Johanyák, Z. C., Kovács S.: Distance based similarity measures of fuzzy sets, SAMI 2005, 3rd Slovakian-Hungarian Joint Symposium on Applied Machine Intelligence, Herl'any, Slovakia, January 21-22 2005, ISBN 963 7154 35 3, pp. 265-276. http://johanyak.hu

Distance based similarity measures of fuzzy sets

Zsolt Csaba Johanyák

Department of Information Technology, Kecskemét College, GAMF Faculty
Kecskemét, H-6001 Pf. 91, Hungary
johanyak.csaba@gamf.kefo.hu

Szilveszter Kovács

Department of Information Technology, University of Miskolc,
Miskolc-Egyetemváros, Miskolc, H-3515, Hungary
szkovacs@iit.uni-miskolc.hu

Abstract: In case of fuzzy reasoning in sparse fuzzy rule bases, the question of selecting the suitable fuzzy similarity measure is essential. The rule antecedents of the sparse fuzzy rule bases are not fully covering the input universe therefore fuzzy reasoning methods applied for sparse fuzzy rule bases requires similarity measures able to distinguish the similarity of non-overlapping fuzzy sets too. The goal of this paper is enumerating some of these distance based similarity measures and briefly introducing them.

Keywords: similarity measure, distance of fuzzy sets, vague environment

1 Distance based similarity measure

The most obvious way of calculating similarity of fuzzy sets is based on their distance. There are more approaches on how the relation between the two notions in form of a function can be expressed. Two of them are presented below.

The first function is the following [7]:

$$\mathrm{SM}(\mathrm{A},\mathrm{B}) = \frac{1}{1+\mathrm{DM}(\mathrm{A},\mathrm{B})}, \tag{1}$$

where SM is the similarity measure, DM is the distance measure of two fuzzy sets, and A respective B are the examined fuzzy sets.

Another way of distance based similarity assessment is proposed by Williams and Steele in [1]. The suggested formula (2) contains an exponential expression.

$$SM(A,B) = e^{-\alpha \cdot DM(A,B)} \quad (2)$$

where α is a steepness measure. The value α=7 was found suitable for the practice in case of a one dimensional universe of discourse.

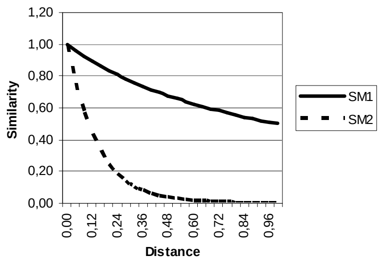

Fig. 1.

The functions (1) and (2) marked with SM1 and SM2 are presented in Fig. 1. using normalized distances. SM1 has a uniform sensitivity opposing to SM2 which has a far higher sensitivity and capability for distinction in the first quarter of the interval.

In case of a multi-dimensional universe of discourse the approximation should be started with a universal distance measure. Forming it needs the normalization of all linguistic variables for e.g. the interval [0,1]. It can be done by the help of Lipschitz functions [2].

The universal distance measure is determined as a weighted mean of the distances measured along each dimension (3).

$$DM_U(A,B) = \sum_{i=1}^{n} w_i \cdot DM_i(A,B) \quad (3)$$

where n is the number of the input linguistic variables, w_i is the weighting for the i^{th} linguistic variable and DM_i is the distance measured along the i^{th} dimension.

$$\alpha = \frac{7}{\sum_{i=1}^{n} w_i} \quad (4)$$

In a multi-dimensional case the value of α in (2) is determined by the formula (4) [1].

Instead of calculating similarities from distances, by a small re-explaining of the meanings of the fuzzy rules, we can use the distances of fuzzy sets directly for approximate fuzzy reasoning.

Using distance based approximate fuzzy reasoning has an important precondition. The distance of fuzzy sets can be defined only on universes where it is possible to define full ordering and metrics on every component of the universe of discourse of the fuzzy sets (any other case the notion of distance is meaningless).

A distance function DM: X x X $\rightarrow$ R can be considered as metrics, if the conditions specified below are fulfilled [4]:

- $DM(A,B) \geq 0 \ \forall \ A,B \in X$
- $DM(A,B)=0 \Leftrightarrow A=B \ \forall \ A,B \in X$
- $DM(A,B)=DM(B,A) \ \forall \ A,B \in X$
- $DM(A,B)+DM(B,C) \geq DM(A,C) \ \forall \ A,B,C \in X$

The City Block (5) and the Euclidean (6) are often used as metrics for distance measure in case of crisp values.

$$DM = \sum_{i=1}^{n} \left| A_i - B_i \right|, \tag{5}$$

$$DM = \sqrt{\sum_{i=1}^{n} \left(A_i - B_i \right)^2}, \tag{6}$$

where n is the number of dimensions and i is the serial number of the actual dimension.

2 Non α-cut based similarity measures

There are many useful distance definitions of fuzzy sets in the literature. The simplest one is the *Disconsistency Measure* (S_D) of the fuzzy sets *A* and *B* (7)

$$S_D = 1 - \sup_{x \in X} \mu_{A \cap B}(x) \tag{7}$$

where $A \cap B$ is the min *t*-norm, $\mu_{A\cap B}(x)=\min\{\mu_A(x), \mu_B(x)\} \ \forall \ x \in X$. It is basically the same measure as used in the min-max composition. The disconsistency measure is one crisp value in range of [0,1].

In the followings, some distance measures, which are used for expressing the similarity of trapezoidal shaped fuzzy sets (or fuzzy sets have membership functions can be traced back to a trapezoid form) will be presented.

In case of trapezoidal shaped fuzzy sets, the fuzzy set can be characterised by a vector of four values, by the upper and lower endpoints of the core and support e.g. $X=[x_1,x_2,x_3,x_4]$.

This case the similarity between sets A and B can be described by the formula (8) proposed by Chen [5].

$$SM(A,B) = 1 - \frac{\sum_{i=1}^{4} |a_i - b_i|}{4} \tag{8}$$

If the universe of the fuzzy sets are normalized, then SM(A,B)$\in$[0,1]. The advantage of (8) is its simplicity and low computational complexity. However, its drawback is that it can easily lead to the same grade of similarity in case of different shapes, too.

For instance if the trapezoid fuzzy set A=[0.2,0.4,0.6,0.8] is compared to the trapezoid term B=[0.4,0.6,0.8,1.0] and to the triangle shaped set C=[0.4,0.7,0.7,1.0] and to the D=[0.7,0.7,0.7,0.7] crisp value, the similarity measure is 1.6 in each case.

Chen and Chen proposed a method in [6], which can be used in case of generalized trapezoid shaped fuzzy sets, too. This similarity measure (9) is based on the calculation of the Center Of Gravity.

$$SM(A,B) = \left[1 - \frac{\sum_{i=1}^{4} |a_i - b_i|}{4}\right] \times \left(1 - \left|x_A^* - x_B^*\right|\right)^{C(A,B)} \times \frac{\min\left(y_A^*, y_B^*\right)}{\max\left(y_A^*, y_B^*\right)} \tag{9}$$

where C(A,B) is defined as follows:

$$C(A,B) = \begin{cases} 1 & a_4 - a_1 + b_4 - b_1 > 0 \\ 0 & a_4 - a_1 + b_4 - b_1 = 0 \end{cases} \tag{10}$$

x_A^* and y_A^* are the coordinates of the COG of the set A, respective x_B^* and y_B^* determine the COG of the set B. The disadvantage of this method is that it can not handle cases when the examined sets have the same COG, but their shape is different. The increased computational complexity can be considered as an additional drawback.

3 α-cut based similarity measures

3.1 Simple distance measures

Most of the distance definitions are based on the α-cuts of the two fuzzy sets, for example:

Hausdorff Measure (∞):

$$\mathrm{HM}_\infty(A,B)=\sup_{\alpha\geq 0}\mathrm{HM}(A_\alpha,B_\alpha) \tag{11}$$

Hausdorff Measure (*):

$$\mathrm{HM}_*(A,B)=\mathrm{HM}(A_1,B_1) \tag{12}$$

where

$$\mathrm{HM}(U,V)=\max\left\{\sup_{v\in V}\inf_{u\in U}\mathrm{d}(u,v),\sup_{u\in U}\inf_{v\in V}\mathrm{d}(u,v)\right\} \tag{13}$$

and $\mathrm{d}(u,v)$ is the Euclidean distance.

Kaufmann and Gupta Measure (∞):

$$\Delta_\infty(A,B)=\sup_{\alpha\geq 0}\Delta(A_\alpha,B_\alpha) \tag{14}$$

Kaufmann and Gupta Measure (*):

$$\Delta_*(A,B)=\Delta(A_1,B_1) \tag{15}$$

where

$$\Delta(A_\alpha,B_\alpha)=\frac{(|a_1-b_1|+|a_2-b_2|)}{2\cdot(\beta_2-\beta_1)} \tag{16}$$

and $[a_1, a_2]$, $[b_1, b_2]$ are the supports of A_α, B_α, respectively $[\beta_1, \beta_2]$ is the support of both A_α and B_α, $\alpha\in[0,1]$.

Both the Hausdorff Measure and Kaufmann and Gupta Measure are a crisp value in range of $[0,\infty]$.

3.2 Kóczy's distance measure

The main problem of the distance definitions presented above is, that the information of the shape of the membership function of the fuzzy sets is mostly lost. It is impossible to reconstruct from a given fuzzy set *A* and from a given Hausdorff or Kaufmann and Gupta distance measure of two fuzzy sets *A* and *B*, the fuzzy set *B*. This type of reconstruction, at least in the one dimensional case, has a great importance in rule interpolation, because without it, from the distances of the rule consequents and the fuzzy conclusion we are looking for, it is impossible to reconstruct the shape of the fuzzy conclusion.

Solving these difficulties a useful definition is introduced by Kóczy [7]. This distance is based on the α-cuts of the two fuzzy sets too, but the distance is not aggregated to one crisp value, so from this kind of distance and from one of the fuzzy sets the other set can be reconstructed.

The distance of two fuzzy sets is expressed by means of a fuzzy set which is defined over the interval [0,1]. In the course of calculations the Euclidean distances between the end points of the α-cuts are considered. These are called lower (d_L^α) and upper (d_U^α) distances and are calculated by formulas (17) and (18) (Fig. 2.).

$$d_L^\alpha(A,B) = \inf\{B^\alpha\} - \inf\{A^\alpha\} \tag{17}$$

$$d_U^\alpha(A,B) = \sup\{B^\alpha\} - \sup\{A^\alpha\} \tag{18}$$

If the universe of discourse is multi-dimensional, the distances between $\inf\{A_\alpha\}$, $\inf\{B_\alpha\}$ and $\sup\{A_\alpha\}$, $\sup\{B_\alpha\}$ can be defined in the Minkowski sense:

$$d_L^\alpha(A,B) = \left(\sum_{i=1}^{k} \mathrm{d}_L^\alpha(A_i, B_i)^w\right)^{1/w} \tag{19}$$

$$d_U^\alpha(A,B) = \left(\sum_{i=1}^{k} \mathrm{d}_U^\alpha(A_i, B_i)^w\right)^{1/w} \tag{20}$$

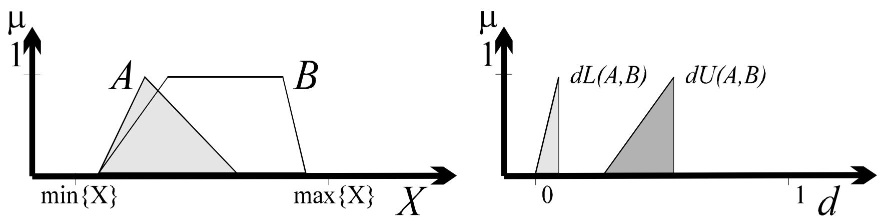

Fig. 2. Normalised fuzzy distance between the fuzzy sets ***A*** and ***B***

An important restriction for the existence of the Kóczy Distance is that all the comparable fuzzy sets should be convex and normal, otherwise some α-cuts are not connected or do not exists at all, which makes the distance corresponding to these α-cuts meaningless. The only disadvantage of using the Kóczy Distance for interpolative fuzzy reasoning is that it is little bit difficult to handle.

4 Vague distance of points in a vague environment

In the case of rule interpolation it would be useful such kind of distance definition, which is easy to handle, for example the distance of two fuzzy sets could be characterised by one crisp number, and give the chance of the reconstruction of the membership function of a fuzzy set from another set and from their distance, at least in the one dimensional case.

These seem to be two contradictory conditions, but they can be satisfied, if we can find a way for handling the distance of the fuzzy sets and a kind of shape description separately.

4.1 Connection between similarity of fuzzy sets and vague distance of points in a vague environment

The concept of vague environment is based on the similarity or indistinguishability of the elements. The x_1 and x_2 values in the vague environment are ε-distinguishable if their distance ($\delta(x_1,x_2)$) is greater than ε (21). The distances in vague environment are weighted distances. The weighting factor or function is called scaling function (s(x)).

$$\delta_s(x_1,x_2)=\left|\int_{x_2}^{x_1}\mathrm{s(x)dx}\right|>\varepsilon \tag{21}$$

For finding connections between fuzzy sets and a vague environment we can introduce the membership function $\mu_A(x)$ as a level of similarity of **a** to *x*. The α-cuts of the fuzzy set described by membership function $\mu_A(x)$ (23) form the set which contains the elements that are (1–α)-indistinguishable from **a** (Fig. 3.) (22):

$$\delta_s(\mathrm{a,b})\le 1-\alpha \tag{22}$$

$$\mu_A(x)=1-\min\{\delta_s(\mathrm{a,b}),1\}=1-\min\left\{\left|\int_{\mathrm{a}}^{\mathrm{b}} s(x)dx\right|,1\right\} \tag{23}$$

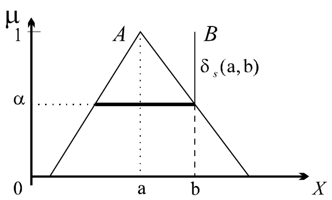

Fig. 3.

The vague distance of points a and b (δ(a,b)) is basically the *Disconsistency Measure* (24) of the fuzzy sets *A* and *B* (where *B* is a singleton):

$$S_D=1-\sup_{x\in X}\mu_{A\cap B}(x)=\delta_s(\mathrm{a,b})\quad \text{if}\ \ \delta(\mathrm{a,b})\in[0,1] \tag{24}$$

Thus disconsistency measures between member fuzzy sets of a fuzzy partition and a singleton can be calculated, as vague distances of points in the vague environment of the fuzzy partition. The main difference between the disconsistency measure and the vague distance is, that the vague distance is a crisp value in range of [0,∞], while the disconsistency measure is limited to [0,1]. That is why it is useful in interpolative reasoning with insufficient evidence.

So if it is possible to describe all the fuzzy partitions of the antecedent and consequent universes of the fuzzy rule-base, and the observation is a singleton, one can calculate the disconsistency measures of the antecedent fuzzy sets of the rule-base and the observation, and the disconsistency measures of the consequent fuzzy sets and the consequence (we are looking for) as vague distances of points.

4.2 Generating vague environments from fuzzy partitions

The vague environment is described by its scaling function. For generating a vague environment we have to find an appropriate scaling function, which describes the shapes of all the terms in the fuzzy partition [8]. The method proposed by Klawonn [9], for choosing the scaling function *s*(*x*) (25), gives an exact description of the fuzzy terms after their reconstruction from the scaling function.

$$s(x) = \left|\mu'(x)\right| = \left|\frac{d\mu}{dx}\right| \qquad (25)$$

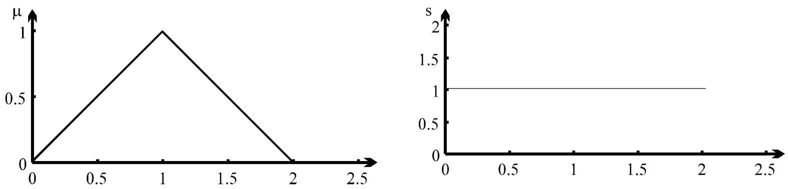

Fig. 4. A fuzzy set and its scaling function

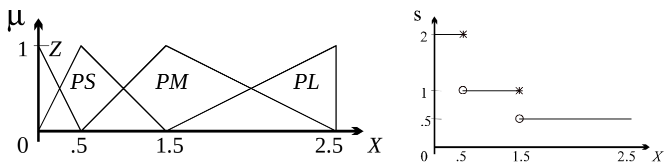

Fig. 5. Scaling function describing all the fuzzy sets

A scaling function always can be found, if there is only one fuzzy set in the fuzzy partition (Fig. 4.). Usually the fuzzy partition contains more than one fuzzy set, so this method requires some restrictions (26) [9].

$$\text{if } \min\{\mu_i(x),\mu_j\}>0 \Rightarrow \left|\mu'_i(x)\right| = \left|\mu'_j(x)\right| \quad \forall i,j\in I \qquad (26)$$

Generally the above condition is not fulfilled, so the use of an approximate scaling function is proposed as a "universal" function describing all the fuzzy sets of a fuzzy partition.

4.3 The approximate scaling function

The approximate scaling function is an approximation of the original scaling functions describing the fuzzy sets separately. The simplest way of generating this function is the linear interpolation. Supposing that the fuzzy sets are triangles, each of them can be characterised by three values, two constant scaling functions, which are the scaling factors of the left and the right slope of the triangle and the value of the core point (Fig. 6.).

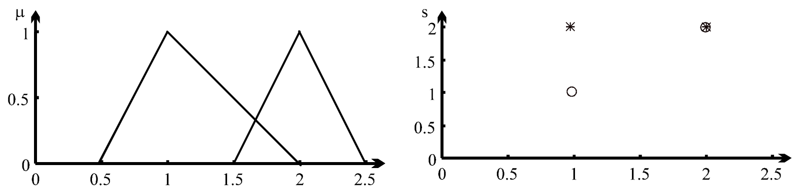

Fig. 6.

Thus the approximation (s(x)) is a piecewise linear function (27), which interpolates the right side scaling factor of the left neighbouring term and the left side scaling factor of the right neighbouring term (Fig. 7.).

$$s(x)=\left\{\frac{s_{i+1}^{L}-s_{i}^{R}}{x_{i+1}-x_{i}}\cdot(x-x_{i})+s_{i}^{R}\mid x\in[x_{i},x_{i+1}),\forall i\in[1,n-1]\right\} \qquad (27)$$

where

x_i is the core of the i^{th} term of the approximated fuzzy partition

s_i^L, s_i^R are the left and right side scaling factors of the i^{th} term

n is the number of the terms in the approximated fuzzy partition

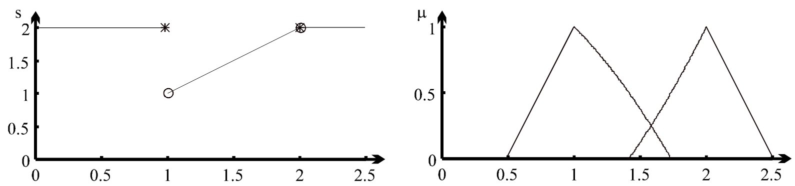

Fig. 7.

The drawback of the approximation presented above is that it can not handle the big differences between neighbouring scaling factors or crisp fuzzy sets correctly.

In case of big differences, the bigger scaling factor “dominates” the smaller one (Fig. 8., 9.). If one of the neighbouring fuzzy set is crisp (its scaling factor is infinite), the slope of the linearly interpolated scaling function is infinite too, so both the fuzzy sets described by this scaling function will be crisp.

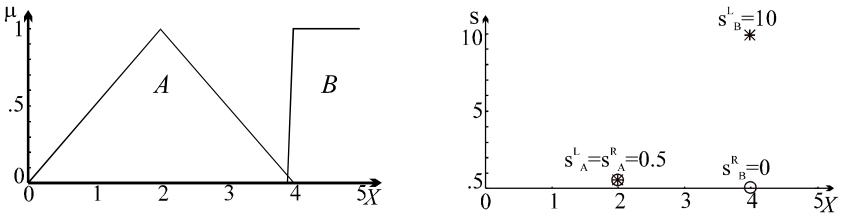

Fig. 8. $s^R_A << s^L_B$

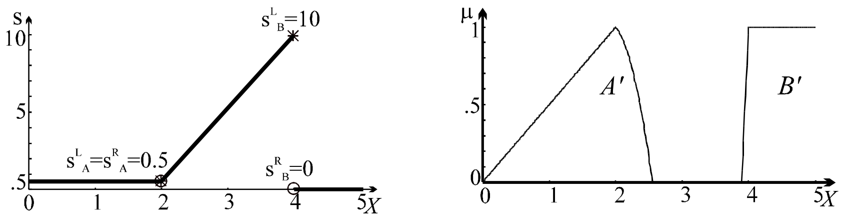

Fig. 9. Linearly interpolated scaling function of fuzzy sets shown in Fig. 8., and these sets as the approximate scaling function describes them (*A’*,*B’*)

As a solution of this problem the adoption of a non-linear interpolative function (28) is suggested [8].

$$
s(x)=\begin{cases} \dfrac{w_i}{(x_{i+1}-x_i+1)^{k\cdot w_i}-1}\cdot\left(\dfrac{(x_{i+1}-x_i+1)^{k\cdot w_i}}{(x-x_i+1)^{k\cdot w_i}}-1\right)+s^L_{i+1} & \mid s^R_i \ge s^L_{i+1} \\ \dfrac{w_i}{(x_{i+1}-x_i+1)^{k\cdot w_i}-1}\cdot\left(\dfrac{(x_{i+1}-x_i+1)^{k\cdot w_i}}{(x_{i+1}-x+1)^{k\cdot w_i}}-1\right)+s^R_i & \mid s^R_i < s^L_{i+1} \end{cases} \tag{28}
$$

$$
w_i=\left|s^L_{i+1}-s^R_i\right| \tag{29}
$$

where $x\in[x_i,x_{i+1}),\forall\, i\in[1,n-1]$,

s(*x*) is the approximate scaling function,

x_i is the core of the ith term of the approximated fuzzy partition,

s^L_i, s^R_i are the left and right side scaling factors of the ith triangle shaped term of the approximated fuzzy partition,

k constant factor of sensitivity for neighbouring scaling factor differences,

n is the number of the terms in the approximated fuzzy partition.

The above function has same useful properties. If the neighbouring scaling factors are equals, s(*x*) is linear. If one of the neighbouring scaling factors (e.g. S_i^R) $s_i^R \to \infty$ and the other one is finite, in case of

$$x \in \left[x_i, x_{i+1}\right) \Rightarrow s(x) \to \begin{Bmatrix} \infty \mid x \to x_i \\ 0 \mid x \neq x_i \end{Bmatrix}$$ and similarly

if $s_{i+1}^L \to \infty$ and s_i^R is finite and $x \in \left[x_i, x_{i+1}\right) \Rightarrow s(x) \to \begin{Bmatrix} \infty \mid x \to x_{i+1} \\ 0 \mid x \neq x_{i+1} \end{Bmatrix}$

Fig. 10. and 11. show some examples for the application of the proposed non-linear function.

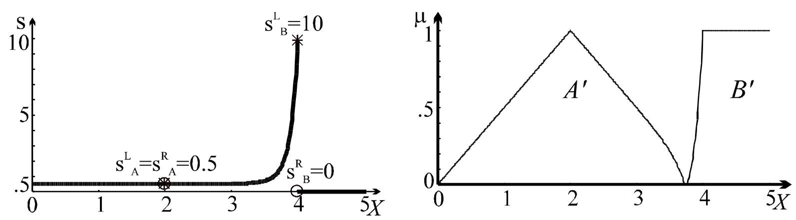

Fig. 10. Approximate scaling function generated by the non-linear function with *k*=1, and the original fuzzy partition (*A*,*B*) as this scaling function describes it (*A’*,*B’*)

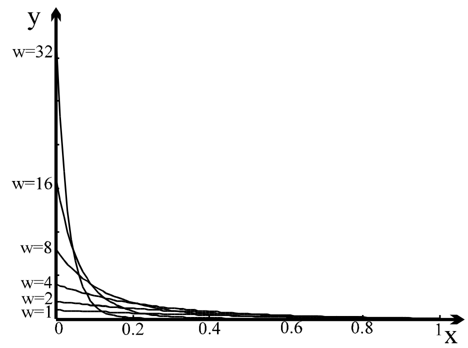

Fig. 11. x_1=0, x_2=1, $s_1^R \geq s_2^L$, *k*=1

5 Conclusions

Distance based similarity measures of fuzzy sets have a high importance in reasoning methods handling sparse fuzzy rule bases. The rule antecedents of the sparse fuzzy rule bases are not fully covering the input universe. Therefore the applied similarity measure has to be able to distinguish the similarity of non-overlapping fuzzy sets, too. The distance based similarity measures are such a measures.

To give an overview of the distance based similarity measures of fuzzy sets, some of the main existing concepts are briefly introduced in this paper.

6 References

[1] J. Williams, N. Steele: Difference, distance and similarity as a basis for fuzzy decision support based on prototypical decision classes, Fuzzy Sets and Systems 131 (2002) 35-46.

[2] P. Diamond, P. Kloeden: Metric Spaces of Fuzzy Sets: Theory and Application, World Scientific, Singapore, 1994.

[3] L. T. Kóczy, K. Hirota: Ordering, distance and closeness of fuzzy sets, Fuzzy Sets And Systems, 60:281-293, 1993.

[4] W. A. Shuterland: Introduction to metric and topological spaces, Oxford University Press, Oxford, 1977.

[5] S. M. Chen: New methods for subjective mental workload assesment and fuzzy risk analysis, Cybernet, Systems: Internat, 1996, J. 27, 449-472.

[6] S. J. Chen, S. M. Chen: Fuzzy risk analysis based on similarity measures of generalized fuzzy numbers, IEEE Trans. Fuzzy Systems 11,45-56, 2003.

[7] Kóczy T. László, Tikk Domonkos: Fuzzy rendszerek, Typotex, 2000.

[8] Sz. Kovács, L. T. Kóczy.: The use of the concept of vague environment in approximate fuzzy reasoning, Tatra Mt. Math. Publ. 12, 477–482, 1997.

[9] F. Klawonn: Fuzzy sets and vague environments, Fuzzy Sets and Systems 66:207-221, 1994.