mathbf{X}$  always has negative sign or positive sign, is called *PN-sign* and is defined recursively as below:

**Definition 2** (Sign function).  $Sgn: X \rightarrow \{-1, 0, 1\}$  is the signum function defined as follows, where  $h, h' \in H$ , and  $c \in \{c^-, c^+\}$ :

- (1)  $Sgn(c^-) = -1, Sgn(c^+) = +1$ .
- (2)  $Sgn(h'hx) = 0$ , if  $hhx = hx$ , otherwise  
 $Sgn(h'hx) = -Sgn(hx)$ , if  $hhx \neq hx$  and  $h'$  is negative with  $h$   
 $Sgn(h'hx) = +Sgn(hx)$ , if  $hhx \neq hx$  and  $h'$  is positive with  $h$ .

**Proposition 2.** *with  $\forall x \in \mathbf{X}$ , it yields:  $\forall h \in \mathbf{H}$ , if  $Sgn(hx) = +1$  then  $hx > x$ , if  $Sgn(hx) = -1$  then  $hx < x$  and if  $Sgn(hx) = 0$  then  $hx = x$ .*

From properties of fuzziness and sign function, semantically quantifying mapping of *HA* is defined as below.

**Definition 3.** Let  $\mathcal{AX} = (X, G, H, \Sigma, \Phi, \leq)$  is a complete, linear and free  $HA$ ,  $fm(x)$  and  $\mu(h)$  are the corresponding fuzziness measures of linguistic and the hedge  $h$  satisfying properties in Proposition 1. Then, we say that  $v$  is the mapping induced by fuzziness measure  $fm$  of the linguistic if it is determined as follows:

- (1)  $v(W) = fm(c^-)$ ,  $v(c^-) = W - \alpha fm(c^-) = \beta fm(c^-)$ ,  $v(c^+) = W + \alpha fm(c^+)$ .
  - (2)  $v(h_j x) = v(x) + Sgn(h_j x) \{ \sum_{i=Sgn(j)}^j \mu(h_i) fm(x) - \omega(h_j x) \mu(h_j) fm(x) \}$ , where
 
$$\omega(h_j x) = \frac{1}{2} [1 + Sgn(h_j x) Sgn(h_p h_j x) (\beta - \alpha)] \in \{\alpha, \beta\}$$
, for all  $j$ ,  $-q \leq j \leq p$  and  $j \neq 0$ .
  - (3)  $v(\Phi c^-) = 0$ ,  $v(\Sigma c^-) = k = v(\Phi c^+)$ ,  $v(\Sigma c^+) = 1$ , and for all  $j$ ,  $-q \leq j \leq p$  and  $j \neq 0$ .
- we have:  $v(\Phi h_j x) = v(x) + Sgn(h_j x) \{ \sum_{i=sign(j)}^{j-1} \mu(h_i) fm(x) \}$  and

$$v(\Sigma h_j x) = v(x) + Sgn(h_j x) \{ \sum_{i=sign(j)}^j \mu(h_i) fm(x) \}.$$

**Lemma 1.** Let  $fm$  be a fuzziness measure on  $X$ . For each  $v$  on  $X$  associated with  $fm$  defined as above, there exists an  $fm$ -decomposition system  $\mathfrak{S}$  associated with  $X$  such that the following statement holds for all  $x \in X$ :  $v(x) \in Im(x)$  and  $v(x)$  divides the interval  $\mathfrak{S}(x)$  into two subinterval in proportion  $\alpha$  to  $\beta$ . Moreover, if  $Sgn(h_j x) = +1$ , then the subintervals of the length  $\beta fm(x)$  is greater than the other one of the length  $\alpha fm(x)$ ; And if  $Sgn(h_j x) = -1$ , then the subinterval of length  $\beta fm(x)$  is less than the other one.

**Proposition 3.**

- (1) For all  $x \in X$ ,  $0 \leq v(x) \leq 1$
- (2) For all  $x, y \in X$ ,  $x < y$  implies  $v(x) < v(y)$

**Example 1.** Let  $\mathcal{AX} = (X, G, H, \leq)$ .

Where  $H^+ = \{\text{More, Very}\}$  with  $\text{More} < \text{Very}$  and  $H^- = \{\text{Little, Possibly}\}$  with  $\text{Little} > \text{Possibly}$ .  $C = \{\text{low, high}\}$  with  $\text{low}$  is negative term,  $\text{high}$  is positive term.

Assuming let  $W=0.5$ ,  $fm(\text{Little}) = 0.4$ ,  $fm(\text{Possibly}) = 0.1$ ,  $fm(\text{More}) = 0.1$ ,  $fm(\text{Very}) = 0.4$ .

Now, we give some examples of computing some values of the quantified semantic mapping  $v$ .

+ For  $x = c^- = \text{low}$ , we have  $v(\text{low}) = W - \alpha fm(\text{low}) = 0.5 - 0.5 \times 0.5 = 0.25$ .

+ For  $x = \text{Verylow}$ , we have  $j = p = 2$ ,  $Sgn(h_2 \text{low}) = -1$ ,  $Sgn(h_2 h_2 \text{low}) = -1$  and  $\omega(h_2 \text{low}) = \frac{1}{2} [1 + (-1)(-1)(\beta - \alpha)] = 0.5$  and  $v(\text{Verylow}) = v(\text{low}) + (-1) \{ fm(h_1 \text{low}) + fm(h_2 \text{low}) - 0.5 fm(h_2 \text{low}) \} = v(\text{low}) + (-1) \{ \mu(h_1) fm(\text{low}) + 0.5 \mu(h_2) fm(\text{low}) \} = 0.25 - \{ 0.1 \times 0.5 + 0.5 \times 0.4 \times 0.5 \} = 0.10$ .

+ For  $x = \text{LittleVerylow}$ , we have  $j = -q = -2$ ,  $Sgn(h_{-2} \text{Verylow}) = 1$ ,  $Sgn(h_2 h_{-2} \text{Verylow}) = 1$  and  $\omega(h_{-2} \text{Verylow}) = 0.5$ . Hence,  $v(h_{-2} \text{Verylow}) = v(\text{Verylow}) + (1) \{ fm(h_1 \text{Verylow}) + fm(h_{-2} \text{Verylow}) - 0.5 fm(h_{-2} \text{Verylow}) \} = v(\text{Verylow}) + \{ \mu(\text{Possibly}) \mu(\text{Very}) fm(\text{low}) + 0.5 \mu(\text{Little}) \mu(\text{Very}) fm(\text{low}) \} = 0.1 - \{ 0.10 \times 0.4 \times 0.5 + 0.5 \times 0.4 \times 0.4 \times 0.5 \} = 0.16$ .

The other values of the quantified semantic mapping  $v$  are computed in a similar way and the results are given in Table 1.

Linguistic value	function v	Linguistic value	function v
Very Very low	0.04	Very Very high	0.96
Very low	0.10	Very high	0.90
Possibly Very low	0.11	Possibly Very high	0.89
Little Very low	0.16	Little Very high	0.84
low	0.25	high	0.75
Very Possibly low	0.26	Very Possibly high	0.74
Little low	0.40	Little high	0.60
More Little low	0.41	More Little low	0.59
Very Little low	0.46	Very Little high	0.54

Table 1: Values function  $v$

**2.2. Fuzzy intervals of two fuzzy concepts**

Assuming that attribute A has a real reference domain of the interval  $[a, b]$  to standardize, by a linear transformation, we assume all such domain is the interval  $[0, 1]$ . Then, property (2) in Proposition 1 allows us to build two fuzzy intervals of two primitive concepts  $c^-$  and  $c^+$ , denoted by  $I(c^-)$  and  $I(c^+)$  with length respectively  $fm(c^-)$  and  $fm(c^+)$  such that they form a partition of the reference domain  $[0, 1]$ , and  $I(c^-)$  and  $I(c^+)$  are covariance with  $c^-$  and  $c^+$ , i.e.  $c^- \leq c^+$  implicate  $I(c^-) \leq I(c^+)$ .

The fuzzy interval is established basing on inductive method. Suppose  $\forall x \in X_{k-1} = \{x \in \mathbf{X} : x, \text{ fuzzy interval system } \{I(x) : x \in X_{k-1} \text{ and } |I(x)| = fm(x)\}$  has been built up so that they are *covariance* and form a *partition* of the interval  $[0, 1]$ . Then, on each fuzzy interval  $I(x)$  with length  $fm(x)$ , of  $x \in X_{k-1}$ , due to the properties (4) in proposition 2.1, we can make  $\{I(h_i x) : -q \leq i \leq p, i \neq 0, |I(h_i x)| = fm(h_i x)\}$  so that they are a partition of fuzzy interval  $I(x)$ . Its easy to see that  $\{I(h_i x) : -q \leq i \leq p, i \neq 0, |I(h_i x)| = fm(h_i x) \text{ and } x \in X_{k-1}\} = \{I(y) : y \in X_k \text{ and } |I(y)| = fm(y)\}$  are a partition of  $[0, 1]$ . Theses intervals are called *fuzzy intervals of depth k*.

**Definition 4.** Let  $P^k = \{I(x) : x \in X_k\}$  with  $X_k = \{x \in \mathbf{X} : |x| = k\}$  is a partition of  $[0, 1]$ . It is to say that  $u$  is equals  $v$  by k level  $k$  in  $P^k$ , denoted as  $u \approx_k v$ , if and only if  $I(u)$  and  $I(v)$  belong to the same interval  $P^k$ . That is  $\forall x, y \in \mathbf{X}, u \approx_k v \Leftrightarrow \exists \Delta^k \in P^k : I(u) \subseteq \Delta^k \text{ and } I(v) \subseteq \Delta^k$ .

**Lemma 2.** *Relation  $\approx_k$  is an equivalence relation*

*Proof.*

(1): *Reflexivity*

The proof is inductive method.

-  $\forall x \in \text{Dom}(A_i)$ , if  $|x| = 1$  then  $x = c^+$  or  $x = c^-$ .

It yields,  $\exists \Delta^1 = I(c^+) \in P^1 : I(c^+) = I(x) \subseteq \Delta^1$  or  $\exists \Delta^1 = I(c^-) \in P^1 : I(c^-) = I(x) \subseteq \Delta^1$ . So,  $\approx_k$  is true with  $k = 1$ , or  $x \approx_1 x$ .

- Assuming  $|x| = n$  is true, to mean  $\approx_k$  is true with  $k = n$ , or  $x \approx_n x$ , it is needed to prove  $\approx_k$  is true with  $k = n + 1$ . Let  $x = h_1 x'$ , with  $|x'| = n$ . Because,  $x \approx_n x$  by definition it yields:  $\exists \Delta^n \in P^n : I(x) \subseteq \Delta^n$ . On the other hands, it yields  $P^{n+1} = \{I(h_1 x'), I(h_2 x'), \dots\}$ , with  $h_1 \neq h_2 \neq \dots$  is a partition of  $I(x')$ . Hence,  $\exists \Delta^{(n+1)} = I(h_1 x') \in P^{(n+1)} : I(h_1 x') = I(x) \subseteq \Delta^{(n+1)}$ . So,  $\approx_k$  is true with  $k = n + 1$ , or  $x \approx_{(n+1)} x$ .

(2): *Symmetry*

$\forall x, y \in Dom(A_i)$ , if  $x \approx_k y$  then by definition  $\exists \Delta^k \in P^k : I(x) \subseteq \Delta^k$  and  $I(y) \subseteq \Delta^k$  or  $\exists \Delta^k \in P^k : I(y) \subseteq \Delta^k$  and  $I(x) \subseteq \Delta^k$ . So,  $y \approx_k x$  then  $x \approx_k y$ .

(3): *Transitivity*

The proof is by inductive method.

- Case  $k = 1$ : it results in  $P^1 = \{I(c^+), I(c^-)\}$ , if  $x \approx_1 y$  then  $y \approx_1 z$  then  $\exists \Delta^1 = I(c^+) \in P^1 : I(x) \subseteq \Delta^1$  and  $I(y) \subseteq \Delta^1$  and  $I(z) \subseteq \Delta^1$  or  $\exists \Delta^1 = I(c^-) \in P^1 : I(x) \subseteq \Delta^1$  and  $I(y) \subseteq \Delta^1$  and  $I(z) \subseteq \Delta^1$ , the mean is  $\exists \Delta^1 \in P^1 : I(x) \subseteq \Delta^1$  and  $I(z) \subseteq \Delta^1$  or  $x \approx_1 z$ . So,  $\approx_k$  is true with  $k = 1$ .

- Assuming relation  $\approx_k$  is true with case  $k = n$ , the mean, it yields  $\forall x, y, z \in Dom(A_i)$  if  $x \approx_n y$  and  $y \approx_n z$  then  $x \approx_n z$ .

- It is needed to prove relation  $\approx_k$  is true with case  $k = n + 1$ . The mean is  $\forall x, y, z \in Dom(A_i)$  if  $x \approx_{n+1} y$  and  $y \approx_{n+1} z$ . Assuming that if  $x \approx_{n+1} y$  and  $y \approx_{n+1} z$  then  $\exists \Delta^{(n+1)} = I(x) \subseteq \Delta^{(n+1)}$  and  $I(y) \subseteq \Delta^{(n+1)}$  and  $I(z) \subseteq \Delta^{(n+1)}$ , the mean is  $\exists \Delta^{(n+1)} \in P^{n+1} : I(x) \subseteq \Delta^{(n+1)}$  and  $I(z) \subseteq \Delta^{(n+1)}$ . So,  $x \approx_{n+1} z$ .  $\square$

### 3. FUZZY OBJECT-ORIENTED DATABASE MODEL

FOODB model is first proposed as a similarity-based data model. FOODB model focuses on the representation of ambiguous information, in other words, the fuzziness in FOODB is shown at three levels: attribute level, object/class level and class/superclass level. In addition, in this paper, we continue to expand study of issues related to the fuzzy object-oriented database model that we have presented in [2]

#### 3.1. Attribute level uncertainty

##### 3.1.1. Attribute uncertainty

FOODB deals with 3 types of uncertainty at the attribute level.

- (a) The first type is the type of incomplete occurs when the value of the attribute is determined to be a range of values. For example, the number of viewers watching a football match is about 10000-20000 people. This type uncertainty is called "incompleteness".
- (b) The second type of uncertainty occurs when the value of the attribute is unknown (unk), does not exist (dne) or there is no information on whether a value exists or not (ni). For instance, the description of a video which may be unknown (unk), the description of a video does not exist (dne) or we may not know whether a description for a video exists or not (ni). This type of uncertainty is called "Null".
- (c) The third type of uncertainty occurs when the value of the attribute is vaguely determined. This type of uncertainty is called "fuzzy". For example, weather condition in a football match can be described with a fuzzy term "very hot".

A similar relationship or the fuzzy equivalence relationship, is represented by a similarity matrix, the basis for similarity-based FOODB model. A similarity matrix defines the similarity of each pair of other elements in fuzzy domain. In next section, an method similar matrix construction based on semantic approximation of  $HA$  is presented.

### 3.1.2. Similarity matrix

**Definition 5.** To evaluate the semantic approximation of each pair of elements in fuzzy domain of a attribute, we construct function SP (Semantic Proximity) as follows:

$$SP(x, y) = 1 - |v(x) - v(y)|$$

where,  $v(x)$  and  $v(y)$  respectively is quantitative semantics value of the linguistic  $x$  and  $y$ .

The SP function have the following properties:

1.  $0 \leq SP(x, y) \leq 1$
2.  $SP(x, x) = 1$
3.  $SP(x, y) = SP(y, x)$

**Example 2.** Build quantitative semantics for attribute in the case where attribute values are linguistic values. Consider HA of linguistic variable *height*, where  $D_{height} = [0, 2]$ , generating elements  $\{0, \text{low}, \text{W}, \text{high}, 1\}$ , the set of hedges is  $\{\text{little, possibly, more, very}\}$  (correspond L, P, M, V),  $FD_{height} = H_{height}(\text{old}) \cup H_{height}(\text{young})$ . Choose  $\text{fm}(\text{low}) = 0.5$ ,  $\text{fm}(\text{high}) = 0.5$ ,  $\mu(P) = 0.3$ ,  $\mu(L) = 0.2$ ,  $\mu(M) = 0.3$  and  $\mu(V) = 0.2$ .

Based on the Definition 2, there are values of the quantitative values of linguistics term for attribute *height*, results as follows:  $v(\text{V low}) = 0.05$ ;  $v(\text{M low}) = 0.175$ ;  $v(\text{low}) = 0.25$ ;  $v(\text{P low}) = 0.425$ ;  $v(\text{L low}) = 0.3$ ;  $v(\text{L high}) = 0.55$ ;  $v(\text{P high}) = 0.675$ ;  $v(\text{high}) = 0.75$ ;  $v(\text{M high}) = 0.825$ ;  $v(\text{V high}) = 0.95$ .

Now, based on the Definition 5, we have the similarity relationship matrix between each pair of elements in the fuzzy value domain of the attribute *height* as follows:

+ For  $v(\text{V low}) = 0.05$ , and  $v(\text{V low}) = 0.05$ , then  $SP(v(\text{V low}), v(\text{V low})) = 1 - |v(\text{V low}) - v(\text{V low})| = 1 - |0.05 - 0.05| = 1$ .

+ For  $v(\text{V low}) = 0.05$ , and  $v(\text{V high}) = 0.95$ , then  $SP(v(\text{V low}), v(\text{V high})) = 1 - |v(\text{V low}) - v(\text{V high})| = 1 - |0.05 - 0.95| = 0.2$ .

The other values of the semantic approximation  $SP$  are computed in a similar way and the results are given in Table 2.

height	V low	M low	low	P low	L low	L high	P high	high	M high	V high
V low	1	0.875	0.8	0.625	0.75	0.5	0.375	0.3	0.225	0.1
M low	0.875	1	0.925	0.75	0.875	0.625	0.5	0.425	0.35	0.225
low	0.8	0.925	1	0.825	0.95	0.7	0.575	0.5	0.425	0.3
P low	0.625	0.75	0.825	1	0.875	0.875	0.75	0.675	0.6	0.475
L high	0.5	0.625	0.7	0.875	0.75	1	0.875	0.8	0.725	0.6
P high	0.375	0.5	0.575	0.75	0.625	0.875	1	0.925	0.85	0.725
high	0.3	0.425	0.5	0.675	0.55	0.8	0.925	1	0.925	0.8
M high	0.225	0.35	0.425	0.6	0.475	0.725	0.85	0.925	1	0.875
V high	0.1	0.225	0.3	0.475	0.35	0.6	0.725	0.8	0.875	1

Table 2: Similarity matrix for attribute *height*

In the Fuzzy Object-Oriented model, fuzzy attribute is a multi-valued attribute, and so it can get a set of values and these values are linked by one of the semantic components AND, OR, or XOR. The attributes can have a set of values (leading to multi-valued attributes) connected with a logical operator AND/OR/XOR. The attribute value sets are differentiated according to their semantics. The following representations are used for multi-valued attributes.

Logical operator	Indicate
AND	<...>
OR	{...}
XOR	[...]

Assuming the domain of weather attribute has the value domain = {cloudy, sunny, cool, cold, hot}. The following interpretations are valid:

AND: The value of the attribute weather is <cloudy, cold>, i.e. the weather is cloudy and cold.

OR: The value of the attribute weather is {cool, hot}, i.e. the weather is cool or hot, or maybe both.

XOR: The value of the attribute weather is [sunny, cold], i.e. the weather is sunny or cold, but not both. This different interpretation from the logical XOR operator, which returns true for an odd number of true's, has to be noted.

### 3.2. Object/Class Level Uncertainty

The level of uncertainty at the object/class refers to the existence of a part of its members object to the class. In FOODB model, the boundaries of a class may be uncertain because it has fuzzy attributes. The range of a fuzzy attribute shows the ideal value for that attribute.

Since a fuzzy attribute may take any value from its value domain regardless of its range definition, some objects are full members of their classes with a membership degree of 1 whereas some other objects are member of their classes with a membership degree changing between 0 and 1. The values of fuzzy attributes of an object determine the membership degree of that object to its class. The closer the value of fuzzy attributes of an object to range definitions is, the higher the object membership degree is. The relevance of the fuzzy attributes and the similarity between the fuzzy attributes values and define their range determine the membership degree of an object with class.

Based on the considerations of relevance and inclusion of attribute values, to find the membership degree of object  $o_j$  in class C (class C have n attribute) is defined as:

$$\mu_C(o_j) = \frac{\sum_{i=1}^n INC(rng_C(a_i)/o_j(a_i)) * RLV(a_i, C)}{\sum_{i=1}^n RLV(a_i, C)}$$

Where:

- $INC(rng_C(a_i)/o_j(a_i))$  denotes the degree of inclusion of the attribute values of  $o_j$  in the formal range  $a_i$  in the class C. The degree of inclusion, determines the extent of similarity between a value (or a set of values) in the denominator with the value in the numerator (or a set of values).
- $RLV(a_i, C)$  is the relevance of the attribute  $a_i$  to the class C. This membership degree show the importance of attribute  $a_i$  to the class C. This membership degree is determined based on the concept of attribute affinity proposed by Hoffer and Severance and cosin measure.

In addition, the weighted average is used to calculate the membership degree of objects. All attributes, therefore, affect the membership degree proportionally to their relevance.

INC has two following cases:

- Case 1: If  $o_j(a_i) = \emptyset$  then  $INC = 0$ .
- Case 2: INC will consider for each case the semantics of attribute values.

Now, consider and calculate the degree of inclusion for the different semantics are explained below:

AND semantics

AND semantics is strong because it requires all values to exist simultaneously. AND semantics arise when data is nested, the INC formulation for AND semantics as follows

$$INC(rng_C(a_i)/o_j(a_i)) = Min[Min[Max(SP(x, y))], Min[Max(SP(z, w))], \\ \forall x \in rng_C(a_i), \forall y \in o_j(a_i), \forall z \in o_j(a_i), \forall w \in rng_C(a_i).$$

Note that the basis of comparison is the range definition, comparing all of the elements in the object attribute to each of the entries in the range definition. Then the order is reversed (i.e.  $x \in rng_C(a_i)$ ). That is, the attribute definition is the reference point.

**Example 3.** Consider Employee class is definition as follows:

class Employee

```
{   FuzzyInteger age;
    FuzzyString coloreye;
}
```

Given the relevance of the attribute, define the range values of the attributes *age* and value of  $o_1, o_2$  for the attribute *age* as follows:

RLV(age, employee)=0.8, RLV(coloreye, employee)=0.2.

$rng_{employee}(age) = \langle \text{old, very old} \rangle$ .

$o_1(\text{age}) = \langle \text{young, old} \rangle$ .

$o_2(\text{age}) = \langle \text{old} \rangle$ .

The value degree of inclusion for attribute is calculated as follows:

$INC(rng_{employee}(age)/o_1(age)) = Min[Min[Max(SP(\text{old, young}), SP(\text{old, old}))], Max(SP(\text{very old, young}), SP(\text{very old, old}))], Min[Max(SP(\text{young, old}), SP(\text{young, very old}))], Max(SP(\text{old, old}), SP(\text{old, very old}))]] = Min[Min[Max(0.5, 1), Max(0.35, 0.85)], Min[Max(0.5, 0.35), Max(1, 0.85)]] = 0.5.$

$INC(rng_{employee}(age)/o_2(age)) = Min[Min[Max(SP(\text{old, old}), Max(SP(\text{very old, old}))), Min[Max(SP(\text{old, old}), Max(SP(\text{old, very old}))]] = Min[Min[Max(1), Max(0.85)], Min[Max(1), Max(0.85)]] = 0.85.$

OR semantics

Under OR semantics an attribute takes more than one value, all or some of which may exist simultaneously (are true). Reformulation is required since the model treats both the connectives the same way ignoring the semantic difference between them. So here we find the value of to be

$$INC(rng_C(a_i)/o_j(a_i)) = Min[Max(SP(x, z)), Threshold(o_j(a_i))], \\ \forall x \in o_j(a_i), \forall z \in rng_C(a_i)$$

Here, the threshold value indicates that the minimum level of similarity between the values of an attribute and is built as follows

$$Threshold(o_j(a_i)) = Min[SP(x, z)], \forall x, \forall z \in o_j(a_i)$$



**Example 4.** Given class Employee and similarity matrix above, and define the values range of the attribute *age* and value of the two objects  $o_1, o_2$  for the attribute *age* as follows:

$$rng_{employee}(age) = \{\text{old, very old}\}.$$

$$o_1(age) = \{\text{young, old}\}.$$

$$o_2(age) = \{\text{old}\}.$$

$$Threshold(o_1(age)) = \text{Min}[\text{SP}(\text{young, young}), \text{SP}(\text{old, young}), \text{SP}(\text{old, old})] = \text{Min}[1, 0.5, 1] = 0.5.$$

$$Threshold(o_2(age)) = 1.$$

$$INC(rng_{employee}(age)/o_1(age)) = \text{Min}[\text{Max}(\text{SP}(\text{young, old}), \text{SP}(\text{young, very old}), \text{SP}(\text{old, old}), \text{SP}(\text{old, very old}), \text{Threshold}(o_1(age)))] = \text{Min}[\text{Max}(0.5, 0.35, 1, 0.85), 0.5] = 0.5.$$

$$INC(rng_{employee}(age)/o_2(age)) = \text{Min}[\text{Max}(\text{SP}(\text{old, old}), \text{SP}(\text{old, very old}), \text{Threshold}(o_2(age)))] = \text{Min}[\text{Max}(1, 0.85), 1] = 1.$$

XOR semantics

For XOR semantics, only one of the attribute values exists at a time. Assume the equal probabilities for the values of an attribute. The INC formulation for XOR semantics as follows

$$INC(rng_C(a_i)/o_j(a_i)) = \text{Avg}[\text{Max}(\text{SP}(x, y))], \forall x \in o_j(a_i), \forall y \in rng_C(a_i)$$

**Example 5.** Given class Employee and similarity matrix above, define the range values of the attributes *age* and value of  $o_1, o_2$  for the attribute *age* as follows:

$$rng_{employee}(age) = [\text{old, very old}].$$

$$o_1(age) = [\text{young, old}].$$

$$o_2(age) = [\text{old}].$$

$$INC(rng_{employee}(age)/o_1(age)) = \text{Avg}[\text{Max}(\text{SP}(\text{young, old}), \text{SP}(\text{young, very old}), \text{Max}(\text{SP}(\text{old, old}), \text{SP}(\text{old, very old})))] = \text{Avg}[\text{Max}(0.5, 0.35), \text{Max}(1, 0.85)] = 0.5.$$

$$INC(rng_{employee}(age)/o_2(age)) = 1.$$

### 3.3. Class/Subclass Level Uncertainty

The level of uncertainty of the class/superclass refers to the existence of a member of a class part of its superclass. This type shows fuzzy uncertainty occurs in the hierarchy of inheritance from a class hierarchy can not be built correctly in some cases. The membership degree of a class C to its superclass  $C_j$  is calculated using the following formula:

$$\mu_{C_j}(C) = \frac{\sum_{i=1}^n INC(rng_{C_j}(a_i)/rng_C(a_i)) * RLV(a_i, C_j)}{\sum_{i=1}^n RLV(a_i, C_j)}$$

INC has two following cases:

- Case 1: If  $rng_C(a_i) = \emptyset$  then  $INC = 0$ .

- Case 2:  $INC = \text{Min}(\text{Max}[\text{SP}(x, y)]), \forall x \in rng_{C_j}(a_i), \forall y \in rng_C(a_i)$ .

**Example 6.** Consider class hierarchies as follows:

Build quantitative semantics for attribute in the case where attribute values are linguistic values. Consider HA of linguistic variable *salary*, where  $D_{salary} = [0, 50]$ , generating elements are  $\{0, \text{low}, \text{W}, \text{high}, 1\}$ , the set of hedge is  $\{\text{little, possibly, more, very}\}$  (correspond L, P, M, V),  $FD_{salary} = H_{salary}(\text{high}) \cup H_{salary}(\text{low})$ . Chosse  $\text{fm}(\text{high}) = 0.5, \text{fm}(\text{low}) = 0.5, \mu(\text{P}) = 0.3, \mu(\text{L}) = 0.2, \mu(\text{M}) = 0.3$  and  $\mu(\text{V}) = 0.2$ .

Based on the Definition 2, we can calculate quantitative values of linguistic term for attribute *salary*, results as follows:  $v(\text{V low}) = 0.005; v(\text{M low}) = 0.175; v(\text{low}) = 0.25; v(\text{P low}) = 0.325;$

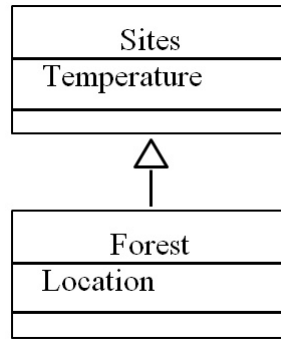


Figure 1: Inherited relationship model

$v(L \text{ low}) = 0.45$ ;  $v(L \text{ high}) = 0.55$ ;  $v(P \text{ high}) = 0.675$ ;  $v(\text{high}) = 0.75$ ;  $v(M \text{ high}) = 0.825$ ;  $v(V \text{ high}) = 0.95$ .

Based on the Definition 5, there is the similarity matrix for attribute *salary* as follows

salary	V low	M low	low	P low	L low	L high	P high	high	M high	V high
V low	1	0.875	0.8	0.725	0.6	0.5	0.375	0.3	0.225	0.1
M low	0.875	1	0.925	0.85	0.725	0.625	0.5	0.425	0.35	0.225
low	0.8	0.925	1	0.925	0.8	0.7	0.575	0.5	0.425	0.3
P low	0.725	0.85	0.925	1	0.875	0.775	0.65	0.575	0.5	0.375
L low	0.6	0.725	0.8	0.875	1	0.9	0.775	0.7	0.625	0.5
L high	0.5	0.625	0.7	0.775	0.9	1	0.875	0.8	0.725	0.6
P high	0.375	0.5	0.575	0.65	0.775	0.875	1	0.925	0.85	0.725
high	0.3	0.425	0.5	0.575	0.7	0.8	0.925	1	0.925	0.8
M high	0.225	0.35	0.425	0.5	0.625	0.725	0.85	0.925	1	0.875
V high	0.1	0.225	0.3	0.375	0.5	0.6	0.725	0.8	0.875	1

Table 3: Similarity matrix for attribute *salary*

We have the value range and membership degree for attributes as follows:

$$rng_{person}(salary) = \{\text{little high, high}\}.$$

$$RLV(salary, person) = 0.5.$$

$$rng_{student}(salary) = \{\text{possibly low, little high}\}.$$

We calculate the degree of inclusion for attributes as follows:

$$INC(rng_{person}(salary)/rng_{student}(salary)) = \text{Min}[\text{Max}(1, 0.8), \text{Max}(0.9, 0.575)] = 0.9.$$

Applying formula calculating membership degree of class student to the class person, we have:

$$\mu_{person}(student) = [INC(rng_{person}(salary)/rng_{student}(salary)) * RLV(salary, person)] / [RLV(salary, person)] = (0.9*0.5)/0.5=0.9.$$

In some applications, we may be faced with a problem of multiple inheritance when multiple superclasses for an object with different values for a field. This problem is not specific to the model FOODB, but in all manifestations decentralized system. In the document, it was pointed out that there is no general solution to fully resolve all cases of multiple inheritance conflict.

One approach to the problem of conflict in multiple inheritance might be to define a default method in the most general superclass are inherited by other classes. The subclass then override

the default class to handle the conflict between the values possible. An alternative approach may inherit values from the layer in which the object has the largest number of members. However, other solutions may be used in combination of two weighted value of the inherited.

The conflicts arise when more than one super class and subclass common attributes not explicitly declare the class where the properties are inherited (see Figure 2).

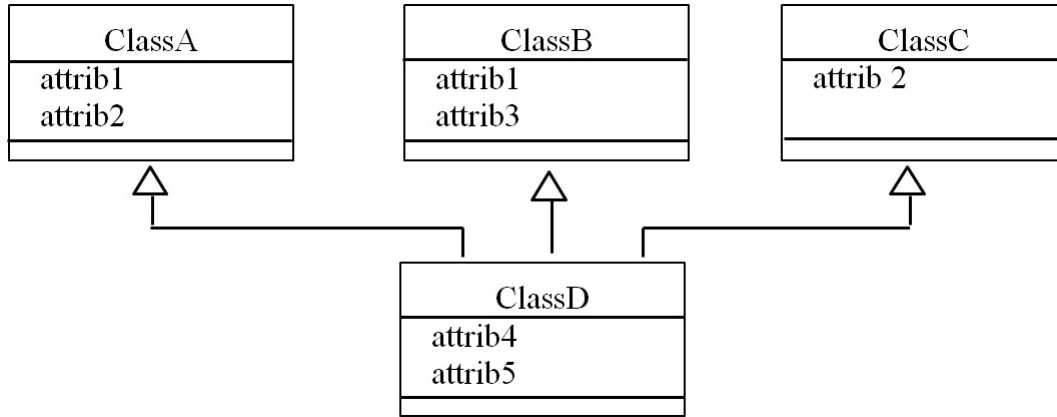


Figure 2: multiple inheritance model

The attributes of ClassD is attrib1, attrib2, attrib3, attrib4, and attrib5, Attrib4, and attrib5 defined in class and others inherited from the superclass. The contradictions exist for attrib1 and attrib2. We suggest that there is no need to consider the super class from which a attribute is not inherited, if it does not have such attributes in the class/superclass calculating membership degree.

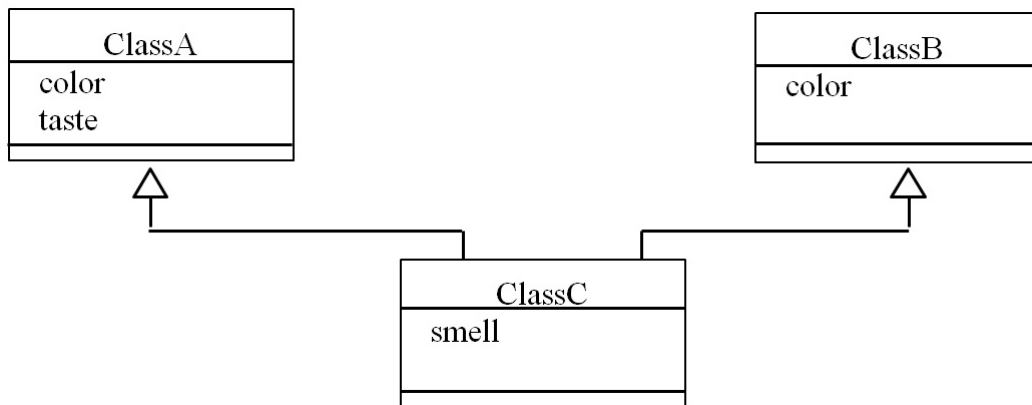


Figure 3: multiple inheritance of ParTimeStudent class

With this review may reflect the change from the defined range of the superclass. In principle, each class was forced to declare its range to all its translucent attributes. This ensures a defined scope for a property published contradictory as attrib1 and attrib2 available in the class definition. The rest is handled by the new formula for calculating the membership degree of the ClassD to ClassA, ClassB and ClassC. For example, consider multiple inheritance of ParTimeStudent class in Figure 3. All of the superclasses have the same attribute salary. The range definitions

are shown, with  $rng_{salary}(employee) = \langle \text{high} \rangle$ ,  $rng_{salary}(student) = \langle \text{little high, little low} \rangle$  and  $rng_{salary}(parttimestudent) = \langle \text{low, little low} \rangle$ .

The salary attribute of the parttimestudent class appears to be inherited from employee class. When applying the INC formula on each superclass to determine the class/subclass membership, this fact is showed in the results. The values included in the range of parttimestudent salary in employee salary will result in a value INC is one, as expected. However the results for the student class are much less

$INC = \text{Min}(\text{Max}[\text{SP}(\text{little high, low}), \text{SP}(\text{little high, little low})], \text{Max}(\text{SP}(\text{little low, low}), \text{SP}(\text{little low, little low})))$

shows the bond between employee and parttimestudent is more powerful than student and parttimestudent for attribute salary. Of course, one important point is that the relevant rules defined in the class. If the salary values of the RVL for employee or student is high, the effect of any deviation will be higher.

#### 4. CONCLUSION

In the context of modeling fuzzy object-oriented database, to have a data query language, we first need to build the membership functions for the relationships in the FOOD model. Based on approach to hedge algebras to represent the fuzzy values and semantic approximation measure for fuzzy data. In this paper, the authors focus on building some of the most important member functions in the fuzzy object-oriented database model, that is membership functions for attribute, object/class, class/superclass relationships. However, for multiple inheritance relationship, we have not mentioned in this paper. On the basis of the definitions of member functions of this, the methods the fuzzy object query and the membership functions for multiple inheritances will be presented in next studies.

#### REFERENCES

- [1] V. Biazzo, R. Giugno, T. Lukasiewicz, and V. Subrahmanian, "Temporal probabilistic object bases," *Knowledge and Data Engineering, IEEE Transactions on*, vol. 15, no. 4, pp. 921–939, 2003.
- [2] D. V. Thang and D. C. Quoc, "Defining membership functions in fuzzy object-oriented database model," in *Future Data and Security Engineering (FDSE), The 2nd International Conference on*. LNCS - Springer Verlag, 2015, pp. 314–322.
- [3] D. V. Ban, T. C. Tuan, and D. V. Thang, "Querying data with fuzzy information in object-oriented databases based on hedge algebraic semantic," in *Knowledge and Systems Engineering (KSE), 2012 Fourth International Conference on*. IEEE, 2012, pp. 39–45.
- [4] D. V. Thang, "Dependence fuzzy objects," in *Advanced Technologies for Communications (ATC), 2014 International Conference on*. IEEE, 2014, pp. 160–166.
- [5] V. B. Doan, C. T. Truong, and V. T. Doan, "Approximate answers to null queries with incomplete fuzzy information in object oriented database," *Journal of Computer Science and Cybernetics*, vol. 27, no. 4, pp. 329–340, 2012.
- [6] D. V. Ban, T. C. Tuan, and D. V. Thang, "Fuzzy dependencies in approximation semantic based on the hedge algebras approach." *Journal of Computer Science and Cybernetics*, vol. 29, no. 1, 2013, (in Vietnamese). [Online]. Available: <http://vjs.ac.vn/index.php/jcc/article/view/2885>

- [7] N. Ho, “Quantifying hedge algebras and interpolation methods in approximate reasoning,” in *Proc. of the 5th Inter. Conf. on Fuzzy Information Processing, Beijing*, 2003, pp. 105–112.
- [8] C. H. Nguyen and W. Wechler, “Hedge algebras: an algebraic approach to structure of sets of linguistic truth values,” *Fuzzy sets and systems*, vol. 35, no. 3, pp. 281–293, 1990.
- [9] C. H. Nguyen, “A topological completion of refined hedge algebras and a model of fuzziness of linguistic terms and hedges,” *Fuzzy Sets and Systems*, vol. 158, no. 4, pp. 436–451, 2007.
- [10] C. H. Nguyen and V. L. Nguyen, “Fuzziness measure on complete hedge algebras and quantifying semantics of terms in linear hedge algebras,” *Fuzzy Sets and Systems*, vol. 158, no. 4, pp. 452–471, 2007.
- [11] C. H. Nguyen, D. K. Tran, H. Van Nam, and H. C. Nguyen, “Hedge algebras, linguistic-value logic and their application to fuzzy reasoning,” *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, vol. 7, no. 04, pp. 347–361, 1999.
- [12] C. H. Nguyen, T. S. Tran, and D. P. Pham, “Modeling of a semantics core of linguistic terms based on an extension of hedge algebra semantics and its application,” *Knowledge-Based Systems*, vol. 67, pp. 244–262, 2014.
- [13] C. H. Nguyen, W. Pedrycz, T. L. Duong, and T. S. Tran, “A genetic design of linguistic terms for fuzzy rule based classifiers,” *International Journal of Approximate Reasoning*, vol. 54, no. 1, pp. 1–21, 2013.
- [14] N. C. Ho and N. V. Long, “Fuzziness measure on complete hedge algebras and quantifying semantics of terms in linear hedge algebras,” *Fuzzy Sets and Systems*, vol. 158, no. 4, pp. 452–471, 2007.
- [15] A. Yazici, R. George, and D. Aksoy, “Design and implementation issues in the fuzzy object-oriented data model,” *Information Sciences*, vol. 108, no. 1, pp. 241–260, 1998.

*Received on May 19 - 2015*  
*Revised on November 10 - 2015*