8-1-2008

Computing Degrees of Subsethood and Similarity for Interval-Valued Fuzzy Sets: Fast Algorithms

Hung T. Nguyen

Vladik Kreinovich
University of Texas at El Paso, vladik@utep.edu

Comments:
Technical Report: UTEP-CS-08-27a
Published in *Proceedings of the 9th International Conference on Intelligent Technologies InTech'08*, Samui, Thailand, October 7-9, 2008, pp. 47-55.

Computing Degrees of Subsethood and Similarity for Interval-Valued Fuzzy Sets: Fast Algorithms

Hung T. Nguyen
Department of Mathematical Sciences
New Mexico State University
Las Cruces, NM 88003, USA
hunguyen@nmsu.edu

Vladik Kreinovich
Department of Computer Science
University of Texas at El Paso
El Paso, TX 79968, USA
vladik@utep.edu

***Abstract*—Subsethood $A \subseteq B$ and set equality $A = B$ are among the basic notions of set theory. For traditional ("crisp") sets, every element a either belongs to a set A or it does not belong to A, and for every two sets A and B, either $A \subseteq B$ or $A \not\subseteq B$. To describe commonsense and expert reasoning, it is advantageous to use *fuzzy sets* in which for each element a, there is a *degree* $\mu_A(a) \in [0,1]$ to which a belongs to this set. For fuzzy sets A and B, it is reasonable to define a *degree* of subsethood $d_{\subseteq}(A,B)$ and degree of equality (degree of similarity) $d_{=}(A,B)$. In practice, it is often difficult to assign a definite membership degree $\mu_A(a)$ to each element a; it is more realistic to expect that an expert describes an *interval* $[\underline{\mu}_A(a), \overline{\mu}_A(a)]$ of possible values of this degree. The resulting *interval-valued* fuzzy set can be viewed as a class of all possible fuzzy sets $\mu_A(a) \in [\underline{\mu}_A(a), \overline{\mu}_A(a)]$.**

For interval-valued fuzzy sets $\mathbf{A}$ and $\mathbf{B}$, it is therefore reasonable to define the degree of subsethood $d_{\subseteq}(\mathbf{A},\mathbf{B})$ as the range of possible values of $d_{\subseteq}(A,B)$ for all $A \in \mathbf{A}$ and $B \in \mathbf{B}$ – and similarly, we can define the degree of similarity $d_{=}(\mathbf{A},\mathbf{B})$. So far, no general algorithms were known for computing these ranges. In this paper, we describe such general algorithms. The newly proposed algorithms are reasonably fast: for fuzzy subsets of an n-element universal set, these algorithms compute the ranges in time $O(n \cdot \log(n))$.

I. Formulation of the Problem

Subsethood and set equality are important notions of set theory. In traditional set theory, among the basic notions are the notions of set equality and subsethood:

- two sets A and B are equal if they contain exactly the same elements, and
- a set A is a subset of the set B if every element of the set A also belongs to B.

Because of this importance, it is desirable to generalize these notions to fuzzy sets.

In fuzzy set theory, it is reasonable to talk about degrees of subsethood and equality (similarity). In traditional set theory, for every two sets A and B, either A is a subset of B, or A is not a subset of B. Similarly, either the two sets A and B are equal or these two sets are different.

The main idea behind fuzzy logic is that for fuzzy, imprecise concepts, everything is a matter of degree; see, e.g., [3], [9]. Thus, for two fuzzy sets A and B, it is reasonable to define degree of subsethood and degree of similarity.

How to describe degree of subsethood: main idea. In fuzzy logic and fuzzy set theory, there is no built-in notion of degree of subsethood or degree of equality (similarity) between the sets. Instead, the standard descriptions of fuzzy logic and fuzzy set theory start with the notions of union and intersection.

The simplest way to describe the union of the two sets is to take the maximum of the corresponding membership functions: $\mu_{A\cup B}(x) = \max(\mu_A(x), \mu_B(x))$. Similarly, the simplest way to describe the intersection of the two sets is to take the minimum of the corresponding membership functions: $\mu_{A\cap B}(x) = \min(\mu_A(x), \mu_B(x))$. Thus, to describe the degrees of subsethood and equality (similarity), it is reasonable to express the notions of subsethood and set equality in terms of union and intersection. This expression is well known in set theory: it is known that

- in general, $A \cap B \subseteq A$, and
- $A \subseteq B$ if and only if $A \cap B = A$.

So, for crisp finite sets, to check whether A is a subset of B, we can consider the ratio

$$\frac{|A \cap B|}{|A|},$$

where $|A|$ denotes the number of elements in a set A:

- in general, this ratio is between 0 and 1, and
- this ratio is equal to 1 if and only if A is a subset of B.

The smaller the ratio, the more there are elements from A which are not part of the intersection $A \cap B$, and thus, not part of the set B. Thus, for crisp sets, this ratio can be viewed as a reasonable measure of degree to which A is a subset of B.

A similar definition can be used to define degree of subsethood of two fuzzy sets. Specifically, for finite fuzzy sets, we can use a natural fuzzy extension of the notion of cardinality: $|A| \stackrel{\text{def}}{=} \sum \mu_A(x)$. Let us describe the resulting formulas.

Since we only consider finite fuzzy sets, we can therefore consider a finite universe of discourse. Without losing generality, we can denote the elements of the universe of discourse by their numbers $1, 2, \ldots, n$. The values of the membership function corresponding to the fuzzy set A can be therefore denoted by $a_1, \ldots, a_n$. Similarly, the values of the membership function corresponding to the fuzzy set B can be denoted by $b_1, \ldots, b_n$. In these notations,

- the membership function corresponding to the intersection $A \cap B$ has the values $\min(a_1, b_1), \ldots, \min(a_n, b_n)$,
- the cardinality $|A|$ of the fuzzy set A is equal to $\sum\limits_{i=1}^{n} a_i$, and
- the cardinality $|A \cap B|$ of the intersection $A \cap B$ is equal to $\sum\limits_{i=1}^{n} \min(a_i, b_i)$.

Thus, the degree of subsethood of fuzzy sets A and B can be defined as the ratio

$$d_{\subseteq}(A, B) = \frac{\sum\limits_{i=1}^{n} \min(a_i, b_i)}{\sum\limits_{i=1}^{n} a_i}.$$

Comment. An alternative (probabilistic) justification of this formula is given in the Appendix.

How to describe degree of equality (similarity). It is known that

- in general, $A \cap B \subseteq A \cup B$, and
- $A = B$ if and only if $A \cap B = A \cup B$.

So, for crisp finite sets, to check whether A is equal to B, we can consider the ratio

$$\frac{|A \cap B|}{|A \cup B|}.$$

- In general, this ratio is between 0 and 1, and
- this ratio is equal to 1 if and only if A is a subset of B.

The smaller the ratio, the more there are elements from $A \cup B$ which are not part of the intersection $A \cap B$, and thus, elements from one of the sets A and B which do not belong to the other of these two sets. Thus, for crisp sets, this ratio can be viewed as a reasonable measure of degree to which A is equal to B.

A similar definition can be used to define degree of equality (similarity) of two fuzzy sets A and B:

$$d_{=}(A, B) = \frac{\sum\limits_{i=1}^{n} \min(a_i, b_i)}{\sum\limits_{i=1}^{n} \max(a_i, b_i)}.$$

There exist many alternative ways of describing degrees of subsethood and similarity. The above expressions are the simplest and probably most frequently used. However, there exist other expressions for the similar degrees; see, e.g., [1], [3], [9]. For example, an alternative way to describe the degree of subsethood is based on the fact that

- in general, B is a subset of the union $A \cup B$, and
- the set B is equal to the union $A \cup B$ if and only if A is a subset of B.

Thus, as an alternative degree of subsethood, we can take a ratio

$$\frac{|B|}{|A \cup B|} = \frac{\sum\limits_{i=1}^{n} b_i}{\sum\limits_{i=1}^{n} \max(a_i, b_i)}.$$

Need for interval-valued (and more general type-2) fuzzy sets. In the above text, we consider the situation in which the values a_i and b_i of the membership function are numbers from the interval $[0, 1]$. Is this the most adequate description?

The main objective of fuzzy logic is to describe uncertain ("fuzzy") knowledge, when an expert cannot describe his or her knowledge by an exact value or by a precise set of possible values. Instead, the expert describes this knowledge by using words from natural language. Fuzzy logic provides a procedure for formalizing these words into a computer-understandable form – as fuzzy sets.

In the traditional approach to fuzzy logic, the expert's degree of certainty in a statement – such as the value $\mu_A(x)$ describing that the value x satisfies the property A (e.g., "small") – is characterized by a number from the interval $[0, 1]$. However, we are considering situations in which an expert is unable to describe his or her knowledge in precise terms. It is not very reasonable to expect that in this situation, the same expert will be able to meaningfully express his or her degree of certainty by a precise number. It is much more reasonable to assume that the expert will describe these degrees also by words from natural language.

Thus, for every x, a natural representation of the degree $\mu(x)$ is not a number, but rather a new fuzzy set. Such situations, in which to every value x we assign a fuzzy set $\mu(x)$, are called *type-2* fuzzy sets.

Successes of type-2 fuzzy sets. Type-2 fuzzy sets are actively used in practice; see, e.g., [5], [6]. Since type-2 fuzzy sets provide a more adequate representation of expert knowledge, it is not surprising that such sets lead to a higher quality control, higher quality clustering, etc., in comparison with the more traditional type-1 sets.

It is therefore desirable to extend the above formulas for the degrees of subsethood and similarity to type-2 fuzzy sets.

The main obstacle to using type-2 fuzzy sets. If type-2 fuzzy sets are more adequate, why are not they used more? The main reason why their use is limited is that the transition from type-1 to type-2 fuzzy sets leads to an increase in computation time. Indeed, to describe a traditional (type-1) membership function function, it is sufficient to describe, for each value x, a single number $\mu(x)$. In contrast, to describe a type-2 set, for each value x, we must describe the entire membership function – which needs several parameters to describe. Since we need more numbers just to store such information, we need more computational time to process all the numbers representing these sets.

Interval-valued fuzzy sets. In line with this reasoning, the most widely used type-2 fuzzy sets are the ones which require the smallest number of parameters to store. We are talking about *interval-valued* fuzzy numbers, in which for each x, the degree of certainty $\mu(x)$ is an interval $[\underline{\mu}(x), \overline{\mu}(x)]$. To store each interval, we need exactly two numbers – the smallest possible increase over the single number needed to store the type-1 value $\mu(x)$.

It is therefore desirable to extend the above definitions to interval-valued fuzzy sets and to come up with algorithms (ideally, efficient algorithms) for computing the corresponding degrees. This problem was formulated in [7].

Comment. It is worth mentioning that once we have efficient algorithms for interval-valued fuzzy sets, these algorithms can be naturally extended to efficient algorithms for arbitrary type-2 fuzzy sets [4].

Let us describe the corresponding computational problems in precise mathematical terms.

Computing the degree of subsethood: precise formulation of the problem. For each i from 1 to n, we know the intervals $[\underline{a}_i, \overline{a}_i]$ and $[\underline{b}_i, \overline{b}_i]$ of possible membership degrees. For each combination of values $a_i \in [\underline{a}_i, \overline{a}_i]$ and $b_i \in [\underline{b}_i, \overline{b}_i]$, we can compute the subsethood degree

$$d_{\subseteq} = \frac{\sum\limits_{i=1}^{n} \min(a_i, b_i)}{\sum\limits_{i=1}^{n} a_i}.$$

The objective is to find the range $[\underline{d}_{\subseteq}, \overline{d}_{\subseteq}]$ of possible values of the above subsethood degree, i.e.,

- to compute the smallest possible value $\underline{d}_{\subseteq}$ of the subsethood degree $d_{\subseteq}$ when $a_i \in [\underline{a}_i, \overline{a}_i]$ and $b_i \in [\underline{b}_i, \overline{b}_i]$; and
- to compute the largest possible value $\overline{d}_{\subseteq}$ of the subsethood degree $d_{\subseteq}$ when $a_i \in [\underline{a}_i, \overline{a}_i]$ and $b_i \in [\underline{b}_i, \overline{b}_i]$.

Computing the degree of equality (similarity): precise formulation of the problem. For each i from 1 to n, we know the intervals $[\underline{a}_i, \overline{a}_i]$ and $[\underline{b}_i, \overline{b}_i]$ of possible membership degrees. For each combination of values $a_i \in [\underline{a}_i, \overline{a}_i]$ and $b_i \in [\underline{b}_i, \overline{b}_i]$, we can compute the similarity degree

$$d_{=} = \frac{\sum\limits_{i=1}^{n} \min(a_i, b_i)}{\sum\limits_{i=1}^{n} \max(a_i, b_i)}.$$

The objective is to find the range $[\underline{d}_{=}, \overline{d}_{=}]$ of possible values of the above similarity degree, i.e.,

- to compute the smallest possible value $\underline{d}_{=}$ of the similarity degree $d_{=}$ when $a_i \in [\underline{a}_i, \overline{a}_i]$ and $b_i \in [\underline{b}_i, \overline{b}_i]$; and
- to compute the largest possible value $\overline{d}_{\subseteq}$ of the similarity degree $d_{=}$ when $a_i \in [\underline{a}_i, \overline{a}_i]$ and $b_i \in [\underline{b}_i, \overline{b}_i]$.

What we plan to do in this paper. In this paper, we design fast algorithms for computing the desired bounds $\underline{d}_{\subseteq}$, $\overline{d}_{\subseteq}$, $\underline{d}_{=}$, and $\overline{d}_{=}$.

II. Computing $\underline{d}_{\subseteq}$: Analysis of the Problem

First observation: minimum of $d_{\subseteq}$ is attained when $b_i = \underline{b}_i$. First, let us notice that for every i, the degree of subsethood is a (non-strictly) increasing function of b_i. Thus, its smallest possible value is attained when each of the n variables b_i attains its smallest possible value $\underline{b}_i$. So, to find the smallest possible value $\underline{d}_{\subseteq}$ of the degree of subsethood $d_{\subseteq}$, it is sufficient to consider the values $b_i = \underline{b}_i$. In other words, it is sufficient to find the minimum of the ratio

$$r = \frac{\sum\limits_{i=1}^{n} \min(a_i, \underline{b}_i)}{\sum\limits_{i=1}^{n} a_i}$$

when $a_i \in [\underline{a}_i, \overline{a}_i]$.

Second observation: minimum of $d_{\subseteq}$ is attained for some $a_i \in [\underline{a}_i, \overline{a}_i]$. When at least one of the lower bounds $\underline{a}_i$ is positive, the ratio r is a continuous function on a compact set

$$[\underline{a}_1, \overline{a}_1] \times \ldots \times [\underline{a}_n, \overline{a}_n].$$

Thus, the minimum of this function are attained at some values $a_i \in [\underline{a}_i, \overline{a}_i]$.

The degenerate case when $\underline{a}_1 = \ldots = \underline{a}_n = 0$ can be obtained as a limit case of the genetic situation. Thus, without losing generality, we can consider the case when one of the values $\underline{a}_i$ is positive and thus, the minimum $\underline{d}_{\subseteq}$ is attained.

Two possible cases. Let us consider the values a_i for which this minimum is attained. For each i, there are two possible cases:

- the case when $a_i \le \underline{b}_i$ and thus, $\min(a_i, \underline{b}_i) = a_i$, and
- the case when $\underline{b}_i \le a_i$ and thus, $\min(a_i, \underline{b}_i) = \underline{b}_i$.

Let us consider these two cases one by one.

Case when $a_i \le \underline{b}_i$. In this case, both the numerator and the denominator of the ratio r contain the term a_i. Thus, the dependence of the ratio r on a_i takes the form

$$r = \frac{a_i + m_i}{a_i + M_i},$$

where $m_i \stackrel{\text{def}}{=} \sum\limits_{j \ne i} \min(a_j, \underline{b}_j)$ and $M_i \stackrel{\text{def}}{=} \sum\limits_{j \ne i} a_j$. Since $\min(a_j, \underline{b}_j) \le a_j$ for all j, we have $m_i \le M_i$. Thus, the ratio can be reformulated as follows:

$$r = 1 - \frac{M_i - m_i}{a_i + M_i}.$$

When a_i increases, the sum $a_i + M_i$ also increases, hence the ratio $\dfrac{M_i - m_i}{a_i + M_i}$ decreases, and the difference $r = 1 - \dfrac{M_i - m_i}{a_i + M_i}$ increases. So, the ratio r attains its smallest value when a_i is the smallest, i.e., when $a_i = \underline{a}_i$.

Case when $\underline{b}_i \le a_i$. In this case, only the denominator of the ratio r contains a_i. Specifically, the dependence of r on a_i takes the form

$$r = \frac{m}{a_i + M_i},$$

where M_i is the same sum as above, and

$$m \stackrel{\text{def}}{=} \sum_{j=1}^{n} \min(a_j, \underline{b}_j)$$

does not depend on a_i, since $m = \underline{b}_i + m_i$. This expression for r clearly decreases with a_i. Thus, in this case, the smallest possible value of r is attained when a_i attains the largest possible value, i.e., when $a_i = \overline{a}_i$.

Comparing the two cases. From the analysis of the above two cases, we conclude that for each i, we have two cases:

- either $a_i = \underline{a}_i$ and $\underline{a}_i \le \underline{b}_i$,
- or $a_i = \overline{a}_i$ and $\underline{b}_i \le \overline{a}_i$.

If $\overline{a}_i < \underline{b}_i$, then we cannot have the second case, and thus, the minimum is attained when $a_i = \underline{a}_i$.

If $\underline{b}_i < \underline{a}_i$, then we cannot have the first case and thus, the minimum is attained when $a_i = \overline{a}_i$.

In the remaining cases, when $\underline{a}_i \le \underline{b}_i \le \overline{a}_i$, in principle, both cases $a_i = \underline{a}_i$ and $a_i = \overline{a}_i$ are possible. Which of these two cases minimizes the ratio r? E.g., when does the value $a_i = \underline{a}_i$ lead to the minimum? Since we have already shown that there are only two possible values a_i for which the ratio attains its minimum, the minimum is attained at $a_i = \underline{a}_i$ if and only if replacing $a_i = \underline{a}_i$ with $a_i = \overline{a}_i$ increases the ratio r.

The original ratio is equal to $r = \dfrac{m}{M}$, where the sum m is defined above, and $M \stackrel{\text{def}}{=} \sum\limits_{j=1}^{n} a_j$. When we replace $\underline{a}_i \le \underline{b}_i$ with $\overline{a}_i \ge \underline{b}_i$, then in the numerator, the i-th term $\min(a_i, b_i)$ changes from $\underline{a}_i$ to $\underline{b}_i$. Thus, instead of the original value m, the numerator gets the new value $m + (\underline{b}_i - \underline{a}_i)$.

In the denominator, the term $\underline{a}_i$ is replaced by a new term $\overline{a}_i$. Thus, instead of the original value M, the denominator gets the new value $M + (\overline{a}_i - \underline{a}_i)$. Hence, instead of the original ratio $\dfrac{m}{M}$, we get the new ratio

$$\frac{m + (\underline{b}_i - \underline{a}_i)}{M + (\overline{a}_i - \underline{a}_i)}.$$

This new ratio is larger than the original one when

$$\frac{m}{M} < \frac{m + (\underline{b}_i - \underline{a}_i)}{M + (\overline{a}_i - \underline{a}_i)}.$$

Multiplying both sides by both (positive) denominators and canceling a term $m \cdot M$ which is common to both sides of the resulting inequality, we conclude that

$$m \cdot (\overline{a}_i - \underline{a}_i) < M \cdot (\underline{b}_i - \underline{a}_i).$$

Dividing both sides of this inequality by positive values $\overline{a}_i - \underline{a}_i$ and M, we conclude that

$$\frac{m}{M} < \frac{\underline{b}_i - \underline{a}_i}{\overline{a}_i - \underline{a}_i}.$$

This is the condition under which the minimum is attained for $a_i = \underline{a}_i$; if this inequality is not satisfied, then the minimum is attained for $a_i = \overline{a}_i$.

Towards an algorithm for computing $\underline{d}_{\subseteq}$. If we know the optimal value $\underline{d}_{\subseteq}$ of the ratio $r = m/M$, then for each i, we can determine whether the minimum is attained for $a_i = \underline{a}_i$ or for $a_i = \overline{a}_i$. In reality, we do not need to know the exact value of the ratio r, we just need to know where it is located in comparison with n ratios $r_i \stackrel{\text{def}}{=} \dfrac{\underline{b}_i - \underline{a}_i}{\overline{a}_i - \underline{a}_i}$. For each of $n+1$ possible locations, we compute the corresponding values a_i, and then find the location for which the ratio is the smallest possible.

Specifically, we analyze the values a_i one by one, and divide the corresponding indices into three groups:

- The group I^- consists of all the indices i for which
$$\overline{a}_i < \underline{b}_i.$$
For this group, as we have shown, the minimum is attained when $a_i = \underline{a}_i$; in this case, $\min(a_i, \underline{b}_i) = \underline{a}_i$.
- The group I^+ consists of all the indices i for which
$$\underline{b}_i < \underline{a}_i.$$
For this group, as we have shown, the minimum is attained when $a_i = \overline{a}_i$; in this case, $\min(a_i, \underline{b}_i) = \underline{b}_i$.

The remaining indices i (i.e., all the indices for which $\underline{b}_i \in [\underline{a}_i, \overline{a}_i]$) form the third group I. For all $i \in I$, we compute the ratios

$$r_i \stackrel{\text{def}}{=} \frac{\underline{b}_i - \underline{a}_i}{\overline{a}_i - \underline{a}_i}.$$

We then sort the indices $i \in I$ in the increasing order of this ratio. Let k denote the number of elements in the group I; it is clear that $k \le n$. Then, in the resulting new ordering, we have $r_1 \le r_2 \le \ldots \le r_k$. We can place indices from the groups I^- and I^+ after k indices from the group I.

The ordering of the ratios r_i subdivides the real line into $k+1$ "zones"

$$z_0 \stackrel{\text{def}}{=} (-\infty, r_1], z_1 \stackrel{\text{def}}{=} [r_1, r_2], \ldots,$$
$$z_{k-1} \stackrel{\text{def}}{=} [r_{k-1}, r_k], z_k \stackrel{\text{def}}{=} [r_k, \infty).$$

For each zone z_j, we take the assignment corresponding to the case when the actual (unknown) value $r = \underline{d}_{\subseteq}$ is in this zone. In this case, as we have shown,

- for $i \le j$, minimum is attained for $a_i = \overline{a}_i$ and $\min(a_i, \underline{b}_i) = \underline{b}_i$;
- for $i > j$, minimum is attained for $a_i = \underline{a}_i$ and $\min(a_i, \underline{b}_i) = \underline{a}_i$.

Thus, the corresponding value $r^{(j)}$ of the ratio r can be computed as

$$r^{(j)} = \frac{m^{(j)}}{M^{(j)}},$$

where

$$m^{(j)} \stackrel{\text{def}}{=} \sum_{i \le j} \underline{b}_i + \sum_{i=j+1}^{k} \underline{a}_i + \sum_{i \in I^-} \underline{a}_i + \sum_{i \in I^+} \underline{b}_i;$$

$$M^{(j)} \stackrel{\text{def}}{=} \sum_{i \le j} \overline{a}_i + \sum_{i=j+1}^{k} \underline{a}_i + \sum_{i \in I^-} \underline{a}_i + \sum_{i \in I^+} \overline{a}_i.$$

The values $m^{(j)}$ and $M^{(j)}$ do not need to be recomputed every time, since the next value differs from the previous one only by two terms:

$$m^{(j+1)} = m^{(j)} + (\underline{b}_{j+1} - \underline{a}_{j+1});$$

$$M^{(j+1)} = m^{(j)} + (\overline{a}_{j+1} - \underline{a}_{j+1}).$$

Thus, we arrive at the following algorithm.

III. Algorithm for Computing $\underline{d}_{\subseteq}$

Description of the algorithm. First, we divide n indices into three groups:

- The group I^- consists of all the indices i for which
$$\overline{a}_i < \underline{b}_i.$$
- The group I^+ consists of all the indices i for which
$$\underline{b}_i < \underline{a}_i.$$

The remaining indices i form the third group I.

For all $i \in I$, we compute the ratios

$$r_i \stackrel{\text{def}}{=} \frac{\underline{b}_i - \underline{a}_i}{\overline{a}_i - \underline{a}_i}.$$

We then sort the indices $i \in I$ in the increasing order of this ratio. In the resulting new ordering, we have

$$r_1 \le r_2 \le \ldots \le r_k,$$

where k is the number of elements in the group I.

We then compute

$$m^{(0)} = \sum_{i \in I} \underline{a}_i + \sum_{i \in I^-} \underline{a}_i + \sum_{i \in I^+} \underline{b}_i$$

and

$$M^{(0)} = \sum_{i \in I} \underline{a}_i + \sum_{i \in I^-} \underline{a}_i + \sum_{i \in I^+} \overline{a}_i.$$

For j from 0 to k, we then compute

$$m^{(j+1)} = m^{(j)} + (\underline{b}_{j+1} - \underline{a}_{j+1});$$

$$M^{(j+1)} = m^{(j)} + (\overline{a}_{j+1} - \underline{a}_{j+1}).$$

For all j from 0 to $k+1$, we compute $r^{(j)} = \dfrac{m^{(j)}}{M^{(j)}}$. The smallest of these values $r^{(j)}$ is the desired smallest value $\underline{d}_{\subseteq}$.

Computational complexity of this algorithm. Let us describe how much computation time is needed for each stage of the above algorithm.

- Dividing indices into 3 groups requires linear time $O(n)$.
- Computing the ratios r_i for $i \in I$ also requires linear time $O(n)$.
- Sorting $k \le n$ requires time $O(k \cdot \log(k)) = O(n \cdot \log(n))$.
- Computing $m^{(0)}$ and $M^{(0)}$ requires linear time $O(n)$.
- Computing each new pair $m^{(j+1)}$ and $M^{(j+1)}$ requires a constant number of computation steps, so computing $\le n$ such pairs requires $\le \text{const} \cdot n = O(n)$ steps.
- Finally, finding the smallest of $n+1$ numbers also requires linear time $O(n)$.

Thus, for this algorithm, the overall computation time is

$$O(n) + O(n) + O(n \cdot \log(n)) + O(n) + O(n) + O(n) =$$
$$O(n \cdot \log(n)).$$

IV. Computing $\overline{d}_{\subseteq}$: Analysis of the Problem

First observation: maximum of $d_{\subseteq}$ is attained when $b_i = \overline{b}_i$. We have already mentioned that for every i, the degree of subsethood $d_{\subseteq}$ is a (non-strictly) increasing function of b_i. Thus, its largest possible value $\overline{d}_{\subseteq}$ is attained when each of the n variables b_i attains its largest possible value $\underline{b}_i$. So, to find the largest possible value $\overline{d}_{\subseteq}$ of the degree of subsethood $d_{\subseteq}$, it is sufficient to consider the values $b_i = \overline{b}_i$. In other words, it is sufficient to find the maximum of the ratio

$$r' = \frac{\sum\limits_{i=1}^{n} \min(a_i, \overline{b}_i)}{\sum\limits_{i=1}^{n} a_i}$$

when $a_i \in [\underline{a}_i, \overline{a}_i]$.

Two possible cases. Let us consider the values a_i for which this maximum is attained. For each i, there are two possible cases:

- the case when $a_i \le \overline{b}_i$ and thus, $\min(a_i, \overline{b}_i) = a_i$, and
- the case when $\overline{b}_i \le a_i$ and thus, $\min(a_i, \overline{b}_i) = \overline{b}_i$.

Let us consider these two cases one by one.

Case when $a_i \le \overline{b}_i$. In this case, both the numerator and the denominator of the ratio r contain the term a_i. Thus, the dependence of the ratio r' on a_i takes the form

$$r' = \frac{a_i + m'_i}{a_i + M_i},$$

where $m'_i \stackrel{\text{def}}{=} \sum\limits_{j \ne i} \min(a_j, \overline{b}_j)$ and $M_i \stackrel{\text{def}}{=} \sum\limits_{j \ne i} a_j$. Since $\min(a_j, \overline{b}_j) \le a_j$ for all j, we have $m'_i \le M_i$. Thus, the ratio can be reformulated as follows:

$$r' = 1 - \frac{M_i - m'_i}{a_i + M_i}.$$

When a_i increases, the sum $a_i + M_i$ also increases, hence the ratio $\dfrac{M_i - m'_i}{a_i + M_i}$ decreases, and the difference $r' = 1 - \dfrac{M_i - m'_i}{a_i + M_i}$ increases. So, the ratio r' attains its largest value when $a_i \in [\underline{a}_i, \overline{a}_i]$ is the largest within the requirement $a_i \le \overline{b}_i$, i.e., when

- either $a_i = \overline{a}_i \le \overline{b}_i$,
- or $a_i = \overline{b}_i$ and thus, $\underline{a}_i \le \overline{b}_i < \overline{a}_i$.

Case when $\overline{b}_i \le a_i$. In this case, only the denominator of the ratio r' contains a_i. Specifically, the dependence of r' on a_i takes the form

$$r' = \frac{m'}{a_i + M_i},$$

where M_i is the same sum as above, and

$$m' \stackrel{\text{def}}{=} \sum_{j=1}^{n} \min(a_j, \overline{b}_j)$$

does not depend on a_i, since $m' = \overline{b}_i + m'_i$. This expression for r' clearly decreases with a_i. Thus, in this case, the smallest possible value of r' is attained when $a_i \in [\underline{a}_i, \overline{a}_i]$ attains the smallest possible value within the requirement $\overline{b}_i \le a_i$, i.e., when

- either $a_i = \underline{a}_i$ and $\overline{b}_i \le \underline{a}_i$,
- or $a_i = \overline{b}_i$ and thus, $\underline{a}_i < \overline{b}_i \le \overline{a}_i$.

Synthesizing the analysis of the two cases. From the analysis of the above two cases, we conclude that for each i, we have three possibilities:

- $a_i = \overline{a}_i \le \overline{b}_i$;
- $a_i = \underline{b}_i \ge \overline{b}_i$; and
- $a_i = \overline{b}_i$ and $\underline{a}_i \le \overline{b}_i \le \overline{a}_i$.

Thus, we can conclude the following:

- when $\overline{a}_i < \overline{b}_i$, we cannot have the second and the third possibilities and thus, we have $a_i = \overline{a}_i$;
- when $\overline{b}_i < \underline{a}_i$, we cannot have the first and the third cases and thus, we have $a_i = \underline{a}_i$;
- in the remaining cases, when $\underline{a}_i \le \overline{b}_i \le \overline{a}_i$, we must have the third possibility and thus, $a_i = \overline{b}_i$.

Thus, we arrive at the following algorithm.

V. Algorithm for Computing $\overline{d}_{\subseteq}$

Description of the algorithm. For every i from 1 to n, we select the following value $a_i \in [\underline{a}_i, \overline{a}_i]$:

- when $\overline{a}_i < \overline{b}_i$, we select $a_i = \overline{a}_i$;
- when $\overline{b}_i < \underline{a}_i$, we select $a_i = \underline{a}_i$;
- when $\underline{a}_i \le \overline{b}_i \le \overline{b}_i$, we select $a_i = \overline{b}_i$.

Using the selected values a_i, we compute the desired value $\overline{d}_{\subseteq}$ as

$$\overline{d}_{\subseteq} = \frac{\sum\limits_{i=1}^{n} \min(a_i, \overline{b}_i)}{\sum\limits_{i=1}^{n} a_i}.$$

Computational complexity of this algorithm. Selecting a_i requires linear time, and computing $\overline{d}_{\subseteq}$ also requires linear time, so overall the algorithm requires linear time $O(n)$.

VI. Computing $\underline{d}_=$: Analysis of the Problem

When is the ratio $d_= = \dfrac{\sum \min(a_i, b_i)}{\sum \max(a_i, b_i)}$ the smallest? Let us pick an index i; for this i, either $a_i \le b_i$ or $b_i \le a_i$.

In the first case $a_i \le b_i$, we have $\min(a_i, b_i) = a_i$ and $\max(a_i, b_i) = b_i$, so the ratio $d_=$ takes the form

$$d_= = \frac{a_i + m_i^=}{b_i + M_i^=},$$

where $m_i^= \stackrel{\text{def}}{=} \sum\limits_{j \ne i} \min(a_j, b_j)$ and $M_i^= \stackrel{\text{def}}{=} \sum\limits_{j \ne i} \max(a_j, b_j)$ do not depend on a_i and b_i.

The above expression $d_=$ increases with a_i and decreases with b_i, so its smallest possible value is attained when a_i is the smallest possible and b_i is the largest possible, i.e., when $a_i = \underline{a}_i$ and $b_i = \overline{b}_i$.

Similarly, for the case when $a_i \ge b_i$, the smallest possible value of $d_=$ is attained when $a_i = \overline{a}_i$ and $b_i = \underline{b}_i$.

Thus, for each i, we have two possibilities:

- $a_i = \underline{a}_i$, $b_i = \overline{b}_i$, and $\underline{a}_i \le \overline{b}_i$; and
- $a_i = \overline{a}_i$, $b_i = \underline{b}_i$, and $\underline{b}_i \le \overline{a}_i$.

When the intervals $[\underline{a}_i, \overline{a}_i]$ and $[\underline{b}_i, \overline{b}_i]$ do not intersect, we only have one of these possibilities:

- if $\overline{a}_i < \underline{b}_i$, then we have $a_i < b_i$ and thus, $a_i = \underline{a}_i$ and $b_i = \overline{b}_i$; and
- if $\overline{b}_i < \underline{a}_i$, then we have $b_i < a_i$ and thus, $a_i = \overline{a}_i$ and $b_i = \underline{b}_i$.

When the interval intersect, then, in principle, we have both options:

- $a_i = \underline{a}_i$ and $b_i = \overline{b}_i$, and
- $a_i = \overline{a}_i$ and $b_i = \underline{b}_i$.

Which of them leads the smaller value of the similarity degree $d_=$? For example, when is the first option better? It is better if the simultaneous replacing $\underline{a}_i$ with $\overline{a}_i$ and replacing $\overline{b}_i$ with $\underline{b}_i$ increases the value of $d_=$. The original value of $d_=$ is $\dfrac{m^=}{M^=}$, where $m^= \stackrel{\text{def}}{=} \sum\limits_{j=1}^{n} \min(a_j, b_j)$ and $M^= \stackrel{\text{def}}{=} \sum\limits_{j=1}^{n} \max(a_j, b_j)$. After the replacement,

- the numerator takes the new value $m^= + (\underline{b}_i - \underline{a}_i)$;
- the denominator takes the new value $M^= + (\overline{a}_i - \overline{b}_i)$;
- thus, the similarity degree takes the new value

$$\frac{m^= + (\underline{b}_i - \underline{a}_i)}{M^= + (\overline{a}_i - \overline{b}_i)}.$$

Thus, the new value of the similarity degree is larger than the original value if and only if

$$\frac{m^=}{M^=} < \frac{m^= + (\underline{b}_i - \underline{a}_i)}{M^= + (\overline{a}_i - \overline{b}_i)}.$$

Multiplying both sides by both positive denominators and canceling the common term $m^= \cdot M^=$ in both sides of the resulting inequality, we conclude that

$$m^= \cdot (\overline{a}_i - \overline{b}_i) < M^= \cdot (\underline{b}_i - \underline{a}_i).$$

If $\overline{a}_i < \overline{b}_i$, then we swap these two values, it will not affect the resulting similarity degree. After such swapping, we get $\overline{a}_i \ge \overline{b}_i$. Dividing both sides by positive numbers $\overline{a}_i - \overline{b}_i$ and $M^=$, we conclude that

$$\frac{m^=}{M^=} < \frac{\underline{b}_i - \underline{a}_i}{\overline{a}_i - \overline{b}_i}.$$

So, if the right-hand side ratio exceeds the ratio $\dfrac{m^=}{M^=}$, we get $a_i = \underline{a}_i$ and $b_i = \overline{b}_i$; otherwise, we get $a_i = \overline{a}_i$ and $b_i = \underline{b}_i$.

Similarly to the computation of $\underline{d}_{\subseteq}$, it is sufficient to know where the ratio $\dfrac{m^=}{M^=}$ stands in comparison with different values $\dfrac{\underline{b}_i - \underline{a}_i}{\overline{a}_i - \overline{b}_i}$. Thus, similarly to the computation of $\underline{d}_\subseteq$, we arrive at the following algorithm.

VII. Algorithm for Computing $\underline{d}_=$

Description of the algorithm. Before performing computations, for each i, we check whether $\overline{a}_i < \overline{b}_i$. If this inequality is indeed satisfied, then we swap the corresponding intervals $[\underline{a}_i, \overline{a}_i]$ and $[\underline{b}_i, \overline{b}_i]$. After these swaps, we have $\overline{a}_i \geq \overline{b}_i$ for all i.

Now come the actual computations. First, we divide n indices into two groups:

- the group I^- consists of all the indices i for which
$$\overline{b}_i < \underline{a}_i,$$
and
- the group I consisting of all remaining indices, i.e., of all the indices i for which
$$\overline{b}_i \geq \underline{a}_i.$$

For all $i \in I$, we compute the ratios
$$r_i^= \stackrel{\text{def}}{=} \frac{\underline{b}_i - \underline{a}_i}{\overline{a}_i - \overline{b}_i}.$$
We then sort the indices $i \in I$ in the increasing order of this ratio. In the resulting new ordering, we have
$$r_1^= \leq r_2^= \leq \ldots \leq r_k^=,$$
where k is the number of elements in the group I.

We then compute $m^{(0)} = \sum\limits_{i=1}^{n} \underline{b}_i$ and $M^{(0)} = \sum\limits_{i=1}^{n} \overline{a}_i$. For j from 0 to k, we then compute
$$m^{(j+1)} = m^{(j)} + (\underline{a}_{j+1} - \underline{b}_{j+1});$$
$$M^{(j+1)} = m^{(j)} + (\overline{b}_{j+1} - \overline{a}_{j+1}).$$

For all j from 0 to $k+1$, we compute $r^{(j)} = \dfrac{m^{(j)}}{M^{(j)}}$. The smallest of these values $r^{(j)}$ is the desired smallest value $\underline{d}_=$.

Computational complexity of this algorithm. Similarly to the algorithm for computing $\underline{d}_\subseteq$, this new algorithm requires
$$O(n) + \ldots + O(n) + O(n \cdot \log(n)) = O(n \cdot \log(n))$$
computation steps.

VIII. Computing $\overline{d}_=$: Analysis of the Problem

When is the ratio
$$d_= = \frac{\sum\limits_{i=1}^{n} \min(a_i, b_i)}{\sum\limits_{i=1}^{n} \max(a_i, b_i)}$$
the largest? Let us pick an index i; for this i, either $a_i \leq b_i$ or $b_i \leq a_i$.

For $a_i \leq b_i$, the ratio $d_=$ takes the form
$$d_= = \frac{a_i + m_i^=}{b_i + M_i^=}.$$
The above expression $d_=$ increases with a_i and decreases with b_i, so its largest possible value is attained when a_i is the largest possible and b_i is the smallest possible within the limitation $a_i \leq b_i$.

In other words, if $\overline{a}_i \leq \underline{b}_i$, then we should take $a_i = \overline{a}_i$ and $b_i = \underline{b}_i$. In all other cases (i.e., when $\overline{a}_i > \underline{b}_i$), we should take $a_i = b_i$. In this case, we have
$$d_= = \frac{a_i + m_i^=}{a_i + M_i^=},$$
with $m_i^= \leq M_i^=$. We already know, from the analysis of $\overline{d}_\subseteq$, that this expression increases with a_i. Thus, its largest value is attained when $a_i = b_i$ attains the largest possible value, i.e., when $a_i = b_i = \min(\overline{a}_i, \overline{b}_i)$.

Similarly, when $a_i \geq b_i$, the largest possible value of $d_=$ is attained when a_i is the smallest and b_i is the largest – within the constraint $a_i \geq b_i$. So, if $\overline{b}_i < \underline{a}_i$, the maximum is attained when $a_i = \underline{a}_i$ and $b_i = \overline{b}_i$. In all other cases, the maximum is attained when $a_i = b_i = \min(\overline{a}_i, \overline{b}_i)$.

Thus, depending on the relation between the intervals $[\underline{a}_i, \overline{a}_i]$ and $[\underline{b}_i, \overline{b}_i]$, we get three possible situations:

- if the interval $[\underline{a}_i, \overline{a}_i]$ is completely to the left of the interval $[\underline{b}_i, \overline{b}_i]$, i.e., if $\overline{a}_i < \underline{b}_i$, then $a_i < b_i$ and thus, in the optimal assignment, $a_i = \overline{a}_i$ and $b_i = \underline{b}_i$;
- if the interval $[\underline{a}_i, \overline{a}_i]$ is completely to the right of the interval $[\underline{b}_i, \overline{b}_i]$, i.e., if $\overline{b}_i < \underline{a}_i$, then $b_i < a_i$ and thus, in the optimal assignment, $a_i = \underline{a}_i$ and $b_i = \overline{b}_i$;
- finally, if the intervals $[\underline{a}_i, \overline{a}_i]$ and $[\underline{b}_i, \overline{b}_i]$ intersect, then in the optimal assignment, $a_i = b_i = \min(\overline{a}_i, \overline{b}_i)$, i.e., both a_i and b_i are equal to the upper endpoint of the intersection interval.

As a result, we arrive at the following algorithm.

IX. Algorithm for Computing $\overline{d}_=$

Description of the algorithm. For every i from 1 to n, we select the following value $a_i \in [\underline{a}_i, \overline{a}_i]$ and $b_i \in [\underline{b}_i, \overline{b}_i]$:

- when $\overline{a}_i < \underline{b}_i$, we select $a_i = \overline{a}_i$ and $b_i = \underline{b}_i$;
- when $\overline{b}_i < \underline{a}_i$, we select $a_i = \underline{a}_i$ and $b_i = \overline{b}_i$;
- in all other cases, we select $a_i = b_i = \min(\overline{a}_i, \overline{b}_i)$.

Using the selected values a_i and b_i, we compute the desired value $\overline{d}_=$ as
$$\overline{d}_= = \frac{\sum\limits_{i=1}^{n} \min(a_i, b_i)}{\sum\limits_{i=1}^{n} \max(a_i, b_i)}.$$

Computational complexity of this algorithm. Selecting a_i and b_i requires linear time, and computing $\overline{d}_=$ also requires linear time, so overall the algorithm requires linear time $O(n)$.

X. Conclusion

To adequately capture commonsense and expert reasoning, we must, in particular, capture the ambiguous, imprecise character of this reasoning. One of the most successful ways to describe such reasoning is the technique of fuzzy sets, a technique that generalizes the traditional set theoretic techniques to situations in which for each element a, there is a degree $\mu_A(a) \in [0,1]$ to which this element belongs to the set A (i.e, to which this element satisfies the property that defines the set A). These degrees, in turn, can only be determined with uncertainty, so in practice, we only know intervals $[\underline{\mu}_A(a), \overline{\mu}_A(a)]$ of possible values of these degrees. In other words, we practice, we only have an interval-valued fuzzy set.

Among the most important concepts of set theory are the notions of subsethood and equality. Thus, to extend set theoretic techniques to fuzzy sets and interval-valued fuzzy sets, it is desirable to be able to efficiently compute degrees of subsethood and degrees of equality (similarity) for fuzzy sets and for interval-valued fuzzy sets. There exist efficient algorithms for computing these degrees for fuzzy sets. In this paper, we showed that these algorithms can be extended to a more realistic case of interval-valued fuzzy sets.

Acknowledgments

This work was supported in part by NSF grant HRD-0734825, by Grant 1 T36 GM078000-01 from the National Institutes of Health, by the Japan Advanced Institute of Science and Technology (JAIST) International Joint Research Grant 2006-08, and by the Max Planck Institut für Mathematik.

The authors are thankful to Bernadette Bouchon-Meunier and Jerry M. Mendel for valuable discussions, and to the anonymous referees for important suggestions.

References

[1] B. Bouchon-Meunier, "Similarities in Fuzzy Data Mining: From a Cognitive View to Real-World Applications", *Proceedings of the IEEE World Congress on Computational Intelligence WCCI'2008*, Hong Kong, China, June 1–6, 2008.

[2] E. T. Jaynes, *Probability Theory: The Logic of Science*, Cambridge University Press, 2003.

[3] G. Klir and B. Yuan, *Fuzzy sets and fuzzy logic*, Prentice Hall, New Jersey, 1995.

[4] V. Kreinovich and G. Xiang, "Towards Fast Algorithms for Processing Type-2 Fuzzy Data: Extending Mendel's Algorithms From Interval-Valued to a More General Case", *Proceedings of the 27th International Conference of the North American Fuzzy Information Processing Society NAFIPS'2008*, New York, New York, May 19–22, 2008.

[5] J. M. Mendel, *Uncertain Rule-Based Fuzzy Logic Systems: Introduction and New Directions*, Prentice-Hall, 2001.

[6] J. M. Mendel, "Type-2 Fuzzy Sets and Systems: an Overview", *IEEE Computational Intelligence Magazine*, 2007, Vol. 2, pp. 20–29.

[7] J. M. Mendel, "Tutorial on the Uses of the Interval Type-2 Fuzzy Set's Wavy Slice Representation Theorem", *Proceedings of the 27th International Conference of the North American Fuzzy Information Processing Society NAFIPS'2008*, New York, New York, May 19–22, 2008.

[8] H. T. Nguyen, *An Introduction to Random Sets*, Chapman & Hall/CRC, Bocar Raton, Florida, 2006.

[9] H. T. Nguyen and E. A. Walker, *A First Course in Fuzzy Logic*, CRC Press, Boca Raton, Florida, 2006.

Appendix

Let us provide an alternative probabilistic justification of the above formulas for the degrees of subsethood and similarity.

Probabilistic justification of the formula for the degree of subsethood. Let us start with the degree of subsethood. From the probabilistic viewpoint, A is a subset of B if and only the conditional probability $P(B\,|\,A)$ that B holds under the condition that A holds is equal to 1. Thus, we can view this conditional probability $P(B\,|\,A)$ as the desired degree of subsethood.

By definition of conditional probability, we can describe this probability as

$$P(B\,|\,A) = \frac{P(A \cap B)}{P(A)}.$$

How do we get these probabilities?

For finite crisp sets, it is reasonable to assume that all the elements of the universe of discourse are equally probable; see, e.g., [2]. In this case, the probability $P(A)$ of a set A is simply proportional to the number of elements $|A|$ in this set. Thus, for finite crisp sets, we get the exact same ratio $\dfrac{|A \cap B|}{|A}$ as before.

How can we extend this idea to the fuzzy case? One reasonable way to gauge the values $\mu_A(x)$ of the membership function is to poll experts and to select, as $\mu_A(x)$, the proportion of experts who believe that the value x indeed satisfies the property A [3], [9]. In this case, $\mu_A(x)$ is the probability that, according to a randomly selected expert, the value x satisfies the corresponding property.

The above description can be reformulated in more mathematical term. Every expert has a set of values that, according to this expert's belief, satisfy the property A. We consider the experts to be equally valuable, so these sets are equally probable.

Thus, we have, in effect, a probability distribution on the class of all possible sets. Similarly to the fact that the probability distribution on the set of all possible numbers is called a random number, the probability distribution on the class of all possible sets is called a *random set.* Thus, a membership function $\mu_A(x)$ can be interpreted as the probability that a given element x belongs to the random set.

This interpretation of fuzzy sets as random sets has been known and used for several decades; see, e.g., [8] and references therein.

A fuzzy set means, in effect, that instead of a deterministic set, we have a random set, i.e., that we have a probability measure on the class of all possible sets. Thus, it is reasonable to describe the probability $P(A)$ as the probability that a random element x belongs to the corresponding random set. Due to the formula of full probability, this probability can be described as the integral $P(A) = \sum\limits_x P_A(x \in S) \cdot p(x)$, where

- $P_A(x \in S)$ is the probability that a given element x belongs to the corresponding random set, and
- $p(x)$ is the probability of the element x.

According to the above random set interpretation of a fuzzy set, the probability $P_A(x \in S)$ that a given element x belongs to the randomly selected set is equal to the corresponding value of the membership function $\mu_A(x)$. We also know that all the probabilities $p(x)$ are the same: $p(x) = c$ for some constant c. Thus, the desired probability has the form $P(A) = c \cdot \sum\limits_x \mu_A(x)$.

Similarly, the probability $P(A \cap B)$ is equal to

$$P(A \cap B) = c \cdot \sum_x \mu_{A \cap B}(x).$$

Therefore, the desired ratio

$$P(B \mid A) = \frac{P(A \cap B)}{P(A)}$$

is equal to

$$P(B \mid A) = \frac{\sum\limits_x \mu_{A \cap B}(x)}{\sum\limits_x \mu_A(x)},$$

i.e., to

$$P(B \mid A) = \frac{\sum\limits_{i=1}^{n} \min(a_i, b_i)}{\sum\limits_{i=1}^{n} a_i}.$$

So, the above formula for the degree of subsethood is indeed justified.

Probabilistic justification of the formula for the degree of similarity. The sets A and B are equal if every element that belongs to one of them belongs to both. In other words, $A = B$ means that every element of the union $A \cup B$ also belongs to the intersection $A \cap B$. From the probabilistic viewpoint, the sets A and B are equal if the conditional probability

$$P(A \cap B \mid A \cup B) = \frac{P(A \cap B)}{P(A \cup B)}$$

is equal to 1. Thus, we can view this conditional probability as the desired degree of similarity.

For fuzzy sets, we get

$$P(A \cap B) = c \cdot \sum_x \mu_{A \cap B}(x)$$

and

$$P(A \cup B) = c \cdot \sum_x \mu_{A \cup B}(x).$$

Thus, the desired ratio takes the form

$$P(A \cap B \mid A \cup B) = \frac{c \cdot \sum\limits_x \mu_{A \cap B}(x)}{c \cdot \sum\limits_x \mu_{A \cup B}(x)} = \frac{\sum\limits_{i=1}^{n} \min(a_i, b_i)}{\sum\limits_{i=1}^{n} \max(a_i, b_i)}.$$

Thus, the above formula for the degree of similarity has also been justified.