On Fuzzy Similarity Relations for Heterogeneous Fuzzy Sets

Leonid Hulianytskyi and Iryna Riasna

V.M. Glushkov Institute of Cybernetics of National Academy of Sciences of Ukraine, Academician Glushkov Avenue, 40, Kyiv, 03187, Ukraine

Abstract

The fundamental problem of development of formal models and methods of ill-structured problems of combinatorial optimization under uncertainty requires utilizing of fuzzy concepts of informal models and different types of scales for measurement of quality and quantity characteristics. In the paper, we introduce a concept of a fuzzy similarity scale. An inadequacy of traditional building a fuzzy similarity scale based on operators of fuzzy logic is shown. A concept of a linguistic correlation coefficient is offered, and conditions of its adequacy in scales of measurement of empirical objects properties on Stevens' classification are derived. For determining fuzzy similarity relations on heterogeneous fuzzy sets, we use the concept of the linguistic correlation coefficient.

Keywords 1[1]

Fuzzy relation, similarity, fuzzy similarity measure, fuzzy set, fuzzy logical operator, fuzzy measurement scale, empirical system with relations, mathematical system with relations.

1. Introduction

Issues of formalization of ill-structured problems under quantitative and qualitative information arise in various areas of human activity. Often, in problems under uncertainty we obtain the information about data from verbal expert assessments, conclusions, judgments, and generalizations. Developing adequate mathematical models and methods for solving this type of problems and approaches providing a meaningful interpretation of their modeling results are relevant; see [1]. The examples are classification and clustering, other ones reducible to combinatorial optimization problems. To formalize such ill-structured problems, methods of representative measurement theory and fuzzy sets theory are normally used. The paper considers two ways of building a fuzzy similarity measure on homogeneous and heterogeneous fuzzy sets under uncertainty in measurements of membership functions values caused by not statistical in nature and subjective expert assessments.

In clustering and classification, much attention is devoted to the identification of fuzzy similarities and differences, as well as to metrics of fuzzy sets; see [2, 3, 4, 5, 6, 7, 8, 9]. In papers [10] and [11], fuzzy logic is used in methodologies where objects under consideration as words and sentences of natural languages – computing with words (CWW). They provide an overview of approaches to constructing similarity measures for interval fuzzy sets of type-1 and type-2. Non-similarity measures on graphs are proposed in [12]. These questions are discussed in [13] for fuzzy intuitionistic sets.

Fuzzy nominal scales are applied in [14] and [15] for generating signals on fuzzy subsets of natural language words. The application of cluster analysis in intelligent systems is considered in [16], thoroughly outlining the application of crisp and fuzzy algorithms. In [9], an overview of approaches to formalization of fuzzy cluster analysis problems is given. There the comprehensive list of methods and algorithms for their solution is also given.

Many scientific studies are devoted to formalizing and solving combinatorial optimization problems on fuzzy sets; see [17].

II International Scientific Symposium «Intelligent Solutions» IntSol-2021, September 28–30, 2021, Kyiv-Uzhhorod, Ukraine
EMAIL: leonhul.icyb@gmail.com (L. Hulianytskyi); riasnaia@gmail.com (I. Riasna)
ORCID: 0000-0002-1379-4132 (L. Hulianytskyi); 0000-0003-1370-3066 (I. Riasna)

The approach to formalization of Euclidean combinatorial configurations dealing with finite sets embedded in Euclid space is developed in [18].

2. Way 1 to define a fuzzy similarity measure

Consider the case where a fuzzy scale defines a fuzzy similarity relation or measure. The scale will further referred to as a fuzzy similarity scale. The problem of building a scale is important, for example, in fuzzy clustering under uncertainty due to a verbal description of properties of empirical objects.

2.1. Basic concepts of representative measurement theory

The subject of measurements is empirical objects properties (features, characteristics) and relations between them. The basic concepts of representative measurement theory are a system with relations and a measurement scale, however, fuzzy sets are not involved in the main works of representative measurement theory [19], [20], and [21].

The study [21] presents three main problems of the theory, i.e., presenting the results of measuring properties or relations in numerical form, the problem of uniqueness of measurement results, and the one of adequacy (meaningfulness [19], [22]).

A subset of the Cartesian product $A \times A$ is the binary relation on a set *A*, and a subset of the Cartesian product or degree of the set A^k is the k -ary relation.

Definition 1 A tuple $\mathrm{M} = \langle A; R_1, ..., R_n \rangle$ is called a system with relations, where *A* is a non-empty set called a domain (support) of a system with relations, and $R_1, ..., R_n$ are the relations given on A.

Let R_i is a k_i - ary relation given on A, $i = 1, ..., n$.

Definition 2 An ordered sequence of positive integers $\langle k_i \rangle_{i=1}^{n}$ we call the type of a system M with relations.

Specifying the system type singles out significant relations in the domain of defining a given system and is determining by the structure of data or its theoretical-set model. We call two systems with relations are similar if they possesses of the same type.

A k -ary relation R is characterized by set of ordered collections $(a_1, ..., a_k) \in A^k$ such that

$$R = \{(a_1, ..., a_k) \in A^k \mid R(a_1, ..., a_k) = 1\}.$$

If a set *A* consists of empirical objects and relations on A empirically defined, then we call the system an empirical system with relations (ESR). It is not important how these empirical relations are determined in practice. They can be defined by physical objects (for example, consecutive connection of measuring apparatuses), or by human answers stated verbally or derived from qualitative and quantitative expert conclusions.

If $A \subseteq R^1$, where R^1 is the set of real numbers, then the system M is called a numerical system with relations (NSR) [20]. If A is a set of non-numeric mathematical objects (for example, symbols, vectors, functions), then M is called a mathematical system with relations (MSR) [22].

Let two similar systems are given:

an ESR $\mathrm{M}_1 = \langle E; S_1, ..., S_n \rangle$, where $E = \{e_j\}_j$ is a set of ESR objects (elements), $S_1, ..., S_n$ are relations given on *E*; and $\mathrm{M}_2 = \langle A; R_1, ..., R_n \rangle$ is a NSR (or MSR), where $A = \{a_j\}_j$ is a set of NSR elements (or MSR), $R_1, ..., R_n$ are relations given on A.

Definition 3 A mapping $f : E \to A$ is called a scale of measurements (in short, a scale) if f is homomorphism, that is,

$$S_i\left(e_1,...,e_j,...,e_{k_i}\right) = R_i\left(a_1,...,a_j,...,a_{k_i}\right),$$

where $\left(e_1,...,e_j,...,e_{k_i}\right) \in E,$ $\left(a_1,...,a_j,...,a_{k_i}\right) \in A,$ $a_j = f\left(e_j\right)$ are the scale values, and $i \in \{1,...,n\}$.

If the mapping f is bijection, then f is isomorphism.

Let us introduce the definition of a "fuzzy scale" of measurements [23]. Let U be a non-empty set of empirical objects, S_i $(i = 1,...,n)$ is a set of relations on U, L is a subset of real numbers, T is a set of fuzzy subsets on L, R_i ($i = 1,...,n$) are k_i-ary relations.

Definition 4 A mapping $\mu : U \to T$ is called a fuzzy scale, if $\forall i \in \{1,...,n\}$ and $\forall\left(u_1,...,u_{k_i}\right) \in U^{k_i}$,

$$S_i\left(u_1,...,u_{k_i}\right) = R_i\left(\mu(u_1),...,\mu\left(u_{k_i}\right)\right),$$

where $\mu(u_1),...,\mu\left(u_{k_i}\right)$ are the corresponding scale values from T.

The questions about permissible transformations, adequate and invariant functions and other relevant issues are not investigated yet [23] for fuzzy scales.

2.2. Using fuzzy operators for building a fuzzy similarity scale

If there are several properties of empirical objects, a formal ESR model, i.e., an MSR, for which measurements of properties of empirical objects represented by a membership function of a fuzzy set, is determined, using concept of "fuzzy measurement scale". The scale we define as a homomorphic mapping of an ESR onto MSR if relations in the MSR we define on a set of fuzzy subsets. Defining the scale in representative measurement starts with building an ESR model.

Let X be a finite set of objects (elements) of the empirical system and W be a finite set of fuzzy properties of elements defined verbally. The result of measurements of values of membership functions for a set of properties $W = \{w_1, ..., w_n\}$, where n is the number of properties, can be a homogeneous or heterogeneous fuzzy set. Membership function is defined as follows: $\mu_W : X \to L_1 \times ... \times L_n$, and L_i is some lattice, i.e., $\mu_W(x) = (\mu_1(x),...,\mu_n(x))$, $\mu_i : X \to L_i$, and $i = \{1,...,n\}$. Based on heterogeneous fuzzy sets, it is possible to build models under different types of properties, which are measured by both quantitative and qualitative scales.

Let the property (for convenience, an index i will be omitted below) takes a finite set $T_w = \left\{t_1,...,t_{m(w)}\right\}$ of verbal values (features, gradations, linguistic terms), where $m(w)$ is the number of these values. If values $\left\{t_1,...,t_{m(w)}\right\}$ of property $w \in W$ are measured on a certain scale, the result of the property measurement is a homogeneous fuzzy subset of T_w. If, in verbal measurements, different scale types or scales of same type are applied, while the valid transformations are independent, the measurements results $\left\{t_1,...,t_{m(w)}\right\}$ of membership function for any $w \in W$ is a heterogeneous fuzzy set. Let for any $w \in W$ be defined $\tilde{R} : X \times T_w \to [0,1]$ – a fuzzy binary relation on $X \times T_w$, then $\mu_{\tilde{R}}(x,t)$ be a membership function defining this fuzzy relation and $\mu_{\tilde{R}}(x,t) \in [0,1]$, $x \in X$, $t \in T_w$.

Denote by F_{T_w} a set of all fuzzy subsets of T_w and let $\eta : X \to F_{T_w}$ be a mapping such that for it $\forall t_j \in T_w \; \forall x \in X \; \eta_x(t_j) = \mu_{\tilde{R}}(x, t_j)$ is met. Here and beyond, in similar cases, $j = 1, ..., m(w)$.

Let for any $x \in X$ the set $\tilde{D}_x = \left\{ \left(t_j, \eta_x\left(t_j\right) \right) \mid t_j \in T_w \right\}$ be a fuzzy subset from T_w, that we call a measurement description of $w \in W$. The meaning of $\eta_x(t_j)$ will be considered as degree of truth of statement "x has the meaning t_j of a fuzzy property w". Let $\forall x \in X \;\; \tilde{D}_x$ be a homogeneous fuzzy set.

Let $\Omega_{\tilde{D}} = \left\{ \tilde{D}_x \mid x \in X \right\}$ be a set of fuzzy measurement descriptions. We consider $\Omega_{\tilde{D}}$ as a vector-valued fuzzy set [24] or as a type-2 fuzzy set:

$$\Omega_{\tilde{D}} = \left\{ \left(x, \vec{\eta}(x) \right) \middle| x \in X \right\} = \left\{ \left(x, \left(\eta_x(t_1), ..., \eta_x\left(t_{m(w)}\right) \right) \right) \middle| x \in X \right\},$$

where $\eta_x\left(t_j\right) \in [0,1]$, $j = 1, ..., m(w)$, and $m(w)$ is the number of property values.

For any $x \in X$, measurements of a certain property $w \in W$ are not a real numbers, but are a fuzzy subset of set of values from the corresponding T_w, to which the studying of problems of representation, uniqueness and adequacy is carried out taking into account the structure of fuzzy measurement scale. However, in this case, we change known definition of fuzzy scale: we define fuzzy subsets on a set of verbal values $T_w = \left(t_1, ..., t_{m(w)} \right)$ of a certain property w, and not on a subset of real numbers, as is certain in [23].

A fuzzy similarity measure on $X \times X$ is a function $\tilde{\tau}(x, y)$, satisfying following conditions:

1) $\forall (x, y) \in X \times X \;\; \tilde{\tau}(x, y) \in [0,1]$;
2) $\forall x \in X \;\; \tilde{\tau}(x, x) = 1$ (reflexivity);
3) $\forall (x, y) \in X \times X \;\; \tilde{\tau}(x, y) = \tilde{\tau}(y, x)$ (symmetry).

A function $\tilde{\tau}(x, y)$ defines a fuzzy set in sense of L. Zadeh. A fuzzy similarity measure specifies on the finite set X a fuzzy similarity relation $R_{\tilde{\tau}} : X \times X \to [0,1]$.

A fuzzy similarity measure $\tilde{\tau}(x, y)$ should be determined taking into account the results of measuring for all values of fuzzy property $w \in W$ for any $x \in X$, that is, by aggregating fuzzy logical operators.

A method of calculating $\tilde{\tau}(x, y)$ values is determined both by type of scale of measuring values of $w \in W$ for objects of an empirical system, and by theoretical-set interpretation of semantic link AND and OR in form of fuzzy logical operators.

A fuzzy analogue of semantic link AND is triangular norm and fuzzy analogue of OR is an operator of triangular conorm; see [24]. In fuzzy sets theory, there are many variants for choosing of norm and conorm operators.

We use fuzzy operators of Zadeh norms and conorms:

$$T(a,b) = \min(a,b), \; S(a,b) = \max(a,b), \; a, b \in [0,1],$$

and fuzzy operators of norms and conorms (Ia. Lukasiewicz)

$$T(a,b) = \max(0, a + b - 1), \; S(a,b) = \min(1, a + b), \; a, b \in [0,1].$$

Let $\tilde{\tau}(\cdot,\cdot)$ be a fuzzy similarity measure defined on $X \times X$ and $\tilde{\tau}^*(\cdot,\cdot)$ be a fuzzy similarity measure defined on $F_{T_w} \times F_{T_w}$. The Homomorphic mapping $\eta : X \to F_{T_w}$ for $w \in W$ determines a fuzzy measurement scale. However, the existence of a scale, i.e., a homomorphism of η mapping we must prove [20].

Denote by F_X a set of all subsets of *X*. Let $\delta : T_w \to F_X$ be the mapping such that $\forall x \in X \;\; \forall t_j \in T_w \;\; \delta_{t_j}(x) = \mu_{\tilde{R}}(x, t_j)$ and let $\tilde{E}_j = \left\{ \left(x,\ \delta_{t_j}(x) \right) \mid x \in X \right\}$ be a fuzzy subset of *X*, called a meaning of $t_j \in T_w$.

Obviously, a condition for similarity $(\tilde{\tau}(x, y) > 0)$ of $x, y \in X$ is an existence of at least one common property, i.e.:

$$\tilde{\tau}(x, y) > 0 \Leftrightarrow \left(\exists t_j \in T_w \right) : \left(x \in Supp\, \tilde{E}_j \right) \wedge \left(y \in Supp\, \tilde{E}_j \right). \quad (1)$$

Since $\delta_{t_j}(x) = \eta_x(t_j)$, then it is from (1) follows that intersection of supports, i.e., $Supp\ \tilde{D}_x \cap Supp\ \tilde{D}_y \neq \varnothing$ is also a condition of similarity of empirical objects. If $Supp\ \tilde{D}_x \cap Supp\ \tilde{D}_y = \varnothing$ case is valid, then $\tilde{\tau}(x, y) = 0$ we obtain. Thus,

$$\tilde{\tau}(x, y) > 0 \Leftrightarrow \left(\exists t_j \in T_w \right) : \left(\eta_x\left(t_j\right) > 0 \right) \wedge \left(\eta_y\left(t_j\right) > 0 \right). \quad (2)$$

Considering (2), we define $\varphi(\tilde{D}_x, \tilde{D}_y)$ as follows:

$$\varphi\left(\tilde{D}_x, \tilde{D}_y\right) = \left(\left(\eta_x(t_1) > 0 \right) \wedge \left(\eta_y(t_1) > 0 \right) \right) \vee \ldots \vee \left(\left(\eta_x(t_m) > 0 \right) \wedge \left(\eta_y(t_m) > 0 \right) \right). \quad (3)$$

Replacing in (3) inequalities with corresponding values of membership function, and logical operators with triangular norms T and conorms *S*, we get:

$$\tilde{\varphi}\left(\tilde{D}_x, \tilde{D}_y\right) = \underset{t_j \in T_w}{S} \left(T\left(\eta_x\left(t_j\right), \eta_y\left(t_j\right) \right) \right),\ \tilde{D}_x, \tilde{D}_y \in \Omega_{\tilde{D}}. \quad (4)$$

Obviously, $\tilde{\varphi}(\tilde{D}_x, \tilde{D}_y) \in [0,1]$. Symmetry $\tilde{\varphi}(\tilde{D}_x, \tilde{D}_y)$ follows from the symmetry of operators of norm and conorm operators.

If (4) satisfies reflexivity condition, i.e., if the condition, $\tilde{\varphi}(\tilde{D}_x, \tilde{D}_y) = 1$ is met, then we define a fuzzy similarity measure on $F_{T_w} \times F_{T_w}$ as

$$\tilde{\tau}^*\left(\tilde{D}_x, \tilde{D}_y\right) = \tilde{\varphi}\left(\tilde{D}_x, \tilde{D}_y\right).$$

Therefore,

$$\tilde{\tau}^*\left(\tilde{D}_x, \tilde{D}_y\right) = \underset{t_j \in T_w}{S} \left(T\left(\eta_x\left(t_j\right), \eta_y\left(t_j\right) \right) \right),\ \tilde{D}_x, \tilde{D}_y \in \Omega_{\tilde{D}}. \quad (5)$$

Such a method of defining a fuzzy similarity measure, which is an analogue of method of calculating a similarity measure for crisp properties based on direct replacement of logical operators with their fuzzy analogues, we call as "direct" method.

We suppose, that $\tilde{\tau}(x, y) = \tilde{\tau}^*\left(\tilde{D}_x, \tilde{D}_y\right)$ is met. If reflexivity condition $\tilde{\tau}^*\left(\tilde{D}_x, \tilde{D}_x\right)=1$ is met, then formula (5) defines a fuzzy similarity measure.

Obviously, mapping $\eta : X \to \Omega_{\tilde{D}}$ gives rise to an equivalence relation θ, i.e.:

$$x\theta y \Leftrightarrow \tilde{D}_x = \tilde{D}_y,$$

that, under the reflexivity of $\tilde{\tau}^*\left(\tilde{D}_x, \tilde{D}_y\right)$ is a congruence relation:

$$\left(x\theta x' \right) \wedge \left(y\theta y' \right) \Rightarrow \tilde{\tau}(x, y) = \tilde{\tau}(x', y').$$

Thus, we define two systems with relations: an ESR, $\mathrm{M}_1 = \langle X, \theta, \tilde{\tau} \rangle$ and a MSR, $\mathrm{M}_2 = \langle \Omega_{\tilde{D}}, =, \tilde{\tau}^* \rangle$. A tuple $\langle \mathrm{M}_1, \mathrm{M}_2, \eta \rangle$ we call a fuzzy scale for measuring a partial fuzzy similarity measure (further – a fuzzy similarity scale).

2.3. Representation theorem

We formulate a representation theorem determining conditions for existence of measurement scale of a fuzzy similarity measure (5).

When proving representation theorem, we believe that an absolute measurement scale we use, while only identical transformations [20] are permissible for it.

Theorem 1 A reflexivity condition for fuzzy similarity measure (5) is in generally not satisfied.

Proof. If we use in (5) Zadeh norm and conorm operators, we get

$$\tilde{\tau}^*\left(\tilde{D}_x, \tilde{D}_y\right) = \max_{t_j \in T_w}\left\{\min\left\{\eta_x\left(t_j\right), \eta_y\left(t_j\right)\right\}\right\},\ \tilde{D}_x, \tilde{D}_y \in \Omega_{\tilde{D}}\ .$$

Because of idempotency of $\min$ operator, we get

$$\tilde{\tau}^*\left(\tilde{D}_x, \tilde{D}_x\right) = \max_{t_j \in T_w}\left\{\eta_x\left(t_j\right)\right\}.$$

In turn,

$$\tilde{\tau}^*\left(\tilde{D}_x, \tilde{D}_x\right) = \max_{t_j \in T_w}\left\{\eta_x\left(t_j\right)\right\} = 1$$

will be the reflexivity condition of fuzzy similarity measure.

That is, we can write

$$\forall x \in X\ \exists t_j : \eta_x(t_j) = 1. \tag{6}$$

Therefore, measurement results must consist of normal [24] fuzzy sets. In general, this condition may not meet.

The Lukasevich norm operator cannot used in (5), since in case where the sum of values of membership functions is less than one, the value of this norm is zero, then

$$\tilde{\tau}^*(\tilde{D}_x, \tilde{D}_x) = 0,$$

and the reflexivity condition is not fulfilled. If in (5) we use Zadeh norm and Lukasevich conorm, then we get

$$\tilde{\tau}^*(\tilde{D}_x, \tilde{D}_x) = \min\left\{1, \sum_{t_j \in T_w} \eta_x(t_j)\right\}.$$

In this case, the reflexivity condition is as follows:

$$\sum_{t_j \in T_w} \eta_x(t_j) \geq 1\ \forall x \in X, \tag{7}$$

which in practice would not always be fulfilled.

Theorem proved.

2.4. Investigating of adequacy of fuzzy similarity measure

Let us study an adequacy of fuzzy similarity measure.

Theorem 2 Fuzzy similarity measure (5) is not adequate while measuring values of membership function of a fuzzy property $w \in W$ in ratio, interval, and order scales.

Proof. Truthiness of theorem falls out of absence of invariance of conditions (6) and (7) for measurements in ratio and interval scales.

Condition (7) we cannot use in order scale because sum operation is invalid. It is obvious that condition (6) violates in monotonic transformations permissible for order scale.

Theorem proved.

Thus, fuzzy similarity measure in "direct" method of determining (5), used, for example, in [25], is generally unsuitable for constructing an adequate formal model, while component of it is a fuzzy similarity relation.

3. Way 2 to define a fuzzy similarity measure

Let X be a finite set of empirical system objects (elements), $W = \{w_1, ..., w_n\}$ is a finite set of fuzzy properties of X defined verbally. Let property $w_i \in W$ takes a finite set of verbal values $T_{w_i} = \{t_1^i, ..., t_{m(w_i)}^i\}$, where $m(w_i)$ is the number of values of $w_i \in W$.

Definition 5 A linguistic correlation coefficient (LCC, K_{lingv}) and a partial linguistic correlation coefficient (k_i) we call, respectively,

$$K_{\text{lingv}}(x, y) = \frac{1}{n}\sum_{i=1}^{n} k_i(x, y), \tag{8}$$

$$k_i(x, y) = \tilde{\tau}_i^*\left(\tilde{D}_x^i, \tilde{D}_y^i\right) = \left(\sum_{j=1}^{m(w_i)} \min\left(\eta_x(t_j^i), \eta_y(t_j^i)\right) \Big/ \sum_{j=1}^{m(w_i)} \max\left(\eta_x(t_j^i), \eta_y(t_j^i)\right)\right), \tag{9}$$

where $\tilde{\tau}_i^*\left(\tilde{D}_x^i, \tilde{D}_y^i\right)$ is a partial similarity measure in MSR of $w_i \in W$, and $\tilde{D}_x^i$, $\tilde{D}_y^i$ are measurements of $w_i \in W$ for elements $x, y \in X$, respectively, n is the total number of properties; $\eta_x(t_j^i)$ determines a measure of belonging of value t_j^i to w_i of $x \in X$, $j \in \{1, ..., m(w_i)\}$, $i \in \{1, ..., n\}$.

Therefore, a partial LCC determines a value of a partial similarity measure on a set of empirical objects. Obviously, when measuring values of membership functions in an absolute scale, we get,

$$k_i(x, x) = 1,\ k_i(x, y) = k_i(y, x);\ K_{\text{lingv}}(x, x) = 1,\ K_{\text{lingv}}(x, y) = K_{\text{lingv}}(y, x);$$

$$k_i(x, y) \in [0,1],\ K_{\text{lingv}}(x, y) \in [0,1].$$

That is, according to formulas (8) and (9), a fuzzy similarity measure: $\tilde{\tau}(x, y) = K_{\text{lingv}}(x, y)$ on $X \times X$ is determined. Unlike "direct" method of determining of similarity measure (5) in the case of measurements of values of membership functions on an absolute scale, in order to ensure reflexivity condition of fuzzy similarity measure with help of LCC, no restrictions on value of membership functions of fuzzy objects properties are necessary. Since min and max operators are valid for order, interval, ratio, and absolute scale [23], such theorems are valid.

Theorem 3 Partial fuzzy similarity measure of (9) is invariant when measuring values of membership function of a fuzzy qualitative property $w_i \in W$ in ratio scales.

Proof. A valid transformation of membership function values in ratio scale is a similarity transformation, i.e., $y = \alpha x$, where $0 < \alpha < 1$. It is not hard to see that

$$k_i(x, y) = \tilde{\tau}_i^*\left(\tilde{D}_x^i, \tilde{D}_y^i\right) = \left(\sum_{j=1}^{m(w_i)} \min\left(\eta_x(t_j^i), \eta_y(t_j^i)\right) \Big/ \sum_{j=1}^{m(w_i)} \max\left(\eta_x(t_j^i), \eta_y(t_j^i)\right)\right) =$$

$$= \left(\sum_{j=1}^{m(w_i)} \min\left(\alpha\eta_x(t_j^i), \alpha\eta_y(t_j^i)\right) \Big/ \sum_{j=1}^{m(w_i)} \max\left(\alpha\eta_x(t_j^i), \alpha\eta_y(t_j^i)\right)\right).$$

Theorem proved.

Theorem 4 When measuring values of membership functions of a fuzzy qualitative property in an order scale and an interval one, there exist a permissible monotonic transformation $\psi\left(\eta_x(t_j^i)\right)$,

$t_j^i \in T_{w_i}$, leading to invariance of values of membership function of a partial fuzzy similarity measure (9), $i \in \{1,...,n\}$.

Proof. For a set of empirical objects $X = \{x_1,...,x_N\}$, result t_j^i of measurements of term values of property w_i we represent as a sequence $A = \left(\eta_{x_1}(t_j^i),...,\eta_{x_k}(t_j^i),...,\eta_{x_N}(t_j^i)\right)$. For simplicity, we denote by $a_k = \eta_{x_k}(t_j^i)$ and $a_k \in [0,1]$. If we order the elements of sequence in ascending order, then the collection $A' = \left(a_{k_1} \le ... \le a_{k_r} \le ... \le a_{k_N}\right)$ is a ranked sequence, where the number $1 \le r \le N$ we call the rank of $a_k \in A$. If there are no identical elements in sequence *A*, then $a_{k_r} \ne a_{k_{r+1}}$; if there are equal elements in sequence *A* ($a_{k_p} = a_{k_{p+1}} = ... = a_{k_{p+m}}$), then rank value in the interval $[p, p+m]$ is $(p+...+(p+m))/(m+1)$ (fractional ranks). In both cases, the sum of ranks is equal to $N(N+1)/2$. Denote by $r(a_k)$, $r(a_k) \in [1, N]$, rank value of a_k. In order scale, it is allowing a monotone transformation of φ that does not change ratio relation, i.e., $a_k = a_m \Rightarrow \varphi(a_k) = \varphi(a_m)$, $a_k > a_m \Rightarrow \varphi(a_k) > \varphi(a_m)$. Then obviously, in such transformations, rank values of $r(a_k)$ in the sequence A' do not change. We define transformation $\psi(a_k) = r(a_k)/N$; because $r(a_k) \in [1, N]$, then $0 < \psi(a_k) \le 1$. Value $\psi(a_k)$ does not change under any valid monotone φ transformation. Therefore, membership function value of partial fuzzy similarity measure

$$k_i(x,y) = \left(\sum_{j=1}^{m(w_i)} \min\left(\psi\left(\eta_x(t_j^i)\right), \psi\left(\eta_y(t_j^i)\right)\right) \Big/ \sum_{j=1}^{m(w_i)} \max\left(\psi\left(\eta_x(t_j^i)\right), \psi\left(\eta_y(t_j^i)\right)\right) \right)$$

is invariant under ψ transformation. This transformation is also valid in interval scales.

Theorem proved.

Invariance of partial similarity measures leads to invariance of LCC, using which as a similarity measure provides an adequacy of formal model of ESR as example in fuzzy clustering.

4. Examples of LCC Calculation

Example 1 Let us give a method of calculating of LCC for the case where elements of a MSR are fuzzy combinatorial configurations (objects) of 1st order of the first type [26]

$$\underset{\sim}{K}_{(1)}^1 = \left\langle Y, \varphi, \underset{\sim}{X}_{(1)}, \Omega \right\rangle,$$

where a base set $\underset{\sim}{X}_{(1)}$ coincides with fuzzy generating set $\underset{\sim}{\mathbf{Z}} = (\underset{\sim}{z}_1,...,\underset{\sim}{z}_n)$, $\underset{\sim}{X}_{(1)} \equiv \underset{\sim}{\mathbf{Z}}$, $\varphi: Y \to \underset{\sim}{X}_{(1)}$, and $Y = \{1,..., p\}$. Let be $p < n$, and $\varphi: Y \to \underset{\sim}{X}_{(1)}$ – some crisp mapping. Then ordered fuzzy set $\underset{\sim}{A} = (\varphi(1),...,\varphi(l),...,\varphi(p)) = (a_1,...,a_l,...,a_p)$ is an arrangement with the repetitions of fuzzy elements of $\underset{\sim}{\mathbf{Z}}$ where $\varphi(l) = a_l$ and $a_l = \underset{\sim}{z}_i \in \underset{\sim}{\mathbf{Z}}$.

Let $\underset{\sim}{A} = (a_1,...,a_p)$ and $\underset{\sim}{B} = (b_1,...,b_p)$ be two arrangements. A LCC (K_{lingv}) and a partial LCC (k_i) by the property $w_i \in W$ we calculate as follows:

$$K_{\text{lingv}}(\underset{\sim}{A}, \underset{\sim}{B}) = \frac{1}{n}\sum_{i=1}^{n} k_i(\underset{\sim}{A}, \underset{\sim}{B}), \; k_i(\underset{\sim}{A}, \underset{\sim}{B}) = \frac{1}{p}\sum_{l=1}^{p} \tau_i^*\left(\tilde{D}_{a_l}^i, \tilde{D}_{b_l}^i\right),$$

where $\tilde{\tau}_i^*\left(\tilde{D}_{a_l}^i, \tilde{D}_{b_l}^i\right) = \sum_{j=1}^{m(w_i)} \min\left(\eta_{a_l}(t_j^i), \eta_{b_l}(t_j^i)\right) \Big/ \sum_{j=1}^{m(w_i)} \max\left(\eta_{a_l}(t_j^i), \eta_{b_l}(t_j^i)\right)$ is partial similarity measure of a_l, b_l for $w_i \in W$; $T_{w_i} = \left\{t_1^i, ..., t_{m(w_i)}^i\right\}$ is a set of verbal values of $w_i \in W$, and $\tilde{D}_{a_l}^i$, $\tilde{D}_{b_l}^i$ are measurements of $w_i \in W$ for elements $a_l \in \underset{\sim}{A}$, $b_l \in \underset{\sim}{B}$, respectively, n is the total number of properties, and $\eta_{a_l}(t_j^i)$, $\eta_{b_l}(t_j^i)$ determine the measure of membership value t_j^i of w_i for elements $a_l \in \underset{\sim}{A}$, $b_l \in \underset{\sim}{B}$, respectively, $\eta_{a_l}(t_j^i) \in [0,1]$, $\eta_{b_l}(t_j^i) \in [0,1]$, $j \in \{1, ..., m(w_i)\}$, and $i \in \{1, ..., n\}$. Obviously, when measuring values $\eta_{a_l}(t_j^i)$ and $\eta_{b_l}(t_j^i)$ of membership functions in an absolute scale then $k_i(\underset{\sim}{A}, \underset{\sim}{A}) = 1$, $k_i(\underset{\sim}{A}, \underset{\sim}{B}) = k_i(\underset{\sim}{B}, \underset{\sim}{A})$; $K_{\text{lingv}}(\underset{\sim}{A}, \underset{\sim}{A}) = 1$, $K_{\text{lingv}}(\underset{\sim}{A}, \underset{\sim}{B}) = K_{\text{lingv}}(\underset{\sim}{B}, \underset{\sim}{A})$; and $k_i(\underset{\sim}{A}, \underset{\sim}{B}) \in [0,1]$, $K_{\text{lingv}}(\underset{\sim}{A}, \underset{\sim}{B}) \in [0,1]$, that is $K_{\text{lingv}}(\underset{\sim}{A}, \underset{\sim}{B})$ determines a fuzzy similarity in an ESR. Proof of invariance of a $K_{\text{lingv}}(\underset{\sim}{A}, \underset{\sim}{B})$ under permissible transformations in applied measurement scales is the same with proof of theorems 3 and 4.

Example 2 Let the manager of the enterprise personnel department be face with the problem of selecting several of the best representatives from the candidates for vacant positions of engineers of the department producing equipment, for example, for a new model of the aircraft.

To determine the best applicants, it is necessary to evaluate their competitiveness according to the totality of two fuzzy qualitative characteristics, i.e., (w_1, w_2), where w_1 – the work experience, w_2 – the level of qualification. Characteristics w_1, w_2 are fuzzy multidimensional expert assessments measured in order scales.

We believe that the description of the work experience has two gradations, i.e., $w_1 = \left(t_1^1,\, t_2^1\right)$, where t_1^1 – the work experience according to the profile of the enterprise, t_2^1 – the work experience on a computer with programmable logic matrices. We assume that the qualification level has three gradations, i.e., $w_2 = \left(t_1^2,\, t_2^2,\, t_3^2\right)$, where by t_1^2 denote knowledge of foreign languages, by t_2^2 denote knowledge of a specify set of programming languages, and by t_3^2 denote duration of work in leading companies in the industry.

Let m ($m = 2$) be the number of vacancies, $X = \{x_i\}_{i=1}^p$ be the set of applicants, and p ($p = 4$) be the number of applicants.

Let the manager give the following gradations of t_1^1, t_2^1 for characteristic w_1 for applicants: $x_1 - (0,6;0,9)$, $x_2 - (0,5;0,4)$, $x_3 - (0,3;0,7)$, $x_4 - (0,4;0,5)$. The ideal applicant has the maximum value of gradations of characteristics, i.e., $x_{id} - (1;1)$. The variation series of manager assessments of gradation of t_1^1, t_2^1 for the characteristic w_1 are respectively:

$$\left(t_1^1(x_1), t_1^1(x_2), t_1^1(x_3), t_1^1(x_4), t_1^1(x_{id})\right) = (0,6;0,5;0,3;0,4;1),$$

$$\left(t_2^1(x_1), t_2^1(x_2), t_2^1(x_3), t_2^1(x_4), t_2^1(x_{id})\right) = (0,9;0,4;0,7;0,5;1).$$

Obviously, the corresponding rank vector of the assessments of t_1^1 (ordered in ascending order) for the sequence $(x_1, x_2, x_3, x_4, x_{id})$ has the form: $\left(r_1^1(x_1), r_1^1(x_2), r_1^1(x_3), r_1^1(x_4), r_1^1(x_{id})\right)$

$=(4;3;1;2;5)$. Similarly, the rank vector of the assessments of t_2^1 is $\left(r_2^1(x_1), r_2^1(x_2), r_2^1(x_3), r_2^1(x_4), r_2^1(x_{id})\right) = (4;1;3;2;5)$.

Since rank values are invariant to permissible (monotonic) transformations in the order scale, the value of the membership functions of fuzzy characteristics w_1 and w_2 is calculated by dividing the rank assessments by a maximum rank value of 5.

Thus, we obtain the values of the membership functions to the assessments t_1^1 and t_2^1 for the ordered totality $(x_1, x_2, x_3, x_4, x_{id})$, respectively,

$$\left(\eta_{x_1}(t_1^1), \eta_{x_2}(t_1^1), \eta_{x_3}(t_1^1), \eta_{x_4}(t_1^1), \eta_{x_{id}}(t_1^1)\right) = (0,8;0,6;0,2;0,4;1),$$

$$\left(\eta_{x_1}(t_2^1), \eta_{x_2}(t_2^1), \eta_{x_3}(t_2^1), \eta_{x_4}(t_2^1), \eta_{x_{id}}(t_2^1)\right) = (0,8;0,2;0,6;0,4;1).$$

Let the manager give the following gradations of t_1^2, t_2^2, t_3^2 for w_2 characteristic for applicants: $x_1 - (0,8;0,7;0,3)$, $x_2 - (0,3;0,6;0,8)$, $x_3 - (0,5;0,9;0,2)$, and $x_4 - (0,6;0,8;0,6)$.

The ideal applicant has the maximum value of gradations of w_2 characteristics, i.e., $x_{id} - (1;1;1)$. The variation series of gradation manager assessments of t_1^2, t_2^2, t_3^2 for w_2 have the form:

$$\left(t_1^2(x_1), t_1^2(x_2), t_1^2(x_3), t_1^2(x_4), t_1^2(x_{id})\right) = (0,8;0,3;0,5;0,6;1),$$

$$\left(t_2^2(x_1), t_2^2(x_2), t_2^2(x_3), t_2^2(x_4), t_2^2(x_{id})\right) = (0,7;0,6;0,9;0,8;1),$$

$$\left(t_3^2(x_1), t_3^2(x_2), t_3^2(x_3), t_3^2(x_4), t_3^2(x_{id})\right) = (0,3;0,8;0,2;0,6;1).$$

It is not difficult to show that the corresponding rank vector of assessments of t_1^2 for the sequence $(x_1, x_2, x_3, x_4, x_{id})$ is $\left(r_1^2(x_1), r_1^2(x_2), r_1^2(x_3), r_1^2(x_4), r_1^2(x_{id})\right) = (4;1;2;3;5)$.

Similarly, the rank assessment vector for t_2^2 is $\left(r_2^2(x_1), r_2^2(x_2), r_2^2(x_3), r_2^2(x_4), r_2^2(x_{id})\right)$ $=(2;1;4;3;5)$, and $\left(r_3^2(x_1), r_3^2(x_2), r_3^2(x_3), r_3^2(x_4), r_3^2(x_{id})\right) = (2;4;1;3;5)$ is the ranking vector for t_3^2.

The value of the membership functions of fuzzy gradations of t_1^2, t_2^2, t_3^2 we calculate by dividing the rank assessments by the maximum rank value (5).

Thus, we get the values of the membership functions of the assessments for t_1^2, t_2^2, and t_3^2 of ordered sequence $(x_1, x_2, x_3, x_4, x_{id})$:

$$\left(\eta_{x_1}(t_1^2), \eta_{x_2}(t_1^2), \eta_{x_3}(t_1^2), \eta_{x_4}(t_1^2), \eta_{x_{id}}(t_1^2)\right) = (0,8;0,2;0,4;0,6;1),$$

$$\left(\eta_{x_1}(t_2^2), \eta_{x_2}(t_2^2), \eta_{x_3}(t_2^2), \eta_{x_4}(t_2^2), \eta_{x_{id}}(t_2^2)\right) = (0,4;0,2;0,8;0,6;1),$$

$$\left(\eta_{x_1}(t_3^2), \eta_{x_2}(t_3^2), \eta_{x_3}(t_3^2), \eta_{x_4}(t_3^2), \eta_{x_{id}}(t_3^2)\right) = (0,4;0,8;0,2;0,6;1).$$

We use $K_{\text{lingv}}(x_p, x_{id})$ to evaluate the measure of similarity of the applicant $x_p \in X$ $(p \in \{1,2,3,4\})$ with the ideal element x_{id} calculated as follows

$$K_{\text{lingv}}(x_p, x_{id}) = \frac{1}{n}\sum_{i=1}^{n} k_i(x_p, x_{id}),\ k_i\left(x_p, x_{id}\right) = \tilde{\tau}_i^*\left(\tilde{D}_{x_p}^i, \tilde{D}_{x_{id}}^i\right),$$

$$\tilde{\tau}_i^*\left(\tilde{D}_{x_p}^i, \tilde{D}_{x_{id}}^i\right) = \left(\sum_{j=1}^{m(w_i)} \min\left(\eta_{x_p}(t_j^i), \eta_{x_{id}}(t_j^i)\right) \Big/ \sum_{j=1}^{m(w_i)} \max\left(\eta_{x_p}(t_j^i), \eta_{x_{id}}(t_j^i)\right)\right),$$

where $k_i(x_p, x_{id})$ is a partial LCC which defines the value of the similarity measure on a set X of empirical objects by property $w_i \in W$; $\tilde{\tau}_i^*\left(\tilde{D}_{x_p}, \tilde{D}_{x_{id}}\right)$ is a partial similarity measure in MSR by $w_i \in W$, $\tilde{D}_{x_p}^i$ and $\tilde{D}_{x_{id}}^i$ are measurements of $w_i \in W$ for elements x_p and x_{id}, respectively, n is the total number of properties; $\eta_{x_p}(t_j^i)$ and $\eta_{x_{id}}(t_j^i)$ determine a measure of belonging of value t_j^i to w_i, respectively, of x_p, x_{id}; $m(w_1) = 2$, $m(w_2) = 3$.

After simple calculations according to the above formulas, we get
$K_{\text{lingv}}(x_1, x_{id}) = 0,67$; $K_{\text{lingv}}(x_2, x_{id}) = 0,40$; $K_{\text{lingv}}(x_3, x_{id}) = 0,44$; $K_{\text{lingv}}(x_4, x_{id}) = 0,50$.

According to calculations,

$$K_{\text{lingv}}(x_1, x_{id}) > K_{\text{lingv}}(x_4, x_{id}) > K_{\text{lingv}}(x_3, x_{id}) > K_{\text{lingv}}(x_2, x_{id}).$$

Thus, LCC assessments of the competitiveness of applicants, which use expert assessments of multidimensional fuzzy characteristics measured in order scales, show that one of the two vacancies should be offered to the applicant x_1, and the second vacancy should be offered to the applicant x_4.

5. Conclusions

The paper concerns the features of combinatorial optimization problems on fuzzy sets under multidimensional qualitative and quantitative information. Uncertainty caused by measurements of membership functions values of fuzzy sets in different scales leads to accounting the base problems studied in representative measurement theory, i.e., a presentation problem, uniqueness one and an adequacy. We offer a concept of fuzzy similarity scale. An inadequacy of traditional building a fuzzy similarity scale based on fuzzy logic operators we prove considering the representative measurement theory positions. A concept of a linguistic correlation coefficient is introduced. Conditions of its adequacy in different scales of measuring of empirical objects properties, i.e., order, ratio, intervals, and absolute scales (according to Stevens' Classification) are derived. As further step of research, based on linguistic correlation coefficient, a fuzzy binary relation of difference on homogeneous and heterogeneous fuzzy sets will be determined.

6. References

[1] L. F. Hulianytskyi, I. I. Riasna, Automatic classification method based on a fuzzy similarity relation, Cybernetics and Systems Analysis 52.1 (2016) 30–37. doi:10.1007/s10559-016-9796-3.

[2] V. Balopoulos, A. G. Hatzimichailidis, B. K. Papadoupoulos, Difference and similarity measures for fuzzy operators, Information Sciences (2007) 2336–2348. doi:10.1016/j.ins.2007.01.005.

[3] B. B. Chaudur, A. Rosenfeld, On a metric distance between fuzzy sets, Pattern Recognition Letters 17.11 (1996) 1157–1160. doi:10.1016/0167-8655(96)00077-3.

[4] Andrew Gardner, et al., Measuring distance between unordered sets of different sizes, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2014, pp. 137–143. doi:10.1109/CVPR.2014.25.

[5] S. Kosub, A note on the triangle inequality for the Jaccard distance, Pattern Recognition Letters (2019) 36–38. URL: https://doi.org/10.1016/j.patrec.2018.12.007.
[6] J. S. Williams, N. Steele, Difference, distance and similarity as a basis for fuzzy decision support based on prototypical decision classes, Fuzzy Sets and Systems (2002) 35–46. URL: https//doi.org/10.1016/S0165-0114(01)00253-6.
[7] A. Bogiatzis, B. Papadopoulos, Producing fuzzy inclusion and entropy measures and their application on global image thresholding, Computer science. Evolving Systems (2018) 331–353. doi:10.1007/s12530-017-9200-1.
[8] A. Bogiatzis, B. Papadopoulos, Global image thresholding adaptive neuro-fuzzy inference system trained with fuzzy inclusion and entropy measures, Symmetry 11.2 (2019) 286. URL: https://doi.org/10.3390/sym11020286.
[9] P. D'Urso, M. A. Gil, Fuzzy data analysis and classification, Adv. Data Anal. Classif. (2017) 645–657. URL: https://doi.org/10.1007/s11634-017-0304-z.
[10] D. Wu, J. M. Mendel, A comparative study of ranking methods, similarity measures and uncertainty measures for interval type-2 fuzzy sets, Information Sciences 179.8 (2009) 1169–1192. doi:10.1016/j.ins.2008.12.010.
[11] D. Wu, J. M. Mendel, A vector similarity measure for linguistic approximation: Interval type-2 and type-1 fuzzy sets, Information Sciences 178.2 (2008) 381–402. doi:10.1007/978-3-540-72950-1_57.
[12] A. V. Kuznetsov, Nonsimilarity measure on graph set and its application, Vestnik VGU, Seria: System analysis and Information Technology (2017) 125–131. URL: http://www.vestnik.vsu.ru/pdf/analiz/2017/01/2017-01-17.pdf. (In Russian).
[13] E. Szmidt, J. Kacprzyk, Distances between intuitionistic fuzzy sets, Fuzzy Sets and Systems 114 (2000) 505–518. doi:10.1016/S0165-0114(98)00244-9.
[14] T. Allevard, E. Benoit, L. Foulloy, The transportation distance for fuzzy descriptions of measurement, Metrol. Meas. Syst. XIV.1 (2007) 25–27. URL: https://hal.archives-ouvertes.fr/hal-00581046/document.
[15] E. Benoit, L. Foulloy, Towards fuzzy nominal scales, Measurement (2003) 49–55. URL: https://doi.org/10.1016/S0263-2241(03)00019-8.
[16] M. Z. Zgurovsky, Y. P. Zaychenko, The Fundamentals of Computational Intelligence: System Approach, Springer, Switzerland, 2016.
[17] P. M. Pardalos, et al, Fuzzy combinatorial optimization problems, Handbook of Combinatorial Optimization, Springer New York, 2013, pp. 1357–1413. URL: https://doi.org/10.1007/978-1-4419-7997-1_68.
[18] O. S. Pichugina, S. V. Yakovlev, Continuous functional representations in combinatorial optimization problems, Zolotaya milya, Kharkov, 2018. (In Russian).
[19] R. Luce, D. Krantz, P. Suppes, and A. Tversky, Vol. III: Foundations of Measurement, Dover Publications, Inc., New York, 2007.
[20] I. Pfanzagl, Measurement Theory, Mir, Moscow, 1976. (In Russian).
[21] P. Suppes, J. Zinnes, The fundamentals of measurement theory. Psychological measurements, Mir, Moscow, 1967, pp. 9–110. URL: http://www.libex.ru/detail/book525723.html. (In Russian).
[22] Yu. N. Tolstova, Measurement in sociology: course of lectures, Infra-M, Moscow, 1998. URL: https://www.isras.ru/publ.html?id=6544. (In Russian).
[23] A. F. Blishun, Comparative analysis of fuzzy measurement methods, Izv. of Academy of Sciences of the USSR. Technical cybernetics, volume 5, 1988, pp. 152–175. (In Russian).
[24] A. Kaufmann, Fuzzy sets theory, Radio and svyaz, Moscow, 1982. (In Russian).
[25] S. V. Ovchinnikov, T. R'era, Fuzzy classifications, in: R. R. Yager (Ed.), Fuzzy sets and possibility theory. Recent developments, Radio and svyaz, Moscow, 1986, pp. 100–113. (In Russian).
[26] L. F. Hulianytskyi, I. I. Riasna, Formalization and Classification of Combinatorial Optimization Problems, in: S. Butenko, P. M. Pardalos, V. Shylo (eds.), Optimization Methods and Applications, Cham, Springer International Publishing AG, 2017, pp. 239–250. doi:10.1007/978-3-319-68640-0_11.