IBA-based framework for modeling similarity

Pavle Milošević[1], Ana Poledica[1], Aleksandar Rakićević[1], Vladimir Dobrić[1], Bratislav Petrović[1], Dragan Radojević[2]

[1] Faculty of Organizational Sciences, University of Belgrade,
Jove Ilića 154
11000 Belgrade, Serbia
E-mail: {pavle.milosevic, ana.poledica, aleksandar.rakicevic, vdobric, bratislav.petrovic}@fon.bg.ac.rs

[2] Mihajlo Pupin Institute,
Volgina 15
11000 Belgrade, Serbia
E-mail: dragan.radojevic@automatika.imp.bg.ac.rs

Abstract

In this paper, we introduce a logic-driven framework for modeling similarity based on interpolative Boolean algebra (IBA). It consists of two main steps: data preprocessing and similarity measuring by means of IBA similarity measure and logical aggregation. The purpose of these steps is to detect dependencies and model interactions among attributes and/or similarities using an appropriate operator. The proposed framework is general, providing different approaches to multi-attribute object comparison: attribute-by-attribute comparison, object-level comparison and their combination. It is also a generic framework since various similarity measures can be easily derived. The proposed IBA-based similarity framework has a solid mathematical background, which ensures all necessary properties of similarity measure are satisfied. It is interpretable and close to human perception. The framework's applicability is illustrated by two numerical examples that confirm the need for a different level of aggregations. Furthermore, the example of similarity-based classification demonstrates the descriptive power and transparency of the framework on real financial data.

Keywords: Multi-attribute object comparison, IBA similarity measure, Logical aggregation, Modeling similarity, Similarity-based classification.

1. Introduction

Similarity plays a fundamental role in human experience and learning. It refers to some degree of closeness between two physical objects or ideas/concepts.[1] When dealing with the problem of comparing multi-attributes objects, similarity is usually measured by distance functions in metric spaces (e.g. Euclidean distance), statistical measures based on correlation (e.g. cosine similarity) and logic-based measures (e.g. Refs. 2, 3), along with chosen aggregation function. Many logical formalizations have been developed in order to model vagueness, uncertainty and truthlikeness such as many-valued logics (including fuzzy logic), probabilistic logic, etc.

As the notion of similarity is graded in its nature, these logic-based approaches are natural formalizations for similarity-based reasoning.[4] On the other hand, logic-based aggregators are not commonly used in similarity measuring, although they are able to model both linear and non-linear relations among attributes. Dealing with correlation and logical dependences among attributes prior to measuring similarities may be crucial for appropriate similarity modeling. Therefore, it is particularly important to consider the idea of the similarity framework that is entirely based on logic.

In Refs. 5-8, the extensive overview and comparison of similarity measures among fuzzy concepts have been provided. For measuring similarity, the authors[3] recognized a group of distance measures specified for

membership functions of fuzzy sets (e.g. Refs. 5, 9), set-theoretic approach (e.g. Refs. 10, 11), the approach based on logical concepts of fuzzy implication (e.g. Refs. 12, 13). In terms of logic, possible formalizations of similarity are both implication and equivalence relations: fuzzy equivalence (e.g. Refs. 14, 15), fuzzy implication (e.g. Ref. 16), IBA equivalence[17], etc.

Along with the choice of similarity measure, an appropriate aggregation operator is crucial for the comparison of multi-attribute objects. The most common aggregation operators are related to the means (e.g. arithmetic mean, weighted means, order weighted aggregation (OWA) operators, etc.)[18], integrals (e.g. discrete Choqeut[19], generalized Choquet[20], etc.), t-norms (e.g. *min*, product, etc.)[21]. In case not all similarities are equally important, weighting techniques are used, e.g. weighted Euclidean distance[22]. Combining various similarity measures (Jaccard index, cosine, proximity or correlation similarity measure) with the *max-min* composition or weighted sum as aggregation operators was proposed in Ref. 23 for content-based recommender systems. Enhancing generalized Lukasiewicz structure using weights was proposed and evaluated on several well-known medical classification problems.[24] In Ref. 25, the authors introduced a similarity measure based on probabilistic equivalence relations and generalized mean. Other mean-related aggregation functions, such as Bonferreni mean[26], are also used for defining similarities. Similarity measures obtained by OWA operators are improved in the sense that they can model particular types of interactions of similarities. OWA operators were used to aggregate similarities obtained by different measures into a resulting value.[27] A family of order weighted distance operators (OWAD) was introduced by combining OWA operator and Hamming distance.[28] OWAD operators were further extended using other distance functions such as Euclidean and Minkowski.[29,30] Choquet integral is also considered to be a promising tool for modeling various interactions of attributes/similarities, while t-norms and t-conorms have the property of compensation in marginal cases when realized as *min* and *max* functions. In Ref. 31, the authors used the discrete Choquet integral as a basis for defining a metric which is applied in semi-supervised clustering. Furthermore, a similarity measure of intuitionistic fuzzy sets was built using Sugeno integral.[32] In Ref. 33, the similarities are aggregated by combining a conjunction of chosen attributes with weighted sum. Generalization of geometrical approach to measuring similarity by including interaction of attributes was proposed in Ref. 34. Some of the most distinguished similarity measures and aggregation operators for ordinary fuzzy sets and their extensions (including intuitionistic, type-2, interval-valued, hesitant fuzzy sets, etc.) are given in Ref. 35.

In this paper, we present a novel logic-driven framework for modeling similarity between multi-attribute objects based on interpolative Boolean algebra (IBA). It is based on two main components: IBA similarity measure[17] and logical aggregation (LA)[36]. IBA similarity measure is based on logic which makes it easily interpretable and close to human perception of similarity, which is a desirable property for clustering[37], classification[38], etc. Further, the proposed framework uses LA capable of modeling any logical (non-linear) interaction among attributes which is not the case with the previously mentioned approaches. Moreover, the proposed framework allows comparison of objects on both attribute-by-attribute and the object-level. The usual way to compare objects is attribute-by-attribute, when all attributes are compared separately and afterwards aggregated into a resulting value. However, there are occasions when it is necessary to consider object's attributes jointly in order to capture the meaning of the object. On these occasions, one should use suggested similarities defined on the object-level, where aggregation of attributes using LA is realized prior to measuring similarities. The proposed framework is flexible, allowing the decision maker to account for different perspectives of modeling similarity. From the mathematical point of view, this approach is generic since LA is a generalization of simple mean, weighted sum, OWA, Choquet integral, etc. Further, we illustrate the applicability of the proposed approach with two simple numerical examples. In the first illustrative example, we demonstrate the importance and benefits of the object-level comparison. In the second example, the proposed framework is applied for similarity-based classification with k-Nearest Neighbors (k-NN) algorithm on a real financial dataset.

This paper is organized as follows. In Section 2, we outline basic definitions and provide a brief overview of IBA. The framework for modeling similarity based on IBA similarity measure and logical aggregation is introduced in Section 3. Further, the applicability of the proposed framework is illustrated in Section 4. In the final section, we summarize the main findings and give guidelines for future work.

2. Theoretical background

2.1. Mathematical prerequisites

The notion and some basic mathematical prerequisites concerning similarity measure and an aggregation operator are given in this section.

Definition 1.[1] *Let X be a set. A function $S : X^2 \to R$ is called similarity on X if, for all $x, y \in X$, there holds:*
- *$S(x, y) \geq 0$ (non-negativity);*
- *$S(x, y) = S(y, x)$ (symmetry);*
- *$S(x, y) \leq S(x, x)$ with equality iff $x = y$ (limited range).*

Problems of modeling similarity and distance are dual concepts. From the standpoint of logic, similarities are the negation of distances. In other words, if the distance function D satisfies properties of non-negativity, symmetry and reflexivity on $[0,1]$ interval, the similarity function may be derived as $S = f(D)$, where f is some fuzzy negation.[1]

Definition 2.[39] *A function $A : [0,1]^2 \to [0,1]$ is called aggregation operator if, for all $x_i, y_i \in [0,1]$, $i = 1, ..., m$, there holds:*
- *$A(0, ..., 0) = 0$ (lower boundary condition);*
- *$A(1, ..., 1) = 1$ (upper boundary condition);*
- *$A(x_1, ..., x_m) \leq A(y_1, ..., y_m)$ when $x_i \leq y_i$, $i = 1, ..., m$ (monotonicity).*

2.2. Interpolative Boolean algebra

Interpolative Boolean algebra (IBA) is a Boolean consistent real-valued realization of finite Boolean algebra (BA).[40] It is a two-level logic that consists of symbolic and valued level.
On the symbolic level, any element of BA:

$$\varphi \mapsto \varphi^\otimes, \ \varphi(x_1, ..., x_m) = \varphi^\otimes(x_1, ..., x_m) \quad (1)$$

where φ is a logical function of attributes $x_1, ..., x_m$, that is mapped to the generalized Boolean polynomial (GBP) $\varphi^\otimes$.
The transformation procedure of logical functions to corresponding GBPs is introduced in Ref. 36. Operators in GBP are standard $+$, standard $-$ and generalized product $\otimes$ (GP), and its variables are observed attributes.

Definition 3.[36] *A function $\otimes : [0,1]^2 \to [0,1]$ is called generalized product if, for all attributes $x, y, z \in [0,1]$, there holds:*
- *$x \otimes y = y \otimes x$ (commutativity);*
- *$x \otimes (y \otimes z) = (x \otimes y) \otimes z$ (associativity);*
- *$x \otimes y \leq x \otimes z$ when $y \leq z$ (monotonicity);*

- *$x \otimes 1 = x$ (neutral element/boundary condition);*
- $\sum_{K \in P(\Omega \setminus S)} (-1)^{|K|} \underset{x_i \in S \cup K}{\otimes} x_i \geq 0$ *(non-negativity condition),*

where $\Omega = \{x_1, ..., x_m\}$, $P(\Omega)$ is a partition of Ω and $S \in P(\Omega)$ are elements of a partition included in a logical function.

The most important property of GP on the symbolic level is idempotency. In fact, it is one of IBA transformation rules[36]:

$$x \otimes x = x \quad (2)$$

Once the structural transformations have been conducted, the valued level is introduced. On the valued level, GP is a subclass of *t*-norms since it should satisfy the additional non-negativity condition. According to the defined properties, GP may be realized as any *t*-norm that produces the result from the following interval[36]:

$$\max(x + y - 1, 0) \leq x \otimes y \leq \min(x, y) \quad (3)$$

In accordance with the nature of the attributes to be aggregated, we may distinguish three realizations of GP. Attributes of the same nature should be aggregated using the *min* function. The product is a proper operator in the case of independent/not correlated attributes. GP is realized as Lukasiewicz *t*-norm if attributes are negatively correlated attributes. Attributes of the same nature should be aggregated first, followed by negatively correlated attributes if more attributes are to be aggregated using GP.
Although *t*-norms and realizations of generalized products are formally similar, their roles are crucially different: while a *t*-norm in conventional fuzzy approaches has the role of a logical (and/or algebraic) operator, GP is only an arithmetic operator on a value level without any influence on algebra.[36] Algebra is wholly determined by the structural transformation procedure. Therefore, GP is idempotent on the symbolic level, regardless of its realization on the valued level.

2.2.1. IBA similarity measure

From the standpoint of logic, the similarity of two attributes is commonly perceived as logical relations of equivalence (bi-implication). Thus, the relation of equivalence in the IBA framework is defined as similarity measure.[17] GBP of IBA similarity measure and its realization on the valued level are:

$$S_{IBA}(x,y) = (x \Leftrightarrow y)^{\otimes} =$$
$$= 1 - x - y + 2 \cdot x \otimes y = \qquad (4)$$
$$= 1 - x - y + 2 \cdot \min(x,y)$$

Theorem 1. *Let* $S_{IBA} : [0,1]^2 \rightarrow [0,1]$ *be the relation of equivalence in the sense of IBA. For any* $x, y \in [0,1]$, $S_{IBA}(x,y)$ *is a similarity measure.*

Proof. The proof can be found in Ref. 17. $\qquad \square$

IBA similarity on the valued level is mathematically identical to Lukasiewicz bi-implication $S_{LUK}(x,y) = 1 - |x - y|$. However, IBA equivalence is valuable from the aspect of interpretation and the possibility to be combined with other IBA-based tools such as logical aggregation.

IBA similarity measure is realized as the union of the intersection of attributes ("both having the trait") and the intersection of their complements ("both not having the trait").[17] Therefore, this measure treats all values from the unit interval equally, which is very important for many applications. Also, IBA similarity has a clear-cut meaning and a comprehensible graphical interpretation. The proposed measure was applied for modeling consensus in group decision making[41], formalizing the human categorization process[38], logical clustering[42], etc. In Ref. 17, it is shown that IBA similarity overcomes the drawbacks of classical logic-based similarity measures (*min* and product bi-implication) on the problem of modeling consensus.

2.2.2. Logical aggregation

For the purpose of aggregating information in IBA framework, Radojevic introduced logical aggregation (LA). LA is a transparent, a Boolean consistent procedure for aggregating attributes in a resulting value with significant meaning.[36] It consists of two steps:

(i) Normalization of attributes' values to the unit interval;
(ii) Aggregation of normalized values into the resulting value by means of a logical/pseudo-logical function.

If the aggregation function is a logical expression, attributes are aggregated using corresponding GBP. In the general case, the aggregation function may be

pseudo-logical, i.e. the linear convex combination of GBP.

Definition 4.[36] *A logical aggregation operator* $LA : [0,1]^m \rightarrow [0,1]$ *in general case is a pseudo-logical function, i.e. a linear convex combination of generalized Boolean polynomials*

$$\underset{i=1}{\overset{m}{LA}}(x_i) = \sum_{j=1}^{m} w_j \cdot \varphi_j^{\otimes}(x_1,...,x_m)$$

where $\sum_{j=1}^{m} w_j = 1$, $w_j \geq 0$, $j = 1,...,m$ *and* $\varphi_j^{\otimes}(x_1,...,x_m)$ *are GBPs.*

In this way, it is possible to aggregate partial demands expressed as logical expressions using weighted sum. Thus, arithmetic mean, weighted sum, *min* function, discrete and generalized Choquet integral can be obtained as special cases of LA.[36]

In the general case, logical aggregation is not monotone, e.g. IBA similarity measure. Although violation of monotonicity has its benefits, e.g. a possibility to model various interactions of attributes[43,44], LA does not satisfy all properties to be an aggregation operator. On the other hand, some special cases of LA are monotone. For instance, GP as a subclass of *t*-norms satisfies all conditions for aggregation function.

3. Modeling similarity: IBA-based framework

The aim of this section is to introduce different approaches to modeling similarity using IBA. Relying on IBA similarity measure and logical aggregation, we present several relations for multi-attribute objects. Special attention will be devoted to the mathematical background/properties of the framework, its validity and applicability. Also, we formalize the IBA-based procedure for modeling similarity explained in several steps.

3.1. IBA-based similarity framework

There are two main steps to perform when modeling similarity using the IBA-based framework:

(i) **Data preprocessing** (data normalization, detection of attribute nature and their potential interaction);
(ii) **IBA similarity modeling** (attribute-by-attribute comparison, comparison on the level of the object and general (combined) approach).

3.1.1. Data preprocessing

Normalization of attribute data is a starting point for data analysis and preprocessing. Data normalization to [0,1] interval is a mathematical prerequisite for similarity modeling using IBA. Standard normalization functions used are *min-max* normalization, scaling with maximum value, etc. where the attribute data are scaled so as to fall within a unit interval. On the other hand, normalization functions may be adapted depending on data range and distribution, e.g. so as to be able to transform both positive and negative values to unit interval. Sometimes from the aspect of the decision maker it is appropriate to use a normalization function in the opposite direction, e.g. the lower price the better.

In the next step, it is important to examine the *nature of the attributes* and their potential *interaction*. Generally, the existence of significant correlation in attribute data will overemphasize certain attributes and/or cause incoherent model results.[45] In our framework we make use of correlation to detect similar nature between attribute data. Further, the *min* function is used as GP for these attributes. The idea is also to examine whether some attributes can compensate for others or to model other meaningful logical dependence discovered by the experts.

3.1.2. IBA similarity modeling

Two distinctive cases when modeling similarity of *m*-attribute objects using IBA are presented. The first one is based on a conventional attribute-by-attribute comparison, while in the second comparison is performed on the level of the object. Also, we introduce the general model that includes both approaches to comparison and even combines them. All proposed models are novel except the ABA comparison 1. This model is used in practice[17] although it is not mathematically formalized and its properties are not investigated so far.

Attribute-by-attribute comparison (ABA). In this approach, the conventional ABA comparison using IBA similarity measure is followed by suitable logical aggregation.

$$S_{IBA}^{A}(X,Y) = \overset{m}{\underset{i=1}{LA}}\big(S_{IBA}(x_i,y_i)\big) \quad (5)$$

where $x_i, i=1,...,m$ and $y_i, i=1,...,m$ are attributes of objects X and Y. Bearing in mind that logical

aggregation is not a monotone function in the general case, we also investigate two special cases of this relation.

ABA comparison 1. In the first one, the similarity between two objects is modeled as a conjunction of the attribute similarities. As it is explained in Ref. 17, there are situations in group decision making when it is necessary to detect the guaranteed (minimal) level of similarity between two objects. These situations may be modeled as LA of the attribute similarities, where LA operator is realized as the conjunction. In the sense of IBA, this model corresponds to a generalized product of the attribute similarities.

$$S_{IBA}^{GP}(X,Y) = \overset{m}{\underset{i=1}{\wedge}}\big(S_{IBA}(x_i,y_i)\big) = \\ = \overset{m}{\underset{i=1}{\otimes}}\big(S_{IBA}(x_i,y_i)\big) \quad (6)$$

This function is monotone, since it inherits the properties of a generalized product, i.e. it is a valid aggregation operator.

ABA comparison 2. The second one is the convex linear combination of GPs of similarities. There are situations when a conjunction of similarities should be observed as partial demands that are further aggregated using weighted sum. For instance, in determining similarity of two students, we may be interested in two aspects: 1) similarity of their math grades and 2) similarity of their social science grades modeled as a conjunction of similarities between history grades and similarities between geography grades. This may be modeled as weighted sum of GPs of similarities:

$$S_{IBA}^{WGP}(X,Y) = w_1 \cdot \overset{p}{\underset{i=1}{\otimes}}\big(S_{IBA}(x_i,y_i)\big) + ... + \\ + w_k \cdot \overset{r}{\underset{i=1}{\otimes}}\big(S_{IBA}(x_i,y_i)\big) \quad (7)$$

where $\overset{p}{\underset{i=1}{\otimes}}\big(S_{IBA}(x_i,y_i)\big),...,\overset{r}{\underset{i=1}{\otimes}}\big(S_{IBA}(x_i,y_i)\big)$ are aggregated similarities of $p,...,r$ attributes and $w_1,...,w_k > 0, \sum_{j=1}^{k} w_j = 1$ are weights for k partial demands. The number of partial demands k and the number similarities to be aggregated $p,...,r$ must be less than or equal to the number of observed attributes m. For $p=1$, element $\overset{p}{\underset{i=1}{\otimes}}\big(S_{IBA}(x_i,y_i)\big)$ is a simple

similarity between attributes, while for $p > 1$ it is a compound element that consists of several similarities. This model generalizes similarity aggregation using weighted sum, OWA, *min* function, etc. Since this model sums GPs of similarities, this relation also generalizes the similarity relation given in Eq. (6). Both of them are special cases of Eq. (5) since these aggregations are special cases of LA. Since $k, p, ..., r \leq m$, this similarity relation may be presented in the most general form:

$$S_{IBA}^{WGP}(X,Y) = \sum_{j=1}^{m} w_j \cdot S_j^{GP}(X,Y) \qquad (8)$$

where $\quad S_j^{GP}(X,Y) = \overset{m}{\underset{i=1}{\otimes}}\left(S_{IBA}(x_i, y_i)\right) \quad$ and $\quad w_j > 0$,

$\sum_{j=1}^{m} w_j = 1$ are weights for m partial demands.

Comparison on the object-level (OL). This approach involves comparison on the level of the object. First of all, the object is uniquely represented by a single value obtained using a logical aggregation function. These values are compared afterwards using the IBA similarity measure.

$$S_{IBA}^{O}(X,Y) = S_{IBA}\left(\overset{m}{\underset{i=1}{LA}}(x_i), \overset{m}{\underset{i=1}{LA}}(y_i)\right) \qquad (9)$$

In this case, we first model the different interactions and dependencies between attributes in order to represent the object properly. The nature/statistical dependencies of attributes are modeled by a different realization of GP, while various LA functions are applied to model logical (nonlinear) dependencies. An aggregated outlook on the object may be more informative since comparing attributes separately is misleading and insufficient in some situations. For instance, LA functions may be used to describe situations when a large value of certain attribute compensates for another. The necessity for this approach to similarity modeling will be further elaborated on the examples in Section 4.

General model. The general approach combines Eq. (8) and Eq. (9) in order to incorporate the benefits of both previously presented models. It should be used in occasions when it is necessary to model both interaction of attributes, e.g. to appropriately present the object, and interaction of similarities, e.g. to model partial demands and guaranteed level of similarity:

$$S_{IBA}^{gen}(X,Y) = \sum_{j=1}^{m} w_j \cdot S_j^{O}(X,Y) \qquad (10)$$

where $\quad S_j^{O}(X,Y) = \overset{m}{\underset{i=1}{\otimes}}\left(S_{IBA}^{O_i}(X,Y)\right) \quad$ and $\quad w_j > 0$,

$\sum_{j=1}^{m} w_j = 1$ are weights for m object representations.

The models given in Eqs. (6), (8) and (9) may be seen as special cases of this relation.

3.2. Mathematical properties

ABA comparison. The conventional ABA comparison using IBA similarity measure followed by logical aggregation is not a similarity measure due to characteristics of LA.

Lemma 1. *Let* $S_{IBA}^{A} : X^2 \rightarrow [0,1]$ *be the logical aggregation of IBA equivalence relations. For any objects* $X = [x_1, ..., x_m]$, $Y = [y_1, ..., y_m]$, $x_i, y_i \in [0,1]$, $S_{IBA}^{A}(X,Y)$ *is not a similarity measure in the general case.*

Proof. We need to investigate if relation S_{IBA}^{A} given in Eq. (5) satisfies properties given in Definition 1.

- Non-negativity: For $x_i, y_i \in [0,1]$ IBA similarity measure $S_{IBA}(x_i, y_i) \in [0,1]$, due to the non-negativity proof of Theorem 1 (see Ref. 17). Hereafter, $\overset{m}{\underset{i=1}{LA}}\left(S_{IBA}(x_i, y_i)\right) \geq 0$ due to the definition of logical aggregation;

- Symmetry: $S_{IBA}^{A}(X,Y) = \overset{m}{\underset{i=1}{LA}}\left(1 - x_i - y_i + 2 \cdot x_i \otimes y_i\right) =$

$= \overset{m}{\underset{i=1}{LA}}\left(1 - y_i - x_i + 2 \cdot y_i \otimes x_i\right) = S_{IBA}^{A}(Y,X) \quad$ due to commutativity of GP;

- Limited range is not fulfilled in the general case.
  For instance, we may compare two 2-attribute objects using the following LA function:

$$\overset{2}{\underset{i=1}{LA}}\left(S_{IBA}(x_i, y_i)\right) = S_{IBA}(x_1, y_1) \vee S_{IBA}(x_2, y_2).$$

Its corresponding GPB is:

$S_{IBA}^{A}(X,Y) = 1 - x_1 \otimes x_2 - y_1 \otimes x_2 - x_1 \otimes y_2 - y_1 \otimes y_2 +$

$+ 2 \cdot \left(x_1 \otimes y_1 \otimes (x_2 + y_2)\right) + 2 \cdot \left(x_2 \otimes y_2 \otimes (x_1 + y_1)\right) -$

$- 4 \cdot x_1 \otimes y_1 \otimes x_2 \otimes y_2.$

Assuming that all attributes of one object are independent and bearing in mind that GP between

corresponding attributes of different objects should be realized as the *min* function, the relation is as follows:

$$S_{IBA}^A(X,Y) = 1 - x_1 \cdot x_2 - y_1 \cdot x_2 - x_1 \cdot y_2 - y_1 \cdot y_2 +$$
$$+ 2 \cdot \min(x_1, y_1) \cdot (x_2 + y_2) + 2 \cdot \min(x_2, y_2) \cdot (x_1 + y_1) -$$
$$- 4 \cdot \min(x_1, y_1) \cdot \min(x_2, y_2).$$

In the case when $X = [1, 0.95]$ and $Y = [0, 0.95]$, $S_{IBA}^A(X,Y) = 1$ thereby proving that the limited range is not fulfilled. In other words, for this aggregation function the complete identity of single attribute is mapped in the complete identity of objects. Therefore, it is not a similarity measure. □

However, there are certain situations when this kind of aggregation is appropriate for similarity modeling. For instance, in consensus models, LA realized as a disjunction (*max* function) may be appropriate in the case when it is satisfactory for a high level of consensus that at least two experts have a high level of agreement.[41] In Ref. 4 authors also refer to similar types of non-monotonic inferences, namely pessimistic and optimistic inferences.

The other proposed relations for comparison of multi-attribute objects satisfy all properties of similarity measure.

Theorem 2. *Let $S_{IBA}^{GP} : X^2 \rightarrow [0,1]$ be the generalized product of the IBA equivalence relations. For any objects $X = [x_1, ..., x_m]$, $Y = [y_1, ..., y_m]$, $x_i, y_i \in [0,1]$, $S_{IBA}^{GP}(X,Y)$ is a similarity measure.*

Proof. We need to investigate if relation S_{IBA}^{GP} given in Eq. (6) satisfies properties given in Definition 1.
- Since GP is a special case of LA, non-negativity and symmetry are followed as a consequence of the proofs in Lemma 1.
- Limited range: Due to the limited range property of IBA similarity and 1 is the neutral element for GP

$$S_{IBA}^{GP}(X,X) = \overset{m}{\underset{i=1}{\otimes}} \left(S_{IBA}(x_i, x_i) \right) = \overset{m}{\underset{i=1}{\otimes}} 1 = 1. \text{ If } X \neq Y, \text{ at}$$

least one similarity is $S_{IBA}(x_i, y_i) < 1$. Therefore,

$$S_{IBA}^{GP}(X,Y) = 1 \otimes 1 \otimes ... \otimes S_{IBA}(x_i, y_i) = S_{IBA}(x_i, y_i) < 1,$$

since 1 is the neutral element for GP. □

Theorem 3. *Let $S_{IBA}^{WGP} : X^2 \rightarrow [0,1]$ be the weighted sum of generalized products of the IBA equivalence relations. For any objects $X = [x_1, ..., x_m]$,*

$Y = [y_1, ..., y_m]$, $x_i, y_i \in [0,1]$, $S_{IBA}^{WGP}(X,Y)$ *is a similarity measure.*

Proof. We need to investigate if relation S_{IBA}^{WGP} given in Eq. (8) satisfies properties given in Definition 1.
- Since the weighted sum of GPs is a special case of LA, non-negativity and symmetry are followed as a consequence of the proofs in Lemma 1.
- Limited range: Due to the limited range property of IBA similarity measure and the idempotency of GP

$$S_{IBA}^{WGP}(X,X) = \sum_{j=1}^{k} \left(w_j \cdot \overset{m}{\underset{i=1}{\otimes}} 1 \right) = \sum_{j=1}^{k} w_j = 1. \text{ If } X \neq Y, \text{ at}$$

least one similarity is $S_{IBA}(x_i, y_i) < 1$. Since the weighted sum is a strictly monotone function and 1 is the neutral element for GP, $S_{IBA}^{WGP}(X,Y) < 1$ for $X \neq Y$. □

Comparison on OL. The proposed IBA-based relation for comparison on OL is a similarity measure in the general case.

Theorem 4. *Let $S_{IBA}^O : X^2 \rightarrow [0,1]$ be the relation of equivalence in the sense of IBA and $X = \overset{m}{\underset{i=1}{LA}}(x_i)$ is a logical aggregation of multi-attribute object. For any objects $X = [x_1, ..., x_m]$, $Y = [y_1, ..., y_m]$, $x_i, y_i \in [0,1]$, $S_{IBA}^O(X,Y)$ is a similarity measure.*

Proof. We need to investigate if relation S_{IBA}^O given in Eq. (9) satisfies properties given in Definition 1.
- Non-negativity: Due to the definition of LA, $\overset{m}{\underset{i=1}{LA}}(x_i) \in [0,1]$ and $\overset{m}{\underset{i=1}{LA}}(y_i) \in [0,1]$. Hereafter, $S_{IBA}\left(\overset{m}{\underset{i=1}{LA}}(x_i), \overset{m}{\underset{i=1}{LA}}(y_i) \right) \geq 0$ due to the non-negativity of IBA similarity measure (see Ref. 17);
- Symmetry: Due to commutativity of GP

$$S_{IBA}^O(X,Y) = 1 - \overset{m}{\underset{i=1}{LA}}(x_i) - \overset{m}{\underset{i=1}{LA}}(y_i) + 2 \cdot \overset{m}{\underset{i=1}{LA}}(x_i) \otimes \overset{m}{\underset{i=1}{LA}}(y_i) =$$
$$= 1 - \overset{m}{\underset{i=1}{LA}}(y_i) - \overset{m}{\underset{i=1}{LA}}(x_i) + 2 \cdot \overset{m}{\underset{i=1}{LA}}(y_i) \otimes \overset{m}{\underset{i=1}{LA}}(x_i) = S_{IBA}^O(Y,X)$$

- Limited range: Due to the idempotency of GP (Eq. (2)), $S_{IBA}^O(X,X) = 1 - \overset{m}{\underset{i=1}{LA}}(x_i) - \overset{m}{\underset{i=1}{LA}}(x_i) + 2 \cdot \overset{m}{\underset{i=1}{LA}}(x_i) \otimes \overset{m}{\underset{i=1}{LA}}(x_i) = 1.$

$S_{IBA}^O(X,X) = 1$ iff $\overset{m}{\underset{i=1}{LA}}(x_i) + \overset{m}{\underset{i=1}{LA}}(y_i) = 2 \cdot \overset{m}{\underset{i=1}{LA}}(x_i) \otimes \overset{m}{\underset{i=1}{LA}}(y_i).$ Due to the monotonicity of GP

$\underset{i=1}{\overset{m}{LA}}(x_i) \preceq \underset{i=1}{\overset{m}{LA}}(x_i) \otimes \underset{i=1}{\overset{m}{LA}}(y_i)$, with equality iff $\underset{i=1}{\overset{m}{LA}}(x_i) = \underset{i=1}{\overset{m}{LA}}(y_i)$ (in the sense that $X = Y$) or $Y = [1,...,1]$ (in the sense that 1 is the neutral element for GP), and $\underset{i=1}{\overset{m}{LA}}(y_i) \preceq \underset{i=1}{\overset{m}{LA}}(x_i) \otimes \underset{i=1}{\overset{m}{LA}}(y_i)$ with equality iff $\underset{i=1}{\overset{m}{LA}}(x_i) = \underset{i=1}{\overset{m}{LA}}(y_i)$ or $X = [1,...,1]$. It follows that $\underset{i=1}{\overset{m}{LA}}(x_i) + \underset{i=1}{\overset{m}{LA}}(y_i) = 2 \cdot \underset{i=1}{\overset{m}{LA}}(x_i) \otimes \underset{i=1}{\overset{m}{LA}}(y_i)$ stands if and only if $X = Y$. $\qquad\square$

General model. The general model is obtained by combining ABA similarity and the comparison on OL, and it is a similarity measure in the general case.

Theorem 5. *Let* $S_{IBA}^{gen} : X^2 \rightarrow [0,1]$ *be the weighted sum of generalized products of IBA similarity relation on the object-level and* $X = \underset{i=1}{\overset{m}{LA}}(x_i)$ *is a logical aggregation of multi-attribute object. For any objects* $X = [x_1,...,x_m]$, $Y = [y_1,...,y_m]$, $x_i, y_i \in [0,1]$, $S_{IBA}^{gen}(X,Y)$ *is a similarity measure.*

Proof. We need to show that relation S_{IBA}^{gen} given in Eq. (10) satisfies properties given in Definition 1.
- Non-negativity: IBA similarity measure of logical aggregations $S_{IBA}(X,Y) \geq 0$ due to the non-negativity proof of Theorem 4. Since the non-negativity property of GP and weighted sum $S_{IBA}^{gen}(X,Y) \geq 0$.
- Symmetry: Due to commutativity of GP

$$S_{IBA}^{gen}(X,Y) = \sum_{j=1}^{m}\left(w_j \cdot \underset{i=1}{\overset{m}{\otimes}}\left(S_{IBA}^{O}(X,Y)\right)\right) =$$
$$= \sum_{j=1}^{m}\left(w_j \cdot \underset{i=1}{\overset{m}{\otimes}}\left(S_{IBA}^{O}(Y,X)\right)\right) = S_{IBA}^{gen}(Y,X)$$

- Limited range: Since $S_{IBA}^{O}(X,X) = 1$ and the limited range proof of Theorem 4 $S_{IBA}^{gen}(X,Y) = \sum_{j=1}^{m}\left(w_j \cdot \underset{i=1}{\overset{m}{\otimes}}1\right) = 1$. In a case if $X \neq Y$, then at least one similarity is $S_{IBA}^{O}(X,Y) < 1$. Since the weighted sum is a strictly monotone and 1 is the neutral element for GP (see the proof of Theorem 3), $S_{IBA}^{gen}(X,Y) < 1$ for $X \neq Y$. $\qquad\square$

4. Numerical examples

In this section we aim to illustrate applicability of the proposed approach and its advantages comparing to traditional methods. In the first example, we highlight the benefits of the proposed IBA-based OL comparison in regard to traditional approaches. In the example concerning similarity-based classification, the classification accuracy is significantly improved as we model existing dependencies and interactions in the data by means of the proposed similarity framework.

Example 1 – Car comparison on the level of the object. The aim of this example is to illustrate, numerically and graphically, the motivation for the novel OL comparison. In this example, we compare cars described by two attributes: speed a_s and comfort a_c, whose values are normalized to the unit interval (Table 1). In order to measure similarity between cars X and Y, and cars U and V, we use common Euclidean distance as well as the proposed IBA comparison approach: ABA and OL.

Table 1. The characteristics of observed cars X, Y, U and V

car	*speed*	*comfort*
X	0.7 (fast)	0.6 (comfortable)
Y	0.3 (slow)	0.4 (not so comfortable)
U	0.7 (fast)	0.4 (not so comfortable)
V	0.3 (slow)	0.6 (comfortable)

On the intuitive level, it seems that cars U and V are more similar than cars X and Y.

If Euclidean metric is used to measure the dissimilarity of X and Y, and U and V the resulting values are equal:

$$S_{Eucl}(X,Y) = \sqrt{(x_s - y_s)^2 + (x_c - y_c)^2} = 0.447$$
$$S_{Eucl}(U,V) = \sqrt{(u_s - v_s)^2 + (u_c - v_c)^2} = 0.447 \tag{11}$$

where attributes x_s, y_s, u_s, v_s represent *speed* of cars X, Y, U, V and attributes x_c, y_c, u_c, v_c represents *comfort* of cars X, Y, U, V.

The similarities between the cars in the example are also equally valued by means of IBA-based ABA comparison. The corresponding GBPs and relations on the valued level are given in Appendix A.
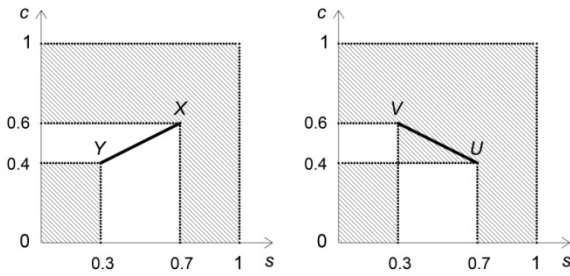
Fig. 1. Geometrical interpretation of similarities $S(X, Y)$ and $S(U, V)$

$$S_{IBA}^{GP}\left(X,Y\right) = S_{IBA}\left(x_s,y_s\right) \otimes S_{IBA}\left(x_c,y_c\right) = 0.48$$
$$S_{IBA}^{GP}\left(U,V\right) = S_{IBA}\left(u_s,v_s\right) \otimes S_{IBA}\left(u_c,v_c\right) = 0.48 \quad (12)$$

However, when we compare these cars on OL using IBA measure we find that the similarity of cars X and Y is smaller than the similarity of U and V. The corresponding GBPs and relations on the valued level are given in Appendix B.

$$S_{IBA}^{O}\left(X,Y\right) = S_{IBA}\left(x_s \otimes x_c, y_s \otimes y_c\right) = 0.70$$
$$S_{IBA}^{O}\left(U,V\right) = S_{IBA}\left(u_s \otimes u_c, v_s \otimes v_c\right) = 0.78 \quad (13)$$

Cars U and V are perceived as more similar than cars X and Y only by means of IBA-base OL comparison. The values of both attributes of car X are larger than both attributes' values of car Y. Therefore, car X is naturally of higher class/price range than car Y. On the other hand, the dissimilarity between objects U and V is not that drastic, since one is faster while the other is more comfortable. In fact, the large value of one attribute compensates for the other to a certain extent. Thus, the similarity values that are provided by IBA comparison on OL are appropriate from the aspect of human perception when comparing two cars.

The graphical interpretation of similarities between the analyzed cars is presented in Fig. 1. Although the geometrical distance between points U and V is equal to the distance between points X and Y, the hatched surfaces that represent the similarity of objects from the aspect of the proposed IBA-based approach are different. More precisely, the hatched surface for U and V is greater than the surface for X and Y. Thus, the graphical interpretation supports the conclusion that objects U and V presented on the right in Fig. 1 are more similar than objects X and Y.

Example 2 – Similarity-based classification.
Similarity-based classification is another example where we apply and evaluate the IBA similarity modeling framework. Basically, we use a simple k-NN algorithm constructed using IBA similarity measure and logical aggregation. The benefits of the proposed aggregation approaches – ABA and OL are also presented. We also compare the results to the traditional k-NN algorithm with Euclidean distance.

Our classification problem is to predict bankruptcy of Serbian middle-sized companies. Using financial reports of 72 Serbian companies, we calculated five financial ratios used in the Z-score model[46] which is widely applied for predicting corporate bankruptcy. These five financial ratios also served as inputs in our classification models, whereas company status (active or bankrupt) is the output. For the purpose of fair comparison among different similarity-based classifiers, we used a balanced dataset. To ensure low bias, we applied the holdout method that is run 100 times as in Ref. 24. Each time data were randomly divided in two equal parts for training and testing.

In the data preprocessing step, attributes' values were scaled to unit interval using standard *min-max* normalization. Further, a high positive correlation was detected in the data between attributes a_2 and a_3 and attributes a_2 and a_4.

For IBA similarity modeling, three groups of IBA similarity classifiers (Eqs. (14), (15), (16)) were constructed on the basis of proposed ABA and OL comparison. The first two classifiers employ a simple average as aggregation operator:

$$S_{IBA}^{A1}\left(X,Y\right) = \frac{1}{5} \cdot \sum_{i=1}^{5} S_{IBA}\left(x_i,y_i\right)$$
$$S_{IBA}^{O1}\left(X,Y\right) = S_{IBA}\left(\frac{1}{5}\cdot\sum_{i=1}^{5}x_i, \frac{1}{5}\cdot\sum_{i=1}^{5}y_i\right) \quad (14)$$

where $x_i, i=1,...,5$ and $y_i, i=1,...,5$ are financial ratios used in the Z-score model of companies X and Y.
The following two classifiers aim to model interaction in the data by using the conjunction of attributes/similarities. Herein, the simple product is used as the operator of GP in aggregation. These models do not include possible correlations.

$$S_{IBA}^{A2}(X,Y) = \prod_{i=1}^{5} S_{IBA}(x_i, y_i)$$

$$S_{IBA}^{O2}(X,Y) = S_{IBA}\left(\prod_{i=1}^{5} x_i, \prod_{i=1}^{5} y_i\right) \tag{15}$$

In the last two models, attributes/similarities are aggregated in accordance with the detected correlations, where *min* function is used as the operator of GP between highly correlated attributes:

$$S_{IBA}^{A3}(X,Y) = S_{IBA}(x_1, y_1) \cdot S_{IBA}(x_5, y_5) \cdot$$
$$\cdot \min\left(S_{IBA}(x_2, y_2), S_{IBA}(x_3, y_3)\right) \cdot$$
$$\cdot \min\left(S_{IBA}(x_2, y_2), S_{IBA}(x_4, y_4)\right) \tag{16}$$
$$S_{IBA}^{O3}(X,Y) = S_{IBA}\big(x_1 \cdot x_5 \cdot \min(x_2, x_3) \cdot \min(x_2, x_4),$$
$$y_1 \cdot y_5 \cdot \min(y_2, y_3) \cdot \min(y_2, y_4)\big)$$

Average accuracy (accy.) and variance (var.) of all presented classification models and $k = \{1, 3, 5, 7, 9\}$ are presented in Table 2.

aggregating similarities. In case of the analyzed data, the described aggregation function on the object level (Eq. (15.2)) improves classification results and classification accuracy is greater than the benchmark.

The highest classification accuracy is achieved when the detected correlations among attributes/similarities are taken into account. Both described classifiers (Eq. (16.1)) and (Eq. (16.2)) improve classification results considerably. However, this time ABA approach outperforms OL comparison, having the best classification accuracy of approximately 90%.

In a nutshell, the classification accuracy is significantly improved as we are able to model existing dependencies and interactions in the data by means of the proposed similarity framework. The constructed classifiers are very descriptive and we can understand the results easily. When it comes to large datasets, IBA-based *k*-NN inherits all well-known pros and cons from a traditional *k*-NN. Additionally, it is obvious that comparison on OL is less computationally demanding

Table 2. Results *k*-NN classification based on Euclidean distance and IBA similarity relations.

k-NN	level	*Euclidean distance (X,Y)*		$S^1(X,Y)$ (average)		$S^2(X,Y)$ (product)		$S^3(X,Y)$ (correlation)	
		accy.	var.	accy.	var.	accy.	var.	accy.	var.
1-NN	ABA	0.7389	0.0028	0.7268	0.0049	0.7486	0.0052	0.8823	0.0027
	OL			**0.5391**	**0.0038**	0.8368	0.0054	0.7873	0.0039
3-NN	ABA	0.7931	0.0019	0.8000	0.0057	0.7986	0.0056	**0.9068**	**0.0029**
	OL			0.5064	0.0008	0.8350	0.0045	0.8441	0.0041
5-NN	ABA	0.8050	0.0019	0.8109	0.0040	0.8055	0.0043	**0.9068**	**0.0029**
	OL			0.4995	0.0000	0.8464	0.0045	0.8677	0.0046
7-NN	ABA	**0.8233**	**0.0016**	0.8164	0.0040	0.8127	0.0042	0.8623	0.0045
	OL			0.5000	0.0000	**0.8441**	**0.0045**	**0.8791**	**0.0042**
9-NN	ABA	0.8189	0.0018	**0.8232**	**0.0055**	**0.8209**	**0.0044**	0.8450	0.0053
	OL			0.5000	0.0000	0.8427	0.0054	0.8673	0.0040

The traditional *k*-NN with Euclidean distance serves as a benchmark with classification accuracy of around 82%. Similar classification results are achieved for the IBA-based classifiers with the simple average of similarities (Eq. (14.1)) for $k = 9$. However, the low accuracy for simple average on OL (Eq. (14.2)) shows that the described relation could not present the object properly. The previous classifiers neither take into account interaction nor the correlations.

The following models (Eq. (15)) that use product function as an aggregation operator try to capture possible interaction among attributes/similarities. The results indicate that modeling interactions as an aggregation of attributes makes more sense than

than ABA comparison.

5. Conclusion

The aim of this paper is to introduce the unified logical framework for modeling similarity based on IBA. It employs the benefits of the existing IBA similarity as interpretable measure and logical aggregation as a descriptive aggregation operator.

In this paper, we have shown that proposed IBA-based similarity framework has a solid mathematical background. More precisely, we have formalized IBA similarity framework and further investigated its properties. In the general case, the proposed IBA similarity models satisfy properties of being similarity

measure. Additionally, IBA-based similarity framework can also be expanded to model non-monotonic inference. Still, the main contribution of our framework lies in its novel approach to multi-attribute object comparison which includes attribute-by-attribute level and object-level comparison, as well as their combination.

The practical advantage of the IBA-based similarity modeling is evaluated on two numerical examples. The first example is illustrative and it confirms motivation and reasoning behind the novel OL comparison. Even though it may not be conventional, OL comparison is important; particularly in the occasions when one object's attribute is logically dependent or can be compensated by another attribute. In the second example, the proposed similarity framework is applied for predicting corporate bankruptcy. Different k-NN classifiers are easily derived on the basis of IBA similarity and LA operators, which has illustrated the generic nature of the proposed IBA-based approach. It is shown that these classification models outperform traditional k-NN with Euclidean distance. They also demonstrate descriptive power and transparency of the proposed similarity modeling.

From the theoretical standpoint the proposed IBA-based similarity framework is rather complete. In future work, the main focus will be on application of the proposed similarity approach to solving problems such as similarity-based classification, clustering, recommender systems as well as TOPSIS based on IBA similarity. Other natural directions for future research include parameterization of IBA similarity measure.

References

1. M. M. Deza and E. Deza, *Encyclopedia of distances* (Springer, Berlin, 2014).
2. H. Le Capitaine, A relevance-based learning model of fuzzy similarity measures, *IEEE T. Fuzzy Syst.* **20**(1) (2012) 57-68.
3. L. A. Zadeh, Similarity relations and fuzzy orderings, *Inform. Sciences* **3**(2) (1971) 177-200.
4. L. Godo and R. O. Rodríguez, Logical approaches to fuzzy similarity-based reasoning: an overview, in *Preferences and similarities*, eds. G. Della Riccia, D. Dubois, R. Kruse and H.-J. Lenz (Springer, Berlin, 2008), pp. 75-128.
5. I. Bloch, Fuzzy relative position between objects in image processing: a morphological approach, *IEEE T. Pattern Anal.* **21**(7) (1999) 657-664.
6. S. M. Chen, M. S. Yeh and P. Y. Hsio, A comparison of similarity measures of fuzzy values, *Fuzzy Set. Syst.* **72** (1995) 79-89.
7. C. P. Pappis and N. I. Karacapilidis, A comparative assessment of measures of similarity of fuzzy values, *Fuzzy Set. Syst.* **56**(2) (1993) 171-174.
8. R. Zwick, E. Carlstein and D. V. Budescu, Measures of similarity among fuzzy concepts: A comparative analysis, *Int. J. Approx. Reason.* **1**(2) (1987) 221-242.
9. L. Kovacs and J. Ratsaby, Analysis of Linear Interpolation of Fuzzy Sets with Entropy-based Distances, *Acta Polytech. Hung.* **10**(3) (2013) 51-64.
10. V. V. Cross and T. A. Sudkamp, *Similarity and compatibility in fuzzy set theory: assessment and applications* (Physica-Verlag, Heidelberg, 2002).
11. A. Tversky, Features of similarity, *Psychol. Rev.* **84**(4) (1977) 327-352.
12. W. Bandler and L. Kohout, Fuzzy power sets and fuzzy implication operators, *Fuzzy Set. Syst.* **4**(1) (1980) 13-30.
13. K. Hirota and W. Pedrycz, Matching fuzzy quantities, *IEEE T. Syst. Man Cybern.* **21**(6) (1991) 1580-1586.
14. B. De Baets and R. Mesiar, Metrics and T-equalities, *J. Math. Anal. Appl.* **267**(2) (2002) 531-547.
15. F. Klawonn and J. L. Castro Peña, Similarity in fuzzy reasoning, *Mathware Soft Comp.* **2**(3) (1995) 197-228.
16. H. Le Capitaine and C. Frelicot, Towards a unified logical framework of fuzzy implications to compare fuzzy sets, in *Proc. of the Joint 2009 Int. Fuzzy Systems Association World Congress and 2009 European Society for Fuzzy Logic and Technologies Conf. (IFSA-EUSFLAT 2009)*, eds. J. P. Carvalho, D. Dubois, U. Kaymak and J. M. C. Sousa (2009), pp. 1200-1205.
17. A. Poledica, P. Milošević, I. Dragović, B. Petrović and D. Radojević, Modeling consensus using logic-based similarity measures, *Soft Comput.* **19**(11) (2015) 3209-3219.
18. M. Grabisch, J. L. Marichal, R. Mesiar and E. Pap, Aggregation functions: means, *Inform. Sciences* **181**(1) (2011) 1-22.
19. J. Bolton and P. Gader, J. N. Wilson, Discrete Choquet integral as a distance metric, *IEEE T. Fuzzy Syst.* **4**(16) (2008) 1107-1110.
20. S. Auephanwiriyakul, J. M. Keller and P. D. Gader, Generalized Choquet fuzzy integral fusion, *Inform. Fusion* **3**(1) (2002) 69-85.
21. E. P. Klement, R. Mesiar and E. Pap, *Triangular norms* (Springer, Berlin, 2013).
22. J. De Leeuw and S. Pruzansky, A new computational method to fit the weighted Euclidean distance model, *Psychometrika* **43**(4) (1978) 479-490.
23. A. Zenebe and A. F. Norcio, Representation, similarity measures and aggregation methods using fuzzy sets for content-based recommender systems, *Fuzzy Set. Syst.* **160**(1) (2009) 76-94.
24. P. Luukka and T. Leppalampi, Similarity classifier with generalized mean applied to medical data, *Comput. Biol. Med.* **36**(9) (2006) 1026-1040.

25. P. Luukka, Similarity classifier using similarities based on modified probabilistic equivalence relations, *Knowl.-Based Syst.* **22**(1) (2009) 57-62.

26. G. Beliakov, T. Calvo and S. James, Consensus measures constructed from aggregation functions and fuzzy implications, *Knowl.-Based Syst.* **55** (2014) 1-8.

27. Q. Ji, P. Haase and G. Qi, Combination of similarity measures in ontology matching using the OWA operator, in *Recent Developments in the Ordered Weighted Averaging Operators: Theory and Practice*, eds. R. R. Yager, J. Kacprzyk and G. Beliakov (Springer, Berlin, 2011), pp. 281-295.

28. N. B. Karayiannis, Soft learning vector quantization and clustering algorithms based on ordered weighted aggregation operators, *IEEE T. Neural Networ.* **11**(5) (2000) 1093-1105.

29. J. M. Merigó and M. Casanovas, Induced aggregation operators in the Euclidean distance and its application in financial decision making, *Expert Syst. Appl.* **38**(6) (2011) 7603-7608.

30. J. M. Merigó and M. Casanovas, A new Minkowski distance based on induced aggregation operators, *Int. J. Comput. Int. Sys.* **4**(2) (2011) 123-133.

31. G. Beliakov, S. James and G. Li, Learning Choquet-integral-based metrics for semisupervised clustering, *IEEE T. Fuzzy Syst.* **19**(3) (2011) 562-574.

32. C. M. Hwang, M. S. Yang, W. L. Hung and M. G. Lee, A similarity measure of intuitionistic fuzzy sets based on the Sugeno integral with its application to pattern recognition, *Inform. Sciences* **189** (2012) 93-109.

33. M. A. Gluck and G. H. Bower, Evaluating an adaptive network model of human learning, *J. Mem. Lang.* **27**(2) (1988) 166-195.

34. D. Danks, Theory unification and graphical models in human categorization, in *Causal Learning: Psychology, Philosophy, and Computation*, eds. A. Gopnik and L. Schulz (Oxford University Press, New York, 2007), pp. 173-189. DOI:10.1093/acprof:oso/9780195176803.003.0012

35. C. Kahraman, B. Oztaysi and S. C. Onar, A Comprehensive Literature Review of 50 Years of Fuzzy Set Theory, *Int. J. Comput. Int. Sys.* **9**(supp 1) (2016) 3-24.

36. D. Radojević, Logical aggregation based on interpolative Boolean algrebra, *Mathware Soft Comp.* **15**(1) (2008) 125-141.

37. A. Rakićević, I. Nešić and D. Radojević, A novel approach to hierarchical clustering based on logical measure of dissimilarity, in *Proceedings of the 11th Balkan Conference on Operational Research*, eds. N. Mladenović, G. Savić, M. Kuzmanović, D. Makajić-Nikolić and M. Stanojević (Faculty of Organizational Sciences, Belgrade, 2013), pp. 147-155.

38. V. Dobrić, D. Kovačević, B. Petrović, D. Radojević and P. Milošević, Formalization of human categorization process using Interpolative Boolean algebra, *Math. Prob. Eng.* **2015** (2015) Article ID 620797.

39. T. Calvo, G. Mayor and R. Mesiar (eds.), *Aggregation operators: new trends and applications* (Physica-Verlag, Heidelberg, 2012).

40. D. Radojević, [0,1]-valued logic: A natural generalization of Boolean logic, *Yugoslav J. Oper. Res.* **10**(2) (2000) 185-216.

41. A. Poledica, P. Milošević, I. Dragović, D. Radojević and B. Petrović, A consensus model in group decision making based on interpolative Boolean algebra, in *Proc. 8th conference of the European Society for Fuzzy Logic and Technology*, eds. G. Pasi, J. Montero and D. Ciucci (Atlantis Press, Amsterdam, 2013), pp. 648-654.

42. A. Rakićević, V. Simeunović, B. Petrović and S. Milić, An automated system for stock market trading based on logical clustering, *Teh. Vjesn.* **25**(4) (2018).

43. I. Dragović, N. Turajlić, D. Radojević, and B. Petrović, Combining Boolean consistent fuzzy logic and AHP illustrated on the web service selection problem. *Int. J. Comput. Int. Sys.* **7**(supp 1) (2014) 84-93.

44. P. Milošević, I. Nešić, A. Poledica, D. Radojević and B. Petrović, Logic-based aggregation methods for ranking student applicants, *Yugoslav J. Oper. Res.* (article in press). DOI:10.2298/YJOR161110007M

45. D. T. Larose, *Discovering knowledge in data: an introduction to data mining* (John Wiley & Sons, Hoboken, 2005).

46. E. I. Altman, Financial ratios, discriminant analysis and the prediction of corporate bankruptcy, *J. Financ.* **23**(4) (1968) 589-609.

Appendix A.

Similarity models given in Eq. (12) are mapped in the following GBP.

$$S_{IBA}^{GP}(X,Y) = S_{IBA}(x_s, y_s) \otimes S_{IBA}(x_c, y_c) =$$
$$= 1 - x_s - y_s - x_c - y_c + 2 \cdot x_s \otimes y_s + 2 \cdot x_c \otimes y_c +$$
$$+ y_s \otimes y_c - 2 \cdot x_s \otimes y_s \otimes x_c - 2 \cdot x_s \otimes y_s \otimes y_c - \quad (A.1)$$
$$+ y_s \otimes x_c - 2 \cdot x_s \otimes x_c \otimes y_c - 2 \cdot y_s \otimes x_c \otimes y_c +$$
$$+ x_s \otimes x_c + x_s \otimes y_c + 4 \cdot x_s \otimes y_s \otimes x_c \otimes y_c$$

The product function is used to aggregate speed and comfort since attributes *s* and *c* are independent, while the *min* function is used to aggregate attributes of the same nature.

$$S_{IBA}^{GP}(X,Y) = S_{IBA}(x_s, y_s) \otimes S_{IBA}(x_c, y_c) =$$
$$= 1 - x_s - y_s - x_c + 2 \cdot \min(x_s, y_s) + 2 \cdot \min(x_c, y_c) -$$
$$- y_c + y_s \cdot y_c - 2 \cdot \min(x_s, y_s) \cdot x_c - 2 \cdot \min(x_s, y_s) \cdot y_c - \quad (A.2)$$
$$+ y_s \cdot x_c - 2 \cdot x_s \cdot \min(x_c, y_c) - 2 \cdot y_s \cdot \min(x_c, y_c) +$$
$$+ x_s \cdot x_c + x_s \cdot y_c + 4 \min(x_s, y_s) \cdot \min(x_c, y_c)$$

Appendix B.

Similarity models given in Eq. (13) are mapped in the following GBP.

$$S_{IBA}^{O}(X,Y) = S_{IBA}(x_s \otimes x_c, y_s \otimes y_c) =$$
$$= 1 - x_s \otimes x_c - y_s \otimes y_c \quad \text{(B.1)}$$
$$+ 2 \cdot x_s \otimes x_c \otimes y_s \otimes y_c$$

The product function is used to aggregate speed and comfort since attributes s and c are independent, while the *min* function is used to aggregate attributes of the same nature.

$$S_{IBA}^{O}(X,Y) = 1 - x_s \cdot x_c - y_s \cdot y_c +$$
$$+ 2 \cdot \min(x_s, x_c) \cdot \min(y_s, y_c) \quad \text{(B.2)}$$