

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

Hung T. Nguyen¹ and Vladik Kreinovich²

¹Department of Mathematical Sciences
New Mexico State University
Las Cruces, NM 88003
hunguyen@nmsu.edu

²Department of Computer Science
University of Texas at El Paso
El Paso, Texas 79968, USA
vladik@utep.edu

Towards the ...
Probabilistic ...
Probabilistic ...
Alternative Ways of ...
Need for Interval- ...
Type-2 Fuzzy Sets: ...
How to Extend d_{\subseteq} to ...
How to Extend $d_{=}$ to ...
What We Do in This Talk
Computing d_{\subseteq} : ...
Computing d_{\subseteq} (cont-d)
Towards an Algorithm ...
Algorithm for ...
Conclusion
Conclusions (cont-d)
Acknowledgments

[Home Page](#)

[Title Page](#)

[⏪](#)

[⏩](#)

[⏴](#)

[⏵](#)

Page 1 of 20

[Go Back](#)

[Full Screen](#)

1. Subsethood and Set Equality – 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 .
- It is desirable to generalize these notions to fuzzy sets.
- In traditional set theory, for every two sets A and B , either $A \subseteq B$ or $A \not\subseteq B$; either $A = B$ or $A \neq B$.
- The main idea behind fuzzy logic is that for fuzzy, imprecise concepts, everything is a matter of degree.
- Thus, it is reasonable to define *degrees* of subsethood $d_{\subseteq}(A, B)$ and similarity $d_{=}(A, B)$.

[Probabilistic...](#)[Alternative Ways of...](#)[Need for Interval...](#)[Type-2 Fuzzy Sets...](#)[How to Extend \$d_{\subseteq}\$ to...](#)[How to Extend \$d_{=}\$ to...](#)[What We Do in This Talk](#)[Computing \$d_{\subseteq}\$...](#)[Computing \$d_{\subseteq}\$ \(cont-d\)](#)[Towards an Algorithm...](#)[Algorithm for...](#)[Conclusion](#)[Conclusions \(cont-d\)](#)[Acknowledgments](#)[Home Page](#)[Title Page](#)[Page 2 of 20](#)

2. How to Describe Degree of Subsethood

- In fuzzy logic and fuzzy set theory, there is no built-in degree of subsethood or degree of equality.
- In fuzzy, we only have union $\mu_{A \cup B}(x) = \max(\mu_A(x), \mu_B(x))$ and intersection $\mu_{A \cap B}(x) = \min(\mu_A(x), \mu_B(x))$.
- So, let us express subsethood in terms of \cup and \cap .
- Such an expression is well known in set theory:
(1) in general, $A \cap B \subseteq A$, and (2) $A \subseteq B \Leftrightarrow A \cap B = A$.
- So, we can take the ratio $d_{\subseteq}(A, B) = \frac{|A \cap B|}{|A|}$, where $|A|$ is the cardinality of A .
- For a fuzzy set, $|A| = \sum_x \mu_A(x) = \sum_{i=1}^n a_i$, so
$$d_{\subseteq}(A, B) = \frac{\sum \min(a_i, b_i)}{\sum a_i}.$$

3. Degree of Equality (Similarity)

- It is known that
 - in general, $A \cap B \subseteq A \cup B$, and
 - $A = B \Leftrightarrow A \cap B = A \cup B$.
- So, it is reasonable to define the degree of similarity as
$$d_{=}(A, B) = \frac{|A \cap B|}{|A \cup B|}.$$
- The smaller the ratio, the more there are elements from one of the sets which do not belong to the other.
- A similar definition can be used to define degree of equality (similarity) of two fuzzy sets A and B :

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

4. Towards the Probabilistic Justification of the Above Formulas: Fuzzy Sets as Random Sets

- *Idea:* we can gauge $\mu_A(x)$ as the proportion of experts who believe that x satisfies the property A .
- So, $\mu_A(x)$ is the probability that, according to a randomly selected expert, x satisfies A .
- 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 – a *random set*.
- Thus, $\mu_A(x)$ can be interpreted as the probability that a given element x belongs to the random set.

[Probabilistic...](#)[Alternative Ways of...](#)[Need for Interval...](#)[Type-2 Fuzzy Sets...](#)[How to Extend \$d_{\subseteq}\$ to...](#)[How to Extend \$d_{=}\$ to...](#)[What We Do in This Talk](#)[Computing \$d_{\subseteq}\$...](#)[Computing \$d_{\subseteq}\$ \(cont-d\)](#)[Towards an Algorithm...](#)[Algorithm for...](#)[Conclusion](#)[Conclusions \(cont-d\)](#)[Acknowledgments](#)[Home Page](#)[Title Page](#)

Page 5 of 20

5. Probabilistic Interpretation of the Formula for the Degree of Subsethood

- $A \subseteq B$ if and only if $P(B | A) = 1$.
- Thus, we define $d_{\subseteq}(A, B) \stackrel{\text{def}}{=} P(B | A) = \frac{P(A \cap B)}{P(A)}$.
- We know the probability $\mu_A(x) = P_A(x \in S)$ that a given element x belongs to the random set.
- Thus, $P(A) = \sum_x p(x) \cdot P_A(x \in S) = \sum_x p(x) \cdot \mu_A(x)$.
- We assume: all x are equally probable: $p(x) = c$.
- Similarly, $P(A \cap B) = c \cdot \sum_x \mu_{A \cap B}(x)$, so

$$P(B | A) = \frac{c \cdot \sum_x \mu_{A \cap B}(x)}{c \cdot \sum_x \mu_A(x)} = \frac{\sum_x \mu_{A \cap B}(x)}{\sum_x \mu_A(x)} = \frac{\sum_{i=1}^n \min(a_i, b_i)}{\sum_{i=1}^n a_i}.$$

6. Probabilistic Justification of the Formula for the Degree of Similarity

- $A = B$ if and only if that every element of the union $A \cup B$ also belongs to the intersection $A \cap B$.
- So, the sets A and B are equal if

$$P(A \cap B \mid A \cup B) = \frac{P(A \cap B)}{P(A \cup B)} = 1.$$

- Thus, we can define $d_{=}(A, B) \stackrel{\text{def}}{=} P(A \cap B \mid A \cup B)$.
- 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

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

7. Alternative Ways of Describing Degrees of Subsethood and Similarity: Comment

- The above expressions are the simplest and probably most frequently used.
- However, there exist other expressions.
- For example, we can use the following idea:
(1) in general, $B \subseteq A \cup B$, and (2) $B = A \cup B \Leftrightarrow A \subseteq B$.
- Thus, as an alternative degree of subsethood, we can take a ratio

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

- In this talk, we only consider the main definitions.

8. Need for Interval-Valued (and More General Type-2) Fuzzy Sets

- An expert often cannot describe his or her knowledge by an exact value or by a precise set of possible values.
- Instead, the expert describe this knowledge by using words from natural language.
- In fuzzy logic, the expert's degree of certainty that “ x is A ” is a number from $\mu_A(x) \in [0, 1]$.
- However, an expert is unable 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.
- Thus, for every x , the degree $\mu_A(x)$ is not a number, but rather a new fuzzy set.

Probabilistic...

Alternative Ways of...

Need for Interval...

Type-2 Fuzzy Sets:...

How to Extend d_{\subseteq} to...

How to Extend $d_{=}$ to...

What We Do in This Talk

Computing d_{\subseteq} ...

Computing d_{\subseteq} (cont-d)

Towards an Algorithm...

Algorithm for...

Conclusion

Conclusions (cont-d)

Acknowledgments

Home Page

Title Page



Page 9 of 20

9. Type-2 Fuzzy Sets: Successes and Challenges

- *Fact:* type-2 fuzzy sets provide a more adequate representation of expert knowledge.
- *Corollary:* they lead to higher quality control, clustering, etc.
- *Main challenge:* transition to type-2 fuzzy sets leads to an increase in computation time:
 - in type-1, each degree $\mu_A(x)$ is a number;
 - in type-2, to describe a fuzzy set $\mu_A(x)$, we need more than one number.
- *Simplest case:* each degree is an interval $[\underline{\mu}_A(x), \bar{\mu}_A(x)]$.
- *Complexity:* we only need 2 numbers to describe each degree.
- *Corollary:* interval-valued fuzzy sets are mostly widely among type-2 sets.

Probabilistic...

Alternative Ways of...

Need for Interval...

Type-2 Fuzzy Sets...

How to Extend d_{\subseteq} to...

How to Extend $d_{=}$ to...

What We Do in This Talk

Computing d_{\subseteq} ...

Computing d_{\subseteq} (cont-d)

Towards an Algorithm...

Algorithm for...

Conclusion

Conclusions (cont-d)

Acknowledgments

Home Page

Title Page

◀◀

▶▶

◀

▶

Page 10 of 20

10. How to Extend d_{\subseteq} to Interval-Valued Fuzzy Sets: Formulation of the Problem

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

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

- The objective is to find the range $[\underline{d}_{\subseteq}, \bar{d}_{\subseteq}]$ of possible values of the above subsethood degree, i.e.,
 - to compute the smallest possible value $\underline{d}_{\subseteq}$ of d_{\subseteq} when $a_i \in [\underline{a}_i, \bar{a}_i]$ and $b_i \in [\underline{b}_i, \bar{b}_i]$; and
 - to compute the largest possible value \bar{d}_{\subseteq} of d_{\subseteq} when $a_i \in [\underline{a}_i, \bar{a}_i]$ and $b_i \in [\underline{b}_i, \bar{b}_i]$.

11. How to Extend $d_{=}$ to Interval-Valued Fuzzy Sets: Precise Formulation of the Problem

- For each i from 1 to n , we know the intervals $[\underline{a}_i, \bar{a}_i]$ and $[\underline{b}_i, \bar{b}_i]$ of possible membership degrees.
- For each combination of values $a_i \in [\underline{a}_i, \bar{a}_i]$ and $b_i \in [\underline{b}_i, \bar{b}_i]$, we can compute the similarity degree

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

- The objective is to find the range $[\underline{d}_{=}, \bar{d}_{=}]$ of possible values of the above similarity degree, i.e.,
 - to compute the smallest possible value $\underline{d}_{=}$ of $d_{=}$ when $a_i \in [\underline{a}_i, \bar{a}_i]$ and $b_i \in [\underline{b}_i, \bar{b}_i]$; and
 - to compute the largest possible value $\bar{d}_{=}$ of $d_{=}$ when $a_i \in [\underline{a}_i, \bar{a}_i]$ and $b_i \in [\underline{b}_i, \bar{b}_i]$.

12. What We Do in This Talk

- *In the paper:* we design four fast algorithms:
 - a fast algorithm for computing $\underline{d}_{\subseteq}$;
 - a fast algorithm for computing \overline{d}_{\subseteq} ;
 - a fast algorithm for computing $\underline{d}_{=}$;
 - a fast algorithm for computing $\overline{d}_{=}$.
- *How fast:* each of these algorithms requires $O(n \cdot \log(n))$ computation steps.
- *In the talk:* we describe, in detail, one of these algorithms: for computing $\underline{d}_{\subseteq}$.

Probabilistic...

Alternative Ways of...

Need for Interval...

Type-2 Fuzzy Sets...

How to Extend d_{\subseteq} to...

How to Extend $d_{=}$ to...

What We Do in This Talk

Computing $\underline{d}_{\subseteq}$...

Computing $\underline{d}_{\subseteq}$ (cont-d)

Towards an Algorithm...

Algorithm for...

Conclusion

Conclusions (cont-d)

Acknowledgments

Home Page

Title Page

◀◀

▶▶

◀

▶

Page 13 of 20

13. Computing d_{\subseteq} : Analysis of the Problem

- *Fact:* $d_{\subseteq} = \frac{\sum \min(a_j, b_j)}{\sum a_j}$ is increasing in b_i .
- *Conclusion:* minimum is attained when $b_i = \underline{b}_i$.
- *Two possible cases:*
 - (1) $a_i \leq \underline{b}_i$ hence $\min(a_i, \underline{b}_i) = a_i$;
 - (2) $\underline{b}_i \leq a_i$ hence $\min(a_i, \underline{b}_i) = \underline{b}_i$.
- *Case 1:* $d_{\subseteq} = \frac{a_i + m_i}{a_i + M_i}$, where $m_i \stackrel{\text{def}}{=} \sum_{j \neq i} \min(a_j, \underline{b}_j)$ and $M_i \stackrel{\text{def}}{=} \sum_{j \neq i} a_j$.
- Since $\min(a_j, \underline{b}_j) \leq a_j$ for all j , we have $m_i \leq M_i$.
- Thus, $d_{\subseteq} = 1 - \frac{M_i - m_i}{a_i + M_i}$ increases when a_i increases.
- Hence, the minimum is when a_i is the smallest: $a_i = \underline{a}_i$.

14. Computing d_{\subseteq} (cont-d)

- Case 2 ($\underline{b}_i \leq a_i$): $d_{\subseteq} = \frac{m}{M} = \frac{\sum \min(a_j, b_j)}{\sum a_j} = \frac{m}{a_i + M_i}$.
- Conclusion: $d_{\subseteq} = \min$ when $a_i = \max = \bar{a}_i$.
- Case 1: $a_i = \underline{a}_i$ and $\underline{a}_i \leq \underline{b}_i$,
- Case 2: $a_i = \bar{a}_i$ and $\underline{b}_i \leq \bar{a}_i$.
- If $\bar{a}_i < \underline{b}_i$, Case 2 is impossible, so $a_i = \underline{a}_i$.
- If $\underline{b}_i < \underline{a}_i$, Case 1 is impossible, so $a_i = \bar{a}_i$.
- In the remaining cases, when $\underline{a}_i \leq \underline{b}_i \leq \bar{a}_i$, both $a_i = \underline{a}_i$ and $a_i = \bar{a}_i$ are possible.
- The value $a_i = \underline{a}_i$ minimizes if replacing it with \bar{a}_i increase the ratio: $\frac{m}{M} \leq \frac{m + (\underline{b}_i - \underline{a}_i)}{M + (\bar{a}_i - \underline{a}_i)}$.
- So, $a_i = \underline{a}_i \Leftrightarrow \frac{m}{M} \leq \frac{m + (\underline{b}_i - \underline{a}_i)}{M + (\bar{a}_i - \underline{a}_i)} \Leftrightarrow \frac{m}{M} \leq \frac{\underline{b}_i - \underline{a}_i}{\bar{a}_i - \underline{a}_i}$.

15. Towards an Algorithm for Computing $\underline{d}_{\subseteq}$

- If we know the optimal value $\underline{d}_{\subseteq}$ of the ratio m/M , then, as we have shown:
 - for each i ,
 - we can determine whether the minimum is attained for $a_i = \underline{a}_i$ or for $a_i = \bar{a}_i$.
- It is sufficient to know where m/M is w.r.t n ratios $r_i \stackrel{\text{def}}{=} \frac{b_i - \underline{a}_i}{\bar{a}_i - \underline{a}_i}$, i.e., $r_{(j)} \leq \frac{m}{M} \leq r_{(j+1)}$ for $r_{(1)} \leq r_{(2)} \leq \dots$
- For each possible location $r_{(j)} \leq m/M \leq r_{(j+1)}$:
 - we compute the corresponding values a_i , and then
 - we compute the resulting ratio $r^{(j)}$.
- Out of the resulting $n + 1$ ratios, we find the smallest.
- This smallest ratio is returned as $\underline{d}_{\subseteq}$.

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

- First, we divide n indices into three groups:

$$I^{-} = \{i : \bar{a}_i < \underline{b}_i\}, I^{+} = \{i : \underline{b}_i < \underline{a}_i\},$$

$$I = \{1, \dots, n\} - I^{-} - I^{+}.$$

- For all $i \in I$, we compute the ratios $r_i \stackrel{\text{def}}{=} \frac{\underline{b}_i - \underline{a}_i}{\bar{a}_i - \underline{a}_i}$ and sort them: $r_1 \leq r_2 \leq \dots \leq r_k$.

- We then compute

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

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

- For all j from 0 to $k + 1$, we compute $r^{(j)} = \frac{m^{(j)}}{M^{(j)}}$.
- The smallest of $r^{(j)}$ is the desired smallest value $\underline{d}_{\subseteq}$.
- Sorting takes $O(n \cdot \log(n))$ steps, rest is linear time $O(n)$, so overall we need $O(n \cdot \log(n))$ steps.

17. Conclusion

- To adequately capture commonsense reasoning, we must capture its imprecise character.
- One of the most successful ways to describe such reasoning is the technique of fuzzy sets.
- In fuzzy sets, for each element a , there is a degree $\mu_A(a) \in [0, 1]$ to which a belongs to the set A .
- These degrees, in turn, can only be determined with uncertainty.
- So, in practice, we only know intervals $[\underline{\mu}_A(a), \bar{\mu}_A(a)]$ 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.

Probabilistic...

Alternative Ways of...

Need for Interval...

Type-2 Fuzzy Sets:...

How to Extend d_{\subseteq} to...

How to Extend $d_{=}$ to...

What We Do in This Talk

Computing d_{\subseteq} :...

Computing d_{\subseteq} (cont-d)

Towards an Algorithm...

Algorithm for...

Conclusion

Conclusions (cont-d)

Acknowledgments

Home Page

Title Page

◀◀

▶▶

◀

▶

Page 18 of 20

18. Conclusions (cont-d)

- *Reminder:* among the most important concepts of set theory are the notions of subethood and equality.
- Thus,
 - to extend set theoretic techniques to fuzzy sets and interval-valued fuzzy sets,
 - we must be able to efficiently compute degrees of subethood 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.
- We have shown how to extend these algorithms to a (more realistic) case of interval-valued fuzzy sets.

Probabilistic...

Alternative Ways of...

Need for Interval...

Type-2 Fuzzy Sets:...

How to Extend d_{\subseteq} to...

How to Extend $d_{=}$ to...

What We Do in This Talk

Computing d_{\subseteq} ...

Computing d_{\subseteq} (cont-d)

Towards an Algorithm...

Algorithm for...

Conclusion

Conclusions (cont-d)

Acknowledgments

Home Page

Title Page

◀◀

▶▶

◀

▶

Page 19 of 20

19. 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.

Towards the ...
Probabilistic ...
Probabilistic ...
Alternative Ways of ...
Need for Interval- ...
Type-2 Fuzzy Sets: ...
How to Extend d_{\subseteq} to ...
How to Extend $d_{=}$ to ...
What We Do in This Talk
Computing d_{\subseteq} : ...
Computing d_{\subseteq} (cont-d)
Towards an Algorithm ...
Algorithm for ...
Conclusion
Conclusions (cont-d)
Acknowledgments

[Home Page](#)[Title Page](#)[◀◀](#)[▶▶](#)[◀](#)[▶](#)[Page 20 of 20](#)[Go Back](#)[Full Screen](#)