Analysis of Linear Interpolation of Fuzzy Sets with Entropy-based Distances

László Kovács[1] and Joel Ratsaby[2]

[1] Department of Information Technology, University of Miskolc, 3515 Miskolc-Egyetemváros, Hungary, kovacs@iit.uni-miskolc.hu

[2] Department of Electrical and Electronics Engineering, Ariel University Center of Samaria, Ariel 40700, Israel, ratsaby@ariel.ac.il

Abstract: An interpolation of fuzzy sets is an important method in development of efficient fuzzy rule systems. An important property of the interpolated set is the distance minimum property. As can be seen, the validity of this property depends on the applied distance metric. The authors analyse the distance relationship among the base and generated fuzzy sets in the case of KH linear interpolation. The paper presents new properties among the entropy-based distances and proposes an appropriate method for distance optimum interpolation.

Keywords: fuzzy interpolation; descriptive complexity; entropy; distance metric

1   Introduction

Interpolation is a widely used method to determine the values of a target function $f()$ at a position x in a real interval $[a,b]$, where $f(a)$ and $f(b)$ are given but $f(x)$ is not known. In a more general approach, the method can be extended for an arbitrary domain D with $a_1, a_2, \dots, a_n, x \in D$ to determine $f(x)$ from $f(a_1), \dots, f(a_n)$. Our investigation focuses on set D of fuzzy sets. The notion of a fuzzy set was introduced by [4]. It is a class of objects with continuous values of membership and hence extends the classical definition of a set (to distinguish it from a fuzzy set we refer to it as a crisp set). Formally, a fuzzy set is a pair *(E, m)* where E is a set of objects and *m* is a membership function $m : E \rightarrow [0, 1]$. Fuzzy set theory can be used in a wide range of domains in which information is incomplete or imprecise, such as pattern recognition and decision theory [2] [3].

In the area of fuzzy rule interpolation (FRI) [7], the goal is to generate new fuzzy rules from existing rules. An important component of FRI is the generation of antecedent and consequent fuzzy sets using a Fuzzy Set Interpolation (FSI) method. In the most widely used approaches, $f(x)$ is generated as a weighted sum

of $f(a_i)$ where the weight value depends on the distance between x and a_i: In the case of linear interpolation, the sum of weights is equal to 1:

$$f(x) = \sum_{i=1}^{n} w_i f(a_i), \quad \sum_{i=1}^{n} w_i = 1.$$

The KH method developed by Kóczy and Hirota [8] uses linear interpolation as a standard FSI method. The position of the generated fuzzy set B^* is calculated with the formula

$$B^*_{\alpha} = \frac{\sum_{i=1}^{n} \dfrac{1}{d(A_{\alpha}^{*}, A_{i,\alpha})} B_{i,\alpha}}{\sum_{i=1}^{n} \dfrac{1}{d(A_{\alpha}^{*}, A_{i,\alpha})}},$$

where A denotes the antecedent set and B is the consequence set. The symbol α denotes a α-cut which is defined as $H_{\alpha} = \{x \in E \mid m_H(x) \geq \alpha\}$ for any H fuzzy set with membership function $m_H()$. In addition to the KH method, several new approaches are available in the literature. In the modified α-cut based interpolation (MACI) [11], fuzzy sets are described with two vectors containing the left (lower) and right (upper) flanks. The improved version of MACI is called the multidimensional modified *α*-cut based interpolation [9], and it extends MACI with the fuzziness conservation technique proposed by [10]. A more detailed survey of FRI methods can be found in [7] [12], among others.

In all versions, the distance value [1] has a central role in the interpolation algorithm. A semi-metric function to measure the distance $d : D \times D \rightarrow \Re$ meets the following conditions:

$$\begin{aligned}
&d(x, y) \geq 0 \\
&d(x, x) = 0 \\
&d(x, y) = d(y, x) \\
&d(x, z) + d(y, z) \geq d(x, y)
\end{aligned} \quad . \tag{1}$$

For the Euclidean space, the most widely used metric is the Minkowski distance between two points x and y in $\Re^n$, which is defined as

$$d_r(x, y) = \left(\sum_{i=1}^{n} |x_i - y_i|^r \right)^{1/r}, \ r \geq 1. \tag{2}$$

For sets in Euclidean space there are several variants for the metric function. The Hausdorff distance $q()$ is defined as

$$q(U,V) = \max \left\{ \sup_{v \in V} \inf_{u \in U} d_2(u,v), \sup_{u \in U} \inf_{v \in V} d_2(u,v) \right\}. \tag{3}$$

This can be extended to fuzzy sets as follows. Let E be a finite set and let $\Phi(E)$ be the set of all fuzzy subsets of E. Then, for two fuzzy subsets $A, B \in \Phi(E)$, the distance in (3) can be extended to the following distance between A and B,

$$q(A,B) = \int_0^1 q(A_\alpha, B_\alpha) d\alpha.$$

A different approach is the Hamming distance for fuzzy sets. Consider two fuzzy subsets $A, B \in \Phi(E)$ with membership functions $m_A, m_B : E \rightarrow [0, 1]$. Then (2) can be extended to the following Hamming distance,

$$d_r(A,B) = \left(\sum_{x \in E} |m_A(x) - m_B(x)|^r \right)^{1/r}, \ r \geq 1. \tag{4}$$

The Euclidean distance has the following nice property: consider two elements A, B in the space, then for every element C that satisfies

$$C = \lambda \cdot A + (1-\lambda) \cdot B, \ \lambda \in [0,1]$$

the following equality holds

$$d(A,C) + d(B,C) - d(A,C) = 0, \tag{5}$$

i.e., the points of the connecting line are extreme points from the viewpoint of distance relationship. This nice property will not in general be met for other distances.

The goal of our investigation is to analyze the relationship between the linear interpolation of fuzzy sets and the distance function in the case of a specific metric, the entropy-based distance function. The analysis shows that the fuzzy sets generated by linear interpolation will not meet (5), and a different generation method should be used to fulfill this extreme condition.

In Section 2, three basic entropy-based distance definitions for fuzzy sets are presented. The first approach corresponds to a global entropy difference, the second method is based on an element-wise entropy difference and the third approach uses a descriptive complexity with symmetric difference of the corresponding membership functions. In Section 3, the property of distance optimality is investigated in KH interpolation for the different distance interpretations. It will be shown that the KH interpolation algorithm is not suitable to generate a fuzzy set lying on the distance optimum middle point between the operand fuzzy sets. To prove the existence of such an optimum fuzzy set, a generation algorithm is also presented in the section. The theoretical considerations are demonstrated with numerical examples in the paper.

2   Entropy-based Distances

Different application areas require different similarity and distance interpretations. In the case of fuzzy sets, there are basically three main aspects of similarity [5]:

- similarity of the support set in E (Hausdorff metric)*;*
- similarity of the values of membership functions (Hamming metric)
- similarity of the fuzziness of membership functions

In the latter, we assume a continuous E domain. The fuzziness of $A \in \Phi(E)$ is defined by De Luca and Termini [6] as

$$entropy(A) = \int_{-\infty}^{\infty} S(m_A(x))dx$$

where

$$S(x) = -x\lg(x) - (1-x)\lg(1-x).$$

One approach to include the fuzziness into the distance calculation is given by the following formula:

$$d_{S1}(A,B) = \sqrt{(entropy(A) - entropy(B))^2} \ . \tag{6}$$

As the *entropy()* function maps the fuzzy sets into $\Re^+$, $d_{s1}()$ meets the requirements of a metric function. Another way is to define an element-wise difference as

$$d_{S2}(A,B) = \left(\int_{-\infty}^{\infty} \left| entropy_A(\{x\}) - entropy_B(\{x\}) \right|^2 dx \right)^{1/2} \tag{7}$$

where

$$entropy_A(S\{x\}) = -A(x)\lg(A(x)) - (1 - A(x))\lg(1 - A(x)).$$

This approach maps the fuzzy sets into a multi-dimensional vector space, where the applied Euclidean distance is a metric; $d_{s2}()$ meets also here the requirements of a metric function.

The third approach uses the distance function that is based on a descriptive-complexity [6]. This distance uses the symmetric difference of the corresponding membership functions and is based on the following considerations. Given two fuzzy subsets $A,B \in \Phi([N])$ with membership functions $m_A(x)$, $m_B(x)$, we denote by

$$m_{A \cap B}(x) = \min\{m_A(x), m_B(x)\}$$

and

$$m_{A \cup B}(x) = \max\{m_A(x), m_B(x)\}.$$

Define by $A \Delta B = (A \cup B) \setminus (A \cap B)$ the symmetric difference between crisp sets A, B. For fuzzy sets $A, B \in \Phi([N])$ define by

$$m_{A\Delta B}(x) = m_{A \cup B}(x) - m_{A \cap B}(x).$$

Define a sequence of Bernoulli random variables $X_A(x)$ for $x \in [N]$ taking the value 1 with respect to $m_A(x)$ and the value 0 with respect to $1 - m_A(x)$. Define by $H(X_A(x))$ the entropy of $X_A(x)$,

$$H(X_A(x)) = -m_A(x)\log m_A(x) - (1 - m_A(x))\log(1 - m_A(x)).$$

Define the random variable

$$X_{A\Delta B}(x)) = \begin{cases} 1 & w.p. \quad m_{A\Delta B}(x) \\ 0 \, w.p. & (1 - m_{A\Delta B}(x) \end{cases}.$$

We define a new distance between $A, B \in \Phi([N])$ as

$$d_{S3}(A,B) = \frac{1}{N}\sum_{x=1}^{N} H(X_{A\Delta B}(x))$$

for discrete domain and

$$d_{S3}(A,B) = \int_{-\infty}^{\infty} H(X_{A\Delta B}(x))dx \tag{8}$$

for continuous domain.

In [6] we proved that the function $d_{S3}(A, B)$ is a semi-metric on $\Phi([N])$; i.e., it is non-negative, symmetric, equals zero if $A = B$, and satisfies the triangle inequality.

Note that for any $x \in [N]$ with a crisp membership value, i.e., $m_A(x)=1$, or $m_A(x)=0$, we have $m_{A\Delta\bar{A}}(x)=1$, and hence in this case $H(X_{A\Delta\bar{A}}(x))=0$. This means that for a crisp set A (for all $x \in A$, $m_A(x) \in \{0,1\}$) our distance has the following property (we call this the complement-property)

$$dist(A, \bar{A}) = 0.$$

From an information theoretic perspective, this property is expected since knowing a set A automatically means that we also know how to describe its complement. Hence, there is no additional description necessary to describe A given its complement. This is what $dist(A, \bar{A}) = 0$ means. It can be seen from the definition that the function $dist(A, B)$ may equal zero even when $A \neq B$.

As an example, consider the fuzzy sets A, B, C and the complement A' with membership functions as shown in Figure 1. Note that A and its complement are crisp sets. The distance matrix $D = [d_{i,j}]$ is shown below; the rows and columns correspond to A, B, C and A' so that for instance the element $d_{2,3} = d_{S3}(B, C) = 0.709$.

$$D = \begin{bmatrix} 0 & 0.354 & 0.354 & 0 \\ 0.354 & 0 & 0.709 & 0.354 \\ 0.354 & 0.709 & 0 & 0.354 \\ 0 & 0.354 & 0.354 & 0 \end{bmatrix}$$

Distance matrix D

As can be seen, *C* is a translated version of *B* and they are both the same distance from *A*. This is due to $H(X_{A\Delta B}(x)) = H(X_{A\Delta C}(x+10))$. *B* and *C* are farther apart than *B* and *A*. Since $d_{S3}(A, A') = 0$ then each one of *B*, *C* is of the same distance to *A* as to *A'*
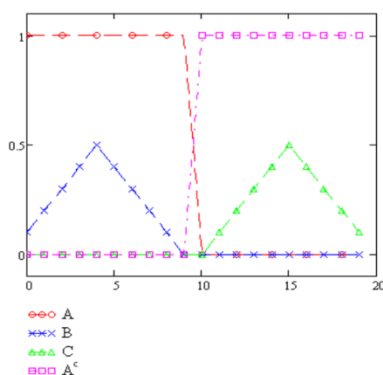

Figure 1 [6]

Fuzzy sets A,B,C and A^c

3   Distance-Optimal Interpolation Algorithm

According to (5), a linear interpolation with Euclidean metric generates elements with optimal distance. In this paper we obtained experimental results using the KH method, which was used to generate the intermediate fuzzy set *C* for given *A,B* $\in$ $\Phi([N])$. In these tests, the λ value runs from 0 to 1. The test results are shown in Figure 2. In the Figure, the x-axis shows the value of λ; on the y-axis the value *ddiff(A,B,C) = d(A,C) + d(B,C) - d(A,B)* is given. The top (red) line is the descriptive complexity distance ($d_{S3}()$), the middle (blue) line is the element-wise entropy distance ($d_{S2}()$) and the bottom (green) line refers to the entropy-difference distance ($d_{S1}()$).

The *ddiff(A,B,C)* value indicates whether the generated *C* element is the closest element to both *A* and *B*. If *ddiff(A,B,C)* is equal to zero, the triangle inequality yields an equality and *C* lies on the line connecting *A* to *B*.
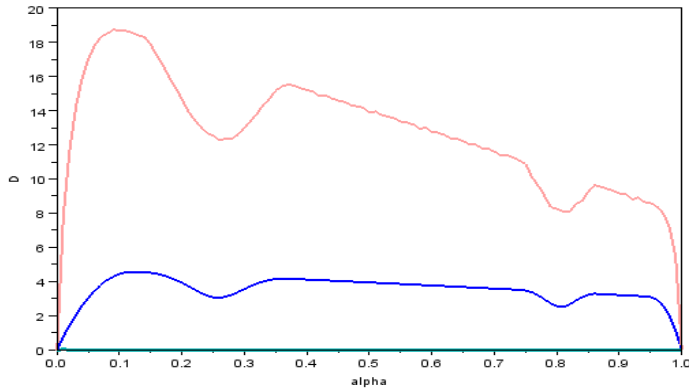
Figure 2

Distance differences for $d_{S1}()$, $d_{S2}()$ and $d_{S3}()$

Based on the test results, we conclude the following:

Property 1: For the entropy-difference distance $d_{S1}()$, for elements generated by KH interpolation, the distance difference $ddiff(A,B,C)$ is equal to zero.

Proof. Let us take trapezoid membership functions with the following parameters for a set A:

$$A_1 = \inf\{A_{\alpha=0}\}$$
$$A_2 = \inf\{A_{\alpha=1}\}$$
$$A_3 = \sup\{A_{\alpha=1}\}$$
$$A_4 = \sup\{A_{\alpha=0}\}$$

where symbol $A_{\alpha=c}$ denotes the set of points with the membership function A equal to c. The *entropy(A)* differs from zero only on the intervals (A_1,A_2) and (A_3,A_4). The entropy value *entropy((A1,A2))* is calculated with

$$-\int_0^{A_2-A_1}\left(\frac{x}{A_2-A_1}\log\left(\frac{x}{A_2-A_1}\right)+(1-\frac{x}{A_2-A_1})\log(1-\frac{x}{A_2-A_1})\right)dx\ .$$

With corresponding substitutions, the integral can be transformed into the form

$$-2(A_2-A_1)\int_0^1 z\log(z)dz = -2(A_2-A_1)\left[\frac{2z^2\log(z)-z^2}{4}\right]_0^1 = \frac{A_2-A_1}{2}\ .$$

Thus, the entropy value for the set A, is equal with

$$entropy(A(A_1,A_2,A_3,A_4)) = \frac{(A_2-A_1)+(A_4-A_3)}{2}\ ,$$

i.e., it is equal to the length of it non-crisp parts. Taking a *C* KH-interpolated set with parameter λ, the *C* will be also a trapezoid fuzzy set with the following parameters:

$$C_i = \lambda \cdot A_i + (1 - \lambda) \cdot B_i .$$

It follows from the linearity that also

$$entropy(C) = \lambda \cdot entropy(A) + (1 - \lambda) \cdot entropy(B)$$

holds. Thus,

$$entropy(C) \in [\min\{entrpoy(A), entropy(B)\}, \max\{entrpoy(A), entropy(B)\}]$$

and *ddiff(A,B,C)* = 0 is met.

Assuming the membership function can be approximated with a chain of linear segments, the *ddiff(A,B,C)* = 0 condition is fulfilled for fuzzy sets of arbitrary shapes. ■

Property 2. For every *A,B* $\in$ *Φ(E)*, the $d_{S3}(A,B) \geq d_{S2}(A,B)$ inequality holds.

Proof. Consider first the following inequality,

$$H(X_{A\Delta B}(x)) \geq \left| entropy_A(\{x\}) - entropy_B(\{x\}) \right| . \tag{9}$$

for every *x* $\in$ *E*. The inequality in (9) can be converted into the following expression:

$$K(x) = entropy_{A\Delta B}(\{x\}) - \left| entropy_A(\{x\}) - entropy_B(\{x\}) \right| \geq 0 .$$

The *entropy()* function can be substituted with its definition:

$$K(x) = -\left|x_A - x_B\right| \log(\left|x_A - x_B\right|) - (1 - \left|x_A - x_B\right|) \log(1 - \left|x_A - x_B\right|) - \tag{10}$$
$$\left| -x_A \log(x_A) - (1 - x_A) \log(1 - x_A) + x_B \log(x_B) + (1 - x_B) \log(1 - x_B) \right| .$$

where x_A denotes $m_A(x)$.

Let us fix x_b to a value *b* and simplify notation x_a to *x*. As (10) contains two absolute value expressions, four different subdomains should be defined:

$$R1 : x < b, entropy(x) < entropy(b)$$
$$R2 : x > b, entropy(x) < entropy(b)$$
$$R3 : x > b, entropy(x) > entropy(b)$$
$$R4 : x < b, entropy(x) > entropy(b)$$

In subdomain R1, formula (10) can be written as

$$K(x) = -(b - x) \log(b - x) - (1 - b + x) \log(1 - b + x) +$$
$$b \log(b) + (1 - b) \log(1 - b) - x \log(x) + (1 - x) \log(1 - x) .$$

The extreme point of $K()$ meets the following equation

$$\frac{\partial K}{\partial x} = -\log(b-x) + \log(1-b+x) - \log(x) + \log(1-x) = 0.$$

This yields in

$$\frac{(1-b+x)x}{(b-x)(1-x)} = 1$$

and

$$x = \frac{b}{2}.$$

In subdomain R1, the extreme points lie on the line $y = 2x$. In a similar way, the extreme points are the following in the other subdomains:

$$R1 : y = 2x$$
$$R2 : y = 2x - 1$$
$$R3 : no\ solution$$
$$R4 : no\ solution$$

As can be easily verified, the following conditions are met:

$$K(0) = 0$$
$$K(1) = 0 .$$
$$K(b) = 0$$

Thus, for every $b \in [0,1]$, the K(x) function has the following function-value segments: zero, increasing, decreasing, zero, increasing, decreasing, zero. From this fact, it follows that

$$K(x) \geq 0$$

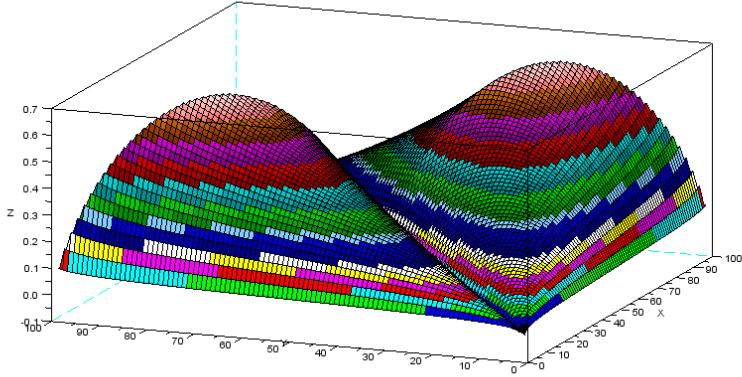for every x and b value. Thus condition (9) is met. The measured K() values are given in Figure 3.

Figure 3
The K() difference function

From the fact

$$H(X_{A\Delta B}(x)) \ge \left| entropy_A(\{x\}) - entropy_B(\{x\}) \right|$$

it follows that

$$\sum H(X_{A\Delta B}(x))^2 \ge \sum \left| entropy_A(\{x\}) - entropy_B(\{x\}) \right|^2$$

and

$$\left(\sum H(X_{A\Delta B}(x)) \right)^2 \ge \sum H(X_{A\Delta B}(x))^2 \ge \sum \left| entropy_A(\{x\}) - entropy_B(\{x\}) \right|^2 .$$

Extending the expression to infinite elements, we get the expected property

$$d_{S3}(A,B) \ge d_{S2}(A,B) . \blacksquare$$

As can be seen, the KH interpolation algorithm is not suitable to generate a fuzzy set C lying on the middle point between A and B, i.e.

$$d_{S3}(A,C) = d_{S3}(B,C) = \frac{d_{S3}(A,B)}{2} .$$

In the next step, an algorithm is presented for generating the required C set.

Property 3. The required C_λ set can be generated from A, B in such a way that every elements of $C_\lambda(x)$ is either equal to $A(x)$ or to $B(x)$.

Proof. For the required element C_λ, the equation

$$ddiff(A,B,C_\lambda) = d_{S3}(A,C_\lambda) + d_{S3}(B,C_\lambda) - d_{S3}(A,B) = 0$$

should be met. It follows from definition (8) that

$$ddiff(A,B,C_\lambda) = \int_{-\infty}^{\infty} H(X_{A\Delta C_\lambda}(x)) + H(X_{B\Delta C_\lambda}(x)) - H(X_{A\Delta B}(x))dx .$$

In a similar way, as was shown in the proof of Property 2, we get

$$H(X_{A\Delta C_\lambda}(x)) + H(X_{B\Delta C_\lambda}(x)) - H(X_{A\Delta B}(x)) \geq 0, \forall x \in [0,1]$$

and

$$dd(x) = (H(X_{A\Delta C_\lambda}(x)) + (H(X_{B\Delta C_\lambda}(x)) - H(X_{A\Delta B}(x)))) = 0$$

if and only if

$$m_{C_\lambda}(x) = m_A(x) \ or$$
$$m_{C_\lambda}(x) = m_B(x) \qquad .$$

If $m_A(x) = 1$ (or $= 0$) then $m_C(x)$ can be equal to zero (or 1) too. The same is true for $m_B(x)$ also. ■

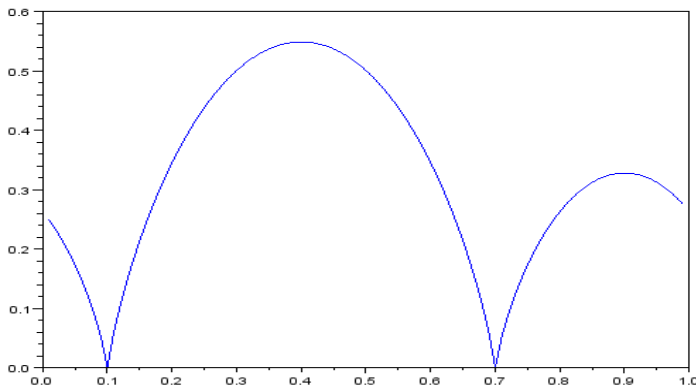Figure 4 shows the $dd(x)$ value for $m_C(x) \in [0..1]$, $m_A(x) = 0.1$, $m_B(x) = 0.7$.

Figure 4
The dd() difference function

Based on this result, a constructive algorithm can be given to generate C_λ from the sets A and B. The algorithm assigns points to C_λ from A in a greedy way, until it reaches the required distance value:

```
Gen(λ,A,B)
C = B
i = 1
while (d_S3(A,C) > λd_S3(A,B)) {
        C =A(0..i) ∪ C (i+1..N)
        i++
}
```

In Figure 5, the fuzzy sets generated by KH and the proposed Gen() function are displayed. The two target trapezoid fuzzy sets *A* and *B* are shown in Figure 5a. The KH interpolated fuzzy set *C′* with λ=0.5 is given in Figure 5b in the middle in a solid blue line. The interpolated fuzzy set C″ generated with Gen() is shown in Figure 5b with a thick brown line.

In the example, the following distance values can be measured:

$$d_{S3}(A, B) = 62.45$$
$$d_{S3}(A, C') = 56.20$$
$$d_{S3}(B, C') = 68.71$$
$$d_{S3}(A, C'') = 31.23$$
$$d_{S3}(B, C'') = 31.23$$

Thus, the Gen() method yields the required distance relationship for the interpolated *C* set using the descriptive complexity distance.
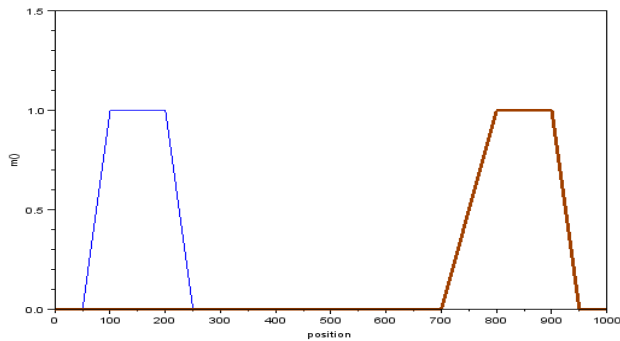

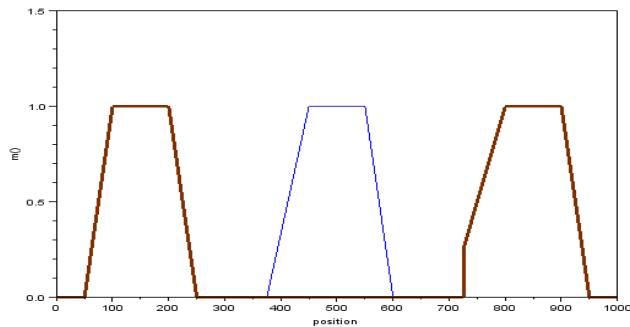Figure 5a

The A and B fuzzy sets

Figure 5b

The interpolated C sets

Conclusion

This paper analyzes the distance relationship among the base and generated fuzzy sets for KH linear interpolation. In the case of Euclidean distance, the usual behavior can be seen, but in the case of entropy-based distances, the new generated sets do not provide the distance optimum. The paper presents new properties among the entropy-based distances and proposes an appropriate method of distance optimum interpolation.

Acknowledgement

References

[1]   M. Deza and E. Deza. *Encyclopedia of Distances*, Vol. 15 of Series in Computer Science. Springer-Verlag, 2009

[2]   J. Ratsaby. Information Efficiency. In *Proc. of 33rd Conference on Current Trends in Theory and Practice of Computer Science (SOFSEM '07)*, Vol. LNCS 4362, pp. 475-487, 2007

[3]   J. Ratsaby, Information Set-Distance, *Proc. of the 2010 Mini-Conference on Applied Theoretical Computer Science (MATCOS 2010)*, pp. 61-64, University of Primorska Press, Koper, Slovenia, 2011

[4]   L. A. Zadeh. Fuzzy Sets. *Information Control*, 8:338-353, 1965

[5]   R. Zwick, E. Carlstein, and D. V. Budescu. Measures of Similarity among Fuzzy Concepts: A Comparative Analysis. *International Journal of Approximate Reasoning*, 1:221-242, 1987

[6]   L. Kovács, J. Ratsaby: Descriptive-Complexity-based Distance for Fuzzy Sets, CoRR abs/1012.3410: (2010)

[7]   Zs. Cs. Johanyák, Sz. Kovács: A Brief Survey and Comparison on Various Interpolation-based Fuzzy Reasoning Methods, *Acta Polytechnica Hungarica*, Vol. 3, No. 1, 2006, pp. 91-105

[8]   Kóczy, L. T., Hirota, K.: Rule Interpolation by α-Level Sets in Fuzzy Approximate Reasoning, In *J. BUSEFAL, Automne, URA-CNRS*, Vol. 46, Toulouse, France, 1991, pp. 115-123

[9]   Wong, K. W., Gedeon, T. D., Tikk, D.: An Improved Multidimensional α-Cut-based Fuzzy Interpolation Technique, In *Proc. Int. Conf Artificial Intelligence in Science and Technology (AISAT 2000)*, Hobart, Australia, 2000, pp. 29-32

[10]   Gedeon, T. D., Kóczy, L. T.: Conservation of Fuzziness in the Rule Interpolation, Intelligent Technologies, *Int. Symposium on New Trends in Control of Large Scale Systems*, Vol. 1, Herl'any, 1996, pp. 13-19

[11]   Tikk, D., Baranyi, P.: Comprehensive Analysis of a New Fuzzy Rule Interpolation Method*, IEEE Trans Fuzzy Syst.*, Vol. 8, June 2000, pp. 281-296

[12]   Perfiliva, I., Wrublova, M., Hodakova, P.: Fuzzy Interpolation According to Fuzzy and Classical Conditions, *Acta Polytechnica Hungarica*, Vol. 7, No. 4, 2010, pp. 39-54