

Distances, Hesitancy Degree and Flexible Querying via Neutrosophic Sets

A.A.Salama

Math and Computer Science
Department, Faculty of Science
Port Said University, EGYPT

Mohamed Abdelfattah

Information System
Department, Faculty of
Computers & Information,
Benha University,
Egypt

Mohamed Eisa

Computer Science Department,
Port Said University,
Egypt

ABSTRACT

Since the world is full of indeterminacy, the neutrosophics found their place into contemporary research. In this paper we, introduce the distances between neutrosophic sets: the Hamming distance, The normalized Hamming distance, the Euclidean distance and normalized Euclidean distance. We will extend the concepts of distances to the case of neutrosophic hesitancy degree. Added to, this paper suggest how to enrich intuitionistic fuzzy querying by the use of neutrosophic values..

General Terms

Your general terms must be any term which can be used for general classification of the submitted material such as Pattern Recognition, Security, Algorithms et. al.

Keywords

Neutrosophic Sets; Hamming distance; Euclidean Distance; Normalized Euclidean Distance; Intuitionistic Fuzzy Querying; Querying Databases; Neutrosophic Querying.

1. INTRODUCTION

Since the world is full of indeterminacy, the neutrosophics found their place into contemporary research. The fundamental concepts of neutrosophic set, introduced by Smarandache in [15, 16], and Salama et al. in [4, 10, 11, 12, 13, 14, 18, 19, 20, 21, 22, 23], provides a natural foundation for treating mathematically the neutrosophic phenomena which exist pervasively in our real world and for building new branches of neutrosophic mathematics. Neutrosophy has laid the foundation for a whole family of new mathematical theories generalizing both their classical and fuzzy counterparts [1, 2, 3, 5, 6, 7, 8, 17, 24, 25] such as a neutrosophic set theory. The traditional query languages, used in the database management systems, require a precise and unambiguous specification of a query. It seems to be a serious limitation since a typical user often formulates his requirements in a natural language using imprecise expressions and vague terms. For this reason several approaches have been proposed to relax the rigidity of the conventional queries and make possible to use queries that allow for a more intelligent and human consistent information retrieval (see, e.g. [5, 8, 9]). The FQUERY for Access in [5, 6, 7, 8, 9] is an example of a computer program that enables to create different kinds of fuzzy queries. Using such fuzzy queries we deal no longer with binary outputs whether a record fulfill given requirement or not – but we get an

Definition 3.1

information on the degree the record complies with the requirement. In this paper we, introduce the distances between neutrosophic sets: the Hamming distance, The normalized Hamming distance, the Euclidean distance and normalized Euclidean distance. We will extend the concepts of distances to the case of neutrosophic hesitancy degree.

2. TERMINOLOGIES

Neutrosophy has laid the foundation for a whole family of new mathematical theories generalizing both their classical and fuzzy counterparts [1, 2, 3, 5, 6, 7, 8, 17, 24, 25] such as a neutrosophic set theory. We recollect some relevant basic preliminaries, and in particular, the work of Smarandache in [15, 16] and Salama et al. [4, 10, 11, 12, 13, 14, 18, 19, 20, 21, 22, 23]. Smarandache introduced the neutrosophic components T, I, F which represent the membership, indeterminacy, and non-membership values respectively, where is nonstandard unit interval. Salama introduced the following: Let X be a non-empty fixed set. A neutrosophic set is an object having the form where and which represent the degree of member ship function (namely), the degree of indeterminacy (namely), and the degree of non-member ship (namely) respectively of each element to the set where and . Smarandache introduced the following: Let T, I, F be real standard or nonstandard subsets of , with

$$\text{Sup}_T = t_{\text{sup}}, \text{inf}_T = t_{\text{inf}}$$

$$\text{Sup}_I = i_{\text{sup}}, \text{inf}_I = i_{\text{inf}}$$

$$\text{Sup}_F = f_{\text{sup}}, \text{inf}_F = f_{\text{inf}}$$

$$n\text{-sup} = t_{\text{sup}} + i_{\text{sup}} + f_{\text{sup}}$$

$$n\text{-inf} = t_{\text{inf}} + i_{\text{inf}} + f_{\text{inf}},$$

T, I, F are called neutrosophic components

3. DISTANCES BETWEEN NEUTROSOPHIC SETS

We will now extend the concepts of distances presented in [17] to the case of neutrosophic sets.

Let $A = \{(\mu_A(x), \nu_A(x), \gamma_A(x)), x \in X\}$ and $B = \{(\mu_B(x), \nu_B(x), \gamma_B(x)), x \in X\}$ in $X = \{x_1, x_2, x_3, \dots, x_n\}$ then

i) The Hamming distance is equal to

$$d_{Ns}(A, B) = \sum_{i=1}^n (|\mu_A(x_i) - \mu_B(x_i)| + |\nu_A(x_i) - \nu_B(x_i)| + |\gamma_A(x_i) - \gamma_B(x_i)|).$$

ii) The Euclidean distance is equal to

$$e_{Ns}(A, B) = \sqrt{\sum_{i=1}^n ((\mu_A(x_i) - \mu_B(x_i))^2 + (\nu_A(x_i) - \nu_B(x_i))^2 + (\gamma_A(x_i) - \gamma_B(x_i))^2)}$$

iii) The normalized Hamming distance is equal to

$$NH_{Ns}(A, B) = \frac{1}{2n} \sum_{i=1}^n (|\mu_A(x_i) - \mu_B(x_i)| + |\nu_A(x_i) - \nu_B(x_i)| + |\gamma_A(x_i) - \gamma_B(x_i)|)$$

iv) The normalized Euclidean distance is equal to

$$NE_{Ns}(A, B) = \sqrt{\frac{1}{2n} \sum_{i=1}^n ((\mu_A(x_i) - \mu_B(x_i))^2 + (\nu_A(x_i) - \nu_B(x_i))^2 + (\gamma_A(x_i) - \gamma_B(x_i))^2)}$$

Example 3.1

Let us consider for simplicity degenerated neutrosophic sets A, B, D, G, F in $X = \{a\}$. A full description of each neutrosophic set i.e. $A = \{(\mu_A(x), \nu_A(x), \gamma_A(x)), a \in X\}$, may be exemplified by $A = \{1, 0, 0\}, a \in X\}$, $B = \{0, 1, 0\}, a \in X\}$,

$D = \{0, 0, 1\}, a \in X\}$, $G = \{0.5, 0.5, 0\}, a \in X\}$, $E = \{0.25, 0.25, 0.05\}, a \in X\}$. Let us calculate four distances between the above neutrosophic sets using i), ii), iii) and iv) formulas ,

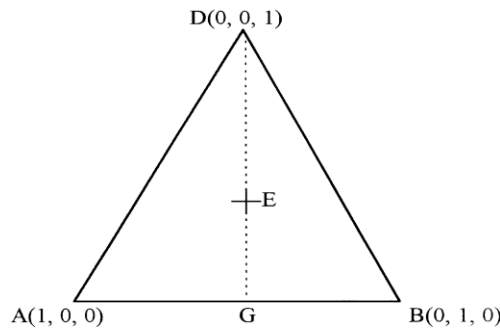


Fig.1A geometrical interpretation of the neutrosophic considered in Example 5.1 .

We obtain $e_{Ns}(A, D) = \frac{1}{2}$, $e_{Ns}(B, D) = \frac{1}{2}$, $e_{Ns}(A, B) = \frac{1}{2}$, $e_{Ns}(A, G) = \frac{1}{2}$, $e_{Ns}(B, G) = \frac{1}{2}$, $e_{Ns}(E, G) = \frac{1}{4}$, $e_{Ns}(D, G) = \frac{1}{4}$,

$NE_{Ns}(A, B) = 1$, $NE_{Ns}(A, D) = 1$, $NE_{Ns}(B, D) = 1$, $NE_{Ns}(A, G) = \frac{1}{2}$, $NE_{Ns}(B, G) = \frac{1}{2}$, $NE_{Ns}(E, G) = \frac{1}{2}$,

$NE_{Ns}(E, G) = \frac{\sqrt{3}}{4}$, and $NE_{Ns}(D, G) = \frac{\sqrt{3}}{2}$,

From the above results the triangle ABD (Fig.1) has edges equal to $\sqrt{2}$ and $e_{Ns}(A, D) = e_{Ns}(B, D) = e_{Ns}(A, B) = \frac{1}{2}$

and $NE_{Ns}(A, B) = NE_{Ns}(A, D) = NE_{Ns}(B, D) = 2NE_{Ns}(A, G) = 2NE_{Ns}(B, G) = 1$, and $NE_{Ns}(E, G)$ is equal to half of the height of triangle with all edges equal to $\sqrt{2}$ multiplied by, $\frac{1}{\sqrt{2}}$ i.e. $\frac{\sqrt{3}}{4}$.

Then $d_{Ns}(A, B) = 3$, $NH_{Ns}(A, B) = 0.43$, $e_{Ns}(A, B) = 1.49$ and $NE_{Ns}(A, B) = 0.55$.

Remark 3.1

Clearly these distances satisfy the conditions of metric space.

Remark 3.2

Example 3.2

Let us consider the following neutrosophic sets A and B in $X = \{a, b, c, d, e\}$,

$A = \{0.5, 0.3, 0.2\}, \{0.2, 0.6, 0.2\}, \{0.3, 0.2, 0.5\}, \{0.2, 0.2, 0.6\}, \{1, 0, 0\}$

$B = \{0.2, 0.6, 0.2\}, \{0.3, 0.2, 0.5\}, \{0.5, 0.2, 0.3\}, \{0.9, 0, 0.1\}, \{0, 0, 0\}$

It is easy to notice that for formulas i), ii), iii) and iv) the following is valid:

vi) The normalized Euclidean distance is equal to

$$NE_{Ns}(A, B) = \sqrt{\frac{1}{2n} \sum_{i=1}^n ((\mu_A(x_i) - \mu_B(x_i))^2 + (\nu_A(x_i) - \nu_B(x_i))^2 + (\gamma_A(x_i) - \gamma_B(x_i))^2 + (\pi_A(x_i) - \pi_B(x_i))^2)}$$

Remark 3.3

It is easy to notice that for formulas i), ii), iii) and iv) the following is valid:

- a) $0 \leq d_{Ns}(A, B) \leq 2n$
- b) $0 \leq NH_{Ns}(A, B) \leq 2$
- c) $0 \leq e_{Ns}(A, B) \leq \sqrt{2n}$
- d) $0 \leq NE_{Ns}(A, B) \leq \sqrt{2}$.

4. QUERYING VIA NEUTROSOPHIC SETS

A query may be treated as a set of searching criteria conceived by a user. A typical query expressed in SQL is written in a following form

```
SELECT < list of attributes >
FROM < list of tables >
WHERE < condition >.
```

Its role is to select records (rows) that satisfy given condition. Each record from the table either satisfies or does not satisfy the condition and as a result we obtain a crisp set of database records that come up to query. However, as it was mentioned above, traditional query syntax requires very rigid formulation of the constraints, while for a human being a common language is a natural medium to form and express his thoughts. Now we will try to construct a query that enables a direct use of linguistic terms modeled by neutrosophic sets, i.e. a query with a following syntax:

```
SELECT < list of attributes >
FROM < list of tables >
WHERE < neutrosophic condition >.
```

Let us consider a crisp relational database with a set of attributes $\Delta = \{A_1, A_2, \dots, A_n\}$ and a set of records

$$R = \{r_1, r_2, \dots, r_m\}.$$

Let X_j denote the universe of discourse for the attribute A_j . Moreover, let

$$Z : R \rightarrow X_1 \times X_2 \times \dots \times X_n$$

denote a function that determines a vector of values of all attributes corresponding to

Let us define a function $U : R \rightarrow NS(R)$ which determines a neutrosophic set R_i for each record r_i in a following way

$$R_i = U(r_i) = \{ \langle A_1, \mu_{R_i}(A_1), \nu_{R_i}(A_1), \sigma_{R_i}(A_1) \rangle, \dots, \langle A_n, \mu_{R_i}(A_n), \nu_{R_i}(A_n), \sigma_{R_i}(A_n) \rangle \};$$

where $\mu_{R_i}(A_j) = \mu_{A_j}^T(z_{ij})$, $\nu_{R_i}(A_j) = \nu_{A_j}^T(z_{ij})$, and $\sigma_{R_i}(A_j) = \sigma_{A_j}^T(z_{ij})$.

In other words $R_i = \{ \langle A_1, \mu_{A_1}(z_{i1}), \nu_{A_1}(z_{i1}), \sigma_{A_1}(z_{i1}) \rangle, \dots, \langle A_n, \mu_{A_n}(z_{in}), \nu_{A_n}(z_{in}), \sigma_{A_n}(z_{in}) \rangle \}$,

It is obvious that an neutrosophic set B corresponding to the best record, i.e. the record satisfying perfectly all requirements of the query, would have a following form

$$B = \{ \langle A_1, 1, 0, 0 \rangle, \dots, \langle A_n, 1, 0, 0 \rangle \},$$

while a neutrosophic set W corresponding to the worst record, i.e. the record that does not satisfy any requirements of the query, would look like :

$$W = \{ \langle A_1, 0, 1, 1 \rangle, \dots, \langle A_n, 0, 1, 1 \rangle \}.$$

We will apply neutrosophic sets B and W in our method of calculating matching degrees. They would simply constitute the upper

each record, i.e. $Z(r_i) = [z_{i1}, \dots, z_{in}]$ where z_{ij} is a value of the attribute A_j for the record r_i . To construct a NS-query, a suitable neutrosophic set must be defined for each attribute used in WHERE clause. Thus, actually, our NS-query is an operator T which transforms each attribute A_j to

$$A_j^T = \{ \langle \mu_j^T(x), \nu_j^T(x), \sigma_j^T(x) \rangle : x \in X_j \}$$

where $\mu_j^T, \nu_j^T, \sigma_j^T : X_j \rightarrow [0,1]$ are the membership, indeterminacy and non-membership function of the defined by the neutrosophic term T for the attribute A_j , respectively.

As soon as we accept vague terms in queries we also have to modify our meaning of matching between the query and a record of database. It would be unreasonable to require the answer for a NS-query to be completely precise, adhering to the classical yes-no or non logic.

Now we expect the system to produce a list of records matching a query to a degree higher than a specified threes hold and to list the records according to the linear semi ordering. However, in our approach utilizing neutrosophic sets we do not have such natural linear ordering, because we have to look on three functions $\mu_j^T, \nu_j^T, \sigma_j^T$. Therefore, we will construct a desired semi ordering using distances mentioned in Sec.3.

hence the upper horizon and the lower horizon, respectively. Hence $d(R_i, B)$ and $d(R_i, W)$ denote the Hamming or

Euclidean distance of the neutrosophic set R_i from the upper and lower horizon, respectively. These two numbers show how close is the record r_i to the best and to the worst possible record, respectively. Of course, while querying database we are looking for records with possibly low $d(., B)$ and possibly high $d(., W)$ Therefore, let us define

$\bar{S}_i = 1 - d(R_i, B)$, $\underline{S}_i = d(R_i, W)$. It is clear that a desired record should have both values \bar{S}_i and \underline{S}_i as high as possible. An easy computation shows that for the Hamming distance we obtain:

$$\bar{S}_i = 1 - d_{Ns}(R_i, B) = 1 - \sum_{i=1}^n (|\mu_{R_i}(x_i) - \mu_B(x_i)| + |\nu_{R_i}(x_i) - \nu_B(x_i)| + |\gamma_{R_i}(x_i) - \gamma_B(x_i)|) \quad (1)$$

$$\underline{S}_i = d(R_i, W) = \sum_{i=1}^n (|\mu_{R_i}(x_i) - \mu_W(x_i)| + |\nu_{R_i}(x_i) - \nu_W(x_i)| + |\gamma_{R_i}(x_i) - \gamma_W(x_i)|) \quad (2)$$

Similarly, we can consider the Euclidean distances $e_{Ns}(R_i, B)$ and $e_{Ns}(R_i, W)$ and corresponding values

$$\bar{S}_i = 1 - e_{Ns}(R_i, B) = 1 - \sqrt{\sum_{i=1}^n ((\mu_{R_i}(x_i) - \mu_B(x_i))^2 + (\nu_{R_i}(x_i) - \nu_B(x_i))^2 + (\gamma_{R_i}(x_i) - \gamma_B(x_i))^2)} \quad (3)$$

$$\underline{S}_i = e_{Ns}(R_i, W) = \sqrt{\sum_{i=1}^n ((\mu_{R_i}(x_i) - \mu_W(x_i))^2 + (\nu_{R_i}(x_i) - \nu_W(x_i))^2 + (\gamma_{R_i}(x_i) - \gamma_W(x_i))^2)} \quad (4)$$

Now the question is how to apply (1), (2), (3) and (4) in matching degrees computation. We suggest here three basic methods for determining matching degrees. Namely, we can calculate the matching degree for the i -th record either as an

average of \bar{S}_i and \underline{S}_i , i.e. $S_i^{AV} = \frac{\bar{S}_i + \underline{S}_i}{2}$, or as a

maximum of these two values $S_i^{MAX} = \max(\bar{S}_i, \underline{S}_i)$,

or as the minimum $S_i^{MIN} = \min(\bar{S}_i, \underline{S}_i)$. It is easily

seen that $\bar{S}_i \leq \underline{S}_i$. Thus we get $S_i^{MAX} = S_i^{MIN}$. Hence

using S_i^{MIN} we restrict our consideration to the distance

from the record which fits best, while using S_i^{MAX} we consider the distance from the worst possibility only. Thus

S_i^{MIN} gives us an optimistic matching degree, S_i^{MAX} a

pessimistic one and S_i^{AV} is a balanced one. We can also

consider a natural family of operators for matching degree

computation. Suppose $q \in [0,1]$ is a constant that

characterizes the subjective weight attributed to the distance

from the upper and the lower horizon. Then, for given q , let us

define the matching degree for the record i -th as follows

$S_i^q = q\bar{S}_i + (1-q)\underline{S}_i$. One can see easily that this

operators discussed above are particular members of the

family $\{S_i^q : q \in [0,1]\}$. Namely, $S_i^{AV} = S_i^{0.5}$,

$S_i^{MIN} = S_i^1$ and $S_i^{MAX} = S_i^0$. Whatever method for

calculating matching degrees (note it briefly as S_i) we choose, this method induces a semi ordering on a set of

records. Hence we may say that a record r_i precedes record

r_j (or is – in some sense – better) if and only if the matching

degree S_i is not smaller than S_j , i.e. $r_i > r_j \Leftrightarrow S_i \geq S_j$.

Of course, this semi ordering strongly depends on the method used for calculating matching degree. We expect the system to

reject the records with matching degree lower than a specified

threshold. Therefore we reject the i -th record if $S_i \leq \xi$,

where ξ is a fixed number from the interval $[0;1]$. Hence we

obtain a following algorithm of querying via neutrosophic values:

1. Take the record from the database.
2. Calculate S_i .
3. Accept the record if $S_i \geq \xi$ (where $\xi \in [0,1]$), otherwise reject.
4. If there are more records go to Step 1, otherwise go to Step 5.
5. List all accepted records from the 'best' to the 'worst'

according to $r_i > r_j \Leftrightarrow S_i \geq S_j$.

5. CONCLUSION

In the present paper we have shown how to enrich fuzzy querying by the use of neutrosophic values. Since a condition in the clause WHERE may involve not only imprecise values but also such linguistic terms as fuzzy relations, and linguistic quantifiers, some other generalizations seem natural. In further work we would try to apply neutrosophic sets for modeling relations and in defining quantifiers too. However, we believe that even limited, our method enables the user to

construct queries in a more flexible way. Some of the properties of the neutrosophic sets, Distance measures and Hesitancy Degree, These measures can be used effectively in image processing and pattern recognition. The future work will cover the application of these measures.

6. ACKNOWLEDGMENTS

I gratefully acknowledge the support and generosity of Benha university, faculty of computers and informatics <http://www.bu.edu.eg/>

7. REFERENCES

- [1] K. Atanassov (1986), Intuitionistic fuzzy sets, *Fuzzy Sets and Systems* 20, 87–96.
- [2] K. Atanassov (1999), *Intuitionistic Fuzzy Sets: Theory and Applications*, Physica -Verlag.
- [3] K. Atanassov, Review and new result on intuitionistic fuzzy sets, preprint IM-MFAIS-1-88, Sofia, 1988.
- [4] S. A. Alblowi, A. A. Salama and Mohamed Eisa, (2014), New Concepts of Neutrosophic Sets, *International Journal of Mathematics and Computer Applications Research (IJMCAR)*, Vol. 4, Issue 1, 59-66, 2014.
- [5] P. Bosc, J. Kacprzyk, Eds. (1995), *Fuzziness in Database Management Systems*, Physica-Verlag, Heidelberg.
- [6] P. Grzegorzewski (1998), Metrics and orders in space of fuzzy numbers, *Fuzzy Sets and Systems* 97, 83-94.
- [7] P. Grzegorzewski (2002), Distances between intuitionistic fuzzy sets based on the Hausdorff metric, submitted to *Fuzzy Sets and Systems*.
- [8] P. Grzegorzewski, E. Mrówka (2002), Soft querying via intuitionistic fuzzy sets, *Proceedings of the 9th International conference on Information Processing and management of Uncertainty in Knowledge-Based Systems IMPU'2002*, Annecy, July 1-5, 2002, pp. 1845 – 1849.
- [9] J. Kacprzyk, S. Zadrozny (2000), On combining intelligent querying and data mining using fuzzy logic concepts, In: Bordogna G., Pasi G. (Eds.): *Recent Issues on Fuzzy Databases*, Physica-Verlag, Heidelberg, pp. 67–81.
- [10] I. M. Hanafy, A.A. Salama and K. Mahfouz, (2013), "Correlation Coefficient of Neutrosophic Sets by Centroid Method" *International Journal of Probability and Statistics*, 2(1), pp 9-12.
- [11] I. M. Hanafy, A. A. Salama and K. Mahfouz (2012), Correlation of neutrosophic Data, *International Refereed Journal of Engineering and Science (IRJES)* , Vol.(1), Issue 2 , PP.39-43.
- [12] A. A. Salama and S.A. Alblowi, (2012), " Neutrosophic Set and Neutrosophic Topological Spaces" *ISOR J. mathematics (IOSR-JM)*, Vol. (3). Issue (4), (Sep-Oct. 2012). pp 31-35.
- [13] A.A. Salama and S.A. Alblowi, (2012), "Generalized Neutrosophic Set and Generalized Neutrosophic Topological Spaces" *Journal Computer Science and Engineering*, Vol. (2) No. (7). pp98-102.
- [14] A. A. Salama, and H. Elagamy, " Neutrosophic Filters" *International Journal of Computer Science Engineering and Information Technology Research (IJCEITR)*, Vol.3, Issue (1), Mar 2013, pp 307-312, 2013.
- [15] F. Smarandach (2003), " Neutrosophic set a generalization of the intuitionistic fuzzy sets" , *Proceedings of the third conference of the European Society for fuzzy logic and Technology*, EUSFLAT, Septemper Zittau Geamany Univ. of Applied Sciences at Zittau Goerlit 2, 141-146.
- [16] F. Smarandach (2013), INTRODUCTION TONEUTROSOPHIC MEASURE, NEUTROSOPHIC INTEGRAL, AND NEUTROSOPHIC PROBABILITY <http://fs.gallup.unm.edu/eBooks-otherformats.htm> EAN: 9781599732534
- [17] Eulalia Szmidi and Janus Koeprzyk (2000), Distance between intuitionistic fuzzy sets , *Fuzzy Sets and Systems* 114(2000) 505-518.
- [18] A. A. Salama, "Neutrosophic Crisp Points & Neutrosophic Crisp Ideals", *Neutrosophic Sets and Systems*, Vol.1, No. 1,(2013) pp 50-54.
- [19] A. A. Salama and F. Smarandache, "Filters via Neutrosophic Crisp Sets", *Neutrosophic Sets and Systems*, Vol.1, No. 1,(2013) pp34-38.
- [20] A.A. Salama, " The Concept of Neutrosophic Set and Basic Properties of Neutrosophic Set Operations" *WASET 2012 PARIS, FRANC*, International University of Science, Engineering and Technology, 2012
- [21] A. A. Salama, F. Smarandache and Valeri Kroumov " Neutrosophic Crisp Sets & Neutrosophic Crisp Topological Spaces", *Neutrosophic Sets and Systems* , Vol.(2) pp25-30, 2014.
- [22] A. A. Salama, F. Smarandache and S. A. Alblowi " The Characteristic Function of a neutrosophic Set " *Neutrosophic Sets and System*, Accepted (2014) .
- [23] A. A. Salama, Mohamed Eisa and M. M . Abdelmoghny, "Neutrosophic Relations Database" *International Journal of Information Science and Intelligent System*, 3(1) (2014).
- [24] Supriya Kumar De Ranjit Biswas and Akhil Ranjan Roy, *Intuitionistic Fuzzy Database Second Int. Conf. on IFS*, Sofia, 3-4 Oct. 1998 NIFS 42, 34-41, 1998.
- [25] L.A. Zadeh, *Fuzzy Sets*, *Inform and Control* 8338-353, 1965